Check for updates

Geophysical Prospecting

Geophysical Prospecting

ORIGINAL ARTICLE

Unsupervised Seismic Data Denoising via SIREN-Guided Diffusion Mode

Ji Li[1] | Dawei Liu[2] | Daniel Trad[1]

[1]Department of Earth, Energy and Environment, University of Calgary, Calgary, Alberta, Canada | [2]the School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China

Correspondence: Dawei Liu (409791715@qq.com)

Received: 26 August 2025 | **Revised:** 12 March 2026 | **Accepted:** 26 March 2026

Keywords: diffusion | implicit neural representations | denoising | signal processing | unsupervised

ABSTRACT

Denoising diffusion probabilistic models have emerged as powerful generative frameworks capable of synthesizing high quality images through iterative denoising of Gaussian noise. However, their unconditional nature limits their effectiveness in inverse problems such as denoising, where reconstruction from a specific observation is required. Recent extensions like denoising diffusion restoration models introduce conditional generation by incorporating known degradation operators and pre-trained denoisers, but these approaches remain heavily reliant on supervised training and large-scale external datasets. This work proposes a fully unsupervised alternative called the sinusoidal representation network (SIREN)-guided diffusion denoising model, which integrates implicit neural representations into the diffusion process. We replace the conventional pre-trained denoiser with a coordinate-based network (SIREN) optimized directly at inference time. Instead of assuming an explicit degradation model, we perform partial noise injection to the observed signal and initialize the reverse diffusion from this intermediate state. At each step, the SIREN is trained to match the current sample. This integration brings out the complementary strengths of both components: the diffusion process provides a structured, probabilistic coarse-to-fine regularization mechanism that stabilizes the optimization trajectory. At the same time, SIREN offers a powerful inductive bias for capturing structured, continuous signals. Crucially, this significantly eliminates the need for early stopping, a known challenge in SIREN-based denoising, and makes the method highly robust to noise without supervision or clean training data. We evaluate our approach on synthetic and real seismic datasets and demonstrate that it achieves better denoising performance compared to traditional and deep learning-based baselines.

1 | Introduction

Seismic data acquisition is inherently contaminated by various types of noise, which degrade the quality of recorded wavefields and obscure subsurface features (Abma and Claerbout 1995). Effective denoising is a critical preprocessing step for seismic interpretation, migration, inversion and other downstream geophysical tasks (Yilmaz 2001). In practice, the recorded signals contain reflections and random noise, coherent noise (e.g., ground roll), ambient environmental disturbances and acquisition-related artefacts (Deighan and Watts 1997; Beasley 2008; Daley et al. 2013). These unwanted components often overlap with the useful signal in time and frequency domains, making separation a nontrivial problem. As seismic surveys become increasingly dense and high resolution, the complexity and volume of noisy data pose significant challenges for traditional signal processing pipelines. Denoising aims to enhance the signal-to-noise ratio (SNR) by suppressing noise while

Abbreviations: INRs, implicit neural representations; DDPMs, Denoising Diffusion Probabilistic Models.

© 2026 European Association of Geoscientists & Engineers.

preserving geological structures such as fault planes, reflectors and stratigraphic features. However, the design of effective denoising methods is complicated by the diverse nature of seismic noise and the variability of subsurface environments across datasets. Therefore, robust and adaptive denoising approaches remain a topic of ongoing research in industry and academia.

Traditional seismic denoising methods are typically based on physical assumptions about the signal and noise characteristics and exploit transform-domain sparsity, predictability or coherence to achieve separation. One widely used category involves sparse representations in transform domains such as the Radon transform (Ibrahim and Sacchi 2014; J. Li and Sacchi 2021), the Fourier transform (Vera Rodriguez et al. 2012; Naghizadeh 2012) and the wavelet transform (Shan et al. 2009; Gaci 2013). These methods assume seismic reflections are compact or sparse in the transform domain, whereas noise is spread out or incoherent. Filtering techniques in the frequency–wavenumber (f-k) domain (Beresford-Smith and Rango 1989) and curvelet domain (Yarham et al. 2006; Herrmann et al. 2008) have also proven effective for removing coherent noise such as ground roll and surface waves. Additionally, prediction-based methods have been applied to suppress predictable noise patterns like Wiener filtering (Robinson and Treitel 1967) and autoregressive modelling (Naghizadeh and Sacchi 2007). More recent approaches include techniques based on singular spectrum analysis (Oropeza and Sacchi 2011) and rank-reduction methods (Y. Chen et al. 2016), which treat the seismic data matrix as a low-rank signal contaminated by random noise. While these traditional techniques are computationally efficient and often interpretable, their performance heavily depends on prior noise or data structure assumptions. They may struggle with complex noise distributions or structural variations in the signal, especially in cases where signal and noise overlap in the transform domain.

Deep learning has emerged as a powerful paradigm for seismic data denoising in recent years, driven by its capacity to learn complex nonlinear mappings directly from data (Yu et al. 2019). Supervised learning approaches train neural networks on paired noisy and clean data, allowing the model to learn an explicit denoising function (Richardson and Feller 2019; J. Li et al. 2021). Convolutional neural networks (CNNs) (Yuqing et al. 2019; Zhang, Lin, et al. 2019), residual networks (ResNets) (Yu et al. 2019; Yang et al. 2021), U-Nets (Mandelli et al. 2019; Sun et al. 2020) and generative adversarial networks (W. Li and Wang 2021) have all been successfully applied to seismic denoising, often achieving superior performance over traditional methods in structured or moderately noisy environments. However, these methods typically require large quantities of clean training labels, which are rarely available in real seismic acquisition settings. To address this limitation, unsupervised and self-supervised learning techniques have been developed. Methods like deep image prior (DIP) (Ulyanov et al. 2018) eliminate the need for labelled data by leveraging the implicit regularization of randomly initialized CNNs, which tend to model structured signal components before fitting noise. Noise2Noise (Lehtinen et al. 2018) and Noise2Void (Krull et al. 2019) further relax supervision by learning from noisy data alone, using masking or expectation-based loss strategies to infer denoising mappings. These methods have been successfully applied to remove seismic noise (Zhang, Liu, et al. 2019; Liu et al. 2020; G. Chen et al. 2022; J. Li et al. 2024). Other approaches integrate regularization techniques, such as total variation, low-rank priors or frequency-domain constraints, into network training to improve robustness in real-world settings (Qiu et al. 2021; Gao et al. 2022; J. Li et al. 2025). Despite their advantages, unsupervised deep learning methods can be sensitive to hyperparameters and may overfit noise in severely corrupted data (Liu et al. 2023). Moreover, their performance is often limited by the absence of a strong prior that models the statistical structure of clean seismic signals. These limitations motivate the exploration of alternative frameworks, such as diffusion-based models, which introduce principled probabilistic mechanisms for progressive refinement and naturally support uncertainty modelling (Hui et al. 2024).

Denoising diffusion probabilistic models (DDPMs) (Ho et al. 2020) have recently gained significant attention as generative models capable of producing high-quality samples from complex data distributions (Sasaki et al. 2021; Wyatt et al. 2022). These models operate by learning to reverse a fixed Markovian forward process that progressively adds Gaussian noise to clean data. The reverse process is modelled as a parameterized denoising network trained to reconstruct the data distribution from noisy inputs iteratively. Unlike conventional deep learning models, diffusion models provide strong theoretical guarantees, stable training, and high-fidelity diverse outputs (Kingma et al. 2021; Yang et al. 2023). In addition to image synthesis, diffusion models have been successfully adapted to a variety of inverse problems, including image super-resolution (Wu et al. 2023), inpainting (Lugmayr et al. 2022), deblurring (Whang et al. 2022) and medical imaging reconstruction (Khader et al. 2023). In these tasks, conditioning mechanisms are introduced to guide the reverse diffusion process towards solutions consistent with the observed measurements. Notably, denoising diffusion restoration models (DDRM) (Kawar et al. 2022) and denoising diffusion null-space models (DDNM) (Wang et al. 2022) utilize pre-trained diffusion priors and incorporate degradation operators to enable posterior sampling from constrained distributions. An alternative yet closely related family of generative methods is the score-based diffusion model (Song et al. 2021; Chung and Ye 2022), which estimates the gradient of the data distribution (i.e., the score function) using denoising score matching. These models interpret the forward-noising process as a stochastic differential equation and use a trained score network to perform reverse-time sampling via Langevin dynamics or probability flow ordinary differential equations (ODEs). Score-based models have shown competitive performance in image generation and restoration tasks and have also been adapted for inverse problems through various conditioning strategies, such as projected Langevin updates and consistency constraints (Batzolis et al. 2021; Tashiro et al. 2021). These approaches demonstrate that diffusion models can act as robust priors for solving ill-posed inverse problems in a data-driven manner. Diffusion models have also begun to be applied in seismic data processing. Recent studies have used diffusion-based methods for seismic interpolation, denoising, and deblending (Durall et al. 2023; Deng et al. 2024; Xiao et al. 2024), demonstrating their potential to handle complex noise and structural patterns. While promising, existing diffusion-based restoration methods often rely on large-scale training data and pre-trained denoisers, which may not generalize well to unseen noise types or domain-specific data such as seismic records. Moreover, these models are typically supervised and require the careful design

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of conditioning mechanisms to maintain consistency with the observations. This motivates the development of diffusion-based frameworks that can operate without external training data or explicit knowledge of the degradation model.

We propose a novel framework called the sinusoidal representation network (SIREN)-guided diffusion denoising model to address the limitations of both supervised diffusion models and traditional unsupervised approaches. This method combines the probabilistic coarse-to-fine refinement mechanism of diffusion processes with the inductive bias of implicit neural representations (INRs), specifically the SIREN. Unlike conventional DDPM-based approaches that require a pre-trained denoising model and large amounts of clean training data, our method is entirely unsupervised and operates directly on the noisy input. The key idea is to embed an untrained SIREN into the reverse diffusion. Instead of starting from pure Gaussian noise, the reverse sampling is initialized from a partially noised version of the observation, allowing signal structures to be preserved in early steps. At each diffusion timestep, the SIREN is optimized to fit the current intermediate sample. This progressive fitting strategy stabilizes the network's optimization and significantly removes the need for early stopping, which is typically required in standalone SIREN or DIP-based denoising methods. The proposed method offers several advantages. First, it is fully unsupervised and requires no external training data or pre-trained models, making it well-suited for field seismic applications where ground truth is unavailable. Second, integrating SIREN with diffusion sampling enhances robustness to heavy noise and improves signal recovery. Third, the flexible framework can be extended to various seismic inverse problems beyond denoising. We validate the proposed approach through experiments on synthetic and real seismic datasets and demonstrate that it consistently outperforms traditional and deep learning baselines regarding denoising quality and structure preservation.

2 | Theory

This section introduces the theoretical foundations behind our proposed method. We begin with an overview of DDPMs, followed by a description of INRs, particularly SIREN. Finally, we present our method, which combines diffusion sampling with untrained coordinate-based networks for unsupervised denoising.

