

# Antileakage least-squares spectral analysis for seismic data regularization and random noise attenuation

Ebrahim Ghaderpour<sup>1</sup>, Wenyan Liao<sup>1</sup>, and Michael P. Lamoureux<sup>1</sup>

## ABSTRACT

Spatial transformation of an irregularly sampled data series to a regularly sampled data series is a challenging problem in many areas such as seismology. The discrete Fourier analysis is limited to regularly sampled data series. On the other hand, the least-squares spectral analysis (LSSA) can analyze an irregularly sampled data series. Although the LSSA method takes into account the correlation among the sinusoidal basis functions of irregularly spaced series, it still suffers from the problem of spectral leakage: Energy leaks from one spectral peak into another. We have developed an iterative method called antileakage LSSA to attenuate the spectral leakage and consequently regularize irregular data series. In this method, we first search for a

spectral peak with the highest energy, and then we remove (suppress) it from the original data series. In the next step, we search for a new peak with the highest energy in the residual data series and remove the new and the old components simultaneously from the original data series using a least-squares method. We repeat this procedure until all significant spectral peaks are estimated and removed simultaneously from the original data series. In addition, we address another problem, which is random noise attenuation in the data series, by applying a certain confidence level for significant peaks in the spectrum. We determine the robustness of our method on irregularly sampled synthetic and real data sets, and we compare the results with the antileakage Fourier transform and arbitrary sampled Fourier transform.

## INTRODUCTION

Regularization (a typical spectral interpolation/extrapolation) of irregularly sampled (unequally spaced) data is a crucial problem in seismology. Marine seismic data sets are usually irregularly sampled along the spatial direction because of cable feathering, editing bad traces, etc. Regularly sampled seismic data are required for various purposes, such as wave-equation migration, seismic inversion, amplitude variation with azimuth or offset analyses, and surface-related multiple elimination (Weglein et al., 1997; Xu et al., 2010; Chen et al., 2015). There are several effective algorithms proposed by researchers to interpolate seismic data such as the  $f$ - $x$  and  $t$ - $x$  prediction error filters (PEFs) (Spitz, 1991; Crawley, 2000; Fomel, 2002; Wang, 2002), masking in the frequency-wavenumber domain (Gülünay, 2003), projection onto convex sets (POCS) and

its modifications (Abma and Kabir, 2006; Yang et al., 2012; Wang et al., 2016), minimum weighted norm interpolation (MWNI) (Liu and Sacchi, 2004; Trad, 2009), shaping regularization (Fomel, 2007; Chen et al., 2015), antileakage Fourier transform (ALFT) (Xu et al., 2005, 2010, 2011), and arbitrarily sampled Fourier transform (ASFT) (Guo et al., 2015).

In the PEF theory, the data must be wide-sense stationary and regular. The  $f$ - $x$  PEFs represent a spectral approach that uses the available traces to predict the linear events, and it is suitable for aliased seismic data. For seismic data, windowing is required to approximately fulfill the stationarity condition of the PEFs. Nonstationary PEF has also been used in seismic data interpolation to solve the interpolation of curved events (Crawley, 2000; Liu and Chen, 2017). The POCS method is based on the Gerchberg-Saxton iterative algorithm (Gerchberg and Saxton, 1972) widely used in industry. Its performance

Manuscript received by the Editor 6 May 2017; revised manuscript received 5 December 2017; published ahead of production 26 January 2018; published online 16 March 2018.

<sup>1</sup>University of Calgary, Department of Mathematics and Statistics, Calgary, Alberta, Canada. E-mail: ebrahim.ghaderpour@ucalgary.ca; wliao@ucalgary.ca; mikel@ucalgary.ca.

© The Authors. Published by the Society of Exploration Geophysicists. All article content, except where otherwise noted (including republished material), is licensed under a Creative Commons Attribution 4.0 Unported License (CC BY-SA). See <http://creativecommons.org/licenses/by/4.0/>. Distribution or reproduction of this work in whole or in part commercially or noncommercially requires full attribution of the original publication, including its digital object identifier (DOI). Derivatives of this work must carry the same license.

on noisy data and its computational efficiency have been recently improved (Ge et al., 2015; Wang et al., 2016). Abma and Kabir (2006) use a linear threshold in the iterations of the POCS method and point out that the threshold is the key parameter in this method. Thereafter, other works have been done on an appropriate selection of a threshold to improve the result of the POCS method (Gao et al., 2013). The MWNI method solves an underdetermined system of equations subject to suitable prior information for the purpose of regularization. The linear shaping regularization is another approach for regularizing an underdetermined geophysical inverse problem that uses a linear shaping operator. The nonlinear shaping regularization allows the shaping operator to be nonlinear making it more flexible in implementing the iterative framework.

The ALFT and ASFT methods estimate the Fourier coefficients of the data series first by searching for the peak with maximum energy and subtracting its component from the data series and then by repeating the procedure on the residual data series until all the significant Fourier coefficients are estimated. This way, the spectral leakages of Fourier coefficients emerging from the nonorthogonality of the global Fourier basis functions will be attenuated. However, these methods usually cannot find the correct location of a peak with maximum energy from a preselected set of wavenumbers because of the correlation between the sinusoids, and so they do not effectively reduce the leakage. This shortcoming becomes more severe when a data series has more spectral components. Moreover, the constituents of known forms such as datum shifts and trends are not explicitly considered in these algorithms. Therefore, if the trends or datum shifts are present in a data series, then the ALFT and ASFT methods implicitly approximate them by sinusoids of various wavenumbers (Xu et al., 2010). Hollander et al. (2012) propose a method that uses an orthogonal matching pursuit (OMP) to improve the ALFT results. In the OMP approach, the coefficients of all previously selected Fourier components are reoptimized, producing better interpolation results compared with the ALFT results.

Least-squares spectral analysis (LSSA) was introduced by Vaníček (1969) to analyze unequally spaced time series. It basically estimates a frequency spectrum based on the least-squares fit of sinusoids to the time series by accounting for measurement errors, trends, and constituents of known forms. In the LSSA method because the selected frequencies are examined one by one (out-of-context), the spectral leakages still appear and result in interpolation inaccuracy, although the nonorthogonality between the sine and cosine basis functions is taken into account for each frequency. The statistical properties of the least-squares spectrum (LSS) were discussed later by Lomb (1976), Craymer (1998), and Pagiatakis (1999). Seismic data are well-sampled along the time direction (regularly sampled), but they are usually are poorly sampled along the spatial direction (irregularly spaced). Note that the LSSA method and its applications are mainly discussed on an unequally spaced time series in the literature; however, we can concurrently apply its concept in the spatial direction that is based on wavenumber (cycles per unit distance) rather than frequency (cycles per unit time). Therefore, when we use the term “data series,” we mean a series of data points sampled along the spatial direction.

In this contribution, we apply the idea of maximum energy in the LSSA method to develop the antileakage LSSA (ALLSSA) method that mitigates the spectral leakages significantly. We further show how the ALLSSA method attenuates the random noise present in data down to a certain confidence level. The main assumption in the

ALFT, ASFT, LSSA, and ALLSSA methods is that the data must be wide-sense stationary (Xu et al., 2010; Ghaderpour and Pagiatakis, 2017). However, we note that the LSSA method considers the constituents of known forms such as trends and/or datum shifts explicitly, so they are somewhat adaptable to the nonstationary data series (Craymer, 1998; Ghaderpour and Pagiatakis, 2017). In the synthetic data examples here, we show that if the data are wide-sense stationary, then the ALLSSA method is an excellent and very accurate method for regularization and random noise attenuation. Özbek et al. (2009) propose an iterative algorithm called interpolation by matching pursuit (IMAP) to regularize irregularly sampled seismic data. The IMAP method improves the original Lomb spectrum (LSS) result, and we will briefly compare it with the ALLSSA method.

However, seismic data are often nonstationary and contain many nonlinear events (Crawley, 2000). It is traditional to use windowing of seismic images in  $t$ - $x$  and  $f$ - $x$  prediction (Abma and Claerbout, 1995). The  $f$ - $x$  domain interpolation algorithms for regularly sampled data also need data windowing. Windowing strategy makes the nonlinear events approximately linear and usually works well when the data do not have a complex geologic structure. Liu et al. (2012) apply regularized nonstationary autoregression in the  $f$ - $x$  domain, which does not require windowing strategies in the spatial direction. Liu et al. (2012) propose a method called local parallel radial-trace time-frequency peak filtering (TFPF), which is less sensitive to the fixed window length used in the TFPF and performs well for random noise attenuation.

Windowing strategy also significantly decreases the computational time in the ALFT method; however, the problem of spatial aliasing may occur more likely within the windows (Schonewille et al., 2009; Xu et al., 2010). Because the events are approximately linear within the windows, an antialiasing scheme (Schonewille et al., 2009) may be used inside the ALFT method (and similarly inside the ASFT and ALLSSA methods) to improve the regularization result. In this paper, we will also use an appropriate weight matrix defined by the Gaussian function in the ALLSSA method to regularize the irregularly spaced seismic data more accurately. In addition, we shall show in our field data example that repeating the ALLSSA method can attenuate the random noise down to a certain level (typically, 10% of seismic data might be random noise).

It is customary to transform each trace from the time domain to the frequency domain using the fast Fourier transform (FFT) (Spitz, 1991; Abma and Claerbout, 1995; Xu et al., 2010; Guo et al., 2015). Then for each frequency, we generate a data series whose data points, located at the trace locations, are the Fourier coefficients of that frequency (a temporal frequency slice). For irregularly spaced seismic data, each temporal frequency slice is an irregularly sampled data series, and the ALFT or ASFT or ALLSSA method may be used to regularize the data series, and then each trace can be transformed back from the frequency domain to the time domain by the inverse FFT. Note that in this work, all the analyses are performed on the temporal frequency slices. We now begin by revisiting the ALFT, ASFT, and LSSA methods in more detail, and then we combine the ideas of these methods to introduce the ALLSSA method.

## METHODS

### Antileakage Fourier transform

Suppose that  $\{f(x_\ell): \ell = 1, \dots, n\}$  is the set of  $n$  ( $n > 1$ ) data points, where the  $x_\ell$ s may be irregularly sampled. We assume that

$x_\ell$  is a real number in the closed interval  $[0,1]$ , where “1” is the unit distance (e.g., 1 km). We may define  $\Delta x_\ell = (1/2)(x_{\ell+1} - x_{\ell-1})$  for  $1 < \ell < n$ . For the two special cases  $\ell = 1$  and  $\ell = n$ , we may define  $\Delta x_1 = x_2 - x_1$  and  $\Delta x_n = x_n - x_{n-1}$ , respectively. Let  $\Delta x = \sum_{\ell=1}^n \Delta x_\ell$ . If  $x_\ell = (\ell/n)d$  for  $\ell = 1, \dots, n$ , then  $\Delta x_\ell = d/n$  and  $\Delta x = d$ , where  $d$  is the trace distance. For simplicity, in this work, we normalize offsets to make the offset of the far offset trace equal to one, and so  $\Delta x = 1$ . The forward discrete summation of Fourier integral (after dividing by the summation range) for wavenumber  $\omega_k \in \mathbf{K}$  is defined as

$$\hat{f}(\omega_k) = \frac{1}{\Delta x} \sum_{\ell=1}^n \Delta x_\ell f(x_\ell) e^{-2\pi i \omega_k x_\ell}, \quad (1)$$

and the contribution of  $\omega_k$  to the inverse is defined as

$$f^k(x_\ell) = \hat{f}(\omega_k) e^{2\pi i \omega_k x_\ell}, \quad (2)$$

where  $\mathbf{K} = \{(-n/2), \dots, (n/2) - 1\}$  if  $n$  is even and  $\mathbf{K} = \{-(n-1)/2, \dots, (n-1)/2\}$  if  $n$  is odd. The wavenumber  $\omega_k$  means the  $\omega_k$  cycles of sinusoids per unit distance. The (spatial) sampling rate is defined as the number of samples (traces) per unit distance. Note that for regularly sampled data series, the Nyquist wavenumber is half of the (spatial) sampling rate; however, the Nyquist wavenumber does not explicitly exist for inherently unequally spaced data series. Therefore, we may implicitly choose the maximum wavenumber for the analysis to be less than half of the sampling rate (the aliasing may not occur for the wavenumber beyond this selection). If a data series is assumed to be regularly spaced, but some samples are missing, then the Nyquist wavenumber may be half of the sampling rate for the regularly spaced data series (Craymer, 1998).

