Signal Processing 137 (2017) 373–386

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Seismic signal recovery based on Earth Q model

Deborah Pereg*, Israel Cohen

Technion – Israel Institute of Technology, Israel

ARTICLE INFO

Article history:
Received 23 August 2016
Revised 4 February 2017
Accepted 25 February 2017
Available online 28 February 2017

Keywords:
Seismic inversion
Seismic signal
Sparse reflectivity
Reflectivity estimation
Earth Q model

ABSTRACT

We consider the problem of recovering the underlying reflectivity signal from its seismic trace, taking into account the attenuation and dispersion propagation effects of the reflected waves, in noisy environments. We present an efficient method to perform seismic time-variant inversion based on the earth Q-model. We derive theoretical bounds on the recovery error, and on the localization error. It is shown that the solution consists of recovered spikes which are relatively close to spikes in the true reflectivity signal. In addition, we prove that any redundant spike in the solution, which is far from the correct support, will have small energy. The robustness of our technique is demonstrated using synthetic and real data examples.

1. Introduction

The problem of decomposing a signal into its building blocks (atoms) [1] is very common in many fields in signal processing, such as: image processing [2], compressed sensing [3], radar [4], ultrasound imaging [5], seismology [6–8] and more. In seismic inversion, a short duration pulse (the wavelet) is transmitted from the earth surface. The reflected pulses from the ground are received by a sensor array and processed into a seismic image [9]. Since reflections are generated at discontinuities in the medium impedance, each seismic trace (a column in the seismic two-dimensional (2D) image) can be modeled as a weighted superposition of pulses further degraded by additive noise. Our task is to recover the earth layers structure (the reflectivity) hidden in the observed seismic image.

Previous works tried to solve the seismic inversion problem by separating the seismic 2D image into independent vertical one-dimensional (1D) deconvolution problems. The wavelet is modeled as a 1D time-invariant signal in both horizontal and vertical directions. Each reflectivity channel (column) appears in the vertical direction as a sparse spike train where each spike is a reflector that corresponds to a boundary between two layers (two different acoustic impedances) in the ground. Then, each reflectivity channel is estimated from the corresponding seismic trace observation apart from the other channels [6–14]. Utilization of sparse seismic inversion methods - based on ℓ_1 minimization problem solving - can yield stable reflectivity solutions [7,12,13,15,16]. These ℓ_1-type methods and their resolution limits are studied thoroughly in signal processing and statistics research [5,17–25].

Multichannel deconvolution methods [26–34] take into consideration the horizontal continuity of the seismic reflectivity. Heimer et al. [31] propose a method based on Markov Bernoulli random field (MBRF) modeling. The Viterbi algorithm [35] is applied to the search of the most likely sequences of reflectors concatenated across the traces by legal transitions. Ram et al. [33] also propose two multichannel blind deconvolution algorithms for the restoration of 2D seismic data. These algorithms are based on the Markov–Bernoulli–Gaussian (MBG) reflectivity model. Each reflectivity channel is estimated from the corresponding observed seismic trace, taking into account the estimate of the previous reflectivity channel or both estimates of the previous and the following neighboring columns. The procedure is carried out using a slightly modified maximum posterior mode (MPM) algorithm [36].

Although the typical seismic wavelet is time-variant, many inversion methods rely on a model which does not take into consideration time-depth variations in the waveform. However, the wave absorption effects are not always negligible as the conventional assumption claim. Seismic inverse Q-filtering [37–40] aims to compensate for the velocity dispersion and energy absorption which causes phase and amplitude distortions of the propagating and reflected acoustic waves. The process of inverse Q filtering consists of amplitude compensation and phase correction which enhance the resolution and increase the signal-to-noise ratio (SNR). Yet, this process is generally computationally expensive and sometimes even impractical.

Nonstationary deconvolution methods aim to deconvolve the seismic data and also compensate for energy absorption, without

* Corresponding author.
E-mail address: deborahp@tx.technion.ac.il (D. Pereg).

http://dx.doi.org/10.1016/j.sigpro.2017.02.016
0165-1684/

knowing Q. Margrave et al. [41] developed the Gabor decovolution algorithm. Chai et al. [42] also propose a method called nonstationary sparse reflectivity inversion (NSRI) to retrieve the reflectivity signal from nonstationary data without inverse Q filtering. Li et al. [43] propose a nonstationary deconvolution algorithm based on spectral modeling [44] and variable-step-sampling (VSS) hyperbolic smoothing.

We propose a novel robust algorithm for recovery of the underlying reflectivity signal from the seismic data without a preprocessing stage of inverse Q filtering. We prove that the solution of a convex optimization problem, which takes into consideration a time-variant signal model, results in a stable recovery. In addition we answer the following questions: To what accuracy can we recover each reflectivity spike? How does this accuracy depend on the noise level, the amplitude of the spike, the medium Q constant and the wavelet's shape? We prove that the recovery error is proportional to the noise level. We also show how the error is affected by degradation. The algorithm is applied to synthetic and real seismic data. Our experiments affirm the theoretical results and demonstrate that the suggested method reveals reflectors amplitudes and locations with high precision.

The paper is organized as follows. In Section 2, we review the basic theory of earth Q model and the seismic inversion problem. In Section 3, we present the main theoretical results. Section 4 presents numerical experiments and real data results. Finally, in Section 5, we conclude and discuss further research.

2. Signal model

2.1. Reflectivity model

We assume the earth structure is stratified, so that reflections are generated at the boundaries between different impedance layers. Therefore, each 1D channel (column) in the unknown 2D reflectivity signal can be formulated as a sparse spike train

$$x(t) = \sum_m c_m \delta(t - t_m), \tag{1}$$

where $\delta(t)$ denotes the Dirac delta function and $\Sigma_m |c_m| < \infty$. The set of delays $T = \{t_m\}$ and the real amplitudes $\{c_m\}$ are unknown.

In the discrete setting, assuming the sampling interval is $1/N$ for a given integer sampling rate N, and that the set of delays $T = \{t_m\}$ lie on the grid k/N, $k \in \mathbb{Z}$, i.e., $t_m = k_m/N$ where $k_m \in \mathbb{Z}$

$$x[k] = \sum_m c_m \delta[k - k_m], \tag{2}$$

where $\delta[k]$ denotes the Kronecker delta function (see [45]).

We consider a seismic discrete trace of the form

$$y[k] = y(k/N) = \sum_m c_m g_{\sigma,m}\left(\frac{t - t_m}{N}\right) = \sum_m c_m g_{\sigma,m}[k - k_m], \tag{3}$$

where $\{g_{\sigma,\,m}\}$ is a known set of kernels (pulses) for a possible set of time delays $T = \{t_m\}$, and a known scaling parameter $\sigma > 0$. In Section 2.3 we discuss specific requirements for $\{g_{\sigma,\,m}\}$.

A time-invariant model assumes for simplicity that all kernels are identical, i.e., $g_{\sigma,m}(t) = g(\frac{t}{\sigma})\ \forall m$ [5,23]. Hence, the model can be represented as a convolution model. However, the shape and energy of each reflected pulse highly depends on its corresponding reflector's depth in the ground. Therefore, an accurate model should take into consideration a set of kernels $\{g_{\sigma,\,m}\}$ which consists of different pulses.

In noisy environments we consider a discrete seismic trace of the form

$$y[k] = \sum_m c_m g_{\sigma,m}[k - k_m] + n[k], \quad |n|_1 \le \delta, \tag{4}$$

where $n[k]$ is additive noise with $|n|_1 = \sum_k |n[k]| \le \delta$. Our objective is to estimate the true support $K = \{k_m\}$ and the spikes' amplitudes $\{c_m\}$ from the observed seismic trace $y[k]$.

2.2. Earth Q model

We assume a source waveform $s(t)$ defined as the real-valued Ricker wavelet.

$$s(t) = \left(1 - \frac{1}{2}\omega_0^2 t^2\right) \exp\left(-\frac{1}{4}\omega_0^2 t^2\right), \tag{5}$$

where ω_0 is the most energetic (dominant) radial frequency [46]. We define the scaling parameter as $\sigma = \omega_0^{-1}$. Wang [47] showed that given a travel time t_m, the reflected wave can be modeled as

$$u(t) = \mathrm{Re}\left\{\frac{1}{\pi}\int_0^\infty S(\omega)\exp[j(\omega t - \kappa r(\omega))]d\omega\right\}, \tag{6}$$

where $S(\omega)$ is the Fourier transform of the source waveform $s(t)$,

$$\kappa r(\omega) \triangleq \left(1 - \frac{j}{2Q}\right)\left|\frac{\omega}{\omega_0}\right|^{-\gamma}\omega t_m, \tag{7}$$

$$\gamma \triangleq \frac{2}{\pi}\tan^{-1}\left(\frac{1}{2Q}\right) \approx \frac{1}{\pi Q}, \tag{8}$$

and Q is the medium quality factor, which is assumed to be frequency independent [37]. Kjartansson defined Q as the portion of energy lost during each cycle or wavelength.

Therefore, the expression of the earth Q filter consists of two exponential operators that express the phase effect (caused by velocity dispersion) and the amplitude effect (caused by energy absorption)

$$U(t - t_m, \omega) = U(t, \omega)\exp\left(-j\left|\frac{\omega}{\omega_0}\right|^{-\gamma}\omega t_m\right) \times \exp\left(-\left|\frac{\omega}{\omega_0}\right|^{-\gamma}\frac{\omega t_m}{2Q}\right). \tag{9}$$

Summing these plane waves we get the time-domain seismic signal

$$u(t - t_m) = \frac{1}{2\pi}\int U(t - t_m, \omega)d\omega. \tag{10}$$

We can now define the known set of kernels (pulses) $\{g_{\sigma,\,m}\}$ for the seismic setting

$$g_{\sigma,m}(t - t_m) = u(t - t_m)|_{\sigma=\omega_0^{-1}}. \tag{11}$$

2.3. Admissible kernels and separation constant

To be able to quantify the waves decay and concavity we recall two definitions from previous works [5,23]:

Definition 2.1. A kernel g is **admissible** if it has the following properties:

1. $g \in \mathbb{R}$ is real and even.
2. Global Property: There exist constants $C_l > 0$, $l = 0, 1, 2, 3$, such that $\left|g^{(l)}(t)\right| \le \frac{C_l}{1+t^2}$, where $g^{(l)}(t)$ denotes the l^{th} derivative of g.
3. Local Property: There exist constants ε, $\beta > 0$ such that
   (a) $g(t) > 0$ for all $|t| \le \varepsilon$ and $g(\varepsilon) > g(t)$ for all $|t| \ge \varepsilon$.
   (b) $g^{(2)}(t) < -\beta$ for all $|t| \le \varepsilon$.

In other words, the kernel and its first three derivatives are decaying fast enough, and the kernel is concave near its midpoint.