2.1 | Diffusion Models for Image Generation and Restoration

Diffusion models are generative frameworks that learn to map simple noise distributions (e.g., Gaussian noise) to complex data distributions through a gradual denoising process. The DDPM (Ho et al. 2020) has become a widely used approach for image generation and restoration. In DDPM (Ho et al. 2020), a forward diffusion process is defined as a Markov chain, where each state $\mathbf{x}_t$ depends only on the previous state $\mathbf{x}_{t-1}$ and not on any earlier steps. This process gradually corrupts a clean image $\mathbf{x}_0$ into pure Gaussian noise over T time steps:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\,\mathbf{x}_{t-1}, \beta_t\mathbf{I}), \tag{1}$$

where $\beta_t \in (0,1)$ controls the variance of added Gaussian noise at step t. Recursively applying this rule yields the marginal distribution:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}), \tag{2}$$

with $\alpha_t = 1-\beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t}\alpha_i$. As t increases, the image $\mathbf{x}_t$ becomes increasingly noisy until it resembles pure Gaussian noise at $t = T$.

The generative process learns the reverse transitions $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$ to reconstruct the original image:

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t)). \tag{3}$$

In most implementations, a U-Net (Ronneberger et al. 2015) is used to model the mean μ_θ by predicting the added noise $\boldsymbol{\epsilon}$ from the noisy input $\mathbf{x}_t$ and the timestep t. The model is trained with the objective:

$$\mathcal{L}_{\text{DDPM}} = \mathbb{E}_{\mathbf{x}_0,\boldsymbol{\epsilon},t}\left[\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\|_2^2\right], \tag{4}$$

where

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}). \tag{5}$$

While DDPM is highly effective at generating samples that resemble the training distribution, they are inherently unconditional. That is, the model learns to generate samples from the marginal distribution $p(\mathbf{x}_0)$ without incorporating any information about a specific corrupted observation. As a result, when provided with a noisy image as input, a trained DDPM cannot recover the original clean image that produced the observation. Instead, it will generate a plausible but unrelated clean sample from the learned data distribution. This makes vanilla DDPM unsuitable for image restoration tasks such as denoising or inpainting, where the goal is to recover a specific ground-truth signal from a degraded input. To adapt diffusion models for inverse problems, it is necessary to introduce conditioning mechanisms that guide the generative process using the observed data.

In image restoration problems, the goal is to recover a clean image $\mathbf{x}_0$ from a degraded observation $\mathbf{y}$, typically modelled as

$$\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \sigma_y^2\mathbf{I}), \tag{6}$$

where $\mathbf{H}$ is a known degradation operator (e.g., masking, blurring or subsampling) and $\mathbf{z}$ is additive noise. To solve this inverse problem, recent methods such as DDRM (Kawar et al. 2022) and DDNM (Wang et al. 2022) adapt the DDPM framework by conditioning the reverse process on the measurement $\mathbf{y}$ using the known degradation model. These methods rely on a pre-trained denoising network (often a U-Net) and incorporate prior knowledge via posterior sampling techniques. The conditioning is typically performed in the reverse steps using tools such as singular value decomposition of $\mathbf{H}$ to align the denoised samples with the known measurements.

Despite their success, traditional diffusion-based restoration methods often require large datasets for training the

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

noise-predictor network, and their performance can degrade when applied to data outside the training distribution. This motivates the development of methods that avoid supervised training, using either self-supervised learning or untrained networks optimized directly at inference time.

2.2 | Implicit Neural Representations

INRs model structured signals such as images or volumes by fitting a neural network directly to the underlying signal, using coordinate locations as inputs. Instead of storing the signal on a discrete grid, an INR learns a continuous function $f_\theta(\mathbf{c})$, where $\mathbf{c} \in \mathbb{R}^d$ represents a spatial or spatiotemporal coordinate. The output approximates the signal value at that coordinate. This formulation allows the network to act as a continuous decoder that maps coordinate positions to signal intensities. The parameters θ are learned by minimizing a reconstruction loss between the network prediction and observed data. INRs are particularly attractive for inverse problems as they can be optimized directly using the corrupted input without requiring supervision from clean ground-truth data. This makes them a natural choice for unsupervised signal recovery.

SIREN (Sitzmann et al. 2020) is a specific INR architecture that uses sinusoidal activation functions to capture high-frequency components in the target signal better. Unlike conventional multilayer perceptrons (MLPs) with ReLU or tanh activations, SIREN can more effectively represent fine textures and smooth oscillations. The key difference lies in the first layer of the network, which applies a scaled sine activation:

$$\mathbf{z}^{(1)} = \sin\left(\omega_0(\mathbf{W}^{(0)}\mathbf{c} + \mathbf{b}^{(0)})\right). \tag{7}$$

Here, ω_0 is a frequency-scaling parameter that controls the network's sensitivity to high-frequency content and $\mathbf{W}^{(0)}, \mathbf{b}^{(0)}$ are learnable parameters. Subsequent layers also use sine activations but without high-frequency scaling. This architecture allows SIREN to overcome the spectral bias of ReLU networks, which tend to prioritize learning low-frequency components.

The effectiveness of SIREN in denoising stems from its structural bias towards smooth, coherent patterns. When fitting to noisy observations, the model typically learns dominant structures before overfitting to random noise. If optimization is stopped early or regularized, the network captures only the meaningful components of the signal. This makes SIREN a natural unsupervised prior for denoising tasks.

2.3 | SIREN-Guided Diffusion for Unsupervised Denoising

Existing diffusion-based restoration methods, such as DDRM (Kawar et al. 2022) and DDNM (Wang et al. 2022), address general inverse problems by incorporating a degradation operator $\mathbf{H}$ into the reverse diffusion process. These methods rely on knowledge of the forward corruption model and use posterior sampling techniques to ensure measurement consistency, often requiring access to a pre-trained denoiser and forward operator.

In our setting, we consider the pure denoising problem, where the corruption model reduces to additive Gaussian noise:

$$\mathbf{y} = \mathbf{x}_0 + \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \sigma^2\mathbf{I}). \tag{8}$$

This corresponds to the special case $\mathbf{H} = \mathbf{I}$, eliminating the need for explicit degradation modelling or posterior sampling. Instead, we impose structural priors by using an implicit neural representation to reconstruct the clean signal $\mathbf{x}_0$. Specifically, we represent $\mathbf{x}_0$ as the output of a coordinate-based neural network f_θ, implemented through a SIREN network, which maps spatial (or spatiotemporal) coordinates $\mathbf{c} \in \mathbb{R}^d$ to signal values:

$$\hat{\mathbf{x}}_0 = f_\theta(\mathbf{c}). \tag{9}$$

The network is randomly initialized and optimized per instance without any supervision or pretraining.

While untrained networks like SIREN exhibit strong architectural biases towards natural signal structures, they can easily overfit noise when trained directly on heavily degraded data. This challenge is particularly acute under high noise levels. To mitigate this, we embed the network within a diffusion-based denoising framework. The diffusion process provides a probabilistic, coarse-to-fine refinement and regularization mechanism that guides the SIREN to progressively refine the signal and avoid overfitting. We adopt the standard DDPM forward process:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0, (1-\bar{\alpha}_t)\,\mathbf{I}), \tag{10}$$

where $\bar{\alpha}_t = \prod_{i=1}^{t}(1-\beta_i)$ is derived from a predefined noise schedule $\{\beta_t\}_{t=1}^{T}$.

Unlike standard DDPM sampling, which initializes from pure noise, we begin the reverse diffusion process from a partially noised version of the observed data:

$$\mathbf{x}_{T'} \sim q(\mathbf{x}_{T'}|\mathbf{y}), \quad T' = T/2. \tag{11}$$

This partial initialization is essential because unsupervised networks such as SIREN lack a learned prior and cannot reconstruct signal content from fully random inputs. Starting from $\mathbf{x}_{T'}$ preserves sufficient signal structure to allow stable optimization and meaningful reconstructions.

From $t = T' - 1$ down to $t = 0$, we iteratively refine the estimate using the DDPM reverse update rule:

$$\mathbf{x}_{t-1} = \mathbf{x}_t + \frac{1-\bar{\alpha}_{t-1}}{\sqrt{1-\bar{\alpha}_t}}(\hat{\mathbf{x}}_0 - \mathbf{x}_t) + \sigma_t\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}), \tag{12}$$

where $\hat{\mathbf{x}}_0 = f_\theta(\mathbf{c})$ is the SIREN prediction. At each reverse step, the network parameters θ are updated by minimizing a simple regression loss:

$$\mathcal{L}_t(\theta) = \|f_\theta(\mathbf{c}) - \mathbf{x}_t\|_2^2. \tag{13}$$

We perform only a few gradient steps per iteration (typically 1–5), reusing parameters from the previous step to warm-start the optimization.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ALGORITHM 1 | Reverse Diffusion with SIREN-based Unsupervised Denoising

Input: Noisy data $\mathbf{y}$, coordinates $\mathbf{c}$, diffusion steps T, parameters $\beta_{1:T}, \sigma_{1:T}$

Output: Reconstructed clean signal $\hat{\mathbf{x}}_0$ Initialize SIREN parameters θ

Generate $\mathbf{x}_{t_0} \sim q(\mathbf{x}_{t_0} \mid \mathbf{y})$, where $t_0 = T/2$**for** $t = t_0 - 1, t_0 - 2, \ldots, 1$ **do**

Update θ by minimizing: $\mathcal{L}_t(\theta) = \|f_\theta(\mathbf{c}) - \mathbf{x}_t\|_2^2$

$\hat{\mathbf{x}}_0 = f_\theta(\mathbf{c})$

Sample noise: $\epsilon_t \sim \mathcal{N}(0, \mathbf{I})$

Compute: $\mathbf{x}_{t-1} = \mathbf{x}_t + \frac{1-\bar{\alpha}_{t-1}}{\sqrt{1-\bar{\alpha}_t}}(\hat{\mathbf{x}}_0 - \mathbf{x}_t) + \sigma_t \epsilon_t$end for

Return: Final reconstruction $\hat{\mathbf{x}}_0 = f_\theta(\mathbf{c})$

Although our approach does not involve explicit posterior sampling or projection onto a measurement-consistent subspace, as commonly used in DDRM-style diffusion restoration methods, it still operates as a conditional diffusion method. Specifically, conditioning is achieved through two key mechanisms. First, we initialize the reverse diffusion process from a partially noised version of the observed data $\mathbf{y}$ rather than pure Gaussian noise, ensuring that the entire denoising trajectory remains guided by the measurement. Second, we update the SIREN parameters at each diffusion step to fit the evolving estimate $\mathbf{x}_t$, allowing the model to capture coherent structure while avoiding overfitting to noise progressively. In this way, our method achieves fully unsupervised denoising, guided solely by the internal inductive bias of the neural representation and the smoothing nature of diffusion dynamics. The full procedure is summarized in Algorithm 1.

