mentation of adaptive filters. In this chapter, we focus on the normalized least mean squares (NLMS) algorithm used with a variable normalized step size. The NLMS has proven indeed to be a simple, robust and effective algorithm that has found a variety of applications in signal processing (Hayes, 1996; Haykin, 2001; Widrow and Stearns, 1985). We will show that the variable normalized step size is necessary for the filter to respond quickly to changes in signal statistics.

First, in Section 6.2 we present the structure of the multiple-input adaptive noise canceller, we develop the theory about the normalized least mean squares algorithm and we derive a method to obtain a variable step size parameter. In Sections 6.3.2 and 6.3.3, we consider an application to swell noise attenuation on two real marine seismic data sets, and compare against two other noise attenuation methods: time-frequency median filtering and a second order high-pass Butterworth filter. We show that the multiple-input adaptive noise canceller achieves a very good attenuation of coherent noise while preserving the seismic reflections in the case of localized, high amplitudes swell noise that differs significantly from seismic reflections. This method yields better results than time-frequency median filtering and second order high-pass Butterworth filter when dealing with such noise. However, the multiple-input adaptive noise canceller can only partly attenuate the coherent noise when this one is somewhat correlated with seismic reflections. Such noise is however much more difficult to handle, and the other two comparative methods also show some limitations in their attempt of attenuating it. However, we demonstrate that combining time-frequency median filtering with the multiple-input adaptive noise canceller yields a very good attenuation of coherent noise. Finally, in Section 6.3.4, we consider an application to seismic interference attenuation, for which the efficiency is admittedly limited.

6.2 Theory

6.2.1 Multiple-input adaptive noise canceller

Let us consider a seismic trace or primary channel whose value at time sample n is denoted by $x_k(n)$. The trace signal consists of the sum of a seismic signal $s_k(n)$ corrupted by noise $\nu_k(n)$, such that

$$x_k(n) = s_k(n) + \nu_k(n). \tag{6.1}$$

The multiple-input adaptive noise canceller uses a set of M noise sequences $\boldsymbol{\nu}_1(n), \ldots,$ $\boldsymbol{\nu}_M(n)$, where $\boldsymbol{\nu}_i(n) = [\nu_i(n)\,\nu_i(n-1)\,\ldots\,\nu_i(n-L+1)]^T$, to predict the noise contained in the primary channel at time sample n, and then subtract it from the primary. The block diagram is shown in Figure 6.1. Thus, if the input noise sequences are correlated with the corrupting noise $\nu_k(n)$ but uncorrelated with the seismic signal $s_k(n)$, then the multiple-input adaptive noise canceller provides an estimate of the noise $\hat{\nu}_k(n) = \sum_{i=1}^M \hat{\nu}_i(n)$, that is subtracted from the primary channel $x_k(n)$ to form an estimate of the

seismic signal $s_k(n)$, given by the error value $e(n)$:

$$e(n) = s_k(n) + \nu_k(n) - \sum_{i=1}^{M} \hat{\nu}_i(n) = \hat{s}_k(n). \tag{6.2}$$

The estimate of the noise component, $\hat{\nu}_k(n)$, is obtained by finding a set of M optimal time-dependent tap-weight vectors $\mathbf{w}_1(n), \dots, \mathbf{w}_M(n)$. As optimality criterion, we mean here the minimization of the mean-square error $E\{|e(n)|^2\}$ at any time. Under the previous assumptions, minimizing the mean-square error $E\{|e(n)|^2\}$ is equivalent to minimizing $E\{|\nu_k(n) - \sum_{i=1}^{M} \hat{\nu}_i(n)|^2\}$, and therefore the error signal $e(n)$ forms the mean-square estimate of $s_k(n)$. In order to determine the set of tap-weight vectors that minimizes the mean-square error, we use the normalized least mean squares algorithm. The choice of an adaptive filter is justified by the nonstationary character of the seismic data where a non-adaptive filter fails to adapt to the changes of the signal statistics. Finally, as input noise sequences we use either pure noise recordings recorded shortly before data has been acquired if they are available, or we extract noise sequences from data records prior the first seismic reflection arrival.
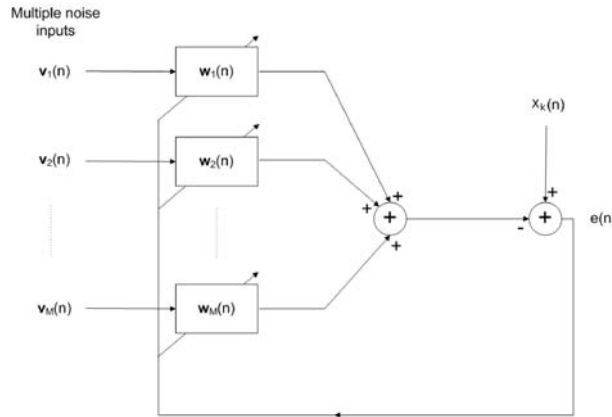

Figure 6.1: Block diagram of the multiple inputs adaptive noise canceller.

6.2.2 The normalized least mean squares algorithm

The least mean squares algorithm is a simple, robust and efficient algorithm that has found a variety of applications in signal processing. In the least mean squares algorithm, the minimization of the mean-square error is actually replaced by the minimization of the instantaneous squared error $|e(n)|^2$, which does not require any ensemble average to be known.

Let us consider an adaptive finite impulse response (FIR) filter of order L that produces an output

$$\hat{\nu}_i(n) = \mathbf{w}_i^H(n)\boldsymbol{\nu}_i(n), \tag{6.3}$$

where $\mathbf{w}_i(n) = [w_{i,0}(n)\, w_{i,1}(n)\, \ldots\, w_{i,L-1}(n)]^T$ is the ith tap-weight vector of the adaptive filter and $\boldsymbol{\nu}_i(n) = [\nu_i(n)\, \nu_i(n-1)\, \ldots\, \nu_i(n-L+1)]^T$ is the ith tap-input vector made of the most recent L samples of input noise sequence $\nu_i(n)$. If we consider M input noise sequences and if we denote the kth primary channel by $x_k(n)$, then the error between the primary channel and the multiple noise inputs is given by

$$e(n) = x_k(n) - \sum_{i=1}^{M} \boldsymbol{\nu}_i^T(n)\mathbf{w}_i^*(n). \tag{6.4}$$

The LMS algorithm updates the ith tap-weight vector by the gradient of the instantaneous squared error (Widrow and Stearns, 1985), such as

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) - \mu \nabla |e(n)|^2 = \mathbf{w}_i(n) - \mu\, e(n)\nabla e^*(n). \tag{6.5}$$

Using the error signal given in equation 6.4, the gradient of the conjugate error is given by

$$\nabla e^*(n) = -[\nu_i^*(n)\, \nu_i^*(n-1)\, \ldots\, \nu_i^*(n-L+1)]^T, \tag{6.6}$$

and therefore, the update equation of the ith tap-weight vector can be written as

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu\, e(n)\boldsymbol{\nu}_i^*(n). \tag{6.7}$$

Because the selection of the step size μ can be tricky, we prefer the use of the normalized LMS instead by setting

$$\mu = \frac{\beta}{\epsilon + ||\boldsymbol{\nu}_i^*(n)||^2} \tag{6.8}$$

where $0 < \beta < 2$ is the normalized step-size and ϵ is a small positive number called regularization parameter. The normalization by $||\boldsymbol{\nu}_i^*(n)||$ makes us avoid the problem of noise amplification, while ϵ represents the regularization parameter added to overcome situations where $||\boldsymbol{\nu}_i^*(n)||$ is too small.

Therefore, we can summarize the normalized LMS algorithm for a FIR adaptive filter of order L using M input noise sequences and applied to the kth seismic trace, as follows:

$$\text{Initialization:} \quad \mathbf{w}_i(0) = \mathbf{0} \qquad \text{for } i = 1, \ldots, M. \tag{6.9}$$

$$\text{Computation:} \quad \text{For } n = 0, 1, 2, \ldots \tag{6.10}$$

$$e(n+1) = x_k(n+1) - \sum_{i=1}^{M} \boldsymbol{\nu}_i^T(n+1)\mathbf{w}_i^*(n) \tag{6.11}$$

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \frac{\beta}{\epsilon + ||\boldsymbol{\nu}_i^*(n+1)||^2}\, e(n+1)\boldsymbol{\nu}_i^*(n+1) \tag{6.12}$$

$$\hat{s}_k(n+1) = x_k(n+1) - \sum_{i=1}^{M} \boldsymbol{\nu}_i^T(n+1)\mathbf{w}_i^*(n+1). \tag{6.13}$$

Equation 6.12 is the tap-weight update equation while equation 6.13 gives the a posteriori error signal that represents the estimate of the seismic signal.

6.2.3 Variable normalized step size parameter

The normalized step-size parameter β of the NLMS algorithm that governs the convergence speed and the steady-state excess of mean-square error must be properly selected. It is shown that the rate of convergence is proportional to the step-size, while the steady-state excess of mean-square error is inversely proportional to the step-size. Thus, a fixed β value means a tradeoff between rate of convergence, steady-state excess of mean-square error and the ability of the filter to track signals as their statistics change (Haykin, 2001; Zhao et al., 2009). Preliminary experiments on seismic data showed us that if β is large, the correction of $\mathbf{w}_i$ is large and as a consequence, the low-frequency noise is significantly attenuated as well as certain seismic reflections. On the opposite, if β is small, the correction of $\mathbf{w}_i$ is small and then, the estimated signal follows rather accurately the seismic reflections but fails to attenuate the low-frequency noise significantly.

Since we seek a filter that conserves the seismic reflections while attenuating coherent low-frequency noise, the use of a variable normalized step size is necessary. We first tested a well-known variable step size algorithm described by Kwong and Johnston (1992) where the step size adjustment is controlled by the square of the prediction error. It appeared to be unsuccessful when applied to seismic data because using the error as a metric yields the same shortcomings as for constant step size. There is, however, one metric that differentiates seismic reflections from the coherent low-frequency noise, which is the instantaneous frequency. Seismic reflections have indeed instantaneous frequencies that are larger than coherent low-frequency noise. Thus, we look for an instantaneous frequency threshold value that differentiates seismic reflections from coherent low-frequency noise. For instantaneous frequencies smaller than the threshold value, low-frequency noise is detected and a larger normalized step size should be used. For instantaneous frequencies larger than the threshold value, seismic reflections are detected and a smaller normalized step size should be used.

