

Unsupervised Deep Learning for Ground Roll and Scattered Noise Attenuation

Dawei Liu^{ID}, Mauricio D. Sacchi^{ID}, Member, IEEE, Xiaokai Wang^{ID}, Member, IEEE, and Wenchao Chen^{ID}

Abstract—The attenuation of coherent noise in land seismic data, specifically ground roll and near-surface scattered energy, remains a longstanding challenge. Although recent advances in deep learning have improved signal separation from coherent noise, supervised methods are limited by the necessity for realistic training samples. To circumvent this issue, we propose an unsupervised deep learning approach to attenuate ground roll and scattered energy, eliminating the requirement for training labels. Our method leverages the inherent low-frequency bias of a generator network, which is naturally prone to learn self-similar features during training. This empowers the network to extract the desired component exhibiting self-similarity in the time-space domain, while disregarding unwanted components. Notably, horizontal components in seismic data exhibit pronounced self-similarity. To enhance the self-similarity of ground roll, we apply a linear moveout (LMO) correction to horizontally align it and utilize the generator network for separation. Additionally, for scattered energy attenuation, we employ the generator network to extract flattened reflections after normal moveout (NMO) correction. Our strategy distinctively merges model-driven procedures, specifically NMO and LMO, anchored in the geological velocity model. The synergy between data-driven deep learning and model-driven processes underscores the success of our approach. We demonstrate the validity of our proposed method using both synthetic and field shot data. The field data examples highlight the superior attenuation capabilities of our method, surpassing conventional denoising techniques by effectively reducing both random and coherent noise.

Index Terms—Ground roll, scattered noise, seismic data denoising, self-similar features, unsupervised learning.

I. INTRODUCTION

GROUND roll and near-surface scattered energy present significant challenges as the most severe forms of coherent noise in land seismic data. Ground roll exhibits distinctive characteristics such as dispersion, high amplitude, low frequency, and low speed [1], [2], [3], while near-surface scattered energy manifests complex diffraction patterns deriving from secondary events [4], [5], [6]. The presence of such coherent noise masks the desired reflections, necessitating

Manuscript received 20 June 2023; revised 6 September 2023; accepted 10 October 2023. Date of publication 19 October 2023; date of current version 27 October 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFA0716904 and in part by the National Natural Science Foundation of China under Grant 41974131 and Grant 41774135. The work of Dawei Liu was supported by the China Scholarship Council. (*Corresponding author: Wenchao Chen*.)

Dawei Liu, Xiaokai Wang, and Wenchao Chen are with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: liudawei2015@stu.xjtu.edu.cn; xkwang@xjtu.edu.cn; wencchen@xjtu.edu.cn).

Mauricio D. Sacchi is with the Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada (e-mail: msacchi@ualberta.ca).

Digital Object Identifier 10.1109/TGRS.2023.3325324

effective attenuation techniques that preserve the integrity of reflections for subsequent processing and interpretation tasks.

Coherent noise attenuation has been extensively studied in previously published studies. They can be broadly classified into four groups based on different prior knowledge regarding the physical representation of useful signals or coherent noise. The first and most commonly employed group is filtering, which exploits a muting region between useful signals and coherent noise in a specific transform domain. Bandpass filtering and $f\text{-}k$ filtering [7], [8] are widely used in the industry to remove coherent noise or adopted as a preprocessing step for subsequent fine separation [9]. Benefiting from advanced multiscale signal representation, wavelet-domain filtering [10], [11] and curvelet transform filtering [12], [13] garner considerable attention among researchers. These methods have demonstrated strong capability to effectively attenuate coherent noise when a distinct boundary exists in the transform domain. However, their performance in real data scenarios is often constrained due to the significant overlap between reflections and coherent noise in the transform domain. Early studies also emphasize the utilization of the signal predictability to attenuate coherent noise and retrieve useful signals, or vice versa. Typical examples include $f\text{-}x$ prediction filtering [14], projection filtering [15], adaptive prediction filtering [16], [17], and nonstationary prediction filtering [18]. However, the presence of large-amplitude coherent noise invariably disrupts reflection continuity, resulting in degraded denoising performance. The third group of methods is characterized by low-rank assumptions on either useful signals or coherent noise, including low-rank factorization [19], [20], [21] and nuclear norm minimization [22], [23]. Most of these methods operate in the Fourier domain, capitalizing on the low-rank property of linear events. Consequently, their effectiveness in denoising is contingent upon the linearity of the events, and they may yield unsatisfactory results when confronted with highly curved useful signals or coherent noise. Finally, sparse priors are increasingly used for coherent noise attenuation, requiring sparse coefficients for the transformed representation of useful signals or coherent noise [24], [25], [26]. Nonetheless, meeting this sparsity assumption can be challenging because useful signals and coherent noise often exhibit similar characteristics and cannot be effectively distinguished in any sparse transform, leading to suboptimal outcomes.

Most of the aforementioned methods are directly applicable to ground roll or scattered noise removal. However, to further improve the separation performance, particularly in strong-energy situations, additional adaptations are necessary to tailor the methods specifically for these noise types. One effective

approach is to integrate normal moveout (NMO) and linear moveout (LMO) corrections, which incorporate geological knowledge from wave propagation velocities. The horizontal nature of primary reflections after NMO provides a valuable advantage for enhancing denoising methods in identifying reflections [27]. For example, Porsani et al. [28] proposed a ground-roll attenuation method based on singular value decomposition (SVD). The SVD computation was performed on the flattened reflections after NMO, yielding better results than $f\text{-}k$ filtering methods. Chiu [29] proposes a method to improve the attenuation of both ground roll and scattered noise by incorporating NMO and a randomizing operator into multichannel singular spectrum analysis (MSSA) [30]. The randomizing operator disrupts the coherence of the noise, thereby transforming it into incoherent noise. Meanwhile, the primary reflections, after undergoing NMO, maintain their nearly horizontal characteristics, which facilitates better differentiation between useful signals and coherent noise for denoising techniques. The LMO contributes another strategy to improving denoising methods by strengthening the coherency of coherent noise, thereby enabling their subtraction from the original data. Chiu and Howell [31] apply LMO to align coherent noise and use an eigenimage filter to extract it. Subsequently, they subtract the extracted noise, resulting in a higher preservation of useful signals. Based on the literature discussed above, we conclude that NMO and LMO can substantially improve denoising performance, inspiring us to incorporate them into recently developed deep learning algorithms.

Recently, deep learning has garnered significant research attention, primarily in the field of image processing, due to its remarkable ability to recognize patterns and reveal hidden correlations from complex data [32], [33]. Naturally, supervised deep learning is introduced to various applications in seismic data processing, including random noise denoising [34], [35], ground roll attenuation [36], [37], scattered noise removal [38], deblending [39], [40], resolution improvement [41], and strong background-noise separation [42]. Constructing training samples and then feeding them to the network for training is a standard workflow for denoising methods based on supervised deep learning. However, this approach relies on a large set of clean-noisy seismic data pairs. Despite the significant improvement in processing speed with supervised deep learning, generating a large number of genuinely realistic-looking synthetic data or pseudo-labels for training purposes remains a challenge for coherent noise attenuation. Additionally, the limited generalization ability has an adverse effect to practical applications.

Unsupervised deep learning methods are able to learn from the noisy data itself and can achieve satisfactory denoising results even with limited training data [43], [44]. However, the majority of these methods are specifically designed for attenuating random noise [45], [46], [47]. Therefore, there is a pressing demand to develop unsupervised deep learning techniques tailored to denoising coherent noise [48], [49], [50], [51]. Motivated by these insights, we propose a two-step method for attenuating ground roll and scattered noise based on unsupervised deep learning. Our approach leverages the intrinsic low-frequency preference [52], [53] of deep learning

and incorporates geological knowledge of approximate moveout velocities. First, we employ LMO correction on the raw seismic data to flatten the ground roll. The LMO correction renders the ground roll nearly horizontal and self-similar, while other signal components remain low self-similarity. Subsequently, we use a convolutional generator network (CGN) to extract the self-similar ground roll. In the second step, NMO correction is applied to align the reflections horizontally. Similarly, another CGN focuses on extracting horizontally self-similar reflections while preserving residual coherent noise. As the first step effectively eliminates linear ground roll, the second step primarily targets suppressing scattered energy. Compared to our previous conference abstract [54], we introduce the application of LMO, leading to significantly improved ground roll attenuation performance. Additionally, experimental comparisons with conventional methods are conducted to validate the advantages of our proposed approach.

II. METHOD

The proposed methodology for attenuating ground roll and scattered noise relies solely on noisy label learning, which is a zero-shot method that eliminates the need for constructing a noisy dataset. In this section, we sequentially describe the denoising principles, the U-shaped architecture with skip connections, and the loss function employed to extract self-similar informative features directly from the noisy data.

A. Denoising Principles

A useful intrinsic preference of neural networks for designing unsupervised deep learning methods is the low-frequency implicit bias, which refers to the tendency of networks to fit signal components from low- to high-frequency during training, also known as the frequency principle [52] or spectral bias [53]. Many recent studies contribute to the discovery of this experimental phenomenon. Zhang et al. [55] draw considerable attention to a non-overfitting puzzle in deep neural networks, contradicting traditional generalization theory which suggests that models with excessive parameters easily overfit the data. Ulyanov et al. [56] also observe that generator networks tend to fit low-frequency and self-similarity features prior to other features during the training process, naming this property as “deep image prior.” Rahaman et al. [53] highlight the learning bias of deep networks toward low-frequency functions through their experimental results.

To address the puzzle of why heavily parameterized neural networks do not overfit the data, Xu [57] contributes a decay rate analysis method of the loss function from a frequency perspective, concluding that the low-frequency component has a faster decay and is easier to fit. Then, Luo et al. [58] extend this analysis to handle infinite samples. However, due to the curse of dimensionality, their 1-D explanation is challenging to comprehend general deep neural networks with high dimensions. Meanwhile, several other idealized models are also introduced to provide rigorous mathematical proofs [59], [60], [61], [62], [63].