If the  $x_\ell$ 's are regularly spaced, then the estimation of the Fourier coefficient  $\hat{f}(\omega_k)$  has no effect on the estimations of other Fourier coefficients because the sinusoids are orthogonal; i.e., the dot product between the sinusoids is zero (see Appendix A). However, the orthogonality condition fails for an irregularly sampled data series resulting in spectral leakages. To mitigate these leakages, Xu et al. (2005) propose the ALFT method, which has the following three steps:

- 1) Compute all Fourier coefficients of the input data series using equation 1. Note that samples of function  $f$  in equation 1 can be the samples of a frequency slice in seismic applications.
- 2) Select the Fourier coefficient with the maximum energy.
- 3) Subtract the contribution of this coefficient (equation 2) from the input data series to obtain the residual data series. Use this residual data series as the new input and go to step 1. Continue this procedure until the L2 norm of the residual data series becomes approximately zeros (below a threshold).

Although the ALFT method is fairly efficient and effective in practical applications, it has some shortcomings. The main shortcoming of the ALFT method is that it only uses wavenumbers in a preselected set of wavenumbers (e.g., set  $\mathbf{K}$ ) and does not search for the actual wavenumbers (real numbers) of components in a data series that may not be in the preselected set of wavenumbers. In other words, the wavenumber context of the data may not fall on one of the wavenumber samples used in set  $\mathbf{K}$ . Therefore, for a better regularization result, a denser set of wavenumbers may be selected in the ALFT method (Xu et al., 2011). For a data series of size  $n$  with  $N$  wavenumbers, the ALFT method requires the com-

putation of  $N$  Fourier coefficients for each iteration, which has an order of  $nN$  floating operations, and it usually takes  $N$  iterations to converge. Thus, the total computational cost in the ALFT method is on the order of  $nN^2$  (Xu et al., 2005, 2010). Therefore, a denser set of wavenumbers will greatly increase the computational cost, and the actual wavenumbers of the components cannot still be accurately estimated. In practice, however, one may need fewer iterations when dealing with data that have by nature only a few different spectral components.

### Arbitrary sampled Fourier transform

Guo et al. (2015) propose a method that is similar to the ALFT method, but it can estimate the wavenumbers of the components in a data series more accurately. In this section, we describe their method called the arbitrary sampled Fourier transform (ASFT). Consider the following cost function:

$$\Lambda(\omega_k) = \sum_{\ell=1}^n |f(x_\ell) - \hat{f}(\omega_k) e^{2\pi i \omega_k x_\ell}|^2, \quad (3)$$

where  $\hat{f}(\omega_k)$  is given in equation 1 and  $|\cdot|$  denotes the absolute value. The ASFT method has the following three steps:

- 1) For each  $\omega_k \in \mathbf{K}$ , compute  $\hat{f}(\omega_k)$  and find the wavenumber  $\omega_k$  such that  $\Lambda(\omega_k)$  is minimum.
- 2) Find wavenumber  $h_k$  in  $I = [\omega_k - b, \omega_k + b]$  such that  $\Lambda(h_k)$  is minimum. Because the wavenumbers in  $\mathbf{K}$  are integer, we may choose  $b = 0.5$ .
- 3) Subtract the contribution  $\hat{f}(h_k) e^{2\pi i h_k x_\ell}$  from the input data series to obtain the residual data series. Use this residual data series as the new input, and go to step 1. Continue this procedure until the L2 norm of the residual data series becomes approximately zero (below a threshold).

Estimating  $h_k$  up to a chosen decimal place in step 2 may simply be done by an appropriate partitioning of interval  $I$  as follows: For  $b = 0.5$ , compute  $\Lambda(h_{k_1})$  for all wavenumbers  $h_{k_1}$  in set  $\{\omega_k - 0.5 + 0.1j; j = 1, \dots, 9\}$  and choose wavenumber  $h_{k_1}$  that  $\Lambda(h_{k_1})$  is minimum. Note that wavenumber  $h_{k_1}$  is estimated up to one decimal place. Next, compute  $\Lambda(h_{k_2})$  for all wavenumbers  $h_{k_2}$  in set  $\{h_{k_1} - 0.1 + 0.01j; j = 1, \dots, 19\}$  and choose wavenumber  $h_{k_2}$  that  $\Lambda(h_{k_2})$  is minimum. Note that wavenumber  $h_{k_2}$  is now estimated up to two decimal places. We may continue this process to estimate the wavenumber  $h_k$  up to any chosen decimal place. Estimating  $h_k$  up to four decimal places by using this technique requires  $9 + 3(19) = 66$  times calculation of equation 3 that is computationally much faster than 9999 (equally spaced numbers in interval  $I$ ) times calculation of equation 3, and it is also practically acceptable. For a fair comparison between the ASFT and ALLSSA methods, we use this partitioning technique for both methods. One may also apply other techniques such as the conjugate gradient or the Broyden-Fletcher-Goldfarb-Shanno algorithm for step 2 (Guo et al., 2015) or other robust minimizers such as simulated annealing (Chen and Luk, 1999).

Note that wavenumbers estimated in the ASFT method are real numbers obtained directly from the cost function minimization that is easier to interpret (Guo et al., 2015). One of the advantages of the ASFT method over the ALFT method is that the wavenumbers are

estimated more accurately at each step. Therefore, if a threshold is defined for the L2 norm of residual data series, then the ASFT method is able to arrive at that threshold in a smaller number of iterations compared with the ALFT method, resulting in more accuracy and less computational cost. In other words, for a data series of size  $n$  with  $n$  preselected wavenumbers, the ASFT method has an order of  $n^2m$  floating operations, where  $m$  is the number of iterations that is usually smaller than  $n$  by selecting an appropriate threshold. For an irregularly sampled data series with several components, although the ASFT method can estimate the wavenumbers of the components more accurately than the ALFT method, its spectrum still shows spectral leakages caused by the correlations between the sinusoids. Therefore, the wavenumbers of the data components cannot be estimated very accurately, and its spectrum contains many wavenumbers.

### Least-squares spectral analysis

Let  $\mathbf{f} = [f(x_\ell)]$  be the column vector of  $n$  data points. Note that  $x_\ell \in [0, 1]$ , where one is the maximum distance. Let  $\Phi_k = [\cos(2\pi\omega_k x_\ell), \sin(2\pi\omega_k x_\ell)]$ ,  $\ell = 1, \dots, n$ , be the  $n \times 2$  matrix of a fixed wavenumber  $\omega_k \in \Omega$ , where  $\Omega$  can be any set of real numbers. When a data series is regularly spaced, for the simultaneous estimation of the amplitudes and phases,  $\Omega$  may be selected as the set of all positive integers less than the Nyquist wavenumber (see Appendix A). Suppose that  $\mathbf{f}$  has been derived from a population of (complex) random variables, and  $\mathbf{C}_f$  is the (regular and Hermitian) covariance matrix associated with  $\mathbf{f}$  defined as

$$\mathbf{C}_f = \begin{bmatrix} \sigma_{f(x_1)}^2 & \sigma_{f(x_1)f(x_2)} & \cdots & \sigma_{f(x_1)f(x_n)} \\ \sigma_{f(x_2)f(x_1)} & \sigma_{f(x_2)}^2 & \cdots & \sigma_{f(x_2)f(x_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{f(x_n)f(x_1)} & \sigma_{f(x_n)f(x_2)} & \cdots & \sigma_{f(x_n)}^2 \end{bmatrix}, \quad (4)$$

where  $\sigma_{f(x_\ell)}^2$  is the variance of (complex) random variable  $f(x_\ell)$  and  $\sigma_{f(x_u)f(x_v)}$  is the covariance between two (complex) random variables  $f(x_u)$  and  $f(x_v)$ . The covariance  $\sigma_{f(x_u)f(x_v)}$  is a measure of the joint variability between random variables  $f(x_u)$  and  $f(x_v)$  defined as

$$\begin{aligned} \sigma_{f(x_u)f(x_v)} &= E[(f(x_u) - E[f(x_u)])(\overline{f(x_v) - E[f(x_v)]})] \\ &= E[f(x_u)\overline{f(x_v)}] - E[f(x_u)]\overline{E[f(x_v)]}, \end{aligned} \quad (5)$$

where  $\overline{f(x)}$  is the conjugate of  $f(x)$  and  $E[X]$  is the expected value of (complex) random variable  $X$  (Eriksson et al., 2009).

Let  $\mathbf{P} = \mathbf{C}_f^{-1}$ . In many practical applications,  $\mathbf{P}$  is approximately a diagonal matrix (the correlations between the data points are negligible), so we may treat it as a vector for computational efficiency (Pagiatakis, 1999; Ghaderpour and Pagiatakis, 2017). In the LSSA method, we minimize the cost function  $\Psi(\mathbf{c}_k) = (\mathbf{f} - \Phi_k \mathbf{c}_k)^T \mathbf{P}(\mathbf{f} - \Phi_k \mathbf{c}_k)$  to estimate  $\mathbf{c}_k$ , so we obtain  $\hat{\mathbf{c}}_k = (\Phi_k^T \mathbf{P} \Phi_k)^{-1} \Phi_k^T \mathbf{P} \mathbf{f}$ , where “T” indicates the transpose and conjugate transpose when  $\mathbf{f}$  is complex. Note that  $\hat{\mathbf{c}}_k$  is a  $2 \times 1$  matrix (a column vector) whose elements are real or complex depending on  $\mathbf{f}$ . The normalized LSS for  $\omega_k \in \Omega$  is defined as

$$s_0(\omega_k) = \frac{\mathbf{f}^T \mathbf{P} \Phi_k \hat{\mathbf{c}}_k}{\mathbf{f}^T \mathbf{P} \mathbf{f}}. \quad (6)$$

Note that  $s_0(\omega_k) \in (0, 1)$ , and it shows how much the sinusoids of wavenumber  $\omega_k$  contribute to the entire data series. Now if we know that there are constituents of known forms in the data series, we may first estimate and remove them from the data series and then study the residual data series. More precisely, let  $\underline{\Phi} = [\Phi_1, \dots, \Phi_q]$  be the  $n \times q$  matrix of constituents of known forms. The constituents can be the column vector of all ones  $[\mathbf{1}]$  and/or the position column vector  $[\mathbf{x}]$  to explicitly account for any linear trend and/or the column vectors of the sinusoids of particular wavenumbers.

We first minimize the cost function  $\Psi_1(\underline{\mathbf{c}}) = (\mathbf{f} - \underline{\Phi} \underline{\mathbf{c}})^T \mathbf{P}(\mathbf{f} - \underline{\Phi} \underline{\mathbf{c}})$  to estimate  $\underline{\mathbf{c}}$  as

$$\hat{\underline{\mathbf{c}}} = \underline{\mathbf{N}}^{-1} \underline{\Phi}^T \mathbf{P} \mathbf{f}, \quad (7)$$

where  $\underline{\mathbf{N}} = \underline{\Phi}^T \mathbf{P} \underline{\Phi}$ , and obtain the residual data series

$$\mathbf{g} = \mathbf{f} - \underline{\Phi} \hat{\underline{\mathbf{c}}}. \quad (8)$$

Next, we minimize the cost function  $\Psi_2(\underline{\mathbf{c}}, \mathbf{c}_k) = (\mathbf{f} - \underline{\Phi} \underline{\mathbf{c}} - \Phi_k \mathbf{c}_k)^T \mathbf{P}(\mathbf{f} - \underline{\Phi} \underline{\mathbf{c}} - \Phi_k \mathbf{c}_k)$  to estimate  $\mathbf{c}_k$ . In Appendix B, we show that

$$\hat{\mathbf{c}}_k = (\Phi_k^T \mathbf{P} \Phi_k - \Phi_k^T \mathbf{P} \underline{\Phi} \underline{\mathbf{N}}^{-1} \underline{\Phi}^T \mathbf{P} \Phi_k)^{-1} \Phi_k^T \mathbf{P} \mathbf{g}. \quad (9)$$

We call the term within the parentheses in equation 9, the matrix of normal equations that is a real square matrix of order two (Craymer, 1998). If  $\mathbf{f}$  is real, then  $\hat{\mathbf{c}}_k$  is a real column vector of order two whose elements are the estimated amplitudes for cosine and sine basis functions. If  $\mathbf{f}$  is complex, then  $\hat{\mathbf{c}}_k$  is a complex column vector whose real and imaginary parts correspond to the estimated amplitudes for cosine and sine basis functions of the real and imaginary parts of  $\mathbf{f}$ , respectively. The normalized LSS of the residual data series for  $\omega_k \in \Omega$  is then defined by

$$s(\omega_k) = \frac{\mathbf{g}^T \mathbf{P} \Phi_k \hat{\mathbf{c}}_k}{\mathbf{g}^T \mathbf{P} \mathbf{g}}. \quad (10)$$

Note that  $s(\omega_k) \in (0, 1)$ , and it shows how much the sinusoids of wavenumber  $\omega_k$  contributes to the residual data series. Ideally, if  $\underline{\Phi}$  is an invertible square matrix of order  $n$ , then  $\mathbf{g} = \mathbf{0}$  (Craymer, 1998).

