

Zero-Shot Denoising for DAS-VSP Data Based on Conditional Diffusion Probabilistic Models

Dawei Liu¹, Zhenyu Wang, Xiaokai Wang², *Member, IEEE*, and Wenchao Chen¹

Abstract—Distributed acoustic sensing (DAS) systems offer a promising framework for advanced subsurface imaging and monitoring. Despite the great potential, the complex noise characteristics inherent in seismic data, such as environmental, mechanical, and instrumental disturbances, pose significant challenges to data fidelity and stability. Conventional noise suppression methods cannot be adequately adapted to dynamic seismic environments due to the need to design filters for different conditions. To overcome these limitations, we introduce the conditional denoising diffusion probabilistic model (C-DDPM), which exhibits strong a priori extraction capabilities and can better cope with seismic signal extraction under different noise combinations. In addition, we incorporate an adaptive FK conditioning approach into the diffusion process, allowing C-DDPM to better learn the data distribution. We also use asymmetric dilated convolution (ADConv) to effectively suppress noise. Our approach is rigorously tested on both synthetic and real-world seismic datasets, demonstrating satisfactory improvements in noise reduction and signal clarity. Comparative analyses with existing classical methods reveal that our framework not only achieves a higher peak signal-to-noise ratio (PSNR) but also reveals waveform details previously obscured by noise, outperforming existing methods in challenging geophysical scenarios.

Index Terms—Conditional denoising diffusion probability model (C-DDPM), distributed acoustic sensing (DAS), generative model, seismic denoising.

I. INTRODUCTION

DISTRIBUTED acoustic sensing (DAS) is an innovative and promising technology for seismic signal acquisition using optical fibers [1]. By injecting coherent light pulses into an optical fiber, DAS detects backscattering caused by impurities in the fiber. The amplitude and phase changes of these reflected signals are measured by photodiodes to acquire seismic data. This process effectively transforms the optical fiber into a dense array of vibration sensors, allowing real-time data collection over extensive areas with unprecedented spatial resolution. DAS offers significant advantages in complex terrains, reducing the need for traditional seismic

equipment [2], [3] and providing continuous monitoring of subsurface phenomena. Its low maintenance requirements and adaptability make DAS a promising replacement or complementary technology to existing seismic array systems [4], [5], [6]. Currently, DAS has been successfully applied in various fields, including signal acquisition in vertical seismic profiles (VSPs) [7], reservoir characterization [8], [9], hydraulic fracture monitoring [10], [11], [12], seismic detection [13], and seismic monitoring [14].

Despite its advantages, DAS is highly susceptible to noise interference from various sources. These include optical and electronic noise in the sensing equipment, inconsistencies in fiber quality and reflectivity, imperfections at the splices and connectors, ambient noise conditions along the fiber path, and inadequate coupling between the fiber and the ground [15]. As depicted in Fig. 1, such interference can significantly degrade the quality of seismic data, making it difficult to extract clean signals with a single denoising method [16], [17], [18]. Given the vast amounts of data generated by DAS, sophisticated denoising techniques are essential for meaningful signal extraction, making it a critical area of research in geophysical data processing.

Traditional methods possess unique advantages in removing noise from DAS data, particularly by leveraging hand-crafted prior knowledge behind noise generation mechanisms and the characteristics of the signals, achieving refined denoising effects and high-quality results. For example, Chen et al. [19] proposed a method for coupled noise suppression by constructing sparse dictionaries for both the signal and noise using continuous wavelet transform and discrete cosine transform, respectively. Lellouch et al. [20] applied median and low-pass filters to clean DAS-VSP data from the Frontier Observatory for Research in Geothermal Energy (FORGE), addressing high-frequency, high-amplitude, and horizontal noise. Chen et al. [21] introduced a comprehensive framework for cleaning the FORGE DAS-VSP dataset, integrating bandpass filters, structure-oriented median filters, and FK-domain dip filters. However, they face several limitations. First, designing filters requires a strong mathematics background and specialized expertise, which can pose a significant challenge, especially when complex mathematical tools are needed to achieve optimal results. Furthermore, selecting the appropriate parameters needs substantial prior knowledge to accurately distinguish between seismic signals and noise. These parameters are fixed and generally assumed to apply to stationary signals. However, in the highly variable environments of DAS systems, such conventional approaches may be effective for specific noise types but often struggle to manage the wide range of noise types simultaneously. Therefore, there is an urgent need for more advanced, tailored

Received 26 February 2025; revised 8 May 2025; accepted 25 May 2025. Date of publication 29 May 2025; date of current version 29 July 2025. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFA0716903, in part by the National Natural Science Foundation of China under Grant 42374135, and in part by the Fundamental Research Funds for the Central Universities under Grant xzy012023073. The work of Dawei Liu was supported by the industrial sponsors of the Signal Analysis and Imaging Group (SAIG) consortium at the University of Alberta. (Corresponding author: Wenchao Chen.)

Dawei Liu is with the School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China, and also with the Signal Analysis and Imaging Group, University of Alberta, Edmonton, AB T6G 2R3, Canada (e-mail: 409791715@qq.com).

Zhenyu Wang, Xiaokai Wang, and Wenchao Chen are with the School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: ZhenyuWang968@stu.xjtu.edu.cn; xkwang@xjtu.edu.cn; wencchen@xjtu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2025.3574981

1558-0644 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded from IEEE Xplore on October 19, 2025 at 08:13:40 UTC. Restrictions apply.

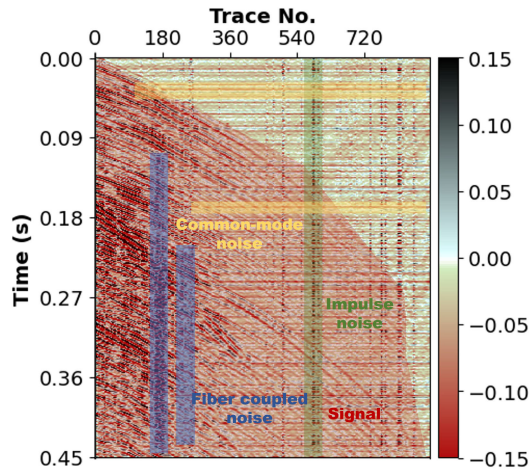


Fig. 1. Segment of real DAS-VSP data contaminated by diverse noise types. The red area indicates seismic signals, the blue areas represent coupled noise caused by cable oscillations, the yellow areas denote common-mode noise resulting from interrogator box vibrations or electromagnetic interference from electronic devices, and the green area shows impulse noise caused by transient faults in the measurement system or instability in optical pulses.

denoising techniques that can adapt to the complexities of DAS systems and handle all noise types collectively.

Recently, deep learning has gained significant attention in seismic signal processing for its efficiency in parallel computation and strong data-driven pattern recognition abilities. Supervised learning can effectively denoise seismic data by learning complex noise patterns directly from training data through end-to-end learning. This approach bypasses the intricate explicit priori design of traditional methods, offering a more flexible way to incorporate prior knowledge. As a result, supervised learning is widely applied to various seismic denoising tasks, including the removal of random noise [22], linear noise [23], ground roll [24], scattering noise [25], and multiples [26]. Other aspects include seismic signal interpolation [27], automatic structural interpretation with realistic folding and faulting features [28], and coherent wave field analysis for large-scale data-driven training [29]. Although fast and effective [30], [31], [32], supervised learning approaches rely on a large number of noisy-clean data pairs, which are unavailable for field data. Unsupervised deep learning methods are capable of learning directly from noisy data and can achieve satisfactory denoising results without clean labels [33], [34], [35]. However, unsupervised methods carry inherent limitations, including a heavy reliance on handcrafted heuristics or traditional assumptions, sensitivity to noise or hyperparameters, and limited generalizability to complex or heterogeneous datasets.

While deep learning methods for seismic signal denoising commonly focus on modeling the signal, we argue that the noise component deserves greater attention. Our work explores this under-emphasized noise perspective, seeking to develop a denoising strategy that effectively removes complex noise while preserving the integrity of the underlying signal. During the acquisition of seismic signals by the DAS-VSP system, when there is no seismic excitation, the DAS constantly receives vibration signals from the environment, and these noises are lower in intensity than the seismic signals. The percentage of background noise is affected by the environment

of the acquisition site (e.g., traffic, industrial activity, climate change). Under certain environmental conditions, the percentage of background noise may be higher, especially if the source is sparse or the seismic wave travels a long distance. Therefore, most of the data in the existing DAS-VSP dataset are noise that does not contain seismic signals. These abundant noise samples can be effectively leveraged to train neural networks, potentially achieving supervised-like training outcomes without explicit labels.

