GEOPHYSICS, VOL. 90, NO. 6 (NOVEMBER-DECEMBER 2025); P. V615–V632, 27 FIGS.
10.1190/GEO2024-0620.1

# Implicit neural representations for self-supervised seismic deblending

Weiwei Xu[1], Dawei Liu[2], Xiaokai Wang[3], Mauricio D. Sacchi[4], and Wenchao Chen[3]

## ABSTRACT

Blended acquisition enables near-simultaneous firing of multiple shots, facilitating faster and denser seismic data collection, which reduces acquisition time, lowers costs, and potentially enhances subsurface illumination. However, these benefits require an additional deblending process to recover the seismic data one would have acquired via a conventional survey. Numerous data-driven deblending algorithms have emerged, predominantly relying on supervised deep learning. Nevertheless, constructing training pairs that are representative and transferable to complex field data remains challenging. Some unsupervised methods have been developed, but they often focus on local similarities within individual seismic data, neglecting the continuity of seismic wavefields and relationships between different data across the entire data set. Implicit neural representation (INR) provides a versatile framework for modeling continuous seismic signals using neural network parameterized functions, capturing the intrinsic relationships throughout the entire data set. Building on this advantage, this study introduces INR for self-supervised seismic deblending. Specifically, to capitalize on the prior knowledge that desired unblended signals in nearby unblended common-receiver gathers (CRGs) vary smoothly due to the continuity of seismic waveforms, a deep neural network is used as an implicit function to parameterize desired unblended CRGs with receiver indices as inputs. By leveraging the blending operator, a physics-informed training strategy is implemented to enforce measurement consistency, enabling the trained network to recover corresponding deblended CRGs accurately. Numerical experiments conducted on synthetic and field data sets demonstrate the efficacy of our approach. Compared with sparsity-promoting inversion with a patched 2D Fourier transform, a widely adopted method in the industry, and the plug-and-play method with blind-spot networks, one of the most widely accepted self-supervised approaches, our method shows superior performance.

## INTRODUCTION

In traditional seismic data acquisition, maintaining receivers to register only the wavefield triggered by a single shot requires setting relatively large intervals between adjacent shots. Such an approach incurs substantial time costs, particularly when striving for high-density, wide-azimuth seismic data surveys. A notable enhancement in acquisition efficiency comes in the form of blended acquisition, which enables multiple seismic shots to approximate synchronous excitation within a brief timeframe, thereby reducing the acquisition cycle time and significantly cutting down costs (Berkhout, 2008; Berkhout et al., 2008). This advancement has made remarkable strides in the realm of seismic data acquisition. However, it comes at the expense of blending together wavefields originating from different shots.

Presently, two approaches are used to manage blended wavefields. One approach involves adjusting conventional imaging and inversion techniques to accommodate blended data (Dai et al., 2011; Verschuur and Berkhout, 2011; Chen et al., 2017; Luo et al., 2022). However, since most mature processing workflows are tailored for unblended data, a more prevalent strategy is to treat interference caused by other shots as blending noise and devise deblending methods to mitigate its effects. From a mathematical standpoint, the blending process can be represented as a linear mapping that transforms conventional unblended seismic data into

Manuscript received by the Editor 27 August 2024; revised manuscript received 17 July 2025; published ahead of production 26 July 2025; published online 15 September 2025.
[1]Chang'an University, School of Geological Engineering and Geomatics, Xi'an, China and Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China. E-mail: 544274644@qq.com.
[2]Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China and University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: 409791715@qq.com (corresponding author).
[3]Xi'an Jiaotong University, School of Information and Communications Engineering, Xi'an, China. E-mail: xkwang@xjtu.edu.cn; wencchen@xjtu.edu.cn.
[4]University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: msacchi@ualberta.ca.

observed blended seismic data using a blending operator. This operator encompasses the firing time of shots at various positions (Berkhout et al., 2008; Berkhout, 2013). Consequently, deblending methods can be broadly classified into two categories based on their utilization of the blending operator: filtering and inversion techniques.

Filtering-based deblending methods rely on the premise that with the use of an appropriate random dithering code in the blended acquisition, the blending noise exhibits coherence in the common-shot domain (CSD) and random distribution in other domains after applying the adjoint of the blending operator to the observed blended data (referred to as pseudo-deblending). In contrast, the desired seismic data maintains coherence across all domains. Consequently, effective filtering techniques can be used for blending noise suppression by leveraging the distinct characteristics between desired seismic data and blending noise. In these domains, widely used methods include the median filter and its various improved versions (Huo et al., 2012; Chen, 2015; Huang et al., 2018; Chen et al., 2020) and prediction-error filters (Spitz et al., 2008; Mahdad et al., 2012). In addition, rank-reduction methods (Cheng et al., 2015; Wu and Bai, 2019) are used as filtering techniques. Notably, filtering the sparse transform domains can be applied not only to blended common-receiver gathers (CRGs) (Ibrahim and Sacchi, 2014) but also to blended common shot gathers (CSGs) (Yu et al., 2017).

Although these traditional methods have proven effective, they often rely on manually designed filters and may struggle with complex noise patterns. Deep-learning approaches have emerged as powerful alternatives for blending noise filtering, allowing the ability to learn intricate data representations automatically. A straightforward approach involves training a deblending network using pseudo-deblending data as inputs and corresponding unblended data as labels through end-to-end training. Advanced network architectures such as convolutional neural networks (CNNs) (Sun et al., 2020), U-nets (Wang et al., 2021), and transformers (Zu et al., 2022) have been explored in this context. Regarding training data sets, a common practice involves collecting suitable field unblended seismic data from other surveys or generating synthetic unblended seismic data for simulating blended acquisition scenarios (Sun et al., 2020; Wang et al., 2021; Zu et al., 2022). However, models trained on such data sets may exhibit limited performance when applied to field blended data with distinct or more complex characteristics. In addition, using conventional processing workflows to deblend blended data and then using the results as training targets introduces significant computational overhead and may compromise training quality due to residual blending artifacts (Sun et al., 2022). To mitigate these issues, several training data construction strategies have been explored. One approach involves using previously acquired unblended data from the same survey area to construct training pairs, ensuring consistency in acquisition conditions (Baardman et al., 2020). Another effective strategy is to incorporate unblended shot gathers acquired at the end of each sail line during the blended acquisition process, which has been demonstrated to be successful in blended-by-acquisition scenarios (Sun et al., 2022). Given the similarity in features between blended CSGs and unblended CRGs, a strategy is proposed that entails using observed blended CSGs or pseudo-deblending data in the CSD for training, followed by filtering in the common-receiver domain (CRD) (Xu et al., 2022b). Furthermore, self-supervised training strategies, which solely rely on observed blended seismic data, have been proposed for filtering-based deblending methodologies (Wang et al., 2023b; Cheng et al., 2024).

The literature demonstrates that inversion-based deblending methods consistently achieve better overall deblending quality than filtering-based approaches (Ibrahim and Sacchi, 2014). The core principle of inversion-based methods is to frame deblending as a linear inverse problem, where the objective function is designed to minimize the residuals between observed blended data and the estimated blended data. However, this inverse problem is inherently underdetermined, which necessitates carefully incorporating priors or regularization techniques to effectively constrain the solution space and enhance the recovery of the desired seismic data. In line with this approach, regularization enforcing sparsity has been extensively used (Mahdad et al., 2011; Mansour et al., 2012; Zhou et al., 2014, 2016; Li et al., 2019; Zu et al., 2019; Kontakis and Verschuur, 2024), leveraging the notion that useful seismic data can be represented in an auxiliary domain with few nonzero coefficients. This auxiliary domain can be obtained through fixed-basis transforms, such as the Fourier transform (Mahdad et al., 2011), generalized windowed transforms (Li et al., 2019), curvelet transform (Mansour et al., 2012), and focal transform (Kontakis and Verschuur, 2024). Alternatively, it can be derived from a dictionary learned directly from the observed blended data (Zhou et al., 2014, 2016; Zu et al., 2019). In addition, the low-rank property of seismic data, after being rearranged into structured matrix forms (e.g., Hankel matrices), has been harnessed as a regularization mechanism (Cheng and Sacchi, 2015, 2016; Zu et al., 2017). Low-rank regularization can be interpreted as imposing sparse constraints on the singular values.

With the capacity of deep learning to extract meaningful features from seismic data, pretrained deep neural networks are used as regularization tools. For instance, a pretrained deblending network can be incorporated into the iterative process to enhance the deblending outcome (Zu et al., 2020). Integrating a pretrained deep Gaussian denoiser into the plug-and-play (PnP) framework has demonstrated very good deblending performance (Xu et al., 2022b). Using the decoder component of a pretrained convolutional autoencoder as a nonlinear preconditioner for the deblending problem has proven effective (Xu et al., 2022a). Unsupervised or self-supervised deep-learning techniques have also been applied to inversion-based deblending to circumvent the need for constructing training pairs. For instance, the concept of deep image prior (DIP) has been leveraged for deblending purposes (Xu et al., 2021). Unsupervised double-deep neural networks have been seamlessly integrated into the iterative process (Wang et al., 2023a). Furthermore, it has been illustrated that the pretrained Gaussian denoiser can be substituted with a novel self-supervised blind-spot network (Luiken et al., 2024). However, though these techniques can capture detailed patterns within individual seismic data, they often overlook the broader relationships among different data across the entire data set, which is crucial prior information for accurate deblending.

The effectiveness of leveraging correlations between adjacent gathers has been demonstrated in several supervised learning-based deblending methods (Richardson and Feller, 2019; Sun, 2022; Sun et al., 2022), motivating the development of a self-supervised strategy that explicitly exploits such spatial dependencies. Recently, implicit neural representations (INRs) (Sitzmann et al., 2020; Chen et al., 2021; Mildenhall et al., 2021; Sun et al., 2021; Brandolin et al., 2024; Li et al., 2024; Liu et al., 2024; Shen et al., 2024) have showcased robust capabilities in representing diverse signals. The core concept involves encoding temporal or spatial information of signals through neural networks, thereby yielding corresponding signal values based on this temporal or spatial information. Inspired by the recent success of INRs for video (Chen et al., 2021), we further incorporate the INR between adjacent unblended CRGs and develop a self-supervised seismic deblending method. In essence, similar to the strong correlation between adjacent frames in a video, seismic data acquisition reveals a certain degree of similarity in features among seismic records collected by adjacent receivers due to the subsurface's gradual variation or overall consistency. Consequently, unblended CRGs can be represented as a function of receiver indices. During training, a deep neural network is used as an implicit function to model desired unblended CRGs, using receiver indices as inputs. Training on individual CRGs allows the network to capture local features within each CRG effectively. By leveraging receiver indices, the network also learns the relationships among different CRGs, encoding data set-wide similarities and acquiring the capability to represent all CRGs. A blending operator is incorporated into the training strategy to enforce measurement consistency. The trained network is a surrogate for all deblended CRGs in the inference phase. It can output the corresponding deblended CRG given the receiver index. Synthetic and field data sets are used to evaluate the performance of the proposed method in comparison to sparsity-promoting inversion with patched 2D Fourier transform, a widely used industry technique, and PnP with blind-spot networks, one of the most accepted self-supervised approaches.

The paper is organized as follows: it begins with a detailed description of the proposed method, followed by experimental results based on synthetic and field seismic data sets. A subsequent section discusses the limitations of the proposed method, its connection to the DIP approach, and the challenges associated with its application in practical blended acquisition scenarios. This is followed by an analysis of the computational efficiency of the proposed method. Finally, the paper concludes the paper and outlines potential directions for future research.

# METHOD

## Deblending problem statement

To retrieve deblended seismic data from blended records, it is necessary to establish a forward model that maps unblended data to blended data. Given the seismic geometry comprising various shots $\{s_j\}_{j=1}^{n_s}$ and receivers $\{r_i\}_{i=1}^{n_r}$, in conventional seismic data acquisition, the initiation of the next shot occurs only after all seismic waves produced by the current shot have been completely received. Hence, a vectorized CRG recorded by receiver $r_i$ can be represented as

$$\mathbf{d}_{r_i} = [\mathbf{d}_{r_i}^1; \mathbf{d}_{r_i}^2; \ldots; \mathbf{d}_{r_i}^{n_s}], \tag{1}$$

where $\{\mathbf{d}_{r_i}^j\}_{j=1}^{n_s}$ represents the time series recorded by receiver $r_i$, resulting from the interaction between seismic waves emitted from shot $s_j$ and the subsurface, comprising $n_t$ time samples. Although maintaining the same seismic geometry as conventional acquisition, blended acquisition triggers multiple shots rapidly based on a predesigned random dithering code. This process results in the incoherent superposition of wavefields from different shots, which are subsequently recorded by the receivers. This process leads to the incoherent blending together of wavefields from different shots, subsequently recorded by receivers. Thus, the blended seismic data $\mathbf{b}_{r_i}$ registered by receiver $r_i$ can be represented as the superposition of all unblended data $\{\mathbf{d}_{r_i}^j\}_{j=1}^{n_s}$

$$\mathbf{b}_{r_i} = \mathbf{\Gamma}\mathbf{d}_{r_i} = \sum_{j=1}^{n_s} \mathbf{\Gamma}_j \mathbf{d}_{r_i}^j, \tag{2}$$

where $\{\mathbf{\Gamma}_j\}_{j=1}^{n_s} \in \mathbb{R}^{f \times n_t}$ represents the time-shift operator, which shifts and pads each $\{\mathbf{d}_{r_i}^j\}_{j=1}^{n_s}$ to match the length of the blended seismic data $\mathbf{b}_{r_i}$. Figure 1 shows the forward model of blended acquisition.