To demonstrate the estimated amplitude and phases simultaneously, we may show the LSS in terms of wavenumber amplitude by  $\hat{\mathbf{c}}_k$  in equation 9. In other words, for each  $\omega_k \in \Omega$ , we calculate the square root of the sum of squares of the first and second elements of  $\hat{\mathbf{c}}_k$  (separately for the real and imaginary parts of  $\hat{\mathbf{c}}_k$  if  $\mathbf{f}$  is complex). This way, the LSS can be compared with the Fourier spectrum expressed in terms of absolute values of the Fourier coefficients as will be described in our synthetic data series example.

It is shown (Pagiatakis, 1999) that if  $\mathbf{f}$  has been derived from a population of random variables following the multidimensional normal distribution, then the spectrum given in equation 10 follows the beta distribution with parameters  $\alpha = 1$  and  $\beta = (n - q - 2)/2$ , where  $q$  is the number of constituents of known forms and  $n$  is the number of data points. We refer the reader to Craig et al. (2013) for a definition of the beta distribution. The critical value at  $(1-c)$  confidence level for the spectrum is then equal to  $1 - c^{1/\beta}$ , where  $c$  is the significance level (usually 0.05 or 0.01). The critical value is independent of the weight matrix and wavenumbers (Pagiatakis, 1999; Ghaderpour and Pagiatakis, 2017).



## Antileakage LSSA

We apply the idea of maximum energy in the algorithm of the ALFT method to develop a new algorithm for the LSSA method that significantly reduces the spectral leakages and consequently regularizes the irregularly sampled data series very accurately. We may assume that  $\Omega = \{1, 2, \dots, \eta\}$ , where  $\eta$  may be chosen as the largest positive integer less than half of the sampling rate. Beginning with an initial  $\Phi$  that may be selected as  $\Phi = [1, \mathbf{x}]$ , the algorithm for the ALLSSA method has the following four steps:

- 1) Given  $\Phi$ , compute  $\hat{\mathbf{c}}$  and  $\mathbf{g}$  (equations 7 and 8).
- 2) For all  $\omega_k \in \Omega$ , calculate  $\hat{\mathbf{c}}_k$  (equation 9) and find  $\omega_u$  in  $\Omega$  such that  $s(\omega_u)$  (equation 10) is maximum.
- 3) Eliminate the cosine and sine basis functions of a wavenumber  $h$  in  $\Phi$  such that  $|h - \omega_u| < b$  (if it exists). Assuming that the difference between any two consecutive actual wavenumbers of the components in the data series is greater than one, we may choose  $b = 0.5$  to resolve the peaks and avoid singularity of the matrix of normal equations.
- 4) Find  $h$  in  $I = [\omega_u - b, \omega_u + b]$  such that  $s(h)$  (equation 10) is maximum. Concatenate the cosine and sine column vectors of wavenumber  $h$  to  $\Phi$  and then go to step 1. Finding  $h$  up to a chosen decimal place can be done by an appropriate partitioning of  $I$ . Terminate the process if  $s(h)$  is no longer significant at certain confidence level (usually 95% or 99%).

When we consider the weight matrix  $\mathbf{P}$ , the method is called the weighted ALLSSA. Removing the constituent of a particular wavenumber with maximum energy from the data series reduces the spectral leakages of the residual spectrum. On the other hand, the elimination of the basis functions of a wavenumber  $h$  in step 3 is crucial in the ALLSSA method because considering the correlation between the sinusoids of different wavenumbers results in a more accurate  $h$  in the next step. This correlation will be taken into account as the column dimension of  $\Phi$  increases in the process. The L2 norm of residual data series  $\mathbf{g}$  approaches zero rapidly because the wavenumbers are accurately estimated recursively that also prevents increasing the column dimension of  $\Phi$  in many practical applications especially when we apply the confidence level. We define the antileakage LSS (ALLSS) in terms of wavenumber amplitude by  $\hat{\mathbf{c}}$  in step 1 after the process is terminated.

Similar to the discussion in the previous section,  $\hat{\mathbf{c}}$  is a real or complex column vector (depending on  $\mathbf{f}$ ) whose elements are the estimated intercept and slope of the linear trend (if the initial  $\Phi$  is  $[1, \mathbf{x}]$ ) and the estimated amplitudes for the sine and cosine basis functions of the estimated wavenumbers. We use this  $\hat{\mathbf{c}}$  for the regularization of a stationary data series that will also result in attenuating the random noise present in the data series down to certain confidence level. For nonstationary data series, we may also treat the final residual data series as a new input data series then perform the ALLSSA method again (by applying the confidence level) and repeat this procedure until the L2 norm of the final residual data series approaches zero (below a threshold). Note that the actual wavenumbers of the data series will be better approximated if the initial  $\Phi$  at least contains the column vector of all ones (Foster, 1996; Ghaderpour and Pagiatakis, 2017).

The computational complexity of the ALLSSA method depends on the estimation of  $\hat{\mathbf{c}}$  and the matrix of normal equations in each iteration. For a data series of size  $n$  with  $\eta$  wavenumbers (e.g.,

$\eta = (n/2) - 1$  if  $n$  is even and  $\eta = (n - 1)/2$  if  $n$  is odd), the process has an order of  $n\eta m$  floating operations that  $m$  mainly depends on the number of significant spectral peaks. When applying the confidence level, the total number of the significant peaks is usually small, making  $m$  small, so the computational efficiency can be much better than the ALFT and ASFT methods for stationary data series.

In the IMAP method, a wavenumber  $k_1$  corresponding to the largest peak in the Lomb spectrum (Lomb, 1976) is selected. Then, the contribution of the sinusoidal basis functions of wavenumber  $k_1$  to the data series is subtracted from it to obtain the residual data series. In the next step, the residual data series is treated as a new input data series to choose a wavenumber  $k_2$  corresponding to the largest peak in the new Lomb spectrum, and this procedure continues until the desired wavenumbers are estimated. This iterative method mitigates the spectral leakages in the Lomb spectrum and results in a faster convergence (Özbek et al., 2009). Although the IMAP method improves the spectral leakages in the Lomb spectrum, it does not explicitly account for the correlations among the sinusoids of different wavenumbers. Vassallo et al. (2010) have also studied estimating many wavenumbers per iteration and have documented several dealiasing schemes.

On the other hand, in the ALLSSA method, the sinusoidal basis functions of different wavenumbers, corresponding to the largest statistically significant peaks in the spectrum, are added to the columns of design matrix  $\Phi$  in an iterative manner. Therefore, the wavenumbers will be more accurately estimated in the next step by considering the correlation among the columns. The accuracy of estimated wavenumbers significantly reduces the number of iterations and results in more accurate regularization.

## SYNTHETIC DATA SERIES EXAMPLE

We use the ALLSSA method to regularize the data series of size 128 given by equation 13 in Xu et al. (2005) with an additional sine wave and a linear trend

$$f(x_\ell) = 5 \sin(25.6x_\ell) + 2.5 \sin(128x_\ell + 1) + \sqrt{3} \sin(140x_\ell) + \sqrt{2} + \pi x_\ell, \quad (11)$$

where  $x_\ell$  ( $\ell = 1, 2, 3, \dots, 128$ ) is a random number in  $[0, 1]$  generated by the MATLAB command “rand.” All the  $x_\ell$ s are sorted in ascending order (see Table 1).

This data series is inherently unequally spaced, and so the Nyquist wavenumber does not explicitly exist. Therefore, there is no restriction for the selection of maximum wavenumber in this example. However, because our goal is to regularize this data series on a series with regular spacing  $1/128$  (the Nyquist wavenumber of this regular series is 64), we choose the initial set of wavenumbers as  $\Omega = \{1, 2, 3, \dots, 63\}$  for the LSSA and ALLSSA methods, and we estimate the wavenumbers up to four decimal places. Note that the data series has three wavenumber components with wavenumbers  $25.6/2\pi \approx 4.0743665$ ,  $128/2\pi \approx 20.3718327$ , and  $140/2\pi \approx 22.2816920$ , rounded to seven decimal places.

For the ALFT and ASFT methods, we also choose  $\mathbf{K} = \{-64, -63, \dots, 63\}$ . We illustrate their spectra with green and blue dots in Figure 1a, respectively. The ALFT and ASFT methods implicitly accounts for the linear trend using the sinusoids of various wavenumbers including zero. Figure 1a clearly shows the presence of the spectral leakages in the ALFT and ASFT methods

(with the threshold 0.5). However, the actual wavenumbers of the data series in the ASFT method are estimated more accurately than the ALFT method (see the numbers shown in blue).

For the LSSA and ALLSSA methods, we select the initial  $\Phi$  as  $\Phi = [\mathbf{1}, \mathbf{x}]$  to explicitly account for the linear trend present in the data series, where  $\mathbf{1}$  and “ $\mathbf{x}$ ” are the column vectors of all ones and the position vector, respectively. Because there is no covariance matrix associated with the data series, we do not consider the square weight matrix  $\mathbf{P}$  in the calculation. We illustrate the LSS (black) and the ALLSS (red) in terms of wavenumber amplitude, shown as “Amplitude (Abs)” on the  $y$ -axis, in Figure 1b. The LSS shows one strong peak at wavenumber 4 that is the closest value to the actual value 4.0743665; however, the other two signal peaks are buried in the leakage spectrum. On the other hand, it can be seen from the ALLSS that only the actual signal peaks are estimated very accu-

rately at a 99% confidence level. Note that we estimated the signal peaks up to only four decimal places, and we choose this accuracy for our synthetic and field data examples in this contribution. The result of the ALLSSA regularization for the data series without the linear trend is also more accurate than the one for the ALFT and ASFT regularization (not shown here).

Table 2 shows the result of the ALLSSA method. In the first iteration, the wavenumber 4.0384 is estimated that is approximately 0.036 different from its actual value 4.0743665. This is a shortcoming of the LSSA method (out-of-context) caused by the presence of other constituents in the data series. In the second and third iterations, the other two wavenumbers are estimated that by removing their corresponding components from the data series simultaneously, the first wavenumber is better approximated and so forth (see the highlighted numbers in Table 2).

