Observation-Driven Method Based on IIR Wiener Filter for Microseismic Data Denoising

NAVEED IQBAL,[1] AZZEDINE ZERGUINE,[1] SANLINN KAKA,[1] and ABDULLATIF AL-SHUHAIL[1]

Abstract—Reliable analysis of low-energy earthquakes (microseismic) depends on how accurately one can detect and pick the arrival times, which are strongly influenced by the noise content. The study of microseismic events becomes even more challenging when the sensors are located on the surface because of the poor signal-to-noise ratio (SNR). Consequently, efficient and robust techniques for denoising microseismic data are necessary. In this study, we propose a method based on an infinite impulse response (IIR) Wiener filter. The proposed method uses statistics based on signal observations (noisy data) and the underlying noise, both recorded by various sensors. The method presented here precludes the need for statistics or prior knowledge of the signal of interest. The second-order statistics of the noise and the noisy data are extracted from the recorded data only. As an advantage, in deriving the filter's impulse response, no underlying structure of noise is assumed. Therefore, our method works for various types of noise, e.g., uncorrelated, spatially correlated, temporally correlated, Gaussian and non-Gaussian noise. Hence, the proposed method can be suitable as well for microseismic data recorded in diverse seismic noise environments. The criteria used to optimize the filter impulse response is the minimization of the mean square error. The proposed method is tested on synthetic and field data sets and found to be effective in denoising microseismic data with very low SNR (−12 dB).

Key words: Microseismic/microearthquake data, IIR filter, Wiener filter, signal-to-noise ratio, autocorrelation.

1. Introduction

A microseismic event is considered to be a small magnitude earthquake, having a magnitude as low as −3 (Maxwell et al. 2008). It can occur as a natural phenomenon or as a result of human activities within the earth. Analysis of artificially induced microseismic events is essential for oil and gas reservoir geophysics (e.g., Kendall et al. 2011) and in geologic carbon dioxide storage (e.g., Verdon 2011). Both these examples of applications of microseismic monitoring are related to conventional reservoirs. However, microseismic analysis is used extensively in unconventional reservoirs as well for imaging the fracture networks. Moreover, microseismic monitoring has been common in the mining industry for over 100 years where it is primarily used for safety from rockbursts and assessing the state of stress within a mine (Mendecki 1993; Castellanos and van der Baan 2013) in the study of water reservoir-induced seismicity for at least 5 decades (Simpson et al. 1988) and in the geothermal industry (Pearson 1981), but its application in the oil and gas industry is relatively new. High-pressure fluid is typically injected to fracture the rock and increase permeability, thus enhancing production. During such a process, microfractures can be induced in the vicinity of the injection well. Monitoring and analyzing these microfractures help understand the rock-breaking mechanisms during the injection process and reservoir exploitation (e.g., Maxwell 2011). To locate the microseismic hypocenters, the accurate identification of microseismic events is crucial.

To monitor these microseismic events, typically 8–12 three-component sensors are placed in nearby wells or on the earth's surface (Eaton et al. 2014; Caffagni et al. 2016). Since it is more cost effective to place geophones on the surface than burying them deeply in a borehole, several hundred sensors can be used in surface arrays (Duncan 2012). Moreover, surface arrays offer other advantages as well. For example, it is known that the accuracy and precision of the hypocenter locations in microseismic monitoring depend on both the signal-to-noise ratio (SNR) of the recorded data and the spatial distribution of the receivers. Usually, downhole monitoring provides better detection because of a higher SNR; however,

[1] Center for Energy and Geo Processing, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. E-mail: naveediqbal@kfupm.edu.sa

the precise location of events might be difficult, especially in the case of a single monitoring well (Eisner et al. 2010). Unlike the accurate depth estimation, epicentral errors for microseismic event locations using downhole arrays increase as a function of distance from the monitoring well. On the other hand, although surface monitoring often suffers from low SNR, the ability to place receivers in multiple azimuths and offsets allows for precise epicenter (horizontal) event location (Mousavi et al. 2016a). The surface microseismic data are characterized by a low SNR (Shemeta and Anderson 2010); consequently, the main challenge in the study of microseismic events is to enhance the SNR by suppressing/removing the noise.

Generally, several denoising or SNR enhancement methods exist in the literature. Seismic interferometry is a well-known technique to enhance the SNR of the seismic data record, which includes cross-correlation, stacking and convolution (Al-shuhail et al. 2012; Mallinson et al. 2011; Bharadwaj et al. 2013). A different approach that allows the reconstruction of signals from noisy observations is based on time-frequency analysis (Mousavi and Langston 2016a; Vera Rodriguez et al. 2012; Mousavi et al. 2016b). This filtering method encodes the noisy signal as the instantaneous frequency of a frequency modulated analytic signal. The signal is recovered by estimating the peak of the time-frequency distribution of the analytic signal. This approach is sensitive to the noise interferences that detract the energy concentration in time-frequency distribution. Furthermore, the wavelet transform is used to decompose the noisy signal into time-frequency components using the appropriate mother wavelet. Here, a threshold is necessary to obtain the enhanced signal. Proper selection of the mother wavelet and the number of decomposition levels are crucial for these methods. A reassignment strategy, together with pre- and post-processing steps, is added in the time-frequency based method for improving the denoising results (Mousavi and Langston 2017). A data-driven approach that derives the basis function from the noisy signal is known as empirical mode decomposition (Han and van der Baan 2015). However, for this decomposition, the basis function might not be accurate because of the strong noise that

affects the denoising results in a low SNR environment. There are other denoising methods that are based on thresholding in time-frequency transforms, e.g., the Radon transform (Sabbione et al. 2013, 2015), reduced-rank filtering (Sabbione and Velis 2013; Iqbal et al. 2016) and damped multichannel singular spectrum analysis (Huang et al. 2016).

Denoising using a Wiener filter approach is also a commonly used method in active seismic surveying and has been used for more than 4 decades (e.g., Peacock and Treitel 1969; Haldorsen et al. 1994). However, the Wiener filter method requires the knowledge of the statistics of the signal, which is not normally available in practice. A solution to this difficulty is to use the wavelet transform to partially differentiate the signal from the noise in an initial stage, and then the Wiener filter is applied after calculating the signal statistics (Aghayan et al. 2016; Kimiaefar et al. 2016). In these methods, wavelets are used to extract the high and low frequency components. The high frequency components are assumed to be noise. Of course a threshold and a basis function (mother wavelet) are needed for this purpose. Proper selection of the mother wavelet is crucial in wavelet transform, and denoising results are greatly affected by the type of wavelet. Improving the SNR of speech signals using the Wiener filter without knowing the signal statistics was proposed by Chen et al. (2006), who calculated the noise statistics using the silence intervals in speech that represent pure noise (including electronically produced noise). Similar approaches are also used in seismology by intuitively finding the noise-only part in data prior to an earthquake or controlled source occurrence (Wang et al.2008, 2009; Coughlin et al. 2014; Mousavi and Langston 2016c; Khadhraoui and Özbek 2013; Baziw and Weir-Jones 2002). Intuitively finding the noise-only part in surface microseismic data is very difficult because of the low SNR. Hence, the main challenge in microseismic surface monitoring is the realistic estimation of the seismic noise. In this study, the noise is estimated blindly, i.e., without the wavelet transform (thus avoiding having to select a proper mother wavelet) or using the assumption that the noise-only portion of the data is available and known. Mousavi and Langston (2016b) proposed minimally controlled recursive averaging in the short-time Fourier transform domain

for estimating the noise. Their method was based on the work of Martin (2001) and Cohen (2003). However, their approach requires a large adaptation time (which, unfortunately, gives incorrect locations of events), and the threshold is also fixed for all the frequencies (see Rangachari and Loizou 2006). Eventually, an IIR Wiener filter is used for the first time in a non-similar approach to that used in the FIR Wiener filter together with a noise estimation method similar to the one proposed by Rangachari and Loizou (2006).

In this study, a method is proposed based on two features of the microseismic data. First, the occurrence of microseismic events is sporadic over time. Therefore, a suitable monitoring period is necessary; hence, portions of the recorded traces are occupied by pure noise. Second, the statistical knowledge (for designing the Wiener filter) of the microseismic event is unknown in advance. Considering these two features, in this study an observation-driven denoising method based on the IIR Wiener filter is proposed. The method works by estimating the statistics of the noise and the observation (signal plus noise) from the received data without any prior knowledge of the signal or the noise statistics. The filter gives promising results when applied to synthetic, semi-synthetic and field data sets at very low SNR of -12 dB, without assuming any specific type of noise. Thus, this makes it suitable for denoising surface microseismic data with any type of noise, e.g., Gaussian, non-Gaussian, correlated, uncorrelated and coherent noise. Note that this article presents a contribution to the SNR enhancement using a data-driven IIR Wiener filter, which, to the best of our knowledge, has never been proposed before in microseismic investigation.

In the following section, a derivation of the IIR Wiener filter is presented. Next, the estimation of the autocorrelation of the noise and the noisy observations from the recorded traces, which is needed for the filter, is discussed. Finally, the proposed method is validated using synthetic and field data sets.

2. IIR Wiener Filter Design for Microseismic Denoising

The Wiener filter is used to statistically estimate the desired signal from the noisy seismic trace, usually in the case of additive noise. Typically, filters are designed for a specific frequency response. However, the Wiener filter adopts a different approach. It is required to have prior statistical knowledge of the desired signal and the noise (or the observation) and the filter is designed so that its output signal matches with the original desired signal as much as possible. With the proper design, a Wiener filter can be used to filter out the noise to get the underlying signal of interest. In this work, we will design the Wiener filter without prior statistical knowledge. To clarify this technique, let A, B and C be the desired signal, the noisy signal and the output of the Wiener filter, respectively. Mathematically, we can write the error e_r as

$$e_r = C - A.$$

The Wiener filter is designed by minimizing the mean-square error, i.e.,

$$\min E\{e_r^2\},$$

where E is the mathematical expectation (Proakis 1985).

In signal processing and mathematics, a discrete-time signal, which is a sequence of complex or real numbers, can be converted to a complex frequency domain representation using the z-transform. This is equivalent to the Laplace transform in continuous time. For more details about the Wiener filter and z-transform, see Proakis and Manolakis (2006), Sayed (2008) and Haykin (2002). In this section, we present the IIR Wiener filter, which is designed to estimate the microseismic signal from the noisy records. The analysis is carried out using the z-transform. In the ensuing, the derivation of the filter is reported.