For the blending operator $\mathbf{\Gamma}$, each $\{\mathbf{\Gamma}_j\}_{j=1}^{n_s}$ can be explicitly represented as a block matrix (Xu et al., 2022b)

$$\mathbf{\Gamma}_j = \begin{bmatrix} \mathbf{0}_{j1} \\ \mathbf{I}_{n_s} \\ \mathbf{0}_{j2} \end{bmatrix}, \tag{3}$$

where $\mathbf{I}_{n_s} \in \mathbb{R}^{n_s \times n_s}$ denotes an identity matrix. In addition, $\mathbf{0}_{j_1}$ and $\mathbf{0}_{j_2}$ represent zero matrices of dimensions $c_j \times n_s$ and $(f - n_s - c_{j_1}) \times n_s$, respectively. The integer $c_j$, where $0 \le c_j \le (f - n_s)$, varies with the firing time of shot $s_j$. It is worth noting that each $\{\mathbf{\Gamma}_j\}_{j=1}^{n_s}$ possesses the following properties:

First, for any $j$, the self-product of $\mathbf{\Gamma}_j$ satisfies the relation

$$\mathbf{\Gamma}_j^T \mathbf{\Gamma}_j = \mathbf{0}_{j_1}^T \mathbf{0}_{j_1} + \mathbf{I}_{n_s}^2 + \mathbf{0}_{j_2}^T \mathbf{0}_{j_2} = \mathbf{I}_{n_s}. \tag{4}$$

Second, for any pair $j \neq k$, the cross-product of $\mathbf{\Gamma}_k$ and $\mathbf{\Gamma}_j$ is given by

$$\mathbf{\Gamma}_k^T \mathbf{\Gamma}_j = \begin{cases} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I}_{n_t + c_k - c_j} & \mathbf{0} \end{bmatrix} & \text{if } c_k + n_t \ge c_j + 1 \\ \begin{bmatrix} \mathbf{0} & \mathbf{I}_{n_t + c_j - c_k} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \text{if } c_k + 1 \le c_j + n_t \\ \mathbf{0}_{n_t} & \text{otherwise} \end{cases}. \tag{5}$$

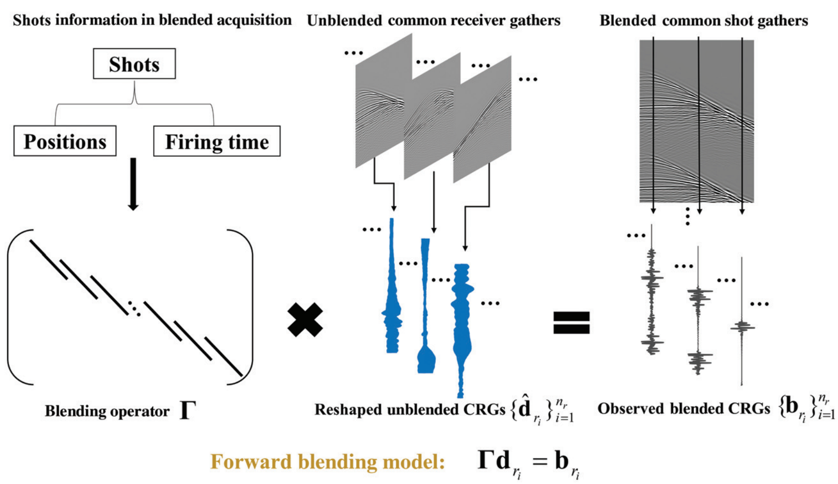



Figure 1. Schematic of the forward model for blended acquisition.

Correspondingly, $\boldsymbol{\Gamma} \in \mathbb{R}^{f\times(n_t\times n_s)}$ signifies the blending operator, constituting a horizontal stack of $\{\boldsymbol{\Gamma}_j\}_{j=1}^{n_s}$

$$\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_{n_s}]. \tag{6}$$

As a consequence of the blending effect, the blended seismic data $\mathbf{b}_{r_i}$ constitutes a singular long vector of length $f$, where $f < (n_t \times n_s)$.

The application of the adjoint of the blending operator $\boldsymbol{\Gamma}$ to the blended data $\{\mathbf{b}_{r_i}\}_{i=1}^{n_r}$ is referred to as pseudo-deblending, which generates the pseudo-deblended data as follows:

$$\boldsymbol{\Gamma}^T\mathbf{b}_{r_i} = \boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{d}_{r_i}$$

$$= \begin{bmatrix} \sum_{j=1}^{n_s}\boldsymbol{\Gamma}_1^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \\ \sum_{j=1}^{n_s}\boldsymbol{\Gamma}_2^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \\ \vdots \\ \sum_{j=1}^{n_s}\boldsymbol{\Gamma}_{n_s}^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \end{bmatrix} = \begin{bmatrix} \mathbf{d}_{r_i}^1 + \sum_{j=2}^{n_s}\boldsymbol{\Gamma}_1^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \\ \mathbf{d}_{r_i}^2 + \sum_{j=1,j\neq 2}^{n_s}\boldsymbol{\Gamma}_2^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \\ \vdots \\ \mathbf{d}_{r_i}^{n_s} + \sum_{j=1}^{n_s-1}\boldsymbol{\Gamma}_{n_s}^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j \end{bmatrix}. \tag{7}$$

It can be observed that the result of pseudo-deblending is a combination of the original CRG and the so-called blending noise. Specifically, by combining equations 5 and 7, it becomes evident that the blending noise on each trace, denoted as $\sum_{j=1,j\neq k}^{n_s}\boldsymbol{\Gamma}_k^T\boldsymbol{\Gamma}_j\mathbf{d}_{r_i}^j$, is produced by extracting and rearranging specific parts of the original CRG using $\boldsymbol{\Gamma}_k^T\boldsymbol{\Gamma}_j$. Studies have shown that, with a suitable random dithering code, the blending noise appears laterally coherent in the CSD but laterally incoherent in nonshot domains, such as the CRD. Consequently, filtering-based methods can be effectively applied.

## Deblending by inversion

In addition to obtaining deblended data $\hat{\mathbf{d}}_{r_i}$ from blended data $\mathbf{b}_{r_i}$ using filtering methods, the deblending process can also be cast as an inverse problem according to equation 2, facilitating the recovery of deblended data $\hat{\mathbf{d}}_{r_i}$ from blended data $\mathbf{b}_{r_i}$. From a Bayesian perspective, this problem can be formulated as the following maximum a posteriori (MAP) estimation:

$$\hat{\mathbf{d}}_{r_i} = \arg\max_{\mathbf{d}_{r_i}} \log\, p(\mathbf{b}_{r_i}|\mathbf{d}_{r_i}) + \log\, p(\mathbf{d}_{r_i}), \tag{8}$$

where $p(\mathbf{b}_{r_i}|\mathbf{d}_{r_i})$ represents the likelihood term and $p(\mathbf{d}_{r_i})$ denotes the prior distribution of unblended data $\mathbf{d}_{r_i}$. Typically, solving the MAP problem in equation 8 corresponds to addressing the following optimization problem:

$$\hat{\mathbf{d}}_{r_i} = \arg\min_{\mathbf{d}_{r_i}} E(\mathbf{b}_{r_i}, \boldsymbol{\Gamma}\mathbf{d}_{r_i}) + \lambda\mathcal{R}(\mathbf{d}_{r_i}), \tag{9}$$

where $E(\mathbf{b}_{r_i}, \boldsymbol{\Gamma}\mathbf{d}_{r_i})$ denotes the data fidelity term, which quantifies the difference between observed blended data $\mathbf{b}_{r_i}$ and the result of the blending operator $\boldsymbol{\Gamma}$ applied to the desired $\mathbf{d}_{r_i}$. The $l_2$ norm is typically associated with Gaussian distribution, while the $l_1$ norm corresponds to Laplace distribution. Here, $\mathcal{R}(\mathbf{d}_{r_i})$ signifies the regularization term, which encapsulates the prior structures of unblended data $\mathbf{d}_{r_i}$. Given the underdetermined nature of the blending model described in equation 2, selecting an appropriate regularization term to confine the solution space is paramount. The term $\lambda$ serves as a trade-off parameter controlling the weight of the regularization term.

## Implicit neural representation for CRGs

Over the years, INR (Sitzmann et al., 2020; Chen et al., 2021; Mildenhall et al., 2021; Sun et al., 2021; Shen et al., 2024; Li et al., 2024; Liu et al., 2024) has garnered significant attention for its notable performance in parameterizing various complex signals in a self-supervised manner. The fundamental concept involves continuously representing the signal of interest as an implicit function, which is approximated using a learnable neural network $f_\theta$. By using a point-wise training scheme, the network effectively maps the signal coordinates $\mathbf{x} \in \mathbb{R}^m$ to their corresponding values $\mathbf{y} \in \mathbb{R}^n$. Once adequately trained, the network can be used as a proxy for the signal, providing flexible representation capabilities in a self-supervised manner.

Due to their resolution-agnostic nature, efficient memory usage, ability to mitigate locality biases, and broad adaptability, multilayer perceptrons (MLPs) are predominantly used in implementing INRs (Molaei et al., 2023). However, adopting conventional rectified linear unit (ReLU)-based MLPs in INRs often makes them learn low-frequency details, manifesting a phenomenon termed spectral bias (Rahaman et al., 2019; Tancik et al., 2020). To address this challenge, Fourier feature mapping (FFM) (Tancik et al., 2020), which expands signal coordinates as a composite of distinct frequency components, can be integrated before MLPs to enhance their capability in capturing high-frequency signal components. Furthermore, the integration of periodic activation functions has been shown to effectively facilitate the learning of high-frequency signal components, exemplified by the utilization of sine functions in sinusoidal representation networks (Sitzmann et al., 2020). In addition, augmenting the output dimensionality of MLPs proves beneficial in mitigating the difficulties encountered by INRs in learning high-frequency components (Aftab et al., 2022). Notably, unlike point-wise INRs, in which output values correspond to individual coordinate points, Chen et al. (2021) propose an image-wise INR for video, termed NeRV. This approach maps frame indices to their corresponding video frames, demonstrating significant advantages in encoding efficiency, decoding speed, and representation quality.

Most existing supervised deep-learning-based deblending methods operating on CRGs need to learn the structural features of all CRGs from the training data and then use this knowledge to process all CRGs. Although not explicitly stated, these methods implicitly assume a certain degree of similarity among all CRGs. This assumption is reasonable due to the gradual variation and relative homogeneity of subsurface geologic structures typically observed in seismic exploration. Usually, the close physical proximity of adjacent receivers leads them to capture reflected signals from comparable subsurface structures and layers. As a result, neighboring CRGs often exhibit more similar morphological characteristics. However, most existing unsupervised or self-supervised deep-learning methods focus solely on local similarities within individual CRGs, failing to exploit this prior information fully. Leveraging this prior could potentially further enhance deblending performance.

Drawing upon the strengths of INR, we introduce a concise and efficient method to represent CRGs, incorporating the prior knowledge that adjacent CRGs share more similar features into the network parameters in a self-supervised manner. Specifically, similar to how a video is represented as a function of frame index $t$ in NeRV (Chen et al., 2021), CRGs can be represented by a neural network as a function of the receiver index. The network $f_\theta$ with parameters $\theta$ can be defined as follows:

$$f_\theta : r_i \in \mathbb{R} \rightarrow \mathbf{d}_{r_i} \in \mathbb{R}^{n_t \times n_s}. \quad (10)$$

The receiver index $\{r_i\}_{i=1}^{n_r}$ is derived from the actual absolute coordinates of the receivers. The output $\{\mathbf{d}_{r_i}\}_{i=1}^{n_r}$ represents the corresponding CRG. Once properly trained, the CRGs can be encoded as follows:

$$\mathbf{d}_{r_i} = f_\theta(r_i), \quad r_i \in \{1, 2, \ldots, n_r\}. \quad (11)$$

It can be seen that the network function $f_\theta$ maps receiver indices to the CRGs, encoding the relevant information of all the CRGs into the network parameters. In addition, since receivers with similar absolute coordinates have close index values, the prior knowledge that adjacent CRGs exhibit more similar morphological characteristics is inherently integrated into the network parameters based on the receiver indices.

## Inversion-based seismic deblending via INR

Building on the preceding analysis, to ensure that the prior knowledge of adjacent CRGs having more similar morphological characteristics informs the optimization process, we propose introducing INR for CRGs to tackle the inversion-based seismic deblending problem. In this approach, deblended CRGs are reformulated as the optimization of the neural network parameters. Figure 2 provides an overview of the proposed framework.

Specifically, the desired unblended CRGs $\{\mathbf{d}_{r_i}\}_{i=1}^{n_r}$ are parameterized as a function of the receiver indices by a neural network $f_\theta$ with parameters $\theta$. Furthermore, to encode the prior knowledge of adjacent CRGs having more similar morphological characteristics, all observed blended CRGs and their corresponding receiver indices are paired as training samples $\{r_i, \mathbf{b}_{r_i}\}_{i=1}^{n_r}$ and included in the training process. Therefore, based on the formulation in equation 9 and adopting the $l_1$ norm for data fidelity term, the inversion-based seismic deblending optimization problem with INR for CRGs can be expressed as

$$\hat{\theta} = \arg\min_\theta \frac{1}{n_r} \sum_{i=1}^{n_r} \| \mathbf{b}_{r_i} - \mathbf{\Gamma} f_\theta(r_i) \|_1. \quad (12)$$

It should be noted that, just as in DIP (Ulyanov et al., 2018), the explicit regularization term $\mathcal{R}(\mathbf{d}_{r_i})$ in equation 9 is replaced by the implicit regularization from neural network parameters $\theta$. After obtaining the optimal neural network parameters $\hat{\theta}$ through end-to-end training, the neural network $f_{\hat{\theta}}$ serves as a proxy for all the deblended CRGs. For a given receiver index $\{r_i\}_{i=1}^{n_r}$, the corresponding deblended CRG can be generated by inferring the trained neural network: $\{\hat{\mathbf{d}}_{r_i} : f_{\hat{\theta}}(r_i)\}_{i=1}^{n_r}$.