**Table 1. The values of the data series given in equation 11.**

$\ell$	$x_\ell$	$f(x_\ell)$	$\ell$	$x_\ell$	$f(x_\ell)$	$\ell$	$x_\ell$	$f(x_\ell)$	$\ell$	$x_\ell$	$f(x_\ell)$
1	0.0012	4.1459	33	0.2554	3.7746	65	0.5114	5.4213	97	0.7257	2.9864
2	0.0098	6.3235	34	0.2568	3.5381	66	0.5246	3.206	98	0.7269	3.6399
3	0.011	6.2455	35	0.2699	3.2506	67	0.5294	4.7075	99	0.7466	5.59
4	0.0249	1.733	36	0.2732	3.9738	68	0.5323	6.158	100	0.7477	5.3266
5	0.0569	10.5501	37	0.2789	5.6463	69	0.5431	11.4665	101	0.7522	4.2909
6	0.0577	10.4264	38	0.2794	5.8081	70	0.5566	9.1851	102	0.7615	4.2617
7	0.0606	9.5492	39	0.2848	7.3645	71	0.5694	3.6382	103	0.7624	4.4445
8	0.061	9.38	40	0.2851	7.4331	72	0.5733	3.663	104	0.789	10.4141
9	0.0614	9.2229	41	0.2874	7.9229	73	0.5761	4.2949	105	0.7897	10.3048
10	0.0887	4.7172	42	0.2915	8.4397	74	0.5761	4.2957	106	0.8106	7.1337
11	0.09	5.3539	43	0.3119	6.6108	75	0.5816	6.5607	107	0.8277	8.5478
12	0.0903	5.5206	44	0.3164	6.2615	76	0.5825	6.9787	108	0.83	8.5057
13	0.0915	6.0724	45	0.3178	6.2017	77	0.5846	7.9723	109	0.8371	7.7299
14	0.0921	6.3641	46	0.3209	6.1423	78	0.5847	8.0061	110	0.8372	7.7104
15	0.1097	5.5084	47	0.3416	6.203	79	0.595	9.7637	111	0.8383	7.5168
16	0.1099	5.4367	48	0.3473	5.8137	80	0.6003	7.8217	112	0.8464	5.7735
17	0.1161	1.33	49	0.3738	0.9117	81	0.6146	-0.6803	113	0.8504	4.899
18	0.1232	-2.1801	50	0.3747	0.75	82	0.6203	-1.4524	114	0.8586	3.415
19	0.1311	-2.2501	51	0.3907	-0.06	83	0.6253	-0.323	115	0.8588	3.3873
20	0.1478	2.7437	52	0.3927	0.0796	84	0.6393	4.3755	116	0.9064	-0.5541
21	0.1519	2.1289	53	0.3971	0.3735	85	0.6423	4.1792	117	0.9111	-1.3885
22	0.1726	-6.4868	54	0.4038	0.3453	86	0.6455	3.2762	118	0.9116	-1.4815
23	0.1858	-2.5312	55	0.4315	-4.1425	87	0.65	1.0609	119	0.9329	0.0564
24	0.1888	-1.2075	56	0.435	-3.4527	88	0.6714	-4.0102	120	0.9462	2.2945
25	0.191	-0.4057	57	0.439	-2.2456	89	0.6723	-3.6663	121	0.9481	2.2368
26	0.193	0.1571	58	0.448	0.6196	90	0.6866	2.4392	122	0.9573	0.8809
27	0.1982	0.5938	59	0.4503	1.0558	91	0.6944	1.8507	123	0.9622	0.1906
28	0.2213	-3.6923	60	0.4522	1.2574	92	0.6996	-0.0216	124	0.9711	0.906
29	0.2299	-0.6893	61	0.4963	6.7645	93	0.6996	-0.0283	125	0.9715	1.0231
30	0.2354	1.8974	62	0.5017	7.6344	94	0.7104	-2.7925	126	0.9727	1.3772
31	0.2369	2.5595	63	0.5044	7.421	95	0.7194	-0.3853	127	0.9798	4.4223
32	0.2553	3.7928	64	0.506	7.1136	96	0.7252	2.71	128	0.9884	8.0317

We illustrate the ideal data series and the regularization results using the three methods in Figure 2a. We also show the differences between each ideal data point and the estimated one for the three methods in Figure 2b. The L2 norm of the ideal regularized data series is 57.382. The L2 norm of the difference between the ideal data series and the regularized data series for the ALFT, ASFT, and ALLSSA methods are 9.864, 1.26, and 0.004, respectively. For the ALLSSA method, we only used the estimated slope and intercept of the linear trend as well as the estimated amplitudes and phases of the three wavenumbers for the regularization. Note that when a smaller threshold is selected for the ASFT method, the computational time significantly increases, and its spectrum will have much more wavenumbers with very low amplitudes (absolute value of Fourier coefficients), and it still cannot reach the accuracy of the ALLSSA method.

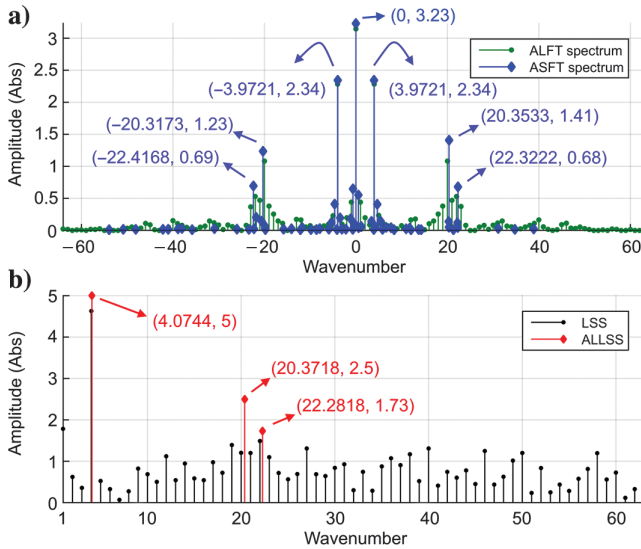


Figure 1. The spectra of the irregularly sampled data series given in equation 11. (a) The ALFT (green) and ASFT (blue) spectra, and (b) the LSS (black) and ALLSS (red).

Table 2. The result of wavenumber estimation of the ALLSSA algorithm on the irregularly sampled data series given in equation 11 after each iteration.

Iteration number	First wavenumber	Second wavenumber	Third wavenumber	Norm of residual
First	4.0384	—	—	23.5040
Second	4.0384	20.3057	—	13.3170
Third	4.0384	20.3057	22.2947	3.5079
Fourth	<b>4.0748</b>	20.3057	22.2947	2.2501
Fifth	4.0748	<b>20.3708</b>	22.2947	0.3068
Sixth	4.0748	20.3708	<b>22.2818</b>	0.0468
Seventh	4.0748	<b>20.3718</b>	22.2818	0.0319
Eighth (final)	<b>4.0744</b>	20.3718	22.2818	0.0036

Note: Better estimated wave numbers are in bold.

## SYNTHETIC DATA EXAMPLE

We simulate four linear seismic events by using an Ormsby wavelet, which is a common type of synthetic wavelet in reflection seismology. The wavelet is defined by the sinc function as

$$A(t) = \frac{(\pi f_4)^2 \text{sinc}^2(\pi f_4 t) - (\pi f_3)^2 \text{sinc}^2(\pi f_3 t)}{\pi f_4 - \pi f_3} - \frac{(\pi f_2)^2 \text{sinc}^2(\pi f_2 t) - (\pi f_1)^2 \text{sinc}^2(\pi f_1 t)}{\pi f_2 - \pi f_1}. \quad (12)$$

We choose  $f_1 = 5$ ,  $f_2 = 10$ ,  $f_3 = 20$ , and  $f_4 = 30$ , and so the wavelet defines a trapezoidal shape in the frequency spectrum with corner frequencies  $5\pi$ ,  $10\pi$ ,  $20\pi$ , and  $30\pi$  Hz. The linear events are designed to have different amplitudes and dips. There are 100 traces that are regularly spaced, and we randomly remove 40 traces (see Figure 3). The sum of the L2 norm of all traces in Figure 3a is 650.86. The time sampling rate is 1000 samples per second, and assuming that the unit distance is in kilometers, the distance sampling rate is 100 samples per kilometer (10 m trace spacing). Therefore, the Nyquist frequency is 500 Hz, and the Nyquist wavenumber is 50 cycles per kilometer (c/k) and 40 traces are removed. The frequency-wavenumber ( $f$ - $k$ ) spectrums of Figure 3a and 3c are shown in Figure 3b and 3d, respectively. It can be observed that the frequencies beyond 70 Hz (and  $-70$  Hz) are aliased for two of the linear events.

We now apply the ALFT, ASFT, and ALLSSA methods to the edited data shown in Figure 3b, and we regularize the temporal frequency slices. For a fair comparison, we apply the same threshold (0.05) for the ALFT, ASFT, and ALLSSA methods. We also amplify the difference between the original synthetic data and the regularized data (the error) by a factor of 30, and so the  $f$ - $k$  spectrums of the errors will correspond to the amplified errors for all the methods. The same partitioning for wavenumbers is used for the ASFT and ALLSSA methods to estimate the wavenumbers up to four decimal

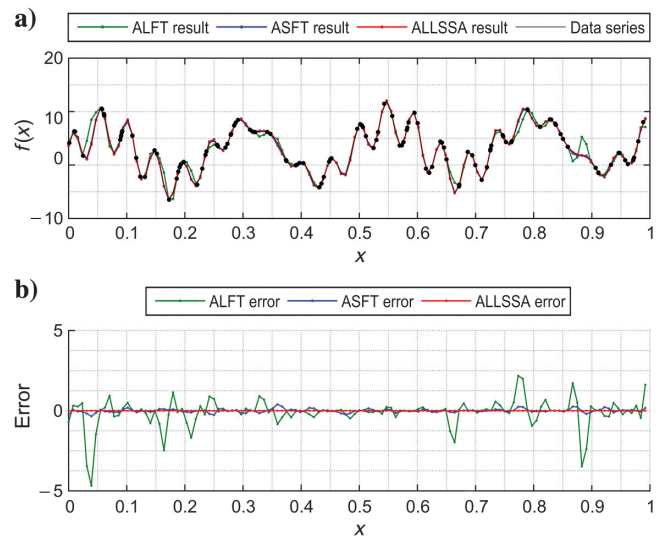


Figure 2. (a) The irregularly sampled data series (black) given in equation 11 and its regularizations using the ALFT (green), ASFT (blue), and ALLSSA (red) methods, and (b) the difference between the actual data series values and the regularized results.

places. All the methods are programmed in MATLAB language, and we run them on a computer with two cores (processors). We also denote the sum of L2 norm of differences between the original and interpolated traces by “ $\epsilon$ .”

First, we choose  $\mathbf{K} = \{-50, -49, \dots, 48, 49\}$  for the ALFT method and show its regularization result and error in Figure 4a and 4c, respectively. The computational time is 7 s, and  $\epsilon = 37.29$ . It can be seen that the missing traces are not very well constructed especially for the extrapolation (the traces toward the edges). This is mainly due to the lack of accuracy of the wavenumbers. We next double the number of wavenumbers, i.e.,  $\mathbf{K} = \{-50, -49.5, -49, \dots, 48.5, 49, 49.5\}$ , and we show the ALFT regularization

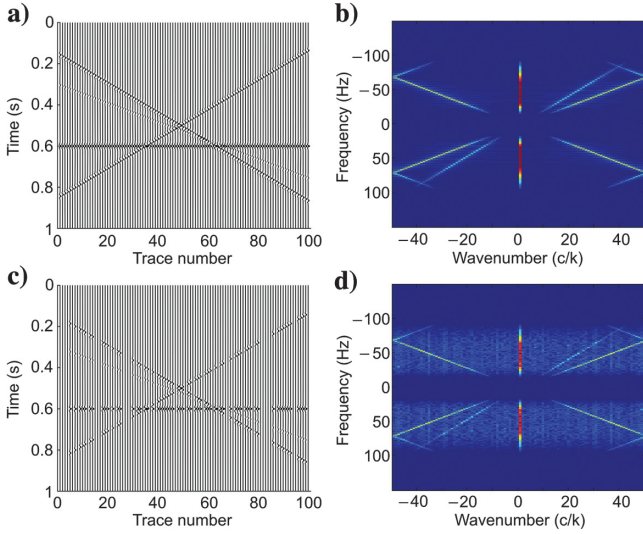


Figure 3. A synthetic seismic data containing four linear events. (a) Original data with 100 traces, (b) the  $f$ - $k$  spectrum of panel a, (c) 40 traces are randomly removed, and (d) the  $f$ - $k$  spectrum of (c).

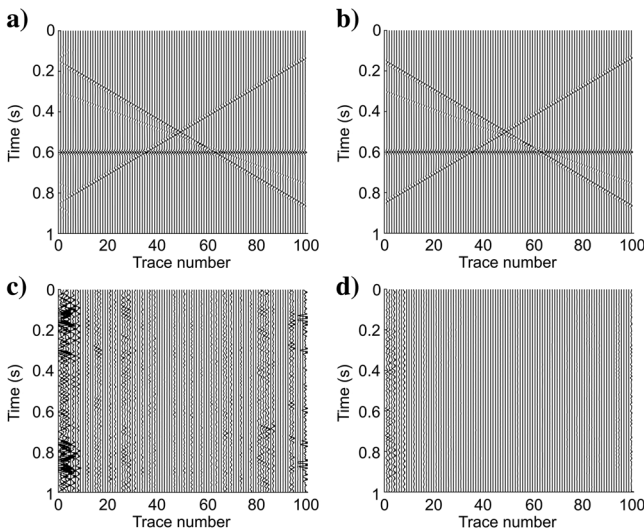


Figure 4. The regularization results using the ALFT method with (a) 100 wavenumbers, (b) a denser set of wavenumbers (200 wavenumbers), and (c and d) the error corresponding to (a and b), respectively (amplified 30 times).

result and error in Figure 4b and 4d, respectively. The computational time increased to 14 s, but  $\epsilon = 6.76$ . We observe that the traces are constructed more accurately than the previous choice for  $\mathbf{K}$ . We also illustrate the  $f$ - $k$  spectrums corresponding to Figure 4 in Figure 5.