Although definitive theoretical studies explaining the low-frequency bias are still under development [64], this

property provides valuable insight and essential guidance for the practical application of unsupervised deep learning. Several algorithms capitalize on this property to extract useful signals from seismic data [65], [66]. These useful signals often consist of low-frequency components or exhibit inherent self-similar features, which can effectively be captured by deep neural networks. Specifically, these algorithms employ a GCN with randomly initialized inputs to fit the noisy seismic data. During the training process, the GCN learns to represent the self-similar useful components at early stages while subconsciously excluding unwanted components. After a specified number of iterations in the optimization process, the GCN successfully extracts all the desired signals. Then, the subtraction of these extracted signals from the raw seismic data leads to effective noise attenuation. Liu et al. [67] highlight that horizontal events exhibit high self-similarities, and NMO correction accordingly enhances the self-similarities of reflections, facilitating successful reflection extraction by the generator. In light of these findings, our proposed method employs two distinct CGNs to attenuate ground roll and scattered noise from noisy seismic data after LMO and NMO corrections, respectively. Specifically, these networks are trained through unsupervised learning to sequentially extract highly self-similar ground roll and reflections, leading to satisfactory denoising results.

B. Model Formulation

We represent the raw seismic data as a matrix \mathbf{Y} , which consists of useful signals, coherent noise, and random noise. Seismic reflections, originating from interfaces between layers with distinct acoustic impedances, serve as the useful signals in seismic exploration. Among the noise components, ground roll, characterized by its low-frequency, high-amplitude events, is a coherent noise that mainly travels along the ground surface. Another dominant form of coherent noise is the near-surface scattered noise. This noise arises from the scattering of seismic waves due to small-scale heterogeneities in the subsurface. In contrast, random noise includes incoherent disturbances from varied sources, such as electrical interference, ambient seismic sounds, or even environmental factors like wind. It is crucial to note that while these signal components may intertwine closely during wave propagation, they are typically considered to be simply additive during the processing phase [2], [11], [68]. Specifically, our model can be expressed as follows:

$$\mathbf{Y} = \mathbf{R} + \mathbf{G} + \mathbf{S} + \mathbf{N}. \quad (1)$$

Here, \mathbf{R} denotes the seismic reflections, \mathbf{G} represents ground roll, \mathbf{S} represents scattered noise, and \mathbf{N} corresponds to the random noise.

To achieve successful separation of ground roll and scattered noise, traditional methods pay particular attention to characterizing the above signal components with different prior knowledge. The objective function is formulated as a penalized cost

$$\min_{\mathbf{R}, \mathbf{G}, \mathbf{S}} E(\mathbf{Y}, \mathbf{R}, \mathbf{G}, \mathbf{S}) + \mathcal{P}_R(\mathbf{R}) + \mathcal{P}_G(\mathbf{G}) + \mathcal{P}_S(\mathbf{S}) \quad (2)$$

where E represents the reconstruction error, in other words, the data fidelity or the data-fitting term; the notation \mathcal{P} denotes a regularization or a priori term, which is a penalization imposed on the signal component of the proposed model.

Different probability distributions of \mathbf{N} have led to the introduction of various E models to represent it, such as the Huber loss or the l_1 loss. In this work, we simply consider \mathbf{N} as Gaussian noise and define E as the mean squared error. Therefore, (2) can be rewritten as

$$\min_{\mathbf{R}, \mathbf{G}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{R} - \mathbf{G} - \mathbf{S}\|_2^2 + \mathcal{P}_R(\mathbf{R}) + \mathcal{P}_G(\mathbf{G}) + \mathcal{P}_S(\mathbf{S}). \quad (3)$$

To achieve successful signal recovery from raw seismic data, the appropriate regularization terms \mathcal{P} are necessary. Traditional methods utilize problem-specific regularization, such as low-rank or sparsity assumptions, to characterize \mathbf{R} , \mathbf{G} , and \mathbf{S} . The optimization process for solving (3) introduces non-linear reconstruction mappings $\hat{\mathbf{R}} = h_R(\mathbf{Y})$, $\hat{\mathbf{G}} = h_G(\mathbf{Y})$, and $\hat{\mathbf{S}} = h_S(\mathbf{Y})$. These traditional methods, categorized as model-based approaches, are generally based on an energy model and explicit regularization with deterministic mathematical formulations.

Different from traditional methods, the low-frequency bias of GCN serves as an implicit regularization. The term “implicit” refers to the fact that the regularization effect cannot be expressed explicitly in the objective function. We incorporate this property into our model through neural network parametrization as follows:

$$\begin{aligned} & \min_{\theta_R, \theta_G} \frac{1}{2} \|\mathbf{Y} - \mathbf{R} - \mathbf{G} - \mathbf{S}\|_2^2 \\ & \text{s.t. } \mathbf{R} = f_{\theta_R}(\mathbf{Z}_R), \quad \mathbf{G} = f_{\theta_G}(\mathbf{Z}_G) \end{aligned} \quad (4)$$

where f_{θ_R} and f_{θ_G} are distinct GCNs used to parameterize \mathbf{R} and \mathbf{G} , respectively. The symbols θ_R and θ_G denote their corresponding weight matrices and biases. Similarly, \mathbf{Z}_R and \mathbf{Z}_G are randomly initialized matrices serving as inputs to the GCNs. Here, we ignore the penalty of \mathbf{S} in (3) due to its complexity. However, this is not an issue since we can obtain \mathbf{S} by subtracting \mathbf{G} and \mathbf{R} from \mathbf{Y} . In this way, the model is considerably simplified and easier to implement. Inserting the neural network parametrization into (4) generates an unconstrained reconstruction problem

$$\min_{\theta_R, \theta_G} \frac{1}{2} \|\mathbf{Y} - f_{\theta_R}(\mathbf{Z}_R) - f_{\theta_G}(\mathbf{Z}_G) - \mathbf{S}\|_2^2. \quad (5)$$

The network parametrization process involves a progressive recovery mapping from \mathbf{Z} to the desired signals through network training iterations. Unfortunately, solving (5) directly through joint optimization of θ_R and θ_G is challenging due to its high non-convexity. To simplify it, we split (5) into two subproblems

$$\min_{\theta_G} \frac{1}{2} \|\mathbf{Y} - \mathbf{R} - f_{\theta_G}(\mathbf{Z}_G) - \mathbf{S}\|_2^2 \quad (6a)$$

$$\min_{\theta_R} \frac{1}{2} \|\mathbf{Y} - \mathbf{G} - f_{\theta_R}(\mathbf{Z}_R) - \mathbf{S}\|_2^2 \quad (6b)$$

and solve them sequentially. To make the former subproblem solvable and reduce uncertainties, we apply LMO to \mathbf{G} to enhance its self-similarity. In other words, the network

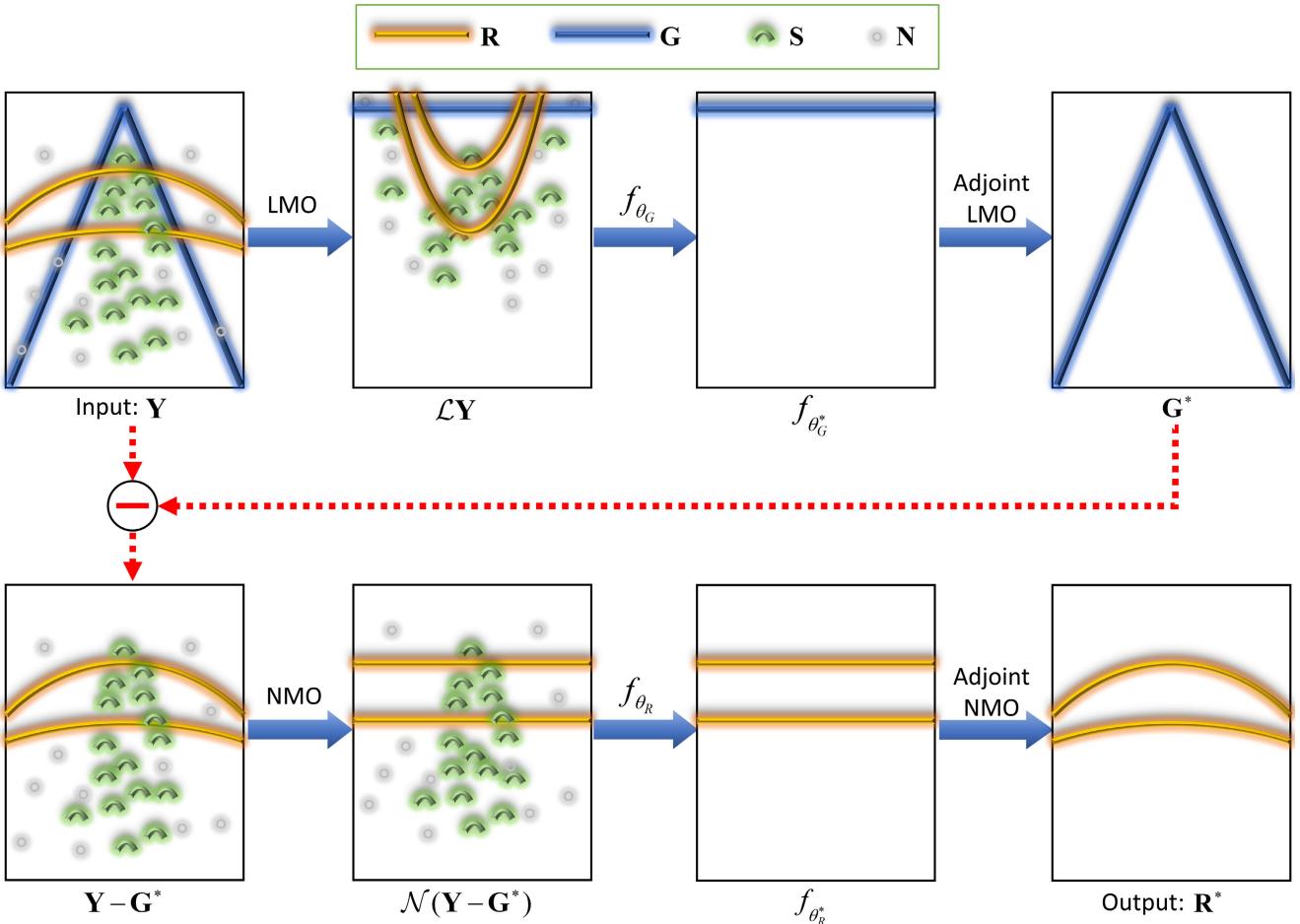


Fig. 1. Illustration of noise attenuation by self-similar feature extraction networks. The reflections, ground roll, scattered noise, and random noise are symbolized by \mathbf{R} , \mathbf{G} , \mathbf{S} , and \mathbf{N} , respectively.

parametrization is performed on the flattened \mathbf{G} , i.e., $\mathcal{L}\mathbf{G} = f_{\theta_G}(\mathbf{Z}_G)$, where \mathcal{L} denotes the LMO operator. By assuming an initialization of $\mathbf{R} = 0$ and $\mathbf{S} = 0$, the objective of ground roll attenuation in (6a) is equivalent to determining the optimal network parameter θ^* that minimizes the energy function

$$\theta_G^* = \arg \min_{\theta_G} \frac{1}{2} \|\mathcal{L}\mathbf{Y} - f_{\theta_G}(\mathbf{Z}_G)\|_2^2 \quad (7)$$

where the only undetermined parameter is θ_G . This parameter can be solved using classical gradient descent introduced in the subsequent Section II-D. Once θ^* is determined, the separated ground roll can be quickly obtained from the output of GCN by $\mathbf{G}^* = \mathcal{L}^\dagger f_{\theta_G^*}(\mathbf{Z}_G)$, where \mathcal{L}^\dagger denotes the adjoint LMO operator.