2.1. Filter Derivation

Consider M sensors placed over a monitoring area for recording microseismic data. Each of these sensors records a time series of sampled measurements, i.e., a microseismic trace, say y_k^i, as

$$y_k^i = s_k^i + w_k^i, i = 1, 2, \ldots, M, \tag{1}$$

where s_k^i and w_k^i represent the signal and noise sample, respectively, at time instant $t = kT$ of the ith

trace (T is the sampling interval). For the derivation, the time series y_k^i, s_k^i and w_k^i are concatenated into vectors as follows:

$$\mathbf{y}_k^i = \left[y_k^i, y_{k-1}^i, y_{k-2}^i, \ldots \right]^T, \tag{2}$$

$$\mathbf{s}_k^i = \left[s_k^i, s_{k-1}^i, s_{k-2}^i, \ldots \right]^T, \tag{3}$$

$$\mathbf{w}_k^i = \left[w_k^i, w_{k-1}^i, w_{k-2}^i, \ldots \right]^T. \tag{4}$$

For our derivations, real numbers are assumed. The target is to design an IIR filter $\mathbf{g}^i$ for each trace to estimate the signal s_k^i as a linear transformation of the measurement y_k^i. The output of the filter is given as

$$\hat{s}_k^i = \mathbf{g}^i \mathbf{y}_k^i, \tag{5}$$

where $\mathbf{g}^i = \left[g_0^i, g_1^i, g_2^i, \ldots \right]$. The filter is of an infinite length duration and so is the data sequence. To estimate the filter coefficients, the mean squared error (MSE) cost function is minimized according to

$$J_k^i = E\left\{ \tilde{\mathbf{s}}_k^i \tilde{\mathbf{s}}_k^{iT} \right\}, \tag{6}$$

where the estimation error is $\tilde{\mathbf{s}}_k^i = \mathbf{s}_k^i - \hat{\mathbf{s}}_k^i$, and $E\{.\}$ and $(.)^T$ represent the mathematical expectation (which gives the most expected value, i.e., the predictor) and the transposition operation, respectively. Here, we use the fact that minimizing the MSE is equivalent to finding the solution by considering the error to be orthogonal to each of the data points in the estimation process. Equation (6) can be solved directly using the correlation of $\tilde{s}_k^i$ and $\mathbf{y}_k^i$ (Sayed 2008, p. 36), which is

$$\begin{aligned} E\left\{ \tilde{s}_k^i \mathbf{y}_k^{iT} \right\} &= E\left\{ \left[s_k^i - \hat{s}_k^i \right] \mathbf{y}_k^{iT} \right\} \\ &= E\left\{ s_k^i \mathbf{y}_k^{iT} - \mathbf{g}^i \mathbf{y}_k^i \mathbf{y}_k^{iT} \right\} \\ &= E\left\{ y_k^i \mathbf{y}_k^{iT} - w_k^i \mathbf{y}_k^{iT} - \mathbf{g}^i \mathbf{y}_k^i \mathbf{y}_k^{iT} \right\} \\ &= E\left\{ y_k^i \mathbf{y}_k^{iT} - w_k^i \mathbf{s}_k^{iT} - w_k^i \mathbf{w}_k^{iT} - \mathbf{g}^i \mathbf{y}_k^i \mathbf{y}_k^{iT} \right\} \\ &= \mathbf{p}_{yy,k}^i - \mathbf{p}_{ws,k}^i - \mathbf{p}_{ww,k}^i - \mathbf{g}^i \mathbf{P}_{yy,k}^i, \end{aligned} \tag{7}$$

where $\mathbf{p}_{ab,k}^i$ denotes the correlation between signals a and b of the ith trace. Assuming that the noise and the signal are uncorrelated, i.e., $\mathbf{p}_{ws,k}^i = 0$ and therefore, (7) looks like:

$$E\left\{ \tilde{s}_k^i \mathbf{y}_k^{iT} \right\} = \mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i - \mathbf{g}^i \mathbf{P}_{yy,k}^i. \tag{8}$$

According to the *principle of orthogonality* (Haykin 2002), i.e., the estimate $\hat{s}_k^i$, that minimizes the MSE

cost function as in Eq. (6), is the orthogonal projection of $\tilde{s}_k^i$ into the space spanned by the observations. This is equivalent to requiring $E\left\{ \tilde{s}_k^i \mathbf{y}_k^{iT} \right\} = 0$, which yields

$$\mathbf{g}^i \mathbf{P}_{yy,k}^i = \mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i. \tag{9}$$

Since Eq. (9) holds for an infinite length of filter, it can not be solved directly using a set of linear equations nor can it be solved using the z-transform because the Wiener filter is causal, i.e., $g_k^i = 0$ for $k < 0$ (Proakis and Manolakis 2006, Chapter 12). To address this issue, the noisy observation $\mathbf{y}_k^i$ will be represented by another equivalent process, $\bar{\mathbf{y}}_k^i$, and this is done by passing it through a noise-whitening filter. Mathematically, this reads as

$$\bar{y}_k^i = \mathbf{v}^i \mathbf{y}_k^i = \underbrace{\mathbf{v}^i \mathbf{s}_k^i}_{\bar{s}_k^i} + \underbrace{\mathbf{v}^i \mathbf{w}_k^i}_{\bar{w}_k^i}. \tag{10}$$

where $\mathbf{v}^i = \left[v_0^i, v_1^i, v_2^i, \ldots \right]$ is the impulse response of the whitening filter. Now $\hat{s}_k^i$ can be written as

$$\hat{s}_k^i = \mathbf{q}^i \bar{\mathbf{y}}_k^i, \tag{11}$$

where $\mathbf{q}^i = \left[q_0^i, q_1^i, q_2^i, \ldots \right]$. The IIR Wiener filter can be seen as a cascade of the whitening filter $V^i(z)$ and another filter $\mathbf{Q}^i(z)$ in the z-domain, where $V^i(z)\{\mathbf{Q}^i(z)\}$ is the z-transform of $\mathbf{v}^i\{\mathbf{q}^i\}$ (Proakis and Manolakis 2006, pp. 818–822). Application of the principle of orthogonality, $E\left\{ \tilde{s}_k^i \bar{\mathbf{y}}_k^{iT} \right\} = 0$, leads to

$$\mathbf{q}^i \mathbf{P}_{\bar{y}\bar{y},k}^i = \mathbf{p}_{y\bar{y},k}^i - \mathbf{p}_{w\bar{w},k}^i. \tag{12}$$

Since $\bar{\mathbf{y}}_k^i$ is white, therefore, $\mathbf{P}_{\bar{y}\bar{y},k}^i$ is a diagonal matrix with $\sigma_{\bar{y}}^2$ as diagonal entries and Eq. (12) becomes

$$\mathbf{q}^i = \frac{1}{\sigma_{\bar{y}}^2} \left(\mathbf{p}_{y\bar{y},k}^i - \mathbf{p}_{w\bar{w},k}^i \right). \tag{13}$$

Now define the z-domain as $\mathbf{z}^+ = [1, z^1, z^2, \ldots]$ and $\mathbf{z}^- = [1, z^{-1}, z^{-2}, \ldots]$. Hence, Eq. (13) in the z-domain is

$$\mathbf{Q}^i(z) = \frac{1}{\sigma_{\bar{y}}^2} \mathbf{z}^- \underbrace{\left(\mathbf{p}_{y\bar{y},k}^i - \mathbf{p}_{w\bar{w},k}^i \right)^T}_{C_{\bar{y}\bar{w},k}^{i+}(z)}. \tag{14}$$

$C_{\bar{y}\bar{w},k}^{i+}(z)$ represents the z-transform of the one-sided autocorrelation sequence $\mathbf{p}_{y\bar{y},k}^i - \mathbf{p}_{w\bar{w},k}^i$ (Proakis and Manolakis 2006, pp. 818–822) and
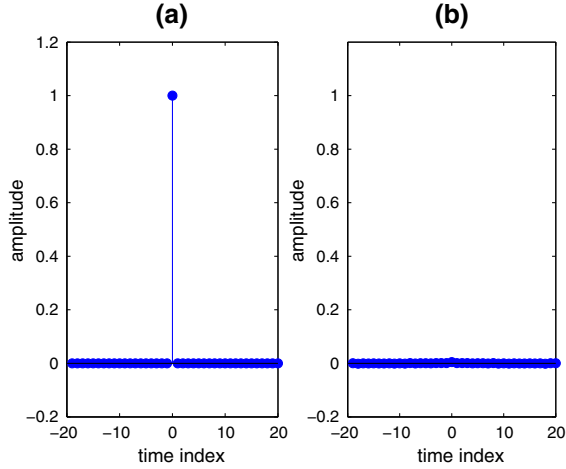
Figure 1
The response of the IIR filter is plotted for SNR $\to \infty$ and SNR $\to 0$. A Ricker wavelet of 30 Hz without noise is used for SNR $\to \infty$, which gives the filter response **a** $g_1^i = 1$, i.e., the filter offers no attenuation to the noise-free signal. Only noise is used for SNR $\to 0$, which gives the filter response, **b** $g_1^i = 0$, i.e., the filter does not allow the noise to pass

$$\mathbf{p}_{y\bar{y},k}^i - \mathbf{p}_{w\bar{w},k}^i = E\{y_k^i \mathbf{y}_k^{iT} - w_k^i \mathbf{w}_k^{iT}\}$$

$$= \left[\mathbf{v}^i E\left\{\left(y_k^i \mathbf{y}_k^{iT} - w_k^i \mathbf{w}_k^{iT}\right)^T\right\}, \mathbf{v}^i E\right.$$

$$\left.\left\{\left(y_k^i \mathbf{y}_{k-1}^{iT} - w_k^i \mathbf{w}_{k-1}^{iT}\right)^T\right\}, \ldots\right]^T$$

$$= \left[\mathbf{v}^i \left(\mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i\right)^T, \mathbf{v}^i \left(\mathbf{p}_{yy,k+1}^i - \mathbf{p}_{ww,k+1}^i\right)^T, \ldots\right]^T. \tag{15}$$

Taking the z-transform of Eq. (15), we get

$$C_{\bar{y}\bar{w},k}^i(z) = \mathbf{z}\left[\ldots, \mathbf{v}^i \left(\mathbf{p}_{yy,k-1}^i - \mathbf{p}_{ww,k-1}^i\right)^T,\right.$$

$$\mathbf{v}^i \left(\mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i\right)^T, \mathbf{v}^i \left(\mathbf{p}_{yy,k+1}^i - \mathbf{p}_{ww,k+1}^i\right)^T, \ldots\right]^T$$

$$= \mathbf{v}^i\left[\ldots, \mathbf{z}\left(\mathbf{p}_{yy,k-1}^i - \mathbf{p}_{ww,k-1}^i\right)^T, \mathbf{z}\left(\mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i\right)^T,\right.$$

$$\mathbf{z}\left(\mathbf{p}_{yy,k+1}^i - \mathbf{p}_{ww,k+1}^i\right)^T, \ldots\right]^T$$

$$= \mathbf{z}^+ \mathbf{v}^{iT} \mathbf{z}\left(\mathbf{p}_{yy,k}^i - \mathbf{p}_{ww,k}^i\right)^T$$

$$= V^i(z^{-1})\left(C_{yy,k}^i(z) - C_{ww,k}^i(z)\right). \tag{16}$$

Using spectral decomposition, we can write $C_{yy,k}^i(z)$ as

$$C_{yy,k}^i(z) = \sigma_{\bar{y}}^2 W(z) W(z^{-1}). \tag{17}$$

$W(z)$ is the minimum-phase part, which is analytic in the region $|z| > r$ and $r < 1$. With spectral factorization the whitening filter becomes $V^i(z) = 1/W(z)$. Therefore,

$$C_{\bar{y}\bar{w}}^{i+}(z) = \left[\frac{C_{yy,k}^i(z) - C_{ww,k}^i(z)}{W(z^{-1})}\right]^+ = \left[\sigma_{\bar{y}}^2 W(z) - \frac{C_{ww,k}^i(z)}{W(z^{-1})}\right]^+. \tag{18}$$

Now,

$$\mathbf{Q}^i(z) = \left[W(z) - \frac{C_{ww,k}^i(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]^+. \tag{19}$$

and finally

$$\mathbf{G}^i(z) = \frac{\mathbf{Q}^i(z)}{W(z)} = \frac{\left[W(z) - \frac{C_{ww,k}^i(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]^+}{W(z)}. \tag{20}$$

For more details on Eqs. (15)–(20), it is recommended to see the book (Proakis and Manolakis 2006, pp. 818–822). In short, to design an IIR Wiener filter we are required to do spectral factorization of $C_{yy,k}^i(z)$, obtain the minimum-phase part $W(z)$ and finally solve for the causal part of $[W(z) - C_{ww,k}^i(z)/\sigma_{\bar{y}}^2 W(z^{-1})]$.