Figure 1 provides a conceptual comparison between DDPM, DDRM and our proposed SIREN-guided diffusion denoising model. In the figure, horizontal arrows indicate the direction of the diffusion process, with left-to-right arrows representing the forward diffusion (noise injection) and right-to-left arrows representing the reverse diffusion (denoising or reconstruction). The greyscale panels illustrate signal realizations at different diffusion steps, and the rightmost panel in each row corresponds to a realization that is (or is close to) pure Gaussian noise. In Figure 1, $\mathbf{y}$ denotes the observed noisy measurement. In DDRM (Figure 1b), $\mathbf{y}$ is assumed to be generated by a known degradation model $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{z}$, where $\mathbf{H}$ is an explicit degradation operator. In contrast, in our proposed method (Figure 1c), $\mathbf{y}$ corresponds to a noisy observation under the special case $\mathbf{H} = \mathbf{I}$ and no explicit degradation operator or projection is used. Instead, conditioning is achieved by initializing the reverse diffusion from a partially noised version of $\mathbf{y}$ and optimizing an implicit neural representation at each step.

All three frameworks are built upon the same formal definition of the forward diffusion process, where a clean signal $\mathbf{x}_0$ is gradually perturbed into Gaussian noise through a Markov chain, expressed as $q(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \mid \mathbf{x}_0)$. However, a key practical difference exists in how (or whether) this forward process is applied. In DDPM (Figure 1a), the model has access to clean training samples $\mathbf{x}_0$

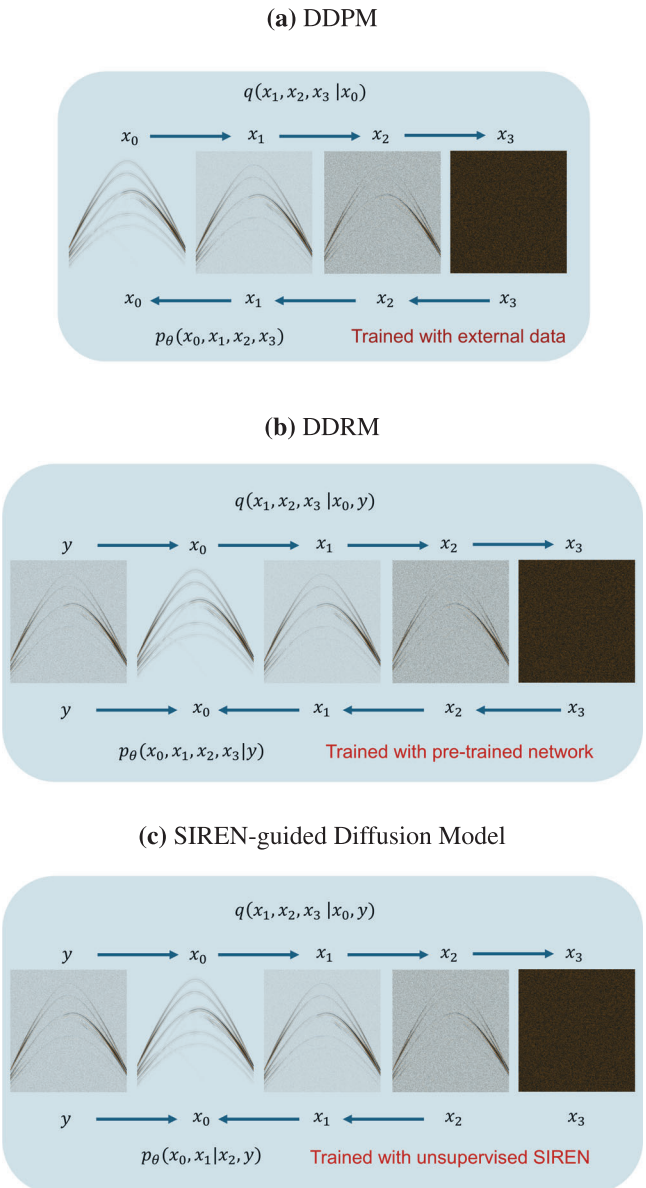

Figure 1 | Comparison of generative and restoration processes across three paradigms. DDPM generates clean samples from pure Gaussian noise without conditioning on observations. DDRM enables conditional restoration by incorporating observed measurements through a pre-trained diffusion model and explicit degradation operators enforced via projection steps during reverse diffusion. In contrast, the proposed SIREN-guided diffusion model achieves fully unsupervised denoising by replacing the pre-trained denoiser with an implicit neural representation (SIREN) and conditioning the reverse diffusion through partial noise injection from the observation and iterative optimization at each step.

and the forward process is explicitly used to simulate corrupted versions $\mathbf{x}_t$ at various timesteps. A neural network is then trained to reverse this corruption by predicting the reverse transition $p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$. Since the reverse process is trained to match the marginal data distribution, DDPM can generate realistic samples from pure Gaussian noise. However, it lacks any conditioning mechanism and cannot recover specific observations. DDRM (Figure 1b) builds on a pre-trained DDPM model and introduces conditioning by incorporating measurements $\mathbf{y}$. During

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

inference, DDRM starts from pure Gaussian noise $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ and performs a reverse diffusion process, but after each step it applies an explicit projection to enforce consistency with the observation $\mathbf{y}$. Although the forward process was used during DDPM pretraining, it is never applied to $\mathbf{y}$ during inference. Thus, while DDRM operates on noisy data, it does not simulate or require forward noising of the measurement. In contrast, our proposed method (Figure 1c) is fully unsupervised and does not rely on any pretraining or clean training data. Since we only have access to noisy observations $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$, we cannot apply the forward process directly from $\mathbf{x}_0$. Instead, we approximate a noised latent $\mathbf{x}_{T'}$ by adding further Gaussian noise to $\mathbf{y}$, that is, $\mathbf{x}_{T'} \sim q(\mathbf{x}_{T'} \mid \mathbf{y})$. This initialization allows us to start the reverse diffusion process from a partially degraded state that retains meaningful signal structure. The denoising function is modelled using an untrained SIREN, which is optimized at each step to fit the evolving variable $\mathbf{x}_t$. In this way, the reverse trajectory $p_\theta(\mathbf{x}_0, \mathbf{x}_1 \mid \mathbf{x}_{T'}, \mathbf{y})$ is implicitly conditioned on the observation, without any explicit projection or measurement model. While the forward diffusion process is not applied to $\mathbf{x}_0$ as in DDPM, our method effectively uses a surrogate version of it applied to $\mathbf{y}$, enabling a fully unsupervised reconstruction pipeline.

While random-noise denoising is a relatively well-posed inverse problem and untrained networks such as DIP or SIREN can already recover coherent signal structures, their practical performance is often limited by instability and strong sensitivity to early stopping, especially under high noise levels. Diffusion-based guidance is introduced in this work not to resolve ill-posedness, but to improve robustness and stabilization in fully unsupervised settings. By imposing a probabilistic noise-scheduling and coarse-to-fine refinement process, diffusion naturally regularizes the optimization trajectory and mitigates premature overfitting to noise. Compared to alternative stabilization strategies for untrained implicit neural representations, such as multiscale training or frequency continuation, diffusion provides a unified and principled framework that does not require explicit scale design, frequency partitioning or problem-specific tuning. This makes diffusion-guided INRs a flexible and robust choice even in settings where the underlying denoising problem is well posed.

3 | Examples

In this section, we evaluate the proposed method on both synthetic and real seismic datasets. We compare its effectiveness with a range of traditional and deep learning-based denoising methods to assess its effectiveness. As a traditional baseline, we include multichannel singular spectrum analysis (MSSA), a widely used technique in seismic data processing for noise suppression and signal enhancement (Oropeza and Sacchi 2010; Carozzi and Sacchi 2021). Among deep learning methods, we consider Noise2Void (Krull et al. 2019), a self-supervised denoising approach that requires no clean labels during training, and conditioned DDPM (Choi et al. 2021), a supervised diffusion-based restoration method guided by partial observations. We also include ablations to analyse the effect of different priors in our diffusion framework. Specifically, we test DIP (Ulyanov et al. 2018) and DIP-guided diffusion, in which DIP replaces SIREN as the implicit denoiser while keeping the rest of the pipeline unchanged. Similarly, we include standalone SIREN and our full SIREN-guided diffusion model. Both DIP- and SIREN-based methods are fully unsupervised and do not rely on any external training data.

3.1 | Synthetic Data Example

We first evaluate the proposed method using a synthetic dataset generated by the finite difference method. The data are based on the open-source Amoco statics test dataset, initially created by Mike O'Brien in 1994 under the supervision of Carl Regone at the Amoco Tulsa Technology Center. This dataset simulates various near-surface geological conditions typically leading to static-related distortions in seismic imaging. It represents a two-dimensional, purely acoustic model with constant density. While synthetic, the dataset is highly realistic and structurally rich, closely mimicking field seismic data – though it lacks ground roll effects due to its acoustic nature. The full dataset includes 65 shots with 1152 time samples and 3008 traces. We extract a small subset from a single shot with dimensions 256×256 for evaluation. For the conditioned DDPM baseline, we pretrain the model using the remaining 64 shots. All other methods, including DIP, SIREN and their diffusion-guided counterparts, are tested solely on the selected 256×256 patch without using any additional data or supervision.

For the first experiment, we evaluate the performance of various denoising methods on synthetic seismic data. The input SNR is set to 0 dB, indicating that the signal and noise power are equal. The corresponding noise profile is shown in Figure 2a. The visual results are summarized in Figure 3, which includes the original clean data, noisy observations, denoised outputs from different methods and their corresponding error panels computed as the difference from the ground truth.

To further evaluate the robustness of the denoising methods under more challenging conditions, we conducted a second experiment where the input SNR was set to −5 dB, meaning the noise power exceeds the signal power by a factor of approximately 3.16. The corresponding noise realization is shown in Figure 2b. This represents a highly noise-contaminated scenario, where coherent signal features are largely buried in random noise. The results are shown in Figure 4, including the noisy input, denoised outputs from each method and their corresponding error panels.

To comprehensively evaluate the denoising performance under varying noise levels, we compare multiple methods using both visual inspection (Figures 3 and 4) and quantitative output SNR values (Table 1) for two input SNR conditions: 0 and −5 dB. The traditional method MSSA and supervised (Noise2Void, Guided DDPM) and unsupervised (DIP, SIREN) learning-based approaches show a limited ability to fully suppress noise, particularly in areas heavily distorted by noise. In these cases, residual noise and signal leakage are still apparent, and some discontinuities in the reflection events can be observed. For instance, while Noise2Void and MSSA preserve the main structure well, they struggle to remove fine-grained noise artefacts, especially under the −5 dB scenario. On the other hand, although unsupervised approaches like DIP and SIREN can recover coherent features, they tend to introduce unwanted artefacts in the output due to their strong internal priors and lack of external guidance. This is particularly evident in the error

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

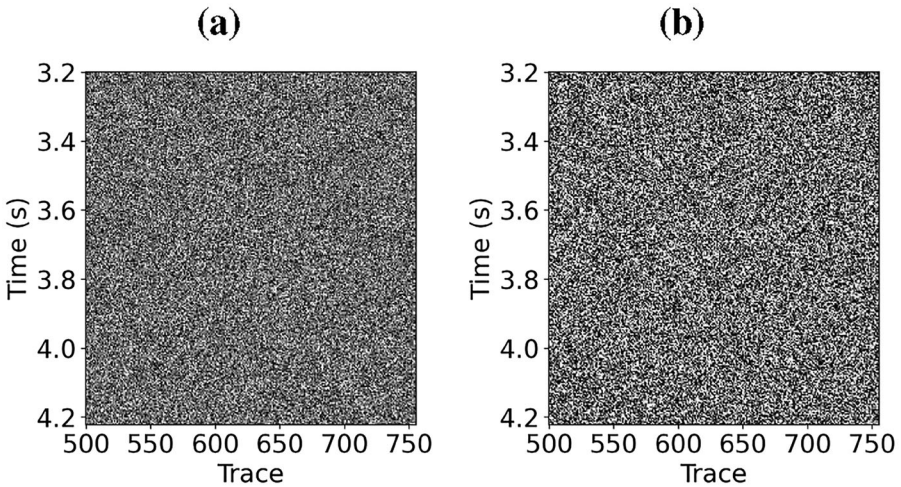

Figure 2 | Synthetic noise profiles used in the previous two examples. (a) Noise realization with SNR = 0 dB. (b) Noise realization with SNR=0dB.