The neural network $f_\theta$, shown in Figure 3, comprises an FFM module $\gamma(r_i)$, an MLP module $M_{\theta_M}$ parameterized by $\theta_M$, and a convolutional decoder module $D_{\theta_D}$ parameterized by $\theta_D$. As a result, the network can be represented as $f_\theta(r_i) = D_{\theta_D}(M_{\theta_M}(\gamma(r_i)))$, where $\theta = \{\theta_M, \theta_D\}$.

We now describe the different modules of $f_\theta$. The FFM module, as previously mentioned, enhances the capability of MLPs in capturing high-frequency signal components (Tancik et al., 2020). More importantly, in the context of seismic deblending, FFM flexibly constrains the correlations between CRGs across different data sets. Therefore, we use an FFM to encode the receiver indices as follows:

$$\gamma(r_i) = [\cos(\omega_1 r_i), \sin(\omega_1 r_i), \ldots, \cos(\omega_K r_i), \sin(\omega_K r_i)]^T, \quad (13)$$

where $K$ denotes the number of frequency components, and $\{\omega_i : \pi\beta^i\}_{i=1}^{K}$ represents the exponential-based frequency mappings. Given a receiver index $r_i$, normalized between $(0, 1]$, the output of FFM $\gamma(r_i)$ is fed into the subsequent neural network.

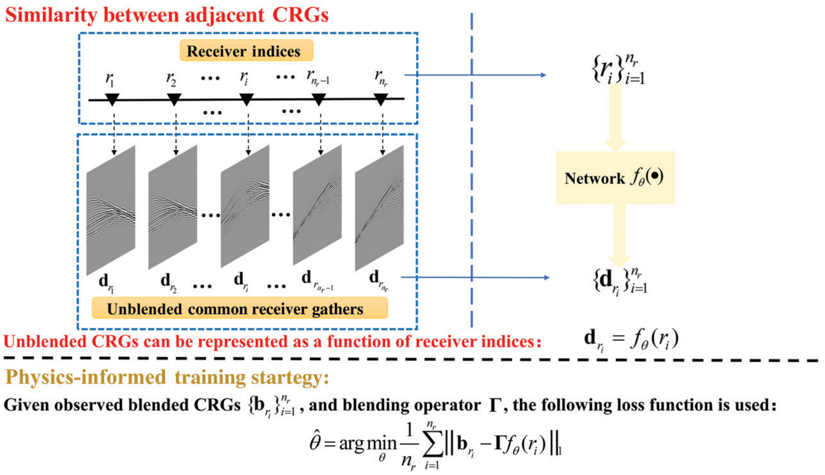



Figure 2. Overview of the proposed framework.

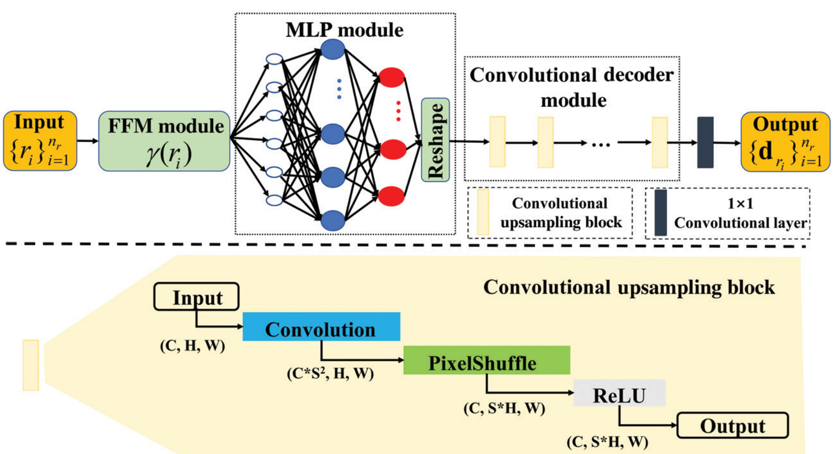



Figure 3. Overview of the proposed network architecture.

The MLP module $M_{\theta_M}$ comprises two fully connected (FC) layers. The number of neurons in each layer are hyperparameters that define the model's representative capacity. Each layer is followed by a ReLU activation function to introduce nonlinearity and mitigate the vanishing gradient problem. Consequently, the mathematical representation of the MLP $M_{\theta_M}$ is as follows:

$$M_{\theta_M}(\gamma(r_i))=\phi(\mathbf{W}_2\cdot\phi(\mathbf{W}_1\gamma(r_i)+\mathbf{b}_1)+\mathbf{b}_2), \tag{14}$$

where $\mathbf{W}_i$ and $\mathbf{b}_i$ represent the weights and biases for the $i$th layer, respectively, with $i=1,2$. And $\phi(\cdot)$ denotes the element-wise ReLU activation function. The MLP module $M_{\theta_M}$ transforms the high-dimensional encoding generated by the FFM into a compact latent representation that captures essential global features. This representation is subsequently reshaped into a 3D feature tensor, which serves as a coarse-resolution approximation of the CRG and acts as the initial input to the convolutional decoder module for progressive refinement.

The convolutional decoder module $D_{\theta_D}$ is applied following the MLP module to enhance feature map resolution and representation. As shown in Figure 3, this module comprises a certain number of convolutional upsampling blocks, each consisting of a convolutional layer, a PixelShuffle (Shi et al., 2016) layer that effectively rearranges the feature map's spatial dimensions to increase resolution, and a ReLU activation function. This hierarchical design facilitates progressive refinement of spatial features while preserving local details. The number of channels in the decoder is adjusted in accordance with a fixed strategy: the first upsampling block retains the same number of channels as the MLP output, while each subsequent block attempts to halve the number of channels. To avoid excessively reducing the feature dimension, a minimum channel threshold is enforced, and once this threshold is reached, all remaining blocks maintain this fixed channel count. Finally, a $1\times1$ convolutional layer reduces the number of channels to one, generating the single-channel CRG output.

The overall network architecture balances the expressive capacity of global implicit representations, enabled by the FFM and MLP modules, with the spatial modeling strengths of decoder module. This architecture is consistently adopted across different data sets.

Once the network architecture is established, the parameters are optimized using the Adam optimizer (Kinga and Adam, 2015), with an initial learning rate of $5\times10^{-4}$ and a cosine annealing learning rate schedule (Loshchilov and Hutter, 2016). Since the proposed method is self-supervised and does not require labeled data, no validation or evaluation metrics are used. Instead, an early stopping strategy is adopted, where the network is trained for a predefined number of epochs without intermediate validation. Upon completion of training, the deblended CRG corresponding to the specified index is directly generated.

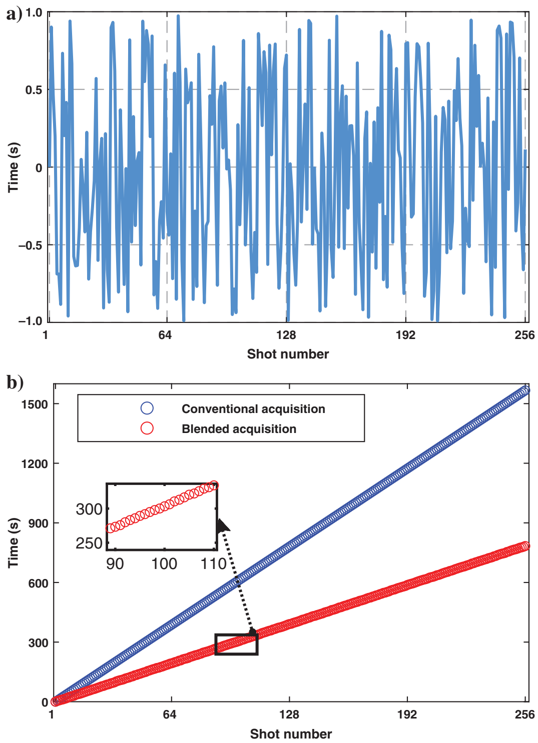



Figure 4. Illustration of shot timing patterns for the Marmousi data example. (a) Random dithering code applied to each shot in the blending simulation and (b) conventional and blended shot firing times.

## EXPERIMENTS

This section presents the proposed method's deblending performance using synthetic and field data sets featuring complex subsurface structures. For these data sets, we simulate continuous blending with random time dithering. In this approach, airguns towed by the acquisition vessel are fired sequentially at regular time intervals, with each shot delayed by a predefined random time. The receivers record the seismic waves continuously, resulting in blended data $\{\mathbf{b}_{r_i}\}_{i=1}^{n_r}$, which is the superposition of unblended data, each shifted in time according to the applied delay. To evaluate the performance, we compare our method with two established approaches, including the sparsity-promoting inversion with patched 2D Fourier transform (Abma et al., 2015), a widely adopted technique in industry. It is important to note that, for this method, we follow the procedure outlined in Luiken et al. (2024). In addition, we compare our approach with the PnP method using blind-spot networks, referred to as SSDeblend (Luiken et al., 2024), one of the most widely recognized self-supervised methods.

The recovery performance is quantitatively assessed using the signal-to-noise ratio (S/N) metric, defined as

$$\mathrm{S/N}=10\log_{10}\frac{\|\mathbf{d}_{r_i}\|_2^2}{\|\hat{\mathbf{d}}_{r_i}-\mathbf{d}_{r_i}\|_2^2}, \tag{15}$$

where $\hat{\mathbf{d}}_{r_i}$ is the deblended CRG, and $\mathbf{d}_{r_i}$ is the original unblended CRG.

In addition to S/N, *f-k* spectrum analysis of the deblending results is performed to validate the proposed approach's effectiveness further.

### The Marmousi data example

The synthetic data set is derived from the central portion of the 2D Marmousi model, comprising 256 shots and 256 receivers. Each shot record consists of 768-time samples with a sampling interval of 8 ms, resulting in a total recording duration of 6.136 s per shot. The regular shot time interval is set to half of the recording duration to simulate continuous blending acquisition, i.e., 3.068 s. A predefined random dithering code, ranging from –1 to 1 s and shown in Figure 4a, is applied to introduce timing variations into the blended acquisition.

Figure 4b shows the shot firing times in both conventional and blended acquisition. Unlike conventional acquisition, blended acquisition reduces the total acquisition time by approximately half.

Figure 5a and 5b shows the 96th unblended CRG and its corresponding pseudo-deblended CRG, respectively. Figure 5c and 5d shows the *f-k* spectra of the data shown in Figure 5a and 5b, respectively. It is evident that in the CRG domain, blended acquisition introduces incoherent blending noise, obscuring the useful signal's structural features. Similarly, Figure 6a and 6b shows the 96th unblended CSG and its corresponding pseudo-deblended CSG, whereas Figure 6c and 6d shows their *f-k* spectra. In the CSG domain, blended acquisition causes overlapping signal energy, resulting in coherent artifacts in the *f-k* domain.

The proposed method and two comparative approaches are applied to all blended data. For the patched 2D Fourier sparse inversion method, patches of size (64, 64) with an overlap of (32, 32) samples are used, and the initial regularization parameter is set to $\lambda_0 = 5.0$ for deblending. For the SSDeblend method, the outer and inner iterations are set to 30 and 10, respectively, with $\rho = 1$, denoiser epochs of 20, and a batch size of eight. For the proposed method, 256 receivers are indexed from one to 256 and normalized to the range (0, 1]. These indices are encoded using an FFM module with hyperparameters $K = 90$ and $\beta = 1.25$, resulting in a 180-dimensional input vector per receiver. The MLP module comprises two FC layers: the first projects the input to a 32-dimensional hidden representation using ReLU activation, and the second generates a 10752-dimensional output. This output is reshaped into a tensor of size $224 \times 12 \times 4$, where 224 denotes the number of feature channels and $12 \times 4$ represents the initial spatial resolution. The resulting tensor is passed to a convolutional decoder with three upsampling stages, using stride settings of eight, four, and two, respectively. These stages progressively increase the spatial resolution from $12 \times 4$ to $96 \times 32$, then to $384 \times 128$, and finally to $768 \times 256$. The number of feature channels is adjusted as follows: the first upsampling block retains the input channel of 224, while each subsequent block halves the number of channels until the minimum of 64 is reached, which is then maintained in the remaining blocks. A final $1 \times 1$ convolutional layer reduces the channel count from 64 to one, producing the single-channel CRG output. The network is trained for 1000 epochs with a batch size of eight. As shown in Figure 7, the training loss indicates successful convergence.

The final S/N results are shown in Figure 8. Although all methods achieve effective deblending, the proposed method significantly outperforms the patched 2D Fourier sparse inversion method in terms of S/N. In addition, it achieves comparable performance to the SSDeblend method, demonstrating its superior deblending capability while preserving more useful signal information. Figure 9 shows the deblending results of different methods for the 96th CRG. From Figure 9a–9c, it can be observed that all three methods successfully recover the useful signal, with blending noise almost entirely suppressed. Figure 9d–9f shows that the proposed method yields the smallest residual error, highlighting its superiority. The *f-k* spectra corresponding to Figure 9 are shown in Figure 10, where the proposed method's deblended result exhibits the closest match to the target data. Similarly, Figure 11 shows the deblending results of various methods for the 96th CSG, and Figure 12 shows the corresponding *f-k* spectra. It is observed that all three methods produce acceptable deblending results, with the proposed method yielding the smallest error, further demonstrating its effectiveness.