2.2. Signal-to-Noise Ratio

The *SNR* is commonly defined as

$$\text{SNR} = \frac{\sigma_s^2}{\sigma_w^2}, \tag{21}$$

where σ_s^2 and σ_w^2 are the signal and noise powers, respectively. Using this definition, the following interpretation of the IIR Wiener filter in terms of the *SNR* can be deduced by considering the two limiting cases of a noise-free signal and an extreme noisy signal, which respectively are given by

$$\lim_{\text{SNR}\to\infty} \mathbf{G}^i(z) = 1, \tag{22}$$

$$\lim_{\text{SNR}\to 0} \mathbf{G}^i(z) = 0. \tag{23}$$

The justification of Eqs. (22) and (23) is as follows. When SNR $\to \infty$, this corresponds to zero noise content and therefore $\sigma_w^2 = 0$ and similarly

$C_{ww,k}^i(z) = 0$ and consequently $\mathbf{G}^i(z) = 1$. However, when SNR $\rightarrow 0$, this corresponds to a zero signal content, i.e., $\sigma_s^2 = 0$, which results in $\sigma_{\bar{y}}^2 = \sigma_{\bar{w}}^2$ and $C_{ww,k}^i(z) = \sigma_{\bar{y}}^2 W(z)W(z^{-1})$ and consequently $\mathbf{G}^i(z) = 0$. This means that at a very high SNR, the filter applies very little or no attenuation to the noise-free signal, whereas when there is only noise, the filter enters the stop band region, i.e., does not allow the input signal (which is only noise) to pass through, since the filter response is $\mathbf{G}^i(z) = 0$. In the time domain, $\mathbf{G}^i(z) = 1$ corresponds to $g_1^i = 1$ and $\mathbf{G}^i(z) = 0$ corresponds to $g_n^i = 0, \forall\, n$ (see the properties of z-transform, Proakis and Manolakis 2006, p. 165). The plot for the two cases highlighted above (Eqs. 22, 23) is shown in Fig. 1. To see the response of the filter, for SNR $\rightarrow \infty$, we have used only a Ricker wavelet of center frequency 30 Hz without noise and for SNR $\rightarrow 0$, we have used only noise.

3. Estimation of the z-Transform of the Observation (Noisy Trace)

In this section, the estimations of the autocorrelation sequence of the observation and its corresponding z-transform in practical scenarios are detailed. To estimate the autocorrelation of the noisy observation (and ultimately the z-transform), first the noisy observation is modeled as an auto-regressive (AR) process. For this purpose, the following model is used for the observed data $\mathbf{y}$ at instant k

$$y_k = -\mathbf{a}_{l_y}\mathbf{y}_{k-1} + \gamma_k, \qquad (24)$$

where $\mathbf{a}_{l_y} = [a_1, a_2, \ldots, a_l]$ are AR coefficients (l is the number of coefficients), $\mathbf{y}_{k-1} = [y_{k-1}, y_{k-2}, \ldots, y_{k-l}]^T$ and γ_k is the white noise with zero mean and variance $\sigma_{\gamma_y}^2$. To find $\mathbf{a}_{l_y}$, the Yule-Walker method (Haykin 2002, Chapter 1) is used. Here, post multiplying Eq. (24) with $\mathbf{y}_{k-1}^T$ and taking the mathematical expectation, we obtain

$$E\{y_k\mathbf{y}_{k-1}^T\} = -E\{\mathbf{a}_{l_y}\mathbf{y}_{k-1}\mathbf{y}_{k-1}^T\} + E\{\mathbf{y}_{k-1}^T\gamma_k\}. \qquad (25)$$

Assuming that data and noise are uncorrelated and the noise has zero mean, Eq. (25) becomes

$$-\mathbf{p}_{yy} = \mathbf{P}_{yy}\mathbf{a}_{l_y}^T, \qquad (26)$$

and $\mathbf{a}_{l_y}^T$ is calculated as

$$\mathbf{a}_{l_y}^T = -\mathbf{P}_{yy}^{-1}\mathbf{p}_{yy}, \qquad (27)$$

where $\mathbf{p}_{yy} = [p_{y,1}, p_{y,2}, \ldots, p_{y,l}]^T$, $\mathbf{P}_{yy} =$ Toeplitz($[p_{y,0}, p_{y,1}, \ldots, p_{y,l-1}], [p_{y,0}, p_{y,-1}, \ldots, p_{y,-l+1}]$), and $p_{y,m} = p_{y,-m}$ (Proakis and Manolakis 2006, p. 798). The first column and first row of the Toeplitz matrix $\mathbf{p}_{yy}$ are $[p_{y,0}, p_{y,1}, \ldots, p_{y,l-1}]$ and $[p_{y,0}, p_{y,-1}, \ldots, p_{y,-l+1}]$, respectively (here we have used Matlab-like syntax to represent a Toeplitz matrix). To ensure that the autocorrelation matrix in Eq. (27) is positive definite, a biased form of the estimator is used for $p_{y,m}$, i.e., $p_{y,m} = \frac{1}{N}\sum_{k=0}^{N-m-1} y_k y_{k+m}, m = 0, 1, \ldots, l-1$, where N is the number of samples in the trace. This estimator results in a stable AR model. To avoid inversion in Eq. (27), the Levinson-Durbin algorithm (Haykin 2002) can be used, which is a recursive and computationally efficient method that utilizes the Toeplitz structure of the correlation matrix. After finding $\mathbf{a}_{l_y}$, then the z-transform of the autocorrelation sequence can be found as follows. Equation (24) can be rewritten as

$$\mathbf{a}_{l_y}'\mathbf{y}_k = \gamma_k, \qquad (28)$$

where $\mathbf{a}_{l_y}' = [1, a_1, a_2, \ldots, a_l]$ and $\mathbf{y}_k = [y_k, y_{k-1}, \ldots, y_{k-l}]^T$. Now, taking the z-transform of Eq. (28), we get

$$Y(z)\left(\mathbf{a}_{l_y}'\mathbf{z}^-\right) = \gamma(z). \qquad (29)$$

Multiplying each side of this equation by its respective time-reversed version (Proakis and Manolakis 2006) gives

$$C_{yy,k}(z)\left(\mathbf{a}_l'\mathbf{z}^-\right)\left(\mathbf{a}_l'\mathbf{z}^+\right) = \sigma_{\gamma_y}^2. \qquad (30)$$

In this way, the noisy observation is modeled by an AR process, and the respective z-transform of the autocorrelation sequence can be obtained as

$$C_{yy,k}(z) = \frac{\sigma_{\gamma_y}^2}{\left(\mathbf{a}_{l_y}'\mathbf{z}^-\right)\left(\mathbf{a}_{l_y}'\mathbf{z}^+\right)}. \qquad (31)$$

Now, $\sigma_{\gamma_y}^2$ can be found by multiplying Eq. (28) by its complex conjugate (*) and taking the mathematical

expectation, and since γ_k is a real-valued number, therefore, $E\{|\gamma_k|^2\} = \sigma_{\gamma_y}^2$. Now from Eq. (28)

$$E\{\gamma_k \gamma_k^*\} = E\{\mathbf{a}'_{l_y} \mathbf{y}_k \mathbf{y}_k^T \mathbf{a}'^T_{l_y}\}, \qquad (32)$$

$$\sigma_{\gamma_y}^2 = \mathbf{a}'_{l_y} E\{\mathbf{y}_k \mathbf{y}_k^T\} \mathbf{a}'^T_{l_y}, \qquad (33)$$

$$\sigma_{\gamma_y}^2 = \mathbf{a}'_{l_y} \mathbf{p}_{yy} \mathbf{a}'^T_{l_y}. \qquad (34)$$

Details of the above derivation can be found in Sayed (2008). From Eq. (31), it is clear that $C_{yy,k}(z)$ can be found using the knowledge of $\mathbf{a}_{l_y}$ and $\sigma_{\gamma_y}^2$. This yields the estimation of the z-transform of the observation autocorrelation sequence.

4. Estimation of z-Transform of Noise

To estimate the noise autocorrelation matrix, we use the approach of Rangachari and Loizou (2006). It consists of five steps, namely, the initial power spectrum estimation, the minimum tracking, the event detection, the smoothing factor calculation and the noise spectrum update. The steps are detailed next.

4.1. Initial Power Spectrum Estimation of Noisy Data

The smoothed power spectrum of the noisy data is estimated using the first-order recursive relation (Rangachari and Loizou 2006; Doblinger 1995) as follows,

$$\mathcal{P}_i(\lambda) = \eta \mathcal{P}_i(\lambda - 1) + (1 - \eta)|\mathcal{Y}_i(\lambda)|^2, \qquad (35)$$

where $\mathcal{Y}_i(\lambda)$ is the short-time Fourier transform of the noisy data, η is the forgetting factor (which gives less weight to older samples), and λ is the frame index (Doblinger 1995). The power spectrum $|\mathcal{Y}_i(\lambda)|^2$ is obtained by taking the absolute value of each element and squaring it. The size of $\mathcal{Y}_i(\lambda)$ is $\bar{N}$, and the step size (or hop size) to calculate the short-time Fourier transform is h.

4.2. Tracking the Minimum of the Noisy Data

The minimum of the noisy data is tracked by a non-linear approach that averages the the past values continuously as follows (Doblinger 1995):

$$\begin{aligned}
\mathcal{P}_{i,min}(\lambda) = {}& [\mathcal{P}_{i,min}(\lambda - 1) < \mathcal{P}_i(\lambda)] \\
& \odot [\gamma \mathcal{P}_{i,min}(\lambda - 1) + \frac{1 - \gamma}{1 - \beta}(\mathcal{P}_i(\lambda) - \beta \mathcal{P}_i(\lambda - 1))] \\
& + [\mathcal{P}_{i,min}(\lambda - 1) > \mathcal{P}_i(\lambda)] \odot \mathcal{P}_i(\lambda),
\end{aligned}$$

$$(36)$$

where $[\mathcal{P}_{i,min}(\lambda - 1) < \mathcal{P}_i(\lambda)]$ represents element-by-element comparison, and its resultant vector contains 1s (if condition is true) and 0s (otherwise), "$\odot$" denotes the Hadamard product (element-by-element multiplication), and $\mathcal{P}_{i,min}(\lambda - 1)$ is the local minimum of the noisy data power spectrum. γ and β are the adaptation constants that are determined experimentally (Rangachari and Loizou 2006; Doblinger 1995).