Then, we incorporate the NMO operator \mathcal{N} to facilitate solving the subproblem (6b). Similarly, we presume an initialization of $\mathbf{S} = 0$. The reflections are horizontally aligned through NMO, allowing us to extract these self-similar reflections by minimizing the following objective function:

$$\theta_R^* = \arg \min_{\theta_R} \frac{1}{2} \|\mathcal{N}(\mathbf{Y} - \mathbf{G}^*) - f_{\theta_R}(\mathbf{Z}_R)\|_2^2. \quad (8)$$

Upon obtaining θ_R^* , we can reconstruct the extracted reflections by $\mathbf{R}^* = \mathcal{N}^\dagger f_{\theta_R^*}(\mathbf{Z}_R)$, where \mathcal{N}^\dagger denotes the adjoint NMO operator. This reconstruction leads to the successful attenuation of scattered noise. Additionally, we can calculate the separated components by $\hat{\mathbf{S}} = \mathbf{Y} - \mathbf{G}^* - \mathbf{R}^*$, which consists

of a mixture of scattered noise and random noise. The above procedures are illustrated in Fig. 1.

C. Network Architecture

The network architecture employed in our work is a U-Net type fully convolutional network, which has proven effective in various image and signal processing tasks. This architecture is composed of three fundamental components: downsampling blocks, skip blocks, and upsampling blocks. These components are constructed using a combination of convolutional layers, batch normalization, downsampling layers, upsampling layers, and activation function layers. The overall structure of the network can be visualized in Fig. 2, providing a clear representation of its components and their connections.

Our network design is primarily motivated by the exploitation of intrinsic self-similarities present in seismic data, which often contain valuable information. By incorporating convolutional filters, the network can leverage the weight replication mechanism, allowing the same filters to recognize common features across different locations. These common features are self-similar and prove particularly beneficial in seismic data analysis. Furthermore, operating at different scales is highly advantageous in this context. Working across multiple scales enables us to capture features at varying levels of granularity. This results in a comprehensive model that

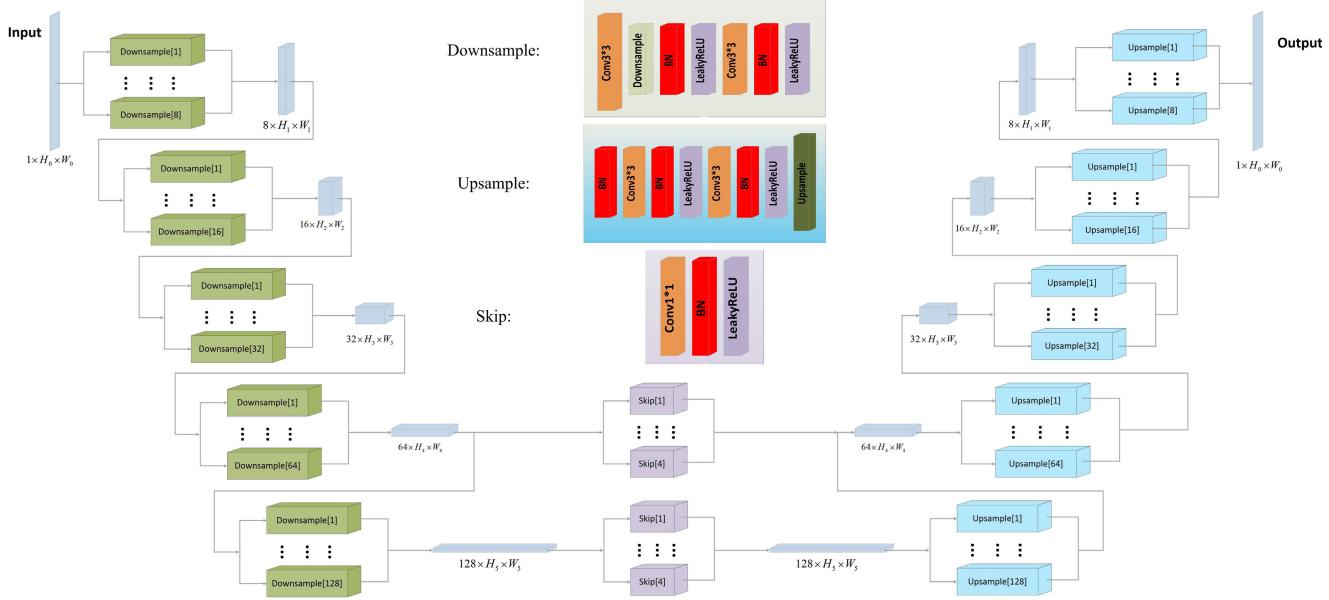


Fig. 2. Network architecture used for self-similar feature extraction.

accurately represents complex patterns and captures nonlinear relationships in seismic signals. Moreover, multiscale analysis permits the identification of finer features that are typically challenging to detect.

To facilitate comprehensive feature extraction across multiple scales, we incorporate five downsample blocks and five upsample blocks into our network architecture. This configuration effectively captures both local and global features within seismic data, significantly enhancing the ability to extract multiscale self-similar patterns with meaningful information. During the downsampling process, we carefully increase the number of feature channels from 8 to 128. This selection strikes a balance between feature extraction capability and computational efficiency. By gradually increasing the number of feature channels, the network can capture more complex and abstract features as the scale decreases. This enhances the overall representational power of our network while maintaining computational tractability. The upsampling process is symmetric to the downsampling process, gradually decreasing the number of feature channels from 128 to 8. This symmetrical upsampling process enables the reconstruction of the original data size while preserving the self-similar features and meaningful multiscale information extracted during the downsampling process.

In addition to the original U-Net architecture [69], we introduce several modifications to further enhance its functionality and address specific challenges encountered in seismic data processing. First, we replace bilinear interpolation, commonly used in traditional U-Net architectures, with transposed convolution. This modification aims to overcome potential checkerboard artifacts that may arise during upsampling. By utilizing transposed convolution, we ensure better spatial coherence and preserve the integrity of the extracted features [70].

Second, we have opted for leaky rectified linear unit (ReLU) activation functions instead of traditional ReLU. This choice is motivated by the need to prevent neuron annihilation and

promote a more robust training process. Leaky ReLU allows for the propagation of small negative gradients, ensuring information retention even in regions with negative activations. This modification enhances the network's ability to capture subtle details and improves the overall gradient flow during training.

Finally, we have implemented a skipping block strategy to address the issue of gradient vanishing. This strategy involves establishing direct connections between the output of each downsample block and its corresponding upsample block. By incorporating these skip connections, we ensure that important gradient information is preserved and propagated throughout the network during the training process. This approach mitigates the problem of gradient vanishing and enhances the stability and convergence speed of our network.

In summary, the modifications we have introduced, coupled with the underlying U-Net architecture, contribute to the effectiveness and performance of our network in extracting self-similar features from seismic data. By leveraging the multiscale feature extraction capability, along with careful adjustments and design choices, our network is well-suited to address the challenges posed by ground roll and scattered noise attenuation.

D. Network Training and Denoising Process

To achieve effective ground roll and scattered noise attenuation, the key lies in solving the optimization subproblems (7) and (8) sequentially, leveraging the inherent low-frequency bias of GCN that suggests a learning progression from simple to complex features [71].

The first step involves solving (7), which focuses on minimizing the reconstruction error under the L_2 norm. In this optimization subproblem, the coding matrix \mathbf{Z}_G is sampled from a uniform distribution in the range of $[-1, 1]$ and possesses the same spatial dimensions as the raw seismic data \mathbf{Y} . This is an unsupervised network training task since

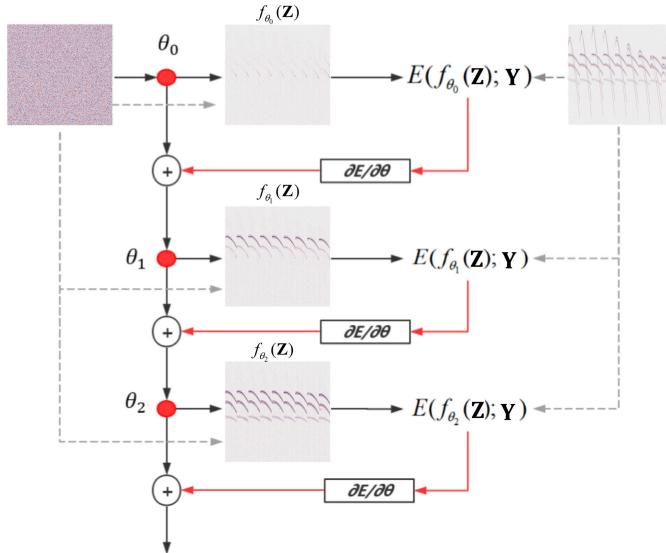


Fig. 3. Illustration of training process for extracting reflections. The subscripts of θ denote the various iterative stages in the self-similarity extraction process by f_{θ_R} .

it relies solely on the determined coding matrix \mathbf{Z} and the observed seismic data \mathbf{Y} . To tackle this optimization challenge, we employ the ADAM optimizer to train the network by gradient descent, iteratively updating the parameters θ_G . At the beginning, the network parameters are randomly initialized as θ_{G_0} . As the training progresses, the reconstruction error in (7) gradually diminishes. The optimization process can be regarded as the reconstruction process of \mathbf{Y} with LMO by GCN. At the i th iteration, the parameters θ_{G_i} are mapped to a network output $f_{\theta_{G_i}}(\mathbf{Z})$.

