GEOPHYSICS, VOL. 90, NO. 2 (MARCH-APRIL 2025); P. V161–V177, 25 FIGS., 6 TABLES.
10.1190/GEO2024-0259.1

Adaptive dictionary identification framework and its application to sparsity-optimized harmonic noise separation

Yanglijiang Hu[1], Weiwei Xu[1], Xiaokai Wang[1], Dawei Liu[2], and Wenchao Chen[1]

ABSTRACT

Coherent noise separation stands as a crucial step in seismic data processing. The morphological component analysis (MCA)-based separation method, which treats coherent noise and signal as distinct components and represents them sparsely with dictionaries, has been widely adopted for noise suppression. Typically, constructing effective fixed dictionaries for MCA-based methods involves necessary expert knowledge to meticulously select appropriate transform basis functions from an extensive dictionary library and fine-tune their parameters. To reduce time consumption and ensure optimal dictionary construction, we introduce an adaptive framework for identifying optimal dictionaries used in MCA-based coherent noise separation. Initially, we define a fixed dictionary library comprising dictionaries constructed using various transform basis functions with their corresponding parameters. Subsequently, we formulate a relative sparsity minimization problem (RSMP) to identify the optimal fixed dictionaries that minimize relative sparsity within this predefined library. Finally, we design a genetic algorithm to solve RSMP. The identified dictionaries are then applied to MCA-based coherent noise separation. Synthetic and field data examples demonstrate the effectiveness of our method.

INTRODUCTION

Coherent noise in seismic data obscures signal, leading to a substantial reduction in the signal-to-noise ratio (S/N) and presenting considerable challenges for seismic inversion and imaging. Therefore, effective separation of coherent noise is a crucial step in seismic data processing.

Sparse representation methods, which assume that seismic data can be sparsely represented in appropriate transform domains, have been widely used for coherent noise separation. These methods are generally divided into two groups. The first group focuses on determining the sparse representation of either signal or coherent noise in a seismic data set using a single dictionary, with the other component being derived through subtraction. In this group, fixed dictionaries constructed using transform basis functions with their corresponding parameters (Richter et al., 2022) are widely used. Examples are the Radon transform (Trad et al., 2003; Fan et al., 2015), the Fourier transform (Yilmaz, 1987; Mostafa, 2012; Felix and Sacchi, 2021), the wavelet transform (Deighan and Watts, 1997; Yu and Garossino, 2005; Almeida et al., 2015), and the chirplet transform (Gregoire et al., 2009; Boßmann and Ma, 2015; Laurent et al., 2015). Recognizing that coherent noise exhibits distinct morphological features compared with signal, the second group uses morphological component analysis (MCA) (Starck et al., 2004). This approach uses multiple dictionaries to sparsely represent signal and coherent noise separately, effectively leveraging their unique morphological features.

Assuming that each seismic datum is a linear combination of signal and coherent noise, the effectiveness of MCA-based separation relies on the fidelity of the dictionaries used. Each dictionary is tailored to represent one specific component of the data sparsely. Following this principle, various fixed dictionaries, constructed using distinct transform basis functions, have been introduced. For instance, Wang et al. (2010) use the local discrete cosine transform (LDCT) and the stationary symlet wavelet transform (SWT) to represent ground roll and reflections separately. Xu et al. (2013) use the stationary Coiflet wavelet transform for reflection representation and the discrete cosine transform (DCT) for monochromatic noise representation. Chen et al. (2017) use two tunable Q-factor wavelet transforms (TQWTs) with low and high Q-values to represent ground roll and body waves, respectively. Chen et al. (2018) use the continuous wavelet transform (CWT) and the DCT for the sparse representation of signal and DAS coupling noise, respectively. Liu et al.

Manuscript received by the Editor 24 April 2024; revised manuscript received 11 November 2024; published ahead of production 9 December 2024; published online 6 March 2025.
[1]Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China. E-mail: 571635961@qq.com; 544274644@qq.com; xkwang@xjtu.edu.cn; wencchen@xjtu.edu.cn (corresponding author).
[2]University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: 409791715@qq.com.

(2022a) use the chirplet transform for harmonic noise representation and the CWT for reflection representation. In the τ-p domain, Liu et al. (2022b) apply a 2D stationary SWT to represent multiple reflection refractions and a shearlet transform to represent reflections. Hu et al. (2024) use the TQWT to represent signal while representing wind turbine noise by the DCT. Learned dictionaries have also been used in MCA-based coherent noise separation. Liu et al. (2021) use K-singular value decomposition to construct dictionaries for signal and footprint, respectively. A notable distinction exists between fixed and learned dictionaries. Atoms in a fixed dictionary, computed using a specific transform basis function and parameters, remain unchanged when applied to data. In contrast, atoms in a learned dictionary are dynamically adapted to better fit the data and are continually updated during the learning process. Although a learned dictionary offers adaptability in representing various components, a fixed dictionary provides a straightforward and rapid separation solution, as it eliminates the need to calculate and update atoms individually.

Currently, in MCA-based coherent noise separation, the traditional approach to constructing fixed dictionaries begins with the selection of appropriate transform basis functions from an extensive library, usually based on observations or feature analysis of the components to be separated. This is followed by necessary expert knowledge to fine-tune the parameters of the chosen basis functions until the dictionaries reach the desired level of effectiveness. However, this process is time consuming and does not ensure the optimal construction of dictionaries. To streamline the construction of fixed dictionaries and guarantee their optimality, we introduce a framework for the adaptive identification of optimal dictionaries used in MCA-based coherent noise separation.

According to MCA theory, relative sparsity can serve as a metric for assessing a dictionary's efficacy, where a highly effective dictionary yields a very small relative sparsity value (Starck et al., 2004). Therefore, in the proposed identification framework, we present a relative sparsity function as the metric to evaluate a fixed dictionary's efficacy in representing signal or coherent noise. Based on this metric, we formalize the task of identifying the optimal fixed dictionary for representing signal or coherent noise as a relative sparsity minimization problem (RSMP) constrained within a defined solution space. We propose the concept of a fixed dictionary library and use it as the solution space for the RSMP. The fixed dictionary library is defined as a collection of distinct dictionaries with each constructed using different transform basis functions and parameters. Dictionaries that share the same transform basis function but differ in parameters form subsets within the library. Each subset is distinguished by its unique basis function. A genetic algorithm (GA) (Eldos, 2008; Auger and Hansen, 2016) is used to solve the RSMP within each subset, allowing for the identification of the optimal dictionary for that subset. Subsequently, we compare the relative sparsity values of these subset-specific optimal dictionaries. The dictionary that achieves the lowest relative sparsity value, whether for representing either signal or coherent noise, is then applied to the MCA-based coherent noise separation.

To evaluate the proposed framework, we apply it to the MCA-based harmonic noise separation problem. Harmonic noise, commonly encountered in seismic data sets acquired by vibroseis systems (Lebedev and Beresnev, 2004), typically features linear frequency modulation across a broad range and has higher energy than the signal from deep layers (Butler and Russell, 2003). These features enable harmonic noise to severely obscure the signal, posing significant challenges for subsequent seismic data processing. Therefore, harmonic noise separation is necessary.

Various filtering methods have been proposed to separate harmonic noise. Li et al. (1995) propose phase-shift filtering, which requires estimating the frequencies of harmonic noise in each data set. Meunier and Bianchi (2002) introduce a harmonic reduction method that does not require frequency estimation, though it demands large amounts of uncorrelated data. Sicking et al. (2009) use a strategy to predict and subtract harmonic noise based on the ground force signal. In the absence of the ground force signal, Wang et al. (2012) develop a trace-by-trace filter to remove cross-harmonic noise in slip-sweep vibroseis data. This method, however, adds computational burden due to the need to estimate a weighted phase-shift operator between traces from successive shots. Karslı and Dondurur (2018) use iterative mean filtering, which risks losing wavefield details (Gonzalez and Woods, 2018). Denisov et al. (2021) propose an optimal recursive filtering method that requires optimization during each recursion. In contrast, the MCA-based harmonic noise separation does not require frequency estimation or parameter optimization during the separation process. Nevertheless, it requires the preconstruction of two well-suited dictionaries to represent the signal and harmonic noise, respectively.

Therefore, we apply our proposed framework to identify effective fixed dictionaries for MCA-based harmonic noise separation adaptively. In the proposed framework, we establish a fixed dictionary library comprising TQWTs, asymmetric Gaussian chirplet transforms (AGCTs), and LDCTs. These dictionary types have been commonly used in seismic noise separation and are all likely to be effective in representing signal and harmonic noise. The TQWT, with a low Q-value, is suitable for sparsely representing low-oscillation body waves, whereas a high Q-value TQWT aptly captures high-oscillation surface waves, a type of coherent noise (Selesnick, 2011; Chen et al., 2017; Hu et al., 2024). The AGCT effectively represents seismic waves across various time-frequency distributions (Boßmann and Ma, 2015, 2016). In addition, the LDCT has been used to represent surface waves and wind turbine noise — a form of coherent noise containing bode waves (Wang et al., 2010; Chen et al., 2018; Hu et al., 2024). The LDCT offers a representation of the wavefield based on its stationarity. Considering that the signal in this study mainly consists of body waves and harmonic noise represents coherent noise, we use the TQWT, the AGCT, and the LDCT as initial dictionary types. Using synthetic and field data examples, we demonstrate the adaptability of the proposed framework in identifying optimal fixed dictionaries for signal or harmonic noise within the given library. A comparison with the high-energy noise attenuation method (Yu and Garossino, 2005) highlights the superior performance of the MCA using identified fixed dictionaries in effectively separating harmonic noise and preserving signal integrity.

First, we describe the fundamental theory of MCA-based coherent noise separation. This is followed by an outline of the adaptive dictionary identification framework. Next, we present examples of harmonic noise separation using synthetic and field data. Finally, the paper discusses the results and concludes with final remarks.

MCA-BASED COHERENT NOISE SEPARATION

The problem of coherent noise separation is resolved in a separation framework where each seismic datum $\mathbf{s}$ is modeled as a linear superposition of the signal $\mathbf{s}_s$ and the coherent noise $\mathbf{s}_n$, that is,

$$\mathbf{s} = \mathbf{s}_s + \mathbf{s}_n. \tag{1}$$

The task involves separation of $\mathbf{s}_s$-type signal and $\mathbf{s}_n$-type coherent noise from $\mathbf{s}$. The MCA (Starck et al., 2004) addresses this by leveraging the distinct morphological features of signal and coherent noise. It recovers signal and coherent noise by their sparse representations. Let $\mathbf{D}_s$ and $\mathbf{D}_n$ be the dictionary for signal and coherent noise, respectively, and $\mathbf{x}_{s,\mathbf{D}_s}$ and $\mathbf{x}_{n,\mathbf{D}_n}$ denote the representation coefficients of $\mathbf{s}_s$ and $\mathbf{s}_n$ with respect to $\mathbf{D}_s$ and $\mathbf{D}_n$, respectively. The relationship between the dimensions of a dictionary and its corresponding representation coefficients depends on the dictionary type used. The coherent noise separation problem can then be formulated as an MCA-based separation problem, expressed as

$$\{\mathbf{x}^*_{s,\mathbf{D}_s}, \mathbf{x}^*_{n,\mathbf{D}_n}\} = \arg\min_{\mathbf{x}_{s,\mathbf{D}_s},\mathbf{x}_{n,\mathbf{D}_n}} \lambda(||\mathbf{x}_{s,\mathbf{D}_s}||_1 + ||\mathbf{x}_{n,\mathbf{D}_n}||_1) + \frac{1}{2}||\mathbf{s} - \mathbf{D}_s\mathbf{x}_{s,\mathbf{D}_s} - \mathbf{D}_n\mathbf{x}_{n,\mathbf{D}_n}||_2^2, \quad (2)$$

where $\mathbf{x}^*_{s,\mathbf{D}_s}$ and $\mathbf{x}^*_{n,\mathbf{D}_n}$ are the optimal sparse representation coefficients of $\mathbf{s}_s$ and $\mathbf{s}_n$, respectively, with respect to $\mathbf{D}_s$ and $\mathbf{D}_n$. In addition, λ is a Lagrange multiplier. Here, $\hat{\mathbf{s}}_s = \mathbf{D}_s\mathbf{x}^*_{s,\mathbf{D}_s}$ and $\hat{\mathbf{s}}_n = \mathbf{D}_n\mathbf{x}^*_{n,\mathbf{D}_n}$ represent the recovered signal and coherent noise, respectively. Equation 2 can be solved by the block coordinate relaxation algorithm (Bruce et al., 2006). The quality of the recovered signal and coherent noise depends on the effectiveness of preconstructed dictionaries $\mathbf{D}_s$ and $\mathbf{D}_n$.

According to MCA theory, each effective dictionary is designed to sparsely represent a specific component type while being inefficient at sparsely representing others. Thus, the relative sparsity is a possible measure of a dictionary's efficacy (Starck et al., 2004). For a dictionary $\mathbf{D}_s$ representing signal, a relative sparsity value $L_{s,\mathbf{D}_s}$ can be calculated by

$$L_{s,\mathbf{D}_s} = \frac{\|\mathbf{x}^*_{s,\mathbf{D}_s}\|_1}{\|\mathbf{x}^*_{n,\mathbf{D}_s}\|_1}, \quad (3)$$

where $\mathbf{x}^*_{s,\mathbf{D}_s}$ and $\mathbf{x}^*_{n,\mathbf{D}_s}$ are the optimal sparse representation coefficients of $\mathbf{s}_s$ and $\mathbf{s}_n$, respectively, with respect to $\mathbf{D}_s$. The smallest $L_{s,\mathbf{D}_s}$ is achieved when $\mathbf{D}_s$ most effectively represents the signal, necessitating the minimization of $\|\mathbf{x}^*_{s,\mathbf{D}_s}\|_1$ while maximizing $\|\mathbf{x}^*_{n,\mathbf{D}_s}\|_1$. A similar metric can be formulated to assess the efficacy of $\mathbf{D}_n$, the dictionary representing coherent noise.

