SYMMETRIC AUTOENCODERS FOR RETRIEVING RINGING MODES AND BACKSCATTER FROM VEHICLE SEISMIC NOISE TO DETECT NEAR-SURFACE VOIDS

A PREPRINT

Sanket Bajad
Centre for Earth Sciences
Indian Institute of Science
Bengaluru, India 560012
sanketbajad@iisc.ac.in

Pawan Bharadwaj
Centre for Earth Sciences
Indian Institute of Science
Bengaluru, India 560012
pawan@iisc.ac.in

ABSTRACT

Detecting deep voids in urban subsurfaces remains a challenging problem for geophysical methods, particularly with conventional active seismic sources that lack the low-frequency energy necessary for greater depth penetration. To address this, we explored an alternative approach using seismic noise generated by the engine of an idle vehicle to supply the needed low frequencies and enhance the detection capabilities for voids at depths beyond 8 m. A key obstacle in processing this seismic noise is the presence of cross-correlation residuals, which occur when traditional linear stacking methods struggle to achieve stable averages, particularly with shorter-than-usual time duration data. These residuals interfere with the identification of void signatures, such as backscatters and ringing patterns, undermining the accuracy of void detection in dynamic urban settings. To tackle this issue, we employ symmetric autoencoders (SymAE), an unsupervised machine learning architecture, applied to vehicle seismic noise. By training neural networks to disentangle coherent subsurface information from source-induced noise, SymAE effectively minimizes cross-correlation residuals and enhances the extraction of backscatters and ringing patterns. By applying this technique in a semi-controlled field experiment, we successfully detected two known voids at approximately 10 m depth. SymAE reduces data requirements while improving detection accuracy, showing potential for rapid, cost-effective, and reliable subsurface void detection in urban environments.

1 Introduction

The field of near-surface geophysics has rapidly evolved, focusing on the detailed imaging of the shallow subsurface, typically spanning from a few meters to depths of hundreds of meters. While it shares foundational physical principles with deeper subsurface investigations, its unique proximity to the Earth's surface and the considerable heterogeneity encountered therein set near-surface geophysics apart as a distinct field, addressing environmental and engineering challenges.

Near-surface geophysics addresses environmental challenges such as accurate groundwater monitoring in densely populated urban areas (Steeples, 2005; Binley et al., 2015; Clements and Denolle, 2018). It is a crucial tool in mitigating water scarcity and detecting risks such as sinkholes and covert tunnels, which threaten public safety and national security (Gutiérrez et al., 2014; Sloan et al., 2015; Youssef et al., 2020). Engineering near-surface geophysics is vital in monitoring existing structures during underground construction, identifying risks, and mitigating issues arising from construction-induced vibrations, essential for urban expansion. Most urban geophysical imaging challenges related to near-surface areas involve investigating changes in geophysical properties both spatially and temporally (Fang et al., 2020).

This paper focuses on detecting clandestine voids as a spatial variation problem, akin to anomaly detection, wherein we hypothesize lateral homogeneity within the subsurface medium, with the exception of the presence of voids. Existing seismic methods for detecting such voids include reflection, refraction, diffraction, and surface waves (Branham and

Steeples, 1988; Dobecki, 1988; Belfer et al., 1998; Landa and Keydar, 1998; Sloan et al., 2013; Schwenk et al., 2014, 2016; Peterie and Miller, 2015). Among these, surface-wave methods, such as multichannel analysis of surface waves (MASW) (Park et al., 1999) and surface-wave backscattering (Schwenk et al., 2016; Ivanov et al., 2017), are preferred for detecting near-surface anomalies, as surface waves attenuate more slowly than body waves. In addition, various studies have focused on identifying resonant ringing patterns emerging from these high-contrast subsurface voids (Korneev, 2009; Rubin et al., 2014; Schneider et al., 2017; Sloan and Feigenbaum, 2020, 2024). However, observing these phenomena necessitates a seismic source capable of generating adequate energy to reach the void. Many active sources, such as sledgehammers, are bandlimited and lack low-frequency energies (Mi and Xia, 2021). As a result, active MASW surveys using a hammer source fail to produce reliable images of the void. For instance, we included the outcomes of an MASW survey in Appendix A that did not succeed in identifying a known void. In contrast, passive seismic survey studies show promise in retrieving low-frequency surface waves from ambient urban noise (Park et al., 2005; Cheng et al., 2015; Morton et al., 2021). Unlike active seismic methods, passive and semi-passive techniques can utilize readily available urban sources, such as controlled vehicle sources, which generate lower frequencies than typical active sources (Hanafy et al., 2015; Li and Hanafy, 2016). However, semi-passive imaging, where the source location is fixed, faces challenges in source repeatability, hindering accurate interpretation essential for time-lapse monitoring (Romero Jr et al., 1997; Lumley, 2001; Gunasekera et al., 2003; Gouedard et al., 2008). In this study, we concentrate on a relatively simpler semi-passive approach, utilizing an idle vehicle with its engine turned on as a seismic source, characterized by uncontrollable signatures but a controlled position. Given our emphasis on an idle vehicle as a seismic source, it is vital to address the challenges associated with source repeatability. Successfully addressing these obstacles will open the door for cost-effective subsurface imaging using urban seismic noise in future projects.

Given the similarities in challenges between semi-passive and passive settings, we build our approach on the method of cross-correlation, commonly used in passive subsurface imaging. Several authors (Weaver and Lobkis, 2001; Campillo and Paul, 2003; Wapenaar, 2004; Snieder, 2004; Weaver, 2005) have demonstrated that the cross-correlation of a random field converges to the Green function, which represents the medium's response at one point when subjected to an impulsive source at another point. A critical aspect of passive seismic imaging is stacking cross-correlated data to enhance the signal-to-noise ratio (S/N).

The conventional stacking method (Mayne, 1962), which computes the arithmetic mean of data over time windows, typically requires random sources and long-duration data for convergence. In the context of urban-noise imaging, collecting data over extended periods is unfeasible; hence, there is a need to design effective stacking methodologies (Zhang et al., 2020). Several researchers (Larose et al., 2004; Snieder, 2004; Sabra et al., 2005; Weaver and Lobkis, 2005; Zhang et al., 2021) have explored the issues encountered in urban seismic imaging when employing linear stacking techniques with shorter time-duration data than usual. Linear stacking with shorter time-duration data leads to residual cross-correlation fluctuations, which hinder convergence to stable long-term averages (De Ridder et al., 2014). In other words, these cross-correlation residuals can obscure crucial surface-wave information (Cheng et al., 2019; Zhang et al., 2020) when processing passive traffic noise (Louie, 2001; Park et al., 2004; Behm et al., 2014) in shorter time intervals.

Nonlinear stacking methods (Schimmel et al., 2011; Rückemann, 2012; Olivier et al., 2015; Xie et al., 2020) improve the coherence of the seismic signal by using nonlinear weighting operations, such as phase weighting, muted weighting, and optimal weighting, to weigh time windows based on signal consistency across the traces, thus enhancing the S/N. Our approach generalizes nonlinear stacking by weighing the time windows with the learned parameters of the neural networks. In other words, we train neural networks to extract coherency from seismic data. Compared to most contemporary deep-learning approaches, the training is unsupervised, where the architectural design of the symmetric autoencoder (SymAE) (Bharadwaj et al., 2020, 2022, 2024) allows for disentanglement of the subsurface-scattering effects from vehicle noise. Specifically, we learn a useful representation of cross-correlated seismic noise, where the features associated with Green's function are separated from the remaining nuisances due to source variability. Compared to conventional stacking, we validated the efficiency of SymAE by locating known voids $8 - 12$ m deep after extracting the associated low-frequency backscattered surface waves and ringing effects. The emphasis lies on detecting the voids through these associated backscattered and ringing signatures to develop a robust void detection system. In summary, our focus is on the detection of voids, with an aim to address two key challenges: 1. the lack of low-frequency energy in active sources like sledgehammers, and 2. the convergence issues faced by conventional stacking methods when processing shorter intervals of vehicle noise data in urban environments. The field experimental setup, using an idle vehicle with its engine turned on as a seismic source, is detailed in the following section.