Table 1 | Output SNR (dB) of denoised results for each method under different input noise levels.

Method	$SNR_{in} = 0$ dB	$SNR_{in} = -5$ dB
MSSA	13.27	9.10
Guided DDPM	10.05	8.23
Noise2Void	14.23	10.74
DIP	12.47	9.18
DIP + Diffusion	14.20	9.97
SIREN	13.86	10.91
SIREN + Diffusion	**16.24**	**11.95**

The best result in each column is shown in bold.

panels and smoother regions where artificial structures appear. The results are significantly improved when diffusion priors are introduced into these unsupervised frameworks (i.e., DIP-guided diffusion and SIREN-guided diffusion). These hybrid methods successfully suppress the noise while preserving continuity in geological features and minimizing hallucinated patterns. This is reflected in both the cleaner appearance of the denoised images and the higher SNR scores in Table 1, where the proposed SIREN-guided diffusion method consistently achieves the best performance across both noise levels

One of the main challenges when using unsupervised methods like SIREN and DIP for denoising is their tendency to overfit the noise, which necessitates careful early stopping during training. As illustrated in Figure 5, both methods initially reconstruct the underlying seismic signal, resulting in a steady increase in output SNR. However, after reaching a peak, the SNR begins to decline as the networks start fitting to the noise in the input data. This behaviour highlights the importance of identifying the optimal stopping point to avoid performance degradation. Finding this early stopping point is particularly difficult in real-world applications, where ground truth data is unavailable for comparison. Furthermore, the figure shows that under more severe noise conditions ($SNR_{in} = -5$ dB), the SNR drops more rapidly, choosing the stopping point even more sensitive and challenging. It is also noteworthy that SIREN exhibits a slower SNR decline than DIP, suggesting that it is slightly more robust to noise overfitting. Nevertheless, both methods require early stopping strategies to prevent the network from fitting noise, especially in high-noise scenarios.

Figure 6 presents the output SNR curves for DIP-guided diffusion (left column) and SIREN-guided diffusion (right column) under two input noise levels ($SNR_{in} = 0$ and -5 dB). Each row corresponds to various reverse diffusion steps from top to bottom. For each step, DIP or SIREN is optimized once. Hence, increasing the number of steps increases the total number of DIP or SIREN optimization epochs rather than increasing the epochs per step. The results demonstrate that integrating the diffusion model significantly improves the robustness of these unsupervised methods. The SNR curves show that, compared to their standalone versions, both DIP and SIREN experience slower SNR decay after reaching their peak. Most notably, SIREN-guided diffusion maintains its peak SNR without dropping, indicating strong resistance to overfitting and noise learning, even with more total optimization steps. Moreover, the figure shows that the network converges more quickly when guided by diffusion: with fewer than 1000 total optimization epochs, both DIP and SIREN can effectively recover the signal. In contrast, their standalone counterparts typically require several thousand epochs to reach a comparable reconstruction quality under the same experimental settings. These results confirm that the SIREN-guided diffusion denoising framework improves robustness and accelerates convergence.

In addition, the results suggest that the proposed framework is not highly sensitive to the exact number of reverse diffusion steps. As shown in Figure 6, increasing the number of steps mainly changes the optimization trajectory while leading to similar peak SNR values, indicating stable performance across different diffusion lengths. In our experiments, the partial noise initialization level is set to $T' = T/2$, which provides a moderate initialization between the observation and heavily noised states. This choice preserves sufficient signal structure for stable reconstruction while still introducing enough stochasticity for diffusion-guided refinement.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

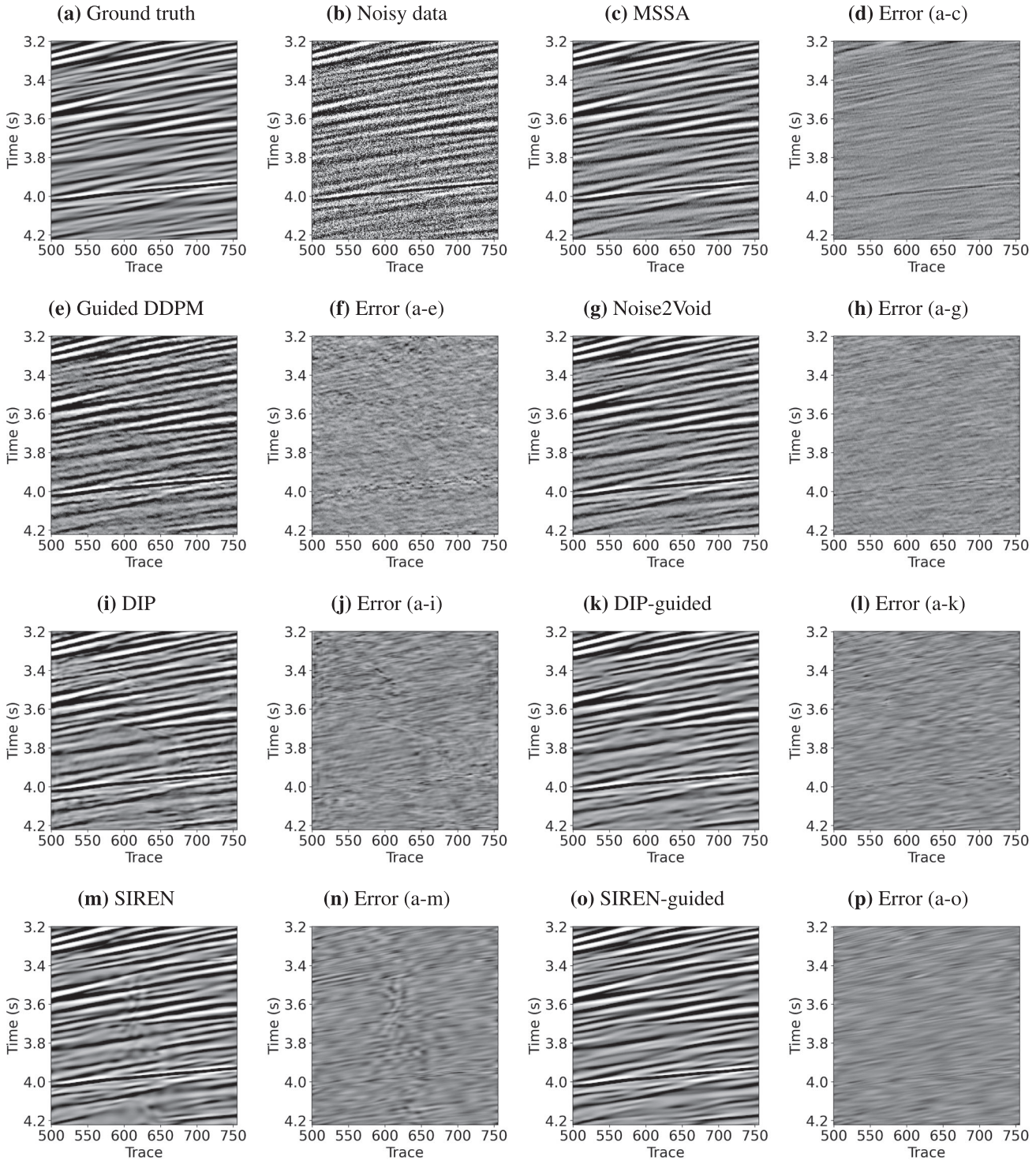

Figure 3 | Denoising performance comparison on synthetic seismic data with input SNR = −5 dB. (a) Ground truth, (b) noisy input and (c)–(p) results of various denoising methods along with their corresponding error panels (difference from ground truth). Traditional method: (c) MSSA. Deep learning-based methods: (e) Guided DDPM, (g) Noise2Void, (i) DIP, (k) DIP-guided diffusion, (m) SIREN and (o) SIREN-guided diffusion. Error panels: (d) MSSA, (f) Guided DDPM, (h) Noise2Void, (j) DIP, (l) DIP-guided diffusion, (n) SIREN and (p) SIREN-guided diffusion.

Traditionally, the original DDPM model (Ho et al. 2020) adopts a linear beta schedule, where the noise variance β_t increases linearly over diffusion steps. This leads to a uniform degradation of the input over time but can be suboptimal for denoising quality and sample efficiency. In our framework, we adopt the cosine beta schedule proposed by Nichol and Dhariwal (2021), which uses a cosine function to control the cumulative noise. This schedule results in a smoother and slower noise accumulation at early steps, improving robustness and stability, especially when using fewer diffusion steps. To evaluate the impact of different beta schedules, we plot the output SNR curves in Figure 7 for five types of schedules: constant, linear, quadratic, sigmoid and