4.3. Microseismic Event Detection Probability

To detect the presence of microseismic events, the ratio of the noisy data spectrum to its local minimum, $\mathcal{S}_i(\lambda)$, is defined as (Cohen and Berdugo 2002),

$$\mathcal{S}_i(\lambda) = \mathcal{P}_i(\lambda) \oslash \mathcal{P}_{i,min}(\lambda), \qquad (37)$$

where "$\oslash$" represents the element-by-element division. The ratio is based on the fact that the power spectrum of the noisy trace will be nearly equal to its local minimum when a microseismic event is absent. The smaller the ratio in Eq. (37) is, the higher the probability that the event is absent. The ratio is compared with a frequency-dependent threshold δ, and consequently the event presence probability $\mathcal{I}_i(\lambda)$ is updated, using first-order recursion, as (Rangachari and Loizou 2006),

$$\mathcal{I}_i(\lambda) = \alpha_p \mathcal{I}_i(\lambda - 1) + (1 - \alpha_p)[\mathcal{S}_i(\lambda) > \delta)], \quad (38)$$

where $[\mathcal{S}_i(\lambda) > \delta)]$ represents a comparison of each element in $\mathcal{S}_i(\lambda)$ with the frequency-dependent threshold δ (to be discussed later in the results section) and its result is a vector with 1s (if the condition is true) and 0s (otherwise). The quantity α_p is a smoothing constant (Rangachari and Loizou 2006).

The recursion in Eq. (38) implicity exploits the correlation among the frames for detecting the event.

4.4. Calculation of the Smoothing Factor

The time-frequency dependent smoothing factor is computed as (Cohen and Berdugo 2002),

$$\alpha_s = \alpha_d + (1 - \alpha_d)\mathcal{I}_i(\lambda). \tag{39}$$

where α_d is a constant and α_s is a time-varying smoothing parameter. Note that α_s has values in the range of $\alpha_d \leq \alpha_s \leq 1$ (Rangachari and Loizou 2006).

4.5. Updating the Estimation of the Noise Spectrum

Finally, the noise power spectrum estimate $\mathcal{N}_i(\lambda)$ is updated according to

$$\mathcal{N}_i(\lambda) = \alpha_s \mathcal{N}_i(\lambda - 1) + (1 - \alpha_s)|\mathcal{Y}_i(\lambda)|^2. \tag{40}$$

The above procedure is done for all frequency bins altogether, as obvious by using the vector notation in equations. Note that constants (mixing parameters) $\eta, \gamma, \beta, \alpha_p$ and α_d can easily be determined experimentally, and their values lie between 0 and 1 (Rangachari and Loizou 2006). The overall algorithm can be summarized as follows. After classifying the frequency bins as event absent/present, the event presence probability is updated using Eq. (38). Using this probability, the time-frequency-dependent smoothing factor is updated as in Eq. (39). Finally, the noise power spectrum is estimated using update Eq. (40). After obtaining the noise power spectrum estimate, it is averaged over all λ's $\mathcal{N}_a = \sum_\lambda N_i(\lambda)$ and converted back to the time domain, which gives the noise autocorrelation estimate $p_{w,0}, p_{w,1}, \cdots, p_{w,l-1}$. Then, from these estimates, $\mathbf{p}_{ww}$ and $\mathbf{P}_{ww}$ are found (as done for $\mathbf{p}_{yy}$ and $\mathbf{P}_{yy}$). To find the z-transform of the noise autocorrelation sequence, the procedure outlined in Sect. 3 has been used.

5. Summary of the Denoising Method

In this section, we summarize our proposed method and suggest some enhancements for estimating the correlation matrices.

5.1. Proposed Denoising Method

The outline of our proposed denoising method can be summarized as follows:

1. Find the autocorrelation of noisy data and noise using $p_{y,m} = \frac{1}{N}\sum_{k=0}^{N-m-1} y_k y_{k+m}$ and $p_{w,m} = \frac{1}{N}\sum_{k=0}^{N-m-1} w_k w_{k+m}, m \geq 0$, respectively, and then form $\mathbf{p}_{yy}, \mathbf{P}_{yy}, \mathbf{p}_{ww}$ and $\mathbf{P}_{ww}$ for each trace.

2. Find the AR parameters for the noisy observation and the noise using $\mathbf{a}_{l_y}^T = -\mathbf{P}_{yy}^{-1}\mathbf{p}_{yy}$ and $\mathbf{a}_{l_w}^T = -\mathbf{P}_{ww}^{-1}\mathbf{p}_{ww}$, respectively.

3. Find the z-transform of the autocorrelation sequence for the observed data and the noise as
$$C_{yy,k}(z) = \frac{\sigma_{\gamma_y}^2}{\left(\mathbf{a}'_{l_y}\mathbf{z}^-\right)\left(\mathbf{a}'_{l_y}\mathbf{z}^+\right)} \quad \text{and}$$
$$C_{ww,k}(z) = \frac{\sigma_{\gamma_w}^2}{\left(\mathbf{a}'_{l_w}\mathbf{z}^-\right)\left(\mathbf{a}'_{l_w}\mathbf{z}^+\right)}, \quad \text{respectively,} \quad \text{where}$$
$\sigma_{\gamma_y}^2 = \mathbf{a}'_{l_y}\mathbf{p}_{yy}^T\mathbf{a}'^T_{l_y}$ and $\sigma_{\gamma_w}^2 = \mathbf{a}'_{l_w}\mathbf{p}_{ww}^T\mathbf{a}'^T_{l_w}$.

4. Find $W(z)$ by calculating the roots of $C_{yy,k}(z)$ that fall inside the unit circle in the z-plane and $\sigma_{\bar{y}}^2 = \sigma_{\gamma_y}^2$.

5. Find the causal part of $\left[W(z) - \frac{C_{ww,k}^{i+}(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]$.

6. Finally, find the transfer function of the IIR filter using $\mathbf{G}^i(z) = \dfrac{\left[W(z) - \frac{C_{ww,k}^{i+}(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]^+}{W(z)}$ and filter the noisy observation to get the clean signal.

7. The procedure is repeated in an iterative fashion to get a clearer signal.

The work flow for the proposed denoising method is presented in Fig. 2. For simplicity of the notation, the superscript i (that corresponds to the ith trace) is omitted, and the work flow is the same for all traces.

5.2. Proposed Enhancement Method

Assuming that the noise has the similar autocorrelation (or power spectral density in frequency-domain) for various traces (Caffagni et al. 2016; Cieplicki et al. 2014), the estimation of the autocorrelation sequences for data and noise can be improved by stacking over the adjacent traces and/or components in case of 3C sensors. This is done as follows.
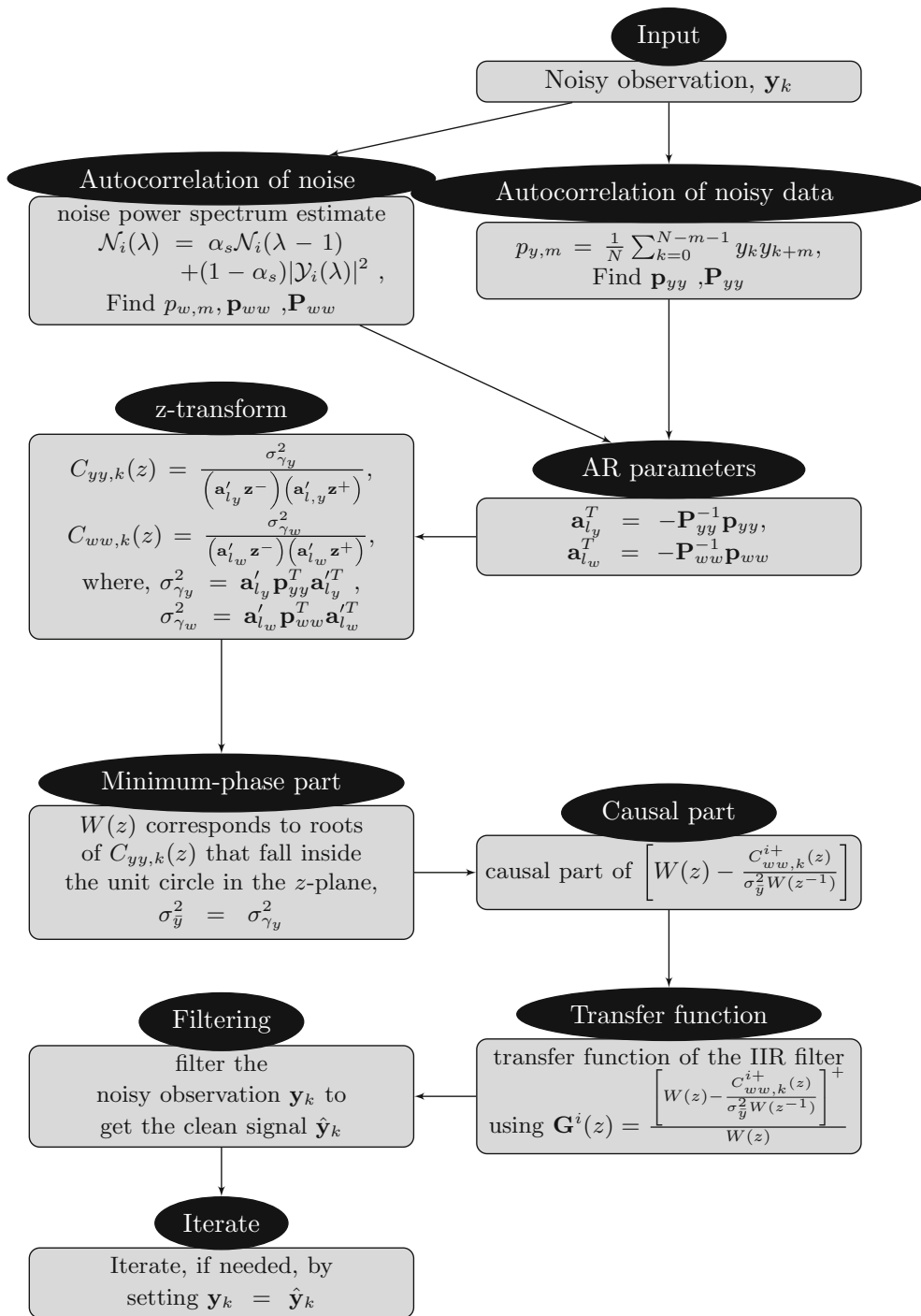
Input

Noisy observation, $\mathbf{y}_k$

Autocorrelation of noise

noise power spectrum estimate
$$\mathcal{N}_i(\lambda) = \alpha_s \mathcal{N}_i(\lambda - 1)$$
$$+ (1 - \alpha_s)|\mathcal{Y}_i(\lambda)|^2 ,$$
Find $p_{w,m}, \mathbf{p}_{ww}, \mathbf{P}_{ww}$

Autocorrelation of noisy data

$$p_{y,m} = \frac{1}{N} \sum_{k=0}^{N-m-1} y_k y_{k+m},$$
Find $\mathbf{p}_{yy}, \mathbf{P}_{yy}$

z-transform

$$C_{yy,k}(z) = \frac{\sigma_{\gamma_y}^2}{\left(\mathbf{a}'_{l_y}\mathbf{z}^-\right)\left(\mathbf{a}'_{l,y}\mathbf{z}^+\right)},$$
$$C_{ww,k}(z) = \frac{\sigma_{\gamma_w}^2}{\left(\mathbf{a}'_{l_w}\mathbf{z}^-\right)\left(\mathbf{a}'_{l_w}\mathbf{z}^+\right)},$$
$$\text{where, } \sigma_{\gamma_y}^2 = \mathbf{a}'_{l_y}\mathbf{p}_{yy}^T\mathbf{a}'^T_{l_y},$$
$$\sigma_{\gamma_w}^2 = \mathbf{a}'_{l_w}\mathbf{p}_{ww}^T\mathbf{a}'^T_{l_w}$$