2 Experimental setup

We conducted geophysical surveys using a roll-along setup with 24-channel vertical-component geophones with a natural frequency of 4.5 Hz connected by leapfrog cables. The surveys were carried out at sites with known voids and our primary objective was to detect voids located approximately at $x = 100$ m along the survey line at site 1 and at

$x = 126$ m along the survey line at site 2. The voids are air filled and have diameters approximately between $1.5 - 2$ m. These sites primarily consist of silt clay with minimal gravel, transitioning to moderately dense clayey silt at varying depths. The S-wave velocity (V_s) for these locations is approximately $300 - 330$ m/s, as determined through multiple MASW surveys. The positions of the survey line were marked by x and the geophones were separated by a distance of 2 m. The source offset, i.e., the distance between the vehicle source and the first geophone, was set at 6 m. As we moved the vehicle along the survey line, we recorded 5-minute noise record generated by the vehicle at each position of the source. At each location, the vehicle was idle with its engine on with a total spread length of 46 m. We covered a total distance of 148 m during each survey. The spectrum of the vehicle source, shown in Figure 1a, indicates the ability of the vehicle to generate low frequencies (in the range of $2 - 15$ Hz) compared to a hammer. The survey, illustrated in Figure 2, involved the movement of a vehicle along the survey line, with each roll-along covering a distance of 6 m. Aside from the void, it is likely that the subsurface exhibits lateral homogeneity at both locations.

For noise preprocessing, we employed several steps. Firstly, we windowed the noise into 2 s time windows with 50% of overlap. Subsequently, we whitened the noise using the average spectrum of the vehicle and applied a low-pass filter with a cutoff frequency of 20 Hz. This filtering aimed to isolate lower frequencies that were conducive to detecting deeper voids within the subsurface. We conducted two data-processing scenarios: analyzing the entire 5-minute recording and analyzing the last 2.5-minute segment. The choice of the 2.5-minute segment is arbitrary, and any 2.5-minute segment can be chosen.

This experimental setup, accompanied by the specified data preprocessing scenarios, provided the framework for our subsequent analyses aimed at isolating and characterizing the void signatures discussed in the next section.

3 Theory

We begin by exploring the principles of cross-correlation and understanding the challenge of cross-correlation residuals in linear stacking. Subsequently, we investigate the void signatures in the seismic wavefield: 1. backscattered correlations, focusing on how surface-wave interactions can reveal hidden subsurface features, and 2. the ringing effect, where trapped seismic energy within voids produces distinctive resonances. In traditional processing, we observed that these void signatures are often obscured by cross-correlation residuals.

3.1 Cross-correlation and linear stacking

Cross-correlated noise gathers (CCNs) are generated by calculating the cross-correlation of overlapping time windows between pairs of receivers. For this purpose, preprocessed 2 s overlapping time windows are used to create the cross-correlation of noise signals recorded at different receiver locations. The equation for cross-correlation between two receivers r and $r + \delta x$ is given as:

$$g^k_{r,s}(\delta x, \tau) = \int_t d^k(r, s, t) d^k(r + \delta x, s, t + \tau) dt, \tag{1}$$

where $d^k(r, s, t)$ represents the noise recorded at the receiver location r due to the vehicle at the source location s along the survey line, with the signature of the vehicle specific to the time window denoted by superscript k. δx represents the inter-receiver distance and τ is the cross-correlation time lag. Each CCN corresponds to a set of cross-correlations computed between a reference receiver and neighboring receivers over a range of time lags. For each reference receiver r, the gathers are computed by cross-correlating the signal at receiver r with signals at adjacent receivers $r + \delta x$, where δx varies over the specified range depending upon the signatures of interest. The green rectangles in Figure 2 refer to the spread of gathers for the reference receiver positions to be considered while discussing the results. For deep learning models to function effectively, they require normalized data during training. Specifically, in this scenario, each CCN is normalized so that $g^k_{r,s}(0, 0) = 1$.

Traditionally, all the CCNs are stacked linearly across time windows to extract the medium's response (i.e., subsurface information). The extraction of void signatures is critically dependent on the time windows utilized for stacking. Therefore, we implement two stacking procedures to extract backscattered correlations and ringing effects:

- **Backscattering experiment**: We stack CCNs over time windows corresponding to multiple source locations that contribute to a specific reference receiver during the roll-along survey. This procedure leverages the independence of the phase terms in the scattered wavefield, as recorded at the reference receiver, from the varying source locations. In this experiment, the range of δx varies from -12 m to 12 m.

  We consider only the CCNs generated from the source locations where the CCNs for the specified range of δx can be reliably computed. In essence, using the reference receiver location and the δx range, these source

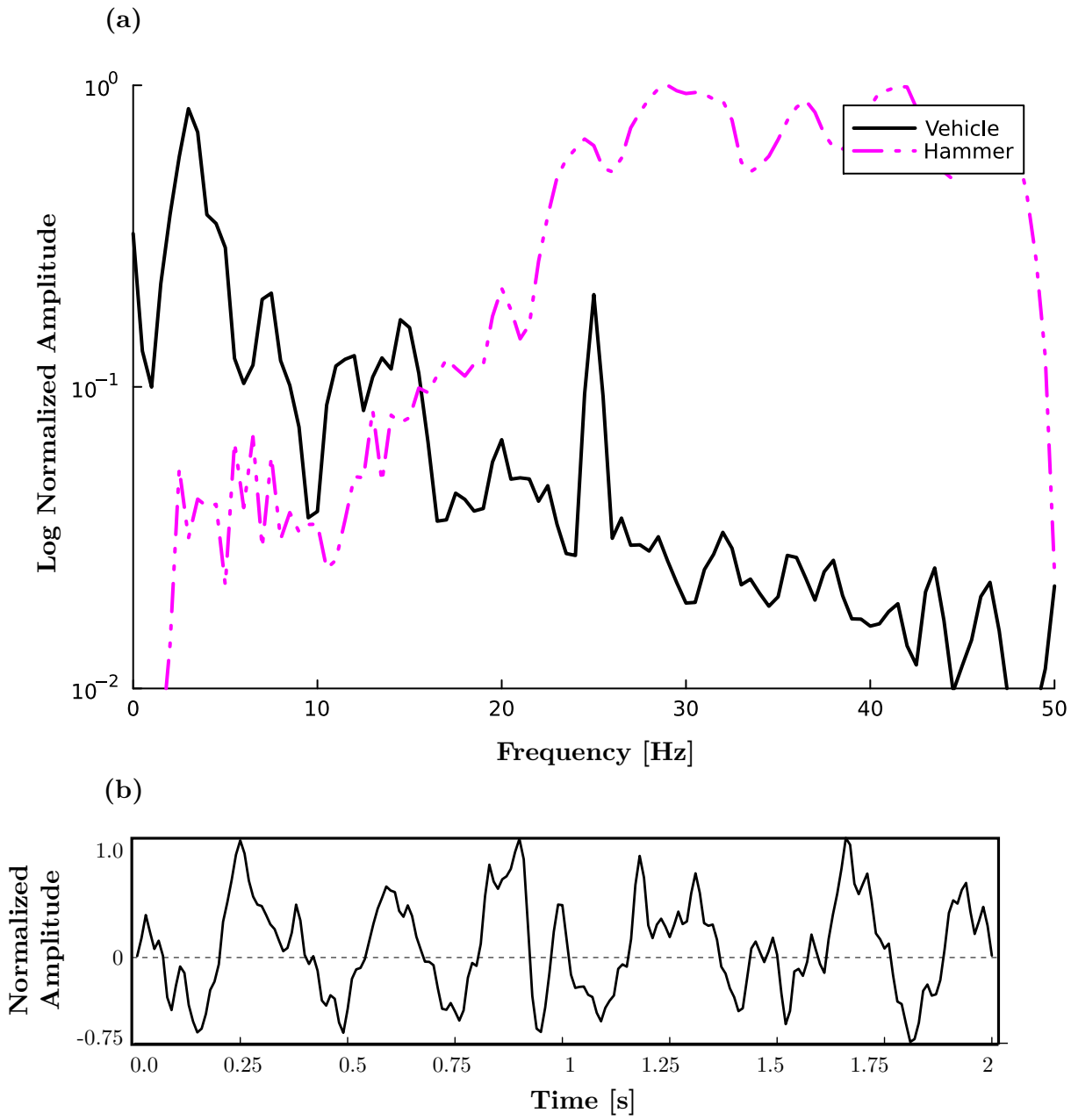

Figure 1: (a): Normalized raw logarithmic amplitude spectra for both vehicle (2 s time window) and hammer sources, demonstrating that the vehicle source is capable of generating lower frequencies in comparison to the hammer source. (b): Seismic trace corresponding to a 2 s time window for the vehicle source.

locations can be determined by examining the Seg-Y file indices along the y-axis in Figure 2. For example, in this experiment, the CCNs associated with the reference receiver at $r = 100$ m, at site 1, are derived from the source locations associated with Seg-Y file indices 2080 through 2083, representing sources positioned at $x = \{118, 124, 130, 136\}$ m. Thus, the linear stacking formula for the backscattering experiment is expressed as:

$$\bar{g}_{r,s}(\delta x, \tau) = \sum_{k,s} g^{k}_{r,s}(\delta x, \tau), \tag{2}$$

where $\bar{g}_{r,s}(\delta x, \tau)$ represents the stacked CCN.