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

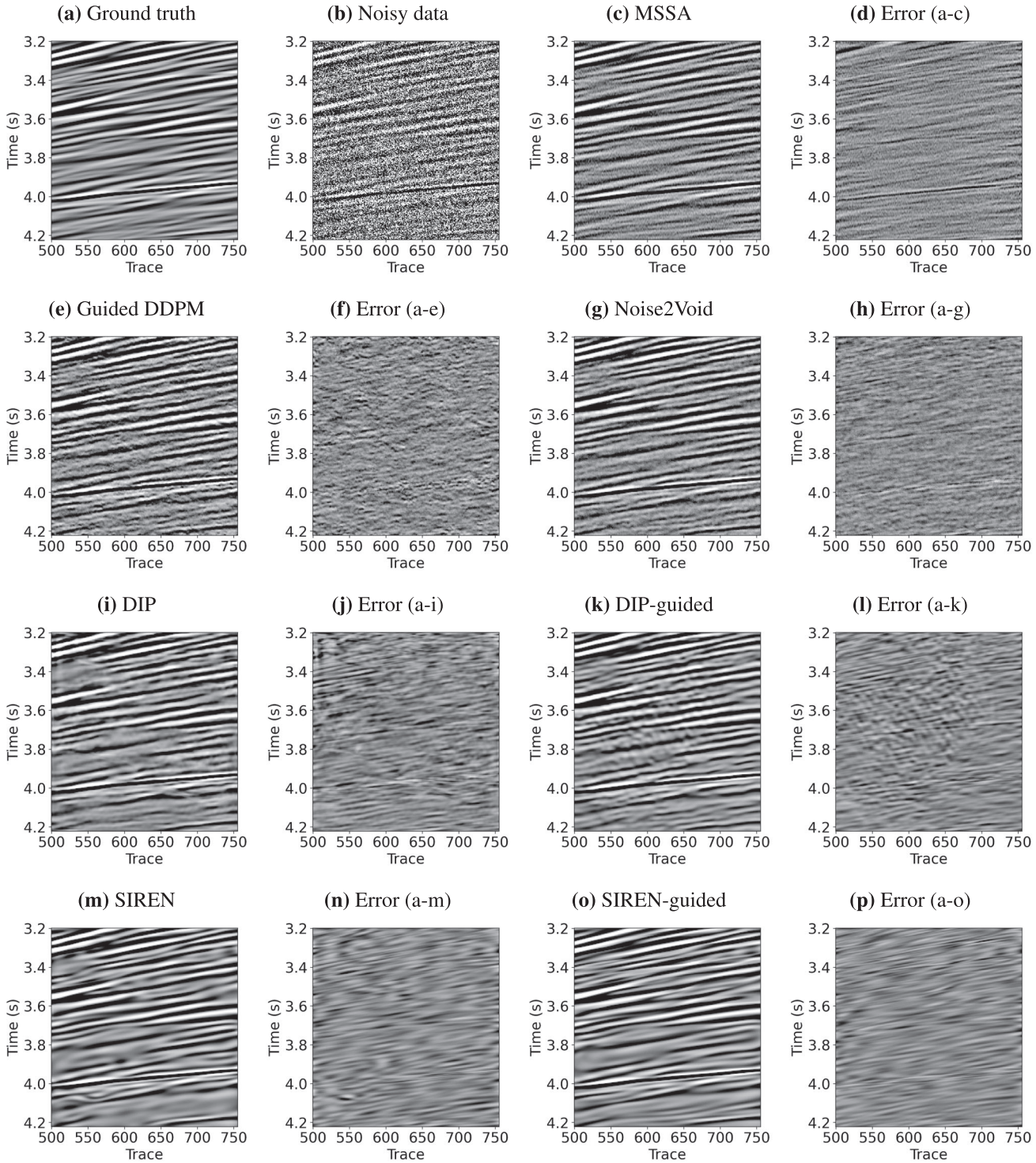

Figure 4 | Denoising performance comparison on synthetic seismic data with input SNR=−5dB. (a) Ground truth, (b) noisy input and (c)–(p) results of various denoising methods along with their corresponding error panels (difference from ground truth). Traditional method: (c) MSSA. Deep learning-based methods: (e) Guided DDPM, (g) Noise2Void, (i) DIP, (k) DIP-guided diffusion, (m) SIREN, (o) SIREN-guided diffusion. Error panels: (d) MSSA, (f) Guided DDPM, (h) Noise2Void, (j) DIP, (l) DIP-guided diffusion, (n) SIREN and (p) SIREN-guided diffusion.

cosine, under two input noise conditions ($SNR_{in} = 0$ and −5 dB). The curves show that the constant schedule performs the worst, producing unstable SNR curves and poor recovery. In contrast, cosine, quadratic and sigmoid schedules demonstrate the fastest convergence and the highest final SNR, particularly under high noise levels. The cosine schedule consistently yields robust and stable reconstruction, justifying its selection in our main experiments. The mathematical expressions and corresponding maximum output SNR values for each beta schedule are listed in Table 2.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

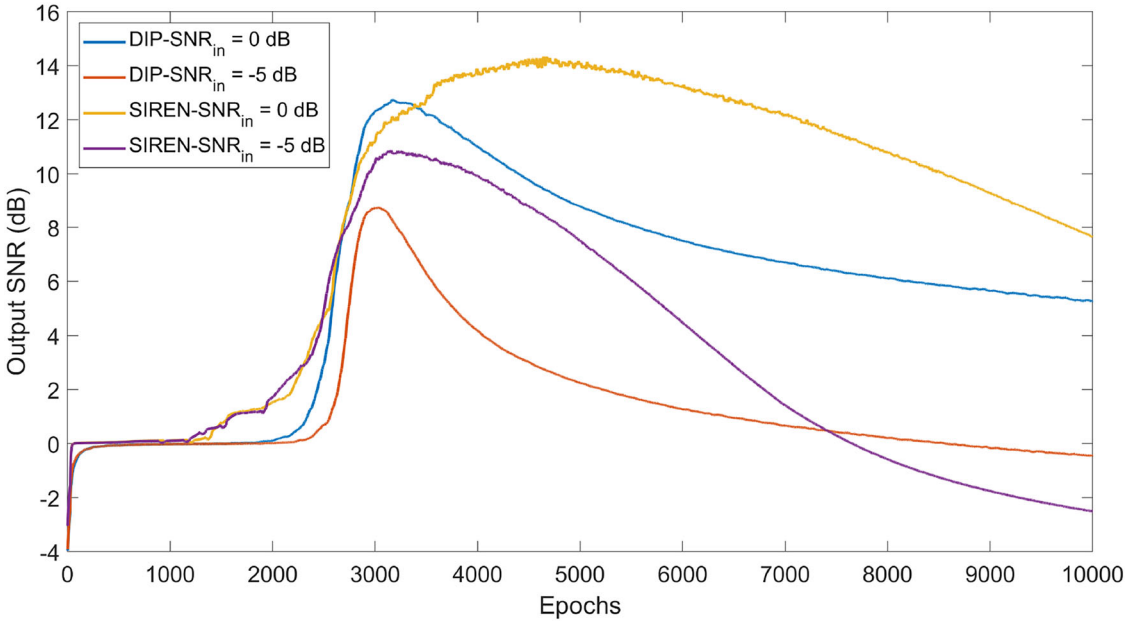

Figure 5 | Output SNR curves of DIP and SIREN under different input noise levels. Both methods initially improve SNR by reconstructing the signal, but overfit to noise over time, causing SNR to drop. SIREN shows slightly better robustness than DIP, especially at lower input SNR

Table 2 | Comparison of different beta schedules and their effect on denoising performance under different input SNR levels.

Beta schedule	$\mathbf{SNR_{in} = 0\ dB}$	$\mathbf{SNR_{in} = -5\ dB}$
Linear $\beta_t = \text{linspace}(\beta_{\text{start}}, \beta_{\text{end}})$	15.62	11.01
Cosine $\alpha_t = \cos^2\left(\frac{(t/T+s)}{1+s} \cdot \frac{\pi}{2}\right), \beta_t = 1 - \frac{\alpha_{t+1}}{\alpha_t}$	**16.24**	**11.95**
Quadratic $\beta_t = \left(\text{linspace}(\sqrt{\beta_{\text{start}}}, \sqrt{\beta_{\text{end}}})\right)^2$	15.91	11.23
Constant $\beta_t = \beta_{\text{const}}$	10.93	8.95
Sigmoid $\beta_t = \sigma(t) \cdot (\beta_{\text{end}} - \beta_{\text{start}}) + \beta_{\text{start}}$	16.13	11.42

The best result in each column is shown in bold.

3.2 | Field Data Example

We now apply the proposed method to prestack field seismic data, using a common reflection point (CRP) gather as our test case. The events within a CRP gather are typically coherent and approximately horizontal, making them well-suited for unsupervised learning methods such as DIP and SIREN, which are known to prioritize structured features early in the optimization process. In this example, we compare the performance of three methods: the traditional MSSA, DIP-guided diffusion and SIREN-guided diffusion. As is typical with real seismic data, no ground truth is available to compute the SNR. Instead, we evaluate denoising quality using the local similarity map between the denoised result and the removed noise (Fomel 2007; Y. Chen and Fomel 2015). The local similarity metric measures the structural alignment between two datasets within a localized window and is commonly used to estimate signal leakage. A high similarity between the denoised output and the residual noise indicates that meaningful signal energy has been mistakenly removed. Figure 8 shows the original noisy CRP gather, where background noise heavily contaminates useful reflections, resulting in weak and blurred signal events. Figure 9 presents the denoised outputs, corresponding noise estimates and local similarity maps for all three methods. The MSSA method shows the most significant signal leakage, as indicated by high local similarity values. In contrast, both DIP-guided and SIREN-guided diffusion models achieve better signal preservation. While the DIP-guided model still has some visible noise near the edges, the SIREN-guided diffusion model produces cleaner edges and a more structurally consistent result.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

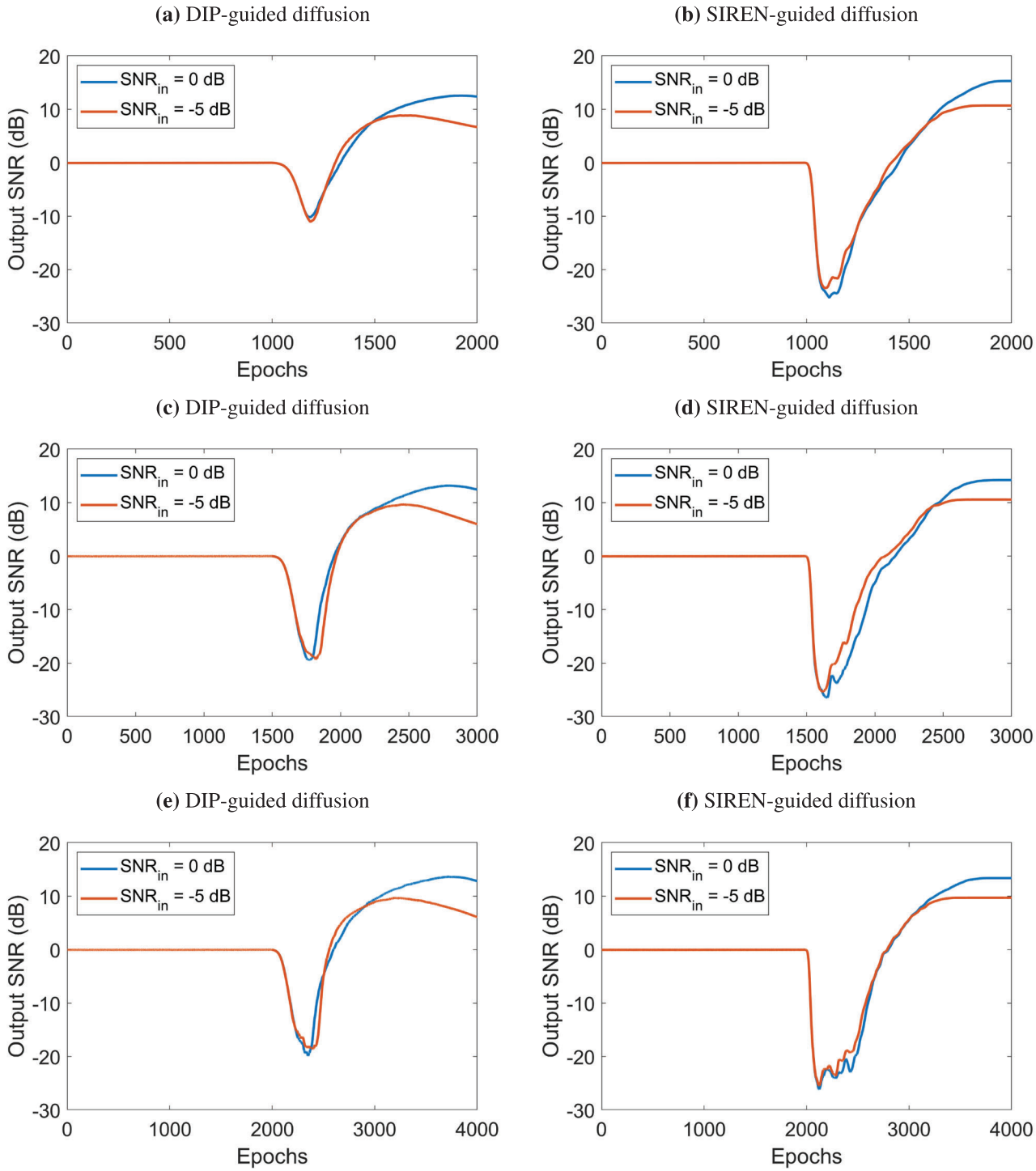

Figure 6 | Output SNR curves of DIP-guided diffusion (left) and SIREN-guided diffusion (right) under different input noise levels and reverse diffusion steps. Each row corresponds to an increasing number of reverse diffusion steps, which increases the total number of DIP or SIREN optimization epochs. Incorporating the diffusion model improves robustness, as evidenced by slower SNR decay and stable SNR after reaching the peak in the case of SIREN-guided diffusion. Additionally, both methods achieve faster convergence, with signal recovery completed in fewer than 1000 epochs.