The denoising diffusion probabilistic model (DDPM) [36] is currently an excellent deep generative model renowned for its effectiveness in image synthesis and image editing and has a very strong ability to learn the signal priors. Conditional denoising diffusion probability models (C-DDPMs), as explored by [37], extend DDPM by conditioning the diffusion process on noisy data through training a conditional denoiser. This allows for more targeted denoising, enabling the model to better distinguish between noise and signals, and also makes C-DDPM receive much attention in seismic signal-processing work [38], [39]. Since it is free of modeling explicit degradation, conditional diffusion models have broader applications in seismic processing, including the removal of ground roll [40] and multiples [41], interpolation [42], and super-resolution [43]. This flexibility makes it a potentially powerful tool for complex denoising challenges in DAS-VSP data.

We propose to use C-DDPM to denoise DAS-VSP data. This is done by training a denoiser obtained by adding synthetic DAS-VSP data with real acquired noise as a training set to recover the real DAS-VSP signal. To better adapt to the nuances of seismic signal processing, we have made several modifications to the original C-DDPM, including the noise schedule, the diffusion process, and the network structure, so that the diffusion model can satisfy the requirement of applying the model trained on synthetic data to real data. Experimental results show that the modified C-DDPM successfully learns the key features of seismic signals and effectively separates the noise. The field data further validate the practicality and strong generalizability of our method. Our contributions are as follows.

- 1) We treat the synthetic data with added noise as additional a priori information and use conditional diffusion probabilistic modeling to mitigate the generative bias inherent in diffusion models. With this approach, we can capture multimodal distributions in the high-dimensional space of DAS-VSP data and effectively suppress noise from various distributions.
- 2) We employ adaptive FK conditioning diffusion to incorporate the idea of filtering into the diffusion process. It allows the condition to gradually guide the generation direction of the model from low to high frequencies as the diffusion process changes with the time step t , thus maximizing the preservation of the DAS-VSP data. We also replace the original linear noise table with a cosine noise table. This can effectively control the fluctuation of the peak signal-to-noise ratio (PSNR), reduce the information loss, and ensure a more stable quality of data denoising.
- 3) We introduce asymmetric dilated convolution (ADConv) into the convolutional layer of the diffusion model U-Net

network. ADConv is inspired by asymmetric convolution and dilated convolution and combines the two by setting special correlation distances to achieve more powerful horizontal and vertical noise removal.

II. METHOD

A. Training Dataset Construction With Real Acquired DAS Noise

Most deep learning methods for seismic signal denoising tend to focus on the signal as the main prior, overlooking the potential benefits of leveraging noise itself. In reality, noise is not merely an obstacle to be removed but a widely available resource that can significantly enhance denoising performance. By reorienting our view to treat noise as an integral component, we seek to remove its complex patterns while preserving the core characteristics of the seismic signal.

In DAS-VSP data acquisition, background noise is continuously recorded even without active seismic sources. The noise primarily originates from the perturbation of detecting equipment and wellbore-coupled vibrations and can be broadly categorized into the following three types based on their different patterns. Vertical noise (the green area in Fig. 1) is mainly impulse noise caused by transient faults in the measurement system or instability in optical pulses. Horizontal noise (the yellow areas in Fig. 1), which is mainly common-mode noise, may originate from vibrations of the DAS interrogator box, such as fan operation or electrical leakage of the interrogator or nearby electronic components. These sources can induce simultaneous vibration signals across all channels, resulting in horizontal strip-like patterns. Fiber-coupled noise (the blue areas in Fig. 1) manifests as spring-wave patterns and is caused by oscillations of loosely coupled fibers bouncing against the borehole wall. Although typically less intense than seismic signals, this noise often dominates the dataset, especially when seismic sources are sparse or wave propagation distances are extended. As a result, vast portions of DAS-VSP recordings are comprised solely of nonseismic noise. This abundance of real-world noise becomes an invaluable training resource for C-DDPM, which can exploit unlabeled noise data to learn robust noise distributions.

By conditioning on noisy samples, the C-DDPM effectively separates the signal from substantial noise content without heavily relying on a large number of labeled datasets. The diffusion process refines an understanding of diverse noise profiles, while the conditional framework ensures that the essential seismic features remain intact. Consequently, this approach transforms noise from a hindrance into a driving force: the more comprehensive and varied the noise dataset is, the more adeptly the network learns to suppress it in new, previously unseen data. This shift in perspective underscores noise's role beyond conventional "clutter," demonstrating how its rich variability and widespread availability can be harnessed to achieve improved seismic signal denoising.

We propose to split the real DAS-VSP dataset into two parts, that is, the part containing seismic signals and the part of pure field noise. The noise properties of these two parts are the same, differing only in whether they contain seismic signals or not. As shown in Fig. 2, we take synthetic seismic signals as $\mathbf{x}_0$ and add real noise to them as $\mathbf{y}$, that is, the condition,

to form the training dataset. Once trained, the optimal network θ^* is directly used for field data denoising. Finally, we apply the a priori information $q(\mathbf{x}_0)$ and the network θ^* learned by the model directly to the denoising of DAS-VSP data containing seismic signals, that is, we use the DAS-VSP data containing seismic signals as $\mathbf{y}^\dagger$ to guide the model to generate denoising results.

B. Conditional DDPM

C-DDPM can learn the prior distribution from the condition and then decode it to generate images through a reverse process. The forward and reverse processes are shown in Fig. 2(a). The forward process of C-DDPM is a deterministic Markov chain. Let the input be $\mathbf{x}_0 \sim q(\mathbf{x}_0)$. In the forward process, predetermined Gaussian noise is gradually added to $\mathbf{x}_0$ to perturb the data distribution, and the addition of noise is constrained by the time step t , $t = 0, \dots, T$. Given $\mathbf{x}_0$ and latent variables $\mathbf{x}_1, \dots, \mathbf{x}_T$ from the same sample space, the diffusion process can be defined as

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0, \mathbf{y}) = \prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}) \quad (1)$$

where $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y})$ represents the mapping from step $t-1$ to step t in the forward process and $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$. That is, $\mathbf{x}_t$ is obtained by adding Gaussian noise to $\mathbf{x}_{t-1}$

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}) = \mathcal{N}\left(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}\right) \quad (2)$$

where $\beta_t \in (0, 1)$ is a predefined variance schedule that increases with time step t . Thus, we can get the forward process $\mathbf{x}_t$ directly from the $\mathbf{x}_0$ iteration

$$q(\mathbf{x}_t|\mathbf{x}_0, \mathbf{y}) = \mathcal{N}\left(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1 - \bar{\alpha}_t)\mathbf{I}\right). \quad (3)$$

A more explicit form is given by

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t}\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}) \quad (4)$$

where $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$. At timestep t , $\mathbf{x}_t$ converges to a prior distribution, that is, a standard normal distribution. Thus, the forward process is a diffusion process that gradually converts a real-world image to a Gaussian noise image.

The reverse process is also defined on a Markov chain, which converts pure noise $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ to the data distribution $\mathbf{x}_0 \sim q(\mathbf{x}_0)$

$$p_\theta(\mathbf{x}_{0:T}, \mathbf{y}) = p(\mathbf{x}_T, \mathbf{y}) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y}) \quad (5)$$

where $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y})$ is the Gaussian transition

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y}) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, \mathbf{y}, t), \Sigma_\theta(\mathbf{x}_t, \mathbf{y}, t)) \quad (6)$$

that has learned the mean $\mu_\theta(\mathbf{x}_t|t)$ and fixed variance $\Sigma_\theta(\mathbf{x}_t|t) = \sigma_t$ [37]

$$\mu_\theta(\mathbf{x}_t, t) = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)\right). \quad (7)$$

Then, the optimization function of the network is

$$L = \mathbb{E}_{\mathbf{x}_t, \mathbf{x}_t, \boldsymbol{\epsilon}_t} [\|\boldsymbol{\epsilon}_t - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)\|^2]. \quad (8)$$