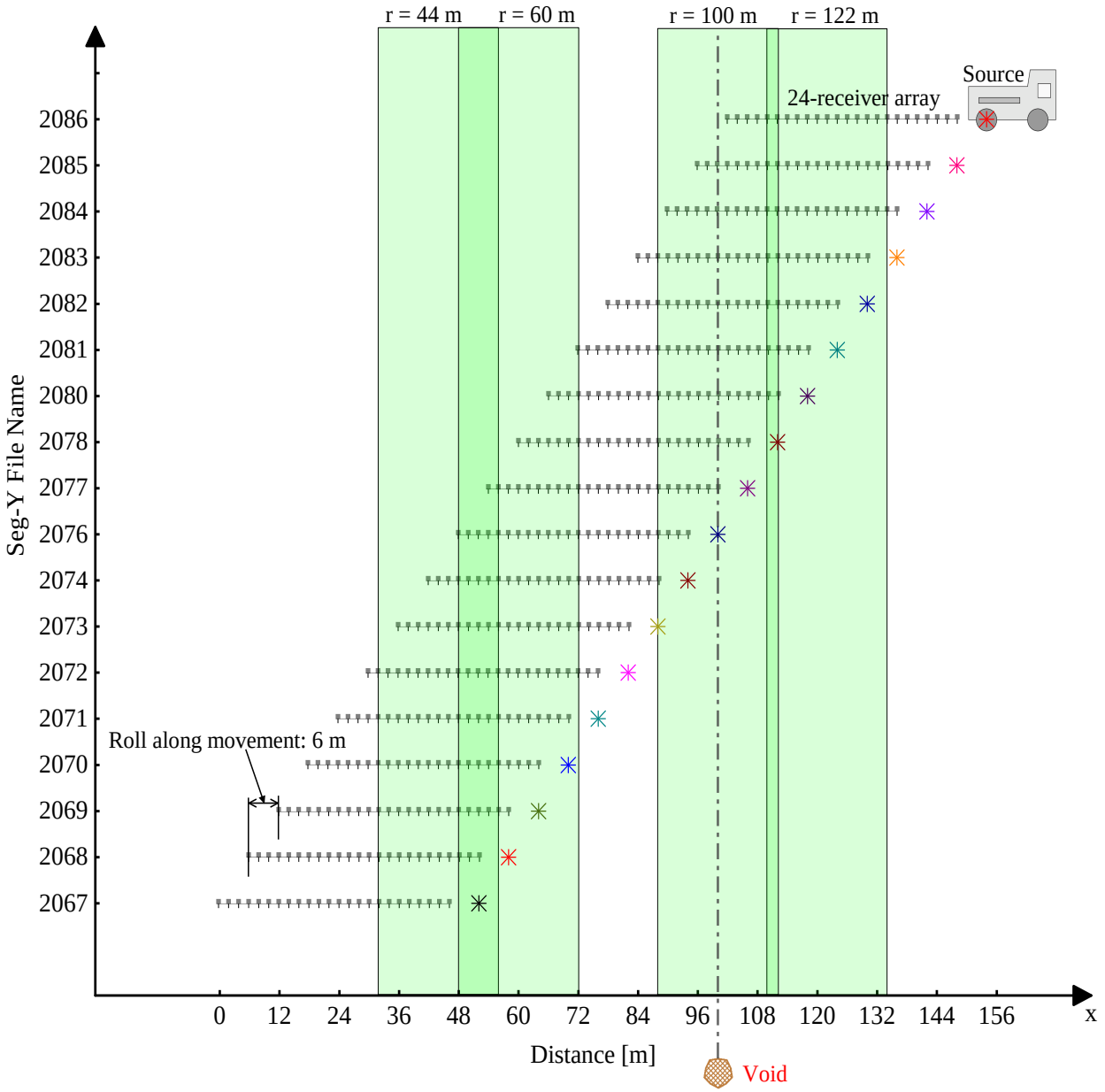

Figure 2: A schematic illustration of a roll-along survey conducted to detect subsurface voids. Variability in the source signature with changes in source position is represented using different colors. The dash-dotted line indicates the void's location at 100 m for site 1. The reference receiver location is indicated by r. Green boxes represent the spread of gathers corresponding to the reference receiver positions considered (subsequently denoted by δx).

- **Ringing experiment**: Unlike backscattering, the resultant ringing pattern might be influenced by the position of the source (Sloan and Feigenbaum, 2024). Therefore, we focus on stacking time windows corresponding to a particular source location, s_r, closest to the reference receiver r. For example, at site 1, the CCNs associated with the reference receiver at $r = 100$ m, are derived from the source location associated with the Seg-Y file numbered 2077 representing the closest source located at $x = 106$ m. Thus, for the ringing experiment, the stacking approach is expressed as:

$$\bar{g}_{r,s_r}(\delta x, \tau) = \sum_k g^k_{r,s_r}(\delta x, \tau). \tag{3}$$

  In the ringing experiment, δx varied from -20 m to 0 m.

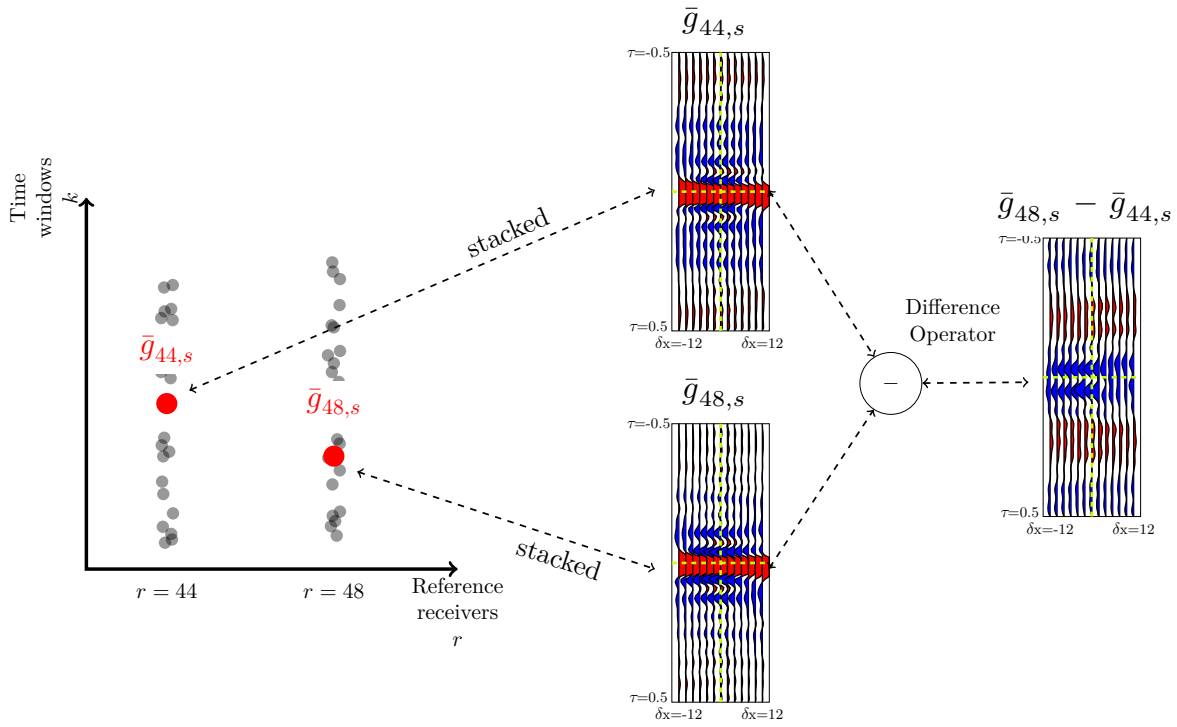

Figure 3: Cross-correlation residuals between two nearby receivers, $r = 44\,\text{m}$ and $r = 48\,\text{m}$. Ideally, the difference should be zero in the absence of residuals. However, due to the unstable stacking average for a shorter processing interval, the difference is significantly high. On the CCNs, the green dashed lines illustrate the location of the reference receiver and the zero lag.

A comprehensive discussion on the choice of δx for both these experiments is provided in Appendix B.

As mentioned earlier, the conventional stacking approach requires long-duration data, typically $20-150$ minutes (Hanafy et al., 2015; Li and Hanafy, 2016; Wang et al., 2023), for a stable convergence. For instance, in the backscattering experiment, we initially anticipated that the stacked cross-correlated data for two closely spaced receiver positions, specifically $r = 44\,\text{m}$ and $r = 48\,\text{m}$, would yield similar results. Since these positions are within a relatively homogeneous medium, as shown in Figure 2, we expected the difference between the two data to be nearly zero. However, due to the insufficient time windows available for analysis, the stacked cross-correlated data for these two receivers did not converge to similar stable averages. Consequently, the observed difference between them remained higher than anticipated, as depicted in Figure 3. A similar issue was encountered during the data processing of the ringing experiment. This underscores the need to eliminate these cross-correlation residuals to effectively extract the desired void signatures (i.e., backscattered correlations and ringing patterns), which are discussed in the following subsections.