Definition 2.2. A set of points $K \subset \mathbb{Z}$ is said to satisfy **the minimal separation condition** for a kernel dependent $\nu > 0$, a given scaling $\sigma > 0$ and a sampling spacing $1/N > 0$ if

$$\min_{k_i, k_j \in K, i \ne j} |k_i - k_j| \ge N\nu\sigma.$$

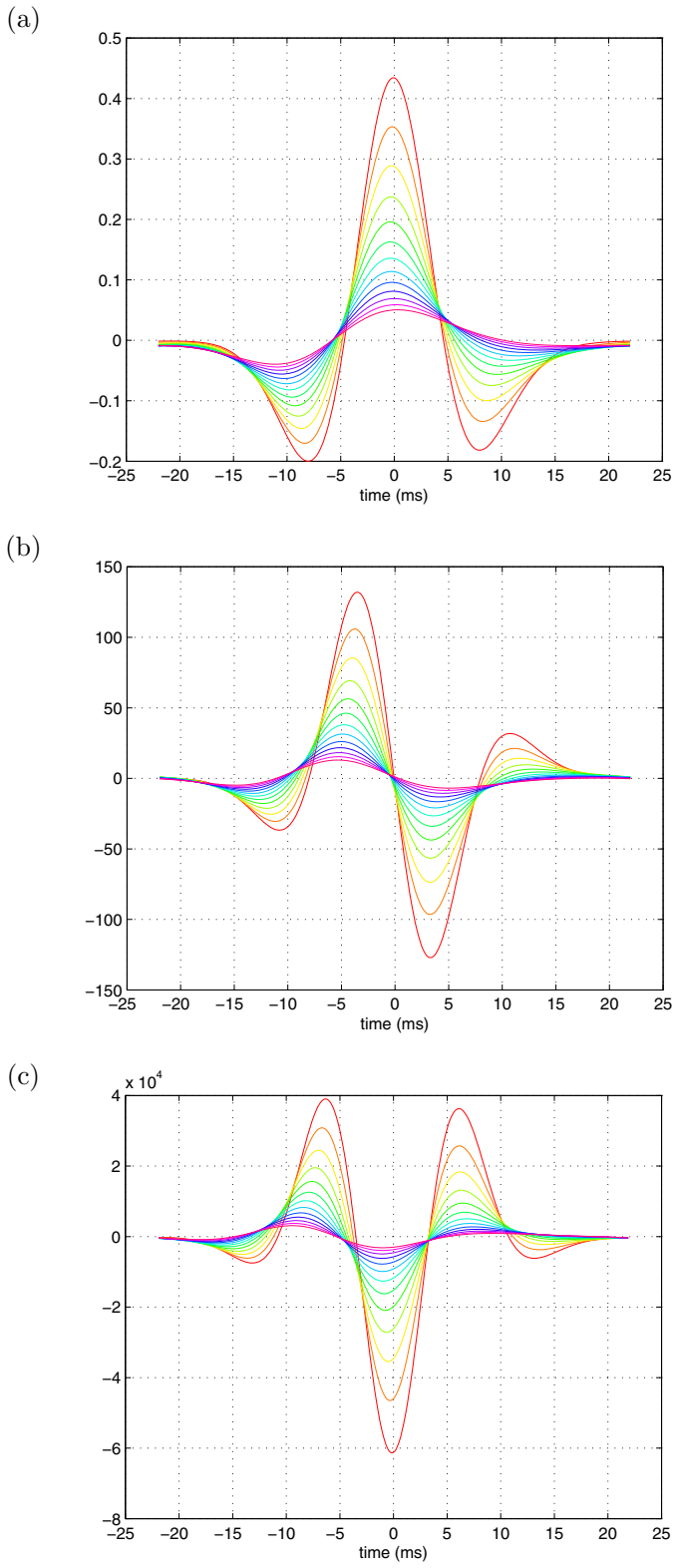

Fig. 1. Centered synthetic reflected wavelets and their derivatives, $Q = 125, \omega_0 = 100\pi$ (50Hz) (a) $g_{\sigma,m}(t)$; (b) $g^{(1)}_{\sigma,m}(t)$; (c) $g^{(2)}_{\sigma,m}(t)$.

where $\nu\sigma$ is the smallest time interval between two reflectors with which we can still recover two distinct spikes, and ν is called the separation constant.

Fig. 1 presents an example of the attenuating wavelets $g_{\sigma,m}(t)$ and their derivatives, $g^{(1)}_{\sigma,m}(t)$ and $g^{(2)}_{\sigma,m}(t)$ for $Q = 125$ and $t_m = 100, 250, 400, ..., 1900$ ms (increment of 150 ms). $\omega_0 = 100\pi$ (50 Hz). The pulses and their derivatives are moved to the origin so that it can be seen that there is a common value of ε and β. Meaning that, for a sequence of kernels $g_{\sigma,m}(t)$ as described in (11), there exist two possible parameters (ε_m, β_m) that determine the concavity of the reflected wave $g_{\sigma,m}(t)$, as defined in Definition 2.1, such that there are two common constants $\varepsilon, \beta > 0$ for all reflected waves. In other words

$$\varepsilon_m = \varepsilon \quad \forall m \tag{12}$$

and

$$\beta_m = \beta \quad \forall m. \tag{13}$$

The reflected waves $g_{\sigma,m}(t)$ are not symmetric, but remain flat at the origin, i.e., $g^{(1)}_{\sigma,m}(t) \approx 0$. So, it can be said that each of the reflected waves $g_{\sigma,m}(t)$ is approximately an admissible kernel, and all of these waves share two common parameters $\varepsilon, \beta > 0$.

We would make one more assumption: $g_{\sigma,m}(\varepsilon) > |g_{\sigma,m}(t)|$ for all $|t| \geq \varepsilon$. Meaning that for $|t| \geq \varepsilon$ the absolute value of the kernel does not increase beyond its value in $t = \varepsilon$.

3. Seismic recovery

3.1. Recovery method and recovery-Error bound

The recovery of the seismic reflectivity could be achieved by solving the optimization problem presented in the following theorem. In addition, we also derive a bound on the recovery error.

Theorem 1. *Let y be of the form of (4) and let $\{g_{\sigma,m}\}$ be a set of admissible kernels as defined in Definition 2.1. If K satisfies the separation condition of Definition 2.2 for $N > 0$ then the solution $\hat{x}$ of*

$$\min_{x \in l_1(Z)} ||x||_{\ell_1} \quad \text{subject to} \quad ||y[k] - \sum_m c_m g_{\sigma,m}[k - k_m]||_{\ell_1} \leq \delta \tag{14}$$

satisfies

$$||\hat{x} - x||_{\ell_1} \leq \frac{4\rho}{\beta\gamma_0}\delta \tag{15}$$

$$\rho \triangleq \max\left\{\frac{\gamma_0}{\varepsilon^2}, (N\sigma)^2\alpha_0\right\}$$

where

$$\alpha_0 = \max_m g_{\sigma,m}(0), \quad \gamma_0 = \min_m g_{\sigma,m}(0).$$

The dependance of x on the time k is not written for simplicity.

Proof. see Appendix A. □

Remarks

- This result guarantees that under the separation condition in Definition 2.2, a signal of the form of (4), can be recovered by solving the ℓ_1 optimization problem formulated in (14). Moreover, a theoretical analysis of the recovered solution ensures that the error is bounded by a relatively small value, which depends mainly on the noise level and on the attenuation of the wavelets and is expressed through the parameters Q and β.
- In the noiseless case where $\delta = 0$, the recovery is perfect. One would probably expect that the recovered solution would slightly deviate from the true one, yet this is not the case. This result does not depend on whether the spikes amplitude are very small or very large.

- If $\gamma_0 = \alpha_0$ we have the time-invariant case

$$||\hat{x} - x||_{\ell_1} \leq \frac{4}{\beta} \max\left\{\frac{1}{\varepsilon^2}, (N\sigma)^2\right\}\delta$$

As expected, in the time-invariant case our result reduces into previous work results [5,23]. The recovery error is proportional to the noise level δ, and small values of β (flat kernels) result in larger errors.

- In the time-variant setting most cases comply with $\frac{\gamma_0}{\varepsilon^2} < (N\sigma)^2\alpha_0$,
Then, the recovery error is bounded by

$$||\hat{x} - x||_{\ell_1} \leq \frac{4(N\sigma)^2}{\beta}\frac{\alpha_0}{\gamma_0}\delta$$

A smaller Q (which corresponds to a stronger degradation) results in higher $\frac{\alpha_0}{\gamma_0}$ ratio and smaller β values. We will hereafter refer to the ratio $\frac{\alpha_0}{\gamma_0}$ as the degradation ratio. Hence, the bound on the error in a time-variant environment implies that the error increases as Q gets smaller, which corresponds to a higher degradation ratio $\frac{\alpha_0}{\gamma_0}$. As in the time-invariant case, the error is linear with respect to the noise level δ. Also, the error is sensitive to the flatness of the kernel near the origin. Namely, small β results in an erroneous recovery.

3.2. Resolution bounds

Theorem 2. *Assume* $\hat{x}[k] = \sum_m \hat{c}_m \delta[k - \hat{k}_m]$ *is the solution of (14) where* $\hat{K} \triangleq \{\hat{k}_m\}$ *is the support of the recovered signal.*

Let y *be of the form of* $y[k] = \sum_m c_m g_{\sigma,m}[k - k_m] + n[k]$, $|n|_1 \leq \delta$ *and let* $\{g_{\sigma,m}\}$ *be a set of admissible kernels with two common parameters* ε, $\beta > 0$, *with* $\varepsilon \geq \tilde{\varepsilon} = \sqrt{\frac{\alpha_0}{\tilde{C}_2 + \beta/4}}$

If K *satisfies the separation condition for* $N > 0$, *then the solution* $\hat{x}$ *satisfies:*

1. $\sum_{\hat{k}_m \in \hat{K}: |\hat{k}_m - k_m| > N\varepsilon, \forall k_m \in K} |c_m| \leq \frac{2D_3\alpha_0}{\beta\varepsilon^2}\delta$
Any redundant spike in $\hat{K}$ *which is far from the correct support* K *will for sure have small energy.*
2. *For any* $k_m \in K$ *if* $|c_m| \geq D_4$, *then there exist* $\hat{k}_m \in \hat{K}$ *such that*

$$(\hat{k}_m - k_m)^2 \leq \frac{2D_3(N\sigma)^2\alpha_0}{\beta\left(|c_m| - D_4\right)}\delta.$$

where

$$D_4 = \frac{2\delta}{\beta}\left(\frac{2\rho}{\gamma_0} + D_3\alpha_0 \max\left\{\frac{1}{\varepsilon^2}, \frac{C_{2,m}}{(N\sigma)^2 g_m(0)}\right\}\right),$$

$$D_3 = \frac{3\nu^2(3\gamma_2\nu^2 - \pi^2\tilde{C}_2) + \frac{12\pi^2\tilde{C}_1^2}{\beta\gamma_0}(1 + \frac{\pi^2}{6\nu^2})\rho}{(3\gamma_2\nu^2 - \pi^2\tilde{C}_2)(3\gamma_0\nu^2 - 2\pi^2\tilde{C}_0)}.$$

and

$$\tilde{C}_l = \max_m C_{l,m}, \quad l = 0, 1, 2, 3.$$

This implies that for any $k_m \in K$ *with sufficiently large amplitude* c_m, *under the separation condition, the recovered support location* $\hat{k}_m \in \hat{K}$ *is close to the original one. The solution* $\hat{x}$ *consists of a recovered spike near any spike of the true reflectivity signal.*

Proof. see Appendix B. □

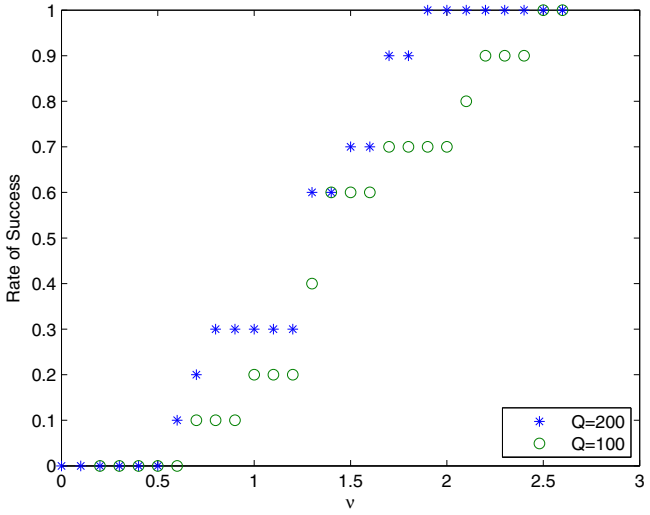

Fig. 2. Support detection vs. the separation constant ν. Rate of success is the average number of perfect recoveries out of 10 experiments.

4. Experimental results

4.1. Synthetic Data

We conducted various experiments in order to confirm the theoretical results. To solve the ℓ_1 minimization in (14) we used CVX [48].

First, we try to estimate the minimal separation constant ν for various Q values. We generate a synthetic reflectivity column, with sampling time $T_s = 2$ ms. The reflectivity is statistically modeled as a zero-mean Bernoulli–Gaussian process [28]. The support was drawn from a Bernoulli process with $p = 0.2$ of length $L_r = 220$ taps, , and the amplitudes were drawn from an i.i.d normal distribution with standard deviation $\nu = 10$. Then, we create the synthetic seismic trace in a noise-free environment, and try to recover the reflectivity by solving (14). Namely, we increase ν until we get an exact recovery in the noise-free setting. Fig. 2 presents the results for $Q = 100, 200$. The initial wavelet was a Ricker wavelet with $\omega_0 = 100\pi$, i.e., 50 Hz. We repeat the experiment 10 times for each value of ν. The success rate is 1 if the support's recovery is perfect for all 10 experiments. As can be seen, the minimal separation constant for $Q = 200$ is $\nu = 1.9$ whereas for $Q = 100$ we have $\nu = 2.5$.