For this application, we calculate the instantaneous frequency $\phi(t)$ by means of the Hilbert transform. The Hilbert transform $\tilde{x}(t)$ of a real signal $x(t)$ is the conjugate of this signal, i.e., a version of this signal with a $90°$ phase shift. They create a so-called analytic signal $z(t)$ that can be written with an amplitude and a phase, and whose phase derivative can be identified as the instantaneous frequency. Thus, we define the analytic signal $z(t)$ of $x(t)$ as

$$z(t) = x(t) + j\tilde{x}(t) \tag{6.14}$$

where j is the square root of -1. The imaginary part $\tilde{x}(t)$ contains the Hilbert transform, defined as

$$\tilde{x}(t) = \mathcal{H}[x(t)] = \frac{1}{\pi} \, p.v. \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} \mathrm{d}\tau \tag{6.15}$$

where p.v. denotes the Cauchy principal value, provided this integral exists as a principal value. Note that the Hilbert transform also satisfies $\mathcal{H}[\mathcal{H}[x(t)]] = -x(t)$. In addition, $x(t)$ and $\tilde{x}(t)$ have the same amplitude and frequency content, and $\tilde{x}(t)$ includes phase information that depends on the phase of $x(t)$ (Boashash, 1992; Claerbout, 1976; Hahn,

1996). The Hilbert transform is useful in calculating instantaneous attributes of a time series and in particular, the instantaneous frequency (Boashash, 1992; Cohen, 1995; Luo et al., 2003). Thus, given the real signal $x(t) = A(t)\cos\theta(t)$, we may express the analytic signal in the compact form $z(t) = A(t)e^{j\theta(t)}$, where

$$A(t) = \sqrt{x^2(t) + \tilde{x}^2(t)} \tag{6.16}$$

is referred to as the instantaneous amplitude and

$$\theta(t) = \arctan\frac{\tilde{x}(t)}{x(t)} \tag{6.17}$$

is referred to as the instantaneous phase of the analytic signal. Therefore, the instantaneous frequency, that is the time rate of change of the instantaneous phase angle, is given by

$$\phi(t) = \frac{\theta(t + T_s) - \theta(t)}{2\pi T_s} \tag{6.18}$$

where T_s denotes the sampling period. Using a discrete-time notation $t = nT_s$, $n \in \mathbb{Z}$, the instantaneous frequency is given by

$$\phi(n) = \frac{\theta(n + 1) - \theta(n)}{2\pi T_s}. \tag{6.19}$$

In our method, we first set up to two threshold values for instantaneous frequency, denoted by Φ_1 and Φ_2. In order to choose these two values, we calculate the instantaneous frequency for all the noise input sequences to obtain the range of possible values, and then choose the thresholds so that they include a high percentage of the frequency values. Thereafter, we need to calculate the instantaneous frequencies $\phi_k(n)$ for each seismic trace $x_k(n)$. Because they usually are very spiky, it is necessary to smooth the frequency time series before applying thresholds. To do so, we use the robust locally weighted regression proposed by Cleveland (1979). This method is an extension of smoothing techniques based on local fitting of polynomials where a robust fitting procedure is used that guards against deviant points distorting the smoothed points. Note that this smoothing function is built in Matlab® (2010). Thus, we are now able to define the normalized step-size parameter value β as follows

$$\beta(n) = \begin{cases} \beta_0 & \text{if } \phi_k(n) \geq \Phi_1 \\ \beta_1 & \text{if } \Phi_2 \leq \phi_k(n) < \Phi_1 \\ \beta_2 & \text{if } \phi_k(n) < \Phi_2. \end{cases} \tag{6.20}$$

Note that the normalized step-size parameter has become time dependent and it is therefore denoted by $\beta(n)$ here. The equation 6.12 must be subsequently updated and becomes

$$\mathbf{w}_i(n + 1) = \mathbf{w}_i(n) + \frac{\beta(n)}{\epsilon + ||\boldsymbol{\nu}_i^*(n + 1)||^2} e(n + 1)\boldsymbol{\nu}_i^*(n + 1). \tag{6.21}$$

6.2.4   Other comparative methods

In order to evaluate the performances of the variable step size multiple-input adaptive noise canceller, we will test our data set against two other low-frequency noise attenuation methods, that are a local time-frequency median filter and a second order high-pass Butterworth filter. In this chapter, we use a time-frequency median filter as described by Elboth et al. (2010). This filter uses a sliding window that scans the input gather both in space and time. Within this window, a spectral estimate of all traces is achieved and the amplitude estimates at each frequency are compared with the amplitude of a presumed good trace within the chosen window. Here, time-frequency median filtering is applied within the frequency range 0 - 15 Hz where every amplitude estimate greater than 3.2 times a reference amplitude corresponding here to the lower quartile of the amplitude spectrum is attenuated.

The Butterworth filter is a well known type of filter designed to have a frequency response which is as flat as possible in the passband and rolls off towards zero in the stopband (Bianchi and Sorrentino, 2007). We use in our applications a second order high-pass Butterworth filter with cutoff frequency at 2 Hz. It is described by its system function $H(z) = \frac{0.97 - 1.93z^{-1} + 0.97z^{-2}}{1 - 1.93z^{-1} + 0.93z^{-2}}$ and its magnitude response is shown in Figure 6.2.
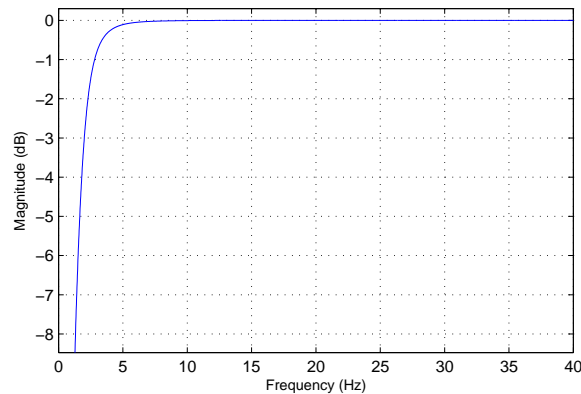

Figure 6.2: Magnitude response of a second order high-pass Butterworth filter with cutoff frequency at 2 Hz (Zoom in the first 40 Hz).

6.3   Applications

6.3.1   Data conditioning

In order to avoid convergence problems with the NLMS algorithm, we wish to have input noise sequences that have the same number of samples as the primary channel. If it is not the case, we replicate the noise sequences as many times as necessary and smooth up to five data samples just before and after the replication time by mean of

cubic spline interpolation to avoid sharp transitions. Finally, we shift each of these new input noise sequences so that the replication time differs from sequence to sequence.

In addition, in spite of the use of a time-adaptive algorithm, it appears that it cannot always follow the rapidly changing statistics of input traces. In that case, a solution to boost our method consists of dividing each input trace into small overlapping segments of N_B samples each and reinitializing the algorithm at the beginning of each segment. The overlapping samples are then discarded to avoid the non-steady state of the first few iterations.

6.3.2   Application to swell noise attenuation: first example

We test the multiple-input adaptive noise canceller on two sets of real marine seismic data. The first one, shown in Figure 6.3 (a), is a marine shot gather that consists of 200 traces and that is greatly corrupted by swell noise in a very localized manner. We extract a data window prior the first seismic reflection arrival in the offset interval $2 - 2.36$ km and in the time interval $0 - 1.12$ s, shown in Figure 6.3 (b). It is then replicated to form 30 input noise sequences of length equal to the traces length.

We first consider a variable step size NLMS algorithm of order $L = 50$. The block length N_B is set to 0.6 s with 10 % overlap between blocks and ϵ is set to 0.0001. In this example, we use two threshold values Φ_1 and Φ_2 by counting up respectively 99 % and 97 % of all instantaneous frequency values. The normalized step size values are set to $\beta_0 = 0.00005$, $\beta_1 = 0.005$ and $\beta_2 = 0.015$. In order to optimize the noise attenuation, we apply the variable step size NLMS algorithm three times successively with the same parameters. The output gather after the first application is shown in Figure 6.4 (a) and after the third application in Figure 6.4 (b). The residual plot, i.e., the difference between the input and output gathers after three applications of the filter is shown in Figure 6.4 (c). We observe that most of the swell noise has been removed while the seismic reflections are unaffected. However, three successive applications of the filter to this data set are necessary to attenuate the largest amplitudes of swell noise.

We now compare the performances of the variable step size multiple-input adaptive noise canceller against the local time-frequency median filter. The time-frequency median filter is applied to the entire data set a first time, and then a second time to data recorded after 1.8 s to avoid the attenuation of seismic reflections. The output gather is shown in Figure 6.5 (a) while the residual plot is shown in Figure 6.5 (b). A few seismic reflections have been attenuated before 2.5 s, but a significant part of swell noise, including the largest swell amplitudes, have been removed. However, some swell noise is still visible in the output gather after 5 s, and in particular for offset distances between 0.3 and 1.0 km. This is caused by the trace-to-trace comparison used by this technique. Applying the time-frequency median filter one more time may remove some more data, which is not desired. In addition, the amount of swell noise left is much more important than for the multiple-input adaptive noise canceller. Let us now apply the second-order high-pass Butterworth filter to remove the low-frequency components of this data set.

Again, the choice of a larger cutoff frequency or a larger filter order, as well as successive applications of this high-pass filter, would remove too much seismic signal. Figure 6.6 (a and b) shows the output gather and the residual plot, respectively. The attenuation of swell noise, especially the highest amplitudes, is very limited and in addition, a few seismic reflections are attenuated as well before 2 s. Such filter is definitively not a valuable alternative to handle this type of noise.