The fixed dictionary, comprising a series of atoms derived from a transform basis function and specific parameters, provides a quick and straightforward MCA-based separation. There is a wide variety of available transform basis functions, and their parameters span a broad range, presenting a challenge in identifying the optimal combination for fixed dictionaries. The traditional fixed dictionary construction approach involves necessary expert knowledge to conduct many experiments. This way can be time consuming and does not guarantee the creation of optimal dictionaries. To address these limitations, we propose a framework aimed at adaptively constructing optimal dictionaries for representing signal and coherent noise. The MCA-based separation method is applicable to 1D, 2D, and higher-dimensional data. In this paper, to concisely demonstrate the proposed method, we use a seismic trace with a dimension of N as $\mathbf{s}$ in equation 1. In this case, the dimensions of $\mathbf{D}_s$, $\mathbf{D}_n$, $\mathbf{x}^*_{s,\mathbf{D}_s}$, and $\mathbf{x}^*_{n,\mathbf{D}_n}$ depend on preconstructed dictionary types. For a specific example, if $\mathbf{D}_s$ is an $N \times M$ matrix, $\mathbf{x}^*_{s,\mathbf{D}_s}$ has a dimension of M. The proposed framework and its implementation are further outlined in the subsequent section.

ADAPTIVE DICTIONARY IDENTIFICATION FRAMEWORK

To identify the optimal fixed dictionaries representing signal $\mathbf{s}_s$ and coherent noise $\mathbf{s}_n$ in equation 2, respectively, we establish a fixed dictionary library. In this library, dictionaries are constructed using a variety of transform basis functions and their respective parameters. By using this predefined fixed dictionary library as the solution space and using a relative sparsity function as the objective, we define the task of identifying the optimal fixed dictionary as an RSMP constrained within the solution space.

To address this constrained RSMP, we propose a strategy that first partitions the library into subsets. Each subset comprises dictionaries that share a transform basis function but differ in parameters. We then design a GA to solve the RSMP within each subset, thereby determining the optimal dictionary specific to that subset. Among these subset-specific optimal dictionaries, we select the one that achieves the lowest relative sparsity to represent either the signal or the coherent noise.

In summary, the proposed framework for adaptive fixed dictionary identification comprises three main parts: a fixed dictionary library, the RSMP, and GA designed to solve the RSMP. Each of these parts is elaborated in the subsequent three sections.

Fixed dictionary library

Let $\mathbf{\Omega}$ be a fixed dictionary library constructed with K different transform basis functions $g_1, g_2, \ldots, g_K$. Taking the transform basis function g_1 as an example, let $\boldsymbol{\omega}_1$ be the index consisting of specific parameter values of g_1. The combination of varied parameter values generates multiple parameter indices $\boldsymbol{\omega}_1^{(1)}, \boldsymbol{\omega}_1^{(2)}, \ldots$, which are essential for constructing diverse dictionaries $\mathbf{D}_{g_1}(\boldsymbol{\omega}_1^{(1)}), \mathbf{D}_{g_1}(\boldsymbol{\omega}_1^{(2)}), \ldots$. Thus, g_1 and its optional parameters can be expressed as a set of dictionaries $\{\mathbf{D}_{g_1}(\boldsymbol{\omega}_1) : \boldsymbol{\omega}_1 \in \{\boldsymbol{\omega}_1^{(1)}, \boldsymbol{\omega}_1^{(2)}, \ldots\}\} \in \mathbf{\Omega}$.

The same expressions can be written for any other transform basis functions and their respective parameter indices. Let $\boldsymbol{\omega}_i, i \in \{1, 2, \ldots, K\}$ be the index consisting of specific parameter values of the ith basis function $g_i \in \{g_1, g_2, \ldots, g_K\}$. Accordingly, the fixed dictionary library $\mathbf{\Omega}$ can be expressed as a set comprising K fixed dictionary subsets:

$$\mathbf{\Omega} = \{\mathbf{D}_{g_i}(\boldsymbol{\omega}_i) : \boldsymbol{\omega}_i \in \{\boldsymbol{\omega}_i^{(1)}, \boldsymbol{\omega}_i^{(2)}, \ldots\}\}_{i=1}^{K}. \quad (4)$$

Each subset can independently be used for identifying a subset-specific optimal dictionary, and these subset-specific optimal dictionaries are then compared to determine the optimal dictionary in $\mathbf{\Omega}$. Notably, the transform basis functions and parameters used to build the library should be comprehensively sufficient to ensure that the library contains effective dictionaries for representing signal and coherent noise.

Relative sparsity minimization problem

In this section, we propose a search strategy aimed at identifying the optimal dictionary among $\mathbf{\Omega}$ (equation 4). We use the pursuit of the optimal dictionary for signal representation as an illustrative example, with similar progress applicable to the pursuit of the optimal fixed dictionary for representing coherent noise.

The first step is to find the subset-specific optimal dictionary in $\mathbf{\Omega}$. Taking the ith subset $\{\mathbf{D}_{g_i}(\boldsymbol{\omega}_i) : \boldsymbol{\omega}_i \in \{\boldsymbol{\omega}_i^{(1)}, \boldsymbol{\omega}_i^{(2)}, \ldots\}\} \in \mathbf{\Omega}$ as

an example, $\mathbf{D}_{g_i}(\boldsymbol{\omega}_i)$ is the fixed dictionary constructed using a transform basis function g_i and an index $\boldsymbol{\omega}_i$. Identifying the subset-specific optimal dictionary $\mathbf{D}_{g_i}(\boldsymbol{\omega}_i^*)$ consists of finding the optimal parameter index of the ith subset $\boldsymbol{\omega}_i^*$.

Following the idea of using the relative sparsity as a metric for a dictionary's efficacy (Starck et al., 2004), a relative sparsity function $L_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)$ is used to evaluate the quality of any index $\boldsymbol{\omega}_i$ for signal representation and is applied on

$$L_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i) = \frac{\|\mathbf{x}^*_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)\|_1}{\|\mathbf{x}^*_{n,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)\|_1}, \tag{5}$$

where $\mathbf{x}^*_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)$ and $\mathbf{x}^*_{n,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)$ are the optimal sparse representation coefficients of signal and coherent noise, respectively, with respect to the fixed dictionary $\mathbf{D}_{g_i}(\boldsymbol{\omega}_i)$. The smaller the value of $L_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i)$, the more effective $\mathbf{D}_{g_i}(\boldsymbol{\omega}_i)$ is in representing the signal. Based on this metric (equation 5), the optimal parameter index $\boldsymbol{\omega}_i^*$ can be expressed as the solution to the following constrained RSMP with

$$\boldsymbol{\omega}_i^* = \arg\min_{\boldsymbol{\omega}_i} L_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i), \text{ subject to } \boldsymbol{\omega}_i \in \{\boldsymbol{\omega}_i^{(1)}, \boldsymbol{\omega}_i^{(2)}, \ldots\}. \tag{6}$$

After obtaining K optimal parameter indices $\{\boldsymbol{\omega}_i^*\}_{i=1}^K$ by solving the RSMPs constrained within K subsets of $\boldsymbol{\Omega}$, we have subset-specific optimal dictionaries $\{\mathbf{D}_{g_i}(\boldsymbol{\omega}_i^*)\}_{i=1}^K$. The second step is to select the signal-representation dictionary $\mathbf{D}^* \in \{\mathbf{D}_{g_i}(\boldsymbol{\omega}_i^*)\}_{i=1}^K$ by the following condition:

$$\mathbf{D}^* : L_{s,\mathbf{D}^*} = \min\{L_{s,\mathbf{D}_{g_i}}(\boldsymbol{\omega}_i^*)\}_{i=1}^K. \tag{7}$$

As shown in equation 7, the key to identifying $\mathbf{D}^*$ is to solve each RSMP. Take the RSMP presented in equation 6 as an example. It can be viewed as a parameter optimization problem. The grid search algorithm (Sundhararajan et al., 1998) and GA (Goldberg, 1989) are two main techniques used to solve geophysical parameter optimization problems, such as residual statics corrections and inversion (Nikravesh et al., 2003; Pierini et al., 2019). Let G, P, and F be the total iteration number, parameter index number per iteration, and relative sparsity evaluation complexity, respectively. The time complexity of the GA is expressed as $O(\mathrm{PGF})$. Let Pn and V be the total number of parameters and discretized values per parameter range, respectively. The time complexity of traditional grid search algorithm is expressed as $O(V^{Pn})$. Solving RSMP involves multiple types of dictionaries with various parameters and extensive parameter spaces. In this case, $O(\mathrm{PGF})$ is lower than $O(V^{Pn})$. Therefore, we design a GA to solve RSMP. Based on a fixed dictionary library, the constrained RSMP and the designed GA, the main flow of the proposed identification framework is summarized in Figure 1. The GA designed to identify the fixed dictionary-representing signal is depicted in the next section.

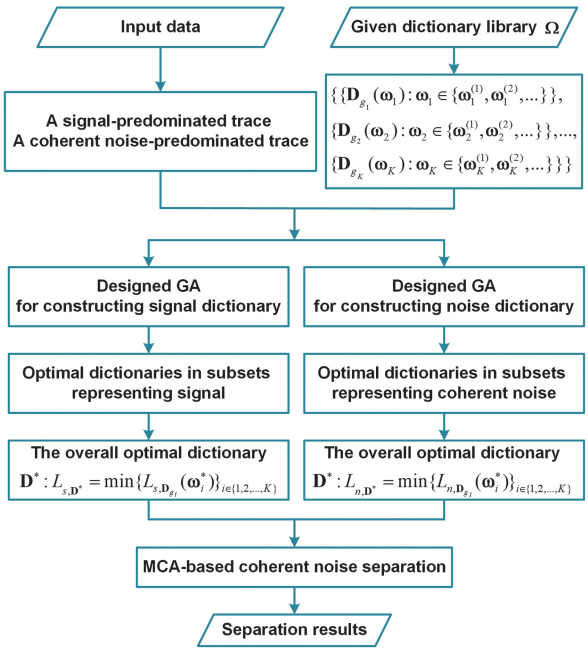

Figure 1. Main flow of the adaptive dictionary identification framework.

GA designed for RSMP solution

As shown in Figure 1, the GA for the RSMP solution requires a fixed dictionary library, a trace with signal predominated and a trace with coherent noise predominated, which are used to compute the relative sparsity. When identifying the dictionary-representing signal, the relative sparsity function $L_{s,\mathbf{D}_{g_i}}(\cdot)$ is as shown in equation 5, assuming that the given library is $\boldsymbol{\Omega} = \{\mathbf{D}_{g_i}(\boldsymbol{\omega}_i) : \boldsymbol{\omega}_i \in \{\boldsymbol{\omega}_i^{(1)}, \boldsymbol{\omega}_i^{(2)}, \ldots\}\}_{i=1}^K$, where $\boldsymbol{\omega}_1$ consists of two parameters Q and r. In this case, the pseudocode of the GA used to solve the RSMP within $\{\mathbf{D}_{g_1}(\boldsymbol{\omega}_1) : \boldsymbol{\omega}_1 \in \{\boldsymbol{\omega}_1^{(1)}, \boldsymbol{\omega}_1^{(2)}, \ldots\}\}$ is shown in Algorithm 1. There are four main steps in Algorithm 1: (1) initialization (line 1): the designed GA starts with a population of randomly generated parameter indices; (2) selection (line 4): selecting some

Algorithm 1. GA used to solve the RSMP within $\{\mathbf{D}_{g_1}(\boldsymbol{\omega}_1) : \boldsymbol{\omega}_1 \in \{\boldsymbol{\omega}_1^{(1)}, \boldsymbol{\omega}_1^{(2)}, \ldots\}\}$.

Input: Max iteration number G, transform basis function g_1, total parameter index number per iteration P, recombination rate σ_1, mutation rate σ_2, and iteration number q

Output: Optimal parameter index $\boldsymbol{\omega}_1^*$.

1: Initialize parameter indices $\{\boldsymbol{\omega}_1^{(q=0,1)}, \boldsymbol{\omega}_1^{(q=0,2)}, \ldots, \boldsymbol{\omega}_1^{(q=0,P)}\}$
2: for q in range (G)
3: calculate relative sparsity values $\{L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}_1^{(q,1)}), \ldots, L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}_1^{(q,P)})\}$
4: for $\boldsymbol{\omega}'^{(q,i)}_1$ in $\{\boldsymbol{\omega}'^{(q,1)}_1, \ldots, \boldsymbol{\omega}'^{(q,\sigma_1 P)}_1\}$
5: if rand() $< \sigma_1$
6: $\boldsymbol{\omega}'^{(q,i)}_1$ = recombination $(\{\boldsymbol{\omega}_1^{(q,1)}, \ldots, \boldsymbol{\omega}_1^{(q,\sigma_1 P)}\}, \boldsymbol{\omega}'^{(q,i)}_1)$
7: if rand() $< \sigma_2$
8: $\boldsymbol{\omega}'^{(q,i)}_1$ = mutation $(\boldsymbol{\omega}'^{(q,i)}_1)$
9: calculate relative sparsity values $\{L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}'^{(q,1)}_1), \ldots, L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}'^{(q,\sigma_1 P)}_1)\}$
10: for $\boldsymbol{\omega}'^{(q,i)}_1$ in $\{\boldsymbol{\omega}'^{(q,1)}_1, \ldots, \boldsymbol{\omega}'^{(q,\sigma_1 P)}_1\}$
11: if $\exists \boldsymbol{\omega}_1^{(q,j)} \in \{\boldsymbol{\omega}_1^{(q,1)}, \ldots, \boldsymbol{\omega}_1^{(q,\sigma_1 P)}\}, L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}_1^{(q,i)}) > L_{s,\mathbf{D}_{g_1}}(\boldsymbol{\omega}'^{(q,i)}_1)$
12: $\boldsymbol{\omega}_1^{(q,i)}$ = reinsertion$(\boldsymbol{\omega}'^{(q,i)}_1, \boldsymbol{\omega}_1^{(q,i)})$
13: $q \leftarrow q + 1$

parameter indices with larger relative sparsity values from the indices generated in the previous iteration; (3) recombination and mutation (lines 4–8): updating parameters in selected indices; and (4) reinsertion (lines 10–12): replace the original indices with better indices. In addition, after calculating the relative sparsity values, we reorder the corresponding indices in descending order according to these values. We repeat the preceding steps until the GA reaches the maximum number of iterations. Figure 2 shows specific processes of recombination, mutation, and reinsertion in Algorithm 1. Specifically, parameters in each index are generated separately, each of which is randomly sampled from its range. The same expression can be written for the GA used to solve RSMPs within other dictionary sets when identifying the dictionary representing signal or coherent noise.