Fig. 3 presents the recovery error $||\hat{x} - x||_{\ell_1}$ as a function of the noise level δ for different Q values - $Q = \infty, 500, 200, 100$. $Ts = 4$ ms and $L_r = 176$. As in Fig. 2, the reflectivity is statistically modeled as a zero-mean Bernoulli-Gaussian process. Under the separation condition, the minimum distance between two spikes satisfies the minimal separation condition. The reflectivity is shown in Fig. 4(a). The initial wavelet was a Ricker wavelet with $\omega_0 = 140\pi$, i.e., 70 Hz. Two seismic traces with SNR$= \infty$ and SNR $= 15.5$dB, are shown in Fig. 4(b) and (c) respectively. The recovered signals from this traces are shown in Fig. 4(d) and (e). As can be seen in Fig. 3 the error is linear with respect to the noise. This implies that the bound we derived in Theorem 1 is reasonable. The theoretical bound is always greater or equal to the empirical error. As Q gets smaller, β - which is common to all reflected pulses - becomes significantly smaller. Hence, the theoretical bound slope becomes significantly larger compared with the empirical one. It can be seen also in the experimental results that as Q gets smaller the error gets bigger. Table 1 presents the theoretical and practical parameters.

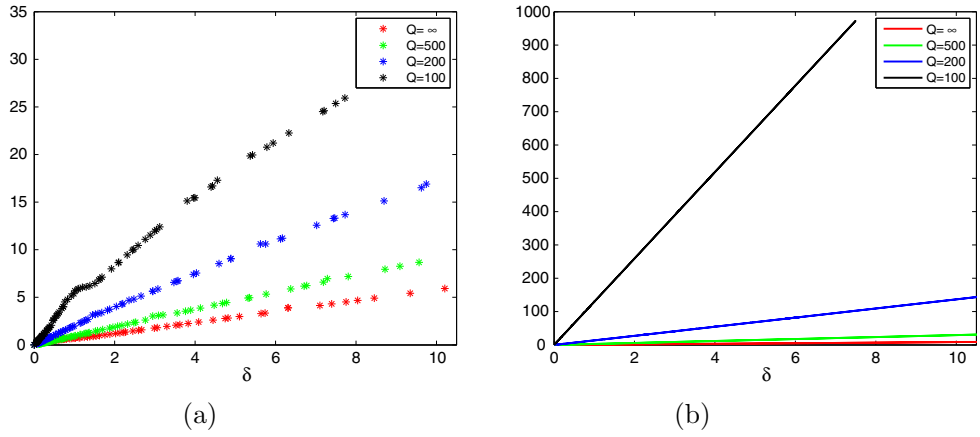

Fig. 3. Recovery error $||\hat{x}-x||_{\ell_1}$ as a function of noise level δ for $Q=\infty, 500, 200, 100$. (a) Experimental results ; (b) Theoretical bounds.

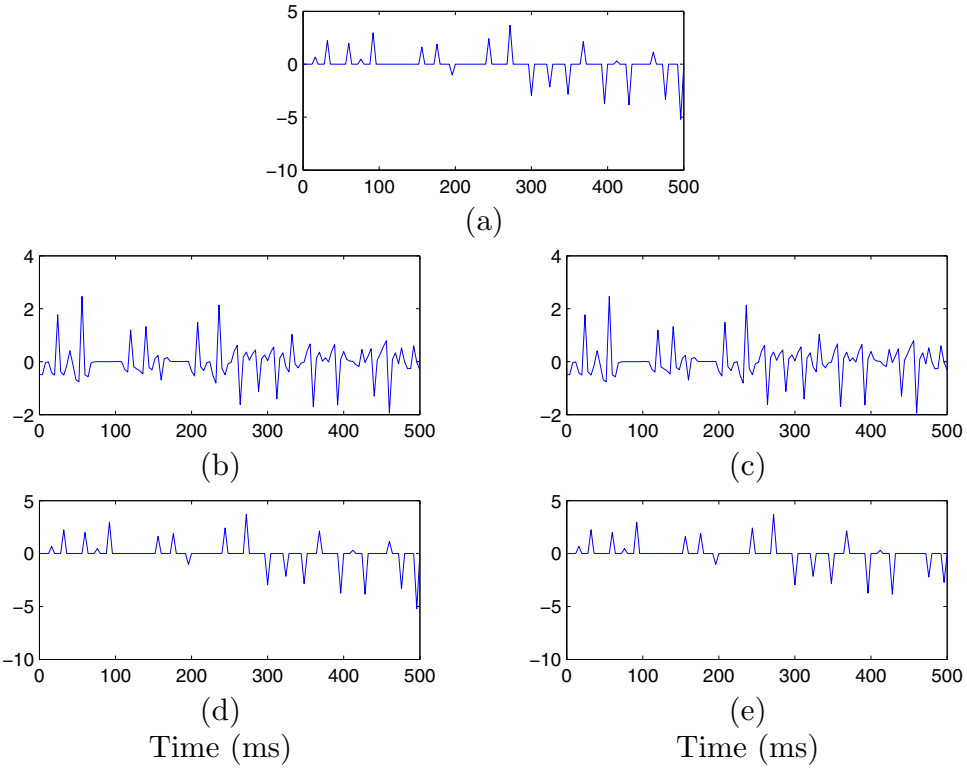

Fig. 4. 1D synthetic tests of (a) True reflectivity. (b),(c) Synthetic trace with 50 Hz Ricker wavelet and SNR= ∞, 15.5 dB respectively, $Q=200$. (d),(e) Recovered 1D channel of reflectivity signal.

Table 1
Synthetic example: theoretical and estimated parameters: Q, the degradation ratio $\frac{\alpha_0}{\gamma_0}$, β, $\frac{4(N\sigma)^2}{\beta}\frac{\alpha_0}{\gamma_0}$ - the bound slope computed from known parameters (by Theorem 1 $||\hat{x}-x||_{\ell_1} \leq \frac{4(N\sigma)^2}{\beta}\frac{\alpha_0}{\gamma_0}\delta$), and the estimated slope computed from the experimental results in Fig. 3(a).

Q	$\frac{\alpha_0}{\gamma_0}$	β	$\frac{4(N\sigma)^2}{\beta}\frac{\alpha_0}{\gamma_0}$	Estimated slope
∞	1	1.5	0.862	0.567
500	1.75	0.77	2.94	0.89
200	3.8	0.36	13.67	1.71
100	9.44	0.094	129.7	3.53

We compare the proposed solution to the blind deconvolution SOOT algorithm of Repetti et al. [16]. Fig 5. presents the results with noise level $\sigma=0.01$ (SNR $=12.9$ dB), $Q=500$, $Ts=4$ ms, and an initial Ricker wavelet with $\omega_0=50\pi$, i.e., 25 Hz. The original reflectivity section is depicted in Fig 5.(a). The estimated reflectivities, obtained by SOOT, and by solving the ℓ_1 minimization in (14) using CVX [48] , for the seismic data in Fig. 5(b), are shown in Fig. 5(c) and (d) respectively. The results demonstrate that sparse recovery methods that do not take into consideration the attenuating and broadening nature of the wavelet, tend to annihilate small reflectivity spikes, especially in the deeper part of the reflectivity section.

Solving the ℓ_1 minimization in (14) using CVX [48], the average processing time of a data set of 100 $\times$ 100 on Intel(R)Core(TM)i7-5600U@2.60 GHz is 40.8 s.

4.2. Real data

We applied the proposed method, to real seismic data from a small land 3D survey in North America (courtesy of GeoEnergy

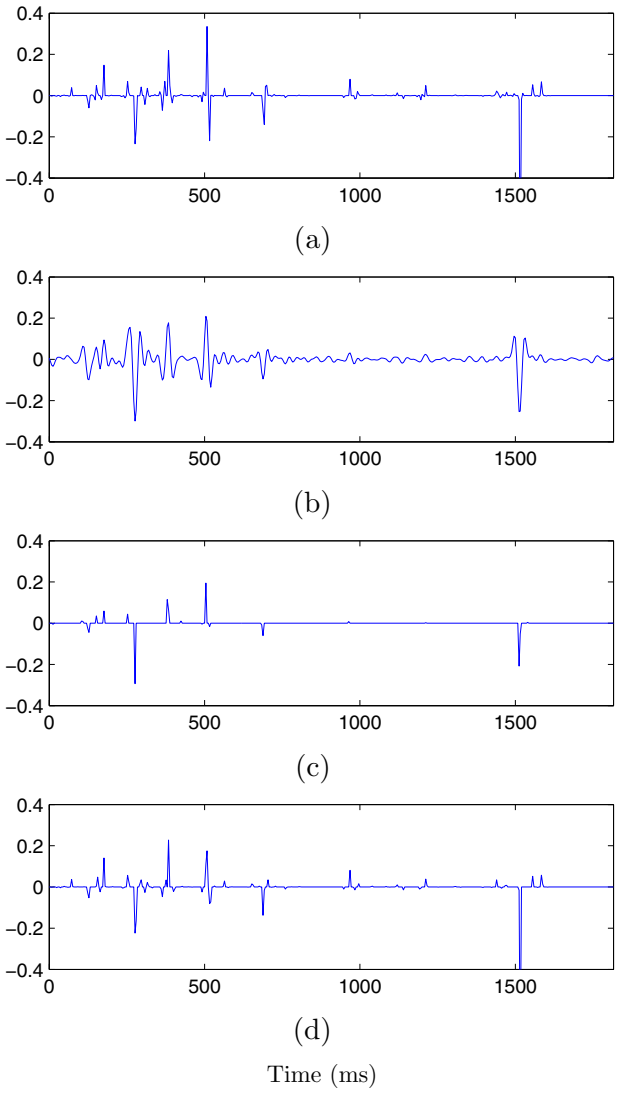

Fig. 5. 1D synthetic tests of (a) True reflectivity. (b) Synthetic trace with 25 Hz Ricker wavelet and SNR= 12.9 dB, $Q = 500$. (c) Recovered 1D channel of reflectivity signal with SOOT. (d) Recovered 1D channel of reflectivity signal with the proposed time-variant model.

Inc., TX) of size 380 × 160, shown in Fig. 6(a). The time interval is 2 ms. Assuming an initial Ricker wavelet with $\omega_0 = 140\pi$ (70 Hz). We estimated $Q = 80$ using common midpoints (CMP) as described in [49]. Then, using (6)–(11) we estimated all possible kernels and solved (14) using CVX [48]. The recovered reflectivity section is shown in Fig. 6(b). The seismic data reconstructed from the estimated reflectivity using the known sequence $\{g_{\sigma,m}(t)\}$, is shown in Fig. 5(c). Visually analyzing this reflectivity section, it can be seen that the layer boundaries in the estimate are clear and quite continuous and smooth. It can also be seen that the reconstructed seismic data fits the original given observation. Since the ground truth is unknown, in order to measure the accuracy in the location and amplitude of the recovered reflectivity spikes we compute the correlation coefficient between the reconstructed data to the given seismic data. In this example we have $\rho_{s,\hat{s}} = 0.967$, which indicates that the reflectivity is estimated with very high precision. Fig. 7(a) shows the estimated reflectivity considering a time-invariant model, using sparse spike inversion (SSI) [6]. The result for a time-varying model is shown in Fig. 7(b). It can be seen, especially in the lower (deeper) half of the image, that the method introduced in this paper produces much clearer results, since it takes in to account the attenuating and broadening nature of the waves as they travel further into the ground and back. Moreover, in terms of correlation coefficients, for SSI we have $\rho_{s,\hat{s}} = 0.89$, implying that considering a time-varying model indeed yields better results.