We now combine the time-frequency median filter with the multiple-input adaptive noise canceller. We think the first method may attenuate the largest swell amplitudes, while the second one should attenuate most of the coherent noise left. Keeping the same parameters for both methods, we first apply the time-frequency median filter to the entire input gather followed by the multiple-input adaptive noise canceller. The output gather and the residual plot are shown respectively in Figure 6.7 (a) and (b). The noise attenuation is not as efficient as in Figure 6.4 (b) where only the multiple-input noise canceller is used, with more leftovers in deep waters for all offsets. In addition, more seismic reflections are attenuated as well. Consequently, there is no benefits to combine the time-frequency median filter with the multiple-input noise canceller for the attenuation of this type of noise, and the multiple-input noise canceller used by itself provides the best solution here.

6.3.3   Application to swell noise attenuation: second example

Let us consider another marine shot gather shown in Figure 6.8 (a) made of 96 traces and corrupted by seismic interferences visible after 1 s, and by swell noise, visible around offsets 0.25 km, 0.5 km and beyond 1 km. Contrary to the previous data set, the frequency content of the swell noise does not differ so dramatically from the frequency content of seismic signals. In addition, it is recorded along the entire traces and characterized by smaller amplitudes. We extract a data window prior the first seismic reflection arrival in the offset interval 0.937 – 1.112 km and in the time interval 0 – 0.6 s, and replicate it to form 15 input noise sequences of length equal to the traces length shown in Figure 6.8 (b).

We consider a variable step size NLMS algorithm of order $L = 50$, where data are treated as a single block, i.e., N_B is equal to the trace length. In this example, we use a unique instantaneous frequency threshold value, $\Phi_1 = \Phi_2 = 70$ Hz. The normalized step size values are $\beta_0 = 0.000005$ and $\beta_1 = \beta_2 = 0.001$, while $\epsilon = 0.0001$. In order to optimize the noise attenuation, we apply the variable step size NLMS algorithm three times successively with the same parameters. The output gather after the first application is shown in Figure 6.9 (a) and after the third application in Figure 6.9 (b). The residual plot, i.e., the difference between the input and output gathers after three applications of the filter is shown in Figure 6.9 (c). After three applications, lots of swell noise has been removed, more specifically around 0.25 km, 0.4 – 0.5 km, 0.7 km and beyond 0.9 km, while the seismic reflections have been preserved. We observed however that the output gather still contains low-frequency noise, especially along a few chaotic traces at
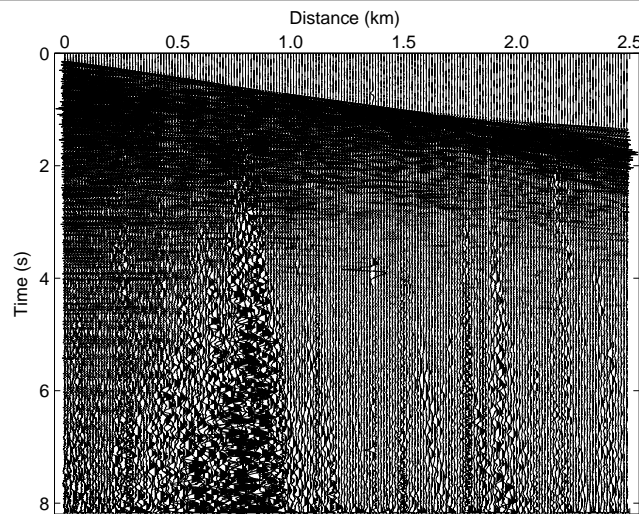
(a)
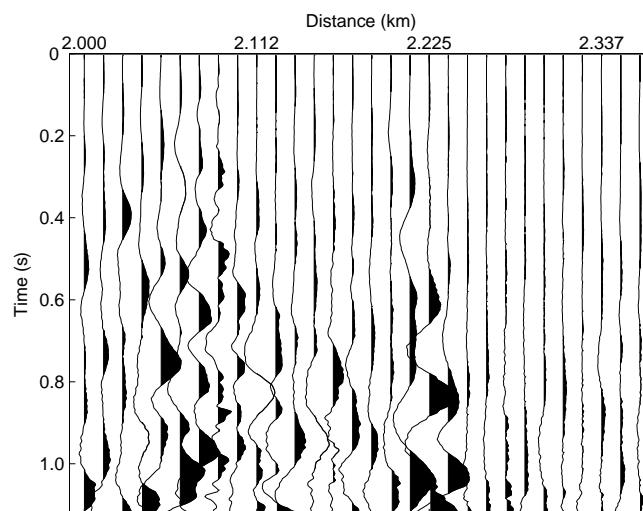
(b)

Figure 6.3: First example: (a) input marine shot gather and (b) 30 extracted input noise sequences.

offsets 0.28 km, 0.4 km, 1.1 km and 1.2 km. Finally, like in the first example, three successive applications of this filter are necessary to attenuate significantly the low-frequency noise. This filter however is not able to attenuate the seismic interferences.

Clearly, distinct swell noise that is localized within an input gather and that is characterized by a high amplitudes, low-frequency content as in the first example, is much more successfully handled by the multiple-input adaptive noise canceller. First, be-

102

(a)

(b)

(c)
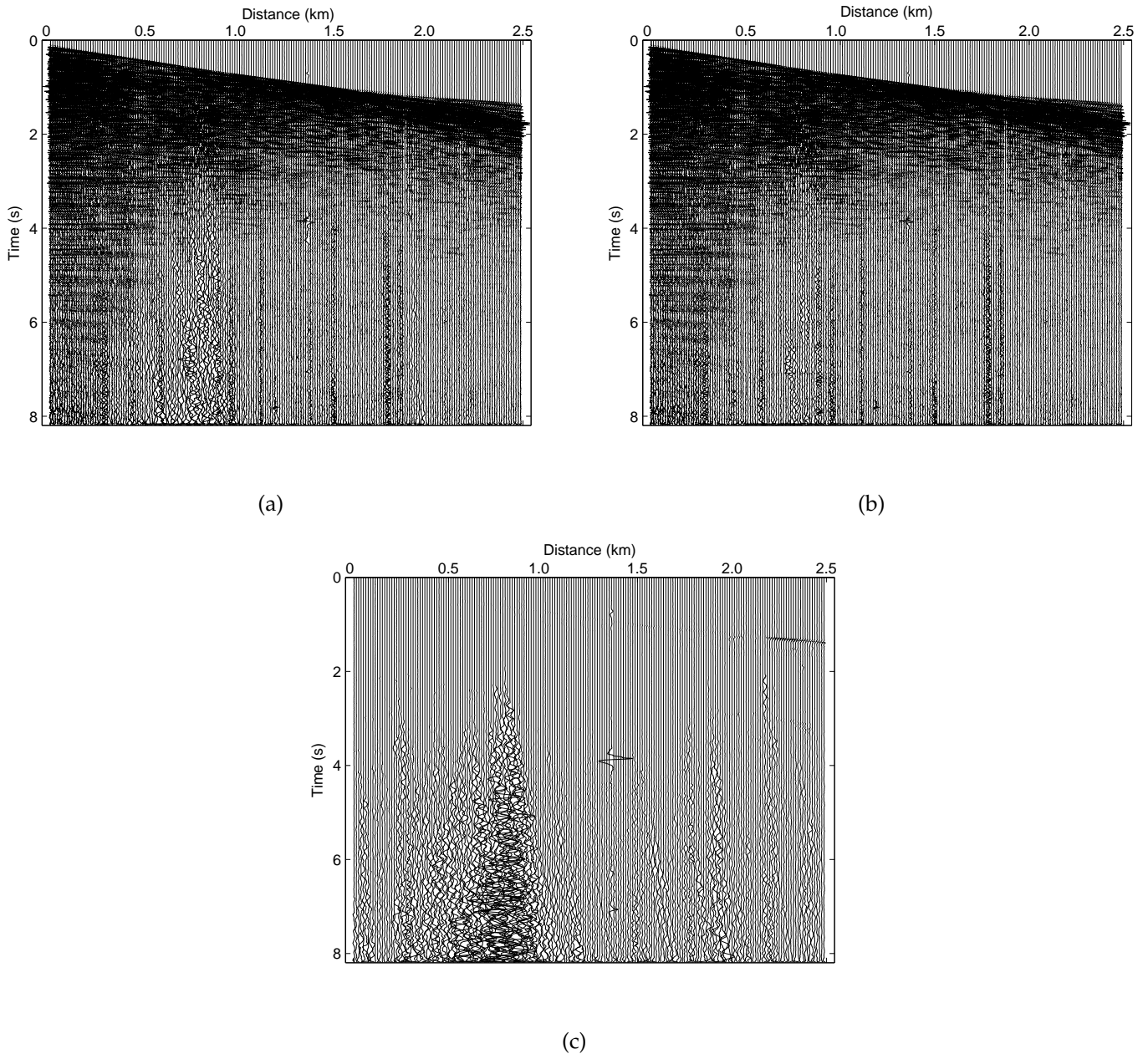
Figure 6.4: Multiple-input noise canceller: output shot gather after (a) one application of the filter, and (b) three applications of the filter. (c) Residual plot after three applications of the filter.

cause swell noise and seismic reflections have a rather different range of instantaneous frequencies, it is easier for the user to set optimal threshold values. Then, because the noise is clearly uncorrelated with the seismic reflections, its attenuation by the multiple-
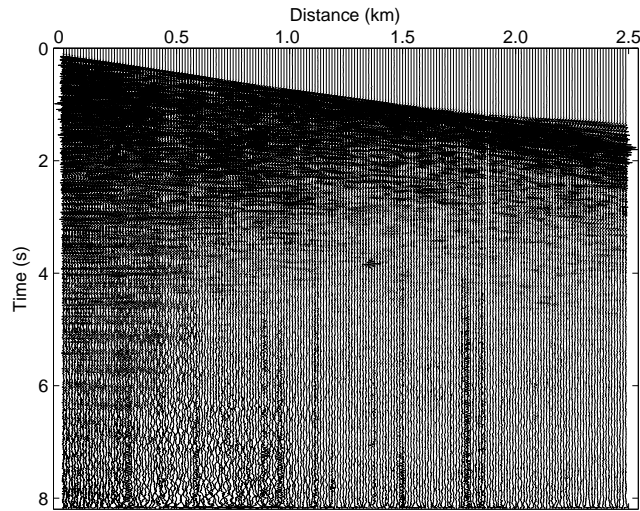
(a)
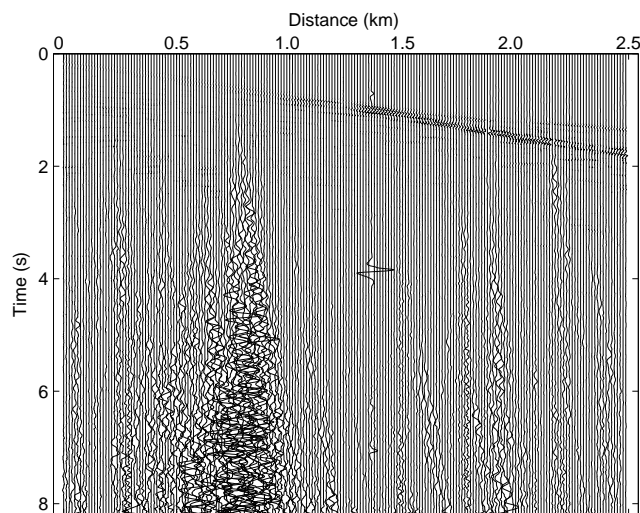
(b)

Figure 6.5: Time-frequency median filtering: (a) output shot gather and (b) residual plot.

input adaptive noise canceller is very efficient. On the contrary, when the low-frequency noise is somewhat correlated with seismic reflections, as in the second example, the multiple-input adaptive noise canceller has a limited efficiency.