The features extracted by GCN exhibit variations across different training stages, influenced by the inherent low-frequency bias. Additionally, the specific network structure depicted in Fig. 2 enhances the network's ability to extract self-similar features at multiple scales. As a result, the network is capable of reconstructing signal components with pronounced self-similar features in earlier iterations. Conversely, reconstructing other signal components with fewer self-similarities necessitates an extended period of iterative training. By virtue of this sequential reconstruction process, GCN is an extremely useful tool for separating signal components.

Before commencing network training, we perform LMO correction on the raw seismic data \mathbf{Y} to flatten ground roll \mathbf{G} . The horizontally aligned ground roll \mathbf{G} exhibits a substantial degree of self-similarity compared to other components. Therefore, leveraging the low-frequency bias of GCN, which aggressively extracts self-similarities during the optimization process, allows for the gradual extraction of \mathbf{G} prior to other signal components. By carefully selecting the number of iterations, we can successfully separate \mathbf{G} from \mathbf{Y} . Notably, we have observed that a broad range of iteration numbers yield satisfactory results, demonstrating the robust adaptability of the approach.

The process of solving (8) follows a similar approach. By applying NMO, the reflections \mathbf{R} are encouraged to exhibit a higher degree of self-similarity compared to other signal

components. The optimization process itself tends to favor solutions that exhibit simpler patterns, implicitly facilitating the extraction of flattened reflections. To achieve a satisfactory separation of \mathbf{R} , we continue this iterative network optimization until the desired flattened reflections are fully extracted. Subsequently, we reverse the NMO correction, effectively isolating \mathbf{R} from the remaining signal components. A schematic visual representation of the training process is depicted in Fig. 3. The extracted reflections consistently exhibit a high level of fidelity and accuracy, thus affirming the efficacy and feasibility of our proposed technique.

III. EXAMPLES

The performance of our method is evaluated in both synthetic and field data. To conduct the numerical experiments, we utilized an NVIDIA GTX 1080-Ti graphics processing unit, which allowed us to process a single gather in approximately 2 min. The iteration parameter for early stopping plays a crucial role in the success of our method. To account for varying noise levels, we fine-tune this parameter according to individual denoising results. Throughout all experiments, stopping iterations range from 1500 to 2500, ensuring a thorough evaluation. To facilitate this stopping iteration selection and ensure training stability, we employ a relatively small learning rate of $5e^{-4}$. Additionally, at each iteration, we introduce Gaussian noise with a specified variance to the input \mathbf{Z} , aiming to further enhance the robustness of our approach. The chosen variance is set to 0.01, providing an appropriate perturbation level. In the field data example, we compare our results with those obtained from the high-resolution Radon transform (HRT) method. The HRT method, renowned for its effectiveness in seismic coherent noise attenuation, is widely employed within the industry. This comparative analysis allows us to assess the performance of our proposed technique and demonstrate its efficacy and superiority in effectively mitigating ground roll and scattered noise.

A. Synthetic Data Example

To evaluate the effectiveness of our method, we generate 3-D seismic synthetic data using an acquisition geometry depicted in Fig. 4. This acquisition geometry serves as the foundation for conducting the following two experiments, allowing us to assess the performance of our network in extracting ground roll and reflections, which are crucial in solving (7) and (8), respectively.

The first dataset is generated to test the performance of our network in extracting ground roll for solving (7). Based on the velocity models shown in Figs. 5 and 6(a), illustrates the synthetic data modeled in the frequency domain, consisting of three hyperbolic reflections representing the desired signals and coherent noise represented by one linear event. Given the relatively low variability in P-wave velocities, a constant S-wave velocity has been used for the sake of simplicity. The central frequency of the adopted Ricker wavelet is 20 Hz. The synthetic dataset comprises a total of nine gathers, each containing 40 traces with a spatial sampling interval of 40 m. The temporal sampling interval is set as 4 ms, and the spatial

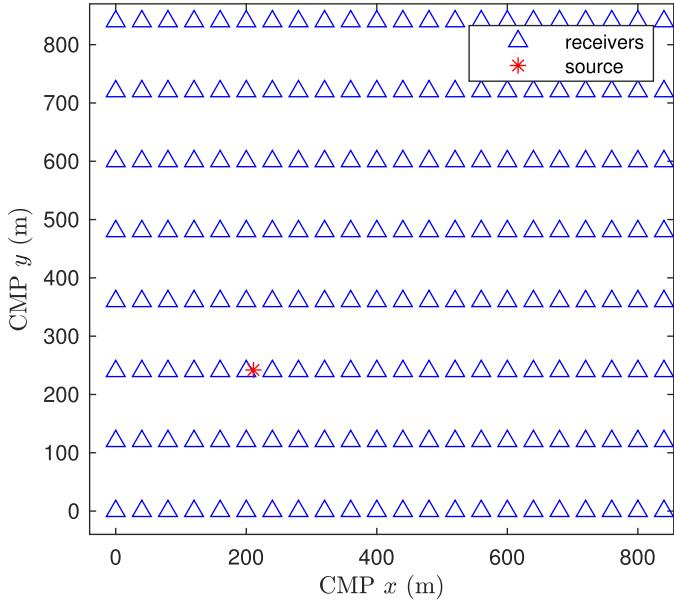


Fig. 4. Designed acquisition geometry employed for generating 3-D synthetic seismic data.

distance between receiver lines is 120 m. To provide a more realistic representation of seismic data, we also incorporate random noise into the synthetic dataset, aligning it with real-world conditions.

Compared to field seismic data, the synthetic example exhibits a relatively simpler structure, enabling us to assess the effectiveness of the LMO-based method alone. The velocity of ground roll and the zero offset travel time are known during data generation, allowing us to apply LMO correction specifically to ground roll and effectively flatten its trajectory. Fig. 6(b) illustrates the extracted ground roll, which remains consistent across all gathers, thereby verifying the efficacy of our method. By subtracting Fig. 6(b) from (a), we obtain the denoised results shown in Fig. 6(c). The subtraction results lead to a clear observation that the GCN effectively eliminates nearly all coherent noise while the significant reflection energy remains intact. Notably, even in the intersection region, our method yields impressive results.

To further emphasize the ground roll separation ability of our approach, we compare Fig. 6(b) with the ground truth depicted in Fig. 6(d). Visual inspection reveals minimal discernible differences between the two, supporting the conclusion that our method successfully removes unwanted ground roll without compromising the integrity of the desired reflections. Their differences are captured in Fig. 6(e), which displays the discrepancies between the modeled and extracted ground roll. To quantify the differences, we calculate the structural similarity index measure (SSIM) and obtain a high value of 0.9835, indicating a strong coherence between the original ground roll and the extracted one.

The second dataset aims to assess the performance of our network in extracting reflections after NMO, which is the underlying assumption for solving (8). Fig. 7(a) presents the original signals generated in this experiment, with parameters identical to the first dataset, except for incorporating an additional linear event to increase the extraction difficulty. The

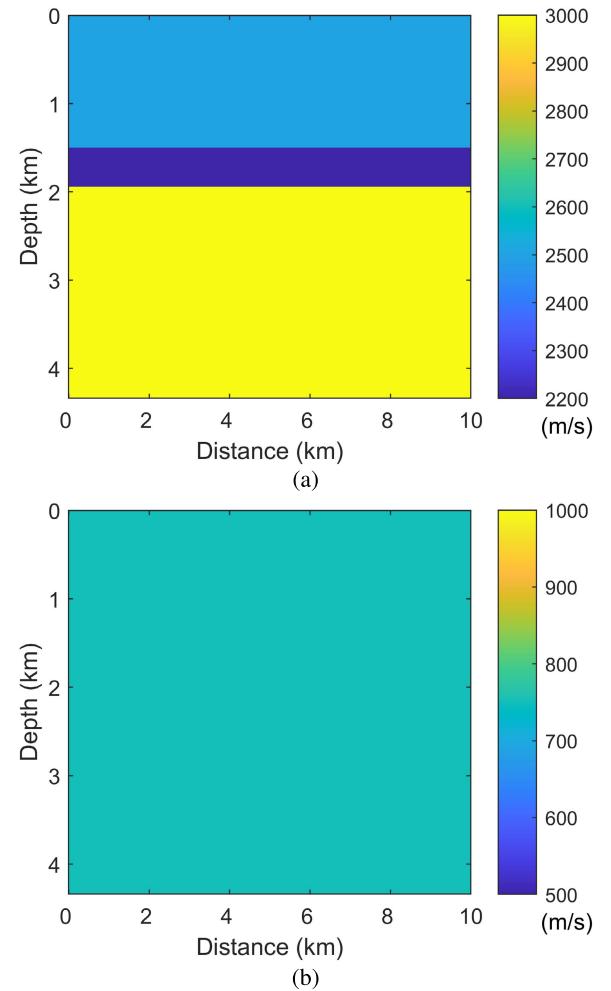


Fig. 5. Velocity models of the first synthetic data. (a) P-wave velocity model for reflections. (b) S-wave velocity model for ground roll.

extracted results by GCN are shown in Fig. 7(b), revealing a distinct and continuous separation of the reflections. Despite the relative simplicity of the synthetic data, the direct extraction of reflections offers an alternative approach to eliminating coherent noise. The difference profiles depicted in Fig. 7(c) further support this finding, where only negligible leakage can be observed. In summary, the results obtained from the synthetic data experiments demonstrate the potential of our approach for real-world applications.