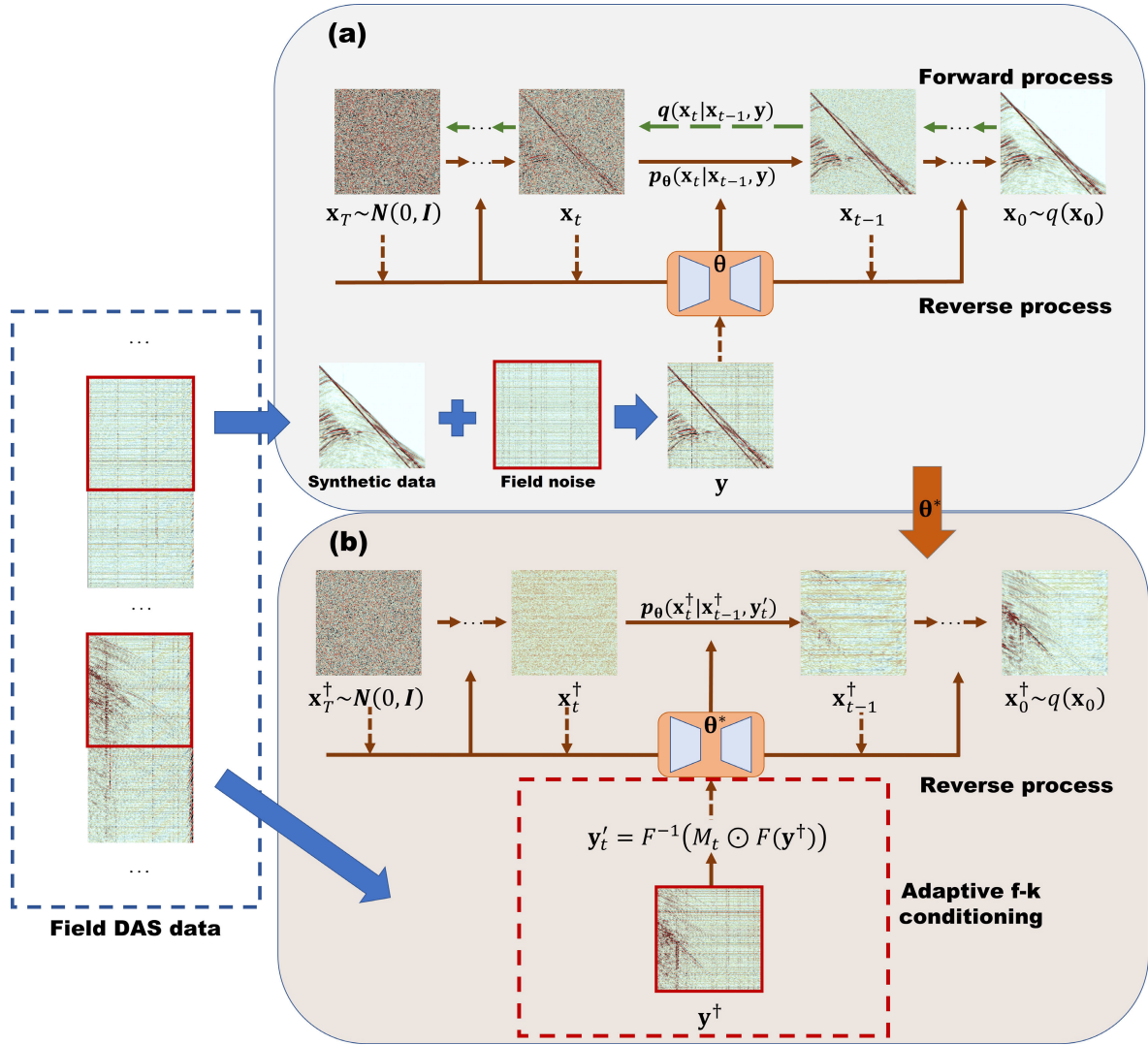


Fig. 2. Overall flow of the proposed algorithm. (a) Training process. The basic C-DDPM is applied to the training of synthetic data. The dashed arrows represent the forward process, while the solid arrows indicate the reverse process. Forward process is a predefined Gaussian transition $q(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y})$. In the reverse process, the learned Gaussian transition $p_{\theta}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y})$ utilizes $\mathbf{y}$ to guide the generation. (b) Denoising process. The trained model is directly applied to the field data denoising. The use of AFKCDM directs the network to learn the field data from low to high frequencies through a low-pass filtering process conditional on the time step t for $\mathbf{y}'$, which is used to obtain better denoising results. And the orange arrows between (a) and (b) indicate that the optimal network θ^* trained by (a) is used directly for (b).

And we can gradually generate $\hat{\mathbf{x}}_0$ during training by

$$\hat{\mathbf{x}}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(\mathbf{x}_t, \mathbf{y}, t) \right) + \sigma_t \mathbf{z} \quad (9)$$

which shows

$$\hat{\mathbf{x}}_0 = \sqrt{\frac{1}{\bar{\alpha}_t}} \mathbf{x}_t - \frac{\sqrt{1 - \bar{\alpha}_t}}{\bar{\alpha}_t} \epsilon_{\theta}(\mathbf{x}_t, \mathbf{y}, t). \quad (10)$$

The training and sampling processes for the synthetic data processing of our method are described in Algorithms 1 and 2. We train the network θ using the label and the field data as $\mathbf{x}_0$ and $\mathbf{y}$. The trained network θ is used at each step of the sampling process to predict the noise that should be removed, and thus samples the denoised results after the whole inverse process. The sampling process for real data processing of our method is described in Algorithm 3. A more detailed discussion is provided in Section II-C.

Algorithm 1 Training of Synthetic Data Processing

```

1 repeat
2    $\mathbf{x}_0, \mathbf{y} \sim q(\mathbf{x}_0, \mathbf{y})$ 
3    $t \sim \text{Uniform}(\{1, \dots, T\})$ 
4    $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ 
5   Take a gradient descent step on:
       $\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(\sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \mathbf{y}, t)\|^2$ 
6 until converged

```

C. Adaptive FK Conditioning Diffusion Model

To improve the noise suppression performance of C-DDPM, better capture the characteristics of seismic signals, and reduce signal leakage, we introduce an adaptive FK conditioning diffusion model (AFKCDM). This approach enhances the conditioning mechanism of C-DDPM by leveraging the FK domain properties of seismic signals, which are critical for

Algorithm 2 Sampling of Synthetic Data Processing

```

1  $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ 
2 for  $t = T, \dots, 1$  do
3  $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$  if  $t > 1$ , else  $\mathbf{z} = 0$ 
4 Take a gradient descent step on:
    $\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(\sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1 - \alpha_t}\epsilon, \mathbf{y}, t)\|^2$ 
5  $\hat{\mathbf{x}}_{t-1} = \frac{1}{\sqrt{\alpha_t}}(\hat{\mathbf{x}}_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}}\epsilon_{\theta}(\hat{\mathbf{x}}_t, \mathbf{y}, t)) + \sigma_t \mathbf{z}$ 
6 until converged

```

Algorithm 3 Sampling of Real Data Processing

```

1  $\mathbf{x}_T^{\dagger} \sim \mathcal{N}(0, \mathbf{I})$ 
2 for  $t = T, \dots, 1$  do
3  $\mathbf{z}^{\dagger} \sim \mathcal{N}(0, \mathbf{I})$  if  $t > 1$ , else  $\mathbf{z}^{\dagger} = 0$ 
4 Transform  $\mathbf{y}^{\dagger}$  into the FK domain:
    $\mathbf{Y}_f = \mathcal{F}(\mathbf{y}^{\dagger})$ 
5 Apply timestep-dependent FK mask  $\mathbf{M}_t$ :
    $\mathbf{Y}'_f = \mathbf{M}_t \odot \mathbf{Y}_f$ 
6 Perform inverse Fourier transform to obtain the filtered
   condition:
    $\mathbf{y}' = \mathcal{F}^{-1}(\mathbf{Y}'_f)$ 
7 Take gradient descent step on:

```

$$\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(\sqrt{\alpha_t}\mathbf{x}_0^{\dagger} + \sqrt{1 - \alpha_t}\epsilon, \mathbf{y}', t)\|^2$$

```

8 Update  $\mathbf{x}_{t-1}^{\dagger}$  using the adaptive FK conditioning
   condition:

```

$$\hat{\mathbf{x}}_{t-1}^{\dagger} = \frac{1}{\sqrt{\alpha_t}}\left(\hat{\mathbf{x}}_t^{\dagger} - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}}\epsilon_{\theta}(\hat{\mathbf{x}}_t^{\dagger}, \mathbf{y}', t)\right) + \sigma_t \mathbf{z}^{\dagger}$$

```

9 until converged

```

effective denoising in geophysical applications. AFKCDM modifies the reverse diffusion process by incorporating adaptive constraints, thereby ensuring that the retained signal components match the dominant seismic frequencies.

An illustration can be seen in Fig. 2(b). At each step t of the reverse process, AFKCDM applies an adaptive FK conditioning transformation to $\mathbf{y}^{\dagger}$ and integrates this information into the denoising procedure. The process begins by transforming $\mathbf{y}^{\dagger}$ into the FK domain using a Fourier transform

$$\mathbf{Y}_f = \mathcal{F}(\mathbf{y}^{\dagger}) \quad (11)$$

where $\mathcal{F}(\cdot)$ denotes the Fourier transform and $\mathbf{Y}_f$ represents the FK spectrum of $\mathbf{y}^{\dagger}$.