For comparison, we apply the time-frequency median filter within the frequency range 0 - 12 Hz to the data set shown in Figure 6.8 (a). Successive applications of this filter do not bring further improvement. Figure 6.10 (a and b) shows the output gather and the residual plot, respectively. The highest amplitudes of swell noise are much better
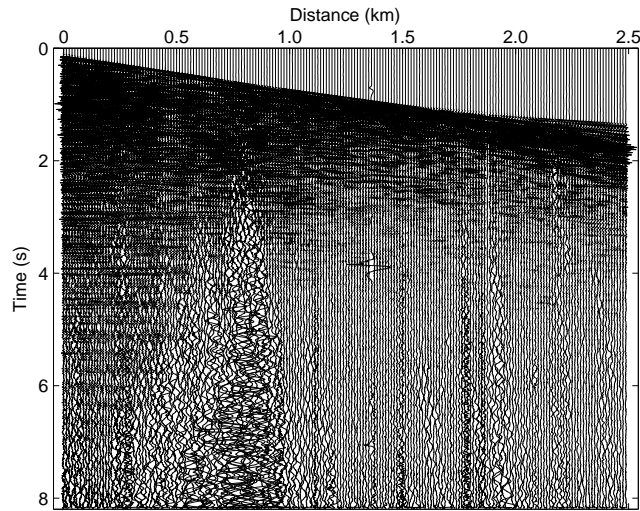
(a)
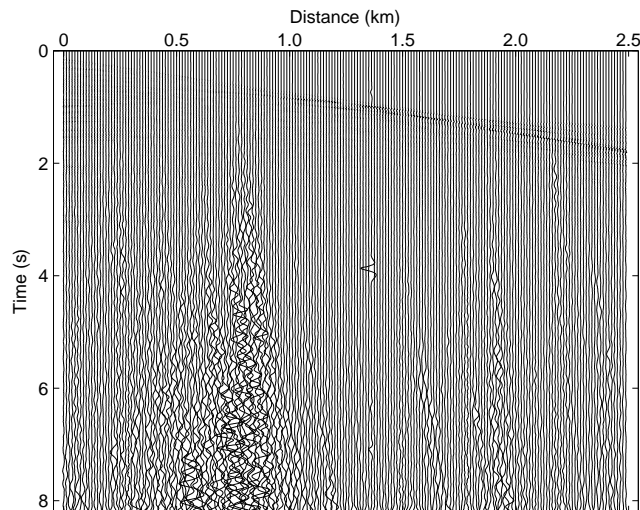
(b)

Figure 6.6: High-pass Butterworth filter: (a) output shot gather and (b) residual plot.

attenuated at for instance distances 0.26 km, 0.40 km and 1.15 km, but the smallest ones are almost left unchanged. This is explained by the fact that the time-frequency median filter compares each frequency amplitude with the ones from the neighbor traces that are also contaminated by noise of similar amplitude spectrum. In such a case, small frequency amplitudes are not considered as abnormal and therefore, they are not attenuated. Since the multiple-input noise canceller better attenuates small swell noise amplitudes, then both methods are potentially complementary. Thus, we investigate

Distance (km)

(a)

Distance (km)
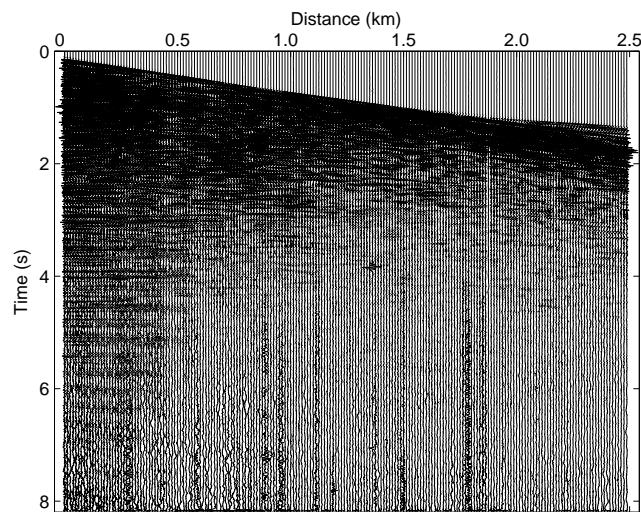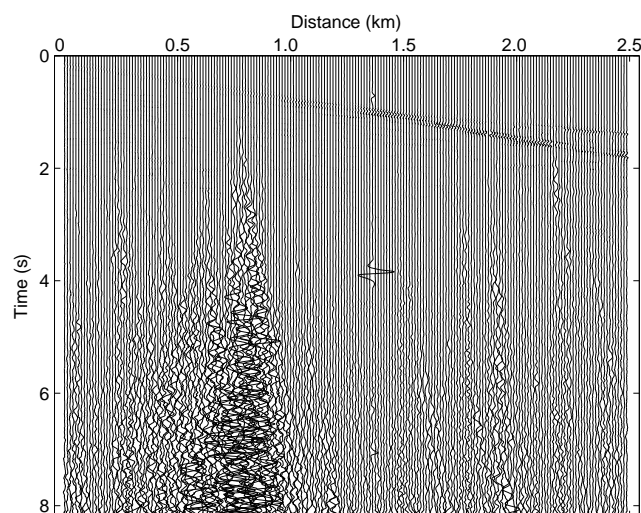
(b)

Figure 6.7: Time-frequency median filtering + multiple-input noise canceller: (a) output shot gather and (b) residual plot.

now the combination of the time-frequency median filter with the multiple-input adaptive noise canceller. We apply the multiple-input adaptive noise canceller to the output of the time-frequency median filter with the same parameters as in Figure 6.9. The resulting output gather and its residual plot are shown in Figure 6.11. This combination yields significant improvement since small swell noise amplitudes are much better attenuated at distances 0.25 – 0.5 km and beyond 1 km. Therefore, the time-frequency

median filter combined with the multiple-input noise canceller yields a real improvement on this data example.

Finally, we apply the second order high-pass Butterworth filter to the second data set to remove the low-frequency content. Again, the choice of a larger cutoff frequency or a larger filter order would remove too much seismic signals, as well as successive applications of this high-pass filter that attenuate certain seismic reflections. Figure 6.12 (a and b) shows the output gather and its residual plot, respectively. The high-pass filter presents the same limitations as the multiple-input adaptive noise canceller concerning the attenuation of high amplitude swell noise. This last method, nevertheless, better attenuates specific traces such as at offsets 0.25 km and 1.2 km, but also attenuates the primary reflection as well. Thus, even if the high-pass Butterworth filter presents the advantage of having a very small computational complexity, it yields some artefact that can be partly resolved by the multiple-input adaptive noise canceller and even better, by the combination of the time-frequency median filter with the multiple-input noise canceller, at the cost however of a larger computational complexity. Finally, we notice that none of the tested methods are able to attenuate the seismic interferences contained in the input data set.
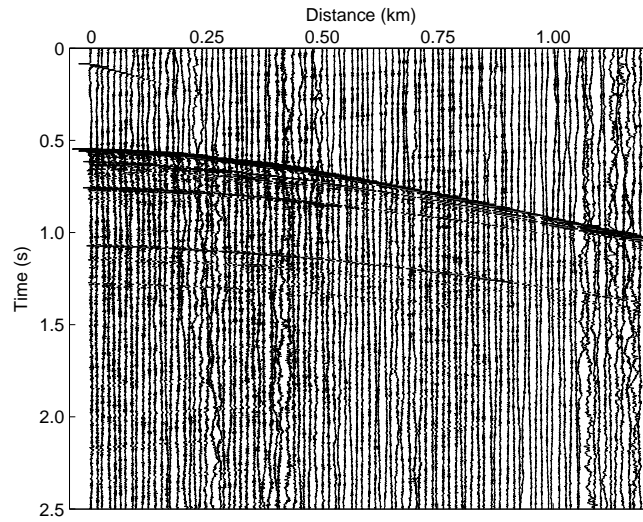
6.3.4   Application to seismic interferences attenuation

The multiple-input noise canceller, in its present form, can also be used for the attenuation of seismic interferences when applied along the dip of the interferences. It basically consists of "rotating" the input gather so that the interferences are aligned vertically, as a normal trace usually is. We then pick-up a few trace segments containing interferences that will be considered as noise sequences, run the multiple-input noise canceller for all the rotated traces, and finally rotate them back to their original configuration.
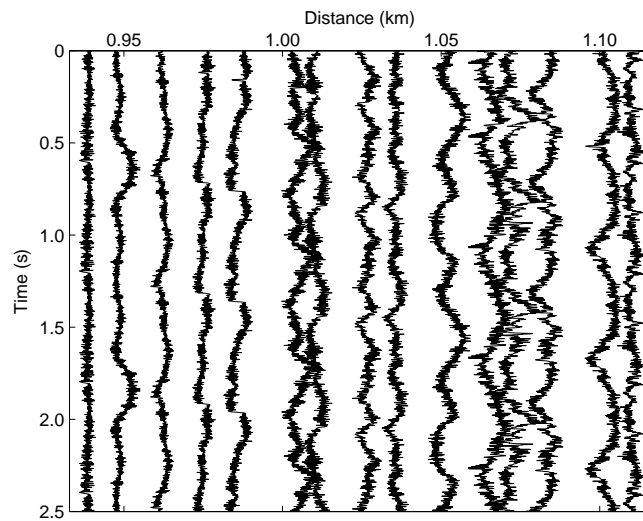
Let us now consider the real data set used in the second example shown in Figure 6.8 (a), and apply the multiple-input noise canceller to attenuate the rather strong seismic interferences recorded after 1.5 s, or the smaller ones recorded before 0.5 s. The dip of the interferences is about 2 ms upward per trace. To form the input noise sequences, we extract 9 rotated traces whose original time location is around 2 s at offset 0 km. Frequency thresholds are set to $\Phi_1 = 125$ Hz and $\Phi_2 = 75$ Hz, and the normalized step size values are $\beta_0 = 0.00005$, $\beta_1 = 0.01$ and $\beta_2 = 0.02$. The other parameters are left unchanged. We apply this filter to the output of time-frequency median filter combined with the multiple-input noise canceller in Figure 6.11 in order to evaluate the best possible result that can be obtained.