AR parameters

$$\mathbf{a}_{l_y}^T = -\mathbf{P}_{yy}^{-1}\mathbf{p}_{yy},$$
$$\mathbf{a}_{l_w}^T = -\mathbf{P}_{ww}^{-1}\mathbf{p}_{ww}$$

Minimum-phase part

$W(z)$ corresponds to roots
of $C_{yy,k}(z)$ that fall inside
the unit circle in the z-plane,
$$\sigma_{\bar{y}}^2 = \sigma_{\gamma_y}^2$$

Causal part

causal part of $\left[W(z) - \frac{C_{ww,k}^{i+}(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]$

Transfer function

transfer function of the IIR filter
using $\mathbf{G}^i(z) = \dfrac{\left[W(z) - \frac{C_{ww,k}^{i+}(z)}{\sigma_{\bar{y}}^2 W(z^{-1})}\right]^+}{W(z)}$

Filtering

filter the
noisy observation $\mathbf{y}_k$ to
get the clean signal $\hat{\mathbf{y}}_k$

Iterate

Iterate, if needed, by
setting $\mathbf{y}_k = \hat{\mathbf{y}}_k$

Figure 2
Work flow of the proposed denoising method. For simplicity of the notation, the superscript i (that corresponds to the ith trace) is omitted. The work flow is the same for all traces

First, the autocorrelation of noisy data and noise is found using $p_{y,m}^{ij} = \frac{1}{N}\sum_{k=0}^{N-m-1} y_k^{ij} y_{k+m}^{ij}$ and $p_{w,m}^{ij} = \frac{1}{N}\sum_{k=0}^{N-m-1} w_k^{ij} w_{k+m}^{ij}, m \geq 0$, respectively, for the ith trace and jth component of a 3C sensor. Second, after finding these autocorrelations for all the traces, the autocorrelations are stacked to improve the estimation, i.e.,

$$p_{w,m} = \frac{1}{3M}\sum_{j=1}^{3}\sum_{i=1}^{M} p_{w,m}^{ij}, \tag{41}$$

A similar procedure is used for estimating the autocorrelation of observation, i.e.,

$$p_{y,m} = \frac{1}{3M}\sum_{j=1}^{3}\sum_{i=1}^{M} p_{y,m}^{ij}, \tag{42}$$

After finding the autocorrelation of the noisy data $p_{y,m}$, and noise $p_{w,m}$, the autocorrelation matrices $\mathbf{p}_{yy}$, $\mathbf{P}_{yy}$, $\mathbf{p}_{ww}$ and $\mathbf{P}_{ww}$ are formed, and the rest of the procedure is the same as for the proposed method.

Importantly, in the proposed enhancement method, we are not stacking the traces prior to the autocorrelation but we are stacking the autocorrelations of the traces, i.e., first the autocorrelations are found for each trace followed by the stacking of the autocorrelations. The autocorrelations do not need to be aligned as they are already aligned. The estimation improves if traces have similar power spectral densities, i.e., traces have white noise or Brownian noise, etc. Now the question arises, if traces have the similar power spectral densities but different noise levels (variances), are the results affected? More importantly, our method compensates automatically for the difference in the variances. The IIR Wiener filter is the ratio of two autocorrelations. Hence, if the noise level changes, the level of the noisy observation also changes and so does the magnitude of the autocorrelations of the noisy observation and noise (remember that the autocorrelations are derived from the traces). Since the filter is a ratio of autocorrelations (Eq. 13) or power spectral densities (Eq. 20), it will compensate for the effect. Even if a bias is present because of a broken channel, this will not affect the results.

In summary, the proposed denoising method estimates the autocorrelation in a trace-by-trace manner, whereas the proposed enhancement method improves the estimation of the autocorrelation by averaging over the adjacent traces.

Other steps of the denoising method remain the same. With the improvement of the autocorrelation sequence estimation, the proposed method will hereafter be called the proposed enhanced method.

6. Results

In this section, the IIR Wiener filter is tested on synthetic, semi-synthetic and field data. To test the robustness of the proposed method, two cases for noise (correlated and uncorrelated noise) are used.

6.1. Synthetic Data Set

For the synthetic data set, a Ricker wavelet with a center frequency of 5 Hz is used as the microseismic source signature to generate the data set. Fifty receivers are placed inline, and the middle receiver is assumed to be the closest to the source. Moreover, a constant medium velocity is assumed, and the sampling frequency is set to 1 kHz. The resulting data are depicted in Fig. 3. For the proposed method, the number of AR coefficients is set to $l = 10$ for all the data sets used. Testing revealed that increasing the
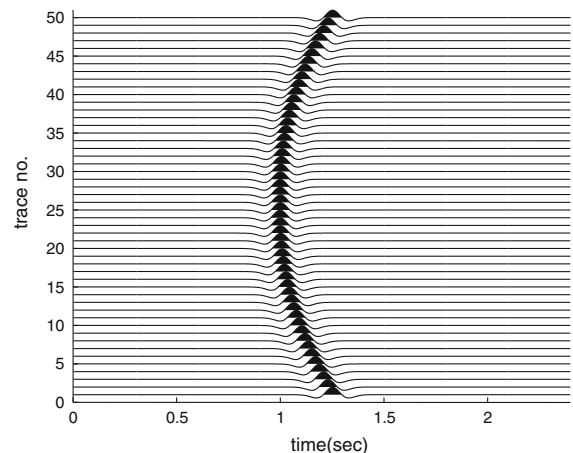

Figure 3
Synthetic data without noise. A Ricker wavelet with a center frequency of 5 Hz is used. Fifty receivers are placed inline, and the middle receiver is assumed to be the closest to the source. A constant medium velocity is assumed, and the sampling frequency is set to 1 kHz

number of parameters does not improve the SNR. However, large values of l increase the complexity because of the inversion of the large matrix in Eq. (27). Finally, N needs to be known for the estimation of the autocorrelation of the observation, $p_{y,m}$. The observation or the observed data in our case are traces; therefore, N is equal to the number of samples in each trace. For long recordings, the data can be divided into windows instead of processing the whole data altogether, and then the proposed method can be applied to these windows. The test revealed that the optimum values of the constants used for noise estimation are $\eta = 0.85, \gamma = 0.998, \beta = 0.85, \alpha_p = 0.2, \alpha_d = 0.95, \bar{N} = N/10,$ and $h = N/20$.

The two kinds of noise (correlated and uncorrelated noise) are added to the raw traces (Fig. 3) to test the proposed method. The SNR [defined as $10 \log_{10}\{\sigma_s^2/\sigma_w^2\}$ in decibel (dB)] of the noisy observation in both cases is equal to -12 dB. It is apparent that the traces are noisy and the microseismic event is difficult to identify, as can be seen in Fig. 4a, b. For generating the correlated noise, white Gaussian noise is filtered using a geophone's impulse response given by Hons and Stewart (2006). After filtering, the spectrum and autocorrelation of noise are shown in Fig. 5a, b, respectively. The spectrum is not flat, and the autocorrelation is not an impulse, unlike white noise (uncorrelated), which has a flat spectrum, and its autocorrelation is that of an impulse-type function

at zero time lag (zero correlation index). The spectrum in Fig. 5a is flat except for frequencies greater than 150 Hz.

Here, the noise is estimated first using the procedure discussed in Sect. 4. For trace 1, the power
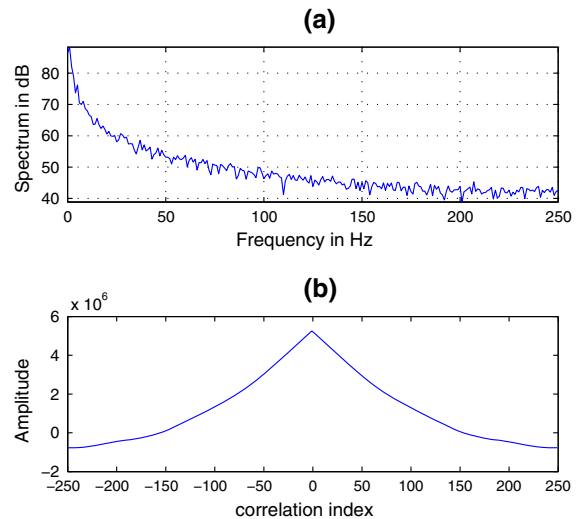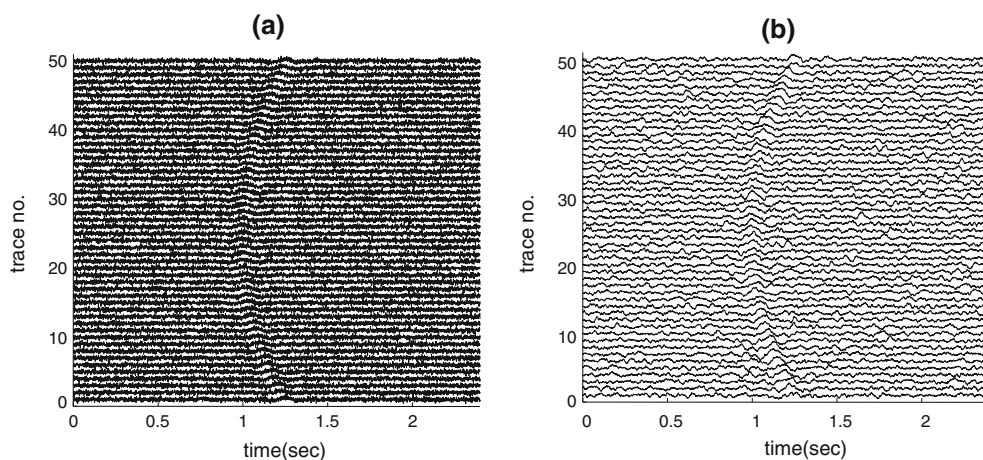

Figure 5
Correlated noise specifications: **a** spectrum of correlated noise. **b** Its corresponding autocorrelation. For generating the correlated noise, white Gaussian noise is filtered using a geophone's impulse response given by Hons and Stewart (2006). The spectrum is not flat and the autocorrelation is not an impulse, unlike white noise (uncorrelated), which has a flat spectrum and its autocorrelation is that of an impulse-type function at zero time lag (zero correlation index)

Figure 4
a Synthetic data with white Gaussian noise. **b** Synthetic data with correlated noise. The SNR of the noisy observation in both cases is set to -12 dB. It is apparent that the traces are noisy, and the microseismic event is difficult to identify. For generating the correlated noise, white Gaussian noise is filtered using a geophone's impulse response given by Hons and Stewart (2006)

spectrum of the noiseless data and event-presence probability $\mathcal{I}_i$ (with white Gaussian noise), calculated using Eq. (38), is shown in Fig. 6a, b. Note that by defining a suitable value of the threshold, the event can be detected with higher confidence. As seen in Fig. 6b, some noise-only regions are detected as the events. This overestimation of events is not likely to affect the enhanced event (after denoising), since the detection probability improves with the iteration, hence eliminating the false event detected regions (Fig. 6b–e). The threshold value used is

$$\delta = \begin{cases} 2, & 1 \leq \lambda \leq F \\ 5, & F \leq \lambda \leq f_s/2, \end{cases} \tag{43}$$

where F is the frequency bin corresponding to 100 Hz frequency. Above 100 Hz, a higher value of the threshold is used because the event will be in the low frequency range (within 100 Hz).