B. Field Data Example

The field dataset used in this study to examine the effectiveness of the proposed method is a common-shot gather with 16 receiver lines, which was acquired in a desert environment of Western China. As depicted in Fig. 8(a), this dataset exhibits a high degree of complexity due to severe contamination from ground roll, scattered noise, industrial noise, and random noise. Highlighted by blue polygons, ground roll manifests as low-frequency, high-amplitude waveforms. These waveforms spread across all seismic gathers and obscure valuable reflection signals from strong coal seams. Since our exploration is conducted in desert loess, the highly heterogeneous subsurface conditions foster wave scattering, as indicated by green polygons. Unlike the ground roll, scattered noise is more

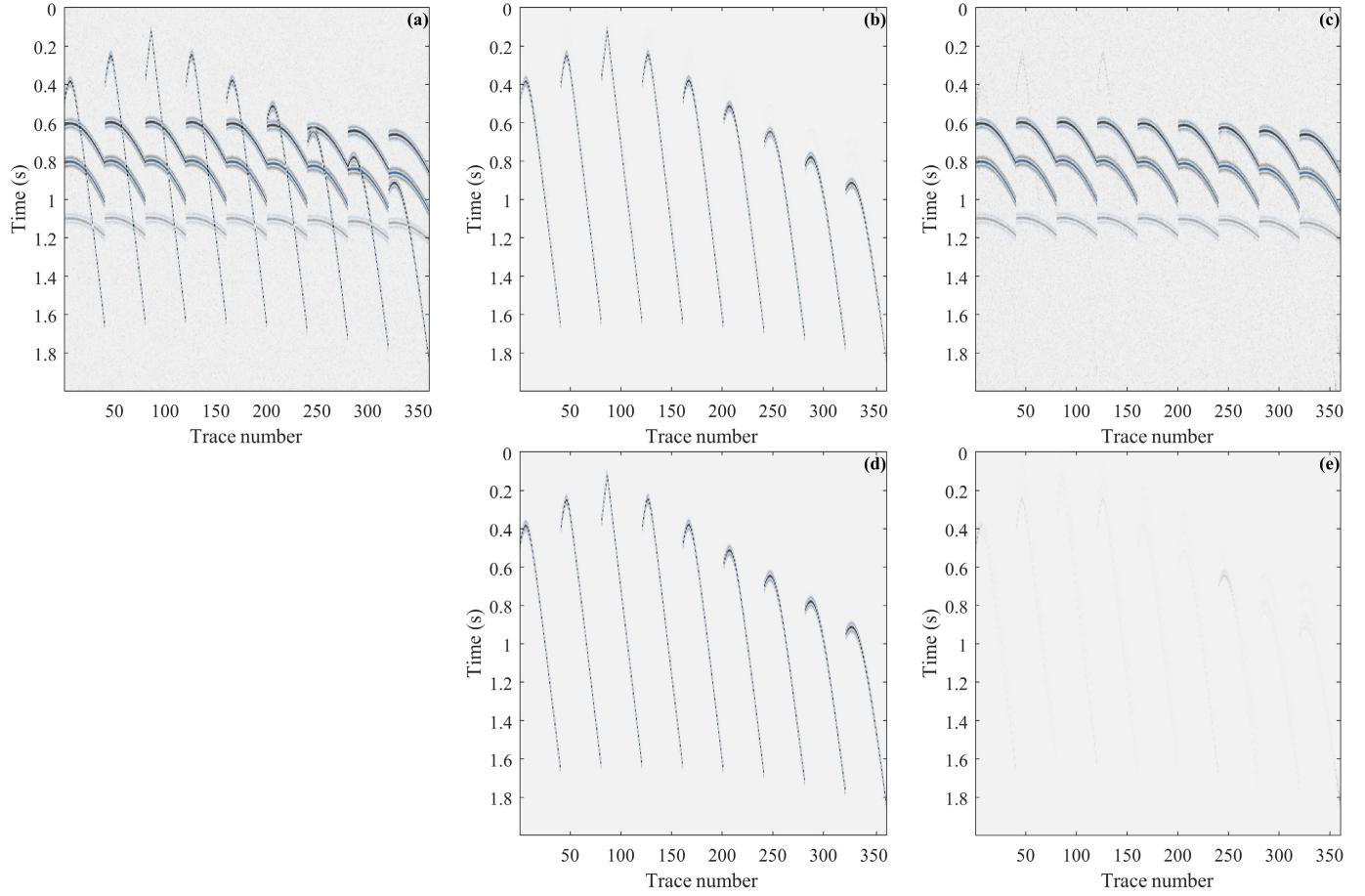


Fig. 6. Synthetic example demonstrating the application of our method for ground roll extraction. (a) Synthetic data composed of reflections (modeled by hyperbolic events) and ground roll (modeled by linear events). (b) Ground roll extracted by our method. (c) Subtraction of (b) from (a). (d) Modeled ground roll serving as the ground truth. (e) Deviations between the modeled and extracted ground roll, highlighting any disparities.

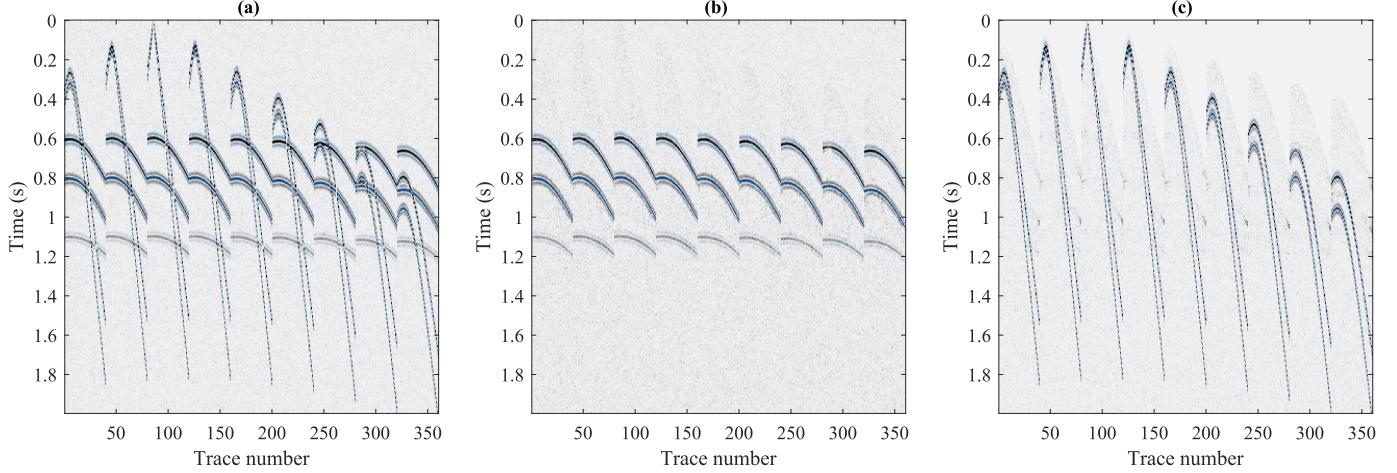


Fig. 7. Synthetic example for extracting reflections with our method. (a) Synthetic data comprising hyperbolic-modeled reflections and linear-modeled ground roll. (b) Reflections extracted by our method. (c) Subtraction of (b) from (a).

difficult to predict. Besides, the main energy of scattered noise is broadband and it can mimic the characteristics of true reflection events, which complicates the data processing. Additionally, our dataset includes a considerable amount of random noise, adding another layer of complexity. Given its unpredictability and broad frequency range, random noise originates from various sources and is notably challenging to manage. As a result, the desired reflections corresponding to the coal seam at approximately 1.5 s are difficult to discern.

To effectively attenuate the unwanted noise and enhance the visibility of the reflections, we employ the proposed two-step approach for data processing.

In the first step, we utilize LMO correction to align ground roll and separate it using the proposed network architecture. To account for the dispersion nature of ground roll, we repeat this procedure with three different velocities, effectively removing three groups of linear ground roll. We prioritize this step because the morphology of ground roll is usually

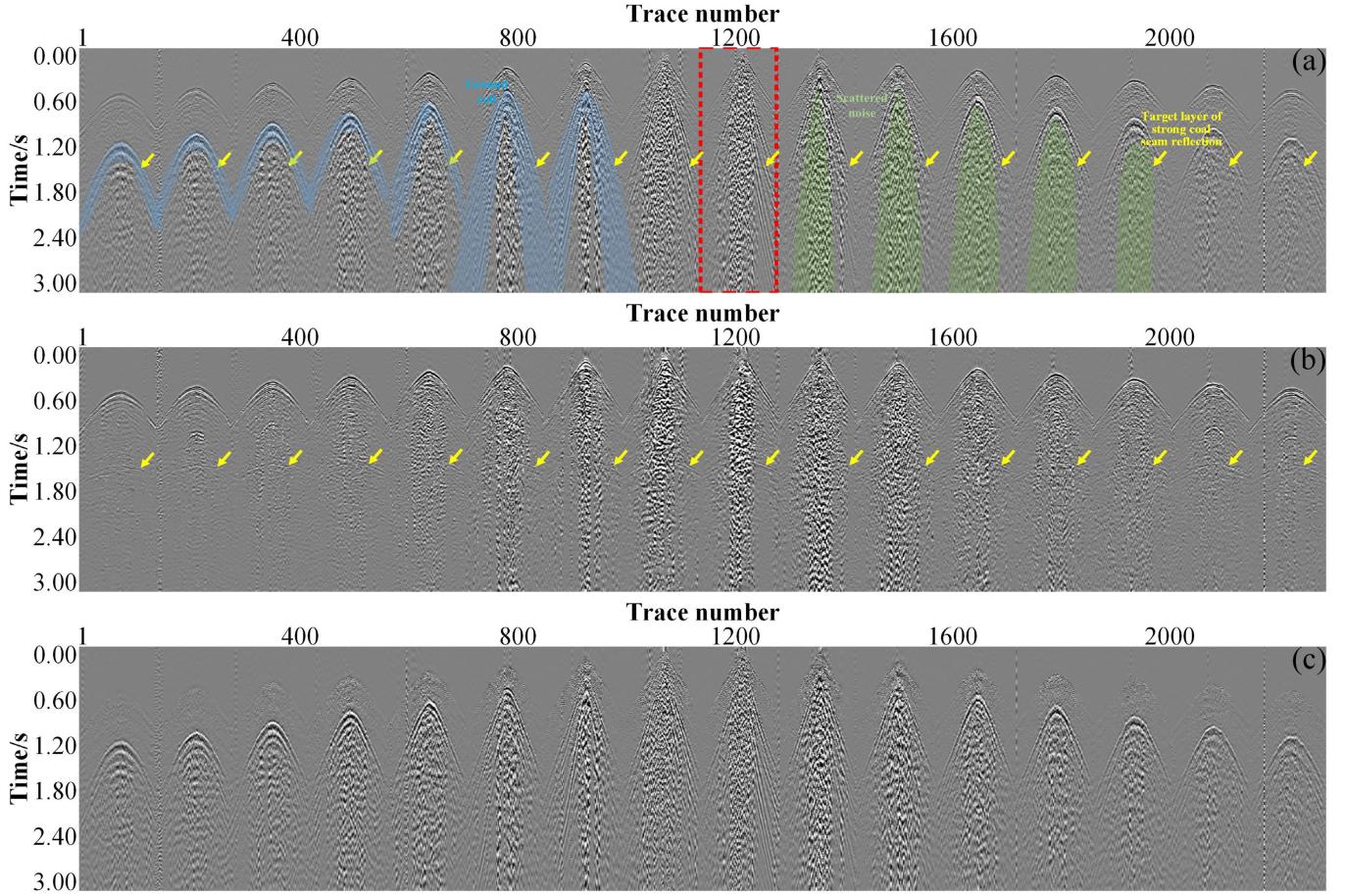


Fig. 8. Field data example. (a) Noisy data. Red-box regions are selected for enlarged evaluation in Fig. 10. (b) Separated reflections by our method. (c) Removed coherent noise by our method.

complete and consistent in raw data, making it more amenable to extraction. In contrast, extracting reflections directly from the raw data proves challenging due to the presence of strong amplitude coherent noise that obscures them, posing difficulties in velocity estimation and precise recognition. Therefore, the application of NMO correction to flatten the reflections is carried out in the second step. Partial coherent noise has been removed in the first step, simplifying the extraction of horizontal reflections. However, it should be noted that the dataset may still be contaminated by strong scattering before extracting the reflections, as observed in the land example used in this experiment. This strong amplitude noise still poses challenges to the accurate recognition of reflections. Hence, we remove partial coherent noise slightly before the second step, following the approach in [11] and [72]. This preprocessing step is conducted carefully to avoid damaging the reflections, resulting in more precise identification of reflections and facilitating the subsequent extraction by GCN.