A timestep-dependent FK mask $\mathbf{M}_t$ is then applied to $\mathbf{Y}_f$

$$\mathbf{Y}'_f = \mathbf{M}_t \odot \mathbf{Y}_f \quad (12)$$

where $\odot$ denotes element-wise multiplication. The mask $\mathbf{M}_t$ is designed to attenuate higher frequencies as the diffusion step t

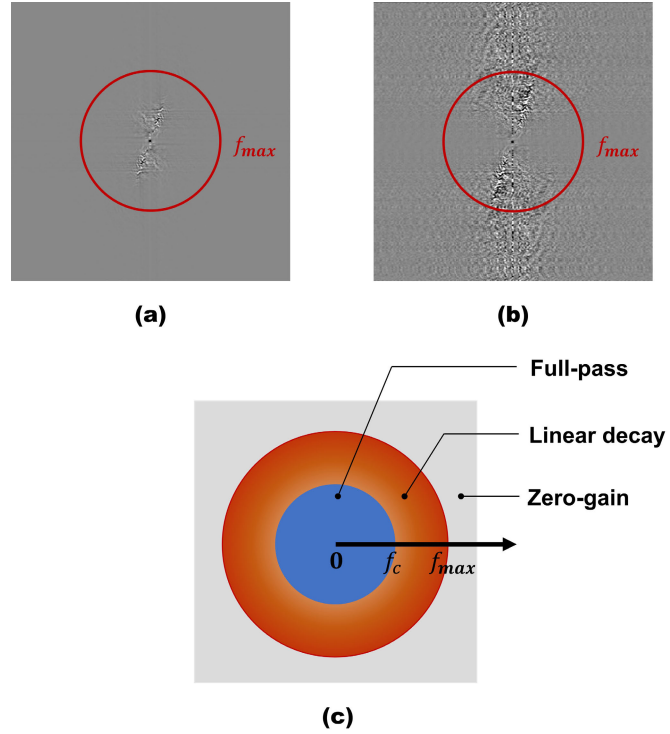


Fig. 3. Schematic illustration of the FK masking mechanism. (a) Clean signal denoised by the integrated framework. (b) Noisy data. (c) FK mask showing smooth transition from pass band to stop band, mitigating boundary artifacts.

increases progressively. Mathematically, $\mathbf{M}_t$ can be expressed as

$$\mathbf{M}_t(f) = \begin{cases} 1, & f \leq f_c(t) \\ \text{linear decay}, & f_c(t) < f \leq f_{\max}(t) \\ 0, & f > f_{\max}(t). \end{cases} \quad (13)$$

The FK mask $\mathbf{M}_t$ is controlled by two parameters. First, the outer radius, $f_{\max}$, defines the filter's cutoff boundary, and components beyond this radius are attenuated to zero. Crucially, $f_{\max}$ increases as the timestep t decreases, progressively expanding the filter's coverage to eventually encompass the full frequency spectrum relevant to the observed seismic data. As illustrated in Fig. 3(a) and (b), the initial $f_{\max}$ (at the start of the reverse process) is set based on the FK spectrum of data preliminarily denoised using an integration framework [21]. This ensures a more informed starting point for the C-DDPM generation.

And then to mitigate potential boundary artifacts arising from abrupt spectral cutoffs, we introduce an inner radius f_c , typically set to $(f_{\max}/2)$. In this radius f_c , the filter gain is unity (i.e., components pass unaltered). As shown in Fig. 3(c), a smooth transition between the full-pass region (inside f_c) and the zero-gain region (outside $f_{\max}$) is achieved by linear decay. This smooth transition minimizes artifacts and enhances the stability of the denoising process.

After masking, an inverse Fourier transform is applied to reconstruct the filtered signal in the time domain

$$\mathbf{y}' = \mathcal{F}^{-1}(\mathbf{Y}'_f) \quad (14)$$

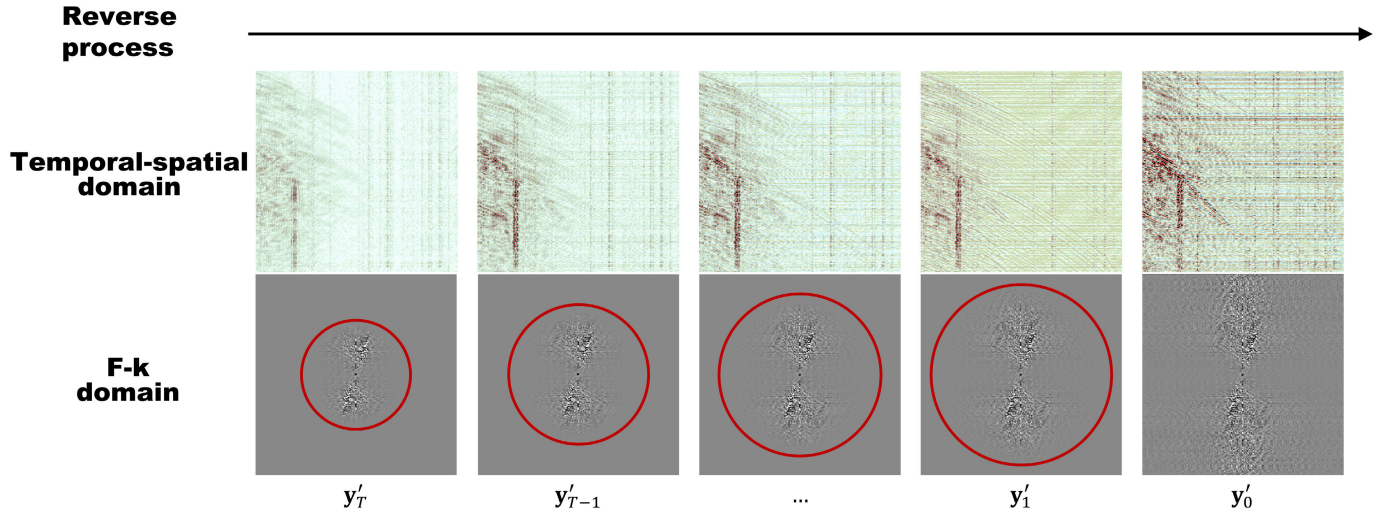


Fig. 4. Condition of the reverse process of AFKCDM. An FK domain mask, indicated by red circles, is applied to $\mathbf{y}^\dagger$ during the reverse process. The mask is initially limited to the low-frequency, low-wavenumber portion of the FK spectrum, and subsequently expands as t decreases. This strategy enables the bootstrap diffusion model to progressively learn the distribution of $\mathbf{y}'$, starting from low frequencies and incrementally incorporating higher frequencies.

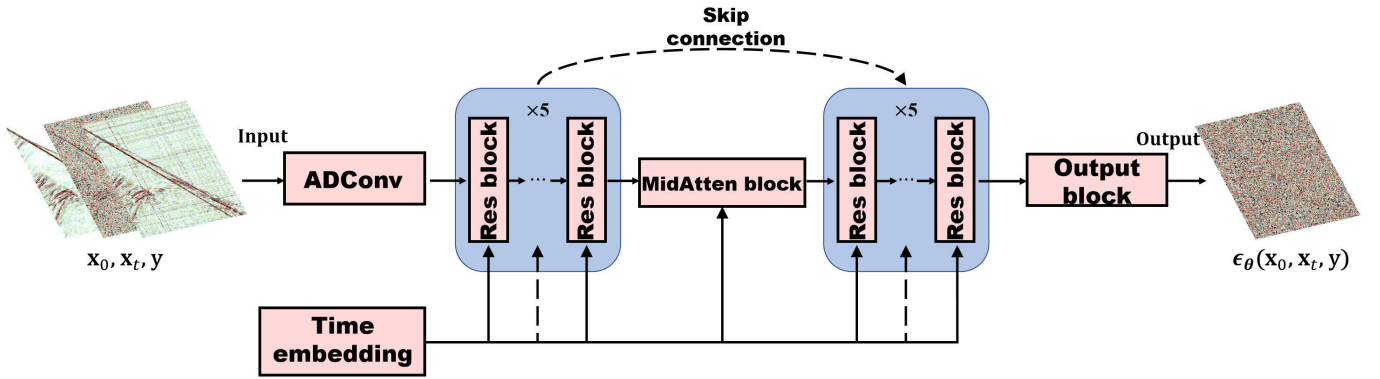


Fig. 5. Illustration of the U-Net architecture used as a backbone of the C-DDPM.

where $\mathcal{F}^{-1}(\cdot)$ is the inverse Fourier transform. The reconstructed signal $\mathbf{y}'$ serves as the updated condition for the reverse diffusion process.

The reverse diffusion step in C-DDPM is updated to incorporate the adaptive FK condition. Let $\mathbf{x}_t^\dagger$ represent the noisy signal at step t . The denoised signal $\mathbf{x}_{t-1}^\dagger$ is computed as

$$p_\theta(\mathbf{x}_{t-1}^\dagger | \mathbf{x}_t^\dagger, \mathbf{y}') \approx p_\theta(\mathbf{x}_{t-1}^\dagger | \mathbf{x}_t^\dagger, \mathbf{y}' = \mathcal{F}^{-1}(\mathbf{M}_t \odot \mathcal{F}(\mathbf{y}^\dagger))). \quad (15)$$

This modification ensures that the reverse process is guided by the FK domain properties of the conditional signal $\mathbf{y}$, effectively aligning the model's output with the dominant seismic signal frequencies.