4 | Discussion

4.1 | Computational Cost and Runtime

As noted in the paper, the SIREN is updated only a small number of times (typically 1–5 iterations) at each reverse diffusion step, and the network parameters are warm-started rather than reinitialized. In the synthetic experiments, we used one SIREN update per reverse step, resulting in a total of $T/2 = 1000$ SIREN optimization steps for $T = 2000$, which is comparable to the number of iterations commonly required when using SIREN alone for denoising. Consequently, the overall computational cost of the proposed method is of the same order of magnitude as standalone SIREN, despite the additional diffusion operations, whose

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

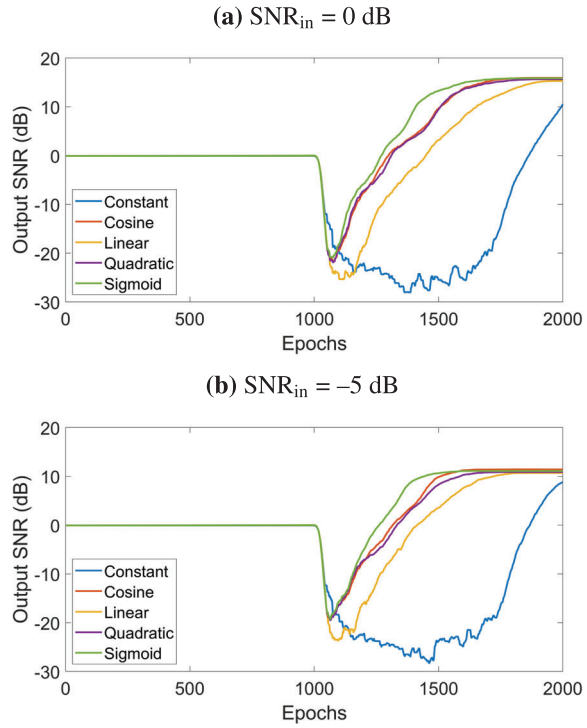

Figure 7 | Comparison of five beta schedules (constant, linear, quadratic, sigmoid and cosine) regarding output SNR performance at two input noise levels. The cosine, sigmoid and quadratic schedules demonstrate faster convergence and higher peak SNR. In contrast, the constant schedule performs poorly, showing unstable SNR and slow recovery.

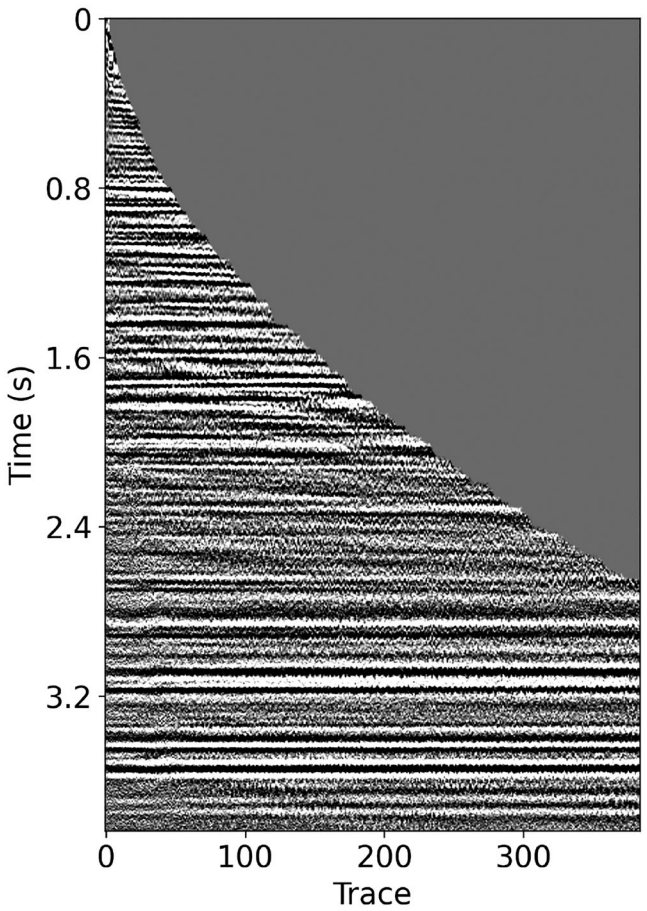

Figure 8 | Noisy common reflection point (CRP) gather from a prestack field dataset. The seismic reflections are contaminated by strong background noise, making some weak signal events difficult to identify.

overhead is lightweight compared to the network optimization. For example, in our 2D 256×256 experiments, the proposed method requires on the order of 1 min on a single GPU.

4.2 | Why SIREN Instead of DIP

Previous real data examples show that both DIP and SIREN-guided diffusion models work for seismic denoising. However, SIREN offers several unique advantages that make it more suitable in certain scenarios. One key advantage is its ability to handle irregularly sampled data. As an INR, SIREN maps continuous spatial coordinates to signal values, allowing it to naturally interpolate data at arbitrary locations-even on non-uniform or sparse grids. In contrast, DIP relies on CNNs, which are inherently defined on regular grids and require data to be evenly sampled in both spatial dimensions. This limitation makes DIP inapplicable to irregularly sampled scenarios such as randomly missing traces in seismic gathers.

Another advantage of SIREN is its flexibility to input resolution. SIREN does not rely on spatial downsampling or upsampling and can be evaluated at any coordinate resolution, making it resolution-agnostic. On the other hand, DIP typically uses an encoder–decoder structure such as a U-Net, which consists of a sequence of downsampling (e.g., strided convolutions) and upsampling (e.g., bilinear or transposed convolutions) operations. Each downsampling layer reduces the spatial dimensions by a factor of 2, and each upsampling layer restores it. This introduces a structural constraint: the input size must be divisible by 2^n, where n is the number of downsampling layers. For example, with five downsampling stages, the input height and width must be divisible by 32. Otherwise, the feature maps cannot be correctly reconstructed, leading to shape mismatches or reconstruction artefacts. This limitation requires users to pad or crop the input data to fit the network's architectural constraints, which can be problematic for data with arbitrary or non-standard sizes. In contrast, SIREN operates over continuous coordinates. It is free from such architectural constraints, making it a more flexible and scalable choice for seismic applications involving irregular sampling or variable resolution.

Figure 10 presents an interpolation example using the same dataset shown in Figure 2a. In this case, two-thirds of the original traces were randomly removed to simulate irregular sampling. Figure 10a shows the clean, irregularly undersampled data, while Figure 10d displays the corresponding noisy version. The interpolation results for both cases are shown in Figure 10b and 10e, respectively. Figure 10c and 10f illustrates the corresponding interpolation errors, computed by comparing the interpolated results with the original fully sampled data in Figure 2a. Despite the severe undersampling, the SIREN-guided diffusion model can effectively recover the missing traces and reconstruct coherent seismic events, demonstrating its strong capability in handling irregularly sampled data.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

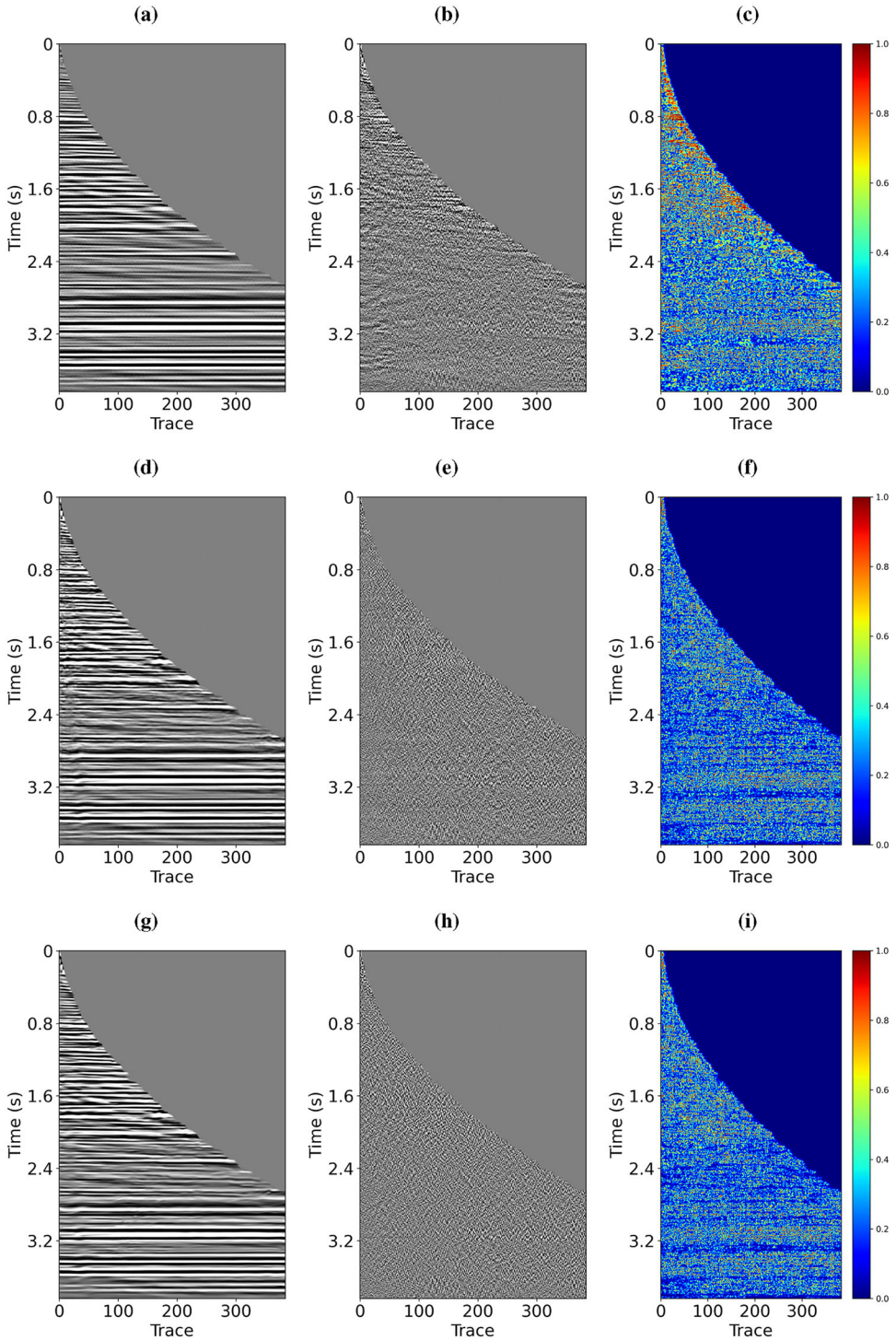

Figure 9 | Comparison of denoising performance using MSSA (top row), DIP-guided diffusion (middle row) and SIREN-guided diffusion (bottom row). Each row shows the denoised output (left), estimated noise (middle) and local similarity map (right). The local similarity map is used to assess signal leakage, with brighter areas indicating higher similarity between the denoised output and residual noise.