The results of our method are plotted in Fig. 8(b), where ground roll and scattered noise are no longer observable, and the target layer becomes clearly visible. The appropriate selection of stopping iterations is crucial for the success of our method. If the number of iterations exceeds the correct number of primary events fully present in the extracted data, residual ground roll and scattered noise may also be extracted, leading

to inadequate denoising and distortion of the target reflections. Conversely, if the number of stopping iterations is too small, the reflections may not be fully recovered, resulting in leakage into the removed noise gathers. By carefully choosing the stopping iterations, our method effectively removes ground roll, scattering energy, and random noise without introducing spatial-amplitude smearing. The efficacy of our method can be further supported by examining the removed coherent noise in Fig. 8(c), where only minimal distortion of the primary reflections can be observed. Our method achieves accurate denoising while preserving the integrity of the target reflections.

Fig. 9 presents two sets of comparative results. The first set is obtained by the unsupervised method described in our previous abstract [54]. As shown in Fig. 9(a), this method proves to be effective as the energy of the scattered noise is significantly reduced. However, although the amplitudes of target reflections are well preserved, conflicting dipping noise trains are still visible, as indicated by the green arrows in Fig. 9(b). There are two main factors contributing to this sub-optimal outcome. First, the presence of strong coherent noise dominates, making accurate velocity estimation challenging. Without accurate velocity estimation, the reflections cannot be adequately flattened, resulting in reduced self-similarities. Second, these remaining dipping noise trains possess a similar

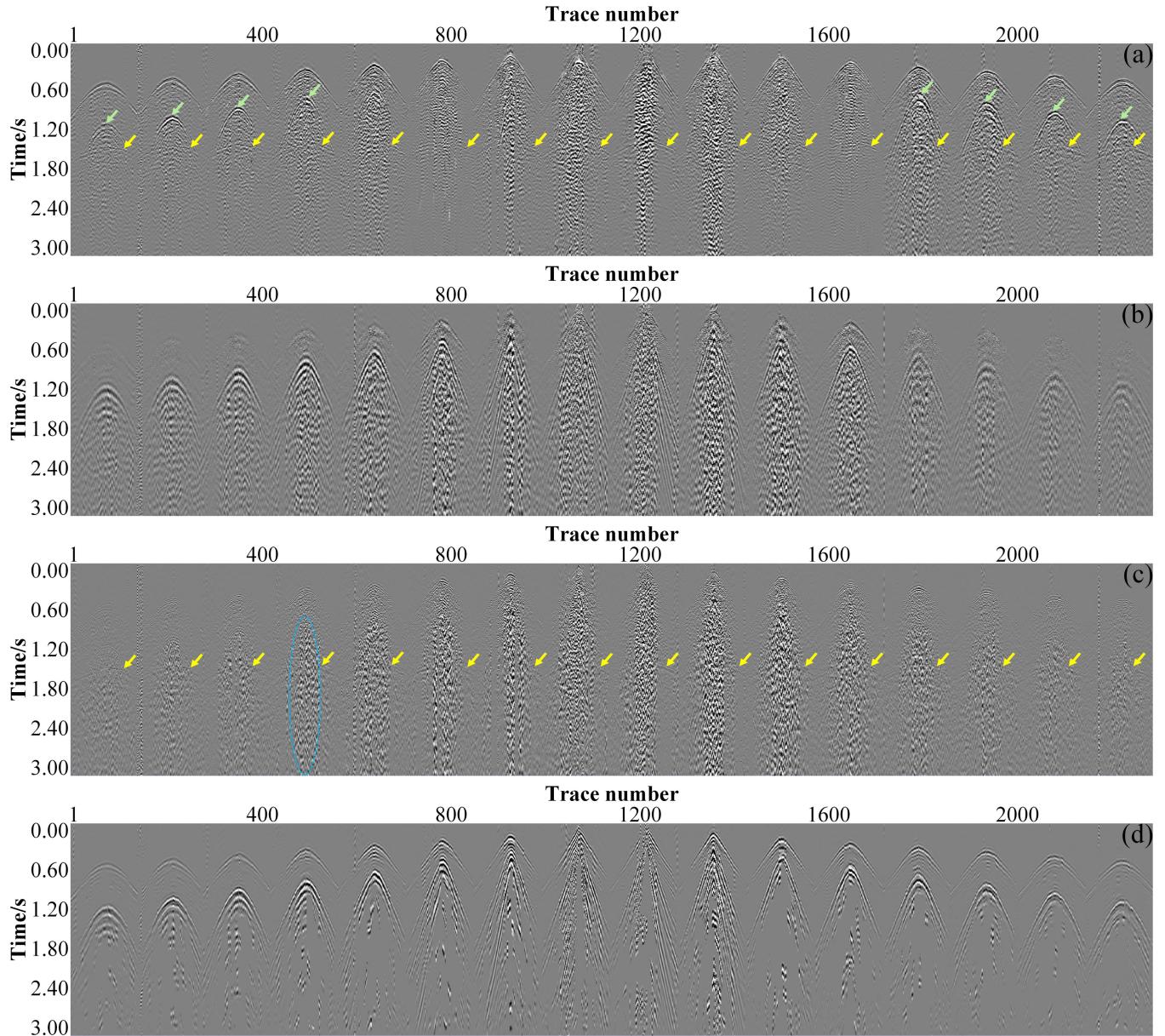


Fig. 9. Field data result comparisons. (a) Separated reflections by the NMO-based unsupervised method [54]. There are still ground roll residues near the apex regions of far-offset gathers, as indicated by green arrows. (b) Removed coherent noise by the NMO-based unsupervised method. (c) Separated reflections by the high-resolution Radon transform method. Scattered noise is not effectively removed, as denoted by the blue circle. (d) Removed coherent noise by the high-resolution Radon transform method.

velocity as reflections, especially near the apex regions in far-offset gathers, unexpectedly allowing them to be simultaneously extracted by the network. These findings underscore the necessity of incorporating LMO to handle linear coherent noise and enhance the flexibility of the unsupervised method.

Fig. 9(c) illustrates the denoised results obtained by the HRT method. Due to the severe masking effect of the dominant ground roll and scattered noise on the target reflective events, complete separation cannot be achieved in the Radon domain. As a result, the application of the high-resolution linear Radon filter attenuates a significant portion of the coherent noise but still leaves residual ground roll. Additionally, there is slight removal of the first arrivals to compromise the filtering region selection. Furthermore, significant scattered noise is

still noticeable, as indicated by the blue circle in Fig. 9(c). This is primarily because the high-resolution Radon filter cannot effectively handle scattered noise that lacks apparent linear features. In Fig. 9(d), a relatively successful ground roll removal is observed. However, scattered noise and random noise remain unresolved.

To further illustrate the effectiveness of our method, we plot the enlarged region indicated by the red box in Fig. 8. The denoised results in Fig. 10(b) show that most reflections become clean and continuous, particularly in the regions indicated by the yellow boxes. Additionally, as denoted by the red arrow in Fig. 10(b), a potential reflection event becomes more prominent after being processed by our method. The removed coherent noise in Fig. 10(c) further confirms that strong