In the first experiment with white Gaussian noise (uncorrelated case), when the filter is applied to the data, the noise is increasingly suppressed with each iteration (Fig. 7), and the microseismic event is dominating. This proves the effectiveness of our proposed method. Figure 8a, b shows the denoising results (after iteration 4) under the white Gaussian noise case using the proposed method and the proposed enhanced method, respectively. In the second experiment with correlated noise, again the filter attenuates the noise, and the event becomes clearer (Fig. 8c, d after iteration 4). Next, to show the improvement in the SNR in each iteration, the

noise level is plotted against the number of iterations, as depicted in Fig. 9. It is apparent from the figure that the SNR is improving with every iteration. With more iterations, we lose the basic assumption that noise and signal are uncorrelated; hence, the SNR starts to decrease after the fourth iteration. Consequently, we used only four iterations throughout our simulation results. To ensure that we get the best results in terms of denoising, after the fourth iteration of applying the filter, the wavelet-based denoising can also be applied before the final output (results shown in Figs. 7 and 8 are without the wavelet-based denoising method). This is necessary to remove any in-band noise present after the filter application. The details of the wavelet denoising method are presented later.

The blind estimation of the noise is clearly an advantage of the proposed method. However, the nature of the noise is really important. With white noise, the spectrum is flat, and when we move to the time-frequency domain for the noise estimation, the noise contents are more or less equally distributed along the frequency. On the other hand, with colored noise, the spectrum is concentrated in some frequency bands (which include the band of the events). Since we have used the same variance (level) of the noise in both cases (white and color), the in-band noise (noise in the band of event) is less in the white noise case (contents are equally distributed along the spectrum) than in the colored noise case (contents are concentrated in some bands). Therefore, the
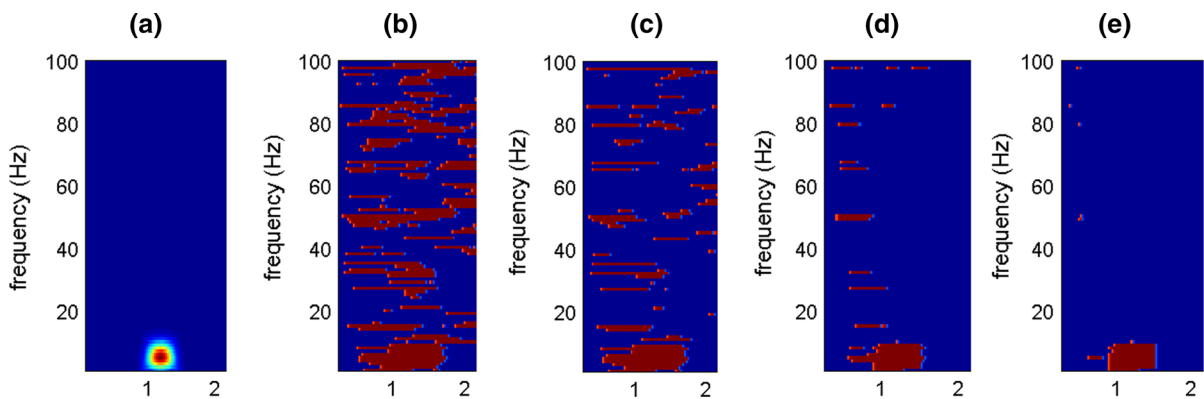

Figure 6
a Power spectrum of the noiseless synthetic data using short-time Fourier transform. Event presence probability $\mathcal{I}_1$ of noisy version of trace 1, iteration 1, **c** iteration 2, **d** iteration 3, **e** iteration 4. The detection probability improves with the iteration, hence eliminating the false event detected region
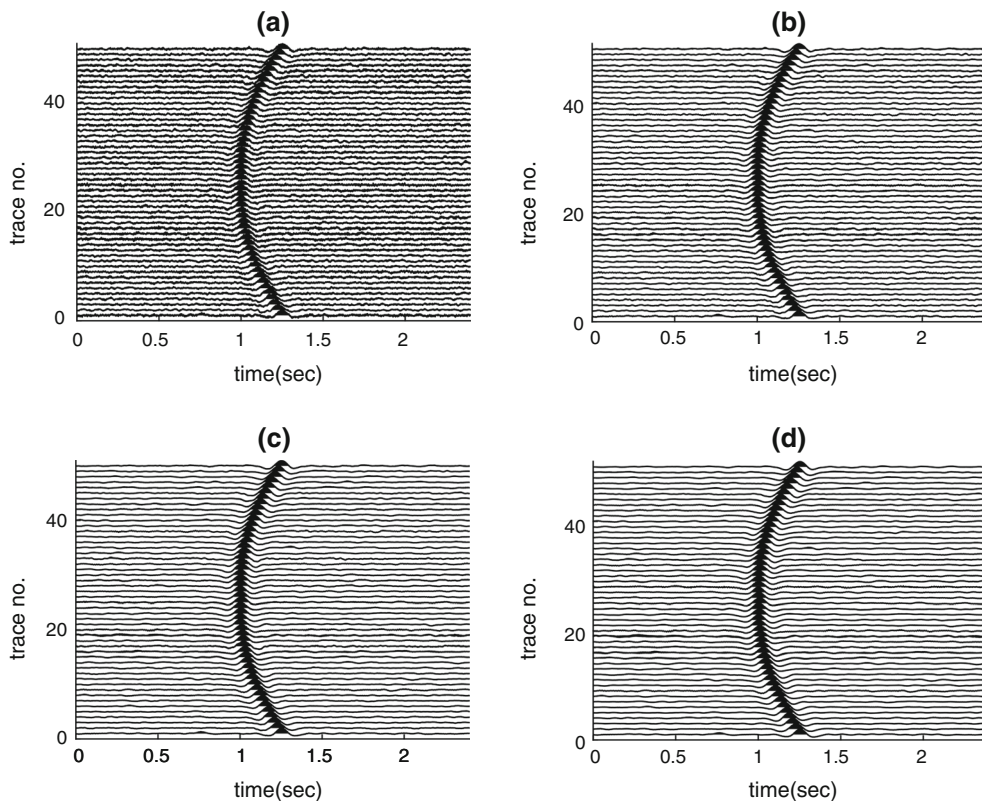
Figure 7
Denoised synthetic data using proposed method: **a** iteration 1, **b** iteration 2, **c** iteration 3, **d** iteration 4. The noise is increasingly suppressed with each iteration, and the microseismic event is dominating

performance of estimation in the colored noise case is somewhat lower than in the white noise case.

6.2. Comparison of Denoising Methods

The quantitative assessment of the proposed method can be verified by comparing the mean square error (MSE), mean absolute error (MAE), SNR, peak-signal-to-noise ratio (PSNR) and maximum correlation coefficient (CC), which are listed in Table 1. In this table, the performance of the proposed method in the case of a synthetic data set with white Gaussian noise is compared with bandpass filtering, wavelet decomposition, empirical mode decomposition and FIR Wiener filtering. For denoising using wavelet decomposition, the '*wden*' function in the wavelet toolbox of MATLAB is used (Misiti et al. 1996). Various wavelet basis functions with their variants, i.e., Daubechies (*db2, db3, db4*), Coiflets (*co2, co4, co5*) and Symlets (*sy2, sy3, sy4*)

(Daubechies 1992; Mallat 1989), are tested on the pseudo-real data set, and then the *coif5* wavelet is selected for comparison based on its best performance over the other wavelets. Furthermore, we used soft thresholding with the principle of Stein's unbiased risk (Stein 1981).[1] Another method used for comparison is the empirical mode decomposition (Rilling et al. 2007) that derives the basis function from the observed data. This is also a well-known method used for denoising in geophysics. There are several different methods for denoising seismic data based on empirical mode decomposition. Here, we have used the ensemble empirical mode decomposition with the adaptive thresholding method (Han and van der Baan 2015). One more widely used method used for enhancing the SNR of noisy seismic data is interferometry (Al-Shuhail et al. 2013). Comparison

---

[1] The same method of wavelet denoising is used in the proposed method after the fourth iteration.

(a)   **(b)**

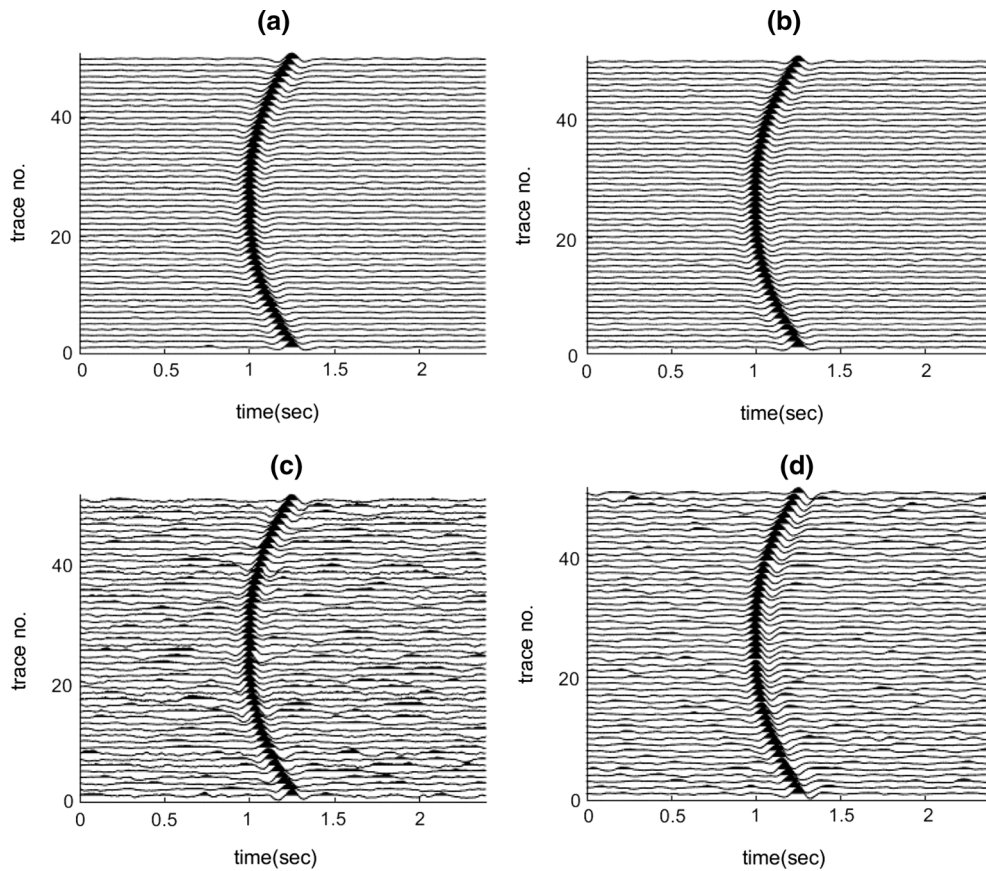**(c)**   **(d)**

Figure 8

Denoised synthetic data: **a** proposed method (white noise case), **b** proposed enhanced method with stacking of adjacent traces (white noise case), **c** proposed method (correlated noise case), **d** proposed enhanced method with stacking of adjacent traces (correlated noise case). In the second experiment with correlated noise, again the filter attenuated the noise, and the event became clearer with the iteration

of the above-mentioned methods reveals that our proposed method is superior to the other methods (see Table 1). Moreover, comparing the performance of the proposed method with the wavelet-based method, we can see that the IIR-based filter has improved the performance by 15% at the cost of doubling the computational complexity. For the proposed enhanced method, stacking is done (see Eq. 42) over the adjacent three traces. In comparing the FIR Wiener filter with the IIR Wiener filter, the length of the FIR Wiener filter is taken to be equal to $N / 10$ (for performance and complexity compromise). Due to the inversion of a large matrix, the FIR Wiener filter took 9 s as compared to 4 s by the IIR Wiener filter as shown in Table 1.