5. Conclusions

We have presented a seismic inversion algorithm under time-variant model. The algorithm both promotes sparsity of the solution and also takes into consideration attenuation and dispersion effects resulting in shape distortion of the wavelet. The inversion results are demonstrated on synthetic and real data, under sufficiently high SNR. We derived a bound on the recovery ℓ_1 error and observed that the error increases as Q gets smaller. As in the time-invariant case, the error is proportional to the noise level. Also, the error is sensitive to the flatness of the kernel near the origin. Simulation results confirm the theoretical bound. We also proved that under the separation condition, for any spike with large-enough amplitude the recovered support location is close to the original one. The solution consists of a recovered spike near every spike of the true reflectivity signal. Any redundant spike in the recov-

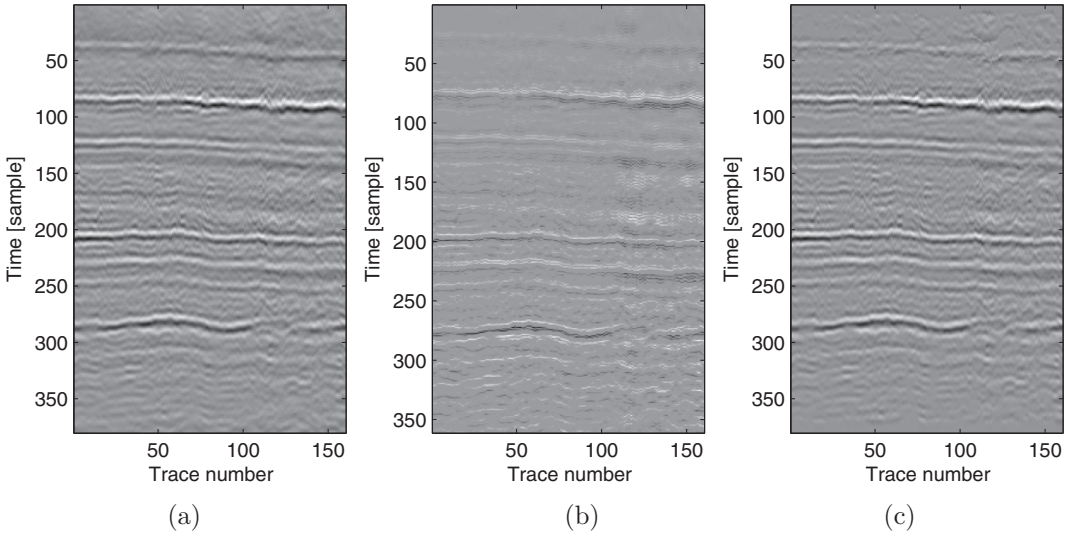

Fig. 6. Real data inversion results: (a) Real seismic data (b) Estimated reflectivity (c) Reconstructed data.

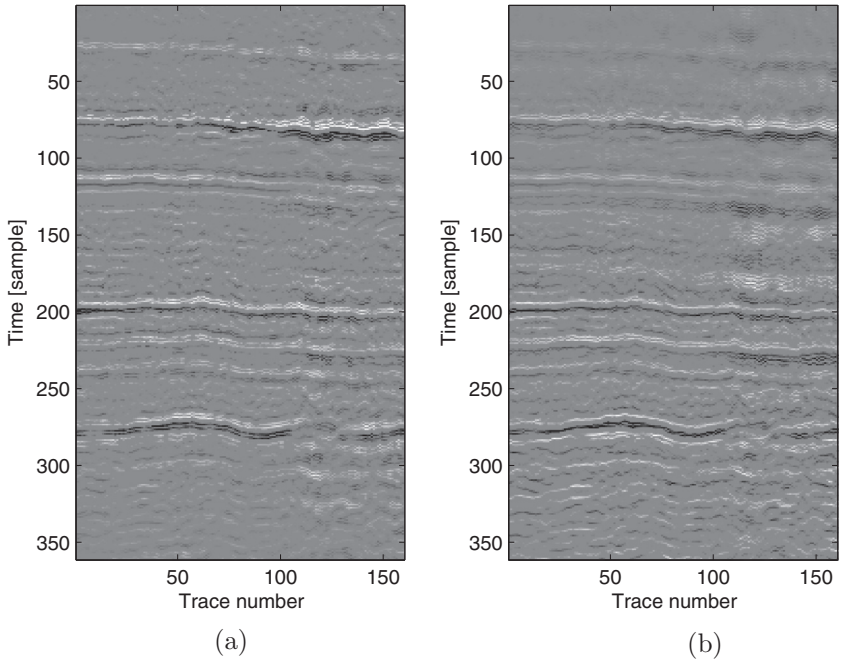

Fig. 7. Real data inversion results: (a) Estimated reflectivity - time-invariant model (SSI) (c) Estimated reflectivity - time-variant model.

ered signal, which is far from the correct support, has small energy. Future research can address the problem of model mismatch. In addition we can elaborate the solution suggested in this paper to non-constant Q layers model.

Appendix A

Proof of Theorem 1. The proof follows the outline of research in [20,23].

Denote $g_m(t) \triangleq g_{\sigma,m}|_{\sigma=1}$. In a similar manner to [20,23], we build a function of the form

$$q(t) = \sum_m a_m g_m(t - t_m) + b_m g_m^{(1)}(t - t_m).$$

The function $q(t)$ satisfies

$$\begin{aligned} q(t_k) &= v_k \quad \forall t_k \in T, \\ q^{(1)}(t_k) &= 0 \quad \forall t_k \in T, \\ |v_k| &= 1; \end{aligned}$$

Its existence enables us to decouple the estimation error at t_m from the amplitude of the rest of the spikes. The magnitude of $q(t)$ reaches a local maximum on the true support. This will in turn enable us to bound the recovery error. □

In the following proof we use the following proposition and two lemmas.

Proposition 3. *Assume a set of delays $T \triangleq \{t_m\}$ that satisfies the separation condition, and let $\{g_m\}$ be a set of admissible kernels as defined in Definition 2.1. Then, there exist coefficients $\{a_m\}$ and $\{b_m\}$ such that*

$$q(t) = \sum_m a_m g_m(t - t_m) + b_m g_m^{(1)}(t - t_m), \tag{16}$$

$$|q(t_k)| = 1 \quad \forall t_k \in T, \tag{17}$$

and

$$q^{(1)}(t_k) = 0 \quad \forall t_k \in T. \tag{18}$$

The coefficients are bounded by

$$||\mathbf{a}||_\infty \leq \frac{3\nu^2}{3\gamma_0\nu^2 - 2\pi^2\tilde{C}_0},$$

$$||\mathbf{b}||_\infty \leq \frac{3\pi^2\tilde{C}_1\nu^2}{(3\gamma_2\nu^2 - \pi^2\tilde{C}_2)(3\gamma_0\nu^2 - 2\pi^2\tilde{C}_0)},$$

where $\mathbf{a} \triangleq \{a_m\}, \mathbf{b} \triangleq \{b_m\}$ are coefficient vectors and

$$\tilde{C}_l = \max_m C_{l,m}, \quad l = 0, 1, 2, 3.$$

We also have

$$a_m \geq \frac{1}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_2 - 2\tilde{C}_2 E(\nu)}}.$$

In other words, if the support is scattered, it is possible to build a function $q(t)$ that interpolates any sign pattern exactly.

Proof. The admissible kernel and its derivatives decay rapidly away from the origin. The proofs of this Proposition, the Theorem and the two Lemmas make repeated use of these facts.

According to (17) and (18)

$$\sum_m a_m g_m(t_k - t_m) + b_m g_m^{(1)}(t_k - t_m) = v_k,$$

and

$$\sum_m a_m g_m^{(1)}(t_k - t_m) + b_m g_m^{(2)}(t_k - t_m) = 0,$$

for all $t_k \in T$, where $v_k \in \mathbb{R}$ so that $|v_k| = 1$.

Therefore, in matrix-vector formulation we express these constraints

$$\begin{bmatrix} G_0 & G_1 \\ G_1 & G_2 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix},$$

where $(G_l)_{k,m} \triangleq g_m^{(l)}(t_k - t_m), l = 0, 1, 2$ and $\mathbf{a} \triangleq \{a_m\}, \mathbf{b} \triangleq \{b_m\}$, $\mathbf{v} \triangleq \{v_k\}$.

We know that the matrix G is invertible if both G_2 and the Schur complement of G_2:

$S = G_0 - G_1 G_2^{-1} G_1$ are invertible [50]. We also know that a matrix A is invertible if there exists $\alpha \neq 0$ such that $||\alpha I - A||_\infty < |\alpha|$, where $||A||_\infty = \max_i \sum_j |a_{i,j}|$. In this case we also know that

$$||A^{-1}||_\infty \leq \frac{1}{|\alpha| - ||\alpha I - A||_\infty}. \tag{19}$$

Denote

$$\alpha_l = \max_m |g_m^{(l)}(0)|, \quad \gamma_l = \min_m |g_m^{(l)}(0)|$$

$$\Delta_l = \alpha_l - \gamma_l = \max_m |g_m^{(l)}(0)| - \min_m |g_m^{(l)}(0)|.$$

We can observe that

$$\begin{aligned} ||\alpha_2 I - G_2||_\infty &= \max_k \Big[\sum_{m \neq k} |g_m^{(2)}(t_k - t_m)| + |g_k^{(2)}(0) - \alpha_2| \Big] \\ &\leq \tilde{C}_2 \max_k \sum_{m \neq k} \frac{1}{1 + (t_k - t_m)^2} + \Delta_2 \\ &\leq \tilde{C}_2 \max_k \sum_{m \neq k} \frac{1}{1 + ((k-m)\nu)^2} + \Delta_2. \end{aligned}$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{1 + (n\nu)^2} \leq E(\nu) \triangleq \frac{\pi^2}{6\nu^2}. \tag{20}$$

$$||\alpha_2 I - G_2||_\infty \leq 2\tilde{C}_2 E(\nu) + \Delta_2 < \alpha_2. \tag{21}$$

Which leads us to,

$$2\tilde{C}_2 \pi^2 < (\alpha_2 - \Delta_2) 6\nu^2. \tag{22}$$

Therefore,

$$\nu^2 > \frac{\tilde{C}_2 \pi^2}{3\gamma_2}. \tag{23}$$

The result is quite intuitive. Small values of γ_2 (small Q values) require a larger separation constant. The flattest kernel determines the global separation requirements for perfect recovery.

Now, we can also derive

$$\begin{aligned} ||\alpha_0 I - S||_\infty &= ||\alpha_0 I - G_0 + G_1 G_2^{-1} G_1||_\infty \\ &\leq ||\alpha_0 I - G_0||_\infty + ||G_1||_\infty^2 ||G_2^{-1}||_\infty. \end{aligned} \tag{24}$$

In a similar manner to (21)

$$||\alpha_0 I - G_0||_\infty \leq 2\tilde{C}_0 E(\nu) + \Delta_0, \tag{25}$$

and since $g_m^{(1)}(0) \approx 0 \ \forall \ m$

$$||G_1||_\infty \leq 2\tilde{C}_1 E(\nu). \tag{26}$$

Using (19) and (21)

$$\begin{aligned} ||G_2^{-1}||_\infty &\leq \frac{1}{|\alpha_2| - ||\alpha_2 I - G_2||_\infty} \leq \frac{1}{\alpha_2 - 2\tilde{C}_2 E(\nu) - \Delta_2} \\ &= \frac{1}{\gamma_2 - 2\tilde{C}_2 E(\nu)}. \end{aligned} \tag{27}$$

So, we have

$$\begin{aligned} ||\alpha_0 I - S||_\infty &\leq 2\tilde{C}_0 E(\nu) + \Delta_0 + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_2 - 2\tilde{C}_2 E(\nu)} \\ &= 2\tilde{C}_0 E(\nu) \Big[1 + \frac{2\tilde{C}_1^2 E(\nu)}{\tilde{C}_0(\gamma_2 - 2\tilde{C}_2 E(\nu))} \Big] + \Delta_0 \\ &\leq 4\tilde{C}_0 E(\nu) + \Delta_0. \end{aligned} \tag{28}$$

The last inequality holds for

$$2\tilde{C}_1^2 E(\nu) \leq \tilde{C}_0(\gamma_2 - 2\tilde{C}_2 E(\nu)).$$

Leading us to

$$3\nu^2 \tilde{C}_0 \gamma_2 \geq \pi^2 (\tilde{C}_1^2 + \tilde{C}_0 \tilde{C}_2).$$

Which yields the condition

$$\nu^2 \geq \frac{\pi^2 (\tilde{C}_1^2 + \tilde{C}_0 \tilde{C}_2)}{3\tilde{C}_0 \gamma_2}. \tag{29}$$

Then, S is invertible if

$$4\tilde{C}_0 E(\nu) + \Delta_0 < \alpha_0,$$

$$4\tilde{C}_0 E(\nu) < \alpha_0 - \Delta_0 = \gamma_0,$$

$$\frac{2\tilde{C}_0 \pi^2}{3\nu^2} < \gamma_0,$$

$$\nu^2 > \frac{2\pi^2 \tilde{C}_0}{3\gamma_0}. \tag{30}$$

Here again it can be observed that small γ_0 values require a larger separation constant ν. Finally we get

$$||\alpha_0 I - S||_\infty \leq 4\tilde{C}_0 E(\nu) + \Delta_0 < \alpha_0. \tag{31}$$

and S is invertible. So we have proved that $q(t)$ exists under certain conditions on the separation constant.