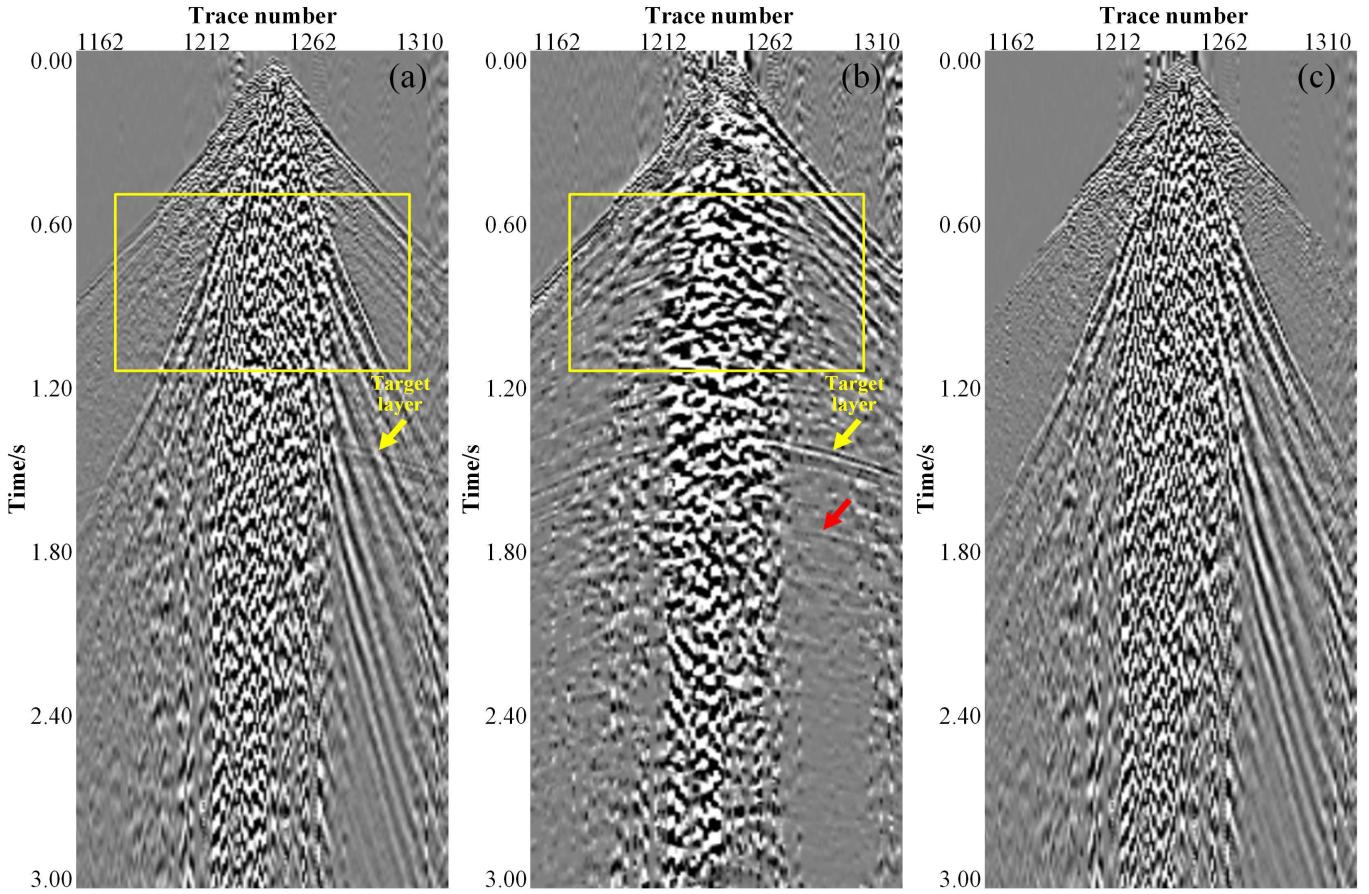


Fig. 10. Zoomed-in view of field seismic data results using our method. (a) Original data. (b) Separated reflections. (c) Separated coherent noise. By comparing the yellow boxes, most of the scattered energy and ground roll are attenuated. A red arrow signifies that this potential reflection event is becoming more prominent.

scattered energy has been attenuated without any detectable loss of reflection energy, thereby demonstrating the fidelity of our method.

It is important to note that NMO flattening in the shot gather is not entirely correct, as reflections in the shot domain can have asymmetric apexes concerning offset. However, this assumption holds valid for sedimentary environments with non-significant structural dips. We would like to mention that the hyperbolic moveout assumption in the shot domain is also utilized by industry-proven methods, such as those proposed by [73] and [74].

IV. CONCLUSION

We propose an unsupervised deep learning method for ground roll and scattered noise attenuation, free for scarce or expensive labels. By leveraging the intrinsic preference of neural networks for low-frequency data during training, we employ two generator networks to reconstruct ground roll and reflections sequentially by fitting raw seismic data, aided by early stopping. Additionally, to ensure successful extraction, LMO and NMO are employed as two auxiliary tools, respectively. Synthetic data experiments demonstrate the satisfactory multiscale self-similar feature extraction capability of the proposed network. This approach significantly attenuates coherent noise while preserving the integrity of the target reflections, outperforming the conventional Radon transform

in complex land field data due to its robust non-linear feature extraction ability. Future work will focus on minimizing human involvement in the workflow and extending its application to other geophysical data types and domains.

REFERENCES

- [1] M. I. Al-Husseini, J. B. Glover, and B. J. Barley, "Dispersion patterns of the ground roll in eastern Saudi Arabia," *Geophysics*, vol. 46, no. 2, pp. 121–137, Feb. 1981.
- [2] R. Saatçilar and N. Canitez, "A method of ground-roll elimination," *Geophysics*, vol. 53, no. 7, pp. 894–902, 1988.
- [3] G. Beresford-Smith and R. Rango, "Suppression of ground roll by windowing in two domains," *1st Break*, vol. 7, no. 2, pp. 55–63, Feb. 1989.
- [4] G. J. O. Vermeer, "Alternative strategies for tackling scattered noise," in *Proc. SEG Tech. Program Expanded Abstr.* Houston, TX, USA: Society of Exploration Geophysicists, 2008, pp. 95–99.
- [5] X. L. Han et al., "Coherent and scattering surface wave attenuation with model-based method: A case study from tight-gas oilfield of China," in *Proc. SPG/SEG Int. Geophys. Conf.* Beijing, China: Society of Exploration Geophysicists and Society of Petroleum Geophysicists, Apr. 2016, pp. 752–754.
- [6] M. Miorali, K. Mills, M. Igwe, and D. McClymont, "Near-surface data-driven methods for surface wave and multiple removal, onshore kenya," in *Proc. 80th EAGE Conf. Exhib.* Copenhagen, Denmark: European Association of Geoscientists & Engineers, 2018, pp. 1–5.
- [7] P. Embree, J. P. Burg, and M. M. Backus, "Wide-band velocity filtering—the pie-slice process," *Geophysics*, vol. 28, no. 6, pp. 948–974, 1963.
- [8] S. Treitel, J. L. Shanks, and C. W. Frasier, "Some aspects of fan filtering," *Geophysics*, vol. 32, no. 5, pp. 789–800, 1967.

- [9] P. Cheng, X. Li, and Z. Liu, "Adaptive ground-roll attenuation with localized eigenimage filtering," in *Proc. 1st Int. Meeting Appl. Geosci. Energy Expanded Abstr.* Denver, CO, USA: Society of Exploration Geophysicists, 2021, pp. 2964–2968.
- [10] A. J. Deighan and D. R. Watts, "Ground-roll suppression using the wavelet transform," *Geophysics*, vol. 62, no. 6, pp. 1896–1903, 1997.
- [11] X. Chen, W. Chen, X. Wang, and W. Wang, "Sparsity-optimized separation of body waves and ground-roll by constructing dictionaries using tunable Q-factor wavelet transforms with different Q-factors," *Geophys. J. Int.*, vol. 211, no. 1, pp. 621–636, Oct. 2017.
- [12] C. Yarham, U. Boeniger, and F. Herrmann, "Curvelet-based ground roll removal," in *Proc. SEG Annu. Meeting*, 2006, pp. 2777–2782.
- [13] M. Naghizadeh and M. Sacchi, "Ground-roll attenuation using curvelet downscaling," *Geophysics*, vol. 83, no. 3, pp. V185–V195, May 2018.
- [14] R. Abma and J. Claerbout, "Lateral prediction for noise attenuation by t-x and f-x techniques," *Geophysics*, vol. 60, no. 6, pp. 1887–1896, Nov. 1995.
- [15] R. Soubaras, "Signal-preserving random noise attenuation by the f-x projection," in *Proc. SEG Tech. Program Expanded Abstr.* Houston, TX, USA: Society of Exploration Geophysicists, 1994, pp. 1576–1579.
- [16] M. D. Sacchi and M. Naghizadeh, "Adaptive linear prediction filtering for random noise attenuation," in *Proc. SEG Tech. Program Expanded Abstr.* Houston, TX, USA: Society of Exploration Geophysicists, 2009, pp. 3347–3351.
- [17] Y. A. S. I. Oboué, Y. Chen, M. Bai, W. Chen, and Y. Chen, "Erratic and random noise attenuation using adaptive local orthogonalization," *Geophysics*, vol. 87, no. 4, pp. V381–V396, Jul. 2022.
- [18] S. Jiao, Y. Chen, M. Bai, W. Yang, E. Wang, and S. Gan, "Ground roll attenuation using non-stationary matching filtering," *J. Geophys. Eng.*, vol. 12, no. 6, pp. 922–933, Dec. 2015.
- [19] S. R. Trickett, "F-xy eigenimage noise suppression," *Geophysics*, vol. 68, no. 2, pp. 751–759, 2003.
- [20] M. Bekara and M. Van der Baan, "Local singular value decomposition for signal enhancement of seismic data," *Geophysics*, vol. 72, no. 2, pp. V59–V65, Mar. 2007.
- [21] D. Liu, M. D. Sacchi, and W. Chen, "Efficient tensor completion methods for 5-D seismic data reconstruction: Low-rank tensor train and tensor ring," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [22] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 2862–2869.
- [23] J. Li, D. Wang, S. Ji, Y. Li, and Z. Qian, "Seismic noise suppression using weighted nuclear norm minimization method," *J. Appl. Geophys.*, vol. 146, pp. 214–220, Nov. 2017.
- [24] X. Dong, Y. Li, and B. Yang, "Desert seismic data denoising and effective signal recovery by using improved shearlet transform based on the deep-learning coefficient selection," *J. Seismic Explor.*, vol. 30, no. 5, pp. 455–479, 2021.
- [25] B. Wang and J. Geng, "Efficient deblending in the PFK domain based on compressive sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 995–1003, Feb. 2020.
- [26] D. Liu, X. Li, W. Wang, X. Wang, Z. Shi, and W. Chen, "Eliminating harmonic noise in vibroseis data through sparsity-promoted waveform modeling," *Geophysics*, vol. 87, no. 3, pp. V183–V191, May 2022.
- [27] C. L. Liner, "Concepts of normal and dip moveout," *Geophysics*, vol. 64, no. 5, pp. 1637–1647, Sep. 1999.
- [28] M. J. Porsani, M. G. Silva, P. E. M. Melo, and B. Ursin, "SVD filtering applied to ground-roll attenuation," *J. Geophys. Eng.*, vol. 7, no. 3, pp. 284–289, Sep. 2010.
- [29] S. K. Chiu, "Coherent and random noise attenuation via multichannel singular spectrum analysis in the randomized domain: Coherent and Random Noise Attenuation in the Randomized Domain," *Geophys. Prospecting*, vol. 61, pp. 1–9, 2013.
- [30] V. Oropeza and M. Sacchi, "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis," *Geophysics*, vol. 76, no. 3, pp. V25–V32, May 2011.
- [31] S. K. Chiu and J. E. Howell, "Attenuation of coherent noise using localized-adaptive eigenimage filter," in *Proc. 78th Annu. Int. Meeting*, 2008, pp. 2541–2545.
- [32] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, F. Bach and D. Blei, Eds. Lille, France, vol. 37, 2015, pp. 448–456.
- [33] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [34] S. Yu, J. Ma, and W. Wang, "Deep learning for denoising," *Geophysics*, vol. 84, no. 6, pp. V333–V350, 2019.
- [35] D. Liu, W. Wang, X. Wang, C. Wang, J. Pei, and W. Chen, "Poststack seismic data denoising based on 3-D convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1598–1629, Mar. 2020.
- [36] H. Li, W. Yang, and X. Yong, "Deep learning for ground-roll noise attenuation," in *Proc. SEG Tech. Program Expanded Abstr.* Anaheim, CA, USA: Society of Exploration Geophysicists, 2018, pp. 1981–1985.
- [37] H. Kaur, S. Fomel, and N. Pham, "Seismic ground-roll noise attenuation using deep learning," *Geophys. Prospecting*, vol. 68, no. 7, pp. 2064–2077, Sep. 2020.
- [38] D. Liu, X. Wang, X. Yang, H. Mao, M. D. Sacchi, and W. Chen, "Accelerating seismic scattered noise attenuation in offset-vector tile domain: Application of deep learning," *Geophysics*, vol. 87, no. 5, pp. V505–V519, Sep. 2022.
- [39] R. H. Baardman and R. F. Hegge, "Machine learning approaches for use in deblending," *Lead. Edge*, vol. 39, no. 3, pp. 188–194, Mar. 2020.
- [40] B. Wang, J. Li, J. Luo, Y. Wang, and J. Geng, "Intelligent deblending of seismic data based on U-Net and transfer learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8885–8894, Oct. 2021.
- [41] H. Chen, M. D. Sacchi, H. Haghshenas Lari, J. Gao, and X. Jiang, "Nonstationary seismic reflectivity inversion based on prior-engaged semisupervised deep learning method," *Geophysics*, vol. 88, no. 1, pp. WA115–WA128, Jan. 2023.
- [42] D. Liu, W. Wang, X. Wang, Z. Shi, M. D. Sacchi, and W. Chen, "Improving sparse representation with deep learning: A workflow for separating strong background interference," *Geophysics*, pp. 1–86, 2022.
- [43] Z. Pan, H. Chen, Y. Wang, B. Huang, and W. Gui, "A new perspective on AE- and VAE-based process monitoring," *Techrxiv*, Apr. 2022, doi: [10.36227/techrxiv.19617534.v1](https://doi.org/10.36227/techrxiv.19617534.v1).
- [44] Z. Pan, Y. Wang, K. Wang, H. Chen, C. Yang, and W. Gui, "Imputation of missing values in time series using an adaptive-learned median-filled deep autoencoder," *IEEE Trans. Cybern.*, vol. 53, no. 2, pp. 695–706, Feb. 2023.
- [45] O. M. Saad, Y. A. S. I. Oboué, M. Bai, L. Samy, L. Yang, and Y. Chen, "Self-attention deep image prior network for unsupervised 3-D seismic data enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [46] W. Fang, L. Fu, H. Li, S. Liu, and Q. Wang, "BSnet: An unsupervised blind spot network for seismic data random noise attenuation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [47] H. Sun, F. Yang, and J. Ma, "Seismic random noise attenuation via self-supervised transfer learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [48] K. Wang, T. Hu, and S. Wang, "Iterative deblending using unsupervised learning with double-deep neural networks," *Geophysics*, vol. 88, no. 3, pp. V187–V205, May 2023.
- [49] K. Wang, T. Hu, and B. Zhao, "An unsupervised deep neural network approach based on ensemble learning to suppress seismic surface-related multiples," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [50] C. Birnie and M. Ravasi, "Explainable blind-spot networks for self-supervised seismic coherent noise suppression," in *Proc. 84th EAGE Annu. Conf. Exhib.*, 2023, pp. 1–5.
- [51] F. Qian et al., "Unsupervised seismic footprint removal with physical prior augmented deep autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023.
- [52] Z.-Q. John Xu, Y. Zhang, and Y. Xiao, "Training behavior of deep neural network in frequency domain," 2018, [arXiv:1807.01251](https://arxiv.org/abs/1807.01251).
- [53] N. Rahaman et al., "On the spectral bias of neural networks," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019, pp. 5301–5310.
- [54] D. Liu, W. Chen, M. D. Sacchi, and H. Wang, "Should we have labels for deep learning ground roll attenuation?" in *Proc. SEG Tech. Program Expanded Abstr.* Houston, TX, USA: Society of Exploration Geophysicists, 2020, pp. 3239–3243.
- [55] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning (still) requires rethinking generalization," *Commun. ACM*, vol. 64, no. 3, pp. 107–115, Mar. 2021.
- [56] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