In the next section, we present the evolution of minimum relative sparsity values when solving the RSMP using the designed GA, which shows the performance of the designed GA.

NUMERICAL EXPERIMENTS

To demonstrate the effectiveness of the proposed adaptive dictionary identification framework, we apply it to the MCA-based separation of harmonic noise. In this part, the high-energy noise attenuation method with the Coiflet wavelet (Yu and Garossino, 2005) is used as the conventional method for comparison.

Following the main flow of the proposed dictionary identification framework (Figure 1), we initially define a fixed dictionary library comprising TQWTs, AGCTs, and LDCTs, which are shown in Appendix A, B, and C, respectively. Table 1 shows the parameter space of the defined library for each data example. Based on Table 1, we express the TQWT constructed by a parameter index (Q, r) as TQWT (Q, r). Here, Q is a Q-value, and r is a redundancy factor. Similarly, AGCT and LDCT are expressed as AGCT $(a, \beta, f, \gamma, \theta)$ and LDCT (l_b, l_o), respectively. The parameters a, β, f, γ, θ, l_b, and l_o are the bandwidth factor, the asymmetry factor, the central frequency, the chirp rate, the phase, the blocking length, and the overlapping length, respectively. In addition, in the GA, the maximum iteration number (stopping criteria) and the parameter index number per iteration are set to 150 and 100, respectively. Our computer processor is an Intel(R) Core(TM) i5-8500 CPU with 3.00 GHz and 6 cores. In the following data examples, the average runtime to obtain the optimized TQWT, AGCT, and 1D LDCT representing signal or harmonic noise is 7.8, 37.5, and 10.0 min, respectively. For data acquired in the same survey area, we generally identify dictionaries once.

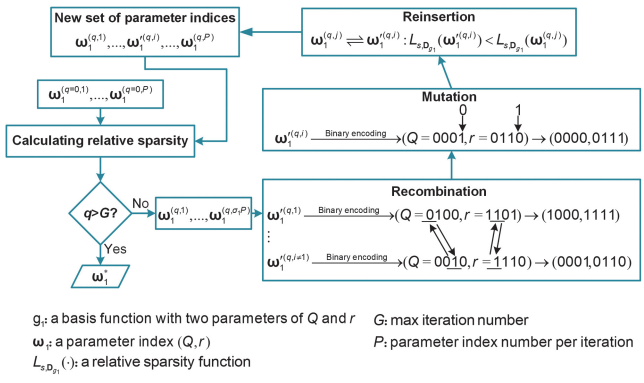

Figure 2. Main flow of the GA used to solve the RSMP within a dictionary set $\{\mathbf{D}_{g_1}(\boldsymbol{\omega}_1) : \boldsymbol{\omega}_1 \in \{\boldsymbol{\omega}_1^{(1)}, \boldsymbol{\omega}_1^{(2)}, \dots\}\}$ when constructing fixed dictionary for signal.

First synthetic data example

In this section, we use two synthetic data sets as examples (Figure 3). The first data set (Figure 3a) comprises the synthetic signal and harmonic noise (Figure 3b and 3c), which are generated by a sweep signal with frequencies increasing linearly from 10 to 40 Hz. The second data set (Figure 3d) comprises the synthetic signal and harmonic noise (Figure 3e and 3f), which are generated by a sweep signal with frequencies increasing linearly from 10 to 60 Hz. Each

Table 1. The given dictionary library comprising TQWTs, AGCTs, and the 1D LDCTs for 1D data with the length of N.

Fixed dictionary	Parameter	Ranges	Parameter index $\boldsymbol{\omega}$
TQWT	Q-value Q	$[1, 6]$	(Q, r)
	Redundancy factor r	$[1, 20]$	—
AGCT	Bandwidth factor a	$(0, 10^{-3})$	$(a, \beta, f, \gamma, \theta)$
	Asymmetry factor β	$(-1, 1)$	—
	Central frequency f	$[0, \pi]$	—
	Chirp rate γ	$[0, \pi]$	—
	Phase θ	$[0, \pi]$	—
1D LDCT	Blocking length l_b	$(1, N)$	(l_b, l_o)
	Overlapping length l_o	$[1, l_b)$	—

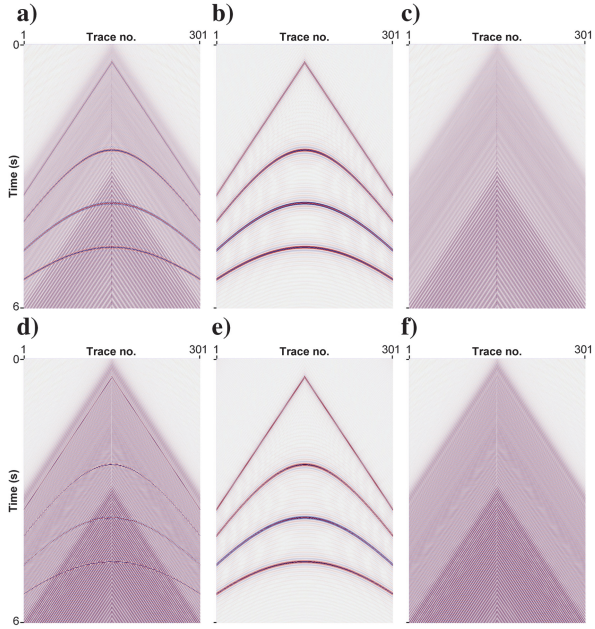

Figure 3. The first synthetic data example contaminated with harmonic noise. The trace interval is 20 m. (a) The first data set with an S/N of −8.05 dB; (b and c) signal and harmonic noise contained in the first data set, respectively; (d) the second data set with an S/N of −11.47 dB; and (e and f) signal and harmonic noise contained in the second data set, respectively.

data set has 301 traces and 3000 samples per trace. The time sampling is 0.002 s, and the trace interval is 20 m. The first and second data sets (Figure 3a and 3b) have S/Ns of −8.05 and −11.47 dB, respectively.

Dictionary identification for the first data set

We extract the 96th–607th samples of the 165th trace in Figure 3b as signal trace (Figure 4a). The last 512 samples of the 75th trace in Figure 3c are extracted as harmonic noise trace (Figure 4b). Their time-frequency spectra are shown in Figure 4c and 4d, respectively. Figure 5a–5c shows the evolution of minimum relative sparsity values in the GA, which identifies the optimized TQWT, AGCT, and 1D LDCT to represent a signal trace (Figure 4a). Figure 6a–6c shows the evolution of minimum relative sparsity values in the GA, which identifies the optimized TQWT, AGCT, and 1D LDCT to represent a harmonic noise trace (Figure 4b), respectively. These minimum relative sparsity values and corresponding index parameters are shown in Figures 5d–5f and 6d–6f, respectively. After 50 iterations, the minimum relative sparsity values and index parameters stabilize at specific values. Table 2 shows these specific values. The optimized TQWT (1.02, 16.15) leads to the smallest relative sparsity value 0.2869, in representing the signal trace. In representing the harmonic noise trace, the optimized AGCT (8.94×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10) leads to the smallest relative sparsity value 0.4109. Figure 7a and 7b shows atoms of the TQWT (1.02, 16.15) and AGCT (8.94×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10),

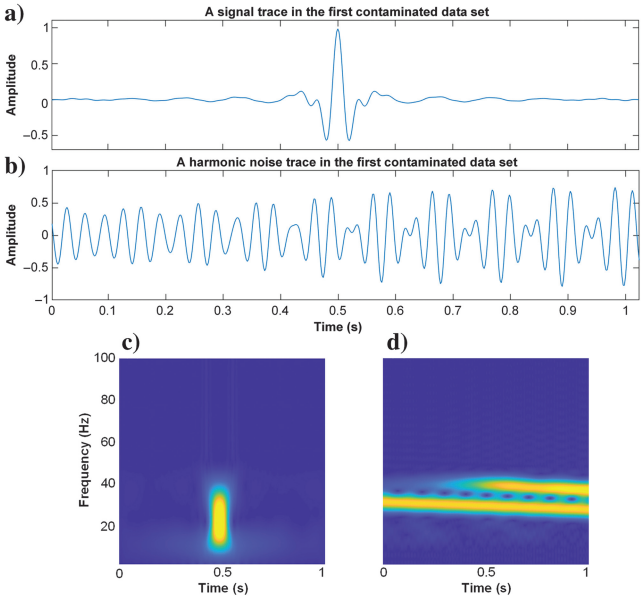

Figure 4. Traces used for fixed dictionary identification. (a) A signal trace extracted from the first data set: 96th–607th samples of the 165th trace in Figure 3b; (b) a harmonic noise trace extracted from the first data set: the last 512 samples of the 75th trace in Figure 3c; and (c and d) time-frequency spectra of (a and b), respectively.

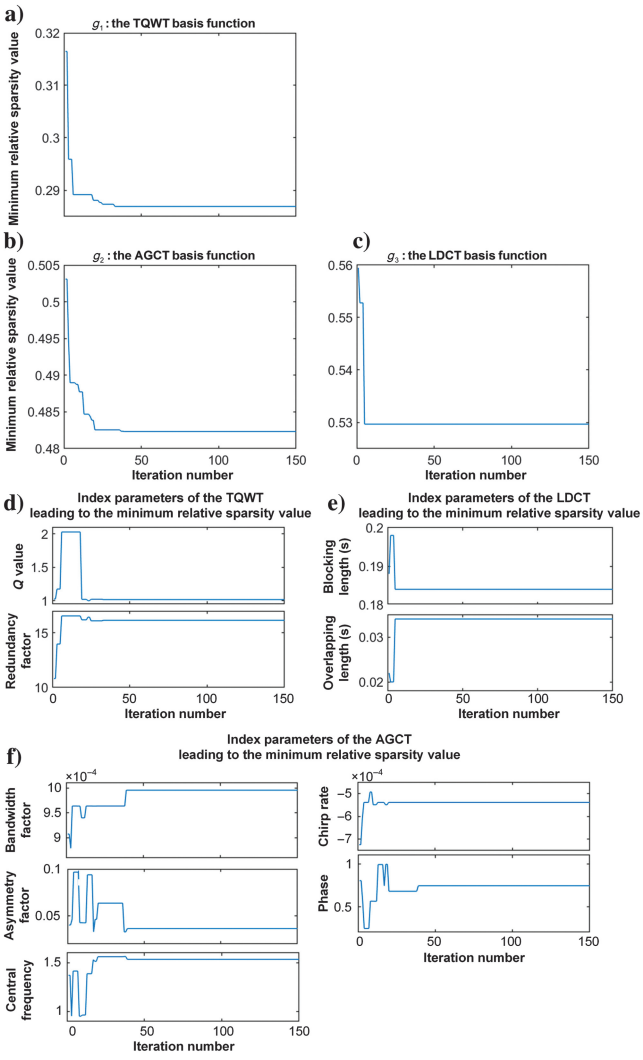

Figure 5. Demonstration of the application of the designed GA when constructing a fixed dictionary for signal in the first synthetic data set. The ratio of the sparsity of the signal trace (Figure 4a) to that of the harmonic noise trace (Figure 4b) is the relative sparsity function. (a–c) Minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) index parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

Table 2. Comparison of the optimized dictionaries representing signal and harmonic noise in the first synthetic data set (Figure 3a).

Dictionary	TQWT	AGCT	LDCT
Optimized parameter index for signal (Figure 4a)	(1.02, 16.15)	$(9.96\times10^{-4}, 0.04, 1.53, -5.39\times10^{-4}, 0.74)$	(0.18, 0.03)
Relative sparsity value	**0.2869**	0.4823	0.5297
Optimized parameter index for harmonic noise (Figure 4b)	(2.79, 3.13)	$(8.94\times10^{-5}, 0.83, 0.44, -9.20\times10^{-5}, 4.10)$	(1.02, 0.20)
Relative sparsity value	1.3339	**0.4109**	0.7212

The bold values represent the smallest relative sparsity value of the signal or coherent noise.

respectively. Their time-frequency spectra are shown in Figure 7c and 7d. Comparing Figure 7c and 7d with Figure 4c and 4d reveals that the TQWT and AGCT atoms (Figure 7a and 7b) display a time-frequency energy distribution similar to that of the signal and harmonic noise trace (Figure 4a and 4b). These findings suggest that the TQWT (1.02, 16.15) and AGCT (8.94×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10) match the signal and harmonic noise in the first data set, respectively. Therefore, they are the first pair of identified dictionaries.