In addition $\mathbf{a}$ and $\mathbf{b}$ are given by

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} G_0 & G_1 \\ G_1 & G_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} S^{-1}\mathbf{v} \\ -G_2^{-1} G_1 S^{-1} \mathbf{v} \end{bmatrix}. \tag{32}$$

Using (19) and (31) we have

$$\begin{aligned} ||\mathbf{a}||_\infty \leq ||S^{-1}||_\infty &\leq \frac{1}{|\alpha_0| - ||\alpha_0 I - S||_\infty} \leq \frac{1}{\alpha_0 - 4\tilde{C}_0 E(\nu) - \Delta_0} \\ &= \frac{1}{\gamma_0 - 4\tilde{C}_0 E(\nu)}. \end{aligned} \tag{33}$$

Using (26) and (27) we also have

$$\begin{aligned} ||\mathbf{b}||_\infty &\leq ||G_2^{-1}||_\infty ||G_1||_\infty ||S^{-1}||_\infty \\ &\leq \frac{2\tilde{C}_1 E(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}. \end{aligned} \tag{34}$$

Assuming $v_k = 1$, we get

$$a_k = (S^{-1}\mathbf{v})_k = \sum_j (S^{-1})_{k,j}.$$

Since $S^{-1}S = I$

$$\sum_j (S^{-1})_{k,j} (S)_{j,k} = 1.$$

We also know

$$\sum_j |(S^{-1})_{k,j} (S)_{j,k}| \leq \sum_j |(S^{-1})_{k,j}| \sum_j |(S)_{j,k}|,$$

which leads us to

$$\sum_j |(S^{-1})_{k,j}| \geq \frac{1}{\sum_j |(S)_{j,k}|}.$$

Also, we can derive

$$\forall k \qquad \sum_j (S)_{j,k} \le ||S||_1 \le ||G_0||_1 + ||G_1||_1^2 ||G_2^{-1}||_1$$

where $||A||_1 = \max_j \sum_i |a_{i,j}|$.

We can observe that

$$||G_0||_1 = \max_m \sum_k |g_m(t_k - t_m)| \le \alpha_0 + \tilde{C}_0 \max_m \sum_{k\neq m} \frac{1}{1+(t_k-t_m)^2}$$
$$\le \alpha_0 + \tilde{C}_0 \max_m \sum_{k\neq m} \frac{1}{1+((k-m)\nu)^2} \le \alpha_0 + 2\tilde{C}_0 E(\nu).$$

Similarly,

$$||\alpha_2 I - G_2||_1 = \max_m \Big[\sum_{k\neq m} |g_m^{(2)}(t_k - t_m)| + |g_m^{(2)}(0) - \alpha_2| \Big]$$
$$\le \tilde{C}_2 \max_m \sum_{k\neq m} \frac{1}{1+(t_k-t_m)^2} + \Delta_2$$
$$\le \tilde{C}_2 \max_m \sum_{k\neq m} \frac{1}{1+((k-m)\nu)^2} + \Delta_2 \le 2\tilde{C}_2 E(\nu) + \Delta_2.$$

Therefore,

$$||G_2^{-1}||_1 \le \frac{1}{|\alpha_2| - ||\alpha_2 I - G_2||_1} \le \frac{1}{\gamma_2 - 2\tilde{C}_2 E(\nu)}.$$

And since $g_m^{(1)}(0) \approx 0 \ \forall\ m$

$$||G_1||_1 \le 2\tilde{C}_1 E(\nu). \tag{35}$$

Which leads us to

$$\sum_j |(S)_{j,k}| \le \alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu)^2)}{\gamma_2 - 2\tilde{C}_2 E(\nu)}$$

So finally

$$a_k \ge \frac{1}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_2 - 2\tilde{C}_2 E(\nu)}}. \tag{36}$$

Hence, a_k's lower bound is inversely proportional to α_0. Numerical experiments have shown that this bound is tight, meaning that the smallest a_k is the exact reciprocal of the amplitude of the strongest kernel in the observation signal. This result is significant since it indicates that the bound on the recovery error is not merely stating the time-invariant result for the kernel with the worst constants. The better kernels are also taken into account. It can be said that the recovery error is proportional to the degradation ratio $\frac{\alpha_0}{\gamma_0}$, which is the ratio between the amplitude of the best kernel to the amplitude of the worst kernel. □

Lemma 4. *Under the separation condition with $\varepsilon < \nu/2$, q as in Proposition 3 satisfies $|q(t) < 1|$ if $0 < |t - t_m| \le \varepsilon$ for some $t_m \in T$.*

Lemma 5. *Under the separation condition with $\varepsilon < \nu/2$, q as in Proposition 3 satisfies $|q(t) < 1|$ if $|t - t_m| > \varepsilon \quad \forall t_m \in T$.*

Proof of Lemma 4. Assume $t \in \mathbb{R}$ and $t_k \le t \le t_k + \varepsilon$ for some $t_k \in T$, and that $q(t_k) = 1$. (The proof is the same for $t_k - \varepsilon \le t \le t_k$ or $v_k = -1$). We also assume that T satisfies the separation condition with $\varepsilon < \nu/2$. Therefore, we have $|t - t_m| > \frac{\nu}{2}$ for $m \neq k$. Then, for $l = 0, 1, 2, 3$ we have

$$\sum_{m\neq k} |g_m^{(l)}(t - t_m)| \le \sum_{m\neq k} \frac{C_{l,m}}{1 + \big(t - t_m\big)^2} \le \tilde{C}_l \sum_{m\neq k} \frac{1}{1 + \big((k-m)\nu/2\big)^2}$$
$$\le 8\tilde{C}_l E(\nu) = \frac{4}{3}\tilde{C}_l \frac{\pi^2}{\nu^2}. \tag{37}$$

>

Using this estimate, as well as (33), (34) and (36), and the admissible kernels' properties we obtain

$$q^{(2)}(t) = \sum_m a_m g_m^{(2)}(t - t_m) + b_m g_m^{(3)}(t - t_m)$$
$$\le a_k g_k^{(2)}(t - t_k) + ||a||_\infty \sum_{m\neq k} |g_m^{(2)}(t - t_m)|$$
$$+ ||b||_\infty \sum_m |g_m^{(3)}(t - t_m)|$$
$$\le -\frac{\beta}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_2 - 2\tilde{C}_2 E(\nu)}} + \frac{8\tilde{C}_2 E(\nu)}{\gamma_0 - 4\tilde{C}_0 E(\nu)}$$
$$+ \frac{16\tilde{C}_3(2E(\nu)+1)\tilde{C}_1 E(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}. \tag{38}$$

For sufficiently large ν that depends on the parameters of $g_m(t)$ we can approximate

$$q^{(2)}(t) < -\frac{\beta}{\alpha_0}. \tag{39}$$

By the Taylor remainder theorem [51], for any $t_k < t < t_k + \varepsilon$ there exists $t_k < \xi \le t$ such that

$$q(t) = q(t_k) + q^{(1)}(t_k)(t - t_k) + \frac{1}{2} q^{(2)}(\xi)(t - t_k)^2. \tag{40}$$

Since by construction $q^{(1)}(t_k) = 0$.

For sufficiently large ν

$$q(t) \le 1 - \frac{\beta}{2\alpha_0}(t - t_k)^2. \tag{41}$$

So we have shown that $q(t) < 1$.

To show that $q(t) > -1$

$$q(t) = \sum_m a_m g_m(t - t_m) + b_m g_m^{(1)}(t - t_m)$$
$$\ge a_k g_k(t - t_k) - ||a||_\infty \sum_{m\neq k} |g_m(t - t_m)|$$
$$- ||b||_\infty \sum_m |g_m^{(1)}(t - t_m)|$$
$$\ge \frac{g_k(\varepsilon)}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_2 - 2\tilde{C}_2 E(\nu)}} - \frac{8\tilde{C}_2 E(\nu)}{\gamma_0 - 4\tilde{C}_0 E(\nu)}$$
$$- \frac{16\tilde{C}_3(2E(\nu)+1)\tilde{C}_1 E(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}.$$

Hence, for sufficiently large ν and γ_0 we've shown that

$$q(t) > -1, \qquad for \quad t_k < t < t_k + \varepsilon, \tag{42}$$

and

$$|q(t)| < 1 \qquad |t - t_k| < \varepsilon, \quad t_k \in T. \tag{43}$$

□

Proof of Lemma 5. Assume $t \in \mathbb{R}$ and $|t - t_m| > \varepsilon$ for all $t_m \in T$, since $\varepsilon < \nu/2$, we have $|t - t_m| > \nu/2$. Then, from (16), the admissible kernel's properties, (33) and (34), we can write

$$|q(t)| \le \| a \|_\infty \sum_m |g_m(t - t_m)| + \| b \|_\infty \sum_m |g_m^{(1)}(t - t_m)|. \tag{44}$$

Let us denote

$$\hat{m} = \arg\min_m |g_m(0)|. \tag{45}$$

By assumption

$$\frac{g_{\hat{m}}(t - t_{\hat{m}})}{\gamma_0} < 1.$$

We recall,

$\gamma_0 = g_{\hat{m}}(0)$,

Moreover, since $|t - t_m| > \varepsilon$

$$0 < \frac{g_{\hat{m}}(t - t_{\hat{m}})}{\gamma_0} < \frac{g_{\hat{m}}(\varepsilon)}{\gamma_0}.$$

By the Taylor remainder theorem and the properties of $g_m(t)$, we know

$$0 < g_m(\varepsilon) \le g_m(0) - \frac{\beta\varepsilon^2}{2}.$$

Therefore,

$$\begin{aligned}|q(t)| &\le ||a||_\infty \Big(g_{\hat{m}}(t - t_{\hat{m}}) + \sum_{m \ne \hat{m}} |g_m(t - t_m)| \Big) \\ &\quad + ||b||_\infty \sum_m |g_m^{(1)}(t - t_m)| \\ &\le \frac{g_{\hat{m}}(t - t_{\hat{m}}) + 8\tilde{C}_0 E(\nu)}{\gamma_0 - 4\tilde{C}_0 E(\nu)} + \frac{16\tilde{C}_1^2 E^2(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}.\end{aligned} \tag{46}$$

Finally, we can conclude that for sufficiently large ν,

$$|q(t)| \le 1 - \frac{\beta\varepsilon^2}{2\gamma_0}. \tag{47}$$

□

Now, we can complete the proof of Theorem 1. Assume $\hat{x}$ is the solution of the optimization problem in (14). $\hat{x}$ obeys $||\hat{x}||_{\ell_1} \le ||x||_{\ell_1}$.

Denote the error $h[k] \triangleq \hat{x}[k] - x[k]$.

Now separate h into $h = h_K + h_{K^c}$, where h_K's support is in the true support $K \triangleq \{k_m\}$. If $h_K = 0$, then $h = 0$, since $h_K = 0$ and $h \ne 0$ would imply that $h_{K^c} \ne 0$ and $||\hat{x}||_{\ell_1} \ge ||x||_{\ell_1}$.

Under the separation condition, the set $T = \{t_m\}$ satisfies $t_i - t_j \ge \nu\sigma$ for $i \ne j$.

We've shown in Proposition 3 that there exists q of the form (16) such that

$$q(t_m) = q\Big(\frac{k_m}{N}\Big) = \mathrm{sgn}(h_K[k_m]) \quad \forall k_m \in T. \tag{48}$$

By assumption, we choose $v_m = \mathrm{sgn}(h_K(t_m))$.

In addition, q also satisfies $|q(t)| < 1$ for $t \notin T$.