3.2 Backscattered correlations as void signatures

We consider a scenario where the noise is predominantly composed of surface waves and their backscatter, thus:

$$d^k(r,s,t) = d^k_{\leftarrow}(r,s,t) + d^k_{\rightarrow}(r,s,t), \tag{4}$$

and

$$d^k(r+\delta x,s,t+\tau) = d^k_{\leftarrow}(r+\delta x,s,t+\tau) + d^k_{\rightarrow}(r+\delta x,s,t+\tau). \tag{5}$$

Here, $d^k_{\rightarrow}$ signifies backscattered surface waves (caused by the void) travelling in the direction of increasing x, while $d^k_{\leftarrow}$ represents direct surface waves moving in the opposite direction, i.e., decreasing x. By incorporating Equations 4 and 5 into Equation 1, the resulting cross-correlated noise is a superposition of four terms:

$$g^k_{r,s} = {}_{\leftarrow,\leftarrow}g^k_{r,s} + {}_{\leftarrow,\rightarrow}g^k_{r,s} + {}_{\rightarrow,\leftarrow}g^k_{r,s} + {}_{\rightarrow,\rightarrow}g^k_{r,s}, \tag{6}$$

where the cross-correlation between the direct surface waves, ${}_{\leftarrow,\leftarrow}g^k_{r,s}$, measured at two receivers (i.e., receivers at r and $r+\delta x$) is the strongest, followed by cross-terms like ${}_{\leftarrow,\rightarrow}g^k_{r,s}$ and ${}_{\rightarrow,\leftarrow}g^k_{r,s}$. The objective is to isolate these cross-terms, which signify backscatter correlations, to facilitate void detection.

The presence of cross-correlation residuals complicates the extraction of these cross-terms in linearly stacked data. This issue was evident in our experiment, for $\bar{g}_{100,s}$ illustrated in Figures 4b and 4e, where the receivers capture both direct and backscattered surface waves. Although we expect to observe weaker backscatter correlations, the stronger cross-correlation residuals obscure these weaker backscatter correlations, as will be discussed in the results section.

3.3 Ringing effects as void signatures

The ringing effects arise from the confinement of seismic or acoustic energy due to disparities in medium properties. As demonstrated by Rubin et al. (2014), in unbounded voids (e.g., cavities, caves), this energy manifests itself as standing acoustic normal modes, while in bounded voids (e.g., boxes, barrels, tubes), it is characterized by flexural vibrations of the boundaries. This trapped energy is subsequently released through resonant emission with the emitted frequencies contingent upon the geometrical and physical attributes of the voids. In the time domain, the ringing effect is observed as a long wave train (Rubin et al., 2014), referred to in this paper as the ringing pattern. Analytically, assuming a spherical void, the resonant frequencies ($f_{m,n}$) for unbounded voids, given by Schneider et al. (2017), are expressed as:

$$f_{m,n} = \frac{j^{'}_{m,n}\, V_p}{\pi d} \tag{7}$$

where the $j^{'}_{n,m}$ is the m^{th} root of the derivative of the n^{th} spherical Bessel function, V_p is the P-wave velocity of the void and d is the diameter of the void. Korneev (2009), through the synthetic experiments, suggested that circumferential waves propagating around the boundary of a barrel are responsible for ringing. Figure 5 illustrates a simulation demonstrating the release of trapped seismic energy from a 10 m deep air-filled void, located at $x = 126$ m, within an elastic homogeneous medium, where the void's P-wave velocity V_p is modeled as 340 m/s, and the S-wave velocity V_s is set to 0.5% of the medium's S-wave velocity. Sloan and Feigenbaum (2024) presented several field examples showing resonances related to voids. However, while the mechanisms or characteristics of the resonant wave propagation phenomena remain unclear, previous studies suggest that: 1. the source should be in proximity to the void for ringing to occur, 2. the ringing pattern is weaker in amplitude, 3. the ringing pattern is a continuous long wave train which can be analyzed once the direct wave has passed, and 4. a void can exhibit a range of resonant frequencies. However, higher frequencies are expected to experience significant attenuation, leading to negligible ringing effects. In contrast, attenuation is less pronounced for low frequencies, making the ringing effect observable primarily in low-frequency seismic noise.

Weaker resonant emissions tend to be comparable to cross-correlation residuals, leading to confusion in interpreting the ringing pattern. This ambiguity was observed at site 2, where the void was located at $x = 126$ m along the survey line. Three reference receivers were considered to explain the ambiguity: $r = 60$ m, $r = 104$ m and $r = 120$ m.

It was anticipated that only the reference receiver at $r = 120$ m, being closest to the void when the source was above it, would exhibit a distinct ringing pattern as shown in Figures 6i and 6l for both the 2.5-minute and 5-minute scenarios. For $r = 104$ m, in $\bar{g}_{104,s_r}$, however, a similar ringing pattern was observed in the 2.5-minute scenario, but this pattern altered in the 5-minute scenario (Figures 6h and 6k). This alteration can be attributed to the decrease in cross-correlation residuals, suggesting that these residuals may be misinterpreted as the ringing pattern. Similar alteration in ringing pattern was visible for $r = 60$ m in Figures 6g and 6j. It is important to note that we have analyzed the ringing pattern for higher positive time lags leaving earlier lags, as ringing should be seen as a continuous long-wave train. Ultimately, similar to the backscattering experiment, this underscores the need to eliminate cross-correlation residuals to prevent misinterpretation. Following the identification of challenges of cross-correlation residuals and the critical need to mitigate them for the effective extraction of backscatter correlations and ringing effects, the subsequent section explores our approach to addressing these challenges and enhancing void detection capabilities.

4 Methodology

Our research leverages the capabilities of neural networks to extract coherency from seismic data. Neural networks can be trained in an unsupervised manner to learn nonlinear transformations that reveal the underlying structures inherent in the data. This process, termed representation learning, is essential for acquiring meaningful representations that enhance our understanding of the phenomena of interest, particularly in complex data sets. Autoencoders are a well-established technique in representation learning; however, conventional autoencoders often encounter difficulties in extracting coherent features, predominantly owing to their deficiency in physical interpretability. Therefore, we utilize