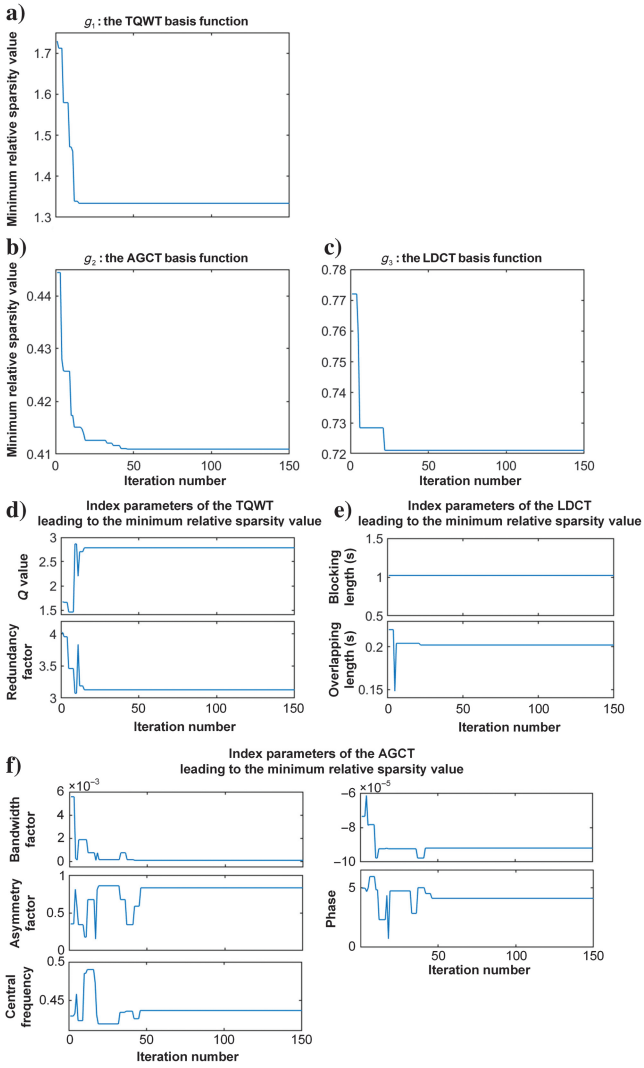

Figure 6. Demonstration of the application of the designed GA when constructing a fixed dictionary for harmonic noise in the first synthetic data set. The ratio of the sparsity of the harmonic noise trace (Figure 4b) to that of the signal trace (Figure 4a) is the relative sparsity function. (a–c) Minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) index parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

Dictionary identification for the second data set

The 91st–602nd samples of the 165th trace in Figure 3e are extracted as a signal trace (Figure 8a). The last 512 samples of the 75th trace in Figure 3f are extracted as a harmonic noise trace (Figure 8b). Figure 8e and 8f shows the time-frequency spectra of the signal trace (Figure 8a) and the harmonic noise trace (Figure 8b), respectively. Table 3 shows the index parameters of optimized AGCT, TQWT, and 1D LDCT, as well as relative sparsity values corresponding to them. The optimized TQWT (1.19, 13.48) leads to the smallest relative sparsity value of 0.2045 in representing a signal trace (Figure 8a). The optimized AGCT (2.68×10^{-5}, 0.31, 0.72, -9.98×10^{-5}, 0.61) leads to the smallest relative sparsity value of 0.3054 in representing a harmonic noise trace (Figure 8b). Figure 8c and 8d shows two atoms from the optimized TQWT and AGCT, respectively. Comparing Figure 8e and 8g with Figure 8f and 8h reveals

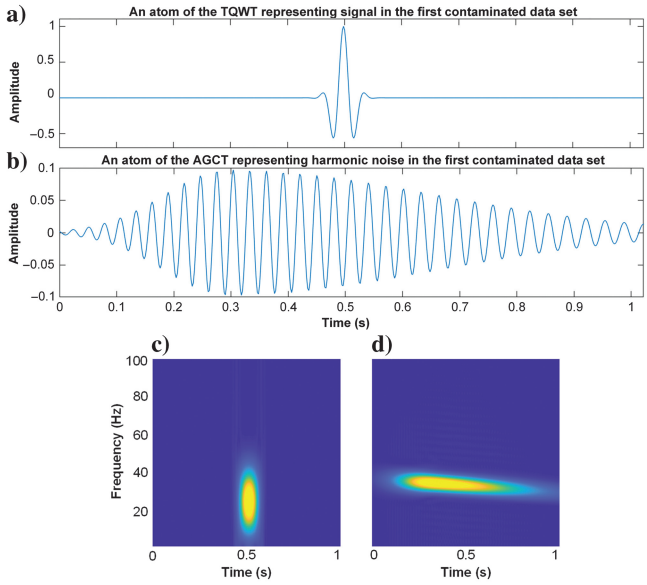

Figure 7. Demonstration of atoms in the optimized TQWT and AGCT representing signal and harmonic noise in the first data set (Figure 3a), respectively. (a) TQWT atom; (b) AGCT atom; and (c and d) time-frequency spectra of atoms shown in (a and b), respectively.

Table 3. Comparison of the optimized dictionaries representing signal and harmonic noise in the second synthetic data set (Figure 3d).

Dictionary	TQWT	AGCT	LDCT
Optimized parameter index for signal trace (Figure 8a)	(1.19, 13.48)	$(9.99 \times 10^{-4}, 3.4 \times 10^{-3}, 1.56, -5.49 \times 10^{-4}, 0.57)$	(0.17, 0.04)
Relative sparsity value	**0.2045**	0.5344	0.5500
Optimized parameter index for harmonic noise trace (Figure 8b)	(2.12, 3.02)	$(2.68 \times 10^{-5}, 0.31, 0.72, -9.98 \times 10^{-5}, 0.61)$	(0.99, 0.25)
Relative sparsity value	1.5960	**0.3054**	0.6580

The bold values represent the smallest relative sparsity value of the signal or coherent noise.

that the TQWT and AGCT atoms (Figure 8c and 8d) exhibit a time-frequency energy distribution similar to that of the signal and harmonic noise trace (Figure 8a and 8b). Therefore, the TQWT (1.19, 13.48) and AGCT (2.68×10^{-5}, 0.31, 0.72, -9.98×10^{-5}, 0.61) are used as the second pair of identified dictionaries.

Separation results

Figure 9a shows the separated signal obtained by applying the MCA using the first pair of dictionaries to the first data set (Figure 3a). Figure 9b shows the separated signal obtained by applying the MCA using the second pair of dictionaries to the second data set (Figure 3d). Figure 9c and 9d shows the separated signal obtained by applying the conventional method to the first and the second data set, respectively. The separated harmonic noise corresponding to the separated signal in Figure 9a–9d are shown in Figure 9e–9h. Figure 9a and 9b shows clearer signals than Figure 9c and 9d. In addition, Figure 9g shows less harmonic noise than Figure 9e around the position marked by a green arrow. Figure 9h shows signal leakage marked by a yellow arrow, whereas Figure 9f shows no visible signal leakage. Moreover, the S/Ns of the separated signal from the proposed method (26.57 and 26.68 dB) are significantly higher than those from the conventional method (10.21 and 11.35 dB). For further comparison, Figure 10a–10d shows the 135th trace of the original signal (four black curves), the separated signal from the proposed method (two red curves), and the conventional method (two purple curves), respectively. Figure 10e and 10f shows the difference between the red curve and the black curve in Figure 10a and 10b, respectively. Figure 10g and 10h shows the difference between the purple curve and the black curve in Figure 10c and 10d, respectively. Comparing Figure 10e and 10f with Figure 10g and 10h, it is observed that the two red curves are closer to the two black curves than the two purple curves. More specifically, Figure 11a and 11b shows time-frequency spectra of the 135th trace in the first and second data sets, respectively. Some parts of harmonic noise energy are marked by two gray arrows, and two red arrows mark some parts of signal energy. The time-frequency spectra of the 135th trace in separation results from the proposed method are shown in Figure 11c–11f. Figure 11g–11j shows the time-frequency spectra of the 135th trace in separation results from the conventional method. There is no visible remaining harmonic noise in Figure 11c and 11d. Figure 11g and 11h shows remaining harmonic noise marked by two gray arrows. Figure 11i and 11j shows signal leakage marked by two red arrows, whereas Figure 11e and 11f shows no visible signal leakage.

These results prove that the MCA using the identified dictionaries has better performance in separating synthetic harmonic noise and preserving synthetic signal than the conventional method. In addition, the proposed identification framework flexibly finds the

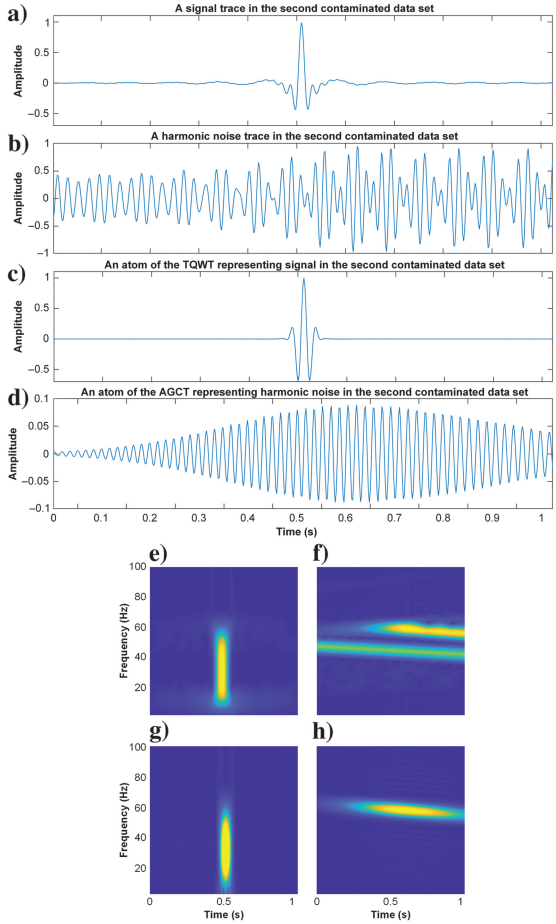

Figure 8. Demonstration of signal trace and harmonic noise trace contained in the second data set, as well as the atoms in the optimized TQWT and AGCT representing them. (a) A signal trace extracted from the second data set: 91st–602nd samples of the 165th trace in Figure 3e. (b) A harmonic noise trace extracted from the second data set: the last 512 samples of the 75th trace in Figure 3f; (c) TQWT atom; (d) AGCT atom; and (e–h) time-frequency spectra of data shown in (a–d), respectively.

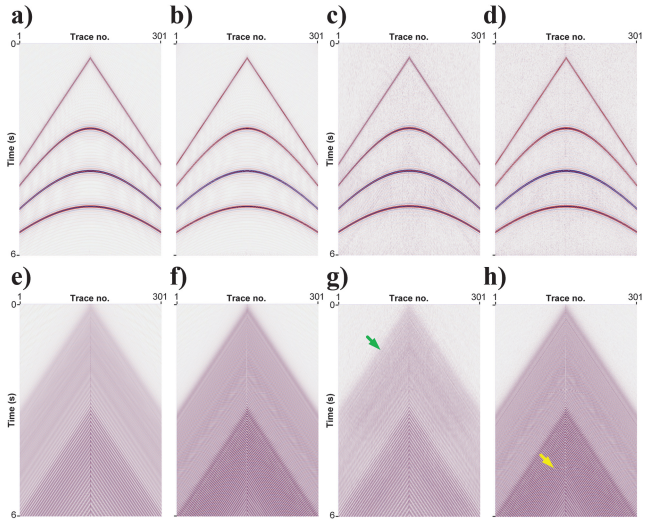

Figure 9. Separated results obtained by applying the proposed method to the first and second data set (Figure 3a and 3d), respectively. (a and e) Separated signal (26.57 dB) and harmonic noise obtained by applying the proposed method to the first data set (Figure 3a), respectively; (b and f) separated signal (26.68 dB) and harmonic noise obtained by applying the proposed method to the second data set (Figure 3d), respectively; (c and g) separated signal (10.21 dB) and harmonic noise obtained by applying the conventional method to the first data set (Figure 3a), respectively; and (d and h) separated signal (11.35 dB) and harmonic noise obtained by applying the conventional method to the second data set (Figure 3d), respectively.

dictionaries representing different signals and harmonic noise among the given fixed dictionary library.

Second synthetic data example

In this section, to assess the impact of signal complexity and noise intensity on the proposed method's separation performance, we use more complex signals and varying noise levels. Notably, when noise has high energy and overlaps with signal within the given data set, we first apply the conventional method to obtain signal-predominated and noise-predominated data, then use the proposed framework for dictionary identification. The high-energy noise attenuation method with the Coiflet wavelet is used as the conventional method.

The synthetic signal and the synthetic harmonic noise are shown in Figure 12a and 12b, respectively. The synthetic signal (Figure 12a) is generated by a two-layer geologic model described by Hu et al. (2022). It consists of direct waves (indicated by a green arrow) and refractions (indicated by two black arrows). The two-layer geologic model uses 30 Hz-Ricker wavelet as an explosive point source. The first layer has a velocity of 800 m/s and a thickness of 80 m. The second layer has a velocity of 2500 m/s and a thickness of 300 m. The layer density is computed by the Gardner formula (Gardner et al., 1974). The synthetic harmonic noise (Figure 12b) is generated by the sweep signal with frequencies increasing linearly from 10 to 30 Hz. Figure 12c and 12d shows two contaminated data sets with S/Ns of −0.70 and −16.26 dB, respectively, which are two different combinations of signal and harmonic noise (Figure 12a and 12b). Each data set has 401 traces and 1001 samples per trace. The time sampling is 0.002 s and the trace interval is 20 m.

Dictionary identification

To obtain two traces predominantly containing signal and harmonic noise, we apply the conventional method to the 195th trace in the second contaminated data set, as shown in Figure 13a.

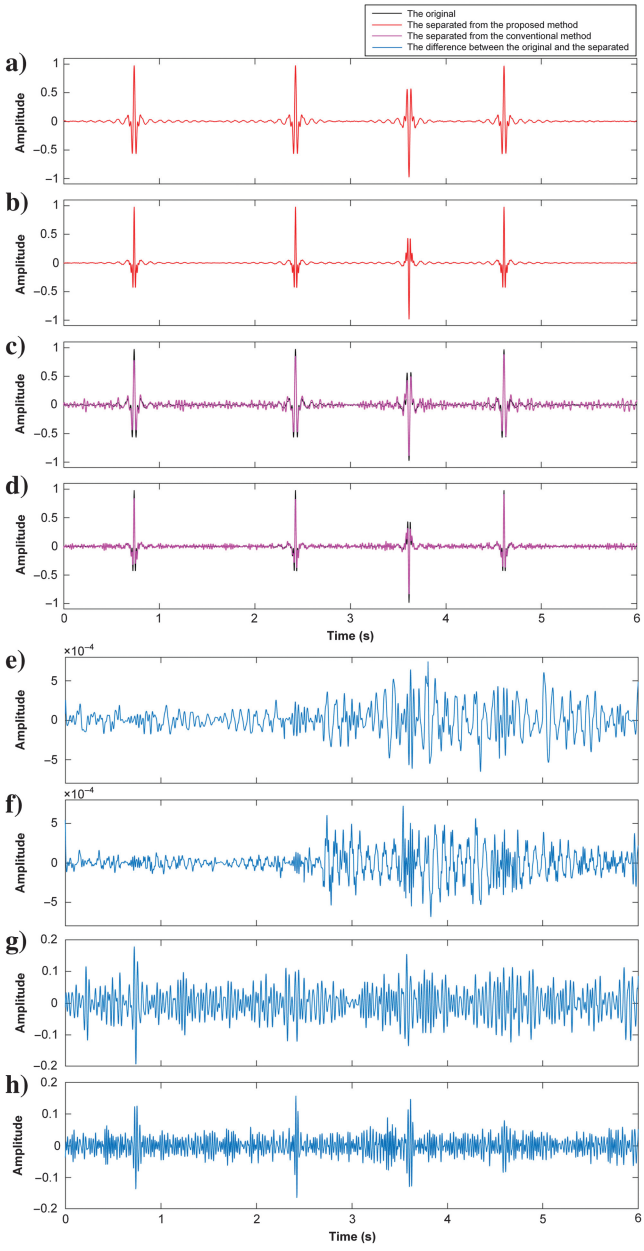

Figure 10. Comparison of the 135th trace in the synthetic signal (Figure 3b and 3e) and separated results from the proposed method (Figure 9a–9h). (a and b) Two black curves represent the original signals, and two red curves represent separated signals from the proposed method; (c and d) two black curves represent the original signals, and two purple curves represent separated signals from the conventional method; and (e–h) the differences between the original signal and the separated signal in (a–d), respectively.