We then define

$$q_\sigma(t) = q\Big(\frac{t}{\sigma}\Big) = \sum_m a_m g_{m,\sigma}\Big(t - \frac{k_m}{N}\Big) + b_m g_{m,\sigma}^{(1)}\Big(t - \frac{k_m}{N}\Big).$$

So that

$$q_\sigma(k_m) \triangleq q_\sigma\Big(\frac{k_m}{N}\Big) = \mathrm{sgn}(h_K[k_m]) \quad \forall k_m \in K,$$

and

$$|q_\sigma(k)| < 1 \quad \forall k \notin K.$$

Denote $g_{m,\sigma}^{(1)}[k] \triangleq g_{m,\sigma}^{(1)}\big(\frac{k}{N}\big)$. Consequently, we can obtain

$$\begin{aligned}&\Big|\sum_{k\in Z} q_\sigma[k]h[k]\Big|_1 \\ &= \Big|\sum_{k\in Z}\Big(\sum_{k_m\in K} a_m g_{m,\sigma}(k - k_m) + b_m g_{m,\sigma}^{(1)}(k - k_m)\Big)h[k]\Big|_1 \\ &\le ||a||_\infty \Big|\sum_{k\in Z}\Big(\sum_m g_{m,\sigma}[k - k_m]h[k]\Big)\Big|_1 \\ &\quad + ||b||_\infty \Big|\sum_{k\in Z}\Big(\sum_m g_{m,\sigma}^{(1)}[k - k_m]h[k]\Big)\Big|_1.\end{aligned} \tag{49}$$

We also have,

$$\begin{aligned}&\Big|\sum_m g_{m,\sigma}[k - k_m]h[k]\Big|_1 \\ &= \Big|\sum_m g_{m,\sigma}[k - k_m]\hat{x}[k] - \sum_m g_{m,\sigma}[k - k_m]x[k]\Big|_1 \\ &= \Big|y[k] - \sum_m g_{m,\sigma}[k - k_m]x[k] - \Big(y[k] - \sum_m g_{m,\sigma}[k - k_m]\hat{x}[k]\Big)\Big|_1 \\ &\le \Big|y[k] - \sum_m g_{m,\sigma}[k - k_m]x[k]\Big|_1 \\ &\quad + \Big|y[k] - \sum_m g_{m,\sigma}[k - k_m]\hat{x}[k]\Big|_1 \le 2\delta.\end{aligned} \tag{50}$$

Since,

$$\big|y[k] - x[k]\big|_1 = \Big|y[k] - \sum_m c_m g_{m,\sigma}[k - k_m]\Big|_1 \le \delta,$$

and also,

$$|y[k] - \hat{x}[k]|_1 = |y[k] - \sum_m \hat{c}_m g_{m,\sigma}[k - \hat{k}_m]|_1 \le \delta.$$

As mentioned above $\{g_m\}$ is a set of admissible kernels. Therefore,

$$|g_{m,\sigma}^{(1)}[k - k_m]| = \Big|g_m^{(1)}\Big(\frac{k - k_m}{N\sigma}\Big)\Big| \le \frac{C_{1,m}}{1 + \big(\frac{k-k_m}{N\sigma}\big)^2}. \tag{51}$$

Under the separation condition we have $|k_i - k_j| \ge N\nu\sigma \quad \forall k_i, k_j \in K$. Hence, for any k we have

$$\sum_{k_m\in K} \frac{1}{1 + \big(\frac{k-k_m}{N\sigma}\big)^2} < 2(1 + E(\nu)). \tag{52}$$

Since we know

$$\sum_{n=1}^{\infty} \frac{1}{1 + (n\nu)^2} \le E(\nu) \triangleq \frac{\pi^2}{6\nu^2}.$$

Then,

$$\begin{aligned}\sum_{k_m\in K}\Big|\sum_{k\in Z}\Big(g_{m,\sigma}^{(1)}[k - k_m]h[k]\Big)\Big|_1 &\le \tilde{C}_1 \sum_{k\in Z}|h[k]|_1 \sum_{k_m\in K}\frac{1}{1 + \big(\frac{k-k_m}{N\sigma}\big)^2} \\ &< 2\tilde{C}_1(1 + E(\nu))||h||_1.\end{aligned}$$

Hence,

$$\Big|\sum_{k\in Z} q_\sigma[k]h[k]\Big|_1 \le 2\delta||a||_\infty + 2\tilde{C}_1(1 + E(\nu))||b||_\infty||h||_1. \tag{53}$$

On the other hand,

$$\begin{aligned}\Big|\sum_{k\in\mathbb{Z}} q_\sigma[k]h[k]\Big|_1 &= \Big|\sum_{k\in\mathbb{Z}} q_\sigma[k](h_K[k] + h_{K^c}[k])\Big|_1 \\ &\ge \sum_{k\in\mathbb{Z}} |q_\sigma[k]h_K[k]|_1 - |q_\sigma[k]h_{K^c}[k]|_1 \\ &\ge \| h_K\|_1 - \max_{k\in\mathbb{Z}\setminus K}|q_\sigma[k]| \ \| h_{K^c}[k]\|_1.\end{aligned} \tag{54}$$

Combining (53) and (54) we get,

$$\begin{aligned}&\| h_K\|_1 - \max_{k\in\mathbb{Z}\setminus K}|q_\sigma[k]| \ \| h_{K^c}\|_1 \\ &\le 2\delta \ \| a\|_\infty + 2\tilde{C}_1(1 + E(\nu)) \ \| b\|_\infty \ \| h\|_1.\end{aligned} \tag{55}$$

We've shown in the proof of lemma 4 that for $|k - k_m| \le \varepsilon N\sigma$, for some $k_m \in K$

$$|q_\sigma[k]| = \Big|q_\sigma\Big(\frac{k}{N\sigma}\Big)\Big| \le 1 - \frac{\beta}{2\alpha_0(N\sigma)^2}.$$

And by (47) for $|k-k_m| > \varepsilon N\sigma$ for all $k_m \in K$

$$|q_\sigma[k]| = \left|q\left(\frac{k}{N\sigma}\right)\right| \le 1 - \frac{\beta\varepsilon^2}{2\gamma_0}.$$

So we can conclude

$$\max_{k\in Z}|q_\sigma[k]| \le 1 - \frac{\beta}{2\rho} \tag{56}$$

where $\rho \triangleq \max\left\{\frac{\gamma_0}{\varepsilon^2}, (N\sigma)^2\alpha_0\right\}$.

Substituting (56) into (55) we get,

$$||h_K||_1 - \left(1-\frac{\beta}{2\rho}\right)||h_{K^c}||_1 \le 2\delta||a||_\infty + 2\tilde{C}_1(1+E(\nu))||b||_\infty||h||_1. \tag{57}$$

Moreover, we know that

$$\begin{aligned}||x||_1 &\ge ||\hat{x}||_1 = ||x+h||_1 = ||x+h_K||_1 + ||h_K^C||_1\\ &\ge ||x||_1 - ||h_K||_1 + ||h_{K^c}||_1.\end{aligned}$$

Which leads us to

$$||h_K||_1 \ge ||h_{K^c}||_1.$$

Combining this with (57) leads us to

$$\begin{aligned}||h||_1 &= ||h_K||_1 + ||h_{K^c}||_1 \le 2||h_{K^c}||_1\\ &\le \frac{4\rho}{\beta}(\delta||a||_\infty + 2\tilde{C}_1(1+E(\nu))||b||_\infty||h||_1).\end{aligned} \tag{58}$$

So we have,

$$||h||_1 \le \frac{4\rho||a||_\infty}{\beta - 4\rho\tilde{C}_1(1+E(\nu))||b||_\infty}\delta. \tag{59}$$

Using (33) and (34)

$$||h||_1 \le \frac{36\rho\gamma_2}{9\beta\gamma_0\gamma_2 - D_1\nu^{-1} - D_2\nu^{-2}}\delta, \tag{60}$$

$$D_1 = 3\pi^2(\beta\tilde{C}_2 + 2\beta\tilde{C}_0 + 4\tilde{C}_1\rho), \quad D_2 = \pi^4(2\rho\tilde{C}_1 - \beta\tilde{C}_2\tilde{C}_0).$$

□

Appendix B

Proof of Theorem 2. To prove Theorem 2 we use the following two Lemmas.

Lemma 6. *Assume a set of delays $T\triangleq\{t_m\}$ that satisfies the separation condition*

$$\min_{k_i,k_j\in K, i\ne j}\left|k_i - k_j\right| \ge N\nu\sigma.$$

Let $\{g_m\}$ be a set of admissible kernels as defined in Definition 2.1 or asymmetric approximately admissible kernels as described in Section 2.2. Then, for any $t_m \in T$ there exist coefficients $\{a_k\}$ and $\{b_k\}$ such that the function

$$q_m(t) = \sum_k a_k g_m\left(\frac{t-t_k}{\sigma}\right) + b_k g_m^{(1)}\left(\frac{t-t_k}{\sigma}\right) \tag{61}$$

obeys

$$\begin{aligned}&q_m(t_m) = 1,\\ &q_m(t_j) = 0 \quad \forall t_j \in T\backslash\{t_m\},\\ &q_m^{(1)}(t_j) = 0 \quad \forall t_j,\end{aligned}$$

$$|1 - q_m(t)| < \frac{C_{2,m}(t-t_m)^2}{g_m(0)\sigma^2} \quad \forall t \ne t_m, \tag{62}$$

$$|q_m(t)| < \frac{C_{2,m}(t-t_j)^2}{g_m(0)\sigma^2}, \quad \forall t_j \in T\backslash\{t_m\}, \quad |t-t_j| \le \varepsilon\sigma, \tag{63}$$

$$|q_m(t)| < 1 - \xi_m\varepsilon^2 \quad |t-t_j| > \varepsilon\sigma, \ \forall t_j \in T, \tag{64}$$

$$\xi_m = \frac{\beta}{4g_m(0)}.$$

These results also for all $0 < \varepsilon' < \varepsilon$.

Remark *Notice that here we have a set of different admissible kernels, each function $q_m(t)$ is based on a different kernel $\{g_m(t)\}$ and it decouples the estimation error at one location t_m from the amplitude of the rest of the support. It is designed to obey $< q_m, x > \triangleq \int q_m(t)x(t)dt = c_m$.*

Lemma 7. *Assume K that satisfies the separation condition of Definition 2.2 for $N > 0$, then*

$$\sum_{\hat{k}_m\in\hat{K}}|\hat{c_m}|\min\left\{\varepsilon^2, \frac{d(\hat{k}_m, K)}{(N\sigma)^2}\right\} \le \frac{2D_3\alpha_0}{\beta}\delta,$$

where

$$d(k,K) = \min_{k_n\in K}(k_n - k)^2.$$

Proof of Lemma 6. We impose

$$\begin{aligned}&q_m(t_m) = 1,\\ &q_m(t_j) = 0 \quad \forall t_j \in T\backslash\{t_m\},\\ &q_m^{(1)}(t_j) = 0 \quad \forall t_j.\end{aligned}$$

$$\sum_k a_k g_m\left(\frac{t_m - t_k}{\sigma}\right) + b_k g_m^{(1)}\left(\frac{t_m - t_k}{\sigma}\right) = 1$$

and

$$\sum_k a_m g_m^{(1)}\left(\frac{t_j - t_k}{\sigma}\right) + b_k g_m^{(2)}\left(\frac{t_j - t_k}{\sigma}\right) = 0, \quad t_j \in T\backslash\{t_m\}.$$

In matrix form,

$$\begin{bmatrix}D_0 & D_1\\ D_1 & D_2\end{bmatrix}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{e_{t_m}}\\ \mathbf{0}\end{bmatrix},$$

where $\mathbf{e}_{t_m}$ is a vector with one nonzero entry at the location corresponding to t_m, $\mathbf{a}\triangleq\{a_m\}, \mathbf{b}\triangleq\{b_m\}$ and $(D_l)_{j,k} \triangleq g_m^{(l)}\left(\frac{t_j - t_k}{\sigma}\right), l = 0,1,2$. As we mentioned in Proposition 3, we know that the matrix D is invertible if both D_2 and the Schur complement of D_2:

$S = D_0 - D_1D_2^{-1}D_1$ are invertible [50]. We also know that a matrix A is invertible if there exists $\alpha \ne 0$ such that $||\alpha I - A||_\infty < |\alpha|$, where $||A||_\infty = \max_i\sum_j|a_{i,j}|$. In this case we have

$$||A^{-1}||_\infty \le \frac{1}{|\alpha| - ||\alpha I - A||_\infty}.$$

Using the properties of an admissible kernel and the separation condition, we can write

$$\begin{aligned}||g_m^{(2)}(0)I - D_2||_\infty &= \max_k\sum_{m\ne k}|g_m^{(2)}\left(\frac{t_k - t_m}{\sigma}\right)|\\ &\le \frac{C_{2,m}}{\sigma^2}\max_k\sum_{m\ne k}\frac{1}{1+\left(\frac{t_k-t_m}{\sigma}\right)^2}\\ &\le \frac{C_{2,m}}{\sigma^2}\max_k\sum_{m\ne k}\frac{1}{1+((k-m)\nu)^2}.\end{aligned}$$

Recall that

$$\sum_{n=1}^{\infty}\frac{1}{1+(n\nu)^2} \le E(\nu) \triangleq \frac{\pi^2}{6\nu^2}.$$

Therefore,

$$||g_m^{(2)}(0)I - D_2||_\infty \le \frac{2C_{2,m}}{\sigma^2}E(\nu). \tag{65}$$

Meaning that D_2 is invertible if $\frac{2C_{2,m}}{\sigma^2}E(\nu) < |g_m^{(2)}(0)|$.