Next, we choose the preselected set of wavenumbers in the ASFT method as  $\mathbf{K} = \{-50, -49, \dots, 48, 49\}$ , and we illustrate its regularization result and error in Figure 6a and 6c, respectively. The computational time is 13 s, and  $\epsilon = 2.6$ , which is more accurate and computationally faster than the ALFT method with the denser set of wavenumbers.

For the ALLSSA method, we choose  $\mathbf{\Omega} = \{1, 2, \dots, 49\}$ . To ensure that all the real wavenumbers are considered up to the Nyquist wavenumber 50, we allow the partitioning for wavenumber 49 to

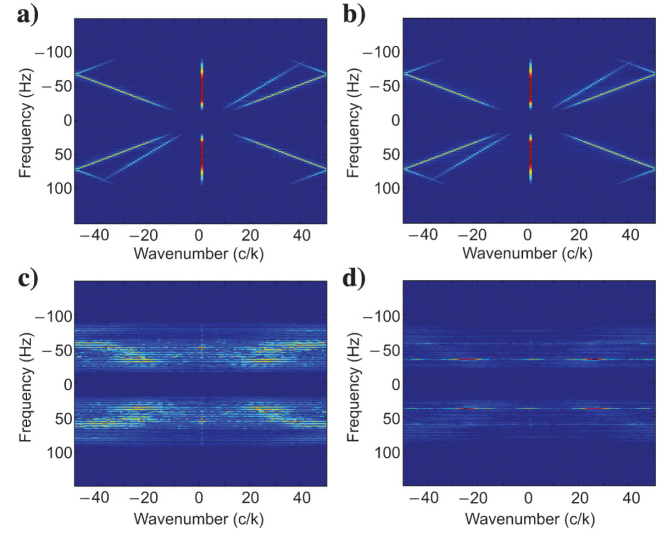


Figure 5. The  $f$ - $k$  spectrums corresponding to (a) Figure 4a, (b) Figure 4b, (c) Figure 4c, and (d) Figure 4d.

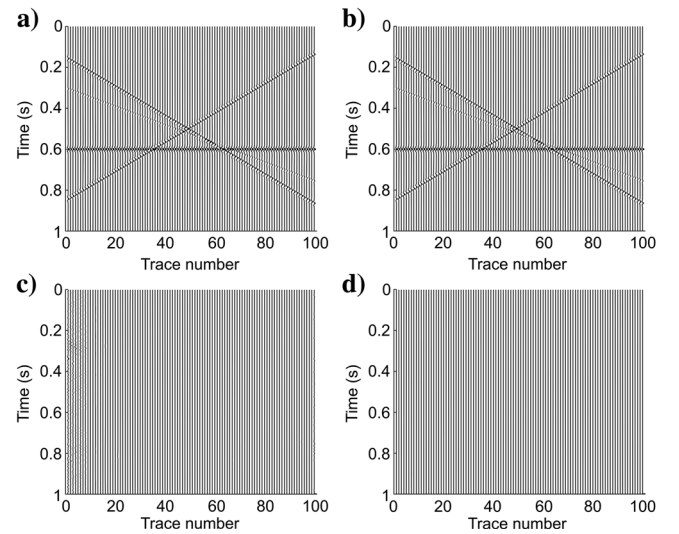


Figure 6. The regularization results using the (a) ASFT method, (b) ALLSSA method, and (c and d) the error corresponding to (a and b), respectively (amplified 30 times).



cover the range [49, 50], and we choose  $\Phi = [\mathbf{1}, \mathbf{x}]$  to explicitly account for any linear trend. If we choose  $\Phi = [\mathbf{1}]$ , we may also allow the partitioning for wavenumber 1 to cover the range (0,1], so the sinusoids of lower wavenumbers can approximate the trends. We show the ALLSSA regularization result and error in Figure 6b and 6d, respectively. The computational time is 3 s, and  $\varepsilon = 0.3$ . It can be seen that the ALLSSA method is able to fill in the missing traces much more accurately and computationally faster than the ALFT and ASFT methods. We show the  $f$ - $k$  spectrums corresponding to Figure 6 in Figure 7.

To show how the ALLSSA method statistically attenuates the random noise, we add a distinct Gaussian random noise to each trace in Figure 3b. Therefore, approximately 20% of the synthetic data is random noise. We illustrate only trace 1 for the extrapolation,

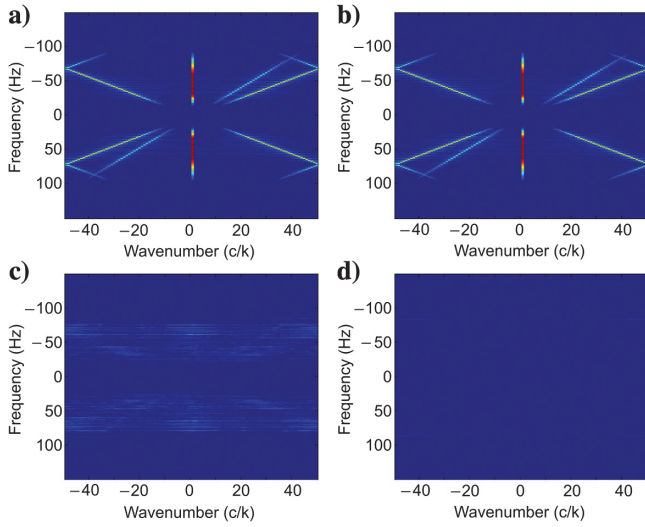


Figure 7. The  $f$ - $k$  spectrums corresponding to (a) Figure 6a, (b) Figure 6b, (c) Figure 6c, and (d) Figure 6d.

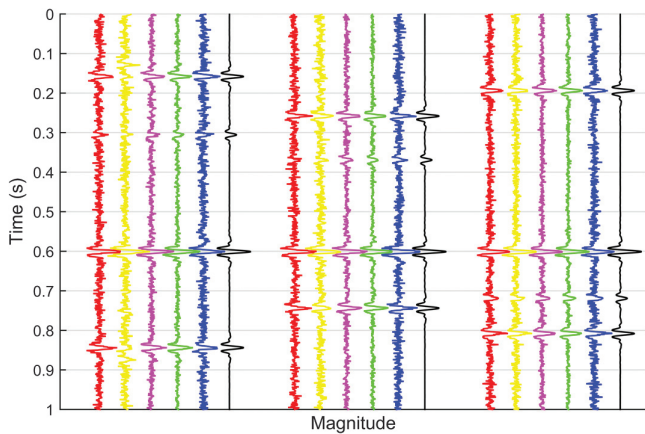


Figure 8. The comparison of three regularized traces using the ALFT, ASFT, and ALLSSA methods. Original traces without random noise in black, with random noise in blue, the ALLSSA results at 99% and 95% confidence levels in green and pink, respectively, the ALFT result in yellow and the ASFT result in red. The left six traces are for trace 1, the middle ones are for trace 15, and the right ones are for trace 92.

trace 15 for the interpolation, and trace 92 for noise attenuation purposes in Figure 8. The set  $\mathbf{K} = \{-50, -49, \dots, 48, 49\}$  is selected for the ALFT and ASFT methods. Comparing the ALLSSA method with the ALFT and ASFT methods, one can observe that the ALLSSA method performs much better in reconstructing the signals (the events) and attenuates the random noise down to a certain confidence level. The result of the ALFT and ASFT methods for traces 1 and 15 looks noisier because aside from the accuracy of the actual wavenumbers, more noise is interpolated in the traces. In the ALLSSA method, when the 99% confidence level is used, the noise attenuated slightly more than when the 95% level is used; however, we recommend using the 95% confidence level for real seismic data because of their nonstationarity behavior. Note that there is no rigorous statistical property of choosing a confidence level for the ALFT and ASFT methods, so a threshold was selected based on the amount of random noise for these methods.

To see the sensitivity of the ALLSSA method to spatially aliased events, we remove the traces of odd numbers except a few of them (e.g., trace numbers 5, 21, 83, 91) and show the result in Figure 9a. We select a dense set of wavenumbers  $\mathbf{K} = \{-50, -49.9, \dots, 49.8, 49.9\}$  (1000 wavenumbers) for the ALFT method. In this case, the ALFT and ASFT methods have approximately the same computational time (57 s) with  $\varepsilon = 109.3$  and  $\varepsilon = 93.4$ , respectively. We only illustrate the ALFT result because the ASFT result is also very similar (see Figure 9). Note that the error is amplified by a

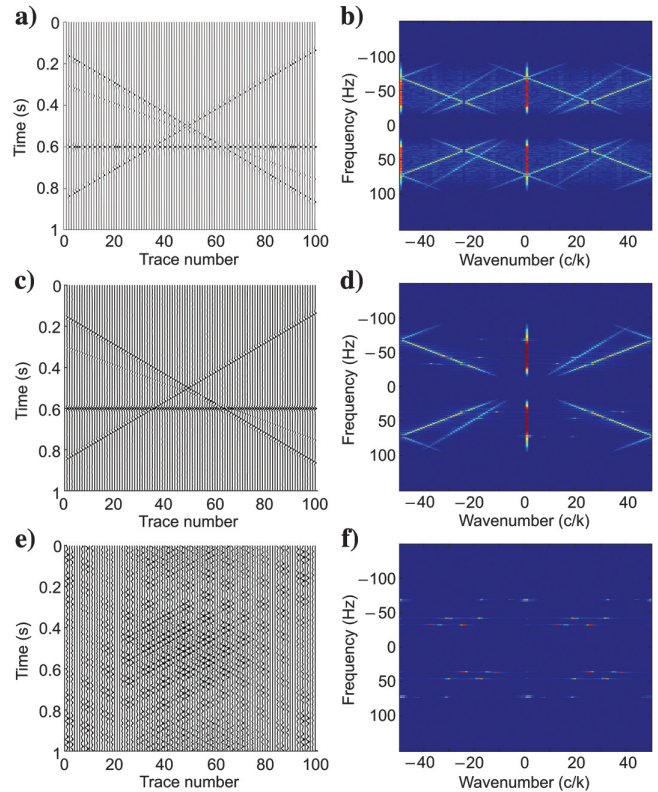


Figure 9. (a) The synthetic data in Figure 3a after setting the traces of odd numbers to zero except trace numbers 5, 21, 83, 91, (b) the  $f$ - $k$  spectrum of (a), (c) the regularization result using the ALFT method with a denser set of wavenumbers (1000 wavenumbers), (d) the  $f$ - $k$  spectrum of (c), (e) the error (amplified 10 times), and (f) the  $f$ - $k$  spectrum of (e).

factor of 10 in Figure 9e. To show the aliasing, we illustrate the  $f$ - $k$  spectrum of this synthetic data in Figure 9b. Because the dominant trace spacing is 20 m, all of the events are spatially aliased. We also show the  $f$ - $k$  spectrums of the ALFT regularization result and its amplified error in Figure 9d and 9f, respectively. One can clearly observe the effect of the alias events in the spectrums. The computational time for the ALLSSA method is 2 s, and  $\varepsilon = 0.9$  which is significantly smaller than the ALFT and ASFT methods. The final regularization result is almost the same as Figure 6b (not shown here). The reason for this accuracy is that the ALLSSA method very accurately estimates the positive wavenumbers, and these few extra traces (four traces) prevent the matrix of normal equations to be singular. In the event that this matrix is singular, one may compute its pseudo-inverse (Rao and Mitra, 1972). Note that if we remove the traces of odd numbers, the Nyquist wavenumber will be 25 c/k and the aliasing will occur for the wavenumbers beyond 25 c/k. In this situation, an antialias technique similar to the technique in Schonewille et al. (2009) may be used.