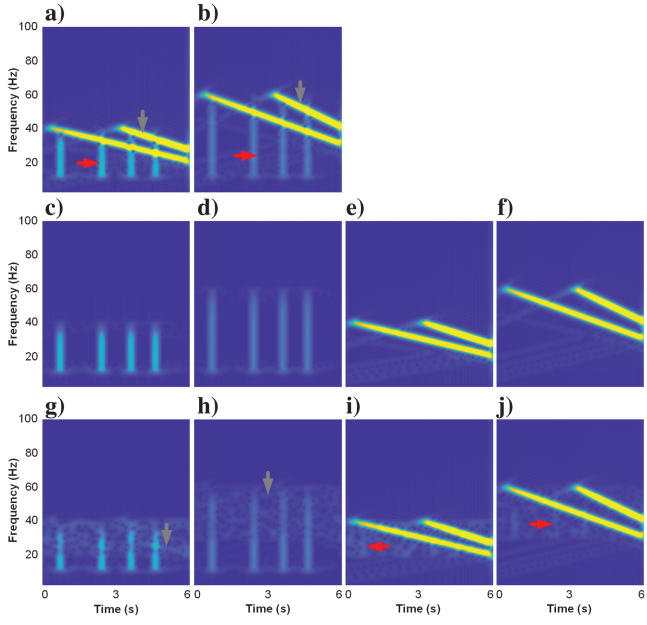

Figure 11. Comparison of time-frequency spectra of the 135th trace in the first and second data set (Figure 3a and 3d) and separated results (Figure 9a–9h). Some parts of harmonic noise energy are marked by two gray arrows and some parts of signal energy are marked by two red arrows in (a and b). (a and b) Time-frequency spectra of original traces; (c–f) time-frequency spectra of the signal and harmonic noise separated from original traces by the proposed method, respectively; and (g–j) time-frequency spectra of the signal and harmonic noise separated from original traces by the conventional method, respectively.

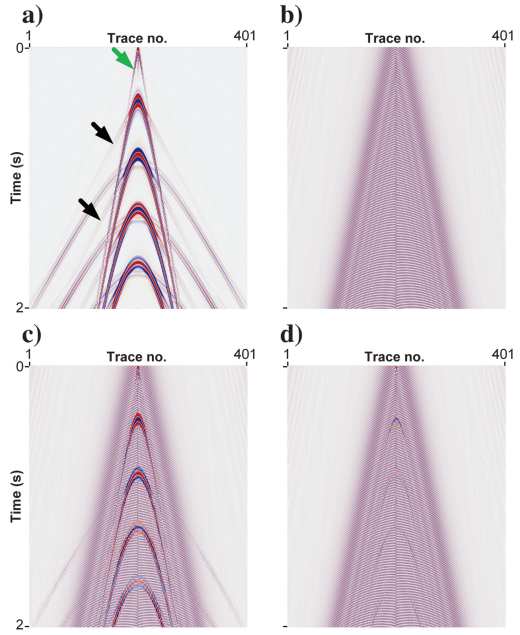

Figure 12. The second synthetic data example contaminated with harmonic noise. The trace interval is 20 m. (a) Signal comprising primaries (a green arrow), refractions (two black arrows), and reflection; (b) harmonic noise, (c) the first contaminated data set (S/N = −0.70 dB) consisting of signal in (a) and harmonic noise in (b), and (d) the second contaminated data set (S/N = −16.26 dB) consisting of signal in (a) and harmonic noise in (b).

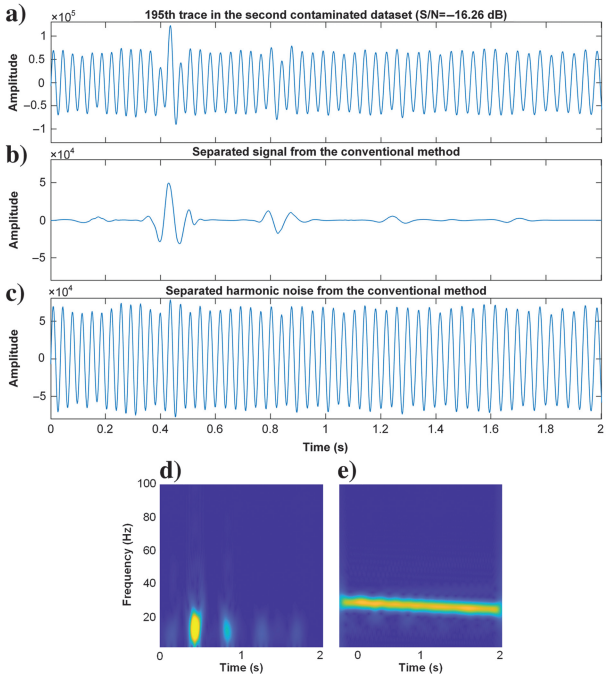

Figure 13. Demonstration of the separated signal and harmonic noise traces obtained by the conventional method, which are used for dictionary identification. (a) The 195th trace in the second contaminated data set; (b and c) the separated signal and harmonic noise obtained by applying the conventional method to the data in (a); and (d and e) time-frequency spectra of the separated signal and harmonic noise, respectively.

Applying to other contaminated traces containing relatively complete signals is also fine. The separated signal and harmonic noise from the conventional method (Figure 13b and 13c) are used for dictionary identification. Figure 14a–14c shows the evolution of minimum relative sparsity values in the GA, which identifies the optimized TQWT, AGCT, and 1D LDCT to represent the separated signal (Figure 13b). Figure 15a–15c shows the evolution of minimum relative sparsity values in the GA, which identifies the optimized TQWT, AGCT, and 1D LDCT to represent the separated harmonic noise (Figure 13c). These minimum relative sparsity values and corresponding index parameters are shown in Figure 14d–14f and Figure 15d–15f. After 50 iterations, the minimum relative sparsity values and index parameters stabilize at specific values. Table 4 shows these specific values. The optimized TQWT (1.01, 7.39) leads to the smallest relative sparsity value, 0.3329, in representing the

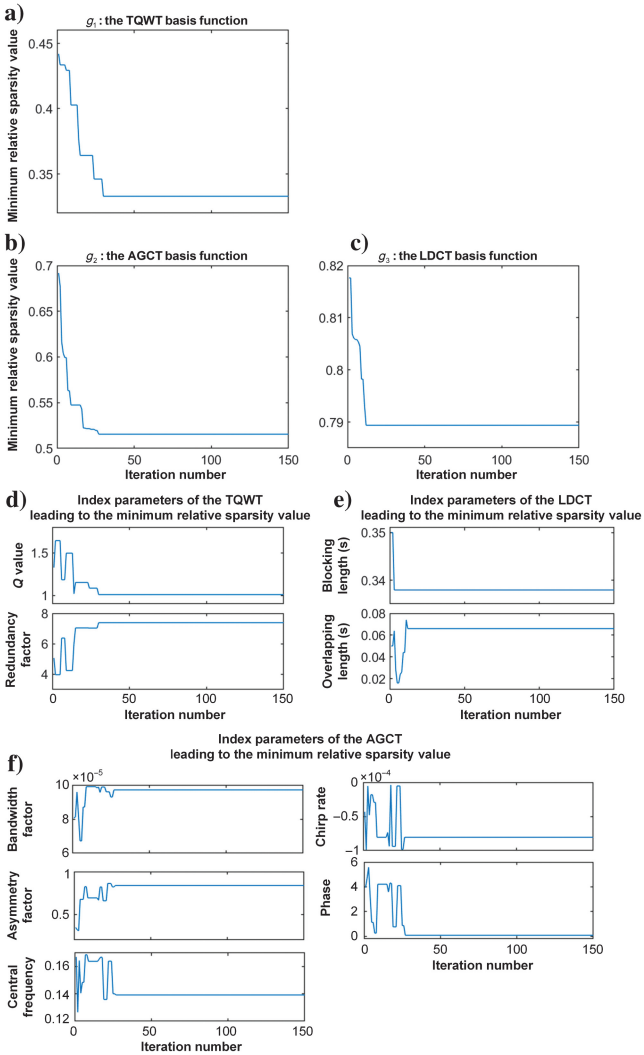

Figure 14. Demonstration of the application of the designed GA when constructing a fixed dictionary for signal. The ratio of the sparsity of the separated signal trace (Figure 13b) to that of the separated harmonic noise trace (Figure 13c) is the relative sparsity function. (a–c) Minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) index parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

separated signal. In representing the separated harmonic noise, the optimized AGCT (8.54×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10) leads to the smallest relative sparsity value, 0.4109. Therefore, the TQWT (1.01, 7.39) and AGCT (8.54×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10) are used as the dictionaries representing signal and harmonic noise, respectively. The MCA using them is the proposed method.

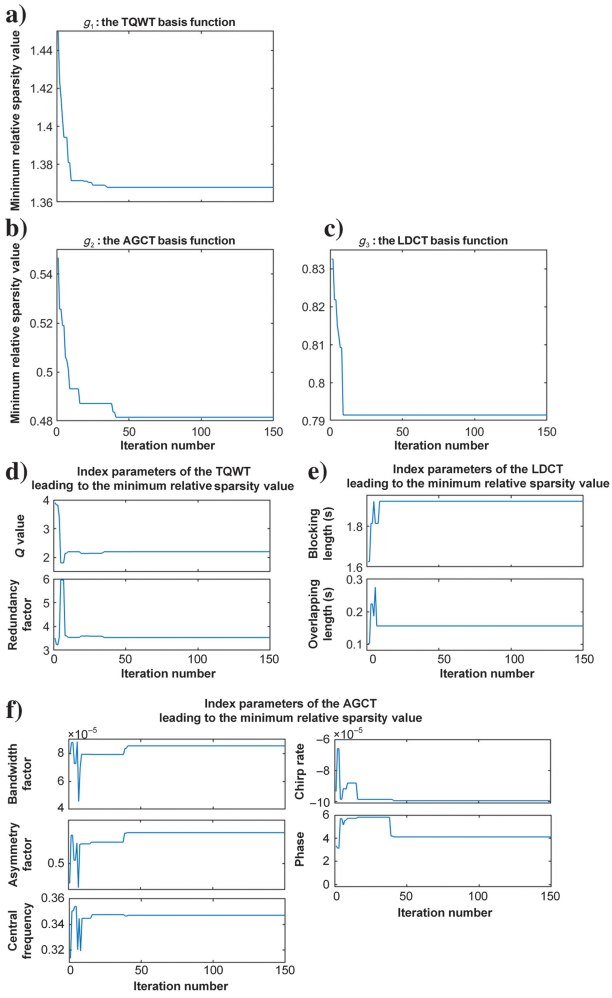

Figure 15. Demonstration of the application of the designed GA when constructing a fixed dictionary for harmonic noise. The ratio of the sparsity of the harmonic noise trace (Figure 13c) to that of the signal trace (Figure 13b) is the relative sparsity function. (a–c) Minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) index parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

Figure 16a and 16b shows an atom in the TQWT (1.02, 16.15) and an atom in the AGCT (8.54×10^{-5}, 0.83, 0.44, -9.20×10^{-5}, 4.10), whose time-frequency spectra are shown in Figure 16c and 16d. Comparing Figure 16c and 16d with Figure 13d and 13e reveals that the TQWT atom and AGCT atom (Figure 16a and 16b) display a time-frequency energy distribution similar to that of the separated signal and harmonic noise (Figure 13b and 13c).

Separation results

Figure 17a and 17b shows the separated signal with S/Ns of 28.47 and 22.38 dB obtained by applying the proposed method to the first and second contaminated data sets (Figure 12c and 12d). Figure 17c and 17d shows the separated signal with S/Ns of 9.30 and 5.13 dB obtained by applying the conventional method to the first and second contaminated data sets, respectively. The separated harmonic noise corresponding to the separated signal shown in Figure 17a–17d is shown in Figure 17e–17h. The separated signal from the proposed method has a much higher S/N than that from the conventional method. In addition, Figure 17a and 17b shows no

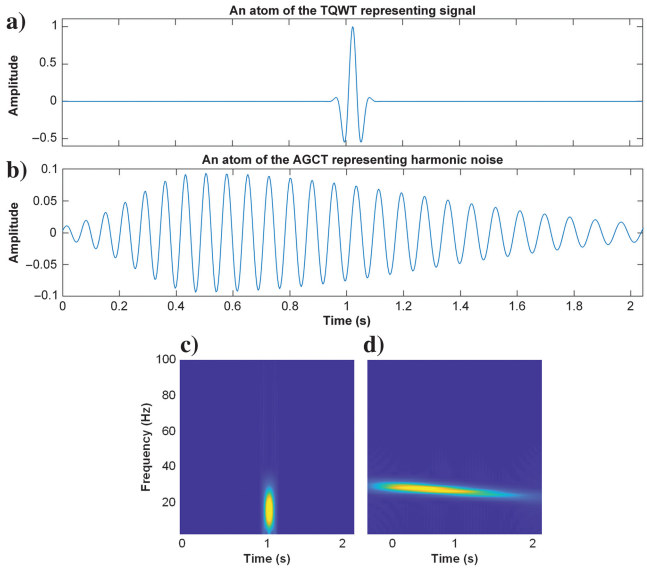

Figure 16. Demonstration of two atoms in the TQWT representing the signal and the AGCT representing harmonic noise. (a and b) TQWT atom and AGCT atom, respectively; and (b and d) time-frequency spectra of atoms shown in (a and b), respectively.