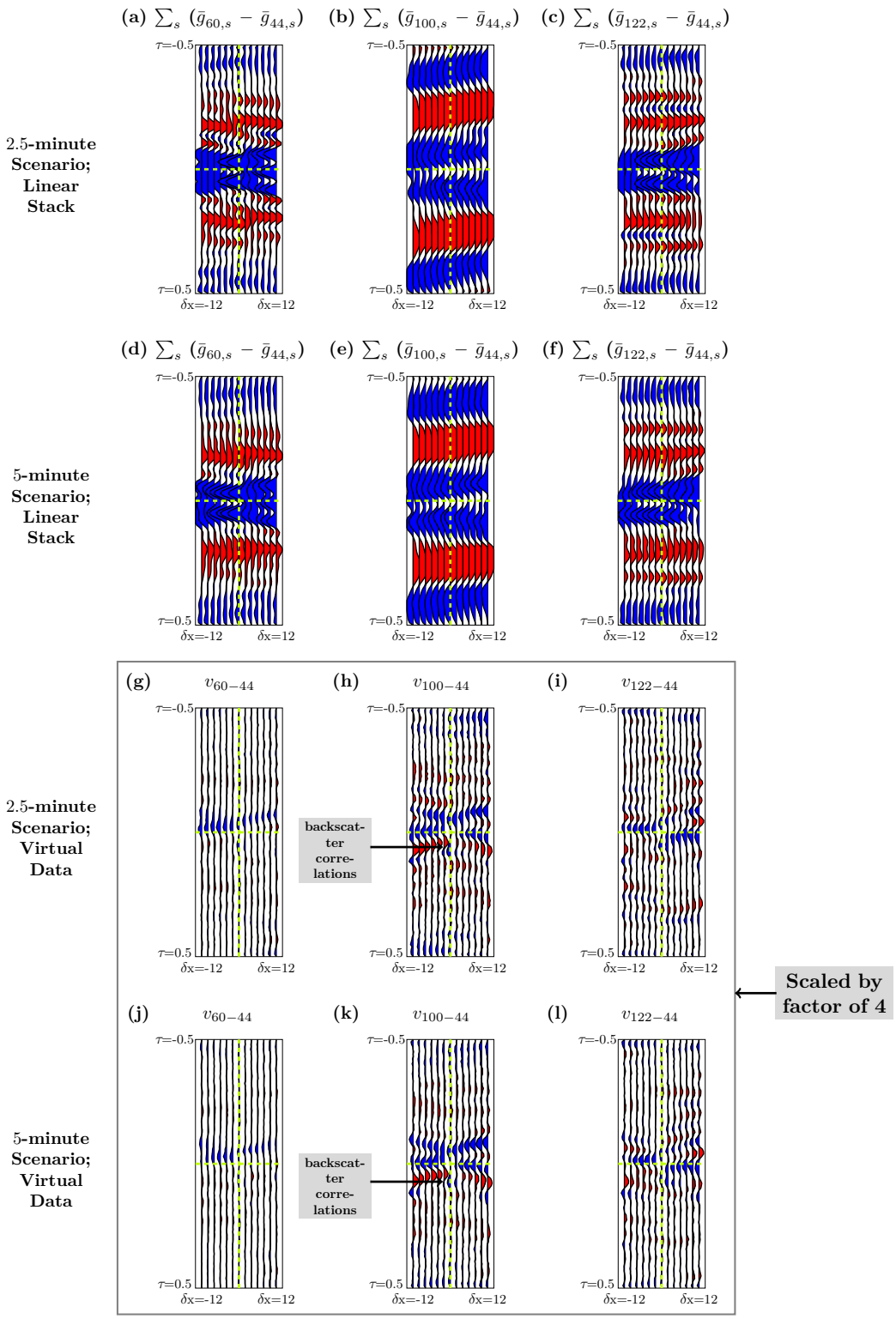

Figure 4: Backscatter correlations as a void signature: Panels (a) to (c) show differential analysis of stacked cross-correlated noise for 2.5-minute and panels (d) to (f) for 5-minute scenarios, both with significant cross-correlation residuals. Here, $r = 44$ m is the baseline reference receiver. DVCCs, from Equation 21, show pronounced backscatter energy in panels (h) and (k), while panels (g) and (j) show reduced cross-correlation residuals. DVCCs (panels (g)–(l)) are scaled by a factor of 4 relative to the differential noise (panels (a)–(f)).

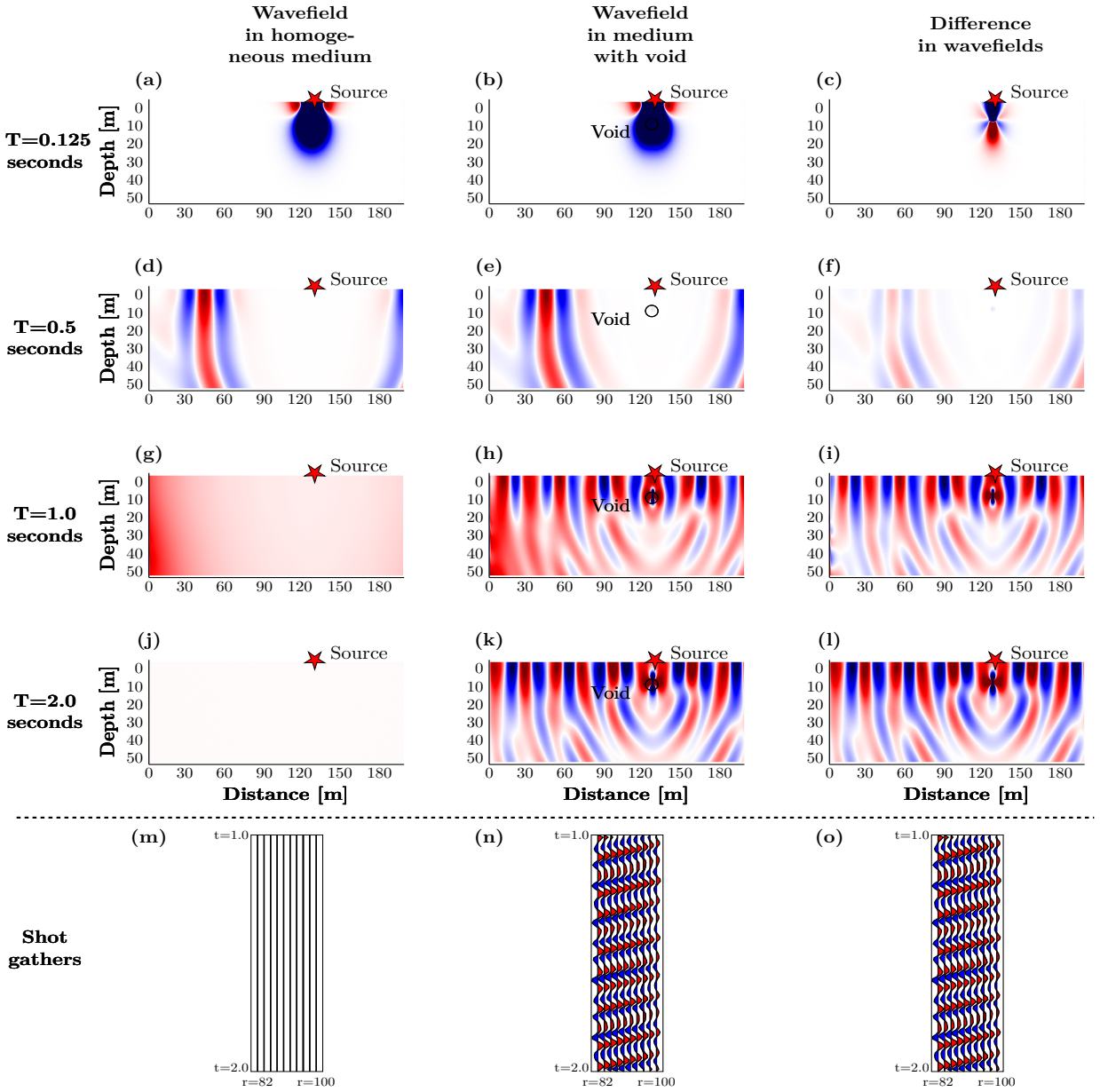

Figure 5: Seismic wavefield snapshots: Panels (a)–(l) display temporal snapshots of seismic wavefields in a homogeneous medium and a medium containing a void, along with the difference between these wavefields. This synthetic simulation demonstrates the progressive release of trapped seismic energy as resonant emissions, with attenuation effects excluded from the model. Panels (m)–(o) show the raw shot gathers for the respective wavefields and their differences. The ringing effect manifests itself as a long wave train.

an innovative neural network architecture known as symmetric autoencoders (SymAE) (Bharadwaj et al., 2020, 2022, 2024). SymAE enables the resynthesis of seismic data, a process referred to as *redatuming*, which entails encoding and altering the latent representation prior to decoding to generate *virtual* data. Unlike traditional autoencoders, SymAE facilitates a physically meaningful redatuming of seismic data–considering both source and subsurface effects–due to its disentangled latent representation, wherein coherent information (related to the subsurface) in seismic waveforms is separated from other nuisance effects (related to the source signature). The method of SymAE introduced in this paper marks a notable advancement from earlier versions. This is the first application of SymAE for training with cross-correlated noise data. This novel approach can be regarded as a generalization of nonlinear stacking techniques, where neural networks replace the ad-hoc nonlinear weighting schemes used in stacking methods.