6.3. Semi-synthetic Data Set

Next, an earthquake data set is used to validate the proposed method. The data set is obtained from Incorporated Research Institutions for Seismology (IRIS). The event occurred on 9 October 2017 in Central Alaska, USA, and had a magnitude of $Ml = 5.2$. The data were recorded on four tri-component (3C) sensors at a sampling frequency of 250 Hz.

The noiseless and noisy data (corrupted with correlated noise, hereafter denominated "semi-real" or "semi-synthetic" data set) are shown in Fig. 10a, b, respectively (for compactness three components are plotted on the same figure). An improvement in SNR of about 13 dB is achieved with the proposed

method (Fig. 10c) and 14-dB improvement with the proposed enhanced method (Fig. 10d).

6.4. Field Data Set

Finally, the proposed method is applied to a real data set. Figure 11a shows the amplitude normalized and mean subtracted version of the data set. This data
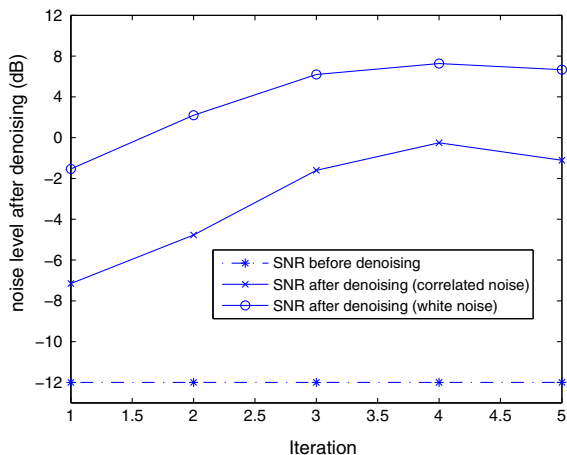

Figure 9
Noise level versus number of iterations. Curves for SNR before and after denoising (for white noise and correlated noise case) using the proposed method. The SNR is improving with every iteration. With more iterations, we lose the basic assumption that noise and signal are uncorrelated; hence, the SNR starts to decrease after the fourth iteration. Consequently, we used only four iterations throughout our simulation results

set is obtained from the IRIS website, and the sampling rate is 4 ms. Similar to Mousavi and Langston (2016c), this microearthquake data set has a magnitude $Ml = 0.4$ and a depth of 4.8 km with the seismometer located on the surface. The earthquake occurred in the California-Nevada Border Region. This scenario is taken to prove the effectiveness of the proposed method in the case of microearthquakes. The parameters used here are the same as those used for the synthetic data set.

The result of the filtering process is shown in Fig. 11b. The results are shown after the fourth iteration and application of the wavelet-based denoising approach. As demonstrated in this figure, our proposed method is indeed able to detect the earthquake signal effectively and attenuated the noise.

The aforementioned test was on a natural microearthquake. The algorithm was originally tested on a real induced microseismic data set from the local oil industry in the Middle East. However, due to legal issues we are unable to report the results here. To demonstrate the effectiveness of the proposed denoising method, we apply it on another field data set that is used by Liu et al. (2017) (Fig. 12a). These data come from the High Resolution Seismic Network (HRSN) operated by Berkeley Seismological Laboratory, University of California, Berkeley. The sampling frequency is 250 Hz. The effectiveness of the proposed method can be appreciated from Fig. 12b, which shows the P- and S-arrivals clearly.

Table 1

Mean square error (MSE), mean absolute error (MAE), signal-to-noise ratio (SNR), peak-signal-to-noise ratio (PSNR) and maximum correlation coefficient (CC) from the synthetic data experiment (white Gaussian noise case) using bandpass filtering, interferometry, wavelet decomposition, empirical mode decomposition, FIR Wiener filtering and the proposed method

Method	MSE	MAE	SNR (dB)	PSNR (dB)	CC	Elapsed time (s)
Noisy data set	0.39	0.50	− 12.03	4.00	0.244	–
Interferometry	0.013	0.085	2.836	18.868	0.805	13.18
Bandpass filtering	0.027	0.109	− 0.286	15.746	0.555	1.32
Wavelet decomposition	0.013	0.084	2.875	18.907	0.797	1.93
Empirical mode decomposition	0.024	0.122	0.160	16.192	0.721	8.47
FIR Wiener filter	0.0256	0.1021	0.9081	15.9256	0.645	9.13
Proposed method	0.011	0.081	3.643	19.675	0.838	4.08
Proposed enhanced method	0.009	0.076	4.285	20.317	0.863	3.98

For a fair comparison, in this table we have not applied the wavelet-based denoising after the fourth iteration of application for the IIR Wiener filter
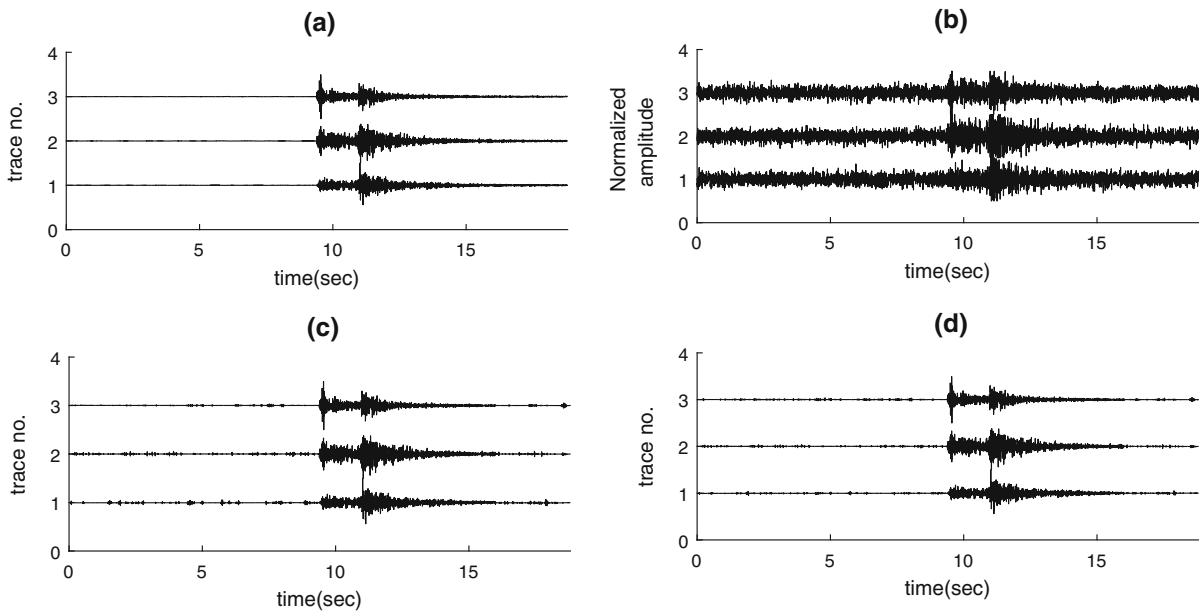
Figure 10

The field data set was obtained from the Incorporated Research Institutions for Seismology (IRIS). The event occurred on 9 October 2017 in Central Alaska, USA, and has magnitude $Ml = 5.2$. The data were recorded on a 3C sensor with sampling frequency of 250 Hz (resampled). **a** Field data set without noise. **b** Field data set with added white Gaussian noise (SNR = −7 dB). **c** Denoised data using the proposed method (SNR= 5.9 dB). **d** Denoised data using the proposed enhanced method (SNR = 7.2 dB). Wavelet-based denoising method is used after forth iteration for these results
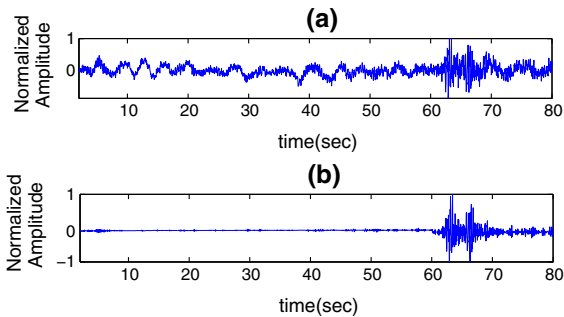
Figure 11

a Noisy field data; **b** denoised field data. The parameters used here are the same as those used for the synthetic data set. Results are shown after the fourth iteration and application of the wavelet denoising method

7. Discussion

Typical microseismic data are characterized by low SNR and highly non-stationary noise. Suppressing noise will drastically improve signal detection, seismogram composition studies, source discrimination for small local/regional seismic sources as well as fracture characterization and monitoring in oil and
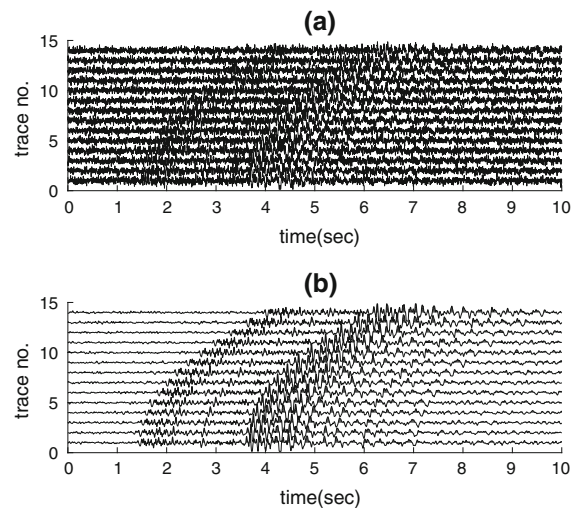

Figure 12

a Field data set (Liu et al. 2017) from the High Resolution Seismic Network (HRSN) operated by Berkeley Seismological Laboratory, University of California, Berkeley. The sampling frequency is 250 Hz. **b** Denoised traces using the proposed denoising method

gas reservoir. The SNR-enhancing methods usually rely on cross-correlation from the seismic traces recorded by geophone arrays. In this work, we

propose a data-driven method to denoise seismic data. To isolate the noise from the signal, we need to acquire knowledge of the second-order statistics of the noise and the noisy signal. Since the occurrence of microseismic events is sporadic, the statistics are estimated directly from the received data. In this study, noise is first estimated and then removed from the receiver record. This makes a practical sense for microseismic denoising, since it is usually possible to estimate the statistics of the noise but not for the signal. The autocorrelations needed for the filter are either estimated from each trace separately or from multiple traces. In the former case, the advantage is that we can use the filter in the case of a single sensor, e.g., microseismic recorded by a single station or use parallel processing as in the case of a sensor array. However, in the latter case, the correlation estimation is improved by stacking the correlation estimates obtained from multiple seismic traces recorded by a geophone array. The stacking does not involve alignment of traces, since the proposed method relies on the autocorrelation. Hence, ambiguities resulting from misalignment are also eliminated here. Another advantage of the proposed method is that it does not impose any assumption on the noise statistics, which makes it suitable for applications with different noise types. Similar denoised results are obtained for the correlated and uncorrelated noise. It is also worth mentioning that our focus in this study is on low SNR seismic signals to prove the effectiveness of the proposed method. However, it is expected to perform well in case of earthquakes with magnitude greater than 2 and active (controlled source) seismic data. The reason is that no underlying assumption about the type of data is used, while designing the filter. Application of the proposed method on synthetic and field microseismic traces with both uncorrelated and correlated noise shows promising results. Our filtering procedure has been applied in an iterative fashion and improvements in SNR shown (see Fig. 9).