Figure 6.13 (a) shows the output gather and (b) the residual plot, i.e., the difference between the input gather shown in Figure 6.8 (a) and the output gather. Only some of the strongest interferences located after 1 s have been been attenuated, while a tiny part of the primary reflection has been attenuated as well. By increasing the value of β_2, we experienced that the attenuation of seismic interferences is more significant but at the cost of a more important attenuation of the primary reflection. Note that no other

(a)

(b)

Figure 6.8: Second example: (a) input marine shot gather and (b) 15 extracted input noise sequences.

seismic reflections are attenuated. We think however that a possible way to overcome this problem consists of applying the multiple-input noise canceller to a limited section of the data that does not include the primary reflection. The attenuation of seismic interferences is more tricky than the attenuation of swell noise, as the frequency spectra of the seismic interferences and the seismic reflections are relatively similar. Thus, using the instantaneous frequency as a metric that can differentiate the seismic interferences
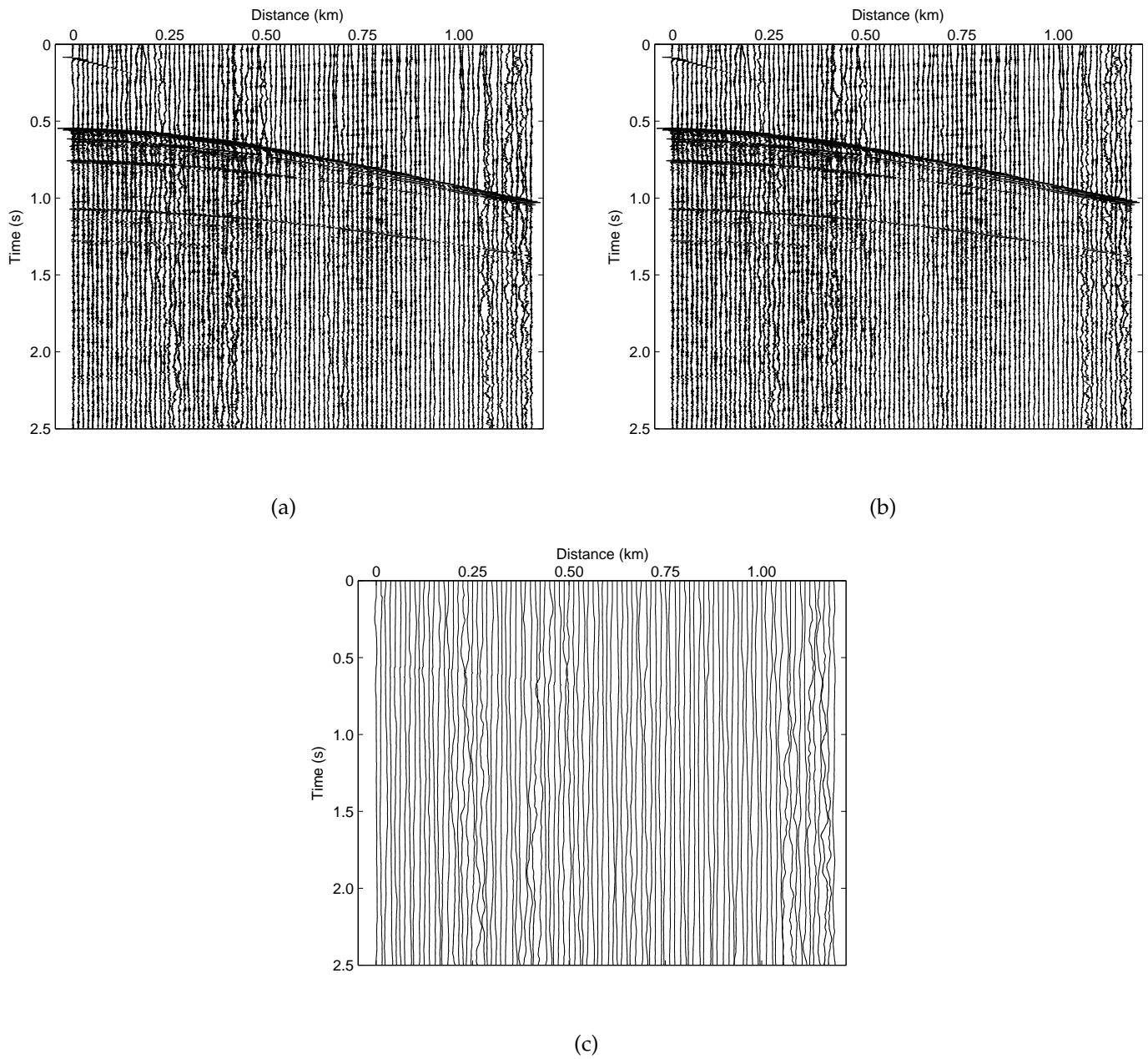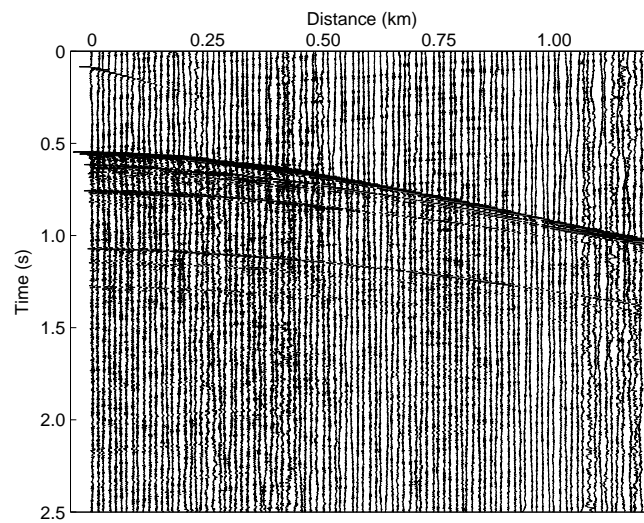
(a)

(b)

(c)

Figure 6.9: Multiple-input noise canceller: output shot gather after (a) one application of the filter and (b) three applications of the filter. (c) Residual plot after three applications of the filter.

from standard seismic reflections is no longer optimal, and therefore the efficiency is limited.

(a)
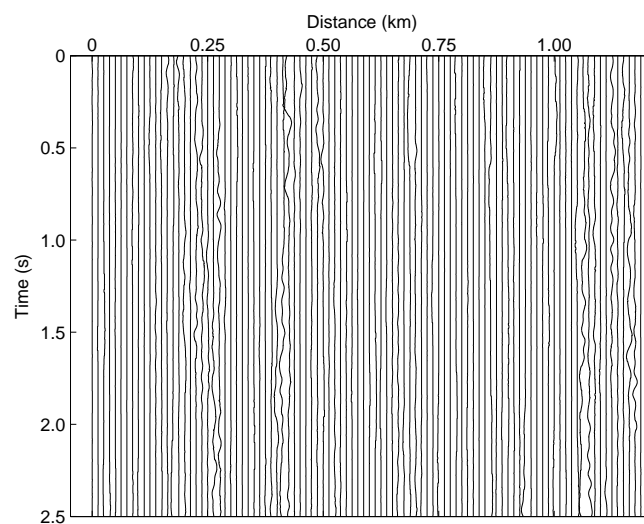
(b)

Figure 6.10: Time-frequency median filtering: (a) output shot gather and (b) residual plot.