As shown in Fig. 4, the incorporation of AFKCDM significantly enhances the denoising ability of C-DDPM by ensuring that only physically meaningful frequencies are retained throughout the reverse process. This not only improves the model's robustness to noise but also reduces the risk of seismic signal leakage.

D. Network Architecture

Most diffusion models are based on the U-Net architecture, which is distinguished by its U-shaped combination of an encoder and a decoder. The encoder shown on the left side of Fig. 5, typically includes max-pooling layers and double convolutions, which reduce the image dimensions while increasing the number of feature channels. On the right side of Fig. 5 is the decoder, which shares a similar structure but incorporates upsampling operations to restore the features to the original image size. Additionally, skip connections integrate encoder features into the decoder to retain spatial information.

1) *Architecture*: We have implemented Saharia et al. [44]'s U-Net architecture with minor modifications. As depicted in Fig. 5, our version of the U-Net features a time embedding block, five Res blocks each in the encoder and the decoder, and a centrally located MidAtten block. Fig. 6 shows the details of these modules. The inclusion of the MidAtten block, which utilizes self-attention, greatly enhances the U-Net's ability to identify relationships between pixels. This mechanism uses three distinct weight matrices: the query matrix $\mathbf{Q}$, the key matrix $\mathbf{K}$, and the value matrix $\mathbf{V}$, each computed independently of the convolutional feature maps by the respective operators $\mathbf{W}_Q$, $\mathbf{W}_K$, and $\mathbf{W}_V$. This self-attention

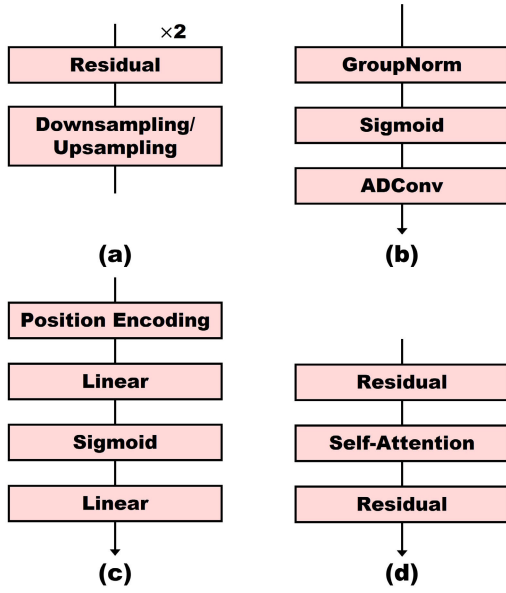


Fig. 6. Detailed internal structure in the U-Net network. (a) Res block. (b) Output block. (c) Time embedding block. (d) MidAtten block.

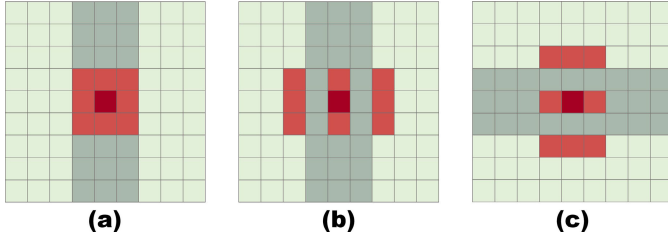


Fig. 7. Illustration of ADConv. In the feature map, the dark green grid points represent vertical or horizontal noise, the red grid points indicate the convolution kernel and the dark red grid point denotes the center of the kernel. (a) Standard convolution kernel. (b) 3×5 ADConv with horizontal dilation. (c) 5×3 ADConv with vertical dilation.

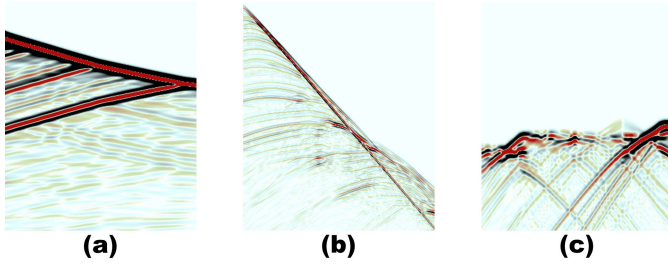


Fig. 8. Synthetic data labels for the training dataset. VSP data created using the Marmousi velocity model in three excitation modes. (a) Surface excitation, (b) interwell excitation, and (c) hydraulic fracturing.

aids in the processing of complex pixel relationships not typically addressed by standard convolutional layers. Given

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (16)$$

where $(d_k)^{1/2}$ is the dimension of the query matrix $\mathbf{Q}$ and the key matrix $\mathbf{K}$.

In the architecture of the Res blocks, each combines residual modules with standard downsampling or upsampling modules, with an exception made for the bottom layer. The inclusion of residual modules specifically targets the degradation issues that are frequently encountered in deep convolutional networks. In both the downsampling and upsampling phases,

two residual modules are systematically utilized to ensure the integrity of information through the depth of the network.

This configuration not only maintains feature richness but also stabilizes learning by mitigating the vanishing gradient problem, common in deeper network structures.

2) *Asymmetric-Dilated Convolution*: Dilated convolution expands the receptive field of convolutional kernels by introducing gaps between the elements of the kernel without increasing the number of parameters or computational complexity. This enables the network to capture a broader context and multiscale features, facilitating the identification and suppression of structural noise patterns in images, such as horizontal and vertical artifacts. This multiscale approach allows the model to adaptively focus on noise of different frequencies and orientations. Larger dilation rates effectively capture wide horizontal or vertical streaks, while smaller dilation rates target finer noise patterns. By systematically adjusting the dilation rate, the network can hierarchically process and eliminate noise at different scales, ensuring complete noise suppression while preserving the essential structures of the image.

Asymmetric convolution employs elongated kernels in either the horizontal or vertical direction (e.g., $1 \times N$ or $N \times 1$) to specifically address directional noise in images. Unlike traditional symmetric kernels, asymmetric convolutions can more effectively suppress noise oriented in specific directions, such as horizontal scan lines or vertical grid artifacts, while preserving the critical structural details of the image. This method enhances computational efficiency by reducing the number of parameters and computational overhead compared to larger symmetric kernels that achieve similar directional sensitivity, enabling efficient operation on resource-constrained devices. Additionally, asymmetric convolution improves the network's ability to capture directional features, aiding in the distinction between noise and genuine image content. This results in clearer and more accurate denoising outcomes.

ADConv can be a good combination of asymmetric convolution and dilated convolution to eliminate vertical and horizontal noise [45]. For noisy data in the vertical direction, we extend the original 3×3 convolution kernel into a 3×5 or 3×7 ADConv, while the vertical direction remains unchanged, as shown in Fig. 7(b). After testing [45], compared with the 3×3 kernel, the receptive field of ADConv in the horizontal direction is increased dramatically, while the receptive field in the vertical direction remains unchanged. For horizontally noisy data, as shown in Fig. 7(c), we expand the convolution kernel size to 5×3 or 7×3 and set the expansion factor in the vertical direction accordingly. With this design, ADConv can filter out the effect of correlated noise to a greater extent and exclude the noise information from the receptive field of the masked point. By combining an asymmetric kernel with the dilation strategy, ADConv effectively expands the model's receptive field in specific directions without significantly increasing the number of parameters. This approach enhances the network's ability to recognize and process spatial correlations over long distances, making it particularly effective in suppressing horizontal and vertical noise in images.