## The Gulf of Suez data example

The Gulf of Suez field data set is used to simulate blended data. This data set comprises 128 shot locations and 128 receivers, with a uniform spatial interval of 12.5 m for both shots and receivers. Each

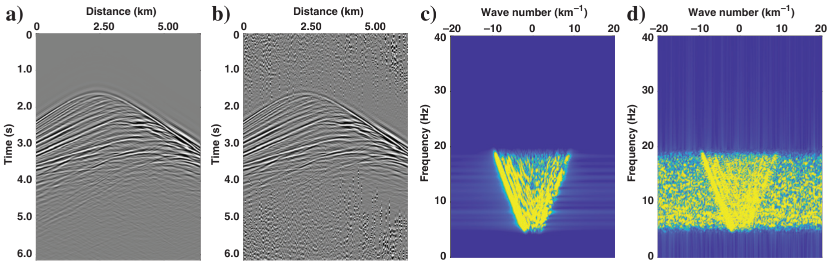



Figure 5. CRGs from the Marmousi data example. (a) The 96th unblended CRG, (b) the corresponding pseudo-deblended CRG, (c) *f-k* spectrum of the unblended CRG, and (d) *f-k* spectrum of the pseudo-deblended CRG.

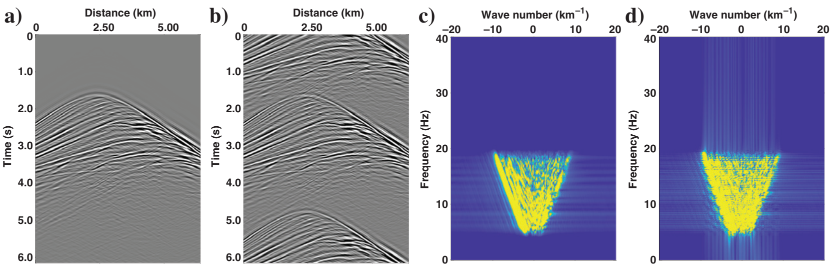



Figure 6. CSGs from the Marmousi data example. (a) The 96th unblended CSG, (b) the corresponding pseudo-deblended CSG, (c) *f-k* spectrum of the unblended CSG, and (d) *f-k* spectrum of the pseudo-deblended CSG.

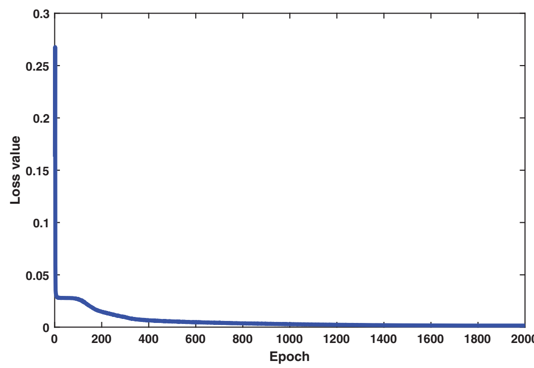



Figure 7. Training loss curve of the proposed method for the Marmousi data example.

shot record consists of 512 time samples at a temporal interval of 4 ms, resulting in a total recording duration of 2.044 s per shot. The regular shot time interval is set to half of the recording duration to simulate continuous blending acquisition, i.e., 1.022 s. A predefined random dithering code, ranging from −1 to 1 s and shown in Figure 13a, is applied to the shots. Shot firing times for conventional and blended acquisition are shown in Figure 13b.

Figure 14a and 14b shows the 96th unblended CRG and its corresponding pseudo-deblended CRG, respectively. Figure 14c and 14d shows the *f-k* spectra of the data shown in Figure 14a and 14b, respectively. Similar to the observations in the Marmousi data set, blending noise in the CRD is incoherent and obscures the structural characteristics of the useful signal. Likewise, Figure 15a and 15b shows the 96th CSG and its corresponding pseudo-deblended CSG, whereas Figure 15c and 15d shows their *f-k* spectra. It is evident that in the CSD, the overlapping shot records introduce mutual interference.

The proposed method and two comparative approaches are applied to all blended data. For the patched 2D Fourier sparse inversion method, patches of size (64, 64) with an overlap of (32, 32) samples are used, and the initial regularization parameter is set to $\lambda_0 = 2.50$ for deblending. For the SSDeblend method, the outer and inner iterations are set to 30 and 10, respectively, with $\rho = 1$, denoiser epochs of 20, and a batch size of eight. For the proposed method, the 128 deployed receivers are indexed from one to 128 and normalized to the range (0, 1). These normalized indices are encoded using the FFM module with hyperparameters $K = 90$ and $\beta = 1.25$, resulting in a 180-dimensional input vector for each receiver. The MLP module comprises two FC layers: the first projects the input to a 32-dimensional hidden representation using ReLU activation, and the second outputs a 3584-dimensional vector. This output is reshaped into a tensor of size $224 \times 8 \times 2$, which is then passed to a convolutional decoder consisting of six upsampling stages, each with stride two. These stages progressively increase the spatial resolution from $8 \times 2$ to $512 \times 128$. The number of feature channels is adjusted during upsampling as follows: the first stage retains the input channel of 224, and each subsequent stage halves the number of channels until the minimum of 64 is reached, which is then preserved in the remaining stages. A final $1 \times 1$ convolutional layer reduces the number of channels from 64 to 1, producing the single-channel CRG output. The network is trained for 2000 epochs with a batch size of 32. As shown in Figure 16, the training loss curve confirms that the model has converged stably.

Various methods are applied to all blended CRGs, and the corresponding S/N results are shown in Figure 17. It can be observed

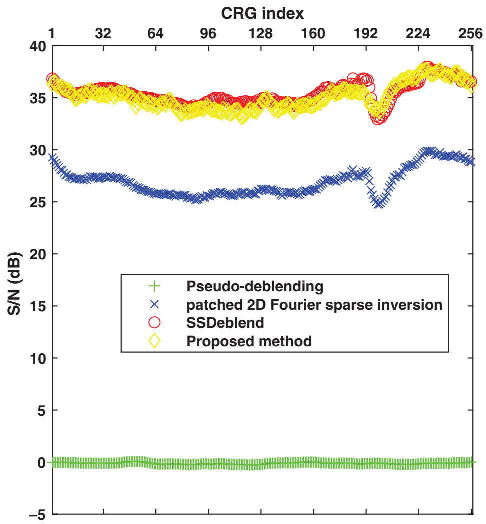



Figure 8. S/N results of different methods on all the blended CRGs for the Marmousi data example.

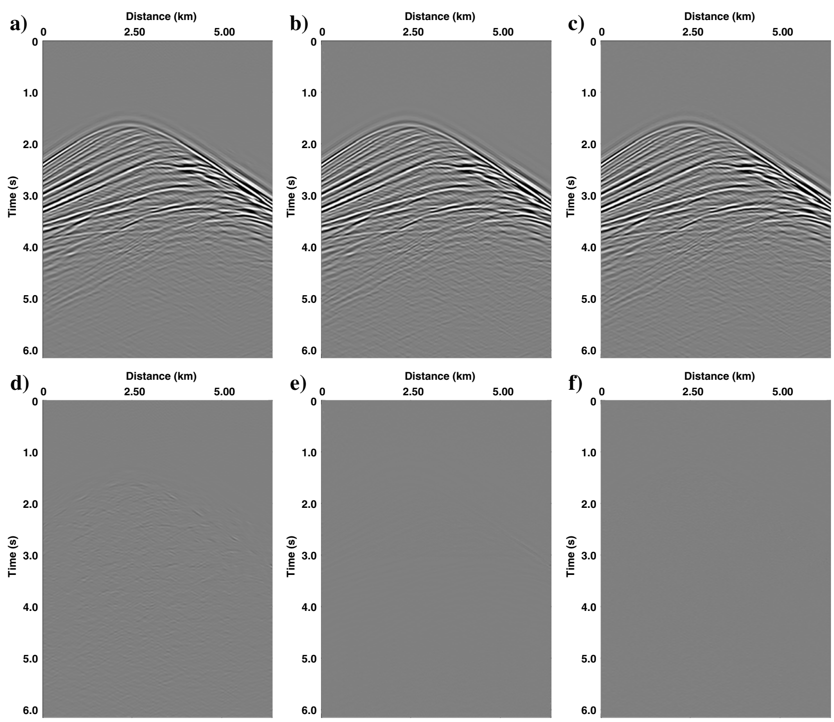



Figure 9. Deblending results of various methods applied to the 96th CRG from the Marmousi data example. (a) Deblended CRG obtained by patched 2D Fourier sparse inversion method, (b) deblended CRG obtained by SSDeblend method, (c) deblended CRG obtained by the proposed method, (d) difference obtained by patched 2D Fourier sparse inversion method, (e) difference obtained by SSDeblend method, and (f) difference obtained by the proposed method.

that while the proposed method does not achieve a higher S/N than the SSDeblend method, it outperforms the patched 2D Fourier sparse inversion method. Figure 18 shows the 96th deblended CRGs obtained using different methods. Figure 18a–18c shows that although all three methods effectively suppress most of the blending noise, the patched 2D Fourier sparse inversion method exhibits the most residual blending noise and introduces discontinuities in the recovered signal. Figure 18d–18f shows that the proposed method preserves more signal information than the patched 2D Fourier sparse inversion method. Although it introduces more signal distortion compared with the SSDeblend method, the level of distortion remains acceptable, demonstrating the superior deblending capability of the proposed approach. The *f-k* spectra corresponding to Figure 18 are shown in Figure 19, further supporting this observation. Figure 20 shows the 96th deblended CSGs obtained using the three methods, whereas Figure 21 shows their corresponding *f-k* spectra. The patched 2D Fourier sparse inversion method causes the most significant signal distortion. Although the proposed method introduces slightly more signal damage than the SSDeblend method, the results remain comparable, validating the effectiveness of the proposed approach in seismic deblending.

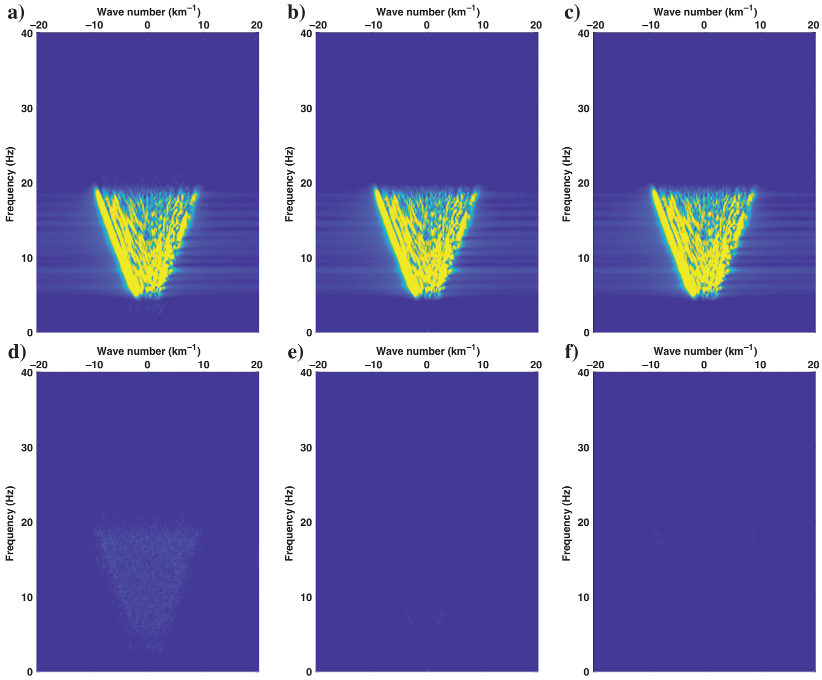



Figure 10. The *f-k* spectra of the deblending results for the 96th CRG from the Marmousi data example. (a) *f-k* spectrum of the deblended CRG obtained by the patched 2D Fourier sparse inversion method, (b) *f-k* spectrum of the deblended CRG obtained by the SSDeblend method, (c) *f-k* spectrum of the deblended CRG obtained by the proposed method, (d) *f-k* spectrum of the difference obtained by the patched 2D Fourier sparse inversion method, (e) *f-k* spectrum of the difference obtained by the SSDeblend method, and (f) *f-k* spectrum of the difference obtained by the proposed method.

## DISCUSSION

This section discusses the parameter selection for the FFM module in the proposed method, its connection to the DIP approach, its performance in irregular shot-distributed blended acquisition, and an analysis of its computational cost.

### Parameter selection for FFM module

As previously discussed, we aim to use the FFM module to encode the correlation between unblended CRGs recorded by adjacent receivers. To validate the feasibility of this approach, Figure 22 shows the FFM-encoded results and their correlation patterns for 200 CRG indices under different parameter settings. It can be observed that varying the number of frequency components $K$ and the frequency scaling factor $\beta$ leads to changes in the correlation patterns of the FFM-encoded results. This demonstrates that the FFM can adapt to different data sets by adjusting its parameters, effectively capturing data set-specific relationships. Furthermore, the ability to reveal diverse correlation patterns highlights the FFM's capability to represent all CRGs comprehensively.