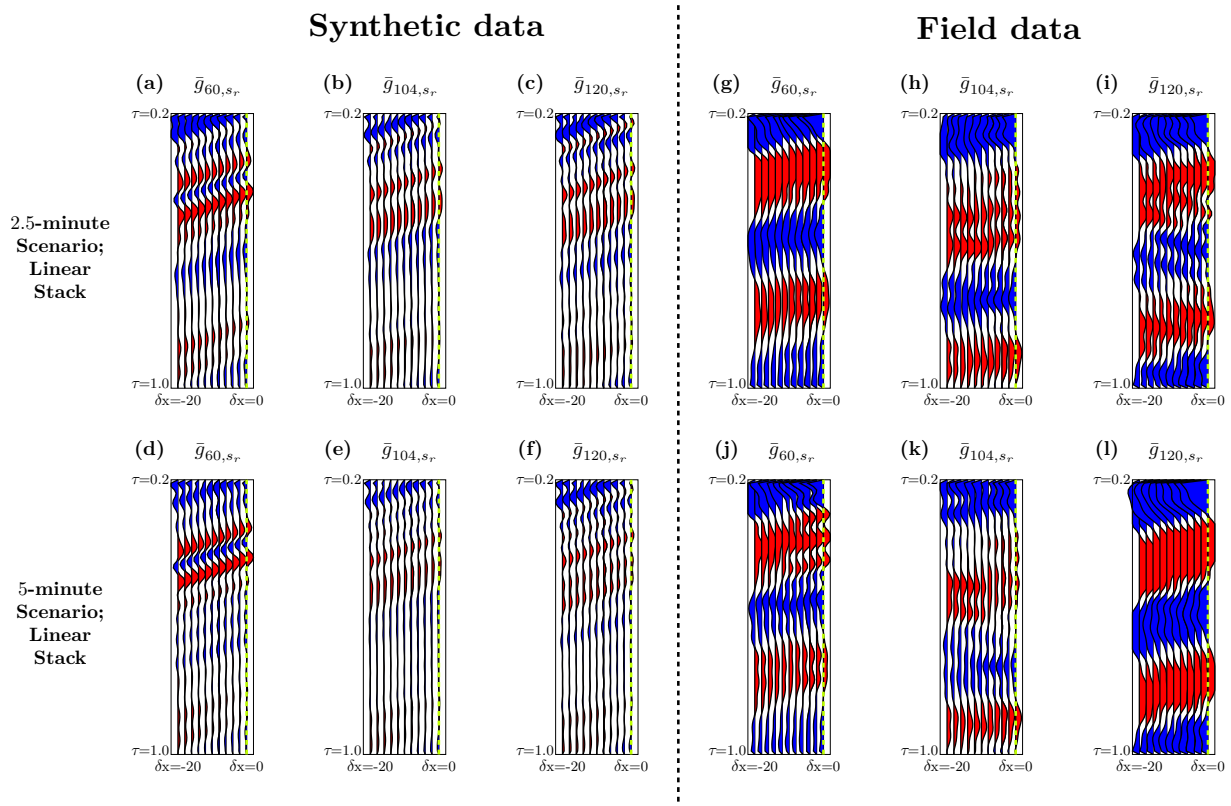

Figure 6: Ringing patterns: Stacked CCNs for the 2.5-minute and 5-minute durations in both synthetic and field ringing experiments. Only positive time lags are shown to emphasize that ringing patterns should appear as continuous long-wave trains. The plots indicate that cross-correlation residuals can be misinterpreted as ringing patterns.

This section encompasses the preparation of the training data set, the understanding of coherent and nuisance features, the architectural specifics of SymAE, and the redatuming process aimed at capturing backscattering and ringing effects in the cross-correlated seismic data. Table 1 provides a summary of the key mathematical notations relevant to subsequent discussions.

4.1 Training data set

The fundamental data structure utilized by SymAE consists of a collection of multiple datapoints. The processed seismic data are systemically organized into these datapoints, each corresponding to a different reference receiver position. For n_r reference receiver positions, there exist n_r corresponding datapoints. SymAE is designed to process each datapoint separately, working on one individual datapoint at a time. The datapoints aimed at extracting backscatter correlations and ringing effects are structured as follows:

- **Datapoints for extracting backscatters:** The elements of each datapoint are CCNs with varying source signatures. For instance, the datapoint B_r corresponding to a reference receiver position r is expressed as:

$$B_r = \left[g^{k_1}_{r,s_1},\, g^{k_2}_{r,s_1},\, g^{k_3}_{r,s_2},\, g^{k_4}_{r,s_2},\, g^{k_5}_{r,s_3},\, g^{k_6}_{r,s_4}, \cdots \right]. \tag{8}$$

- **Datapoints for extracting ringing effects:** Here, each datapoint encapsulates CCNs associated with the closest source location. Thus, the datapoint Y_r for a reference receiver position r with the closest source location s_r is formulated as:

$$Y_r = \left[g^{k_1}_{r,s_r},\, g^{k_2}_{r,s_r},\, g^{k_3}_{r,s_r}, \cdots \right]. \tag{9}$$

For clarity in our subsequent discussions, we will refer to the general datapoint as X_r, which may represent either B_r or Y_r, depending on the feature being extracted. The collection of these datapoints is known as the data set, denoted by $\{X_r\}$, and is used during the training of SymAE.

Notation	Description	Dimension/Quantity
	Data Representation	
$g^k_{r,s}$	Individual CCN / flattened individual CCN	$\mathbb{R}^{n_\tau \times n_{tr}}$ / $\mathbb{R}^n$
X_r	Datapoint: collection of n_w flattened CCNs for the same reference receiver position	$\mathbb{R}^{n \times n_w}$
$X_r[i]$	i-th flattened CCN (simply referred as CCN) in datapoint X_r	$\mathbb{R}^n$
n_τ	Number of time lags in each CCN	Scalar, $n_\tau \in \mathbb{N}$
n_{tr}	Number of traces (receivers) in each CCN	Scalar, $n_{tr} \in \mathbb{N}$
n	$n_\tau \times n_{tr}$	Scalar, $n \in \mathbb{N}$
n_w	Number of CCNs in a datapoint	Scalar, $n_w \in \mathbb{N}$
n_r	Number of reference receiver positions	Scalar, $n_r \in \mathbb{N}$
	Feature Representation	
P_r	Coherent subsurface code: consistent characteristics within a datapoint X_r	$\mathbb{R}^p$
$N^k_{r,s}$	Nuisance code: CCN-specific variability	$\mathbb{R}^q$
	SymAE Architecture	
`PEnc` (`PEnc`$_1$ and `PEnc`$_2$)	Symmetric Encoder: designed to extract coherent subsurface code from all CCNs within a datapoint	$\mathbb{R}^n \to \mathbb{R}^p$
`NEnc`	Nuisance Encoder: designed to extract CCN-specific nuisance code	`NEnc` : $\mathbb{R}^n \to \mathbb{R}^q$
`Dec`	Decoder: reconstructs CCN from coherent subsurface and nuisance codes	`Dec` : $\mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^n$
p	Dimension of the coherent code	Scalar, $p \in \mathbb{N}$
q	Dimension of the nuisance code	Scalar, $q \in \mathbb{N}$
	Training and Evaluation	
$\hat{X}_r$	Reconstructed datapoint	$\mathbb{R}^{n \times n_w}$
L	Expected reconstruction error	Scalar, $L \in \mathbb{R}_{\geq 0}$

Table 1: Summary of key mathematical notations

4.2 Features

Understanding the features present in the CCNs within the data set is essential. We analyze these features by considering a convolutional model with a point noise source. Each CCN can be expressed as a discrete convolution between the cross-correlated Green function, g^0 and the noise source autocorrelation w^k. It is important to note that g^0 remains constant in different time windows. As previously mentioned, each CCN is normalized at $(\delta x = 0, \tau = 0)$. Therefore, the normalized CCN can also be expressed as:

$$g^k_{r,s}(\delta x, \tau) = (g^k_{r,s}(0,0))^{-1} w^k(0)\, g^0_{r,s}(\delta x, \tau) + (g^k_{r,s}(0,0))^{-1} \sum_{t \neq 0} w^k(t)\, g^0_{r,s}(\delta x, \tau - t). \tag{10}$$

We can ignore the variation of $(g^k_{r,s}(0,0))^{-1} w^k(0)$ relative to k and substitute it with a constant c_1. Hence, Equation 10 can be approximated as:

$$g^k_{r,s}(\delta x, \tau) \approx \underbrace{c_1 g^0_{r,s}(\delta x, \tau)}_{\text{coherent term}} + \underbrace{(g^k_{r,s}(0,0))^{-1} \sum_{t \neq 0} w^k(t)\, g^0_{r,s}(\delta x, \tau - t)}_{\text{nuisance terms}}. \tag{11}$$

In this formulation, every CCN is delineated by two components:

- Coherent features: These latent features encapsulate coherent subsurface characteristics shared by all CCNs within a datapoint. They are hypothesized to correspond to the term $c_1\, g^0_{r,s}$, which represents the coherent correlations specific to the reference receiver position r.
- Nuisance features: These latent features capture variations that deviate from the coherent subsurface characteristics. They are hypothesized to correspond to the nuisance terms, represented by $(g^k_{r,s}(0,0))^{-1} w^k\, g^0_{r,s}$ at nonzero time steps, which are specific to the noise window indexed by k. This component reflects the influence of individual source signatures or other factors introducing variability in CCNs.