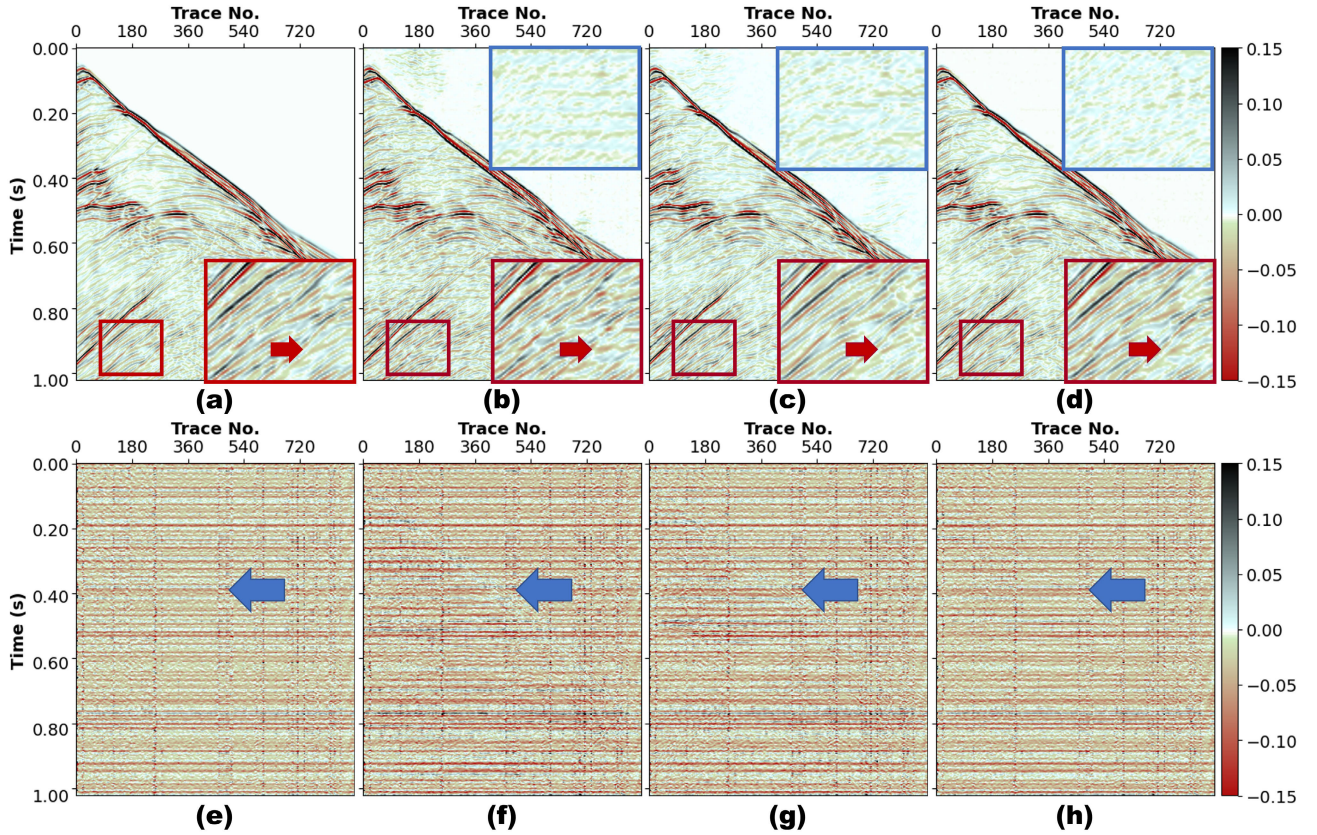


Fig. 9. Comparison of supervised denoising results and residuals of different models on synthetic VSP data with added real noise. (a) Ground truth. (b) DnCNN denoising results. (c) U-Net denoising results. (d) C-DDPM denoising results. (e) The added real noise. (f)–(h) Residuals of the three different networks. The red rectangular boxes represent complex signals. The blue rectangular boxes represent the difference between the individual model results and the label. The red arrows indicate the recovery results for different networks. Blue arrows identify signal leakage.

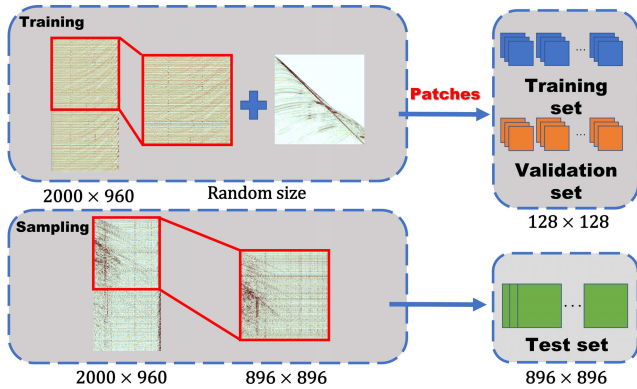


Fig. 10. To ensure robust training and evaluation, the dataset was partitioned into training, validation, and test sets using an 8:1:1 ratio. Our model was trained and evaluated on patches extracted from the training and validation sets, and its performance was then evaluated on the complete, unsegmented data in the test set.

III. EXPERIMENT

We use synthetic data combined with real noise for training and use the trained model for real data denoising. The implementation process is shown in Fig. 2.

In this section, we first validate the effectiveness of C-DDPM using supervised denoising of synthetic data in comparison to other network structures. Second, we perform ablation experiments to demonstrate that the diffusion model

made possible under our innovations is capable of zero-sample unsupervised denoising of real DAS-VSP data.

A. Synthetic Seismic Data

We first tested the proposed C-DDPM using synthetic seismic data constructed based on the Marmousi velocity model. The model is detailed with various geological features and has dimensions of 2301×751 , consisting of 896 traces, each with 896 sample points, and a spatial resolution of 4 m per point. Our simulation setup includes 115 shots, each shot having one source and 384 receivers. The sources are placed every 80 m starting from 40 m across, at a depth of 8 m. Receivers are evenly spaced at 24 m, beginning right at the model's edge, and are also placed at a depth of 8 m.

We utilize a Ricker wavelet with a dominant frequency of 25 Hz as our seismic source. The total recording time for our simulations is 3 s, with a temporal sampling interval of 4 ms. Forward modeling is performed using the Deep-wave package [46], which treats both horizontal and vertical dimensions equally, allowing for precise simulation under set perfectly matched layer (PML) frequency conditions to minimize boundary reflections.

We create VSP data for each of the three excitation modes: surface excitation, interwell excitation, and hydraulic fracturing to form the training dataset labels, which are shown in Fig. 8. The reason for taking multiple excitation modes is to construct a more comprehensive diffusion prior. We add

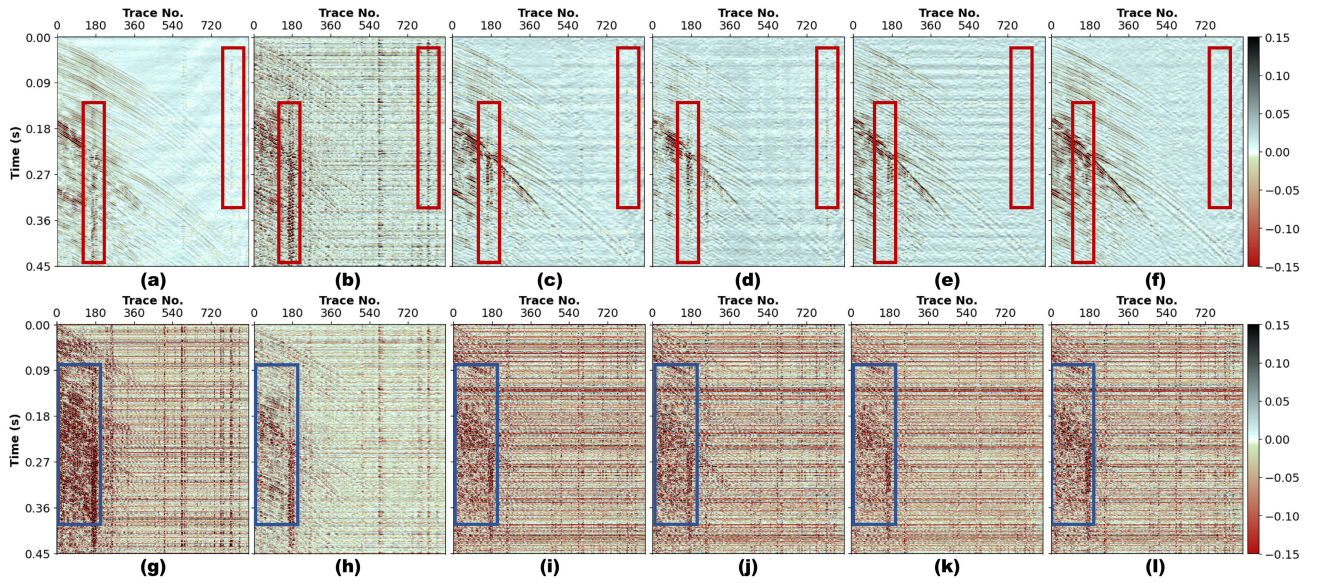


Fig. 11. Field DAS-VSP data denoising results and residuals. (a) Filter framework denoising result. (b) DnCNN denoising result. (c) C-DDPM denoising result. (d) C-DDPM + ADConv denoising result. (e) AFKCDM denoising result. (f) Denoising result of our method. (g)–(l) Residuals of the six denoising methods. The red rectangular boxes in (a)–(f) highlight noise remnants, while the blue rectangular boxes in (g)–(l) indicate signal leakage.