However, the correlation between adjacent CRGs varies across different data sets, making the optimal selection of $K$ and $\beta$ nontrivial. Taking

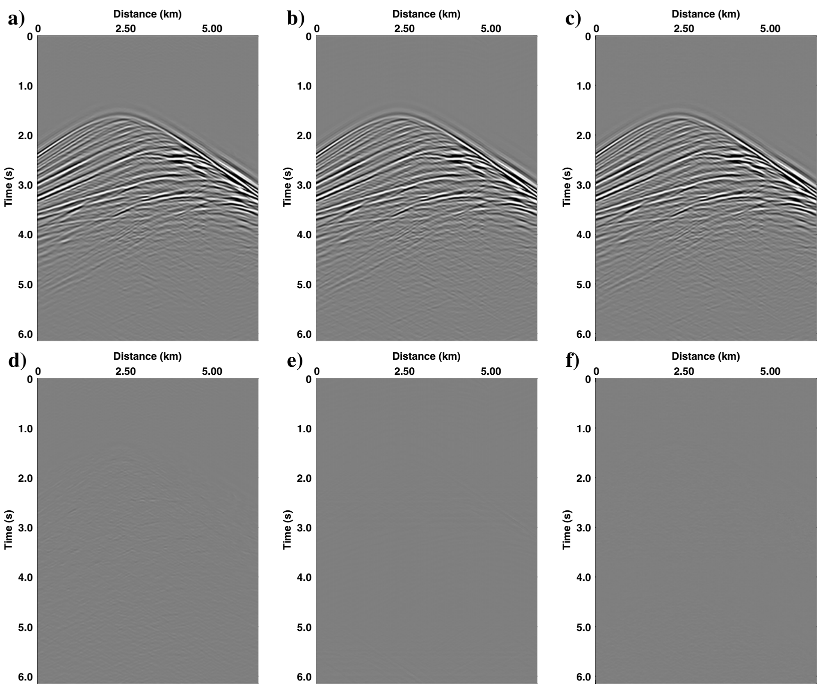
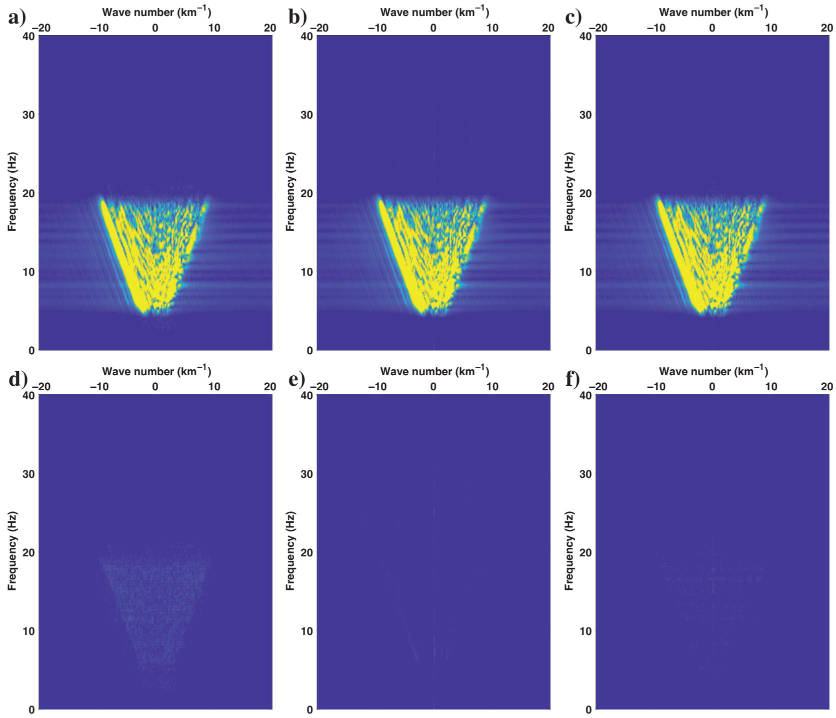



Figure 11. Deblending results of various methods applied to the 96th CSG from Marmousi data example. (a) Deblended CSG obtained by patched 2D Fourier sparse inversion method, (b) deblended CSG obtained by SSDeblend method, (c) deblended CSG obtained by the proposed method, (d) difference obtained by patched 2D Fourier sparse inversion method, (e) difference obtained by SSDeblend method, and (f) difference obtained by the proposed method.

Figure 12. The *f-k* spectra of the deblending results for the 96th CSG from the Marmousi data example. (a) *f-k* spectrum of the deblended CSG obtained by the patched 2D Fourier sparse inversion method, (b) *f-k* spectrum of the deblended CSG obtained by the SSDeblend method, (c) *f-k* spectrum of the deblended CSG obtained by the proposed method, (d) *f-k* spectrum of the difference obtained by the patched 2D Fourier sparse inversion method, (e) *f-k* spectrum of the difference obtained by the SSDeblend method, and (f) *f-k* spectrum of the difference obtained by the proposed method.



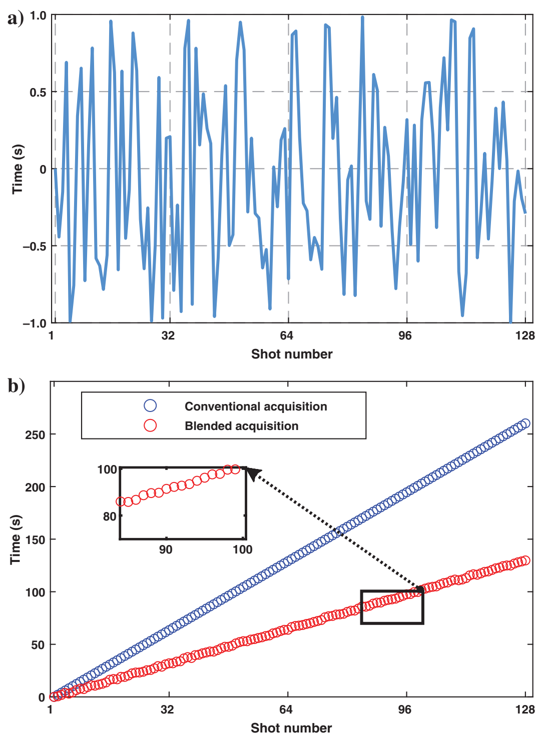




Figure 13. Illustration of shot timing patterns for the Gulf of Suez data example. (a) Random dithering code applied to each shot in the blending simulation and (b) conventional and blended shot firing times.

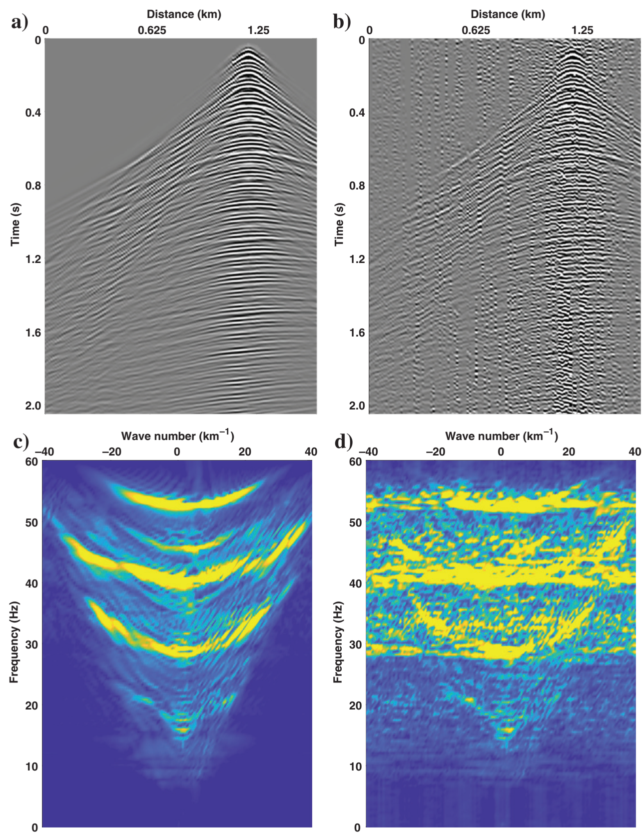



Figure 14. CRGs from the Gulf of Suez data example. (a) The 96th unblended CRG, (b) the corresponding pseudo-deblended CRG, (c) *f-k* spectrum of the unblended CRG, and (d) *f-k* spectrum of the pseudo-deblended CRG.

the Gulf of Suez data set as an example, Figure 23a and 23b shows the loss curves and final S/N values obtained under different $K$ and $\beta$ settings. Although all configurations eventually converge (as indicated by the negligible changes in the loss curves after 1750 epochs), the convergence rates and final loss values exhibit noticeable differences. Furthermore, the resulting S/N varies significantly across different parameter choices, necessitating empirical tuning to achieve optimal performance. This remains a practical limitation of the proposed method.

Despite this limitation, we provide empirical guidelines for selecting $K$ and $\beta$. For data sets with more complex structures, a larger $K$ is generally required, but our experiments suggest that $K$ does not need to exceed 90. In addition, we observe that when $\beta$ exceeds 2.5, the network often struggles to converge. Based on these findings, we recommend a practical range of $\beta$ between 1.0 and 2.5.

## Relation to deep image prior

DIP (Ulyanov et al., 2018) is another widely used unsupervised method incorporating physics-informed strategies. It has been demonstrated that a CNN-based network can recover a clean signal from a degraded one during the optimization process of mapping random noise to the degraded signal, and it has been successfully applied to seismic deblending (Xu et al., 2021). From the perspective of INR, DIP can be regarded as an implicit function that maps random noise to the target signal. This raises a fundamental question: Given the difficulty of selecting optimal FFM parameters and the demonstrated success of DIP in encoding random noise, is it still necessary to use an FFM module based on receiver indices in the proposed method? To investigate this, we retain all other components of the proposed method but replace the encoding results of the FFM module with randomly sampled Gaussian vectors. This results in a DIP-based variant of our approach, termed variant-DIP (V-DIP).

It is crucial to highlight the similarities and differences between V-DIP and the original DIP (O-DIP). The O-DIP method typically adopts an encoder-decoder hourglass architecture, often a U-Net-like structure. In contrast, V-DIP uses an MLP + decoder architecture. In essence, the encoder in O-DIP and the MLP in V-DIP serve as feature encoding modules, processing the input similarly to extract meaningful representations. This structural similarity indicates that V-DIP can still be regarded as part of the DIP framework. Another key distinction is the training

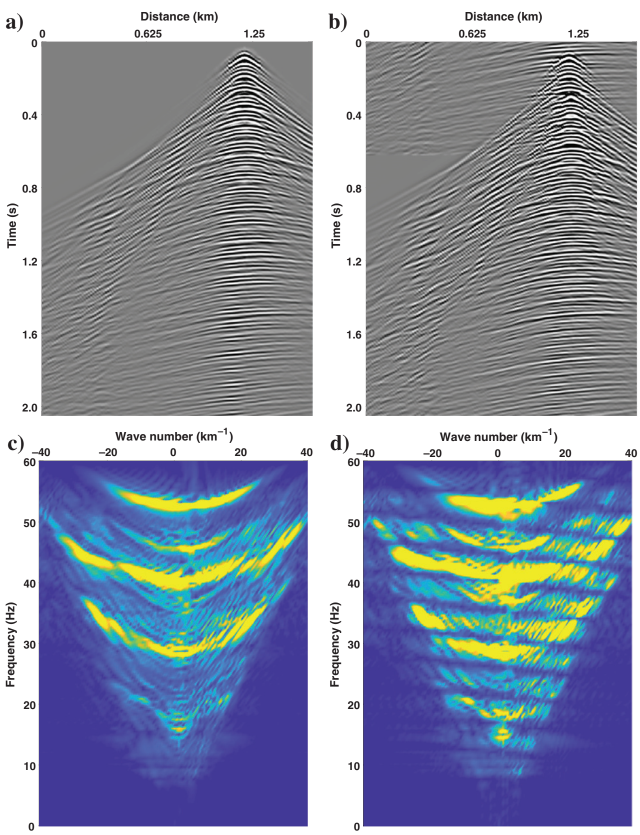



Figure 15. CSGs from the Gulf of Suez data example. (a) The 96th unblended CSG, (b) the corresponding pseudo-deblended CSG, (c) $f$-$k$ spectrum of the unblended CSG, and (d) $f$-$k$ spectrum of the pseudo-deblended CSG.

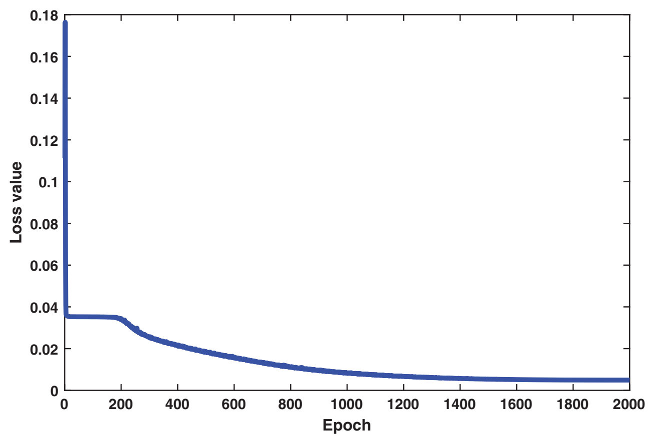



Figure 16. Training loss curve of the proposed method for the Gulf of Suez data example.

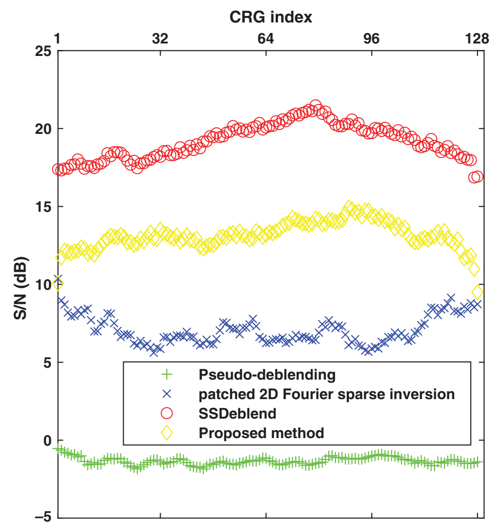



Figure 17. S/N results of different methods on all the blended CRGs for the Gulf of Suez data example.

strategy. O-DIP uses a zero-shot training approach, processing only a single instance from the data set at each training session. In the case of seismic deblending, this corresponds to a single CRG. In contrast, V-DIP preassigns a randomly sampled Gaussian vector to each CRG in the data set (with the same dimensionality as the FFM encoding output to ensure a fair comparison with the proposed method), and all CRGs participate in training simultaneously. This constitutes the primary difference between V-DIP and O-DIP. However, O-DIP could also be modified to train on all CRGs simultaneously by increasing the number of input channels. From this perspective, V-DIP can still be considered a form of DIP.