6.4 Discussion

The user effort to set up the parameters is dependent on the data set to be treated. The key parameters are the instantaneous frequency threshold values, Φ_1 and Φ_2, and the normalized step size values β_0, β_1 and β_2. Here, β_0 is usually low to keep high frequency components unchanged, while β_1 and β_2 are much larger and control the rate
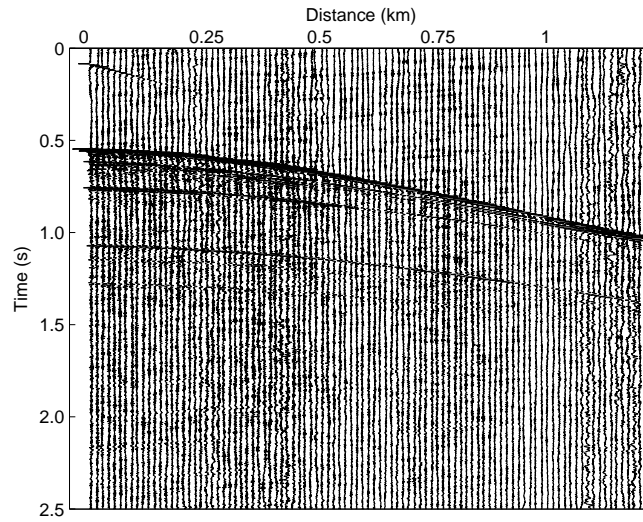
(a)
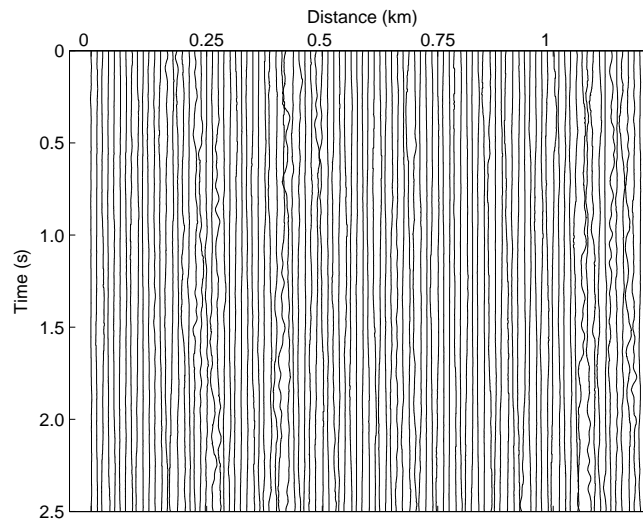
(b)

Figure 6.11: Time-frequency median filtering + multiple-input noise canceller: (a) output shot gather and (b) residual plot.

of convergence, i.e., the level of signal attenuation. In the case where the swell noise is localized within the input gather and characterized by high amplitudes low-frequency signals that are uncorrelated with the seismic reflections, as in the first data example, the instantaneous frequencies of the noise and the seismic reflections differ significantly. Therefore, the threshold values Φ_1 and Φ_2 are easy to find, typically corresponding to a high percentage of all instantaneous frequency values. In the case where the swell
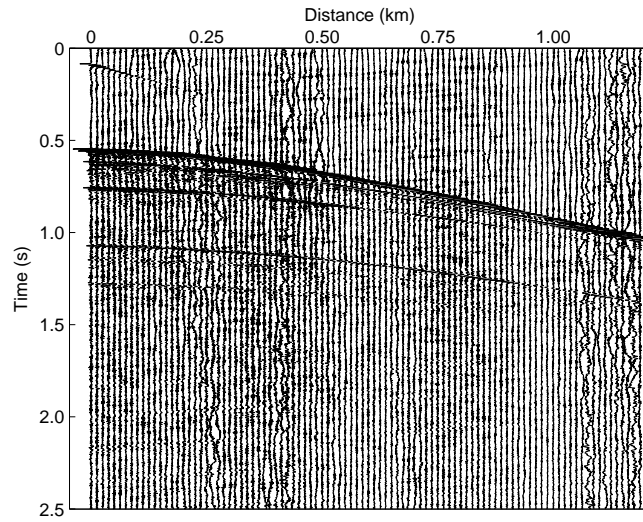
(a)
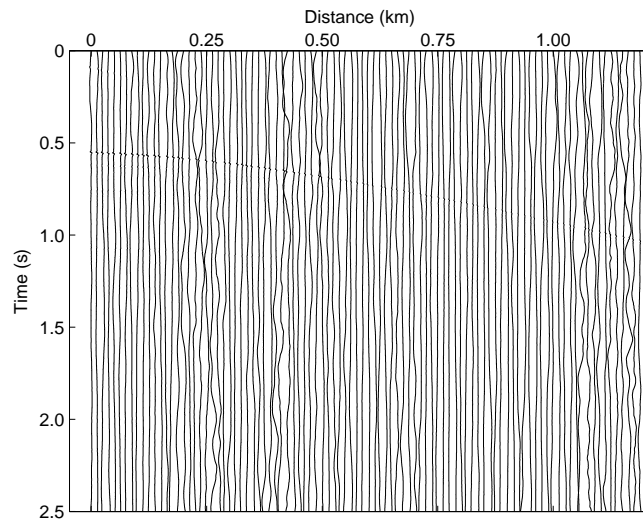
(b)

Figure 6.12: High-pass Butterworth filter: (a) output shot gather and (b) residual plot.

noise is somewhat correlated with seismic reflections, as in the second data example, the determination of the threshold value parameters is more tricky and requires a few trial-and-errors to find optimal values.

The computational cost of the multiple-input adaptive noise canceller is satisfactory. This method is definitely more complex than a standard second order high-pass filter, but it requires less operations than the time-frequency median filter, and it is therefore, less computationally demanding. The main reason is that the multiple-input adaptive
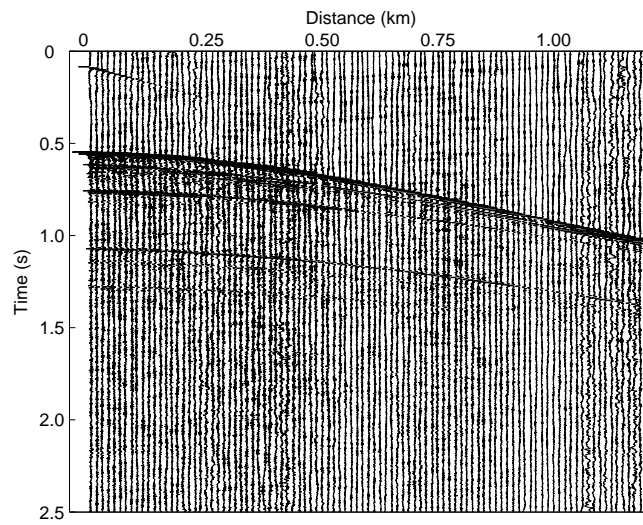
(a)
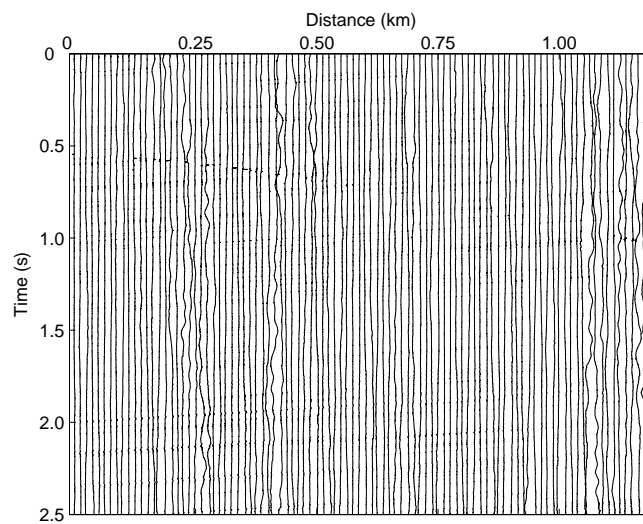
(b)

Figure 6.13: Multiple-input noise canceller application for seismic interferences attenuation: (a) output shot gather and (b) residual plot.

noise canceller treats all the traces independently and does not perform any trace-to-trace comparison as the time-frequency median filter does.

6.5 Summary

We have proposed a method to remove coherent noise that utilizes the filter structure of the multiple-input adaptive noise canceller, and an adaptive normalized least mean squares algorithm with variable step size to resolve the tap-weights coefficients. This method has been compared to the time-frequency median filter and a second-order high-pass Butterworth filter.

We showed that the use a variable step size is a necessary condition for this method to adapt to the changing statistics of the data and therefore, successfully attenuate coherent noise. By choosing the appropriate set of normalized step size values, the multiple-inputs adaptive noise canceller achieves a very good attenuation of coherent noise while preserving the seismic reflections in the case where the swell noise is localized within the input gather and characterized by high amplitudes low-frequency signals that are uncorrelated with the seismic reflections. In such a case, we showed that this method was more efficient than the time-frequency median filter by its ability to attenuate coherent noise even when the neighbor traces are corrupted by similar noise. However, we found that when the low-frequency noise is somewhat correlated with the seismic reflections, the multiple-input adaptive noise canceller affects mostly the small noise amplitudes whereas the time-frequency median filtering was better at attenuating the highest noise amplitudes. We showed then that the combination of the time-frequency median filter with the multiple-input adaptive noise canceller yielded significant improvements. Thus, the multiple-input adaptive noise canceller is a powerful and efficient filter to attenuate swell noise, that can be used either by itself or in combination with the time-frequency median filter, depending of the noise configuration. In addition, we demonstrated in both cases that the high-pass Butterworth filter had limited effects on coherent noise attenuation. Finally, we showed that the multiple-input adaptive noise canceller could find some applications for seismic interferences attenuation, after manipulation of the input data set. However, because the frequency spectra of the seismic interferences and the seismic reflections are relatively similar, the efficiency of this filter is limited.

Finally, this method is relatively easy to implement, it has an acceptable computational cost, and the user effort to set up the parametrization is reasonable. However, one requirement for this method is that the reference noise is available and well correlated with the noise contained in the primary channel. As future work, we may consider its application in the frequency domain.

Chapter 7

Conclusions and suggestions for future work

First, the statistical modeling and analysis of seismic background noise have been investigated. They have revealed that the seismic noise is almost symmetric and slightly platykurtic, and that it yields a composite power-law spectrum. Furthermore, we have proposed a clustering procedure of the noise level as an original quality estimation tool for geophysicists. Thus, the cluster matrix emphasizes defective channels as well as the inhomogeneity of seismic noise along a streamer, and in particular a higher noise level at the head of the streamer and nearby the units hanged to the streamer.