Although this example focuses on interpolation rather than pure denoising, the proposed SIREN-guided diffusion model does not require an explicit degradation operator **H**, which is typically necessary in other diffusion-based inverse problem frameworks. This is because SIREN learns a continuous mapping from spatial coordinates to signal values, enabling it to represent and reconstruct the seismic field at arbitrary locations directly. As a result, the model can naturally handle missing or irregularly sampled data without relying on predefined forward operators to simulate degradation or sampling masks. This inherent property of SIREN makes the method highly flexible and generalizable to various inverse problems beyond denoising, including interpolation, inpainting and reconstruction from sparse observations.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

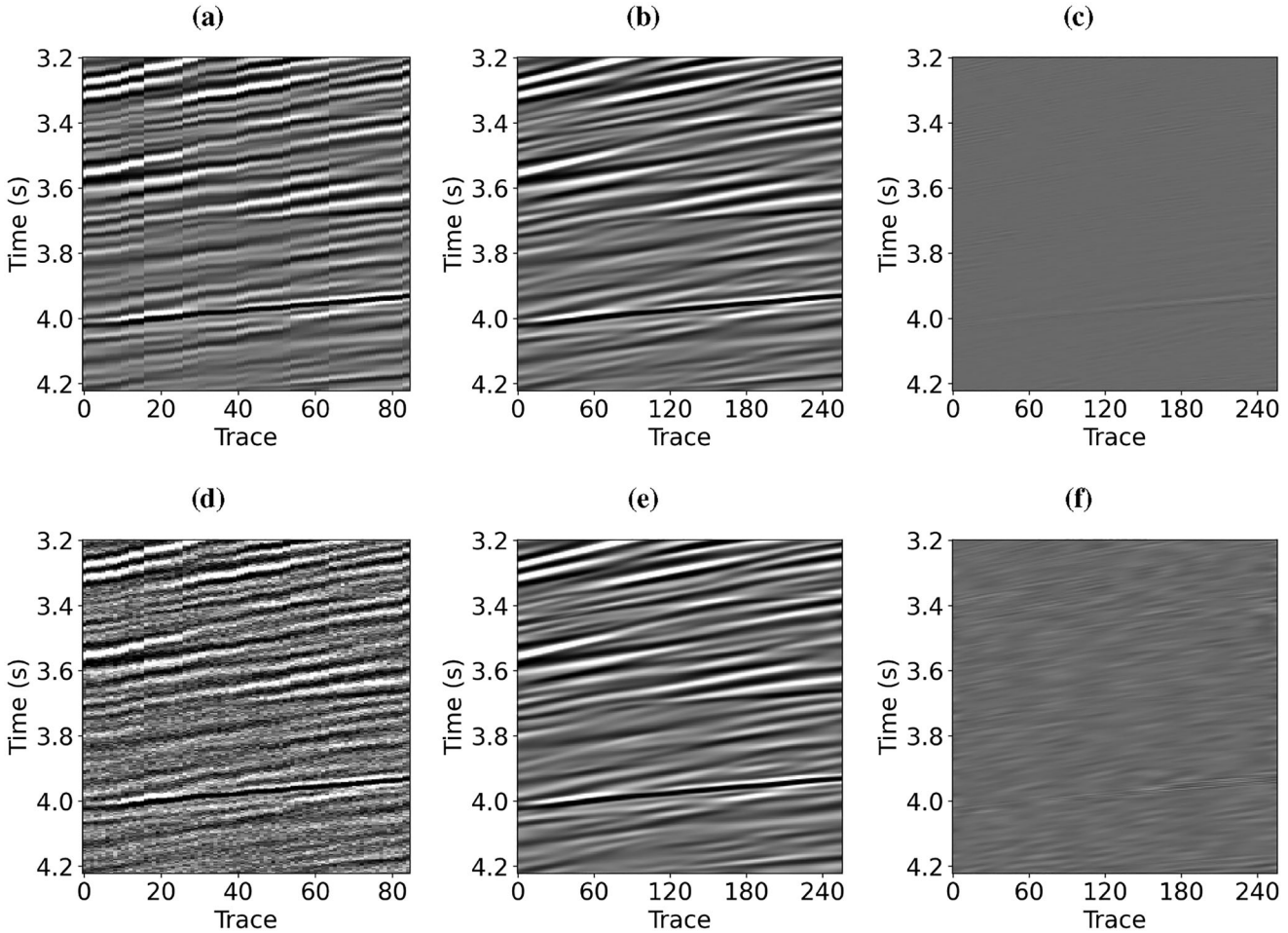

Figure 10 | Interpolation results using the SIREN-guided diffusion model on irregularly sampled seismic data. (a) Clean seismic section with two-thirds of the traces randomly removed. (b) Interpolated result from (a). (c) Interpolation error compared with the original fully sampled data. (d) Noisy seismic section with the same irregular sampling. (e) Interpolated result from (d). (f) Corresponding interpolation error.

4.3 | Extending the Method to Non-Gaussian Noise

This study primarily applies the proposed method to suppress random background noise in seismic data. However, seismic records are often contaminated by another challenging type of interference known as erratic noise, which is characterized by high-amplitude, sparsely distributed spikes that deviate significantly from the Gaussian distribution. Traditional denoising methods based on the ℓ_2 loss function are generally ineffective at mitigating such non-Gaussian outliers. In contrast, unsupervised frameworks such as SIREN and DIP can be adapted to handle erratic noise by employing the ℓ_1 norm, which enhances robustness to outliers. Conventional diffusion models like DDPM and DDRM, however, are built upon the assumption of additive Gaussian noise and typically underperform when facing strongly non-Gaussian corruptions. Modifying diffusion frameworks to accommodate erratic noise, especially when combined with implicit neural representations such as SIREN, remains an open and promising research direction in seismic denoising. Recent work has explored diffusion models with non-Gaussian corruption processes, suggesting that the diffusion framework could potentially be extended beyond the standard Gaussian noise assumption (Nachmani et al. 2021). In the context of the proposed framework, several directions could be explored to improve robustness to erratic noise. For example, one possibility is to incorporate robust loss functions (e.g., ℓ_1 or Huber loss) in the SIREN optimization to improve resilience to sparse outliers. Another approach would be to modify the diffusion corruption model to account for heavy-tailed or mixed noise distributions rather than purely Gaussian noise. Additionally, combining diffusion-guided reconstruction with explicit sparse noise modelling may provide a promising strategy for separating coherent seismic signals from impulsive noise.

5 | Conclusions

In this work, we presented a fully unsupervised seismic denoising framework by integrating an untrained SIREN into the reverse process of a diffusion model. The proposed method leverages both the structural bias of implicit neural representations and the generative strength of diffusion sampling by initializing from a partially noised observation and progressively refining the signal through SIREN optimization. Unlike traditional or supervised deep learning approaches, our framework does not

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

rely on clean labels or pretraining, making it particularly suitable for real seismic data. Experimental results on synthetic and field datasets demonstrate that our method achieves superior denoising performance and preserves structural features more effectively than existing baselines. Future work may extend this approach to other seismic inverse problems such as interpolation, deblending or full waveform inversion.

Acknowledgements

The second author is supported by the Key Program of the National Natural Science Foundation of China (No. 42530802). The first and third authors' work was funded by CREWES industrial sponsors and NSERC (Natural Sciences and Engineering Research Council of Canada) through grant CRDPJ 543578-19. Further support was provided by Emissions Reduction Alberta through the ACT4-SPARSE project.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abma, R., and J. Claerbout. 1995. "Lateral Prediction for Noise Attenuation by Tx and Fx Techniques." *Geophysics* 60, no. 6: 1887–1896.

Batzolis, G., and J. Stanczuk, C.-B. Schönlieb, and C. Etmann. 2021. "Conditional Image Generation with Score-Based Diffusion Models." Preprint arXiv, November 26. https://doi.org/10.48550/arXiv.2111.13606.

Beresford-Smith, G., and R. Rango. 1989. "Suppression of Ground Roll by Windowing in Two Domains." *First Break* 7, no. 2: 55–63.

Carozzi, F., and M. D. Sacchi. 2021. "Interpolated Multichannel Singular Spectrum Analysis: A Reconstruction Method that Honors True Trace Coordinates." *Geophysics* 86, no. 1: V55–V70.

Chen, Y., and S. Fomel. 2015. "Random Noise Attenuation Using Local Signal-and-Noise Orthogonalization." *Geophysics* 80, no. 6: WD1–WD9.

Chen, Y., D. Zhang, Z. Jin, et al. 2016. "Simultaneous Denoising and Reconstruction of 5-D Seismic Data Via Damped Rank-Reduction Method." *Geophysical Journal International* 206, no. 3: 1695–1717.

Choi, J., S. Kim, Y. Jeong, Y. Gwon, and S. Yoon. 2021. "ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models." Preprint arXiv, September 15. https://doi.org/10.48550/arXiv.2108.02938.

Chung, H., and S. J. C. Ye. 2022. "Score-Based Diffusion Models for Accelerated MRI." *Medical Image Analysis* 80: 102479.

Daley, T. M., B. M. Freifeld, J. Ajo-Franklin, et al. 2013. "Field Testing of Fiber-Optic Distributed Acoustic Sensing (DAS) for Subsurface Seismic Monitoring." *The Leading Edge* 32, no. 6: 699–706.

Deighan, A. J., and D. R. Watts. 1997. "Ground-Roll Suppression Using the Wavelet Transform." *Geophysics* 62, no. 6: 1896–1903.

Deng, F., S. Wang, X. Wang, and P. Fang. 2024. "Seismic Data Reconstruction Based on Conditional Constraint Diffusion Model." *IEEE Geoscience and Remote Sensing Letters* 21: 7502305.