This yields the condition

$$\nu^2 \ge \frac{2C_{2,m}\pi^2}{3|g_m^{(2)}(0)|\sigma^2}. \tag{66}$$

This implies that as m increases and $g_m(t)$ loses more energy, in order to achieve the correct recovery by ℓ_1 optimization, the minimum distance between two adjacent spikes should be larger.

$$\begin{aligned}||g_m(0)I - S||_\infty &= ||g_m(0)I - D_0 + D_1D_2^{-1}D_1||_\infty \\ &\le ||g_m(0)I - D_0||_\infty + ||D_1||_\infty^2||D_2^{-1}||_\infty.\end{aligned} \tag{67}$$

In a similar manner to (65)

$$||g_m(0)I - D_0||_\infty \le 2C_{0,m}E(\nu). \tag{68}$$

And since $g_m^{(1)}(0) \approx 0 \ \forall\ m$

$$||D_1||_\infty \le \frac{2C_{1,m}}{\sigma}E(\nu). \tag{69}$$

Using (19) and (65) we get

$$||D_2^{-1}||_\infty \le \frac{1}{|g_m^{(2)}(0)| - ||g_m^{(2)}(0)I - D_2||_\infty} \le \frac{1}{|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2}E(\nu)}. \tag{70}$$

So, we have

$$\begin{aligned}||g_m(0)I - S||_\infty &\le 2C_{0,m}E(\nu) + \frac{(\frac{2C_{1,m}}{\sigma}E(\nu))^2}{|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2}E(\nu)} \\ &= 2C_{0,m}E(\nu)\left[1 + \frac{\frac{2C_{1,m}^2}{\sigma^2}E(\nu)}{C_{0,m}(|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2}E(\nu))}\right] \\ &\le 4C_{0,m}E(\nu).\end{aligned} \tag{71}$$

The last inequality holds for

$$\nu^2 \ge \frac{\pi^2(C_{1,m}^2 + C_{0,m}C_{2,m})}{3\sigma^2C_{0,m}|g_m^{(2)}(0)|}. \tag{72}$$

If in addition

$$\nu^2 > \frac{2\pi^2C_{0,m}}{3\sigma^2g_m(0)}. \tag{73}$$

Here again it can be observed that small $g_m(0)$ values require a larger separation constant ν. Then finally we have

$$||g_m(0)I - S||_\infty \le 4C_{0,m}E(\nu) < g_m(0) \tag{74}$$

and S is invertible.

Furthermore, $\mathbf{a}$ and $\mathbf{b}$ are given by

$$\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}D_0 & D_1\\ D_1 & D_2\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{e_{t_m}}\\ \mathbf{0}\end{bmatrix},$$

$$\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}I\\ -D_2^{-1}D_1S^{-1}\mathbf{e_{t_m}}\end{bmatrix}. \tag{75}$$

Using (19) and (74) we have

$$\begin{aligned}||\mathbf{a}||_\infty \le ||S^{-1}||_\infty &\le \frac{1}{|g_m(0)| - ||g_m(0)I - S||_\infty} \\ &\le \frac{1}{g_m(0) - 4C_{0,m}E(\nu)}.\end{aligned} \tag{76}$$

Using (69) and (70) we also have

$$\begin{aligned}||\mathbf{b}||_\infty &\le ||D_2^{-1}||_\infty||D_1||_\infty||S^{-1}||_\infty \\ &\le \frac{\frac{2C_{1,m}}{\sigma}E(\nu)}{(|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2}E(\nu))(g_m(0) - 4C_{0,m}E(\nu))}.\end{aligned} \tag{77}$$

And we can also derive

$$\begin{aligned}a_k = (S^{-1}\mathbf{e_{t_m}})_k &= \frac{1}{g_m(0)}\left(1 - \left(S^{-1}\left(S - g_m(0)I\right)\mathbf{e_{t_m}}\right)_k\right) \\ &\ge \frac{1}{g_m(0)}\left(1 - ||S^{-1}||_\infty||S - g_m(0)I||_\infty\right).\end{aligned} \tag{78}$$

$$a_k \ge \frac{1}{g_m(0)}\left(1 - \frac{4C_{0,m}E(\nu)}{g_m(0) - 4C_{0,m}E(\nu)}\right). \tag{79}$$

Fix $t_k \in T$ and $|t - t_k| \le \varepsilon\sigma$. Under the separation condition we have $|t - t_j| \ge \frac{\nu}{2}$ for $t_j \in T\backslash\{t_k\}$. Therefore, we have for $l = 0, 1, 2, 3$:

$$\begin{aligned}\sum_{j\ne k}|g_m^{(l)}\left(\frac{t - t_j}{\sigma}\right)| &\le \frac{C_{l,m}}{\sigma^l}\sum_{j\ne k}\frac{1}{1 + \left(\frac{t - t_k}{\sigma}\right)^2} \\ &\le \frac{C_{l,m}}{\sigma^l}\sum_{j\ne k}\frac{1}{1 + \left((k - j)\nu/2\right)^2} \le \frac{4}{3}C_{l,m}\frac{\pi^2}{\sigma^l\nu^2}.\end{aligned} \tag{80}$$

Using this estimate, as well as (61) and the admissible kernels' properties we obtain

$$\begin{aligned}||q_m^{(2)}(t)||_\infty &\le ||a||_\infty\sum_j|g_m^{(2)}(t - t_j)| + ||b||_\infty\sum_j|g_m^{(3)}(t - t_j)| \\ &\le ||a||_\infty\left(\frac{4}{3}C_{2,m}\frac{\pi^2}{\nu^2\sigma^2} + |g_m^{(2)}(t - t_k)|\right) \\ &+ ||b||_\infty\left(\frac{4}{3}C_{3,m}\frac{\pi^2}{\nu^2\sigma^3} + |g_m^{(2)}(t - t_k)|\right) \\ &\le \frac{1}{\sigma^2}\left(1 + \frac{4\pi^2}{3\nu^2}\right) \\ &\times\left(\frac{3C_{2,m}\nu^2\left(3|g_m^{(2)}(0)|\nu^2 - \pi^2\frac{C_{2,m}}{\sigma^2}\right) + C_{1,m}C_{3,m}\frac{\pi^2}{\sigma}}{\left(3|g_m^{(2)}(0)|\nu^2 - \pi^2\frac{C_{2,m}}{\sigma^2}\right)\left(3|g_m(0)|\nu^2 - 2\pi^2C_{0,m}\right)}\right).\end{aligned} \tag{81}$$

For sufficiently large ν that depends on the parameters of $g_m(t)$ we can approximate

$$\left|q_m^{(2)}(t)\right| < \frac{2C_{2,m}}{g_m(0)\sigma^2} \qquad |t - t_k| \le \varepsilon\sigma, \qquad t_k \in T. \tag{82}$$

By the Taylor Remainder theorem, for any $t_m < t < t_m + \varepsilon$ there exists $t_m < \xi \le t$ such that

$$q_m(t) = q_m(t_m) + q^{(1)}(t_m)(t - t_m) + \frac{1}{2}q^{(2)}(\xi)(t - t_m)^2. \tag{83}$$

Since by construction $q_m(t_m) = 1$ and $q_m^{(1)}(t_k) = 0$ we have

$$|1 - q_m(t)| \le \frac{C_{2,m}}{g_m(0)\sigma^2}(t - t_m)^2. \tag{84}$$

In the same manner since $q_m(t_k) = 0$ for all $t_k \in T\backslash\{t_m\}$ there exists $t_k < \xi \le t$ for any $t_k < t \le t_k + \varepsilon$ such that

$$\begin{aligned}q_m(t) &= q_m(t_k) + q^{(1)}(t_k)(t - t_k) + \frac{1}{2}q^{(2)}(\xi)(t - t_k)^2 \\ &= \frac{1}{2}q^{(2)}(\xi)(t - t_k)^2.\end{aligned} \tag{85}$$

Leading to

$$|q_m(t)| \leq \frac{C_{2,m}}{g_m(0)\sigma^2}(t-t_k)^2. \tag{86}$$

Similar arguments hold for $t_m - \varepsilon < t < t_m$ and $t_k - \varepsilon < t < t_k$. □

Proof of Lemma 7. Set $q[k] = q\left(\frac{k}{N}\right)$, $k \in \mathbb{Z}$, where $q(t)$ is given in Proposition 3 and $v_m = \text{sgn}(c_m)$. By the Taylor Remainder theorem, for any $0 < k - k_m \leq \varepsilon N\sigma$ there exists $\frac{k_m}{N} < \eta < \frac{k}{N} + \varepsilon$ such that

$$\begin{aligned} q[k] &= q\left(\frac{k}{N}\right) \\ &= q\left(\frac{k_m}{N}\right) + q^{(1)}\left(\frac{k_m}{N}\right)\left(\frac{k-k_m}{N}\right) + \frac{1}{2}q^{(2)}(\eta)\left(\frac{k-k_m}{N}\right)^2. \end{aligned}$$

Using (39)

$$|q[k]| \leq 1 - \frac{\beta}{\alpha_0(N\sigma)^2}(k-k_m)^2.$$

We observe that

$$<q, \hat{x}> \leq \sum_m |\hat{c}_m||q(\hat{k}_m)| \leq \sum_m |\hat{c}_m|\left(1 - \frac{\beta}{\alpha_0}\min\left\{\varepsilon^2, \frac{d\left(\hat{k}_m, K\right)}{(N\sigma)^2}\right\}\right)$$

where

$$d(k, K) = \min_{k_n \in K}(k_n - k)^2.$$

q is designed to satisfy $< q, x >= \sum_m |c_m| = ||x||_{\ell_1} \geq ||\hat{x}||_{\ell_1}$.

Moreover, we can apply (53) and get

$$\begin{aligned} <q, \hat{x} - x> &= \left|\sum_{k\in Z} q[k]h[k]\right| \\ &\leq 2\delta||a||_\infty + 2\tilde{C}_1(1 + E(\nu))||b||_\infty||h||_1. \end{aligned} \tag{87}$$

Therefore,

$$<q, \hat{x} - x> \leq 2\delta D_3 \tag{88}$$

where

$$D_3 = \frac{3\nu^2(3\gamma_2\nu^2 - \pi^2\tilde{C}_2) + \frac{12\pi^2\tilde{C}_1^2}{\beta\gamma_0}(1 + \frac{\pi^2}{6\nu^2})\rho}{(3\gamma_2\nu^2 - \pi^2\tilde{C}_2)(3\gamma_0\nu^2 - 2\pi^2\tilde{C}_0)}.$$

Then we have

$$\begin{aligned} <q, \hat{x}> &= < q, \hat{x} - x > + < q, x > \\ &\geq ||x||_{\ell_1} - 2\delta D_3 \\ &\geq ||\hat{x}||_{\ell_1} - 2\delta D_3 \\ &\geq \sum_m |\hat{c}_m| - 2\delta D_3. \end{aligned}$$

Which leads us to

$$\sum_{\hat{k}_m \in \hat{K}} |\hat{c_m}| \min\left\{\varepsilon^2, \frac{d(\hat{k}_m, K)}{(N\sigma)^2}\right\} \leq \frac{2D_3\alpha_0}{\beta}\delta.$$

□

The first result of Theorem 2 is a direct corollary of Lemma 7. It ensures that any false spikes in the recovered reflectivity, which is far from the true support, have small energy.

Now, we shall proceed to prove the second result of Theorem 2. Let us denote $q_m[k] \triangleq q_m(\frac{k}{N})$, $k \in \mathbb{Z}$, where $q_m(t)$ is given in Lemma 6. Using $q_m(t)$ we can decouple the support-detection error at k_m, for one spike, from the rest of the support.

We can apply Theorem 1 and get

$$<q_m, \hat{x} - x> \leq |\hat{x} - x|_1 \leq \frac{4\rho}{\beta\gamma_0}\delta. \tag{89}$$

Where we have used that the absolute value of $q_m[k]$ is bounded by one.