In Chapter 3, we focused on single-azimuth stacking methodologies. We have proposed a novel enhanced stacking method that utilizes local correlation between each individual trace and a chosen reference trace, as a measure of weight, and a new weight normalization scheme. Three different reference traces have been proposed based on (i) conventional stacking, (ii) weighed-stacking using S/N-estimation weights, and (iii) Kalman filter. The use of a Kalman filter is an innovative approach for this application as well as the weight normalization scheme in the enhanced local correlation stacking method. We showed that the use of S/N-estimation reference trace and even more significantly, the use of Kalman reference trace yield consistently superior results compared to conventional stacking. In particular, the superior reference traces exhibit cleaner and better defined reflected events, as well as the ability to see a larger number of reflections in deep waters. We documented that the use of an S/N-estimation reference trace has the ability to enhance coherent signals recorded by short offset traces, while a Kalman reference trace has the ability to emphasize any coherent signal recorded by any traces. For this reason, misalignment does not degrade S/N-estimation reference trace as much as Kalman reference trace. Furthermore, we showed that the enhanced stacking method yields superior results compared to conventional stacking. The best combination is obtained when using S/N estimation as reference method. It exhibits more reflected events, and in our application on sub-salt structure, it reveals a few continuous reflected events under the sub-salt structure. However, we showed that enhanced stacking emphasizes the strongest reflections while discriminating weaker

ones, which deteriorates the longitudinal homogeneity of the seismic image. Thus, the use of Kalman filter reference method should be preferred to enhanced stacking using S/N estimation as reference method, since it achieves a better overall seismic image contrast and reveals many more, cleaner and better defined reflected events. However, this result is obtained at the cost of a higher noise level and a larger processing time. Therefore, choosing between one of these two methods is a trade-off between more revealed information and cleaner seismic image. As future work, we suggest to explore the use of mutual information as a generalized statistical measure of signal correlation. Another possible improvement of the presented methods consists of deriving S/N estimates locally instead of globally to improve the S/N-estimation method. Equivalently, deriving the trace amplitudes locally instead of globally should improve the Kalman filter method.

In Chapter 5, we presented an innovative application of sparse code shrinkage and independent component analysis for signal to noise ratio enhancement of seismic signals. More specifically, we proposed to use the enhanced stacking method presented in Chapter 3 applied locally to obtain the "noise-free" realization of the data that allows the derivation of data-driven shrinkages. We showed that using data-fitted shrinkages is a necessary condition for sparse code shrinkage to be a valuable alternative as signal enhancement method. In addition, we investigated several density models, either non-parametric or parametric. Thus, we have derived the shrinkage theory for the non-parametric Gaussian kernel density estimate and for the parametric models using the sparse, the normal inverse Gaussian and the mixture of generalized Gaussian densities. We showed that sparse code shrinkage is an efficient method that targets the noise components to shrink them while preserving the signals. By its ability to enhance signals in both shallow and deep waters, the normal inverse Gaussian density appears to be the best choice of density. However, the sparse density, the mixture of generalized Gaussian density and non-parametric Gaussian kernel estimate have proven to be acceptable choices as well. The flexible parametrization of the normal inverse Gaussian density allows it to model a larger number of unimodal densities, and therefore, it suits seismic data very well. In addition, the level of attenuation of the sparse coded data can be easily adjusted by the user through the noise variance parameter. Finally, the comparison of sparse code shrinkage with local SVD and *f-x* deconvolution has shown that sparse code shrinkage was an efficient method for removing background noise and estimating signals in both shallow and deep waters. However, this is achieved at the cost of a higher computational load, mostly because of the determination of the ICA transformation matrix. As future work, we suggest the investigation of more elaborate "noise-free" realizations, as the signal reconstruction achieved by sparse code shrinkage may be enhanced by improving the quality of signals contained in this realization. In addition, some other super-Gaussian density models can be investigated, and in particular, the generalized Laplacian density and the mixture of two normal inverse Gaussian might good candidates.

Finally, we presented in Chapter 6 an application of the multiple-input adaptive noise canceller to the attenuation of flow-generated nonstationary coherent noise and

seismic interferences. This filter uses a normalized least mean squares algorithm with a variable normalized step size. We showed that the use a variable step size is a necessary condition for this method to adapt to the changing statistics of the data, and therefore successfully attenuates coherent noise. The determination of the variable step size parameter is not however an easy task and classical methods fail because they cannot adapt to the characteristics of the seismic noise. In this chapter, we proposed to derive this parameter as a function of instantaneous frequency for each seismic trace, as it is a metric that differentiates relatively well the seismic reflections from the coherent low-frequency noise. The comparison with a time-frequency median filter and a second-order high-pass Butterworth filter revealed that the multiple-input adaptive noise canceller achieves a very good attenuation of coherent noise while preserving the seismic reflections in the case of localized, high amplitudes swell noise that is uncorrelated with the seismic reflections. This method is indeed more efficient than the time-frequency median filter by its ability to attenuate coherent noise even when the neighbor traces are corrupted by similar noise. However, we found that when the low-frequency noise is somewhat correlated with the seismic reflections, the multiple-input adaptive noise canceller affects mostly the small noise amplitudes whereas the time-frequency median filtering is better at attenuating the highest noise amplitudes. We showed then that the combination of the time-frequency median filter with the multiple-input adaptive noise canceller yields significant improvements. Thus, the multiple-input adaptive noise canceller is a powerful and efficient filter to attenuate swell noise, that can be used either by itself or in combination with the time-frequency median filter, depending of the noise configuration. In addition, we demonstrated in both cases that the high-pass Butterworth filter has limited effects on coherent noise attenuation. Finally, we showed that the multiple-input adaptive noise canceller could find some applications for seismic interferences attenuation, after manipulation of the input data set. However, because the frequency spectra of the seismic interferences and the seismic reflections are relatively similar, the efficiency of the filter remains limited. As future work, we may consider the derivation of such filter in the frequency domain.

Appendix

Proof of equation 5.8: $f(x|s) = f_\nu(x - s)$

Consider the model

$$x = s + \nu$$

where s and ν are two independent random variables. The change of variables (s, x) to (u, v) is defined by the linear transform G such that

$$\begin{cases} u = x - s \\ v = s, \end{cases}$$

which is equivalent to

$$\begin{cases} x = u + v \\ s = v. \end{cases}$$

The probability density of the new variables, $f_{u,v}$, is given by

$$f_{u,v}(u, v) = f_{s,x}(G^{-1}(u, v))|J^{-1}|$$

where J^{-1} is the inverse of the Jacobian matrix given by

$$\begin{pmatrix} \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \\ \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Thus, $|J^{-1}| = 1$ and $f_{u,v}(u, v) = f_{s,x}(G^{-1}(u, v)) = f_{s,x}(v, u + v)$.
Using again the variables (s, x) instead of (u, v) in the latest equation, we obtain

$$f_{s,\nu}(s, \nu) = f_{s,x}(s, x).$$

Moreover, s and ν are independent, which yields

$$f_{s,\nu}(s, \nu) = f_s(s)f_\nu(\nu) = f_{s,x}(s, x).$$

Bayes rule for densities gives

$$f_{s,x}(x|s) = \frac{f_{s,x}(s, x)}{f_s(s)} = \frac{f_s(s)f_\nu(\nu)}{f_s(s)} = f_\nu(\nu) = f_\nu(x - s).$$

Finally, by dropping the subscripts s and x, we obtain

$$f(x|s) = f_\nu(x - s).$$

Bibliography

T. W. Anderson. *Asymptotic theory for principal component analysis. Annals of Mathematical Statistics*, **34**: 122–148, 1963.

H. Andrews and C. Patterson. *Singular value decomposition and digital image processing. IEEE Transactions on Acoustics, Speech, and Signal Processing*, **24**: 26–53, 1976.

J. C. Bancroft, H. D. Geiger and G. F. Margrave. *The equivalent offset method of prestack time migration. Geophysics*, **63**(6): 2042–2053, 1998.

H. B. Barlow. *What is the computational goal of the neocortex?*. In *C. Koch, and J. L. Davis, eds., Large-Scale Neuronal Theories of the Brain*, pp. 1–22. MIT Press, Cambridge, MA, 1994.

J. W. Bayless and E. O. Brigham. *Application of the Kalman filter to continuous signal restoration. Geophysics*, **35**(1): 2–23, 1970.

J. B. Bednar. *Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data. Geophysics*, **48**(12): 1598–1610, 1983.

M. Bekara and M. Van der Baan. *Local singular value decomposition for signal enhancement of seismic data. Geophysics*, **72**: V59–V65, 2007.

M. Bekara, A. Ferreira and M. Van der Baan. *A statistical technique for high amplitude noise detection: Application to swell noise attenuation. SEG Technical Program Expanded Abstracts*, **27**(1): 2601–2605, 2008.

G. Bianchi and R. Sorrentino. *Electronic Filter Simulation and Design*. McGraw-Hill Professional, 2007.

B. Boashash. *Estimating and interpreting the instantaneous frequency of a signal - part 1: Fundamentals. Proceedings of the IEEE*, **80**(4): 520–538, 1992.

P. Brockett, M. Hinich and G. Wilson. *Nonlinear and non-Gaussian ocean noise. Journal of the Acoustical Society of America*, **82**: 1386–1394, 1987.

T. Bronez. *On the performance advantage of multitaper spectral analysis. IEEE Trans. Signal Proc.*, **40**(12): 2941–2946, 1992.

L. L. Canales. *Random noise reduction. SEG Technical Program Expanded Abstracts*, **3**(1): 525–527, 1984.

C. Chui. *An Introduction to Wavelets*. Academic Press, 1992.

J. Claerbout. *Fundamentals of Geophysical Data Processing*. McGraw-Hill, 1976.

W. S. Cleveland. *Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association*, **74**: 829–836, 1979.

L. Cohen. *Time-frequency analysis*. Prentice Hall PTR, Upper Saddle River, N.J, 1995.

P. Comon. *Independent component analysis-a new concept?. Signal Processing*, **36**: 287–314, 1994.

N. D. Crump. *A Kalman filter approach to the deconvonlution of seismic signals. Geophysics*, **39**(1): 1–13, 1974.

C. D'Agosto, K. Marfurt and J. Steven. *Modeling and removal of ground roll from horizontal component of C-waves. SEG Technical Program Expanded Abstracts*, **22**(1): 1981–1983, 2003.

D. L. Donoho, I. M. Johnstone, G. Kerkyacharian and D. Picard. *Wavelet shrinkage: Asymptopia?. Journal of the Royal Statistical Society, Series B*, **57**: 301–337, 1995.

B. Efron and C. Morris. *Data analysis using Stein's estimator and its generalizations. J. American Statistical Association*, **70**: 311–319, 1975.

T. Elboth, I. V. Presterud and D. Hermansen. *Time-frequency seismic data de-noising. Geophysical prospecting*, 2010.

S. L. M. Freire and T. J. Ulrych. *Application of singular value decomposition to vertical seismic profiling. Geophysics*, **53**(6): 778–785, 1988.