TABLE I

PSNR AND SSIM VALUES FOR DIFFERENT DENOISING NETWORKS: DnCNN, U-Net, AND C-DDPM UNDER VARYING NOISE LEVELS

Noise level	DnCNN		U-Net		C-DDPM	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.05	36.12	0.98	39.45	0.92	42.81	0.98
0.10	36.29	0.94	38.11	0.86	39.41	0.95
0.50	34.47	0.90	35.02	0.83	36.39	0.92
Real noise	28.95	0.86	33.78	0.92	39.08	0.96

noise to the synthetic dataset labels to constitute training data pairs for model training including artificial noise and real noise. Artificial noise is synthetic coupled noise, random noise, vertical noise, and horizontal noise superimposed on each other, and the noise variance is set to 0.05, 0.1, and 0.5. The real noise comes from the same FORGE as the real signal to be denoised, where the part that does not contain seismic signals. We explore a numerical comparison of the denoising effects of different module actions. The performance of each configuration in processing noise-enhanced seismic data scenarios is compared quantitatively based on the improvements in the PSNR and SSIM. The results summarized in Table I reveal the different effects of the different modules incorporated into the diffusion model. It can be seen that our proposed combination approach achieves the highest PSNR and SSIM.

The dataset is divided into training, testing, and validation sets with a distribution ratio of 8:1:1. The images are segmented into 128×128 patches with 50% overlap, generating a total of 6714 training patches. For validation, the original images remain unsegmented to preserve completeness. We apply the classical denoising convolutional network DnCNN as well as U-Net to the synthetic seismic signals as a comparison of our proposed C-DDPM method and the denoising results are shown in Fig. 9, respectively.

Fig. 9(a) denotes the ground truth. Fig. 9(b)–(d) present the denoising results for different methods. Fig. 9(e) shows the

added real noise. Fig. 9(f)–(h) are the residuals of the three denoising results. In the denoising results, the red rectangular boxes highlight complex seismic signals, and the red arrows indicate discrepancies between the results obtained by different methods and the true values. In the DnCNN denoising results, the complex signals are not correctly restored. In the U-Net denoising results, the complex signals show significant breaks, whereas the C-DDPM denoising results are the closest to the true values among the three methods. Additionally, in the upper blank areas of the denoising results, the remaining noise from both DnCNN and U-Net can still be observed. The blue rectangular boxes indicate the differences between the denoising results and the true values. The differences in all three methods show that C-DDPM's residuals are the smallest. The blue arrows in the residuals point to signal leakage. In the DnCNN and U-Net, the signal contours are still discernible, while in C-DDPM, signal leakage is almost undetectable.

B. Field Seismic Data

The real dataset originates from the Stage 2C hydraulic fracturing production enhancement at the FORGE site in Utah. The DAS-VSP data were collected from fiber-optic cables installed in monitoring wells 78-32 at the FORGE geothermal site. This well, drilled to a depth of 3274.78 feet, extended approximately 1200 feet into the granite basement rock. The Silixa iDAS v3 interrogator was employed to measure the strain rate as DAS-VSP data. During the initial enhancement of the FORGE site's enhanced geothermal system (test well 58-32) from late April to early May 2019, the fiber-optic cables continuously recorded data for 10.5 days. All DAS-VSP datasets are freely downloadable from the University of Utah High-Performance Computing Center. In this dataset, the gauge length is 10 m and the channel spacing is 1 m. The entire continuous recording is segmented into multiple SEG-Y files, with each segment lasting 15 s, at a sampling interval of

0.5 ms, each file contains 30 000 samples, and the number of channels ranges from 980 to 1088.

We segment the entire DAS-VSP dataset and extract 112 samples. We selectively download the SEG-Y files containing detected seismic events and segment them according to the P -wave arrival sample numbers. The first few and the last few defective channels are removed, leaving 896 seismic traces per segment. Then, to improve processing speed, we extract the portion rich in seismic signals, meaning that we only use 896 sampling points along the time axis (resulting in data of size 896×896). These data are divided into two groups: one containing seismic signals and the other consisting of ambient noise without seismic signals. The ambient noise is then added to the synthetic data to form data pairs for network training. This approach allows the model to learn more realistic noise, thereby better distinguishing between noise and signal and achieving improved denoising performance. The data containing seismic signals is used as test data, with the trained model directly applied for denoising to assess its effectiveness. The construction process of the dataset is shown in Fig. 10.

We test the proposed method on the test set. We compare different combinations of our methods. Comparisons are also made with classical conventional methods. The denoising results of different methods are shown in Fig. 11(a)–(f) and Fig. 11(g)–(l) are the residuals of the denoising results of these methods. The red rectangular boxes in Fig. 11(a)–(f) highlight noise remnants, while the blue rectangular boxes in Fig. 11(g)–(l) indicate signal leakage. Conventional methods such as filters [21] are effective in suppressing noise but cause signal leakage. A single C-DDPM cannot effectively handle the balance between strongly coupled noise and seismic signals, resulting in some of the strong noise being misclassified as signal retention and some of the signal components being removed. C-DDPM with ADConv can significantly suppress strongly coupled noise and horizontal noise, but expanding the receptive field also leads to signal loss. In the result by AFKCDM, the seismic signal is enhanced, but at the cost of heavier horizontal noise residues. Our proposed combination method has the best denoising effect among the four methods, which suppresses the strong noise and maximizes the preservation of the seismic signal.

IV. DISCUSSION

A. Non-Gaussian Forward Processes for Targeted Denoising

DDPMs predominantly use Gaussian noise in the forward process, giving them strong generative capabilities, but their performance in removing other types of noise (e.g., speckle, and structural) is still not as robust as it could be. This limitation is mainly because the models themselves are biased toward Gaussian noise, which may not be able to adequately capture the features of non-Gaussian noise types. Therefore, removing these other types of noise poses a significant challenge, resulting in a degradation of denoising performance.

One potential way to improve the efficiency of DDPMs in dealing with a wider range of noise types is to incorporate these different noise patterns into the diffusion process itself. By integrating non-Gaussian noise perturbations during

the training phase, the model can be specialized to identify and efficiently remove different noise features. For example, incorporating synthetic pretzel-like noise, speckle noise, or correlated noise into the training data allows the model to learn more robust denoising strategies that are not solely dependent on Gaussian noise assumptions. This approach has the potential to enhance the generalization ability of diffusion models, making them more versatile in real-world applications where noise types are diverse and unpredictable.

Further research is needed to explore the best way to integrate these non-Gaussian noise perturbations into a diffusion framework. This may involve experimenting with different noise addition strategies, evaluating their impact on model performance, and fine-tuning the balance between various noise types during training. By expanding the noise range of DDPMs, we can develop more comprehensive denoising solutions that consistently perform well in a wider range of noisy scenarios.

B. Training Acceleration

A significant limitation of DDPMs, including C-DDPM, is that their sampling process requires a significant amount of time. This difficulty is addressed by employing denoising diffusion implicit modeling (DDIM), which accelerates the sampling process by significantly reducing the number of time steps while maintaining competitive generation quality. Unlike unconditional DDPM, conditions in C-DDPM orient the generation process to a specific reference, thus trading diversity for improved relevance and quality of the conditional output. In our implementation, diversity is significantly reduced to ensure that the generated denoised images closely match the DAS-VSP data, with most remaining diversity introduced by random noise z at each sampling step.

To facilitate diverse outputs for other seismic processing tasks, a linear combination of conditional and unconditional noise estimates can be introduced as outlined in (17)

$$\hat{\mathbf{x}}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\hat{\mathbf{x}}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} (\omega \epsilon_\theta(\hat{\mathbf{x}}_t, t)) + (1 + \omega) \epsilon_\theta(\hat{\mathbf{x}}_t, \mathbf{y}, t) \right) + \sigma_t z \quad (17)$$

where $\epsilon_\theta(\hat{\mathbf{x}}_t, t)$ and $\epsilon_\theta(\hat{\mathbf{x}}_t, \mathbf{y}, t)$ represent the estimated noise for unconditional and conditional DDPM, respectively, with ω as the scaling factor controlling the diversity.

V. CONCLUSION

We develop an adaptive FK conditioning zero-shot denoising network for DAS-VSP data based on the C-DDPM. Our approach begins by isolating seismic signal components from the field-acquired DAS-VSP data, separating them from purely ambient noise. The ambient noise is then combined with finely synthesized VSP data to create a robust training set, enabling the network to perform zero-shot denoising on DAS-VSP data containing seismic signals. Experimental results demonstrate that the denoising performance of C-DDPM surpasses that of classical networks like DnCNN and U-Net, on both synthetic

and real-world datasets. Furthermore, ablation studies reveal that integrating AFKCDM and ADConv significantly enhances the model's denoising capability, minimizes signal leakage, and achieves remarkable results under zero-shot conditions.

This proposed method offers valuable insights for advancing seismic signal processing, especially in handling complex noise environments in DAS-VSP applications.