Durall, R., A. Ghanim, M. R. Fernandez, N. Ettrich, and J. Keuper. 2023. "Deep Diffusion Models for Seismic Processing." *Computers & Geosciences* 177: 105377.

Fomel, S. 2007. "Local Seismic Attributes." *Geophysics* 72, no. 3: A29–A33.

Gaci, S. 2013. "The Use of Wavelet-Based Denoising Techniques to Enhance the First-Arrival Picking on Seismic Traces." *IEEE Transactions on Geoscience and Remote Sensing* 52, no. 8: 4558–4563.

Gao, Y., J. Zhang, H. Li, and G. Li. 2022. "Incorporating Structural Constraint Into the Machine Learning High-Resolution Seismic Reconstruction." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–12.

Herrmann, F. J., D. Wang, and D. J. Verschuur. 2008. "Adaptive Curvelet-Domain Primary-Multiple Separation." *Geophysics* 73, no. 3: A17–A21.

Ho, J., A. Jain, and P. Abbeel. 2020. "Denoising Diffusion Probabilistic Models." *Advances in Neural Information Processing Systems* 33: 6840–6851.

Hui, M., Z. Wei, H. Zhu, and F. Xia, and Y. Zhou. 2024. "Microdiffusion: Implicit Representation-Guided Diffusion for 3D Reconstruction From Limited 2D Microscopy Projections." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 11460–11469. IEEE.

Ibrahim, A., and M. D. Sacchi. 2014. "Simultaneous Source Separation Using a Robust Radon Transform." *Geophysics* 79, no. 1: V1–V11.

J Beasley, C. 2008. "A New Look at Marine Simultaneous Sources." *The Leading Edge* 27, no. 7: 914–917.

Kawar, B., M. Elad, S. Ermon, and J. Song. 2022. "Denoising Diffusion Restoration Models." *Advances in Neural Information Processing Systems* 35: 23593–23606.

Khader, F., G. Müller-Franzes, S. Tayebi Arasteh, et al. 2023. "Denoising Diffusion Probabilistic Models for 3D Medical Image Generation." *Scientific Reports* 13, no. 1: 7303.

Kingma, D., T. Salimans, B. Poole, and J. Ho. 2021. "Variational Diffusion Models." *Advances in Neural Information Processing Systems* 34: 21696–21707.

Krull, A., T.-O. Buchholz, and F. Jug. 2019. "Noise2void-Learning Denoising From Single Noisy Images." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2129–2137. IEEE.

Lehtinen, J., J. Munkberg, J. Hasselgren, et al. 2018. "Noise2Noise: Learning Image Restoration Without Clean Data." Preprint arXiv, October 29. https://doi.org/10.48550/arXiv.1803.04189.

Li, J., D. Liu, and M. D. Sacchi. 2025. "Unsupervised Ground-Roll Attenuation via Implicit Neural Representations." *Geophysics* 90, no. 2: V111–V121.

Li, J., and M. D. Sacchi. 2021. "An Lp-Space Matching Pursuit Algorithm and Its Application to Robust Seismic Data Denoising via Time-Domain Radon Transforms." *Geophysics* 86, no. 3: V171–V183.

Li, J., D. Trad, and D. Liu. 2024. "Robust Seismic Data Denoising via Self-Supervised Deep Learning." *Geophysics* 89, no. 5: V437–V451.

Li, J., X. Wu, and Z. Hu. 2021. "Deep Learning for Simultaneous Seismic Image Super-Resolution and Denoising." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–11.

Li, W., and J. Wang. 2021. "Residual Learning of Cycle-GAN for Seismic Data Denoising." *IEEE Access* 9: 11585–11597.

Liu, D., Z. Deng, C. Wang, X. Wang, and W. Chen. 2020. "An Unsupervised Deep Learning Method for Denoising Prestack Random Noise." *IEEE Geoscience and Remote Sensing Letters* 19: 1–5.

Liu, D., M. D. Sacchi, X. Wang, and W. Chen. 2023. "Unsupervised Deep Learning for Ground Roll and Scattered Noise Attenuation." *IEEE Transactions on Geoscience and Remote Sensing* 61: 1–13.

Lugmayr, A., M. Danelljan, A. Romero, and F. Yu, R. Timofte, and L. Van Gool. 2022. "Repaint: Inpainting Using Denoising Diffusion Probabilistic Models." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11461–11471. IEEE.

Mandelli, S., V. Lipari, P. Bestagini, and S. Tubaro. 2019. "Interpolation and Denoising of Seismic Data Using Convolutional Neural Networks." Preprint arXiv, October 21. https://doi.org/10.48550/arXiv.1901.07927.

Nachmani, E., R. S. Roman, and L. Wolf. 2021. "Non Gaussian Denoising Diffusion Models." Preprint arXiv, June 14. https://doi.org/10.48550/arXiv.2106.07582.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Naghizadeh, M. 2012. "Seismic Data Interpolation and Denoising in the Frequency-Wavenumber Domain." *Geophysics* 77, no. 2: V71–V80.

Naghizadeh, M., and M. D. Sacchi. 2007. "Multistep Autoregressive Reconstruction of Seismic Records." *Geophysics* 72, no. 6: V111–V118.

Nichol, A. Q., and P. Dhariwal. 2021. "Improved Denoising Diffusion Probabilistic Models." In *International Conference on Machine Learning*, 8162–8171. PMLR.

Oropeza, V., and M. Sacchi. 2011. "Simultaneous Seismic Data Denoising and Reconstruction via Multichannel Singular Spectrum Analysis." *Geophysics* 76, no. 3: V25–V32.

Oropeza, V. E., and M. D. Sacchi. 2010. "A Randomized SVD for Multichannel Singular Spectrum Analysis (MSSA) Noise Attenuation." In *SEG Technical Program Expanded Abstracts 2010*, 3539–3544. Society of Exploration Geophysicists.

Qiu, C., B. Wu, N. Liu, X. Zhu, and H. Ren. 2021. "Deep Learning Prior Model for Unsupervised Seismic Data Random Noise Attenuation." *IEEE Geoscience and Remote Sensing Letters* 19: 1–5.

Richardson, A., and C. Feller. 2019. "Seismic Data Denoising and Deblending Using Deep Learning." Preprint Arxiv, July 2. https://doi.org/10.48550/arXiv.1907.01497.

Robinson, E. A., and S. Treitel. 1967. "Principles of Digital Wiener Filtering." *Geophysical Prospecting* 15, no. 3: 311–332.

Ronneberger, O., P. Fischer, and T. Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: Proceedings of the 18th International Conference, Munich, Germany, October 5-9, 2015, Part III*, 234–241. Springer.

Sasaki, H., C. G. Willcocks, and T. P. Breckon. 2021. "Unit-DDPM: Unpaired Image Translation with Denoising Diffusion Probabilistic Models." Preprint arXiv, April 12. https://doi.org/10.48550/arXiv.2104.05358.

Shan, H., J. Ma, and H. Yang. 2009. "Comparisons of Wavelets, Contourlets and Curvelets in Seismic Denoising." *Journal of Applied Geophysics* 69, no. 2: 103–115.

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* 33: 7462–7473.

Song, Y., C. Durkan, I. Murray, and S. Ermon. 2021. "Maximum Likelihood Training of Score-Based Diffusion Models." *Advances in Neural Information Processing Systems* 34: 1415–1428.

Sun, J., S. Slang, T. Elboth, T. L. Greiner, S. McDonald, and L.-J. Gelius. 2020. "Attenuation of Marine Seismic Interference Noise Employing a Customized U-Net." *Geophysical Prospecting* 68, no. 3: 845–871.

Tashiro, Y., J. Song, Y. Song, and S. Ermon. 2021. "CSDI: Conditional Score-Based Diffusion Models for Probabilistic Time Series Imputation." *Advances in Neural Information Processing Systems* 34: 24804–24816.

Ulyanov, D., A. Vedaldi, and V. Lempitsky. 2018. "Deep Image Prior." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9446–9454. IEEE.

Vera Rodriguez, I., D. Bonar, and M. Sacchi. 2012. "Microseismic Data Denoising Using a 3C Group Sparsity Constrained Time-Frequency Transform." *Geophysics* 77, no. 2: V21–V29.

Wang, Y., J. Yu, and J. Zhang. 2022. "Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model." Preprint Arxiv, December 7. https://doi.org/10.48550/arXiv.2212.00490.

Whang, J., M. Delbracio, H. Talebi, C. Saharia, A. G. Dimakis, and P. Milanfar. 2022. "Deblurring Via Stochastic Refinement." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16293–16303. IEEE.

Wu, Z., X. Chen, S. Xie, J. Shen, and Y. Zeng. 2023. "Super-Resolution of Brain MRI Images Based on Denoising Diffusion Probabilistic Model." *Biomedical Signal Processing and Control* 85: 104901.

WW, G., Y. Liu, M. Zhang, and H. Zhang. 2022. "Dropout-Based Robust Self-Supervised Deep Learning for Seismic Data Denoising." *IEEE Geoscience and Remote Sensing Letters* 19: 1–5.

Wyatt, J., A. Leach, S. M. Schmon, and C. G. Willcocks. 2022. "Anoddpm: Anomaly Detection with Denoising Diffusion Probabilistic Models Using Simplex Noise." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 650–656. IEEE.

Xiao, Y., K. Li, Y. Dou, W. Li, Z. Yang, and X. Zhu. 2024. "Diffusion Models for Multidimensional Seismic Noise Attenuation and Superresolution." *Geophysics* 89, no. 5: V479–V492.

Yang, L., W. Chen, H. Wang, and Y. Chen. 2021. "Deep Learning Seismic Random Noise Attenuation Via Improved Residual Convolutional Neural Network." *IEEE Transactions on Geoscience and Remote Sensing* 59, no. 9: 7968–7981.

Yang, L., Z. Zhang, Y. Song, et al. 2023. "Diffusion Models: A Comprehensive Survey of Methods and Applications." *ACM Computing Surveys* 56, no. 4: 1–39.

Yarham, C., U. Boeniger, and F. Herrmann. 2006. "Curvelet-Based Ground Roll Removal." In the *SEG International Exposition and Annual Meeting*, SEG–2006. Society of Exploration Geophysicists.

Yilmaz, Ö 2001. *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists.

Yu, S., J. Ma, and W. Wang. 2019. "Deep Learning for Denoising." *Geophysics* 84, no. 6: V333–V350.

Yuqing, W., L. Wenkai, L. JinLin, Z. Meng, and M. YongKang. 2019. "Random Seismic Noise Attenuation Based on Data Augmentation and CNN." *Chinese Journal of Geophysics* 62, no. 1: 421–433.

Zhang, M., Y. Liu, and Y. Chen. 2019. "*Unsupervised Seismic Random Noise Attenuation Based on Deep Convolutional Neural Network*."IEEE Access 7: 179810–179822.

Zhang, Y., H. Lin, Y. Li, and H. Ma. 2019. "*A Patch Based Denoising Method Using Deep Convolutional Neural Network for Seismic Image*."IEEE Access 7: 156883–156894.

13652478, 2026, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1365-2478.70172 by Xian Jiaotong University, Wiley Online Library on [20/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License