## FIELD DATA APPLICATIONS

The field data example is a marine 2D shot gather from the deep-water of Gulf of Mexico that is the same as the one used by Crawley (2000), Fomel (2002), and Chen et al. (2015). The time sampling rate is 250 samples per second. To compare our interpolation approach with the nonlinear shaping regularization method, we remove the same traces (30%) as in Chen et al. (2015) (see Figure 10). Note that the normalized wavenumber axis in the  $f$ - $k$  spectrums is obtained by considering  $\Delta x = 1$ . We illustrate the nonlinear shaping regularization result after 15 iterations and its error in Figure 11a and 11c, respectively.

Because the stationarity of data series is the main assumption of the ALFT, ASFT, and ALLSSA methods (i.e., the seismic events must be approximately linear), we use the windowing strategy. Therefore, we divide the seismic data to six spatial windows and apply the same threshold for the ALFT and ALLSSA methods. Within each window, we choose  $\mathbf{K} = \{-15, -14.9, \dots, 14.8, 14.9\}$  (300 wave-

numbers) for the ALFT method, and  $\mathbf{\Omega} = \{1, \dots, 14\}$  and initial  $\mathbf{\Phi} = [\mathbf{1}]$  for the ALLSSA method. We illustrate the regularization result of the ALFT method and its error in Figure 11b and 11d. The  $f$ - $k$  spectrums corresponding to the nonlinear shaping regularization and ALFT methods are shown in Figure 12. The regularization result using the ALLSSA method and its error are shown in Figure 13a and 13c, respectively. Figure 14a and 14c shows the  $f$ - $k$  spectrums corresponding to Figure 13a and 13c, respectively.

It can be seen that the ALFT and ALLSSA methods performed well in interpolating the missing traces compare with the nonlinear shaping regularization method; however, they still have some small amount of artifacts (compare the magnified sections, for example). The data are very poorly sampled along the spatial direction within some of the windows (e.g., half of the traces are missing within the magnified section), and so the spatial aliasing mainly caused these

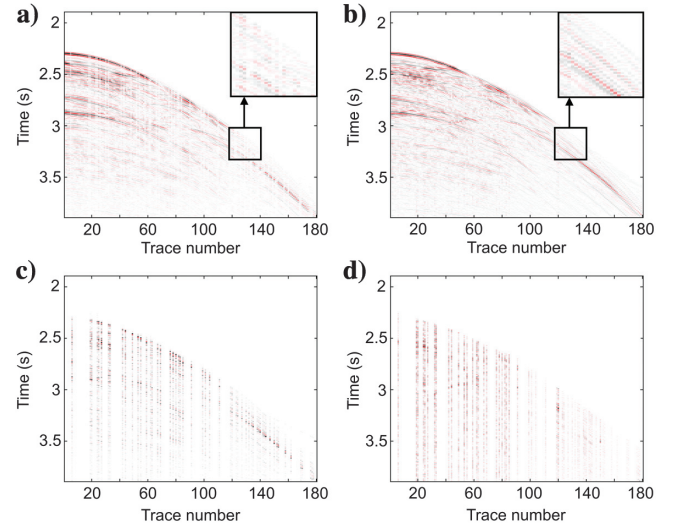


Figure 11. (a) The result of the nonlinear shaping regularization after 15 iterations, (b) the ALFT result, (c) the error of (a), and (d) the error of (b).

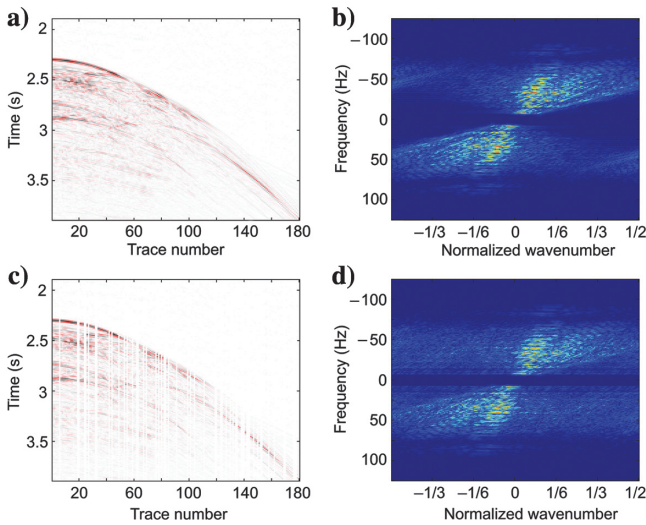


Figure 10. (a) A marine 2D shot gather from the deepwater Gulf of Mexico, (b) the  $f$ - $k$  spectrum of (a), (c) 30% of the traces are randomly removed, and (d) the  $f$ - $k$  spectrum of (c).

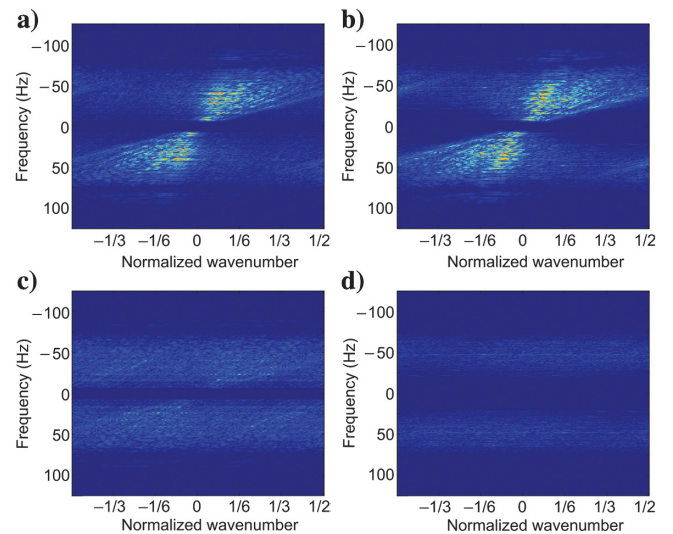


Figure 12. The  $f$ - $k$  spectrums corresponding to (a) Figure 11a, (b) Figure 11b, (c) Figure 11c, and (d) Figure 11d.



artifacts. An antialiasing scheme within the windows may further improve these results (Schonewille et al., 2009; Xu et al., 2010). We now show an alternative approach for a better interpolation result.

In the weighted ALLSSA method, the larger that the values of the diagonal entries  $\mathbf{P}$  are, the more the fitting process is focused around the locations of those values and vice versa. For a faster computation, we may treat the weight matrix in equation 10 as a vector  $\mathbf{P} = [w(\ell)]$  ( $1 \leq \ell \leq n$ ) whose entries are the values of the following Gaussian function:

$$w(\ell) = e^{-\lambda(x_\ell - x_i)^2}, \quad (13)$$

where  $\lambda$  determines the effective width of the Gaussian function (length of the window) and  $x_i$  is the location of a trace being approximated. Therefore, we may apply the weighted ALLSSA method to estimate each trace individually using the data within the spatial window containing that trace, which may not necessarily fall in the center of the window. For existing traces, we may still take advantage of the computational speed of the ALLSSA method and apply it to attenuate the random noise of the traces down to a certain confidence level and so smooth the entire seismic data. By this selection of  $\mathbf{P}$ , the sinusoidal basis functions will be adapted to the Morlet wavelet in the least-squares sense, smoothing the seismic image (Foster, 1996; Ghaderpour and Pagiatakis, 2017). This selection also allows us to incorporate more wavenumbers into the algorithm resulting in a better interpolation result (Crawley, 2000). A similar discussion for selection of the weight function in the ALFT method is also given in Xu et al. (2010).

For a better comparison between the ALLSSA and the weighted ALLSSA methods, we select  $\lambda$  in equation 13 in such away that the Gaussian bell curve is almost truncated to zero in the margins of the translating windows whose lengths are the same as the window length used above. We show the result of the weighted ALLSSA method and its error in Figure 13b and 13d, respectively. The  $f$ - $k$  spectrums corresponding to Figure 13b and 13d are illustrated in

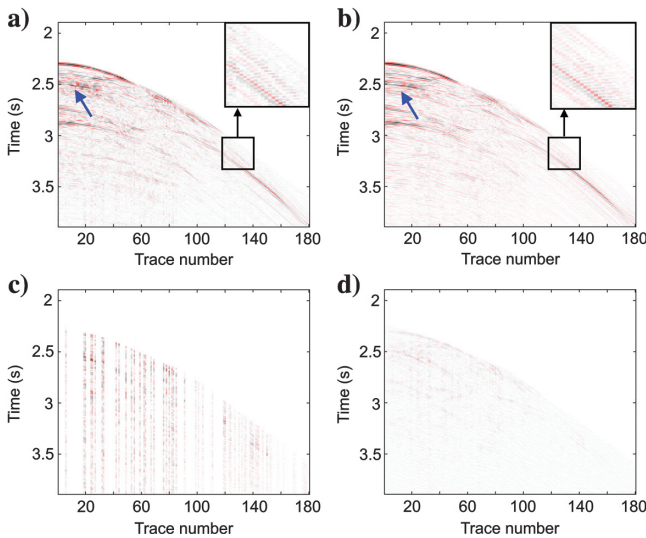


Figure 13. (a) The ALLSSA result, (b) the weighted ALLSSA result, (c) the error of (a), and (d) the error of (b). The events in (b) are constructed more smoothly than in (a) (e.g., the blue arrows and magnified sections).

Figure 14b and 14d, respectively. One can observe that the events are smoothly constructed, and the random noise is also attenuated down to 95% confidence level (e.g., the blue arrows in Figure 13). The almost linear event in the southeast part of the seismic data is clearly constructed (compare the events within the magnified sections shown in Figure 13a and 13b and the corresponding  $f$ - $k$  spectrums shown in Figure 14a and 14b, white arrows). All the methods show small artifacts above the water-bottom reflections that are not significant. We used a mask to set the values of the northeast part of the images shown in Figures 11 and 13 to zero after the interpolation.

We emphasize that the traces within each of the six spatial windows are simultaneously estimated in the ALFT and ALLSSA

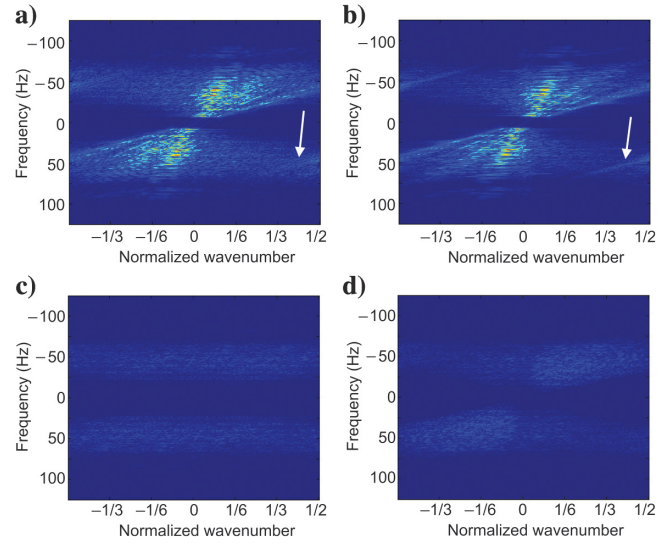


Figure 14. The  $f$ - $k$  spectrums corresponding to (a) Figure 13a, (b) Figure 13b, (c) Figure 13c, and (d) Figure 13d. The white arrows in (a and b) refer to the linear event shown in the magnified sections in Figure 13a and 13b, respectively.

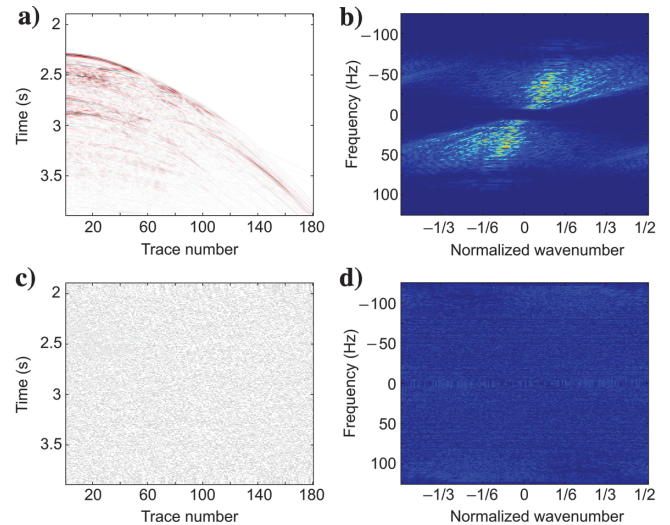


Figure 15. (a) The noise attenuated result of Figure 10a using the ALLSSA method, (b) the  $f$ - $k$  spectrum of (a), (c) the error (amplified by a factor of 20), and (d) the  $f$ - $k$  spectrum of (c).

methods, and so the existing traces are almost fully constructed, resulting in error propagation to other traces (Figures 11b and 13a). However, in the weighted ALLSSA method, each trace is individually approximated by applying a 95% confidence level, and so the events are smoothly constructed and random noise is also attenuated down to a 95% confidence level for the entire seismic data.