REFERENCES

- [1] F. Walter et al., "Distributed acoustic sensing of microseismic sources and wave propagation in glaciated terrain," *Nature Commun.*, vol. 11, no. 1, p. 2436, May 2020.
- [2] G. Marra et al., "Ultrastable laser interferometry for earthquake detection with terrestrial and submarine cables," *Science*, vol. 361, no. 6401, pp. 486–490, Aug. 2018.
- [3] J. B. Ajo-Franklin et al., "Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection," *Sci. Rep.*, vol. 9, no. 1, p. 1328, Feb. 2019.
- [4] A. D. Booth et al., "Distributed acoustic sensing of seismic properties in a borehole drilled on a fast-flowing Greenlandic outlet glacier," *Geophys. Res. Lett.*, vol. 47, no. 13, Jul. 2020, Art. no. e2020GL088148.
- [5] G. Fang, Y. E. Li, Y. Zhao, and E. R. Martin, "Urban near-surface seismic monitoring using distributed acoustic sensing," *Geophys. Res. Lett.*, vol. 47, no. 6, 2020, Art. no. e2019GL086115.
- [6] A. Titov et al., "Modeling and interpretation of scattered waves in interstage distributed acoustic sensing vertical seismic profiling survey," *Geophysics*, vol. 86, no. 2, pp. D93–D102, Mar. 2021.
- [7] A. Mateeva et al., "Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling," *Geophys. Prospecting*, vol. 62, no. 4, pp. 679–692, Jul. 2014.
- [8] K. Harris, D. White, and C. Samson, "Imaging the aquistore reservoir after 36 kilotonnes of CO₂ injection using distributed acoustic sensing," *Geophysics*, vol. 82, no. 6, pp. M81–M96, Nov. 2017.
- [9] F. Cheng et al., "Testing of a permanent orbital surface source and distributed acoustic sensing for monitoring of unconventional reservoirs: Preliminary results from the eagle Ford shale," *Geophysics*, vol. 86, no. 2, pp. P1–P12, Mar. 2021.
- [10] M. W. Becker, C. Ciervo, M. Cole, T. Coleman, and M. Mondanos, "Fracture hydromechanical response measured by fiber optic distributed acoustic sensing at milliHertz frequencies," *Geophys. Res. Lett.*, vol. 44, no. 14, pp. 7295–7302, Jul. 2017.
- [11] A. F. Baird et al., "Characteristics of microseismic data recorded by distributed acoustic sensing systems in anisotropic media," *Geophysics*, vol. 85, no. 4, pp. KS139–KS147, Jul. 2020.
- [12] B. Luo, G. Jin, and F. Stanek, "Near-field strain in distributed acoustic sensing-based microseismic observation," *Geophysics*, vol. 86, no. 5, pp. P49–P60, Sep. 2021.
- [13] T. S. Hudson et al., "Distributed acoustic sensing (DAS) for natural microseismicity studies: A case study from Antarctica," *J. Geophys. Res., Solid Earth*, vol. 126, no. 7, Jul. 2021, Art. no. e2020JB021493.
- [14] I. Lior et al., "On the detection capabilities of underwater distributed acoustic sensing," *J. Geophys. Res., Solid Earth*, vol. 126, no. 3, Mar. 2021, Art. no. e2020JB020925.
- [15] A. Hartog, B. Frignet, D. Mackie, and M. Clark, "Vertical seismic optical profiling on wireline logging cable," *Geophys. Prospecting*, vol. 62, no. 4, pp. 693–701, Jul. 2014.
- [16] D. Molteni, M. J. Williams, and C. Wilson, "Detecting microseismicity using distributed vibration," *First Break*, vol. 35, no. 4, pp. 51–55, 2017.
- [17] M. Karrenbach et al., "Fiber-optic distributed acoustic sensing of microseismicity, strain and temperature during hydraulic fracturing," *Geophysics*, vol. 84, no. 1, pp. D11–D23, Jan. 2019.
- [18] Y. Zhao, Y. Li, and N. Wu, "Distributed acoustic sensing vertical seismic profile data denoiser based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2020, Art. no. 5900511.
- [19] 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.
- [20] A. Lellouch, R. Schultz, N. J. Lindsey, B. Biondi, and W. L. Ellsworth, "Low-magnitude seismicity with a downhole distributed acoustic sensing array—Examples from the FORGE geothermal experiment," *J. Geophys. Res., Solid Earth*, vol. 126, no. 1, 2021, Art. no. e2020JB020462.
- [21] Y. Chen et al., "Denoising of distributed acoustic sensing seismic data using an integrated framework," *Seismol. Res. Lett.*, vol. 94, no. 1, pp. 457–472, Jan. 2023.
- [22] 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.
- [23] S. Yu, J. Ma, and W. Wang, "Deep learning for denoising," *Geophysics*, vol. 84, no. 6, pp. V333–V350, Jul. 2019.
- [24] D. Augusto Borges Oliveira, D. Semin, and S. Zaytsev, "Ground roll suppression using convolutional neural networks," 2020, *arXiv:2010.15209*.
- [25] F. Li, H. Liu, W. Wang, and J. Ma, "Swin transformer for seismic denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, pp. 1–5, 2024.
- [26] R. Durall, A. Ghanim, N. Ettrich, and J. Keuper, "Dissecting U-net for seismic application: An in-depth study on deep learning multiple removal," 2022, *arXiv:2206.12112*.
- [27] B. Wang, N. Zhang, W. Lu, and J. Wang, "Deep-learning-based seismic data interpolation: A preliminary result," *Geophysics*, vol. 84, no. 1, pp. V11–V20, Jan. 2019.
- [28] X. Wu, Z. Geng, Y. Shi, N. Pham, S. Fomel, and G. Caumon, "Building realistic structure models to train convolutional neural networks for seismic structural interpretation," *Geophysics*, vol. 85, no. 4, pp. WA27–WA39, Jul. 2020.
- [29] B. Schwarz, K. Sager, P. Jousset, G. Currenti, C. Krawczyk, and V. Tsai, "Leveraging coherent wave field analysis and deep learning in fiber-optic seismology," in *Proc. EGU Gen. Assem. Conf. Abstracts*, 2021, p. 7856.
- [30] E. R. Martin et al., "A seismic shift in scalable acquisition demands new processing: Fiber-optic seismic signal retrieval in urban areas with unsupervised learning for coherent noise removal," *IEEE Signal Process. Mag.*, vol. 35, no. 2, pp. 31–40, Mar. 2018.
- [31] M. van den Ende, I. Lior, J.-P. Ampuero, A. Sladen, A. Ferrari, and C. Richard, "A self-supervised deep learning approach for blind denoising and waveform coherence enhancement in distributed acoustic sensing data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 7, pp. 3371–3384, Jul. 2023.
- [32] L. Yang et al., "Denoising of distributed acoustic sensing data using supervised deep learning," *Geophysics*, vol. 88, no. 1, pp. WA91–WA104, Jan. 2023.
- [33] Y. Chen, M. Zhang, M. Bai, and W. Chen, "Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning," *Seismol. Res. Lett.*, vol. 10, pp. 1552–1564, May 2019.
- [34] L. Yang et al., "Unsupervised 3-D random noise attenuation using deep skip autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5905416.
- [35] F. Qian et al., "Unsupervised seismic footprint removal with physical prior augmented deep autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5910920.
- [36] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840–6851.
- [37] J. Ho and T. Salimans, "Classifier-free diffusion guidance," 2022, *arXiv:2207.12598*.
- [38] H.-R. Zhang, Y. Liu, Y.-H. Sun, and G. Chen, "SeisResoDiff: Seismic resolution enhancement based on a diffusion model," *Petroleum Sci.*, vol. 21, no. 5, pp. 3166–3188, Oct. 2024.
- [39] J. Yang, Y. Zhang, Y. Sheng, Y. Lin, and L. Yang, "A novel diffusion model for pairwise geoscience data generation with unbalanced training dataset," 2025, *arXiv:2501.00941*.
- [40] Y. Li, H. Zhang, J. Huang, and Z. Li, "Conditional denoising diffusion probabilistic model for ground-roll attenuation," 2024, *arXiv:2403.18224*.
- [41] D. Trappolini et al., "Cold diffusion model for seismic denoising," *J. Geophys. Res., Mach. Learn. Comput.*, vol. 1, no. 2, Jun. 2024, Art. no. e2024JH000179.
- [42] Q. Liu and J. Ma, "Generative interpolation via a diffusion probabilistic model," *Geophysics*, vol. 89, no. 1, pp. V65–V85, Jan. 2024.
- [43] Y. Xiao, K. Li, Y. Dou, W. Li, Z. Yang, and X. Zhu, "Diffusion models for multidimensional seismic noise attenuation and superresolution," *Geophysics*, vol. 89, no. 5, pp. V479–V492, Sep. 2024.
- [44] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, "Image super-resolution via iterative refinement," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4713–4726, Apr. 2022.
- [45] S. Zhu, H. Zhao, T. Bai, W. Chen, and Y. Chen, "SelfTexture: Self-supervised learning for spatially correlated noise removal of DAS VSP data via adaptive texture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5937017.
- [46] A. Richardson, "Deepwave," Zenodo, Tech. Rep., Sep. 2023, doi: 10.5281/zenodo.8381177.