Table 4. Comparison of the optimized dictionaries representing signal and harmonic noise in two synthetic data sets (Figure 12c and 12d).

Dictionary	TQWT	AGCT	LDCT
Optimized parameter index for signal trace (Figure 13b)	(1.01, 7.39)	(9.71×10^{-5}, 0.84,0.14, -8.05×10^{-5}, 0.08)	(0.34, 0.07)
Relative sparsity value	**0.3329**	0.5155	0.7894
Optimized parameter index for harmonic noise trace (Figure 13c)	(2.20, 3.54)	(8.54×10^{-5}, 0.83,0.44, -9.20×10^{-5}, 4.10)	(1.02, 0.20)
Relative sparsity value	1.3679	**0.4815**	0.7212

The bold values represent the smallest relative sparsity value of the signal or coherent noise.

visible signal damage but Figure 17d shows signal damage marked by two black arrows. Figure 17e and 17f shows no visible signal leakage but Figure 17g shows signal leakage marked by two green arrows.

These results indicate that the MCA using the dictionaries identified by the proposed framework achieves better separation performance than the conventional method. Furthermore, applying the conventional method to contaminated traces is an effective way to obtain the signal- or noise-predominated data used in our framework.

Field data example

We exemplify the proposed approach using a shot data set displayed in Figure 18, where some parts of the signal are marked by a yellow arrow and some parts of harmonic noise energy are marked by two green arrows. The data set has 816 traces and 3000 samples per trace. The time sampling is 0.002 s, and the trace interval is 20 m.

Dictionary identification

We extract 200th–711th samples of the 101st trace (the first black arrow in Figure 18) as a signal-predominated trace (Figure 19a). The last 512 samples of the 184th trace (the second black arrow in Figure 18) are extracted as the harmonic-noise-predominated trace (Figure 19b). Their time-frequency spectra are shown in Figure 19c and 19d. Figure 20a–20c shows the evolution of minimum relative sparsity values in the GA, which identifies parameter indices of the optimized TQWT, AGCT, and 1D LDCT representing field signal (Figure 19a). Figure 21a–21c shows the evolution of minimum relative sparsity values in the GA, which identifies parameter indices of the optimized TQWT, AGCT, and 1D LDCT representing field harmonic noise (Figure 19b). These minimum relative sparsity values and corresponding index parameters are shown in Figure 20d–20f and Figure 21d–21f. After 75 iterations,

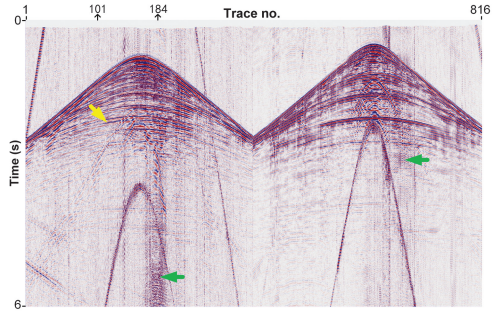

Figure 18. Two-shot data sets contaminated with the harmonic noise, respectively, where some parts of the signal are marked by a yellow arrow and some parts of harmonic noise are marked by two green arrows. The trace interval is 20 m. The 101st trace, marked by the first black arrow, mainly contains a signal above 3 s. The 184th trace, marked by the second black arrow, mainly contains harmonic noise below 3 s.

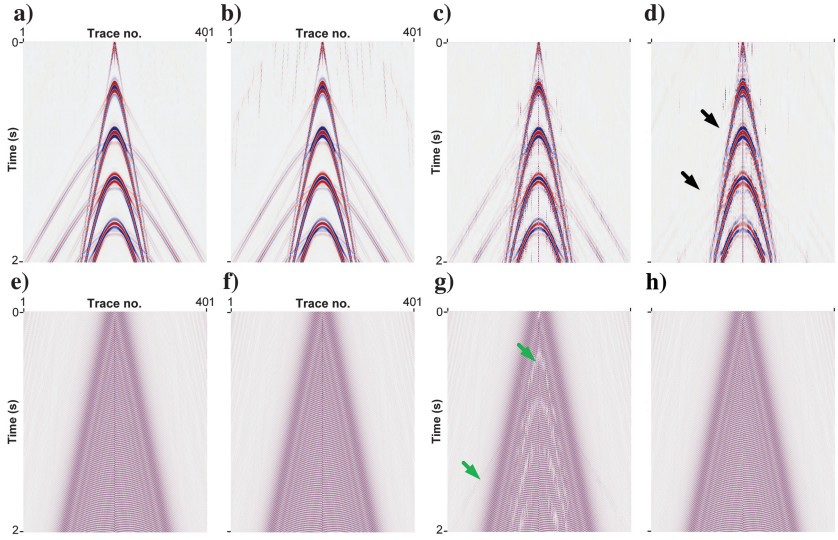

Figure 17. Separation results: two black arrows mark the signal damage, and two green arrows mark the signal leakage. (a and b) The separated signal with S/Ns of 28.47 dB and 22.38 dB obtained by applying the proposed method to the first and second contaminated data sets, respectively; (c and d) the separated signal with S/Ns of 9.30 and 5.13 dB obtained by applying the conventional method to the first and second contaminated data sets, respectively; and (e–h) the separated harmonic noise corresponding to the separated signal shown in (a–d), respectively.

Table 5. Comparison of the optimized dictionaries representing field signal and harmonic noise in the field data set (Figure 18).

Dictionary	TQWT	AGCT	LDCT
Optimized parameter index for signal trace (Figure 19a)	(2.36, 3.16)	$(9.74\times10^{-4}, 0.042, 0.33, -9.80\times10^{-4}, 0.60)$	(0.19, 0.03)
Relative sparsity value	**0.4810**	0.5083	0.6551
Optimized parameter index for harmonic noise trace (Figure 19b)	(2.6, 6.04)	$(1.82\times10^{-4}, 0.11, 0.84, -2.58\times10^{-4}, 0.14)$	(1.01, 0.25)
Relative sparsity value	0.9321	**0.4173**	0.9444

The bold values represent the smallest relative sparsity value of the signal or coherent noise.

the minimum relative sparsity values and index parameters stabilize at specific values. Table 5 shows these specific values. The optimized TQWT (2.36, 3.16) leads to the smallest relative sparsity value, 0.4810, in representing the signal-predominated trace. In representing the harmonic-noise-predominated trace, the optimized AGCT (1.82×10^{-4}, 0.11, 0.84, -2.58×10^{-4}, 0.14) leads to the smallest relative sparsity value, 0.4173. Therefore, the TQWT (2.36, 3.16) and AGCT (1.82×10^{-4}, 0.11, 0.84, -2.58×10^{-4}, 0.14) are the identified dictionaries representing signal and harmonic noise, respectively. The MCA using them is the proposed method.

Two atoms of the TQWT (2.36, 3.16) and the AGCT (1.82×10^{-4}, 0.11, 0.84, -2.58×10^{-4}, 0.14) are shown in Figure 22a and 22b. Their time-frequency spectra are shown in Figure 22c and 22d. Comparing Figure 22c and 22d with Figure 19c and 19d reveals that the TQWT and AGCT atom (Figure 22a and 22b) display a time-frequency energy distribution similar to that of the signal- and harmonic-noise-predominated trace (Figure 19a and 19b). These results suggest that the TQWT (2.36, 3.16) and the AGCT (1.82×10^{-4}, 0.11, 0.84, -2.58×10^{-4}, 0.14) match field signal and harmonic noise, respectively.

Separation results

Figure 23a and 23c shows the separated signal and separated harmonic noise from the proposed method. Figure 23b and 23d shows the separated signal and separated harmonic noise from the conventional method. Figure 23a shows a clearer signal than Figure 23b, especially in a green rectangle. In addition, Figure 23b shows remaining harmonic noise marked by two black arrows. Figure 23c exhibits no visible signal leakage, whereas two green arrows in Figure 23d mark signal leakage. For further comparison, Figure 24 shows the magnified view of local data at the position marked by a green rectangle in Figure 23a. Figure 24a shows the original data set. Figure 24b and 24d shows separated signal and harmonic noise from the proposed method. The separated signal and harmonic noise from the conventional method are shown in Figure 24c and 24e, respectively. There remains harmonic noise in Figure 24c (a green arrow) and signal leakage in Figure 24e (a black arrow). For a detailed comparison, the 671st trace in the original data set (Figure 18) and the separated signal from the proposed and the conventional method (Figure 23a and 23b) are shown in Figure 25a–25c. Figure 25d and 25e shows the 671st trace in separated harmonic noise from the proposed and the conventional method (Figure 23c and 23d). Their time-frequency spectra are shown in Figure 25f–25j, respectively. Figure 25b exhibits a much

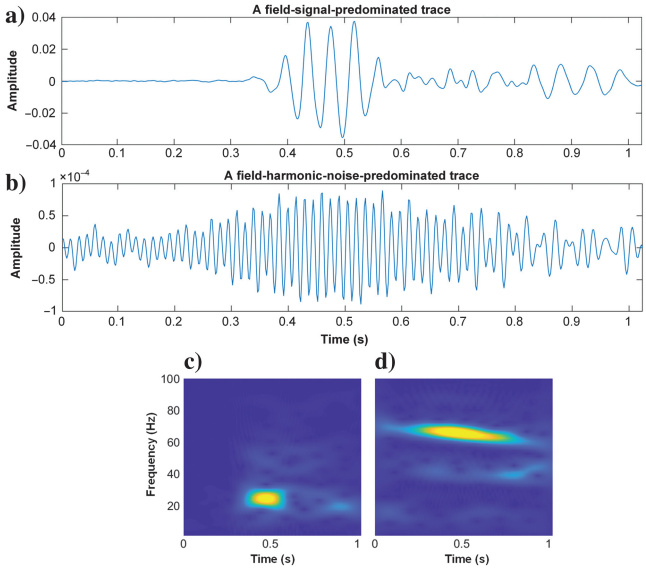

Figure 19. Traces used for fixed dictionary identification. (a) A signal-predominated trace: 200th–711th samples of the 101st trace marked by a black arrow in the original data set (Figure 18); (b) a harmonic-noise-predominated trace: the last 512 samples of the 184th trace marked by a green arrow in the original data set (Figure 18); and (c and d) time-frequency spectra of (a and b), respectively.

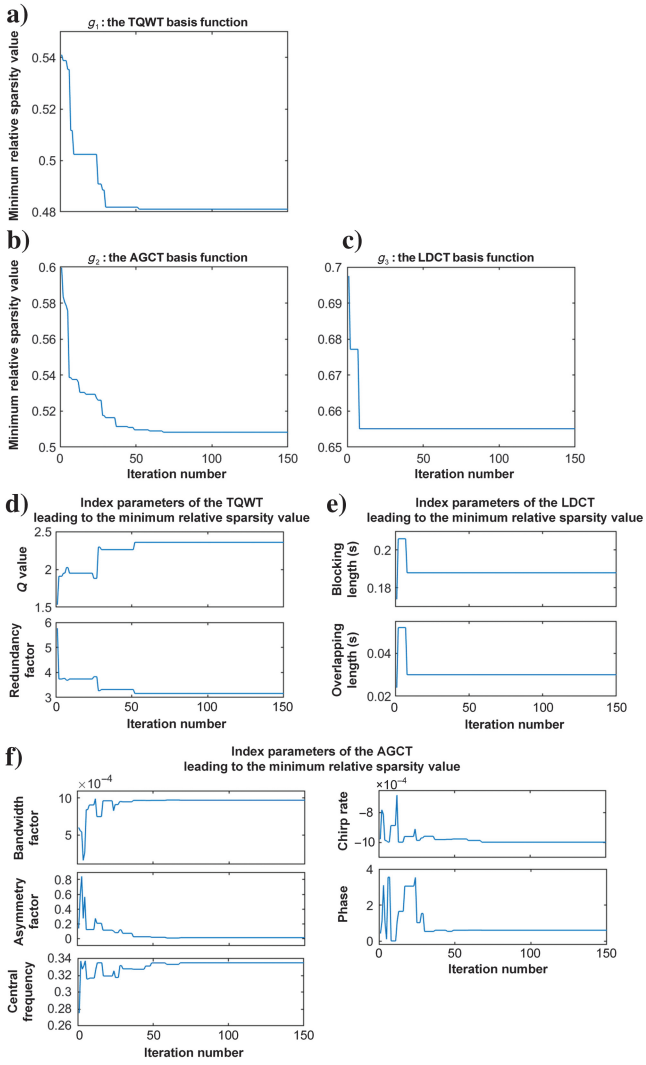

Figure 20. Demonstration of the application of the designed GA when constructing a fixed dictionary for field signal. The ratio of the sparsity of signal-predominated trace (Figure 19a) to that of harmonic-noise-predominated data (Figure 19b) is the relative sparsity function. (a–c) The minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

clearer signal than Figure 25c, whereas Figure 25d shows more harmonic noise than Figure 25e. Correspondingly, Figure 25h shows remaining harmonic noise marked by a gray arrow. These results indicate that the proposed method is more effective than the conventional method in separating field harmonic noise. Furthermore, the proposed framework contributes to the identification of effective dictionaries representing signal and harmonic noise in the field data set.

DISCUSSION

This study aims to develop and evaluate an adaptive identification dictionary framework. The proposed method is applied to 1D traces for MCA-based noise separation and identifies dictionaries from a library that includes AGCTs, TQWTs, and 1D LDCTs. As shown by synthetic and field data examples, the effectiveness of our approach is highlighted by its robustness across diverse noise levels, ranging from −16.26 to −0.70 dB, and its performance with noise across different frequency ranges, specifically from 10 to 80 Hz.