J. Gerbrands. *On the relationship between SVD, KLT, and PCA. Pattern Recognition*, **14**(6), 1981.

G. Golub and C. van Loan. *Matrix computations, Third Edition*. The Johns Hopkins University Press, London, 1996.

R. C. Gonzalez and R. E. Woods. *Digital image processing, Third Edition*. Prentice Hall, 2008.

J. C. Gower. *Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika*, **53**: 325–338, 1966.

S. Grion and A. Mazzotti. *Stacking weights determination by means of SVD and cross-correlation. SEG Technical Program Expanded Abstracts*, **17**(1): 1135–1138, 1998.

S. L. Hahn. *Hilbert transforms in signal processing*. Artech House, Inc., Boston, 1996.

J. B. U. Haldorsen and P. A. Farmer. *Suppression of high-energy noise using an alternative stacking procedure. Geophysics*, **54**(2): 181–190, 1989.

A. Hanssen and T. A. Øigård. *The normal inverse Gaussian distribution: a versatile model for heavy-tailed stochastic processes*. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pp. 3985–3988. Salt Lake City, UT, May 2001a.

A. Hanssen and T. A. Øigård. *The normal inverse Gaussian distribution, as a flexible model for heavy tailed processes*. In *Proceedings IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*. Baltimore, Maryland, USA, June 2001b.

M. H. Hayes. *Statistical digital signal processing and modeling*. Wiley, 1996.

S. Haykin. *Adaptive Filter Theory, 4th Edition*. Prentice Hall, September 2001.

S. Helgason. *The Radon Transform (Progress in Mathematics), Second Edition*. Birkhäuser, 1999.

C. Hemon and D. Mace. *Essai d'une application de la transformation de Karhunen-Loève au traitement sismique. Geophysical Prospecting*, **26**: 600–626, 1978.

H. Hotelling. *Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology*, **24**: 417–441, 498–520, 1933.

A. Hyvärinen and E. Oja. *A fast fixed-point algorithm for independent component analysis. Neural Computation*, **9**: 1483–1492, 1997.

A. Hyvärinen, P. Hoyer and E. Oja. *Sparse code shrinkage: Denoising of nongaussian data by maximum likehood estimation. Neural Computation*, **11**: 1739–1768, 1999.

J. N. R. Jeffers. *Two case studies in the application of principal component analysis. Applied Statistics*, **16**: 225–236, 1967.

R. Jenssen, T. A. Øigård, T. Eltoft and A. Hanssen. *Sparse code shrinkage based on the normal inverse Gaussian density model*. In *Proceedings of International Workshop on Independent Component Analysis and Blind Signal Separation*, pp. 212–217. San Diego, 2001.

I. Jolliffe. *Principal Component Analysis, Second Edition*. Springer, 2002.

S. T. Kaplan, M. D. Sacchi and T. J. Ulrych. *Sparse coding for data-driven coherent and incoherent noise attenuation. SEG Technical Program Expanded Abstracts*, **28**(1): 3327–3331, 2009.

M. Kendall and A. Stuart. *The advanced theory of statistics*. Charles Griffin & Company, 1958.

K. Kreutz-Delgado, B. Rao, K. Engan, T. W. Lee and T. Sejnowski. *Convex/schur-convex (CSC) log-priors and sparse coding*. In *6th Joint Symposium on Neural Computation*, pp. 65–71. 1999.

R. Kwong and E. Johnston. *A variable step size* LMS *algorithm. IEEE Transactions on Signal Processing*, **40**: 1633–1642, 1992.

G. Liu, S. Fomel, L. Jin and X. Chen. *Stacking seismic data using local correlation. Geophysics*, **74**: V43–V48, 2009.

W. Lu. *Adaptive multiple subtraction using independent component analysis. Geophysics*, **71**(5): S179–S184, 2006.

Y. Luo, S. Al-Dossary, M. Marhoon and M. Alfaraj. *Generalized Hilbert transform and its applications in geophysics. The Leading Edge*, **22**(3): 198–202, 2003.

D. Manolakis, V. Ingle and S. Kogon. *Statistical and Adaptive Signal Processing*. Artech House, 2005.

J. Mari, F. Glangeaud and F. Coppens. *Signal processing for geologists and geophysicists*. Technip, Paris, 1999.

Matlab®. The MathWorks, Inc., Boston, MA, R2010a edition, 2010.

W. H. Mayne. *Common reflection point horizontal data stacking techniques. Geophysics*, **27**(6): 927–938, 1962.

O. M. M. Mohamed and M. Jaïdane-Saïdane. *On the parameters estimation of the generalized Gaussian mixture model*. In *Proceedings of the* 17th *European Signal Processing Conference (EUSIPCO 2009)*. Glasgow, Scotland, August 2009.

R. Neelamani, T. A. Dickens and M. Deffenbaugh. *Stack-and-denoise: A new method to stack seismic datasets. SEG Technical Program Expanded Abstracts*, **25**(1): 2827–2831, 2006.

T. A. Øigård, A. Hanssen, R. Hansen and F. Godtliebsen. EM-*estimation and modeling of heavy-tailed processes with the multivariate normal inverse Gaussian distribution. Signal Processing*, **85**: 1655–1673, 2005.

B. A. Olshausen and D. J. Field. *Natural image statistics and efficient coding. Network*, **7**(2): 333–339, 1996.

A. Özbek. *Multichannel adaptive interference cancelling. SEG Technical Program Expanded Abstracts*, **19**(1): 2088–2091, 2000.

K. Pearson. *On lines and planes of closest fit to systems of points is space. Philosophical Magazine*, **2**: 559–572, 1901.

D. Percival and A. T. Walden. *Spectral Analysis for Physical Applications: Multivariate and Conventional Univariate Techniques*. Cambridge: Cambridge Univ. Press, 1993.

L. A. Pflug, P. Jackson, G. E. Ioup and J. W. Ioup. *Variability in higher order statistics of measured shallow-water shipping noise*. In *Proceedings of the 1997 IEEE Signal Processing Workshop on Higher-Order Statistics (SPW-HOS '97)*, p. 400. IEEE Computer Society, Washington, DC, USA, 1997.

R. C. Pruett. *Long period multiple reflection suppression and enhanced velocity discrimination using a weighted stack*. SEG Technical Program Expanded Abstracts, **1**(1): 10–12, 1982.

C. Rao. *The use and interpretation of principal component analysis in applied research. Sankhya A*, **26**: 329 –358, 1964.

M. A. Rashed. *Smart stacking: A new CMP stacking technique for seismic data. The Leading Edge*, **27**(4): 462–467, 2008.

P. Ready and P. Wintz. *Information extraction,* SNR *improvement, and data compression in multispectral imagery*. IEEE Transactions on Communications, **21**(10): 1123 – 1131, October 1973.

E. Rietsch. *Estimation of the signal-to-noise ratio of seismic data with an application to stacking. Geophysical Prospecting*, **28**: 531–550, 1980.

J. C. Robinson. *Statistically optimal stacking of seismic data. Geophysics*, **35**(3): 436–446, 1970.

H. C. Romesburg. *Cluster Analysis for Researchers*. Lulu press, 2004.

M. D. Sacchi and Signal Analysis and Imaging Group. *SeismicLab*. http://www-geo.phys.ualberta.ca/saig/SeismicLab, accessed 9 July 2010, 2008.

C. Sanchis and A. Hanssen. *Statistical analysis of noise in towed streamer arrays*. In *Proceedings of the 2008 European Signal processing conference (EUSIPCO 2008)*, pp. 200–205. Lausanne, Switzerland, August 2008.

W. A. Schneider. *Integral formulation for migration in two and three dimensions. Geophysics*, **43**(1): 49–76, 1978.

M. Schoenberger and J. F. Mifsud. *Hydrophone streamer noise. Geophysics*, **39**(6): 781–793, 1974.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1986.

D. Slepian. *Prolate spheroidal wave functions,* Fourier analysis, and uncertainty - part V*: Th*e discrete case. Bell System Technical Journal*, **57**: 1371–1430, 1978.

S. Solbø and T. Eltoft. *Homomorphic wavelet-based statistical de-speckling of* SAR *images.* IEEE *Transactions on Geoscience & Remote Sensing,* **42**: 711–721, April 2004.

A. Tarantola. *Inverse problem theory.* Elsevier Science Publ. Co., Amsterdam, 1987.

D. J. Thomson. *Spectrum estimation and harmonic analysis.* In *Proc. IEEE*, volume 70 of *Institute of Electrical and Electronics Engineers, Inc. Conference*, pp. 1055–1096. 1982.

S. Trickett. *Maximum-likelihood-estimation stacking.* SEG Technical Program Expanded Abstracts, **26**(1): 2640–2643, 2007.

T. J. Ulrych and M. Sacchi. *Volume 36: Information-based Inversion and Processing with Applications (Handbook of Geophysical Exploration: Seismic Exploration).* Elsevier, 2005.

T. J. Ulrych, S. Freire and P. Siston. *Eigenimage processing of seismic sections. SEG Technical Program Expanded Abstracts,* **7**(1): 1261–1265, 1988.

M. Van der Baan. PP/PS W*avefield separation by independent component analysis. Geophysical Journal International,* **166**(1): 339–348, 2006.

V. D. Vrabie, J. I. Mars and J.-L. Lacoume. *Modified singular value decomposition by means of independent component analysis. Signal Processing,* **84**: 645–652, 2004.

M. Wand and M. Jones. *Kernel Smoothing.* Chapman & Hall, 1995.

J. Wang, F. Tilmann, R. S. White, H. Soosalu and P. Bordoni. *Application of multichannel Wiener filters to the suppression of ambient seismic noise in passive seismic arrays. The Leading Edge,* **27**(2): 232–238, 2008.

T. Watt and J. B. Bednar. *Role of the alpha-trimmed mean in combining and analyzing seismic common-depth-point gathers. SEG Technical Program Expanded Abstracts,* **2**(1): 276–277, 1983.

B. Widrow and S. D. Stearns. *Adaptive signal processing.* Prentice-Hall, Inc., Engelwood Cliffs, NJ, USA, 1985.

S. Zhao, Z. Man, S. Khoo and H. R. Wu. *Variable step-size* LMS *algorithm with a quotient form. Signal Processing,* **89**(1): 67–76, 2009.