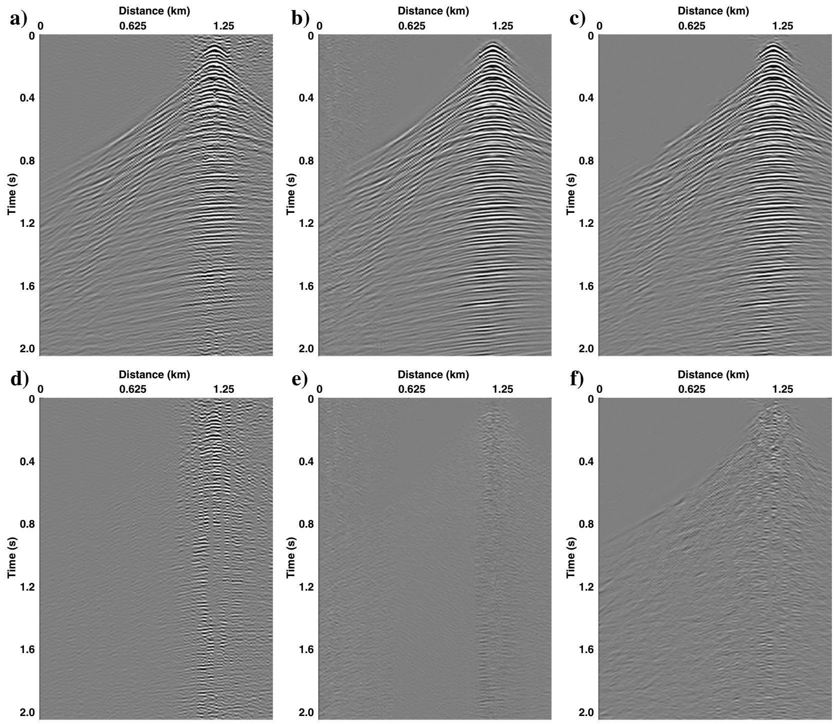



Figure 18. Deblending results of various methods applied to the 96th CRG from Gulf of Suez data example. (a) Deblended CRG obtained by patched 2D Fourier sparse inversion method, (b) deblended CRG obtained by SSDeblend method, (c) deblended CRG obtained by the proposed method, (d) difference obtained by patched 2D Fourier sparse inversion method, (e) difference obtained by SSDeblend method, and (f) difference obtained by the proposed method.

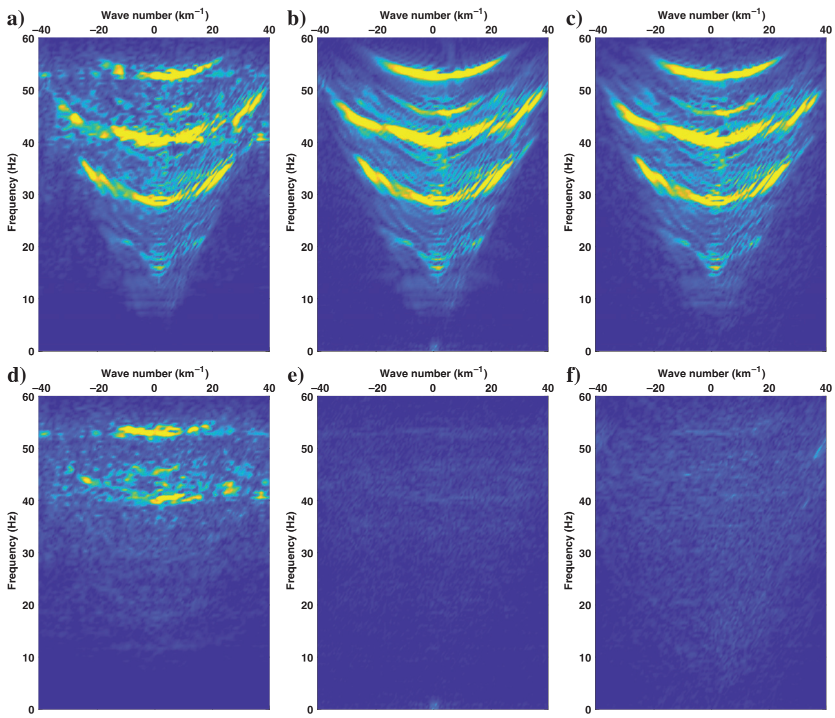



Figure 19. The *f-k* spectra of the deblending results for the 96th CRG from the Gulf of Suez data example. (a) *f-k* spectrum of the deblended CRG obtained by the patched 2D Fourier sparse inversion method, (b) *f-k* spectrum of the deblended CRG obtained by the SSDeblend method, (c) *f-k* spectrum of the deblended CRG obtained by the proposed method, (d) *f-k* spectrum of the difference obtained by the patched 2D Fourier sparse inversion method, (e) *f-k* spectrum of the difference obtained by the SSDeblend method, and (f) *f-k* spectrum of the difference obtained by the proposed method.

Figure 24a and 24b shows the loss convergence curves and final S/N obtained using V-DIP and the proposed method on the Marmousi data set. Similarly, Figure 25a and 25b shows the results for the Gulf of Suez data set. It can be observed that, compared with V-DIP, the proposed method achieves superior deblending performance, demonstrating the necessity and effectiveness of the FFM module.

## Performance of the proposed method in irregular shot-distributed blended acquisition

In practical blended acquisition, the shot positions are often irregularly distributed due to terrain constraints and economic considerations. In other words, they do not always align with a regular grid, which results in insufficient lateral continuity of CRGs and

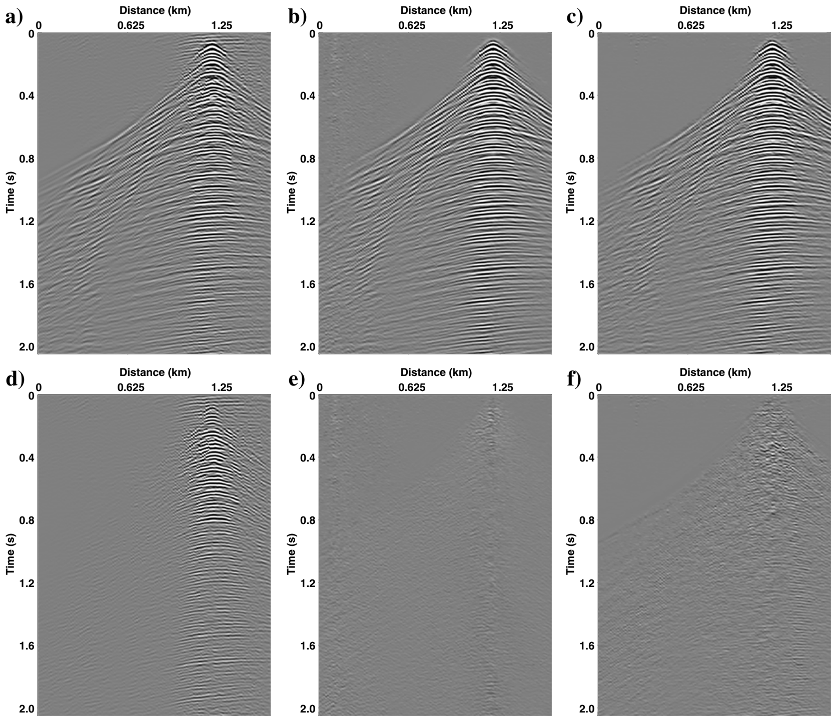



Figure 20. Deblending results of various methods applied to the 96th CSG from Gulf of Suez data example. (a) Deblended CSG obtained by patched 2D Fourier sparse inversion method, (b) deblended CSG obtained by SSDeblend method, (c) deblended CSG obtained by the proposed method, (d) difference obtained by patched 2D Fourier sparse inversion method, (e) difference obtained by SSDeblend method, and (f) difference obtained by the proposed method.

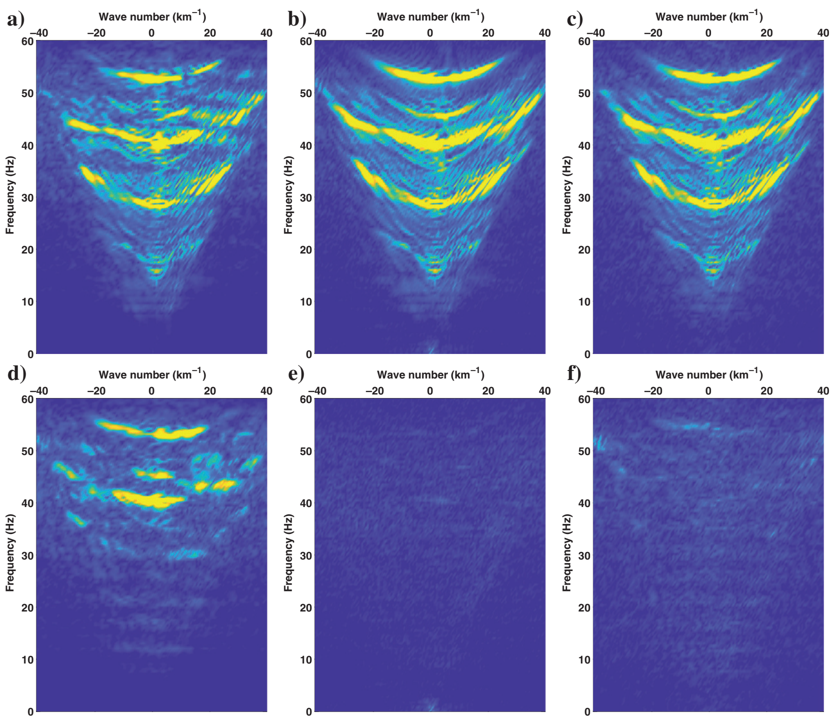



Figure 21. The *f-k* spectra of the deblending results for the 96th CSG from the Gulf of Suez data example. (a) *f-k* spectrum of the deblended CSG obtained by the patched 2D Fourier sparse inversion method, (b) *f-k* spectrum of the deblended CSG obtained by the SSDeblend method, (c) *f-k* spectrum of the deblended CSG obtained by the proposed method, (d) *f-k* spectrum of the difference obtained by the patched 2D Fourier sparse inversion method, (e) *f-k* spectrum of the difference obtained by the SSDeblend method, and (f) *f-k* spectrum of the difference obtained by the proposed method.

poses significant challenges for seismic deblending. To evaluate the capability of the proposed method in handling irregular shot-distributed blended CRGs, we conduct experiments using the Marmousi and Gulf of Suez data sets.

First, to simulate deviations in source vessel activation positions from a regular grid, we randomly remove 50% of the traces to generate irregular-grid CRGs. It is important to note that such deviations are typically less pronounced in real-world scenarios. We then apply continuous blending simulation to construct the blended data, which is subsequently processed using the proposed method.

For the Marmousi data set example, Figure 26a shows an irregular-grid CRG, which serves as the target for deblending. Figure 26b shows the corresponding pseudo-deblended CRG, whereas Figure 26c and 26d shows the deblended CRG obtained using the proposed method and the corresponding difference, respectively. Figure 26e and 26f shows a CSG and its corresponding pseudo-deblended CSG, respectively. Figure 26g and 26h shows the deblended CSG obtained using the proposed method and the corresponding difference. The results indicate that the difference remains within an acceptable range, demonstrating the effectiveness of the proposed method in handling irregular shot-distributed blended acquisition. Similarly, Figure 27 shows an example from the Gulf of Suez data set, further validating the robustness of our approach in irregular acquisition scenarios.

## Computational cost analysis

We conduct all experiments on an Intel(R) Xeon(R) Silver 4110 central processing unit (CPU) @ 2.10 GHz equipped with a single NVIDIA GeForce RTX 3090 graphics processing unit (GPU). The computational cost of seismic deblending algorithms can be categorized into two main components: the cost of forward and adjoint operations using the blending operator in iterative optimization, and the computational overhead associated with the applied priors. Among the three methods evaluated, the SSDeblend method and the proposed approach require network training, which incurs higher computational costs than the patched 2D Fourier sparse inversion method, which relies solely on patched Fourier operators.

The SSDeblend method is essentially based on the alternating direction method of multipliers (ADMM) algorithm, which requires solving a quadratic optimization problem and updating network parameters iteratively. In the provided implementation (Luiken et al., 2024), network optimization is performed on the GPU, while all non-network computations run on the CPU. This leads to slow computations on the CPU and additional overhead from CPU-GPU data transfer. Empirically, we found that the computational time of the SSDeblend method is significantly affected by the CRG size. For the Marmousi and Gulf of Suez data sets, the processing times are approximately 2007 and 142 min, respectively. Notably, while the data volume approximately doubles, the computational time increases by a factor of 15, which suggests that the method exhibits a nonlinear scalability issue, likely due to memory access inefficiencies and increased iteration complexity.

In contrast, the proposed method leverages a blending-operator-informed training strategy and fully uses GPU computations, resulting in significantly lower computational costs than the SSDeblend method. The processing times for the Marmousi and Gulf of Suez data sets are approximately 306 and 104 min, respectively.