There are some aspects for further research. Because many other types of dictionaries and coherent noise exist, the proposed method

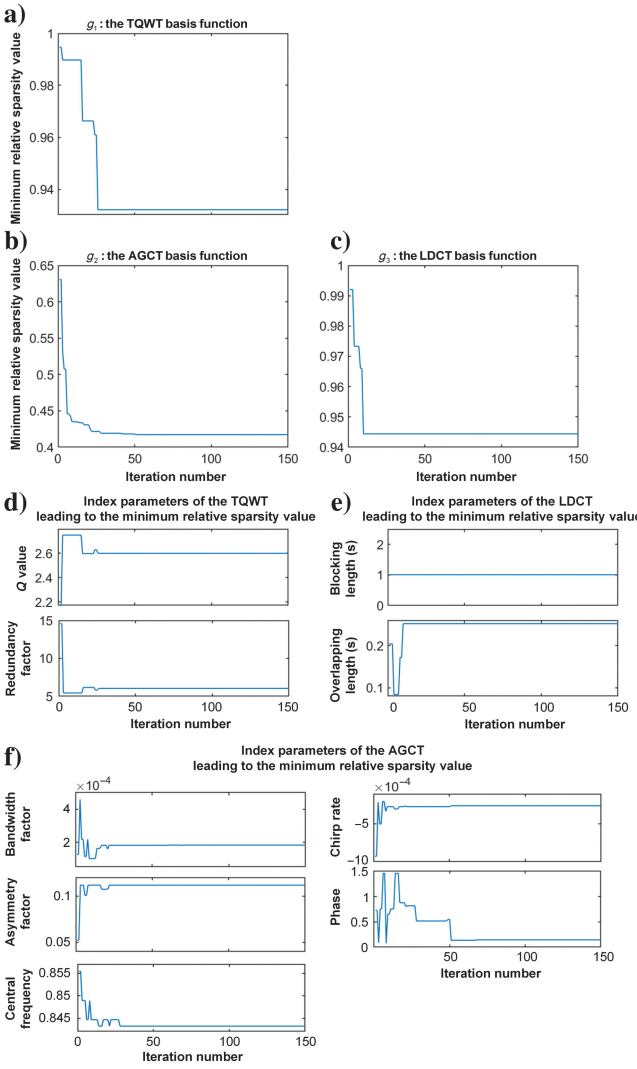

Figure 21. Demonstration of the application of the designed GA when constructing a fixed dictionary for field harmonic noise. The ratio of the sparsity of harmonic-noise-predominated data (Figure 19b) to that of signal-predominated trace (Figure 19a) is the relative sparsity function. (a–c) Minimum relative sparsity values in iterations of the GA applied on the TQWT, AGCT, and LDCT basis functions, respectively; and (d–f) parameters leading to these minimum relative sparsity values shown in (a–c), respectively.

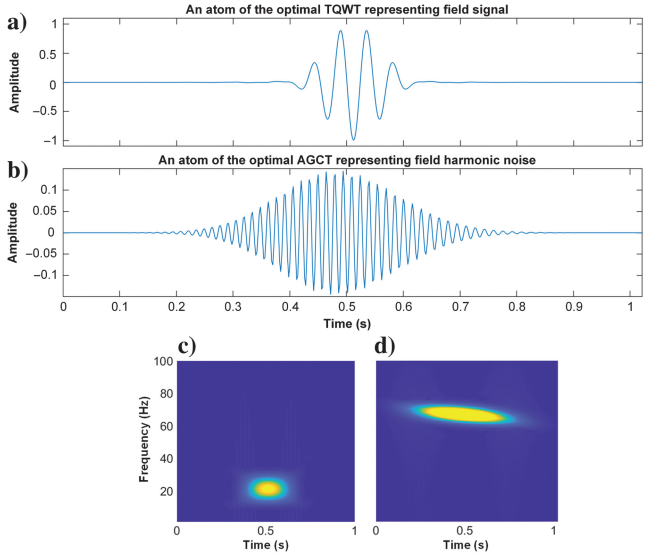

Figure 22. Demonstration of atoms in the optimized TQWT and AGCT representing field signal and harmonic noise, respectively. (a and b) TQWT atom and AGCT atom, respectively; and (c and d) time-frequency spectra of atoms shown in (a and b), respectively.

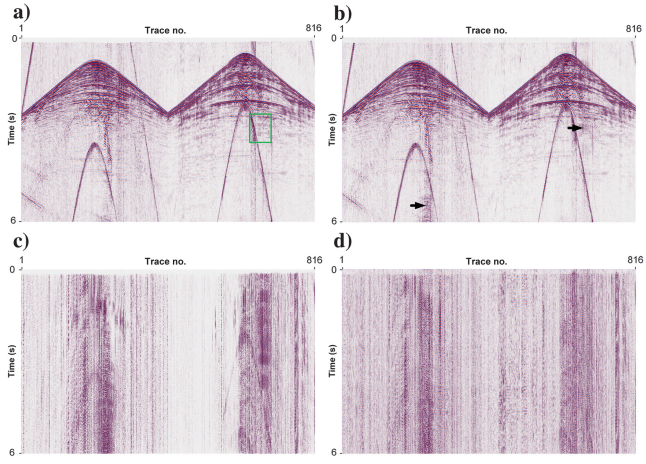

Figure 23. Separated results obtained by the proposed and conventional methods to the contaminated data set (Figure 18). (a and c) Separated signal and harmonic noise from the proposed method, respectively. The green rectangle marks an area with a clearer signal than in (b); and (b and d) the separated signal and harmonic noise from the conventional method, respectively. The two black arrows mark the areas with remaining harmonic noise, whereas two green arrows mark the areas with signal leakage.

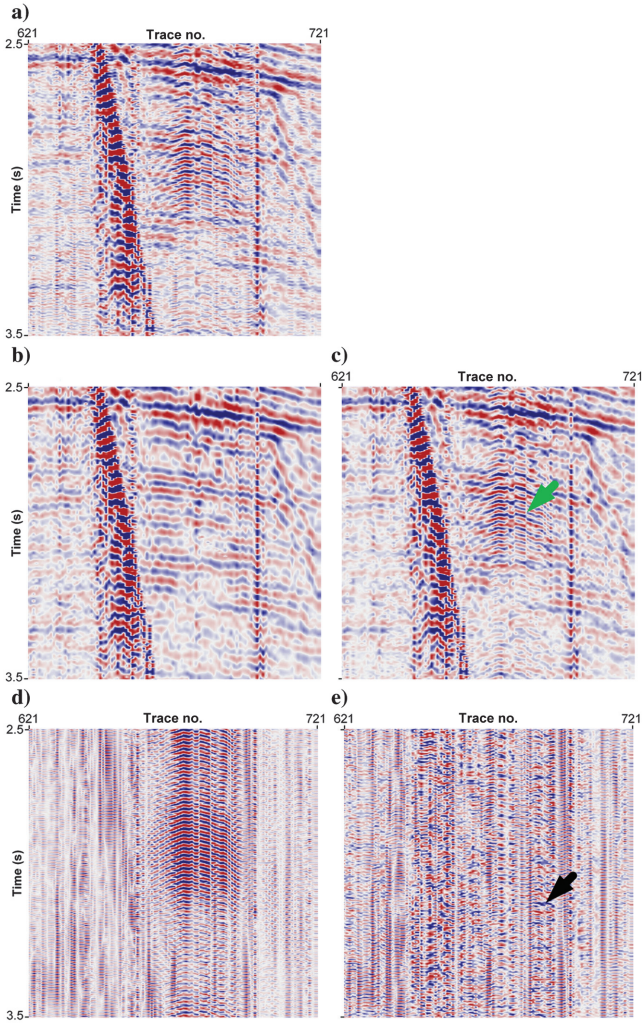

Figure 24. The magnified view of the local data set at the position marked by a green rectangle in Figure 23a. (a) Original data set; (b and d) separated signal and harmonic noise from the proposed method, respectively; and (c and e) separated signal and harmonic noise from the conventional method, respectively. The remained a green arrow marks harmonic noise, and a black arrow marks the signal leakage.

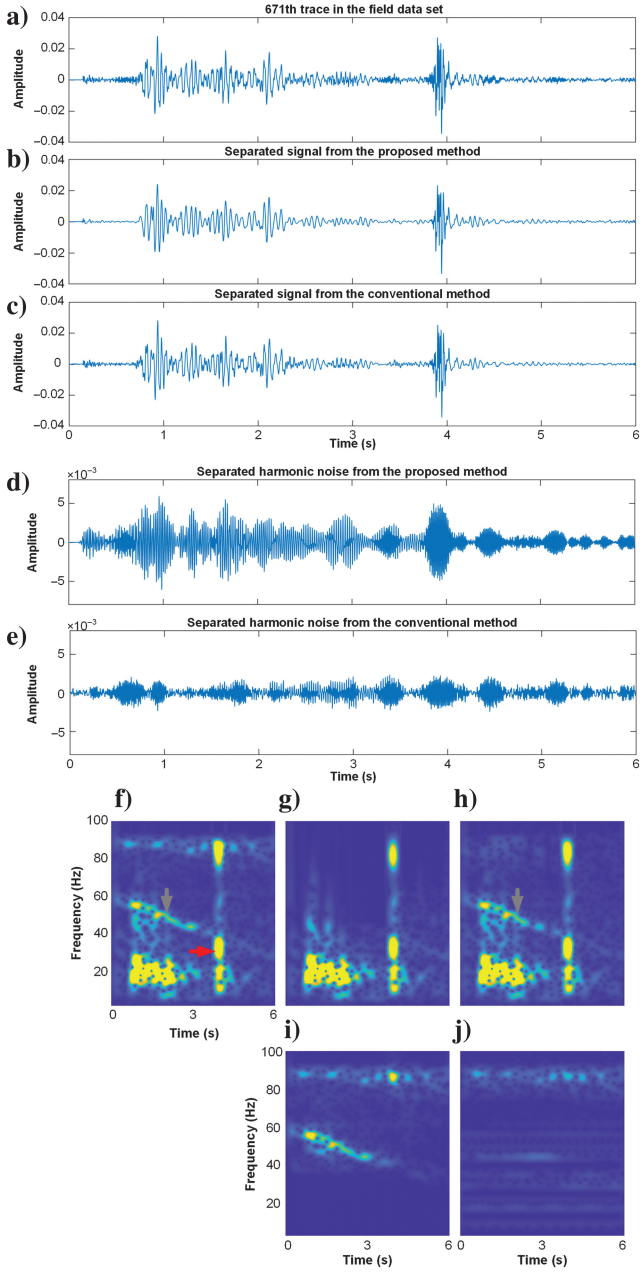

Figure 25. Comparison of the 671st trace of the original data set (Figure 18) and the separated results (Figure 23a–23d). (a–c) Original trace, the separated signal from the proposed and the conventional method, respectively; (d and e) the separated harmonic noise from the proposed and the conventional method, respectively; (f–j) time-frequency spectra of the traces shown in (a–e), respectively. (f) Some parts of the harmonic noise energy are marked by a gray arrow, whereas a red arrow marks the signal energy; and (h) the remained a gray arrow marks harmonic noise energy.

can offer potential improvements in the separation of other coherent noise using the library with more dictionary types. In addition, the MCA-based coherent noise separation can handle 2D or higher-dimensional data sets. By using higher-dimensional dictionaries to a library, this method can be integrated into higher-dimensional seismic data processing workflows. Furthermore, in the cases where signal and coherent noise in different data vary greatly, using the proposed framework to re-identify the fixed dictionary based on local data would be beneficial.

However, the computational demand increases with the dimensionality of data and dictionaries, as well as with the number of re-identifications, which potentially limits the applicability of the proposed method in real-time processing scenarios. Thus, solving the RSMP more efficiently by leveraging existing parameter optimization methods, such as particle swarm optimization (Nuha et al., 2021), is an important direction for future work.

CONCLUSION

We propose an innovative adaptive framework to construct effective fixed dictionaries for MCA-based coherent noise separation. The proposed framework consolidates diverse transform basis functions and parameters into a fixed dictionary library. By using a predefined fixed dictionary library as the solution space and a relative

sparsity function as the objective function, we present a systematic searching strategy to identify the fixed dictionaries representing signal and coherent noise, respectively. The proposed strategy entails characterizing a constrained RSMP within the predefined library, allowing individual optimization of the parameters of each type of fixed dictionary. A specially designed GA facilitates this optimization process. Subsequently, from the optimal dictionaries constructed using optimal parameter indices and corresponding transform basis functions, we select the dictionary with the lowest relative sparsity to represent either signal or coherent noise. The effectiveness of the proposed framework is demonstrated through its application to the separation of harmonic noise. Separation results highlight the superior performance of the proposed method, achieving enhanced coherent noise separation while effectively preserving signal.

ACKNOWLEDGMENTS

This work was supported in part by the National Key Research and Development Program of China under grant 2021YFA0716903, by the National Natural Science Foundation of China under grants 42374135 and 41974131, and by the Fundamental Research Funds for the Central Universities (xzy012023073). The authors would like to express their sincere gratitude to the editors and anonymous reviewers for their valuable insights and constructive suggestions.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

APPENDIX A

TUNABLE Q-FACTOR WAVELET TRANSFORM

The TQWT is a discrete wavelet transform (Selesnick, 2011) using J filter banks to decompose 1D data. Many scholars have demonstrated the effectiveness of TQWTs in representing seismic waves with varying degrees of oscillation (Chen et al., 2017; Hu et al., 2024). Each filter bank is calculated based on the TQWT basis function, such as $G_0(f)$ and $H_0(f)$, which are expressed as

$$\mathrm{H}_0(f) = \begin{cases} 1 & |f| \le (1-b_1)\pi \\ \theta\left(\frac{f+(b_1-1)\pi}{b_2+b_1-1}\right) & (1-b_1)\pi < |f| < b_2\pi \\ 0 & b_2\pi \le |f| \le \pi \end{cases}, \quad \text{(A-1)}$$

$$\mathrm{G}_0(f) = \begin{cases} 0 & |f| \le (1-b_1)\pi \\ \theta\left(\frac{b_2\pi-f}{b_2+b_1-1}\right) & (1-b_1)\pi < |f| < b_2\pi \\ 1 & b_2\pi \le |f| \le \pi \end{cases}, \quad \text{(A-2)}$$

where $\theta(f) = 0.5(1+\cos f)\sqrt{2-\cos f}, |f| \le \pi$, and b_1 and b_2 are two scaling parameters that are calculated by the following equations:

$$b_1 = 2/(Q+1) \in (0,1], \quad b_2 = 1 - b_1/r \in (0,1], \quad \text{(A-3)}$$

where Q is the Q-value affecting the oscillation degree of TQWT atoms, and r is the redundancy factor. Furthermore, r is the total number of wavelet coefficients divided by the length of the data to which the TQWT is applied (Selesnick, 2011). Thus, the TQWT basis function is calculated by Q and r. Here, J does not affect the morphological feature of TQWT atoms. Thus, we set J to 10 and only adjust Q and r. The term $\boldsymbol{\omega} = (Q, r)$ is the parameter index.