8. Conclusion

In this study, we proposed an IIR Wiener filter-based denoising method. The proposed method is directly based on the second-order statistics of the noise and the observations, which can be obtained easily from the recorded time-series data. The proposed method gives a promising performance in a low SNR situation. The filter does not assume any specific noise statistics; this is desirable for applicability of the denoising method to field data recorded in diverse seismic noise environments. More importantly, its computational complexity is much lower in comparison to an equivalent FIR filter approach.

REFERENCES

Aghayan, A., Jaiswal, P., & Siahkoohi, H. R. (2016). Seismic denoising using the redundant lifting scheme. *Geophysics*, *81*(3), V249–V260.

Al-shuhail, A., Aldawood, A., & Hanafy, S. (2012). Application of super-virtual seismic refraction interferometry to enhance first arrivals: A case study from Saudi Arabia. *The Leading Edge*, *31*, 34–39.

Al-Shuhail, A., Kaka, S. I., & Jervis, M. (2013). Enhancement of passive microseismic events using seismic interferometry. *Seismological Research Letters*, *84*(5), 781–784.

Baziw, E., & Weir-Jones, I. (2002). Application of Kalman filtering techniques for microseismic event detection. *Pure and Applied Geophysics*, *159*(1), 449–471.

Bharadwaj, P., Wang, X., Schuster, G., & McIntosh, K. (2013). Increasing the number and signal-to-noise ratio of OBS traces with supervirtual refraction interferometry and free-surface multiples. *Geophysical Journal International*, *192*(3), 1070–1084.

Caffagni, E., Eaton, D. W., Jones, J. P., & van der Baan, M. (2016). Detection and analysis of microseismic events using a Matched Filtering Algorithm (MFA). *Geophysical Journal International*, *206*(1), 644–658.

Castellanos, F., & van der Baan, M. (2013). Microseismic event locations using the double-difference algorithm. *CSEG Recorder*, *38*, 26–37.

Chen, J., Benesty, J., Huang, Yiteng, & Doclo, S. (2006). New insights into the noise reduction Wiener filter. *IEEE Transactions on Audio, Speech and Language Processing*, *14*(4), 1218–1234.

Cieplicki, R., Mueller, M., and Eisner, L. (2014). Microseismic event detection: Comparing P-wave migration with P- and S-wave cross-correlation. In *SEG Technical Program Expanded Abstracts 2014* (pp. 2168–2172). Society of Exploration Geophysicists.

Cohen, I. (2003). Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging. *IEEE Transactions on Speech and Audio Processing*, 11(5), 466–475.

Cohen, I., & Berdugo, B. (2002). Noise estimation by minima controlled recursive averaging for robust speech enhancement. *IEEE Signal Processing Letters*, 9(1), 12–15.

Coughlin, M., Harms, J., Christensen, N., Dergachev, V., DeSalvo, R., Kandhasamy, S., et al. (2014). Wiener filtering with a seismic underground array at the Sanford Underground Research Facility. *Classical and Quantum Gravity*, 31(21), 215003.

Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia: Society for Industrial and Applied Mathematics.

Doblinger, G. (1995). Computationally efficient speech enhancement by spectral minima tracking in subbands. In *Proceedings of Eurospeech* (pp. 1513–1516).

Duncan, P. M. (2012). Microseismic monitoring for unconventional resource development. *Geohorizons*, 26–30.

Eaton, D. W., van der Baan, M., Birkelo, B., & Tary, J.-B. (2014). Scaling relations and spectral characteristics of tensile microseisms: Evidence for opening/closing cracks during hydraulic fracturing. *Geophysical Journal International*, 196(3), 1844–1857.

Eisner, L., Hulsey, B. J., Duncan, P., Jurick, D., Werner, H., & Keller, W. (2010). Comparison of surface and borehole locations of induced seismicity. *Geophysical Prospecting*, 58(5), 809–820.

Haldorsen, J. B. U., Miller, D. E., & Walsh, J. J. (1994). Multichannel Wiener deconvolution of vertical seismic profiles. *Geophysics*, 59(10), 1500–1511.

Han, J., & van der Baan, M. (2015). Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding. *Geophysics*, 80(6), KS69–KS80.

Haykin, S. (2002). *Adaptive filter theory* (4th ed.). Upper-Saddle River: Prentice Hall.

Hons, M. S., & Stewart, R. R. (2006). Transfer functions of geophones and accelerometers and their effects on frequency content and wavelets. CREWES Research Report, p. 18.

Huang, W., Wang, R., Chen, Y., Li, H., & Gan, S. (2016). Damped multichannel singular spectrum analysis for 3D random noise attenuation. *Geophysics*, 81(4), V261–V270.

Iqbal, N., Zerguine, A., Kaka, S., & Al-Shuhail, A. (2016). Automated SVD filtering of time-frequency distribution for enhancing the SNR of microseismic/microquake events. *Journal of Geophysics and Engineering*, 13(6), 964–973.

Kendall, M., Maxwell, S., Foulger, G., Eisner, L., & Lawrence, Z. (2011). Microseismicity: Beyond dots in a box Introduction. *Geophysics*, 76(6), WC1–WC3.

Khadhraoui, B., & Özbek, A. (2013). Multicomponent time-frequency noise attenuation of microseismic data. In *75th EAGE Conference and Exhibition incorporating SPE EUROPEC 2013*, pp. 200–204.

Kimiaefar, R., Siahkoohi, H. R., Hajian, A. R., & Kalhor, A. (2016). Seismic random noise attenuation using artificial neural network and wavelet packet analysis. *Arabian Journal of Geosciences*, 9(3), 234.

Liu, E., Zhu, L., Govinda Raj, A., McClellan, J. H., Al-Shuhail, A., Kaka, S. I., & Iqbal, N. (2017). Microseismic events

enhancement and detection in sensor arrays using autocorrelation-based filtering. *Geophysical Prospecting*.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.

Mallinson, I., Bharadwaj, P., Schuster, G., & Jakubowicz, H. (2011). Enhanced refractor imaging by supervirtual interferometry. *The Leading Edge*, 30(5), 546–560.

Martin, R. (2001). Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Transactions on Speech and Audio Processing*, 9(5), 504–512.

Maxwell, S. C. (2011). What does microseismicity tells us about hydraulic fractures? In *SEG Technical Program Expanded Abstracts 2011*, pp. 1565–1569. Society of Exploration Geophysicists.

Maxwell, S. C., Shemeta, J. E., Campbell, E., & Quirk, D. J. (2008). Microseismic deformation rate monitoring. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Mendecki, A. (1993). Real time quantitative seismology in mines. In *3rd International Symposium on Rockbursts and Seismicity in Mines* (pp. 287–295). Canada: Kingston.

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. (1996). *Wavelet toolbox for use with MATLAB*. Natick, MA: The MathWorks.

Mousavi, S. M., Horton, S. P., Langston, C. A., & Samei, B. (2016a). Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression. *Geophysical Journal International*, 207(1), 29–46.

Mousavi, S. M., & Langston, C. (2016a). Fast and novel microseismic detection using time-frequency analysis. In *SEG Technical Program Expanded Abstracts 2016*, pp. 2632–2636. Society of Exploration Geophysicists.

Mousavi, S. M., & Langston, C. A. (2016b). Adaptive noise estimation and suppression for improving microseismic event detection. *Journal of Applied Geophysics*, 132, 116–124.

Mousavi, S. M., & Langston, C. A. (2016c). Hybrid seismic denoising using higherorder statistics and improved Wavelet block thresholding. *Bulletin of the Seismological Society of America*, 106(4), 1380–1393.

Mousavi, S. M., & Langston, C. A. (2017). Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data. *Geophysics*, 82(4), V211–V227.

Mousavi, S. M., Langston, C. A., & Horton, S. P. (2016b). Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *Geophysics*, 81(4), V341–V355.

Peacock, K. L., & Treitel, S. (1969). Predictive deconvolution: Theory and practice. *Geophysics*, 34(2), 155–169.

Pearson, C. (1981). The relationship between microseismicity and high pore pressures during hydraulic stimulation experiments in low permeability granitic rocks. *Journal of Geophysical Research: Solid Earth*, 86(B9), 7855–7864.

Proakis, J. (1985). Probability, random variables and stochastic processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(6), 1637–1637.

Proakis, J. G., & Manolakis, K. D. (2006). *Digital signal processing* (4th ed.). Upper Saddle River, NJ: Prentice-Hall Inc.

Rangachari, S., & Loizou, P. C. (2006). A noise-estimation algorithm for highly non-stationary environments. *Speech Communication*, 48(2), 220–231.

Rilling, G., Flandrin, P., Gonalves, P., & Lilly, J. (2007). Bivariate empirical mode decomposition. *IEEE Signal Processing Letters*, *14*(12), 936–939.

Sabbione, J. I., Sacchi, M. D., & Velis, D. R. (2015). Radon transform-based microseismic event detection and signal-to-noise ratio enhancement. *Journal of Applied Geophysics*, *113*, 51–63.

Sabbione, J. I., & Velis, D. R. (2013). A robust method for microseismic event detection based on automatic phase pickers. *Journal of Applied Geophysics*, *99*, 42–50.

Sabbione, J. I., Velis, D. R., & Sacchi, M. D. (2013). Microseismic data denoising via an apex-shifted hyperbolic Radon transform. In *SEG Technical Program Expanded Abstracts 2013*, pp. 2155–2161. Society of Exploration Geophysicists.

Sayed, A. H. (2008). *Adaptive filters*. Hoboken, NJ: Wiley.

Shemeta, J., & Anderson, P. (2010). It's a matter of size: Magnitude and moment estimates for microseismic data. *The Leading Edge*, *29*(3), 296–302.

Simpson, D., Leith, W., & Scholz, C. (1988). Two types of reservoir-induced seismicity. *Bulletin of the Seismological Society of America*, *78*(6), 2025–2040.

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Annals of Statistics*, *9*(6), 1135–1151.

Vera Rodriguez, I., Bonar, D., & Sacchi, M. (2012). Microseismic data denoising using a 3C group sparsity constrained time-frequency transform. *Geophysics*, *77*, V21–V29.

Verdon, J. (2011). Microseismic monitoring and geomechanical modeling of storage in subsurface reservoirs. *Geophysics*, *76*(5), Z102–Z103.

Wang, J., Tilmann, F., White, R. S., & Bordoni, P. (2009). Application of frequency-dependent multichannel Wiener filters to detect events in 2D three-component seismometer arrays. *Geophysics*, *74*(6), V133–V141.

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