- [57] Z.-Q.-J. Xu, "Frequency principle: Fourier analysis sheds light on deep neural networks," *Commun. Comput. Phys.*, vol. 28, no. 5, pp. 1746–1767, Jun. 2020.
- [58] T. Luo, Z. Ma, Z.-Q. John Xu, and Y. Zhang, "Theory of the frequency principle for general deep neural networks," 2019, *arXiv:1906.09235*.
- [59] A. M. Saxe et al., "On the information bottleneck theory of deep learning," *J. Stat. Mech., Theory Exp.*, vol. 2019, no. 12, Dec. 2019, Art. no. 124020.
- [60] B. Aubin, A. Maillard, J. Barbier, F. Krzakala, N. Macris, and L. Zdeborová, "The committee machine: Computational to statistical gaps in learning a two-layers neural network," in *Advances in Neural Information Processing Systems*, vol. 31. Red Hook, NY, USA: Curran Associates, 2018.
- [61] J. Sirignano and K. Spiliopoulos, "Mean field analysis of neural networks: A central limit theorem," *Stochastic Processes Their Appl.*, vol. 130, no. 3, pp. 1820–1852, Mar. 2020.
- [62] J. Lee et al., "Wide neural networks of any depth evolve as linear models under gradient descent," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 8570–8581.
- [63] Z. Cheng, M. Gadelha, S. Maji, and D. Sheldon, "A Bayesian perspective on the deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 5438–5446.
- [64] L. Zdeborová, "Understanding deep learning is also a job for physicists," *Nature Phys.*, vol. 16, no. 6, pp. 602–604, Jun. 2020.
- [65] O. M. Saad, Y. A. S. I. Oboue, M. Bai, L. Samy, L. Yang, and Y. Chen, "Self-attention deep image prior network for unsupervised 3-D seismic data enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
- [66] C. Qiu, B. Wu, N. Liu, X. Zhu, and H. Ren, "Deep learning prior model for unsupervised seismic data random noise attenuation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [67] D. Liu, Z. Deng, C. Wang, X. Wang, and W. Chen, "An unsupervised deep learning method for denoising prestack random noise," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [68] F. Oghenekohwo and M. D. Sacchi, "Transform-domain noise synthesis and normal moveout-stack deconvolution approach to ground roll attenuation," *Geophysics*, vol. 86, no. 1, pp. V15–V22, Jan. 2021.
- [69] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Munich, Germany: Springer, Oct. 2015, pp. 234–241.
- [70] J. Gao, L. Weng, M. Xia, and H. Lin, "MLNet: Multichannel feature fusion lozenge network for land segmentation," *J. Appl. Remote Sens.*, vol. 16, no. 01, Mar. 2022, Art. no. 016513.
- [71] D. Arpit et al., "A closer look at memorization in deep networks," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 233–242.
- [72] J. Chen, J. Ning, W. Chen, X. Wang, W. Wang, and G. Zhang, "Distributed acoustic sensing coupling noise removal based on sparse optimization," *Interpretation*, vol. 7, no. 2, pp. T373–T382, May 2019.
- [73] C. Perkins and M. Zwaan, "Ground roll attenuation," in *Proc. 62nd EAGE Conf. Exhib.*, 2000, Paper cp-28–00 172.
- [74] D. Le Meur, N. Benjamin, R. Cole, and M. Al Harthy, "Adaptive groundroll filtering," in *Proc. 70th EAGE Conf. Exhib. Incorporating SPE EUROPEC*. Rome, Italy: European Association of Geoscientists & Engineers, 2008.



Dawei Liu received the bachelor's degree in communication engineering from Chang'an University, Xi'an, China, in 2013, the master's degree in electronic and communication engineering from Xi'an Jiaotong University (XJTU), Xi'an, in 2018, and the Ph.D. degree from the School of Information and Communication Engineering, Xi'an, in 2022.

He is currently a Post-Doctoral Scholar at Purdue University, Lafayette, IN, USA. His research interests include tensor decomposition, deep learning, and time-frequency analysis and their applications in the seismic data processing.

Dr. Liu is a reviewer for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Geophysics*, *Acta Geophysica*, and *PetroleumScience*.



Mauricio D. Sacchi (Member, IEEE) received the Diploma degree in geophysics from the National University of La Plata, La Plata, Argentina, in 1988, and the Ph.D. degree in geophysics from the University of British Columbia (UBC), Vancouver, BC, Canada, in 1996.

He joined the Department of Physics, University of Alberta, Edmonton, AB, Canada, in 1997. He directs the Signal Analysis and Imaging Group, Edmonton, an initiative for advanced research for geophysical signal processing and imaging. He has developed and taught short courses for industry and professional associations. His research interests include geophysical signal analysis and seismic imaging methods.

Dr. Sacchi was a recipient of the 2012 Medal of the Canadian Society of Exploration Geophysicists (CSEG), the 2014 Central and South America Honorary Lecturer for the Society of Exploration Geophysicists, and the 2016 CSEG Distinguished Lecturer. Recently, he became a recipient of the 2019 Virgil Kauffman Gold Medal. Also, he was the Editor-in-Chief of the journal *Geophysics*, from 2016 to 2018.



Xiaokai Wang (Member, IEEE) received the bachelor's degree in information engineering and the Ph.D. degree in communication engineering from Xi'an Jiaotong University (XJTU), Xi'an, China, in 2006 and 2012, respectively.

From 2012 to 2014, he was a Post-Doctoral Fellow with the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China. He joined the Department of Computational Geophysics, XJTU, in 2015, where he was transferred to the School of Information and Communication Engineering, in 2018. He was a Visitor of the Bureau of Economic Geology at The University of Texas at Austin, Austin, TX, USA, from 2016 to 2017. His research interests include seismic attributes extraction, and time-frequency analysis and their applications in seismic data processing.



Wenchao Chen received the B.S. and M.S. degrees in seismic exploration and information technology from Chang'an University, Xi'an, China, in 1993 and 1996, respectively, and the Ph.D. degree in electromagnetic field and microwave technology from Xi'an Jiaotong University, Xi'an, in 2000.

From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Computation Science, Northwestern Polytechnical University, Xi'an. Since 2002, he has been with the School of Information and Communications Engineering, Xi'an Jiaotong University, where he is currently a Professor. His research interests include seismic and ground-penetrating radar (GPR) signal processing, blind signal processing, and sparse representation.

Dr. Chen has served as a member of the Society of Exploration Geophysicists (SEG) and the Chinese Geophysical Society. He is also a Referee for several journals, including the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Geophysics*, *Interpretations*, and *Journal of Applied Geophysics*.