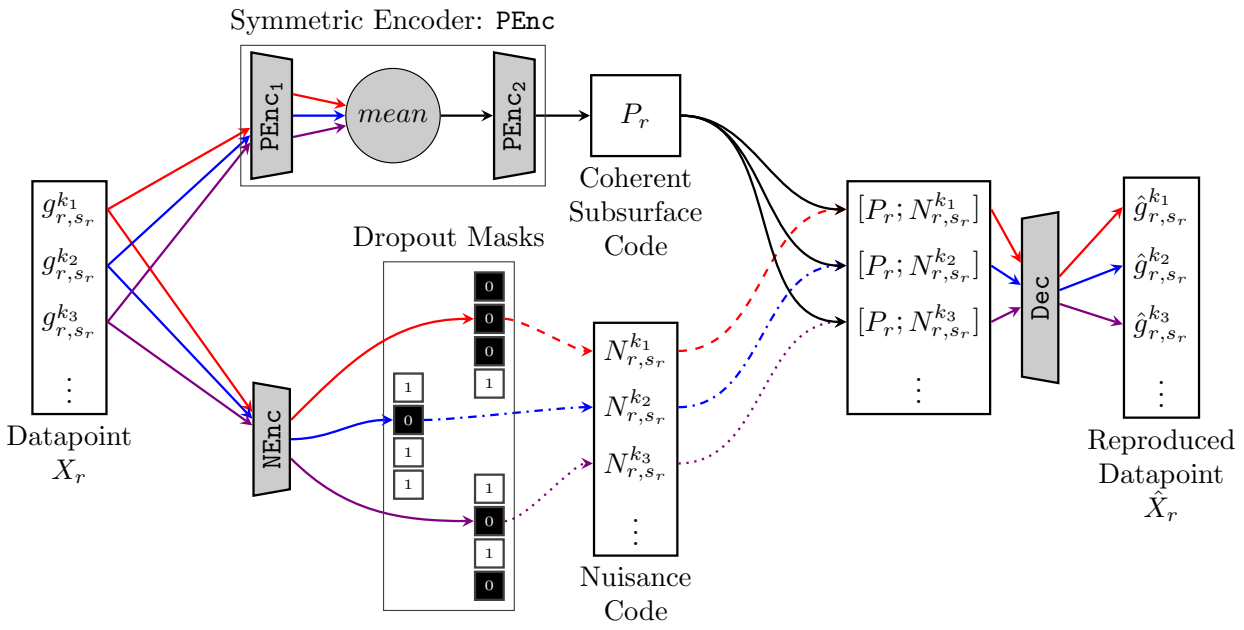

Figure 7: The architecture of a symmetric autoencoder that disentangles subsurface information, coherent across multiple CCNs, from the CCN-specific source effects (denoted by different arrow colors). Coherent information can only propagate through the network via solid arrows with dropout masks inhibit its propagation. Dashed arrows represent excluded coherent features, while varying arrow styles indicate randomness in nuisance data. Consequently, the autoencoder effectively disentangles coherent information from nuisance variations within the latent space.

In the case of extended sources with temporally correlated signatures, more terms might contribute to the coherent features in Equation 11.

4.3 SymAE architecture

SymAE is a statistically constrained autoencoding architecture that disentangles coherent subsurfaces and nuisance effects using permutation invariance and stochastic regularization. The architectural details of SymAE are depicted in Figure 7, which features two encoders dedicated to encoding subsurface and nuisance information, respectively, and a decoder tasked with reconstructing CCNs in an almost lossless manner utilizing both codes. These components are described in detail as follows:

- **Symmetric Encoder:** The symmetric encoder, $\texttt{PEnc}$, is designed to extract coherent features shared by all CCNs within X_r. Essentially, the architecture of $\texttt{PEnc}$ ensures that, when applied to X_r, it generates a single coherent subsurface code $P_r \in \mathbb{R}^p$ that represents only coherent subsurface features, excluding CCN-specific nuisance information. However, as indicated by the universal approximation theorem (Hornik et al., 1989; Cybenko, 1989), an unconstrained $\texttt{PEnc}$ would capture both coherent and nuisance information. To address this, we impose a permutation invariance constraint on $\texttt{PEnc}$. This constraint ensures that extracted coherent features are independent of the ordering of the CCNs in X_r. Mathematically, this means that
$$P_r = \texttt{PEnc}(X_r) = \texttt{PEnc}(\Pi(X_r)). \tag{12}$$
Here, Π is a random permutation function that reorders the elements of X_r according to a random permutation Π of the indices $\{1, 2, \ldots, n_w\}$, i.e.,
$$\Pi(X_r) = [X_r[\Pi(1)], X_r[\Pi(2)], \ldots, X_r[\Pi(n_w)]]\,, \tag{13}$$
where Π is a bijection $\Pi : \{1, 2, \ldots, n_w\} \to \{1, 2, \ldots, n_w\}$ that shuffles the indices of X_r randomly.
In practice, we adopt the network architectures outlined in Zaheer et al. (2017) that can approximate functions with permutation-invariance properties. Specifically, we employ the mean pooling operation to enforce permutation invariance. This is achieved through a series of transformations. First, a neural network $\texttt{PEnc}_1$ is applied to each element of X_r. The transformed elements are then aggregated using a mean pooling operation, and finally, a second network $\texttt{PEnc}_2$ generates the coherent latent code P_r:
$$P_r = \texttt{PEnc}(X_r) = \texttt{PEnc}_2(\text{mean}(\texttt{PEnc}_1 \bullet (X_r))). \tag{14}$$

In this formulation, $f \bullet (X)$ denotes the application of a function f to each element of the vector X. The mean pooling operation maintains symmetry with regard to the ordering of CCNs within a datapoint, thereby ensuring the necessary permutation invariance for `PEnc`, as described in Equation 12. Since the encoder is constrained to capture only features that are coherent across all CCNs, it cannot effectively represent CCN-specific features, such as source-signature effects, without a significant loss of information.

- **Nuisance Encoder:** The second encoder, termed nuisance encoder, `NEnc`, encodes the CCN-specific information for each element of the datapoint X_r. Each element $\in \mathbb{R}^n$ is mapped to an intermediate feature space $\mathbb{R}^q$, where q represents the dimension of the nuisance code. Unlike the symmetric encoder, `NEnc` operates without any architectural constraints. However, if the output of `NEnc` is not constrained during training, a significant concern arises: the decoder `Dec` may prioritize information from the nuisance encoder over the coherent component P_r during reconstruction. To address the potential issue of the decoder neglecting coherent information, SymAE incorporates stochastic regularization through dropout masks during training. This regularization employs Bernoulli dropouts (Srivastava et al., 2014) with a probability of $\mathtt{p} = 0.8$. Unlike the traditional use of dropouts for regularizing functions approximated by deep neural networks, SymAE employs dropouts to: 1. prevent the nuisance encoder from inadvertently capturing coherent subsurface information, and 2. encourage the decoder to prioritize and extract the most meaningful information from the symmetric encoder, as the dropout intentionally obscures certain information. Thus, the nuisance code of a particular CCN, say $g^{k_1}_{r,s_r}$, is

$$N^{k_1}_{r,s_r} \quad = N_{r,s_r}[1] \quad = \mathrm{dropout}(\mathtt{NEnc}(g^{k_1}_{r,s_r}), \mathtt{p}), \tag{15}$$

  which represents the information specific to the source location s_r and the noise in the first time window k_1. Dropout masks are utilized exclusively during the training phase. After training, these masks are deactivated by setting $\mathtt{p} = 0$ to produce the complete $N_{r,s_r}[\cdot]$ code. Evidently, dropout is implemented on each CCN in a datapoint.

- **Decoder:** A network, `Dec`, nonlinearly decodes the latent codes to reconstruct the datapoint as:

$$\hat{X}_r = \mathtt{Dec} \bullet \left(\left[[P_r;\, N^{k_1}_{r,s_r}],\, [P_r; N^{k_2}_{r,s_r}],\, [P_r; N^{k_3}_{r,s_r}],\, \cdots \right] \right), \tag{16}$$

  where $[P;\, N]$ denotes concatenation of codes P and N.

This design of SymAE allows for the simultaneous learning of coherent and nuisance features without prior knowledge of their specific characteristics or distributions.

4.4 Training SymAE

In this study, the encoders (`PEnc` and `NEnc`) and the decoder (`Dec`) are parameterized using fully connected feedforward networks. These networks are trained simultaneously by minimizing the reconstruction loss,

$$L = \frac{1}{n_r \times n_w} \sum_{r=1}^{n_r} \sum_{i=1}^{n_w} \|X_r[i] - \hat{X}_r[i]\|^2, \tag{17}$$

where n_r denotes the total number of datapoints (i.e., reference receiver positions). The use of permutation invariance and stochastic regularization establishes a distinct information flow in SymAE, as depicted in Figure 7, characterized by: 1. a continuous flow of coherent information from the symmetric encoder, represented by the solid black arrow, and 2. outputs from the nuisance encoder that comprise randomly obscured information, with varying line styles (such as dashed, dash-dotted, and dotted) indicating the exclusion of coherent features. Through training, the symmetric encoder becomes proficient in extracting coherent features, as these are consistently useful across all reconstructions. Meanwhile, the nuisance encoder specializes in capturing CCN-specific information, which cannot be encoded by the symmetric encoder but is essential for accurately reconstructing each CCN.

The dimensions p and q of the latent codes are user-defined hyperparameters. The optimal values of these dimensions were chosen by training on the synthetic data and adopted for the field data. Furthermore, keeping q small limits the expressiveness of the nuisance encoder, which further assists the decoder in prioritizing the extraction of relevant coherent features. Effective training of SymAE requires a sufficient number of datapoints, corresponding to reference receiver positions, to prevent overfitting. This ensures that the model learns a robust representation in the latent space, effectively distinguishing between coherent and nuisance features.