Finally, for each temporal frequency slice of the original seismic data (regularly spaced), one may repeat the ALLSSA method until the L2 norm of the residual series becomes less than approximately 10% of the total L2 norm of the data series and use the inverse FFT to reconstruct the traces. We show the result of this process, its difference from the original seismic data (amplified by a factor of 20), and their corresponding  $f$ - $k$  spectrums in Figure 15. One can observe that the geologic structures are clearly reconstructed, and the random noise is attenuated (a 95% confidence level is used in each iteration).

## CONCLUSIONS

The ALLSSA method is a fast and very accurate new method of regularizing stationary data series that can also attenuate the random noise down to a certain confidence level. This method takes into account the correlations among the sinusoidal basis functions as well as the constituents of known forms by considering the covariance matrix associated to the data series. We demonstrated how an appropriate selection of the weight matrix (the inverse of the covariance matrix) adapts the method for regularizing nonstationary data series.

Similar to the ALFT and ASFT methods, the ALLSSA method can be used for seismic data regularization. Because the ALLSSA method estimates the wavenumbers of the data series more accurately than do the ALFT and ASFT methods up to a certain confidence level, it is computationally faster and more accurate than the ALFT and ASFT methods for regularization. The numerical data examples in this contribution show promising results. Applications of the ALLSSA method are not restricted to seismology, and it may show its potential performance in other fields such as geodesy, astronomy, medical sciences, and finance.

Taking into account that the correlation among the sinusoids is the key point in the ALLSSA method. This correlation will be considered when the matrix of normal equations is being inverted. In practice, this matrix is often regular because the traces are usually inherently unequally spaced. However, one may calculate the pseudoinverse of this matrix when it is singular. For seismic data with very complex geologic structures that have severe aliasing, the ALLSSA method may not be reliable for regularization because this method is established mainly for the wide-sense stationary seismic data similar to the ALFT, ASFT, and IMAP methods. In future work, we extend the method to the higher dimension cases and show its capability in regularization. It is expected that the generalization of the ALLSSA method to higher dimensions has better performance in regularization.

## ACKNOWLEDGMENTS

The authors would like to thank the editors, N. Gülünay and I. Moore, and the anonymous reviewers for their very useful and constructive comments that significantly improved the presentation of the paper. We would also like to thank Y. Chen for providing the

seismic data and S. D. Pagiatakis and D. Trad for useful discussion. This research has been financially supported by the Pacific Institute for the Mathematical Sciences, the Natural Sciences and Engineering Research Council of Canada, and the postdoctoral program at the Department of Mathematics and Statistics, University of Calgary.

## APPENDIX A

### LSS WHEN SPATIAL SAMPLING IS UNIFORM

The discrete Fourier transform (DFT) of a sequence of  $n$  regularly spaced complex numbers  $f(x_1), \dots, f(x_n)$  into another sequence of complex numbers (Fourier coefficients) is defined as

$$\hat{f}(\omega_k) = \frac{1}{n} \sum_{\ell=1}^n f(x_\ell) e^{-2\pi i \omega_k x_\ell}, \quad (\text{A-1})$$

where  $x_\ell = \ell/n$ ,  $\omega_k \in \mathbf{K} = \{(-n/2), \dots, (n/2) - 1\}$  if  $n$  is even, and  $\omega_k \in \mathbf{K} = \{-(n-1)/2, \dots, (n-1)/2\}$  if  $n$  is odd. We refer to  $\omega_k$  as the wavenumber that is the number of cycles per unit distance. Considering all the wavenumbers simultaneously, the DFT can be represented in matrix form as  $\mathbf{f} = (1/n)\mathbf{D}^T \mathbf{f}$ , where

$$\mathbf{F} = \begin{bmatrix} \hat{f}(\omega_1) \\ \vdots \\ \hat{f}(\omega_n) \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} e^{2\pi i \omega_1 x_1} & e^{2\pi i \omega_2 x_1} & \dots & e^{2\pi i \omega_n x_1} \\ \vdots & \vdots & \dots & \vdots \\ e^{2\pi i \omega_1 x_n} & e^{2\pi i \omega_2 x_n} & \dots & e^{2\pi i \omega_n x_n} \end{bmatrix},$$

$$\mathbf{f} = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}, \quad (\text{A-2})$$

and  $\mathbf{T}$  denotes the conjugate transpose (Craymer, 1998). It can be seen that  $\mathbf{f} = \mathbf{D}\mathbf{F}$  (when traces are regularly sampled in space). The Euler's formula implies that  $e^{2\pi i \omega_k x_\ell} = \cos(2\pi \omega_k x_\ell) + i \sin(2\pi \omega_k x_\ell)$ , and so the LSSA method expresses the columns of  $\mathbf{D}$  as cosine and sine basis functions. More precisely, let  $\Omega = \{1, \dots, \eta\}$ , where  $\eta = (n/2) - 1$  if  $n$  is even ( $\eta = n/2$  is the Nyquist wavenumber), and  $\eta = (n-1)/2$  if  $n$  is odd. When the real-valued version of the normal equations considered, the least-squares minimization uses the model  $\mathbf{f} = \Phi \mathbf{c}$  to estimate  $\mathbf{c}$  as  $\hat{\mathbf{c}} = \mathbf{N}^{-1} \Phi^T \mathbf{f}$ , where  $\mathbf{N} = \Phi^T \Phi$ ,

$$\Phi = \begin{bmatrix} 1 & \cos(2\pi \omega_1 x_1) & \sin(2\pi \omega_1 x_1) & \dots & \cos(2\pi \omega_\eta x_1) & \sin(2\pi \omega_\eta x_1) \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 1 & \cos(2\pi \omega_1 x_n) & \sin(2\pi \omega_1 x_n) & \dots & \cos(2\pi \omega_\eta x_n) & \sin(2\pi \omega_\eta x_n) \end{bmatrix}, \quad (\text{A-3})$$

and  $\omega_k \in \Omega$ . The dimensions of  $\Phi$ ,  $\mathbf{N}$ , and  $\hat{\mathbf{c}}$  are  $n \times (2\eta + 1)$ ,  $(2\eta + 1) \times (2\eta + 1)$ , and  $(2\eta + 1) \times 1$ , respectively.

Note that the first column of design matrix  $\Phi$  above corresponds to the zero wavenumber. In fact, we could have two column vectors associated to the zero wavenumber: the first column for cosine whose elements are  $\cos(0) = 1$  and the second column for sine whose elements are  $\sin(0) = 0$ . However, we avoid writing the zero



column vector because it will make matrix  $\mathbf{N}$  singular. We emphasize that matrix  $\mathbf{N}$  is a real square matrix, and it is regular as long as the column vectors of  $\Phi$  are linearly independent. We may also include the column vector of  $\cos(2\pi(n/2)x_\ell) = \cos(2\pi(n/2)(\ell/n)) = \cos(\ell\pi) = \pm 1$ ,  $\ell = 1, \dots, n$ , in matrix  $\Phi$  above when  $n$  is even. This way  $\mathbf{f}$  will be reconstructed accurately if  $\mathbf{f}$  contains constituents of the Nyquist wavenumber  $n/2$ . Note that we do not include the column vector of the sine because  $\sin(2\pi(n/2)x_\ell) = \sin(2\pi(n/2)(\ell/n)) = \sin(\ell\pi) = 0$  for  $\ell = 1, \dots, n$ . To avoid the singularity of square matrix  $\mathbf{N}$ , the negative wavenumbers are excluded from matrix  $\Phi$ . Craymer (1998) discusses that when  $\mathbf{f}$  is real, we may only consider the positive wavenumbers. However, not much work is done on the LSSA method for the case in which  $\mathbf{f}$  is complex. Our analyses show that if  $\mathbf{f}$  is complex, we may still consider only the positive wavenumbers for regularization (see below for further details).

Assume that  $n$  is even (for  $n$  odd, we have a similar discussion). When a data series is regularly spaced, we have

$$\sum_{\ell=1}^n \cos(2\pi\omega_k x_\ell) \cos(2\pi\omega_u x_\ell) = \begin{cases} n & \text{for } \omega_k = \omega_u = -\frac{n}{2} \text{ or } 0 \\ \frac{n}{2} & \text{for } \omega_k = \omega_u \neq -\frac{n}{2} \text{ or } 0, \\ 0 & \text{for } \omega_k \neq \omega_u \end{cases} \quad (\text{A-4})$$

$$\sum_{\ell=1}^n \sin(2\pi\omega_k x_\ell) \sin(2\pi\omega_u x_\ell) = \begin{cases} 0 & \text{for } \omega_k = \omega_u = -\frac{n}{2} \text{ or } 0 \\ \frac{n}{2} & \text{for } \omega_k = \omega_u \neq -\frac{n}{2} \text{ or } 0, \\ 0 & \text{for } \omega_k \neq \omega_u \end{cases} \quad (\text{A-5})$$

$$\sum_{\ell=1}^n \cos(2\pi\omega_k x_\ell) \sin(2\pi\omega_u x_\ell) = 0 \text{ for all } \omega \in \mathbf{K}. \quad (\text{A-6})$$

Therefore, including only the positive wavenumbers in the LSSA method also implicitly accounts for the identical response to positive and negative wavenumbers and produces a one-sided spectrum that is physically meaningful. This will also effectively double the amplitude of the least-squares transform with respect to the DFT (Vaníček, 1969; Craymer, 1998). More precisely, it can be seen that

$$\hat{\mathbf{c}} = \mathbf{N}^{-1} \Phi^T \mathbf{f} = \begin{bmatrix} \frac{1}{n} \sum_{\ell=1}^n f(x_\ell) \\ \frac{2}{n} \sum_{\ell=1}^n \cos(2\pi\omega_1 x_\ell) f(x_\ell) \\ \frac{2}{n} \sum_{\ell=1}^n \sin(2\pi\omega_1 x_\ell) f(x_\ell) \\ \vdots \\ \frac{2}{n} \sum_{\ell=1}^n \cos(2\pi\omega_\eta x_\ell) f(x_\ell) \\ \frac{2}{n} \sum_{\ell=1}^n \sin(2\pi\omega_\eta x_\ell) f(x_\ell) \end{bmatrix}, \quad \Phi \hat{\mathbf{c}} \approx \mathbf{f}. \quad (\text{A-7})$$

The elements of column vector  $\hat{\mathbf{c}}$  in equation A-7 from top to bottom are the estimated coefficients for the columns of  $\Phi$  in equation A-3 from left to right, respectively. Note that Vaníček (1969), Craymer (1998), and Pagiatakis (1999) discuss the LSSA method mainly for a real time series (i.e.,  $\mathbf{f}$  is a real vector) with its applications in geodesy and geodynamics. However, because design

matrix  $\Phi$  is a real matrix,  $\mathbf{f}$  can also be a complex vector (e.g., a temporal frequency slice). In this case,  $\hat{\mathbf{c}}$  in equation A-7 will be a complex vector whose real and imaginary parts contain the estimated amplitudes corresponding to the real and imaginary parts of  $\mathbf{f}$ , respectively. One may also use equation A-7 for the real part and imaginary part of  $\mathbf{f}$  individually (treating each part as a real data series) and then form vector  $\hat{\mathbf{c}}$  by combining the estimated amplitudes of each part, producing similar regularization results to the previous case.

For an irregularly sampled data series, equations A-4–A-6 are no longer valid; i.e., the sinusoidal basis functions are no longer orthogonal. However, the correlation between them is taken into account when matrix  $\mathbf{N}$  is being inverted. The selection of wavenumbers is a critical task for the simultaneous estimation of the LSS. The correlation among the sinusoids must be taken into account carefully; otherwise, ill-conditioning in  $\mathbf{N}$  can produce incorrect results (Craymer, 1998). In the ALLSSA method, the wavenumbers are accurately estimated in an iterative manner, producing an accurate one-sided spectrum for stationary data series. We emphasize that in the ALLSSA method, we only estimate the coefficients of significant components of a data series of size  $n$ , and so the number of columns in equation A-3 may be much smaller than  $n$ . Also, note that we are not using the fast methods of the FFT for the ALFT, ASFT, and ALLSSA methods.