Recall that we assumed $\varepsilon \geq \tilde{\varepsilon} = \sqrt{\frac{\alpha_0}{\tilde{C}_2 + \beta/4}}$,

Let us denote

$$\begin{aligned} \hat{K}_{far} &\triangleq \{n : |\hat{k}_n - k_m| > \tilde{\varepsilon}N\}, \\ \hat{K}_{near} &\triangleq \{n : |\hat{k}_n - k_m| \leq \tilde{\varepsilon}N\}, \\ \xi_m &= \frac{\beta}{4g_m(0)}. \end{aligned}$$

In other words, $\hat{K}_{far}$ is the recovered support located far from the true support, whereas $\hat{K}_{near}$ is the recovered support located close to the true support.

We then derive

$$\begin{aligned} &\left|\sum_{\{n:\hat{k}_n\in\hat{K}_{far}\}} \hat{c}_n q_m[\hat{k_n}] - \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \hat{c}_n(1 - q_m[\hat{k_n}])\right| \\ &\leq |\sum_{\{n:\hat{k}_n\in\hat{K}_{far}\}} |\hat{c}_n||q_m[\hat{k_n}]|| + |\sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} |\hat{c}_n||1 - q_m[\hat{k_n}]| \\ &\leq \sum_{\hat{k}_n\in\hat{K}} |\hat{c}_n| \min\left\{1 - \xi_m\tilde{\varepsilon}^2, \frac{C_{2,m}d(\hat{k}_n, k)}{g_m(0)(N\sigma)^2}\right\} \\ &\leq \sum_{\hat{k}_n\in\hat{K}} |\hat{c}_n| \min\left\{1, \frac{4\xi_m C_{2,m}d(\hat{k}_n, k)}{\beta(N\sigma)^2}\right\} \\ &\leq \max\left\{\frac{1}{\varepsilon^2}, \frac{4\xi_m C_{2,m}}{(N\sigma)^2\beta}\right\} \sum_{\hat{k}_n\in\hat{K}} |\hat{c}_n| \min\left\{\varepsilon^2, d(\hat{k}_n, k)\right\} \\ &\leq \max\left\{\frac{1}{\varepsilon^2}, \frac{4\xi_m C_{2,m}}{\beta(N\sigma)^2}\right\} \frac{2D_3\alpha_0}{\beta}\delta \\ &= \frac{2D_3\alpha_0}{\beta}\delta \max\left\{\frac{1}{\varepsilon^2}, \frac{C_{2,m}}{(N\sigma)^2 g_m(0)}\right\}. \end{aligned}$$

$q_m[k]$ is designed to satisfy $< q_m, x >= c_m$.

So we can bound the difference between each spike amplitude to the energy of the estimated spikes clustered tightly around it by

$$\begin{aligned} \left|c_m\right| - \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \left|\hat{c}_n\right| &\leq \left|c_m - \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \hat{c}_n\right| \\ &= \Big|< q_m, x > - \Big[< q_m, \hat{x} > - \sum_{\{n:\hat{k}_n\in\hat{K}_{far}\}} \hat{c}_n q_m[\hat{k_n}] \\ &\quad + \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \hat{c}_n(1 - q_m[\hat{k_n}])\Big]\Big| \\ &= \Big|< q_m, x - \hat{x} > + \sum_{\{n:\hat{k}_n\in\hat{K}_{far}\}} \hat{c}_n q_m[\hat{k_n}] \\ &\quad - \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \hat{c}_n(1 - q_m[\hat{k_n}])\Big]\Big| \\ &\leq \frac{2\delta}{\beta}\left(\frac{2\rho}{\gamma_0} + D_3\alpha_0 \max\left\{\frac{1}{\varepsilon^2}, \frac{C_{2,m}}{(N\sigma)^2 g_m(0)}\right\}\right). \end{aligned}$$

Denote

$$D_4 = \frac{2\delta}{\beta}\left(\frac{2\rho}{\gamma_0} + D_3\alpha_0 \max\left\{\frac{1}{\varepsilon^2}, \frac{C_{2,m}}{(N\sigma)^2 g_m(0)}\right\}\right),$$

Consequently, if $|c_m| \geq D_4$, there exists at least one $\hat{k}_m \in \hat{K}$ so that $|\hat{k}_m - k_m| \leq \tilde{\varepsilon}N$ with $\left|c_m\right| - D_4 \leq \sum_{\{n:\hat{k}_n\in\hat{K}_{near}\}} \left|\hat{c}_n\right|$.

Therefore, using Lemma 7 we get

$$(\hat{k}_m - k_m)^2 \leq \frac{2D_3\alpha_0}{\beta \sum_{\{n:\hat{k}_n \in \hat{K}_{near}\}} \left|\hat{c}_n\right|}\delta \leq \frac{2D_3(N\sigma)^2\alpha_0}{\beta\left(\left|c_m\right| - D_4\right)}\delta. \quad (90)$$

This concludes the proof.

Hence, this bound proves that solving the convex optimization problem in (14) locates the support of the reflectivity with high precision, as long as the spikes are separated by ν and the noise level is small with respect to the spikes amplitude. Moreover, the bound on the support detection error depends only on the amplitude of the corresponding spike c_m, on Q, and on the signal length. It does not depend on the amplitudes of the reflectivity in other locations. □

References

[1] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159.
[2] M. Elad, Sparse and Redundant Representations, Springer, 2010.
[3] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (4) (2006) 1289–1306.
[4] R.G. Baraniuk, P. Steeghs, Compressive radar imaging, in: IEEE Radar Conference, 2007.
[5] T. Bendory, S.D. A. Bar-Zion, A. Feuer, Stable support recovery of stream of pulses with application to ultrasound imaging, IEEE Trans. Sig. Process. 64 (14) (2016) 3750–3759.
[6] H.L. Taylor, S.C. Banks, J.F. McCoy, Deconvolution with the ℓ_1 norm, Geophysics 44 (1979) 39–52.
[7] T. Nguyen, J. Castagna, High resolution seismic reflectivity inversion, J. Seismic Explor. 19 (4) (2010) 303–320.
[8] R. Zhang, J. Castagna, Seismic sparse-layer reflectivity inversion using basis pursuit decomposition, Geophysics 76 (6) (2011) 147–158.
[9] A. Berkhout, The seismic method in the search for oil and gas: current techniques and future development, Proc. IEEE 74 (8) (1986) 1133–1159.
[10] T. Ulrych, Application of homomorphic deconvolution to seismology, Geophysics 36 (6) (1971) 650–660.
[11] R. Wiggins, Minimum entropy deconvolution, Geoexploration 16 (1978) 21–35.
[12] P.V. Riel, A.J. Berkhout, Resolution in seismic trace inversion by parameter estimation, Geophysics 50 (1985) 1440–1455.
[13] A. Gholami, M.D. Sacchi, A fast and automatic sparse deconvolution in the presence of outliers, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 4105–4116.
[14] D. Pereg, D. Ben-Zvi, Blind deconvolution via maximum kurtosis adaptive filtering, in: Proceedings of 2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE), 2016, pp. 1–5.
[15] M.Q. Pham, L. Duval, C. Chaux, J.-C. Pesquet, A primal-dual proximal algorithm for sparse template-based adaptive filtering: application to seismic multiple removal, IEEE Trans. Sig. Process. 62 (16) (2014) 4256–4269.
[16] A. Repetti, M.Q. Pham, L. Duva, E. Chouzenoux, J.-C. Pesquet., Euclid in taxicab: sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization, IEEE Sig. Process. Lett. 22 (5) (2015) 539–543.
[17] V. Duval, G. Peyré, Exact support recovery for sparse spikes deconvolution, Found. Comput. Math. (2015) 1–41.
[18] D.L. Donoho, Super-resolution via sparsity constraints, SIAM J. Math. Anal. 23 (5) (1992) 1309–1331.
[19] C. Dossal, S. Mallat, Sparse spike deconvolution with minimum scale, in: SPARS, 2005, pp. 123–126.
[20] C. Fernandez-Granda, Support detection in super-resolution, in: Proceedings of SampTA, 2013, pp. 145–148.
[21] E.J. Candès, C. Fernandez-Granda, Super-resolution from noisy data, J. Fourier Anal. Appl. 19 (6) (2013a) 1229–1254.
[22] E.J. Candès, C. Fernandez-Granda, Towards a mathematical theory of superresolution, Commun. Pure Appl. Math. (2013b).
[23] T. Bendory, S. Dekel, A. Feuer, Robust recovery of stream of pulses using convex optimization, J. Math. Anal. Appl. 442 (2) (2016) 511–536.
[24] R.J. Tibshirani, The LASSO problem and uniqueness, Electr. J. Stat. 7 (2013) (2013) 1456–1490.
[25] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, Ann. Stat. 32 (2) (2004) 407–499.
[26] J. Idier, Y. Goussard, Multichannel seismic deconvolution, IEEE Trans. Geosci. Remote Sensing 31 (5) (1993) 961–979.
[27] J. Mendel, J. Kormylo, F. Aminzadeh, J. Lee, F. Habibi-Ashrafi, A novel approach to seismic signal processing and modeling, Geophysics 46 (1981) 1398–1414.
[28] J. Kormylo, J. Mendel, Maximum likelihood detection and estimation of Bernoulli–Gaussian processes, IEEE Trans. Inf. Theory 28 (2) (1982) 482–488.
[29] K. Kaaresen, T. Taxt, Multichannel blind deconvolution of seismic signals, Geophysics 63 (6) (1998) 2093–2107.
[30] A. Heimer, I. Cohen, A. Vassiliou, Dynamic programming for multichannel blind seismic deconvolution, in: Proc. Society of Exploration Geophysicist International Conference, Exposition and 77th Annual Meeting, 2007, pp. 1845–1849. San Antonio
[31] A. Heimer, I. Cohen, Multichannel seismic deconvolution using markov-Bernoulli random field modeling, IEEE Trans. Geosci. Remote Sens. 47 (7) (2009) 2047–2058.
[32] A. Heimer, I. Cohen, Multichannel blind seismic deconvolution using dynamic programmming, Sig. Process. 88 (4) (2008) 1839–1851.
[33] I. Ram, I. Cohen, S. Raz, Multichannel deconvolution of seismic signals using statistical MCMC methods, IEEE Trans. Sig. Process. 58 (5) (2010) 2757–2769.
[34] M. Mirel, I. Cohen, Multichannel semi-blind deconvolution (msbd) of seismic signals, Sig. Process. (2017). http://dx.doi.org/10.1016/j.sigpro.2017.01.026.
[35] D. Forney, The viterbi algorithm, in: Proc. IEEE, 1973, pp. 268–278.
[36] B. Chalmond, An iterative gibbsian technique for reconstruction of m-ary images, Pattern Recognit. 22 (6) (1989) 747–761.
[37] E. Kjartansson, Constant Q-wave propagtion and attenuation, J. Geophys. Res. (84) (1979) 4737–4747.
[38] L. Gelius, Inverse Q filtering. a spectral balancing technique, Geophys. Prospect. (35) (1987) 656–667.
[39] D. Hale, An inverse Q filter, Stanford Explor. Project Rep. (26) (1981) 231–243.
[40] Y. Wang, Seismic Inverse Q-filtering, Blackwell Pub., 2008.
[41] G. Margrave, M.P. Lamoureux, D. Henleyl, Gabor deconvolution: estimating reflectivity by nonstationary deconvolution of seismic data, Geophysics 76 (3) (2011) W15–W30.
[42] X. Chai, S. Wang, S. Yuan, J. Zhao, L. Sun, X. Wei, Sparse reflectivity inversion for nonstationary seismic data, Geophysics 79 (3) (2014) V93–V105.
[43] F. Li, S. Wang, X. Chen, G. Liu, Q. Zheng, A novel nonstationary deconvolution method based on spectral modeling and variable-step sampling hyperbolic smoothing, J. Appl. Geophys. 103 (2014) 132–139.
[44] A. Rosa, T.J. Ulyrch, Processing via spectral modeling, Geophysics 56 (8) (1991) 1244–1251.
[45] N. Ricker, The form and nature of seismic waves and the structure of seismogram, Geophysics 5 (4) (1940) 348–366.
[46] Y. Wang, Frequencies of the ricker wavelet, Geophysics 80 (2) (2015) A31–A37.
[47] Y. Wang, A stable and efficient approoach of inverse q filtering, Geophysics 67 (2) (2002) 657–663.
[48] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, 2014, (version 2.1).
[49] C. Zhang, T. Ulrych, Estimation of quality factors from CMP record, Geophysics 67 (5) (2002) 1542–1547.
[50] F. Zhang, The Schur Complement and its Applications, Springer, 2005.
[51] M. Spivak, Calculus, Cambridge, 1994.