APPENDIX B

ASYMMETRIC GAUSSIAN CHIRPLET TRANSFORM

Among many versions of the chirplet transform (Baraniuk and Jones, 1993; Mann and Haykin, 1995), which serve as extensions of the short-time Fourier transform and the wavelet transform, the six-parameter AGCT (Boßmann and Ma, 2015) stands out due to its flexibility in matching seismic waves. The AGCT basis function is written as

$$AGCT(a,\beta,f,\gamma,\theta,\tau_0;t) = e^{-a(1-\beta\tanh(t-\tau_0))(t-\tau_0)^2} \cos(f(t-\tau_0)+\gamma(t-\tau_0)^2+\theta), \quad \text{(B-1)}$$

where t is the time, $AGCT(a,\beta,f,\gamma,\theta,\tau_0;t)$ is the function having six parameters of the bandwidth factor $a > 0$, the asymmetry factor $\beta \in (-1,1)$, the central frequency $f \in [0,2\pi)$, the chirp rate $\gamma \in [0,2\pi)$, the phase $\theta \in [0,2\pi)$, and the time shift $\tau_0 > 0$. Here, τ_0 does not affect the morphological feature of AGCT atoms. Thus, we will only adjust the remaining five parameters. In this case, $\boldsymbol{\omega} = (a,\beta,f,\gamma,\theta)$ is the parameter index.

APPENDIX C

1D LOCAL DISCRETE COSINE TRANSFORM

The 1D DCT is derived from the discrete Fourier transform and is orthogonal but not multiscale. The presentation $\mathbf{x}_{\mathrm{DCT}}$ of a discrete waveform $\mathbf{s}$ with the length of N after a 1D DCT is

$$\mathbf{x}_{\mathrm{DCT}}[i] = A_i \sum_{j=0}^{n-1} \mathbf{s}[j] \cos\frac{(0.5+j)i\pi}{N}, \quad i = 0, \cdots, N-1,$$
$$A_i = \begin{cases} \sqrt{\frac{1}{N}} & i = 0 \\ \sqrt{\frac{2}{N}} & i > 0 \end{cases}. \quad \text{(C-1)}$$

The 1D DCT can be used as an over-complete dictionary when applied in blocks (Starck et al., 2004), namely 1D LDCT. The 1D LDCT has two main parameters more than 1D DCT, the block length $l_b \in (1,N)$ and the overlap length $l_o \in (1,N)$ (Starck et al., 2004). For $\mathbf{s}$, the kth block $\mathbf{s}_k[i]$ is

$$\mathbf{s}_k[i] = \mathbf{s}[k(l_b-l_o)+i], \quad i = 0,1,\ldots,l_b-1. \quad \text{(C-2)}$$

Then, DCT is applied to each block. Because the total number of blocks is rounded to $[(N-l_b)/l_o]+1$, the representation of LDCT is a set of representation $\{\mathbf{x}_{k,\mathrm{DCT}}\}_{k=0}^{[(N-l_b)/l_o]}$. Here, $\boldsymbol{\omega} = (l_b, l_o)$ is the parameter index of 1D LDCT.

REFERENCES

Almeida, L., R. Manenti, and M. Porsani, 2015, Coherent noise attenuation using the wavelet transform on radial basis: 85th Annual International Meeting, SEG, Expanded Abstracts, 4720–4724, doi: 10.1190/segam2015-5830728.1.

Auger, A., and N. Hansen, 2016, Linear convergence of comparison-based step-size adaptive randomized search via stability of Markov chains: SIAM Journal on Optimization, **26**, 1589–1624, doi: 10.1137/140984038.

Baraniuk, R. G., and D. L. Jones, 1993, Shear madness: New orthonormal bases and frames using chirp functions: IEEE Transactions on Signal Processing: Special Issue on Wavelets, **41**, 3543–3549, doi: 10.1109/78.258094.

Boßmann, F., and J. Ma, 2015, Asymmetric chirplet transform for sparse representation of seismic data: Geophysics, **80**, no. 6, WD89–WD100, doi: 10.1190/geo2015-0063.1.

Boßmann, F., and J. Ma, 2016, Asymmetric chirplet transform — Part 2: Phase, frequency, and chirp rate: Geophysics, **81**, 6, V425–V439, doi: 10.1190/geo2015-0696.1.

Bruce, A., and X. Li, 2006, Block coordinate relaxation methods for nonparametric signal denoising: IEEE Transactions on Signal Processing, **54**, 2174–2183, doi: 10.1109/TSP.2006.874267.

Butler, K. E., and R. D. Russell, 2003, Cancellation of multiple harmonic noise series in geophysical records: Geophysics, **68**, 1083–1090, doi: 10.1190/1.1581080.

Chen, J., W. Chen, X. Wang, Y. Zhou, Z. Shi, and G. Zhang, 2018, DAS coupling noise suppression using wavelet and DCT dictionary based on sparse optimization: 88th Annual International Meeting, SEG, Expanded Abstracts, 4938–4942, doi: 10.1190/segam2018-2996038.1.

Chen, X., W. Chen, X. Wang, and W. Wang, 2017, Sparsity-optimal separation of body waves and ground-roll by constructing dictionaries using tunable Q-factor wavelet transforms with different Q-factors: Geophysical Journal International, **211**, 621–636, doi: 10.1093/gji/ggx332.

Deighan, A. J., and D. R. Watts, 1997, Ground-roll suppression using the wavelet transform: Geophysics, **62**, 1896–1903, doi: 10.1190/1.1444290.

Denisov, M. S., A. A. Egorov, and M. B. Shneerson, 2021, Optimization-based recursive filtering for separation of signal from harmonics in vibroseis: Geophysical Prospecting, **69**, 779–798, doi: 10.1111/1365-2478.13084.

Eldos, T., 2008, Distributed genetic algorithms: A scheme for genetic drift avoidance: Journal of Electrical Engineering, **59**, 45–48.

Fan, J., Z. Li, X. Song, K. Zhang, and Q. He, 2015, Application of anisotropic high-resolution Radon transform for multiple attenuation: 85th Annual International Meeting, SEG, Expanded Abstracts, 4580–4584, doi: 10.1190/segam2015-5887857.1.

Felix, O., and M. D. Sacchi, 2021, Transform-domain noise synthesis and normal moveout-stack deconvolution approach to ground roll attenuation: Geophysics, **86**, no. 1, V15–V22, doi: 10.1190/geo2019-0638.1.

Gardner, G. H. F., L. W. Gardner, and A. R. Gregory, 1974, Formation velocity and density — The diagnostic basics for stratigraphic traps: Geophysics, **39**, 770–780, doi: 10.1190/1.1440465.

Goldberg, D., 1989, Genetic algorithms in search, optimization, and machine learning: Addison-Wesley Publishing Company.

Gonzalez, R. C., and R. E. Woods, 2018, Digital image processing, 4th ed.: Pearson.

Gregoire, L., P. Cristini, and J. Blanco, 2009, Microseismic wavefield separation using multi-channel chirplet decomposition: 79th Annual International Meeting, SEG, Expanded Abstracts, 1566–1570, doi: 10.1190/1.3255148.

Hu, Y., X. Wang, Q. Hou, D. Liu, X. Shang, M. Zhang, and W. Chen, 2024, Modeling and sparsity promoting separation of wind turbine noise in common-shot gathers: Geophysics, **89**, no. 2, V87–V101, doi: 10.1190/geo2023-0033.1.

Hu, Y., D. Liu, X. Wang, Z. Zhao, and W. Chen, 2022, Attenuation of the multiple reflection-refraction in 2-D common-shot gather via random-derangement-based FX Cadzow filter: IEEE Geoscience and Remote Sensing Letters, **19**, 1–5, doi: 10.1109/LGRS.2021.3051620.

Karslı, H., and D. Dondurur, 2018, A mean-based filter to remove power line harmonic noise from seismic reflection data: Journal of Applied Geophysics, **153**, 90–99, doi: 10.1016/j.jappgeo.2018.04.014.

Laurent, D., S. Fomel, M. Naghizadeh, and M. Sacchi, 2015, Taking signal and noise to new dimensions — Introduction: Geophysics, **80**, no. 6, WDi–WDii, doi: 10.1190/2015-0921-SPSEINTRO.1.

Lebedev, A. V., and I. A. Beresnev, 2004, Nonlinear distortion of signal radiated by vibroseis sources: Geophysics, **69**, 968–977, doi: 10.1190/1.1778240.

Li, X., W. Soellner, and P. Hubral, 1995, Elimination of harmonic distortion in vibroseis data: Geophysics, **60**, 503–516, doi: 10.1190/1.1443787.

Liu, D., L. Gao, X. Wang, and W. Chen, 2021, A dictionary learning method with atom splitting for seismic footprint suppression: Geophysics, **86**, no. 6, V509–V523, doi: 10.1190/geo2020-0681.1.

Liu, D., X. Li, W. Wang, X. Wang, Z. Shi, and W. Chen, 2022a, Eliminating harmonic noise in vibroseis data through sparsity-promoted waveform modeling: Geophysics, **87**, no. 3, V183–V191, doi: 10.1190/geo2021-0448.1.

Liu, D., H. Zhang, X. Wang, W. Chen, Z. Shi, and Z. Zhao, 2022b, Separation of seismic multiple reflection-refraction based on morphological component analysis with high-resolution linear Radon transform: Geophysics, **87**, no. 4, V367–V379, doi: 10.1190/geo2021-0468.1.

Mann, S., and S. Haykin, 1995, The chirplet transform: physical considerations: IEEE Transactions on Signal Processing, **43**, 2745–2761, doi: 10.1109/78.482123.

Meunier, J., and T. Bianchi, 2002, Harmonic noise reduction opens the way for array size reduction in vibroseis operations: 72nd Annual International Meeting, SEG, Expanded Abstracts, 70–73, doi: 10.1190/1.1817354.

Mostafa, N., 2012, Seismic data interpolation and denoising in the frequency-wavenumber domain: Geophysics, **77**, no. 2, V71–V80, doi: 10.1190/geo2011-0172.1.

Nikravesh, M., F. Aminzadeh, and L. A. Zadeh, 2003, Soft computing and intelligent data analysis in oil exploration: Elsevier Science, 1–754.

Nuha, H., B. Liu, M. A. Mohandes, A. Balghonaim, and F. Fekri, 2021, Seismic dataset modeling and compression using particle swarm optimization: Arabian Journal of Geosciences, **14**, 1–11, doi: 10.1007/s12517-021-08675-y.

Pierini, S., M. Aleardi, and A. Mazzotti, 2019, A method to attenuate genetic drift in genetic-algorithm optimizations: Applications to analytic objective functions and two seismic optimization problems: Geophysics, **84**, no. 2, R295–R310, doi: 10.1190/geo2018-0374.1.

Richter, M. M., S. Paul, V. Këpuska, and M. Silaghi, 2022, Fundamentals of signal transformations, *in* J. Doe and J. Smith, eds., Signal processing and machine learning with applications: Springer.

Selesnick, I. W., 2011, Wavelet transform with tunable Q-factor: IEEE Transactions on Signal Processing, **59**, 3560–3575, doi: 10.1109/TSP.2011.2143711.

Sicking, C., T. Fleure, S. Nelan, and B. McLain, 2009, Slip sweep harmonic noise rejection on correlated shot data: 79th Annual International Meeting, SEG, Expanded Abstracts, 36–39, doi: 10.1190/1.3255636.

Starck, J., M. Elad, and D. L. Donoho, 2004, Redundant multiscale transforms and their application for morphological component separation: Advances in Imaging and Electron Physics, **132**, 287–348, doi: 10.1016/S1076-5670(04)32006-9.

Sundhararajan, S., A. Pahwa, and P. Krishnaswami, 1998, A comparative analysis of genetic algorithms and directed grid search for parametric optimization: Engineering with Computers, **14**, 197–205, doi: 10.1007/BF01215973.

Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399, doi: 10.1190/1.1543224.

Wang, B., H. Li, B. Zhao, Z. Huang, M. Zhang, and L. Mei, 2012, Cross-harmonic noise removal on slip-sweep vibroseis data: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0059.1.

Wang, W., W. Chen, J. Lei, and J. Gao, 2010, Ground roll separation by sparsity and morphological diversity promotion: 90th Annual International Meeting, SEG, Expanded Abstracts, 3705–3710, doi: 10.1190/1.3513621.

Xu, J., W. Wang, J. Gao, and W. Chen, 2013, Monochromatic noise removal via sparsity-enabled signal decomposition method: IEEE Geoscience and Remote Sensing Letters, **10**, 533–537, doi: 10.1109/LGRS.2012.2212271.

Yilmaz, O., 1987, Seismic data processing: SEG, Volume 2 of Investigations in Geophysics.

Yu, Z., and P. G. Garossino, 2005, High-energy noise attenuation of seismic data in the wavelet-transform domain: Integrated Computer-Aided Engineering, **12**, 57–67, doi: 10.3233/ICA-2005-12105.

Biographies and photographs of the authors are not available.