## APPENDIX B

### OPTIMIZATION METHODS

Estimating  $\mathbf{c}_k$  by minimizing the cost function  $\Psi_2(\mathbf{c}, \mathbf{c}_k)$  is in fact solving the overdetermined system  $\mathbf{f} = \Phi \mathbf{c} = \Phi_k \mathbf{c}_k$  by using the least-squares method (Craymer, 1998). Therefore,  $\bar{\mathbf{c}} = \bar{\mathbf{N}}^{-1} \Phi^T \mathbf{P} \mathbf{f}$ , where  $\bar{\mathbf{N}} = \Phi^T \mathbf{P} \Phi$ . Calculation of the inverse of  $\bar{\mathbf{N}}$  for obtaining the spectrum of a data series is computationally tedious especially when the number of constituents of known forms (the column dimension of  $\Phi$ ) increases. Therefore, we may use a sequential technique similar to Craymer (1998) to obtain equation 9. We expand  $\bar{\mathbf{N}}^{-1}$  as

$$\bar{\mathbf{N}}^{-1} = \begin{bmatrix} \Phi^T \mathbf{P} \Phi & \Phi^T \mathbf{P} \Phi_k \\ \Phi_k^T \mathbf{P} \Phi & \Phi_k^T \mathbf{P} \Phi_k \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix}. \quad (\text{B-1})$$

It can be seen that

$$\mathbf{M}_1 = (\Phi^T \mathbf{P} \Phi - \Phi^T \mathbf{P} \Phi_k (\Phi_k^T \mathbf{P} \Phi_k)^{-1} \Phi_k^T \mathbf{P} \Phi)^{-1}, \quad (\text{B-2})$$

$$\mathbf{M}_2 = -\mathbf{M}_1 \Phi^T \mathbf{P} \Phi_k (\Phi_k^T \mathbf{P} \Phi_k)^{-1}, \quad (\text{B-3})$$

$$\mathbf{M}_3 = -\mathbf{M}_4 \Phi_k^T \mathbf{P} \Phi \mathbf{N}^{-1}, \quad (\text{B-4})$$

$$\mathbf{M}_4 = (\Phi_k^T \mathbf{P} \Phi_k - \Phi_k^T \mathbf{P} \Phi \mathbf{N}^{-1} \Phi^T \mathbf{P} \Phi_k)^{-1}, \quad (\text{B-5})$$

where  $\mathbf{N} = \Phi^T \mathbf{P} \Phi$ . Thus, from

$$\begin{bmatrix} \mathbf{x} \\ \hat{\mathbf{c}}_k \end{bmatrix} = \bar{\mathbf{c}} = \bar{\mathbf{N}}^{-1} \bar{\mathbf{\Phi}}^T \mathbf{P} \mathbf{f} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^T \mathbf{P} \mathbf{f} \\ \mathbf{\Phi}_k^T \mathbf{P} \mathbf{f} \end{bmatrix}, \quad (\text{B-6})$$

we obtain

$$\hat{\mathbf{c}}_k = \mathbf{M}_3 \mathbf{\Phi}^T \mathbf{P} \mathbf{f} + \mathbf{M}_4 \mathbf{\Phi}_k^T \mathbf{P} \mathbf{f}. \quad (\text{B-7})$$

Substituting equation B-4 into equation B-7 and simplifying, we obtain

$$\hat{\mathbf{c}}_k = \mathbf{M}_4 \mathbf{\Phi}_k^T \mathbf{P} (\mathbf{f} - \mathbf{\Phi} \bar{\mathbf{N}}^{-1} \mathbf{\Phi}^T \mathbf{P} \mathbf{f}), \quad (\text{B-8})$$

that is equation 9, noting that the term within the parentheses is  $\mathbf{g}$ . Computing the spectrum of a data series using equation B-8 is much more efficient because  $\mathbf{M}_4$  is a square matrix of order 2 for all  $k$ s.

## REFERENCES

- Abma, R., and J. Claerbout, 1995, Lateral prediction for noise attenuation by t-x and f-x techniques: *Geophysics*, **60**, 1887–1896, doi: [10.1190/1.1443920](https://doi.org/10.1190/1.1443920).
- Abma, R., and N. Kabir, 2006, 3D interpolation of irregular data with a POCs algorithm: *Geophysics*, **71**, no. 6, E91–E97, doi: [10.1190/1.2356088](https://doi.org/10.1190/1.2356088).
- Chen, S., and B. L. Luk, 1999, Adaptive simulated annealing for optimization in signal processing applications: *Signal Processing*, **79**, 117–128, doi: [10.1016/S0165-1684\(99\)00084-5](https://doi.org/10.1016/S0165-1684(99)00084-5).
- Chen, Y., L. Zhang, and L. Mo, 2015, Seismic data interpolation using non-linear shaping regularization: *Journal of Seismic Exploration*, **24**, 327–342.
- Craig, A. T., R. V. Hogg, and J. McKean, 2013, *Introduction to mathematical statistics*, 7th ed.: Pearson Education Inc.
- Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. thesis, Stanford University.
- Craymer, M. R., 1998, The least-squares spectrum, its inverse transform and autocorrelation function: Theory and some application in geodesy: Ph.D. thesis, University of Toronto.
- Eriksson, J., E. Ollila, and V. Koivunen, 2009, Statistics for complex random variables revisited: *IEEE International Conference on Acoustics, Speech and Signal Processing*, doi: [10.1109/ICASSP.2009.4960396](https://doi.org/10.1109/ICASSP.2009.4960396).
- Fomel, S., 2002, Applications of planewave destruction filters: *Geophysics*, **67**, 1946–1960, doi: [10.1190/1.1527095](https://doi.org/10.1190/1.1527095).
- Fomel, S., 2007, Shaping regularization in geophysical-estimation problems: *Geophysics*, **72**, no. 2, R29–R36, doi: [10.1190/1.2433716](https://doi.org/10.1190/1.2433716).
- Foster, G., 1996, Wavelet for period analysis of unevenly sampled time series: *The Astronomical Journal*, **112**, 1709–1729, doi: [10.1086/118137](https://doi.org/10.1086/118137).
- Gao, J. J., A. Stanton, M. Naghizadeh, M. D. Sacchi, and X. Chen, 2013, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: *Applied Geophysics*, **61**, 138–151.
- Ge, Z., J. Li, S. Pan, and X. Chen, 2015, A fast-convergence POCs seismic denoising and reconstruction method: *Applied Geophysics*, **12**, 169–178, doi: [10.1007/s11770-015-0485-1](https://doi.org/10.1007/s11770-015-0485-1).
- Gerchberg, R. W., and W. O. Saxton, 1972, A practical algorithm for the determination of phase from image and diffraction plane pictures: *Optik*, **35**, 227–246.
- Ghaderpour, E., and S. Pagiatakis, 2017, Least-squares wavelet analysis of unequally spaced and nonstationary time series and its applications: *Mathematical Geosciences*, **49**, 819–844, doi: [10.1007/s11004-017-9691-0](https://doi.org/10.1007/s11004-017-9691-0).
- Gülünay, N., 2003, Seismic trace interpolation in the Fourier transform domain: *Geophysics*, **68**, 355–369, doi: [10.1190/1.1543221](https://doi.org/10.1190/1.1543221).
- Guo, Z., Y. Zheng, and W. Liao, 2015, High fidelity seismic trace interpolation: 85th Annual International Meeting, SEG, Expanded Abstracts, 3915–3919, doi: [10.1190/segam2015-5923716.1](https://doi.org/10.1190/segam2015-5923716.1).
- Hollander, Y., D. Kosloff, Z. Koren, and A. Bartana, 2012, Seismic data interpolation by orthogonal matching pursuit: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20148144](https://doi.org/10.3997/2214-4609.20148144).
- Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: *Geophysics*, **69**, 1560–1568, doi: [10.1190/1.1836829](https://doi.org/10.1190/1.1836829).
- Liu, G., and X. Chen, 2017, Seismic data interpolation using frequency domain complex nonstationary autoregression: *Geophysical Prospecting*, **66**, 478–497, doi: [10.1111/1365-2478.12499](https://doi.org/10.1111/1365-2478.12499).
- Liu, G., X. Chen, J. Du, and K. Wu, 2012, Random noise attenuation using f-x regularized nonstationary autoregression: *Geophysics*, **77**, no. 2, V61–V69, doi: [10.1190/geo2011-0117.1](https://doi.org/10.1190/geo2011-0117.1).
- Lomb, N. R., 1976, Least squares frequency analysis of unequally spaced data: *Astrophysics and Space Science*, **39**, 447–462, doi: [10.1007/BF00648343](https://doi.org/10.1007/BF00648343).
- Özbek, A., A. K. Özdemir, and M. Vassallo, 2009, Interpolation by matching pursuit: 79th Annual International Meeting, SEG, Expanded Abstracts, 3254–3258, doi: [10.1190/1.3255534](https://doi.org/10.1190/1.3255534).
- Pagiatakis, S., 1999, Stochastic significance of peaks in the least squares spectrum: *Journal of Geodesy*, **73**, 67–78, doi: [10.1007/s001900050220](https://doi.org/10.1007/s001900050220).
- Rao, C. R., and S. K. Mitra, 1972, Generalized inverse of a matrix and its applications: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* 1, 601–620.
- Schonewille, M., A. Klaedtke, and A. Vigner, 2009, Antialias antileakage Fourier transform: 79th Annual International Meeting, SEG, Expanded Abstracts, 3249–3253, doi: [10.1190/1.3255533](https://doi.org/10.1190/1.3255533).
- Spitz, S., 1991, Seismic trace interpolation in the f-x domain: *Geophysics*, **56**, 785–794, doi: [10.1190/1.1443096](https://doi.org/10.1190/1.1443096).
- Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: *Geophysics*, **74**, no. 6, V123–V132, doi: [10.1190/1.3245216](https://doi.org/10.1190/1.3245216).
- Vaníček, P., 1969, Approximate spectral analysis by least squares fit: *Astrophysics and Space Science*, **4**, 387–391, doi: [10.1007/BF00651344](https://doi.org/10.1007/BF00651344).
- Vassallo, M., A. Özbek, A. K. Özdemir, and K. Eggenberger, 2010, Crossline wavefield reconstruction from multicomponent streamer data: Multichannel interpolation by matching pursuit: 80th Annual International Meeting, SEG, Expanded Abstracts, 3594–3598, doi: [10.1190/1.3513597](https://doi.org/10.1190/1.3513597).
- Wang, B., X. Chen, J. Li, and J. Cao, 2016, An improved weighted projection onto convex sets method for seismic data interpolation and denoising: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9**, 228–235, doi: [10.1109/JSTARS.2015.2496374](https://doi.org/10.1109/JSTARS.2015.2496374).
- Wang, Y., 2002, Seismic trace interpolation in the f-x-y domain: *Geophysics*, **67**, 1232–1239, doi: [10.1190/1.1500385](https://doi.org/10.1190/1.1500385).
- Weglein, A. B., F. A. Gasparotto, P. M. Carvalho, and R. H. Stolt, 1997, An inverse scattering series method for attenuating multiples in seismic reflection data: *Geophysics*, **62**, 1975–1989, doi: [10.1190/1.1444298](https://doi.org/10.1190/1.1444298).
- Xu, S., Y. Zhang, and G. Lambaré, 2010, Antileakage Fourier transform for seismic data regularization in higher dimensions: *Geophysics*, **75**, no. 6, WB113–WB120, doi: [10.1190/1.3507248](https://doi.org/10.1190/1.3507248).
- Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, no. 4, V87–V95, doi: [10.1190/1.1993713](https://doi.org/10.1190/1.1993713).
- Xu, S., Y. Zhang, and B. Tang, 2011, 3D angle gathers from reverse time migration: *Geophysics*, **76**, no. 2, S77–S92, doi: [10.1190/1.3536527](https://doi.org/10.1190/1.3536527).
- Yang, P., J. Gao, and W. Chen, 2012, Curvelet-based POCs interpolation of nonuniformly sampled seismic records: *Journal of Applied Geophysics*, **79**, 90–99, doi: [10.1016/j.jappgeo.2011.12.004](https://doi.org/10.1016/j.jappgeo.2011.12.004).