Our proposed method requires significantly less GPU memory than SSDeblend. Taking the Marmousi data set as an example, the SSDeblend method consumes 23,785 MiB, whereas our method only requires 2992 MiB, demonstrating a significant reduction in memory footprint.

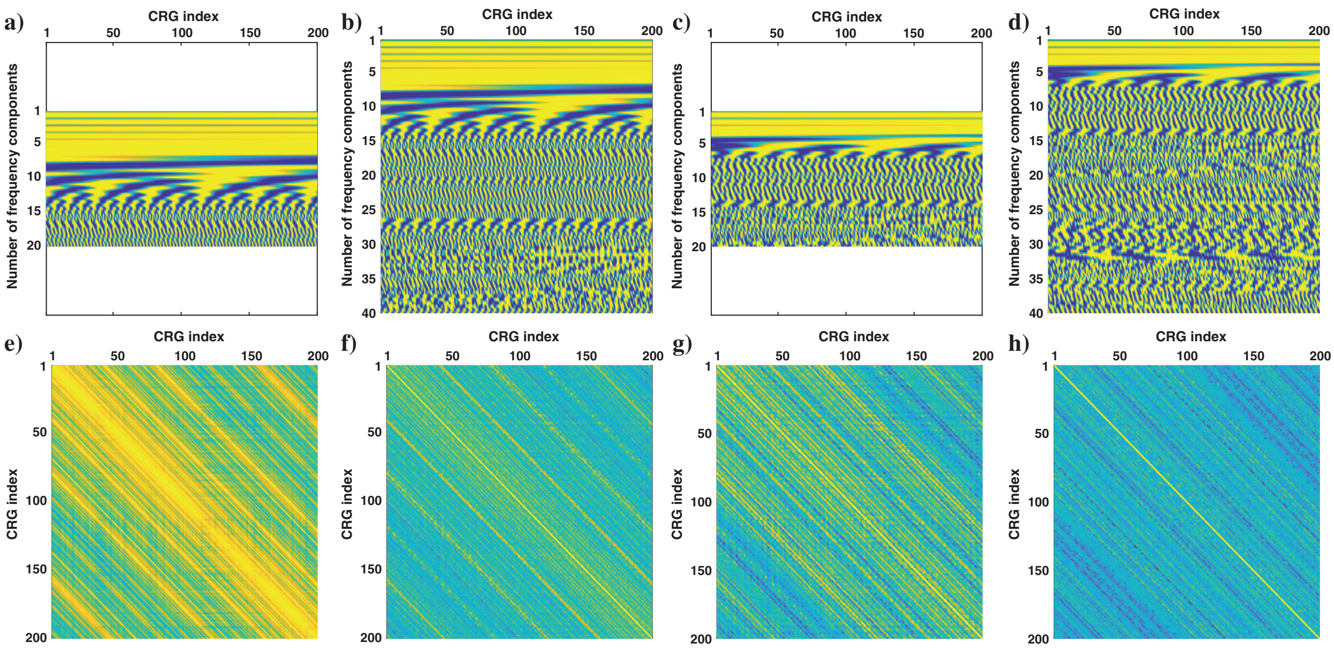



Figure 22. Results and correlations of the FFM for 200 CRG indices with various parameter selections. $\{\gamma(r_i)\}_{i=1}^{200}$ with (a) $\beta=2, K=20$, (b) $\beta=2, K=40$, (c) $\beta=4, K=20$, and (d) $\beta=4, K=40$. Correlations of $\{\gamma(r_i)\}_{i=1}^{200}$ with (e) $\beta=2, K=20$, (f) $\beta=2, K=40$, (g) $\beta=4, K=20$, and (h) $\beta=4, K=40$.

## Limitation of shot firing time accuracy

In the proposed self-supervised framework, the blending operator is constructed based on the assumed accurate shot firing time, which plays a critical role in enforcing measurement consistency during network training. Although shot timing information is generally well recorded, it may still be partially missing or inaccurate in certain challenging acquisition scenarios due to factors such as clocking errors or field operational delays (Chen et al., 2023). If the provided timing information deviates from the actual shot schedule, the blending operator used in training may not faithfully reflect the true blending process, which can lead to degraded deblending performance. Therefore, while the method does not require unblended data for training, it relies on the availability and accuracy of shot time metadata. Future work may explore strategies to improve robustness against shot timing uncertainty, such as incorporating timing estimation or adopting data-driven approaches to infer blending patterns.

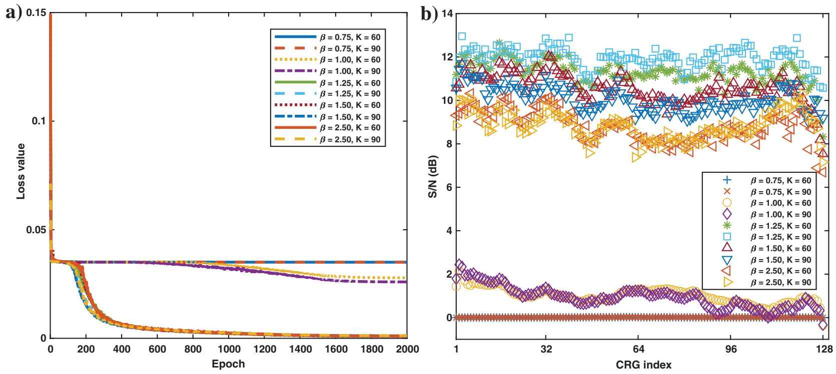



Figure 23. Loss curves and final S/N values for the Gulf of Suez data example under different $\beta$ and $K$ settings in FFM module. (a) Loss curves under different $\beta$ and $K$ settings and (b) final S/N values under different $\beta$ and $K$ settings.

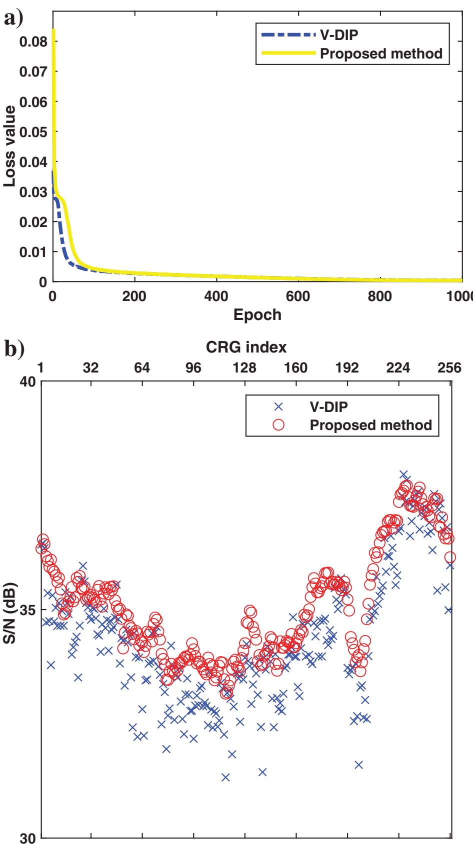



Figure 24. Loss curves and final S/N obtained using V-DIP and the proposed method on the Marmousi data example. (a) Loss curves and (b) final S/N values.

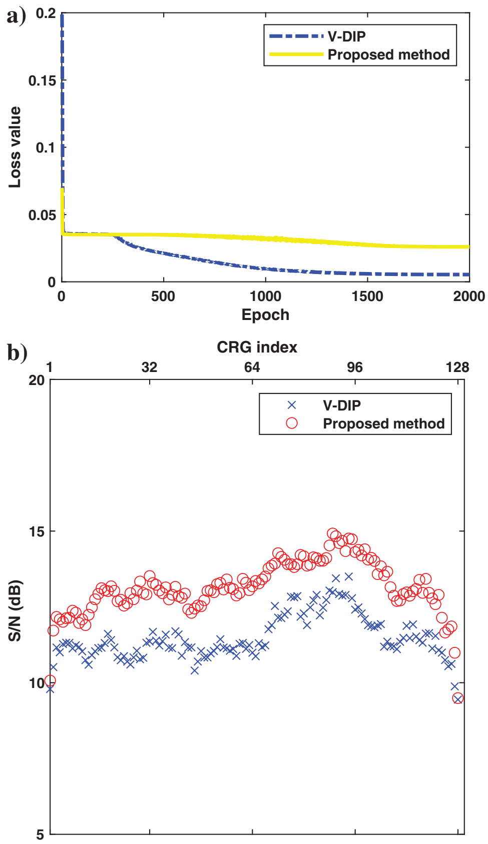



Figure 25. Loss curves and final S/N obtained using V-DIP and the proposed method on the Gulf of Suez data example. (a) Loss curves and (b) final S/N values.

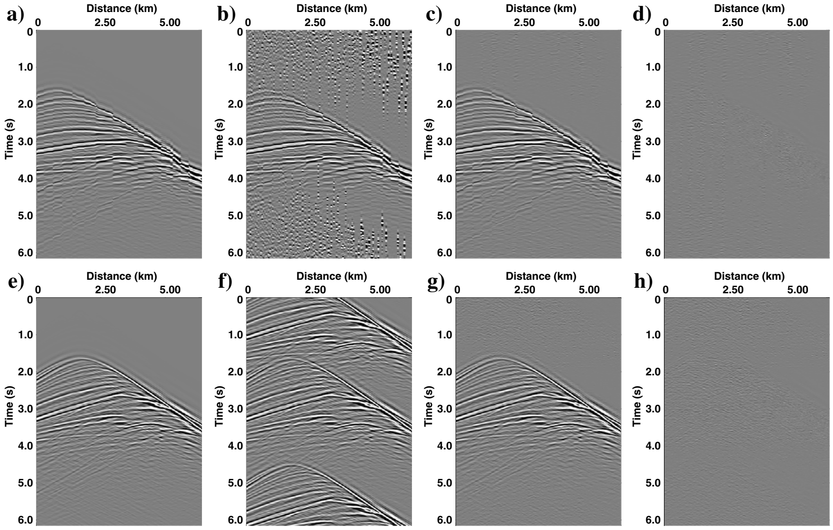



Figure 26. Deblending results of the proposed method on the irregular shot-distributed Marmousi data example. (a) An unblended CRG, (b) the corresponding pseudo-deblended CRG, (c) the deblended CRG obtained using the proposed method, (d) difference between the unblended and deblended CRG, (e) an unblended CSG, (f) the corresponding pseudo-deblended CSG, (g) the deblended CSG obtained using the proposed method, and (h) difference between the unblended and deblended CSG.

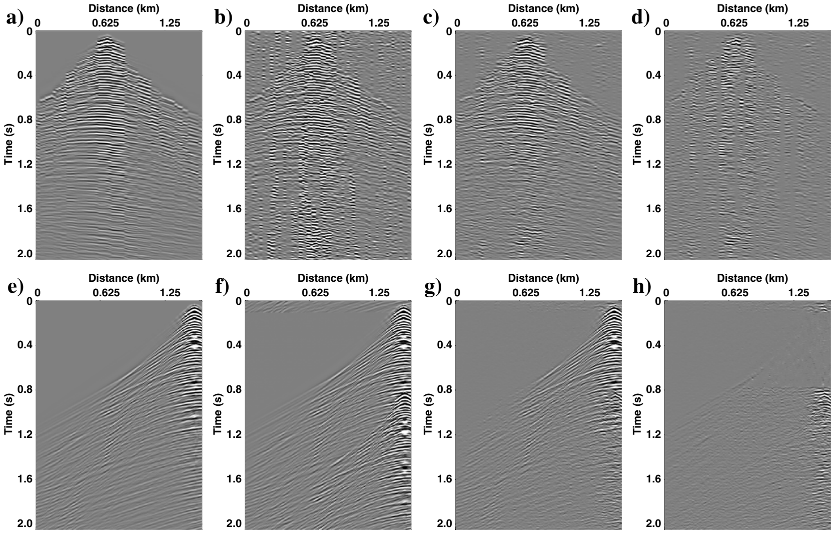



Figure 27. Deblending results of the proposed method on the irregular shot-distributed Gulf of Suez data example. (a) An unblended CRG, (b) the corresponding pseudo-deblended CRG, (c) the deblended CRG obtained using the proposed method, (d) difference between the unblended and deblended CRG, (e) an unblended CSG, (f) the corresponding pseudo-deblended CSG, (g) the deblended CSG obtained using the proposed method, and (h) difference between the unblended and deblended CSG.

## CONCLUSION

This study introduced an advancement in seismic deblending by proposing an INR-based approach. By modeling the deblending process as an implicit function within a deep neural network, the proposed INR-based method effectively exploits feature similarities among adjacent unblended CRGs, enabling robust data recovery despite the complexities introduced by blending operators. In addition, the blending-operator-informed training strategy enhances measurement consistency, thereby improving the accuracy of the deblended CRGs. Numerical experiments on synthetic and field data sets demonstrated that the proposed method outperforms the patched 2D Fourier sparse inversion method and achieves comparable deblending performance to the SSDeblend method. Notably, the proposed method operates self-supervised, eliminating the need for labeled training data. This characteristic makes it particularly advantageous for large-scale applications, where a small subset of data can be initially processed to generate training pairs. It can then train a superior network architecture with a supervised learning strategy. By leveraging the rapid inference capability of supervised deep learning, the remaining data set can be processed more efficiently, significantly accelerating the deblending workflow. Moreover, the current study focuses on 2D seismic deblending and assumes that noise components such as direct waves and random noise have been effectively suppressed before deblending. Future work will extend the proposed approach to 3D seismic deblending while accounting for the presence of noise, further enhancing its applicability to real-world seismic data processing.

## ACKNOWLEDGMENTS

This work was supported in part by the National Key Research and Development Program of China under grant no. 2021YFA0716903, in part by the National Natural Science Foundation of China under grant no. 42374135, in part by the Fundamental Research Funds for the Central Universities under grant no. xzy012023073 and in part by the Innovation Capability Support Program of Shaanxi under the grant no. 2021TD-08.

## DATA AND MATERIALS AVAILABILITY

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

## REFERENCES

Abma, R., D. Howe, M. Foster, I. Ahmed, M. Tanis, Q. Zhang, A. Arogunmati, and G. Alexander, 2015, Independent simultaneous source acquisition and processing: Geophysics, **80**, no. 6, WD37–WD44, doi: 10.1190/geo2015-0078.1.

Aftab, A., A. Morsali, and S. Ghaemmaghami, 2022, Multi-head ReLU implicit neural representation networks: IEEE International Conference on Acoustics, Speech and Signal Processing, 2510–2514.

Baardman, R., R. Hegge, and P. Zwartjes, 2020, Deblending via supervised transfer learning-DSA field data example: 81st Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.202010935.

Berkhout, A. G., 2013, Decentralized blended acquisition: 83rd Annual International Meeting, SEG, Expanded Abstracts, 7–11, doi: 10.1190/segam2013-0845.1.

Berkhout, A. G., G. Blacquière, and E. Verschuur, 2008, From simultaneous shooting to blended acquisition: 78th Annual International Meeting, SEG, Expanded Abstracts, 2831–2838, doi: 10.1190/1.3063933.

Berkhout, A. J. 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938, doi: 10.1190/1.2954035.

Brandolin, F., M. Ravasi, and T. Alkhalifah, 2024, PINNslope: Seismic data interpolation and local slope estimation with physics informed neural networks: Geophysics, **89**, no. 4, V331–V345, doi: 10.1190/geo2023-0323.1.

Chen, H., B. He, H. Wang, Y. Ren, S. N. Lim, and A. Shrivastava, 2021, NeRV: Neural representations for videos: Advances in Neural Information Processing Systems, Curran Associates, Inc., 21557–21568.

Chen, Y., 2015, Deblending using a space-varying median filter: Exploration Geophysics, **46**, 332–341, doi: 10.1071/EG14051.

Chen, Y., H. Chen, K. Xiang, and X. Chen, 2017, Preserving the discontinuities in leastsquares reverse time migration of simultaneous-source data: Geophysics, **82**, no. 3, S185–S196, doi: 10.1190/geo2016-0456.1.

Chen, Y., S. Fomel, and R. Abma, 2023, Joint deblending and source time inversion: Geophysics, **88**, no. 1, WA27–WA35, doi: 10.1190/geo2022-0149.1.

Chen, Y., S. Zu, Y. Wang, and X. Chen, 2020, Deblending of simultaneous source data using a structure-oriented space-varying median filter: Geophysical Journal International, **222**, 1805–1823, doi: 10.1093/gji/ggaa189.

Cheng, J., K. Chen, and M. D. Sacchi, 2015, Application of robust principal component analysis (RPCA) to suppress erratic noise in seismic records: 85th Annual International Meeting, SEG, Expanded Abstracts, 4646–4651, doi: 10.1190/segam2015-5869427.1.

Cheng, J., and M. D. Sacchi, 2015, Separation and reconstruction of simultaneous source data via iterative rank reduction: Geophysics, **80**, no. 4, V57–V66, doi: 10.1190/geo2014-0385.1.

Cheng, J., and M. D. Sacchi, 2016, Fast dual-domain reduced-rank algorithm for 3D deblending via randomized QR decomposition: Geophysics, **81**, no. 1, V89–V101, doi: 10.1190/geo2015-0292.1.

Cheng, S., Z. Cheng, C. Jiang, W. Mao, and Q. Zhang, 2024, An effective self-supervised learning method for attenuating various types of seismic noise: Geophysics, **89**, no. 6, V589–V604, doi: 10.1190/geo2023-0656.1.

Dai, W., X. Wang, and G. T. Schuster, 2011, Least-squares migration of multisource data with a deblurring filter: Geophysics, **76**, no. 5, R135–R146, doi: 10.1190/geo2010-0159.1.

Huang, W., R. Wang, X. Gong, and Y. Chen, 2018, Iterative deblending of simultaneous-source seismic data with structuring median constraint: IEEE Geoscience and Remote Sensing Letters, **15**, 58–62, doi: 10.1109/LGRS.2017.2772857.

Huo, S., Y. Luo, and P. G. Kelamis, 2012, Simultaneous sources separation via multidirectional vector-median filtering: Geophysics, **77**, no. 4, V123–V131, doi: 10.1190/geo2011-0254.1.

Ibrahim, A., and M. D. Sacchi, 2014, Simultaneous source separation using a robust radon transform: Geophysics, **79**, no. 1, V1–V11, doi: 10.1190/geo2013-0168.1.

Kinga, D., and J. B. Adam, 2015, A method for stochastic optimization: International Conference on Learning Representations.

Kontakis, A., and D. J. Verschuur, 2024, Focal deblending: Marine data processing experiences: Geophysical Prospecting, **72**, 333–352, doi: 10.1111/1365-2478.13404.

Li, C., C. C. Mosher, and Y. Ji, 2019, An amplitude-preserving deblending approach for simultaneous sources: Geophysics, **84**, no. 3, V185–V196, doi: 10.1190/geo2018-0240.1.

Li, J., D. Liu, D. Trad, and M. Sacchi, 2024, Robust unsupervised 5D seismic data reconstruction on both regular and irregular grid: Geophysics, **89**, no. 6, V537–V549, doi: 10.1190/geo2024-0098.1.

Liu, D., W. Gao, W. Xu, J. Li, X. Wang, and W. Chen, 2024, 5-D seismic data interpolation by continuous representation: IEEE Transactions on Geoscience and Remote Sensing, **62**, 1–11, doi: 10.1109/TGRS.2024.3431439.

Loshchilov, I., and F. Hutter, 2016, SGDR: Stochastic gradient descent with warm restarts: arXiv preprint, doi: 10.48550/arXiv.1608.03983.

Luiken, N., M. Ravasi, and C. Birnie, 2024, Integrating self-supervised denoising in inversion-based seismic deblending: Geophysics, **89**, no. 1, WA39–WA51, doi: 10.1190/geo2023-0131.1.

Luo, J., X. Wang, and W. Chen, 2022, Subsurface elastic parameter reconstruction based on seismic data from the high-speed trains using full waveform inversion: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–8, doi: 10.1109/TGRS.2022.3151352.

Mahdad, A., P. Doulgeris, and G. Blacquiere, 2011, Separation of blended data by iterative estimation and subtraction of blending interference noise: Geophysics, **76**, no. 3, Q9–Q17, doi: 10.1190/1.3556597.

Mahdad, A., P. Doulgeris, and G. Blacquière, 2012, Iterative method for the separation of blended seismic data: Discussion on the algorithmic aspects: Geophysical Prospecting, **60**, 782–801, doi: 10.1111/j.1365-2478.2012.01084.x.

Mansour, H., H. Wason, T. T. Lin, and F. J. Herrmann, 2012, Randomized marine acquisition with compressive sampling matrices: Geophysical Prospecting, **60**, 648–662, doi: 10.1111/j.1365-2478.2012.01075.x.

Mildenhall, B., P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, 2021, NeRF: Representing scenes as neural radiance fields for view synthesis: Communications of the ACM, **65**, 99–106, doi: 10.1145/3503250.

Molaei, A., A. Aminimehr, A. Tavakoli, A. Kazerouni, B. Azad, R. Azad, and D. Merhof, 2023, Implicit neural representation in medical imaging: A comparative survey: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2381–2391.

Rahaman, N., A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, 2019, On the spectral bias of neural networks: Proceedings of the 36th International Conference on Machine Learning, 5301–5310.

Richardson, A., and C. Feller, 2019, Seismic data denoising and deblending using deep learning: arXiv preprint, doi: 10.48550/arXiv.1907.01497.

Shen, L., J. Pauly, and L. Xing, 2024, NeRP: Implicit neural representation learning with prior embedding for sparsely sampled image reconstruction: IEEE Transactions on Neural Networks and Learning Systems, **35**, 770–782, doi: 10.1109/TNNLS.2022.3177134.

Shi, W., J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, 2016, Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Sitzmann, V., J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, 2020, Implicit neural representations with periodic activation functions: Advances in Neural Information Processing Systems, Curran Associates, Inc., 7462–7473.

Spitz, S., G. Hampson, and A. Pica, 2008, Simultaneous source separation: A prediction subtraction approach: 78th Annual International Meeting, SEG, Expanded Abstracts, 2811–2815, doi: 10.1190/1.3063929.

Sun, J., 2022, Deep learning-based seismic data processing for attenuation of interference noise and deblending in the shot domain: Ph.D thesis, University of Oslo.

Sun, J., S. Hou, V. Vinje, G. Poole, and L.-J. Gelius, 2022, Deep learning-based shot-domain seismic deblending: Geophysics, **87**, no. 3, V215–V226, doi: 10.1190/geo2020-0865.1.

Sun, J., S. Slang, T. Elboth, T. L. Greiner, S. McDonald, and L.-J. Gelius, 2020, A convolutional neural network approach to deblending seismic data: Geophysics, **85**, no. 4, WA13–WA26, doi: 10.1190/geo2019-0173.1.

Sun, Y., J. Liu, M. Xie, B. Wohlberg, and U. S. Kamilov, 2021, CoIL: Coordinate-based internal learning for tomographic imaging: IEEE Transactions on Computational Imaging, **7**, 1400–1412, doi: 10.1109/TCI.2021.3125564.

Tancik, M., P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, 2020, Fourier features let networks learn high frequency functions in low dimensional domains: Advances in Neural Information Processing Systems, Curran Associates, Inc., 7537–7547.

Ulyanov, D., A. Vedaldi, and V. Lempitsky, 2018, Deep image prior: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Verschuur, D. J. E., and A. J. G. Berkhout, 2011, Seismic migration of blended shot records with surface-related multiple scattering: Geophysics, **76**, no. 1, A7–A13, doi: 10.1190/1.3521658.

Wang, B., J. Li, J. Luo, Y. Wang, and J. Geng, 2021, Intelligent deblending of seismic data based on U-Net and transfer learning: IEEE Transactions on Geoscience and Remote Sensing, **59**, 8885–8894, doi: 10.1109/TGRS.2020.3048746.

Wang, K., T. Hu, and S. Wang, 2023a, Iterative deblending using unsupervised learning with double-deep neural networks: Geophysics, **88**, no. 3, V187–V205, doi: 10.1190/geo2022-0299.1.

Wang, S., W. Hu, P. Yuan, X. Wu, Q. Zhang, P. Nadukandi, G. O. Botero, and J. Chen, 2023b, A self-supervised deep learning method for seismic data deblending using a blindtrace network: IEEE Transactions on Neural Networks and Learning Systems, **34**, 3405–3414, doi: 10.1109/TNNLS.2022.3188915.

Wu, J., and M. Bai, 2019, Iterative deblending based on the modified singular spectrum analysis: Journal of Seismic Exploration, **28**, 1–20.

Xu, W., V. Lipari, P. Bestagini, M. Ravasi, W. Chen, and S. Tubaro, 2022a, Intelligent seismic deblending through deep preconditioner: IEEE Geoscience and Remote Sensing Letters, **19**, 1–5, doi: 10.1109/LGRS.2022.3193716.

Xu, W., Y. Zhou, D. Liu, X. Wang, and W. Chen, 2021, Seismic simultaneous source separation via an unsupervised deep learning method: 82nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.202112628.

Xu, W., Y. Zhou, D. Liu, X. Wang, and W. Chen, 2022b, Seismic intelligent deblending via plug and play method with blended CSGS trained deep CNN Gaussian denoiser: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–13, doi: 10.1109/TGRS.2022.3166029.

Yu, Z., R. Abma, J. Etgen, and C. Sullivan, 2017, Attenuation of noise and simultaneous source interference using wavelet denoising: Geophysics, **82**, no. 3, V179–V190, doi: 10.1190/geo2016-0240.1.

Zhou, Y., W. Chen, and J. Gao, 2014, Separation of seismic blended data by sparse inversion over dictionary learning: Journal of Applied Geophysics, **106**, 146–153, doi: 10.1016/j.jappgeo.2014.04.010.

Zhou, Y., J. Gao, W. Chen, and P. Frossard, 2016, Seismic simultaneous source separation via patchwise sparse representation: IEEE Transactions on Geoscience and Remote Sensing, **54**, 5271–5284, doi: 10.1109/TGRS.2016.2559514.

Zu, S., J. Cao, S. Qu, and Y. Chen, 2020, Iterative deblending for simultaneous source data using the deep neural network: Geophysics, **85**, no. 2, V131–V141, doi: 10.1190/geo2019-0319.1.

Zu, S., C. Ke, C. Hou, J. Cao, and H. Zhang, 2022, End-to-end deblending of simultaneous source data using transformer: IEEE Geoscience and Remote Sensing Letters, **19**, 1–5, doi: 10.1109/LGRS.2022.3174106.

Zu, S., H. Zhou, W. Mao, D. Zhang, C. Li, X. Pan, and Y. Chen, 2017, Iterative deblending of simultaneous-source data using a coherency-pass shaping operator: Geophysical Journal International, **211**, 541–557, doi: 10.1093/gji/ggx324.

Zu, S., H. Zhou, R. Wu, W. Mao, and Y. Chen, 2019, Hybrid-sparsity constrained dictionary learning for iterative deblending of extremely noisy simultaneous-source data: IEEE Transactions on Geoscience and Remote Sensing, **57**, 2249–2262, doi: 10.1109/TGRS.2018.2872416.

Biographies and photographs of the authors are not available.