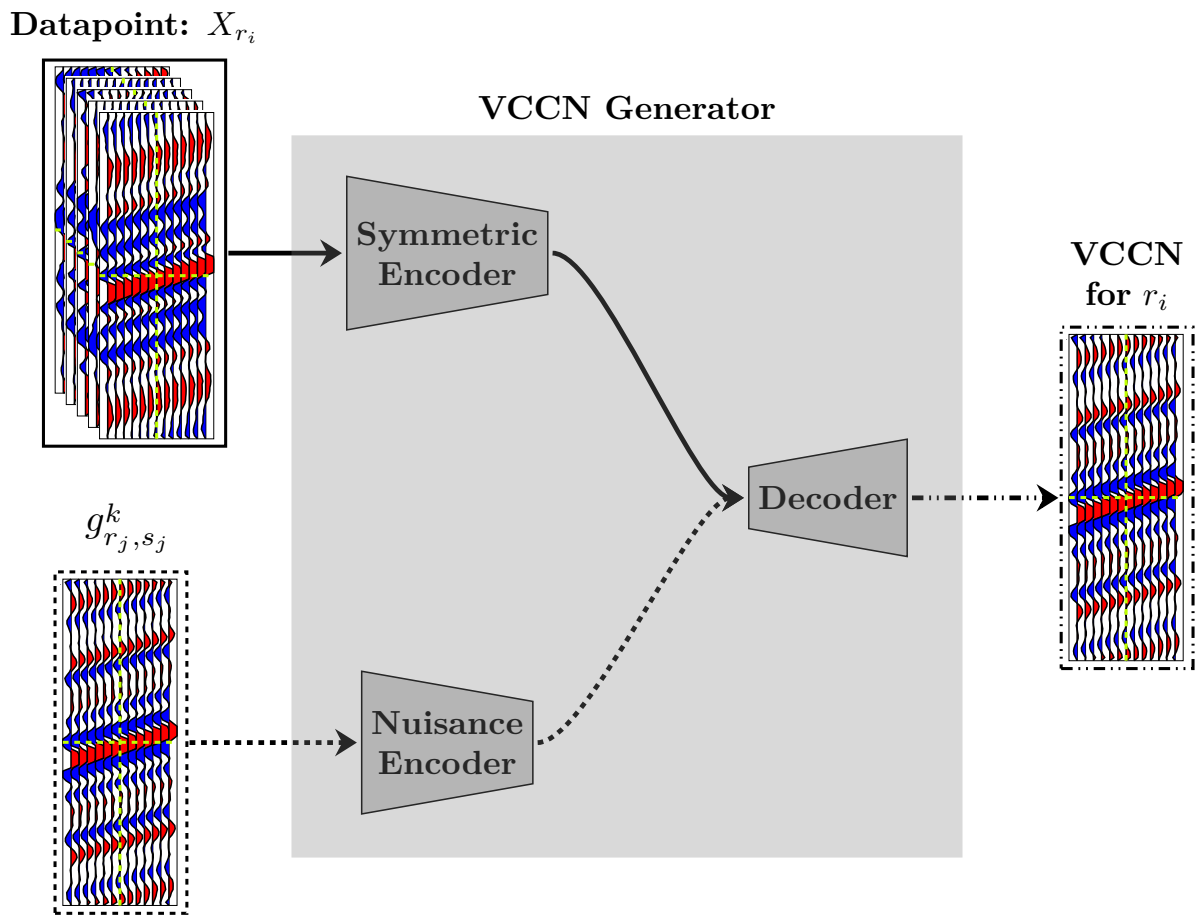

Figure 8: Illustration of the redatuming process used to generate a VCCN via Equation 18. The arrows depict the flow of information during this process. The symmetric encoder extracts the coherent latent code from the datapoint X_{r_i}, while the nuisance encoder encodes the nuisance latent code from the CCN $g^k_{r_j,s_j}$. These latent codes are concatenated to form an altered latent code, which is decoded by the decoder to generate the VCCN for r_i. All CCNs and the VCCN depicted in this plot share the identical axes as those depicted in Figure 4.

4.5 Redatuming

The trained SymAE model can be utilized for redatuming, a process that generates virtual cross-correlation gathers (VCCNs) by decoding altered latent code. This altered latent code is formed by concatenating the coherent latent code from a single datapoint (specifically, a reference receiver position) with the nuisance latent code derived from a specific CCN associated with another datapoint. The decoder then processes the concatenated latent code to synthesize a VCCN, which lies outside the training data set. This process is illustrated in Figure 8 for generating a VCCN at the reference receiver position r_i. The symmetric encoder extracts the coherent latent code P_{r_i} from the datapoint X_{r_i}, while the nuisance encoder encodes the nuisance latent code $N^k_{r_j,s_j}$ from the CCN $g^k_{r_j,s_j}$. These latent codes are concatenated and then decoded to produce the VCCN for r_i. Once the symmetric encoder, nuisance encoder, and decoder have been trained, they can be employed to build the VCCN generator. Mathematically, the resulting VCCN can be expressed as:

$$\texttt{Dec}\Big(\Big[P_{r_i};\, N^k_{r_j,s_j}\Big]\Big) \approx c_1\, g^0_{r_i,s_j} + (g^k_{r_j,s_j}(0,0))^{-1} \sum_{t\neq 0} w^k(t)\, g^0_{r_j,s_j}\,(\tau - t). \tag{18}$$

In the RHS, the first term represents the coherent information from the reference receiver position r_i, while the second term captures the nuisance variations introduced by the noise source indexed by k at a different source location s_j.

Earlier research on SymAE applied redatuming to generate virtual data and showcased its applications. However, the VCCN still contains nuisance information, as shown in Equation 18, making it challenging to identify void signatures

within the generated VCCN. Therefore, removing nuisance information is crucial. This paper suggests generating differential virtual cross-correlations (DVCCs) for each reference receiver position, employing one of two approaches:

- **Subtracting a baseline VCCN**: This approach involves subtracting two VCCNs, corresponding to two separate data points, both produced using identical nuisance information. The subtraction extracts the difference in subsurface information between those two reference receiver positions. To demonstrate this mathematically, let us consider two datapoints X_{r_l} and X_{r_m} corresponding to the reference receiver positions, r_l and r_m, respectively. Initially, we determine the coherent subsurface codes, P_{r_l} and P_{r_m}, for these datapoints by applying Equation 14. Likewise, the identical nuisance code $N^k_{r_j,s_j}$ of a specific CCN $g^k_{r_j,s_j}$ is formulated using Equation 15. By employing the VCCN generator, as depicted in Figure 8, the two VCCNs for r_l and r_m can be created. Following the approach of Equation 18, these two VCCNs can then be mathematically represented as:

$$\texttt{Dec}\Big(\Big[P_{r_l};\, N^k_{r_j,s_j}\Big]\Big) \approx c_1\, g^0_{r_l,s_j}\,(\delta x,\tau) + (\, g^k_{r_j,s_j}(0,0))^{-1} \sum_{t\neq 0} w^k(t)\, g^0_{r_j,s_j}\,(\tau - t), \tag{19}$$

$$\texttt{Dec}\Big(\Big[P_{r_m};\, N^k_{r_j,s_j}\Big]\Big) \approx c_1\, g^0_{r_m,s_j}\,(\delta x,\tau) + (\, g^k_{r_j,s_j}(0,0))^{-1} \sum_{t\neq 0} w^k(t)\, g^0_{r_j,s_j}\,(\tau - t). \tag{20}$$

  Since the identical nuisance code is employed during the generation of these VCCNs, the nuisance terms in RHS of Equations 19 and 20 can be easily eliminated. This elimination is achieved by subtracting the VCCNs, resulting in a DVCC:

$$v_{r_l - r_m}(\delta x,\tau) \approx c_1\, g^0_{r_l,s_j}\,(\delta x,\tau) - c_1\, g^0_{r_m,s_j}\,(\delta x,\tau). \tag{21}$$

  The DVCC captures the difference in coherent subsurface features between the two reference receiver positions, r_l and r_m, with the nuisance effects eliminated. This enables the identification of void signatures in the subsurface. As demonstrated in Figure 9, we apply this approach to synthetic data from a backscattering experiment. Here, two VCCNs are generated for reference receiver positions $r = 100\,\text{m}$ and $r = 44\,\text{m}$, using the identical nuisance code from a specific CCN $g^1_{10,s_{10}}$. The subtraction of these VCCNs removes the nuisance, resulting in the DVCC v_{100-44}. This DVCC reveals the backscatter correlations, as shown in the figure. Consequently, this approach enables the extraction of differential subsurface characteristics between a baseline reference receiver and a reference receiver position of interest, thereby facilitating the identification of void signatures.

- **Subtracting an average VCCN**: Rather than relying on a specific reference receiver position as a baseline, this alternative approach leverages VCCNs from all reference receiver positions within the data set. In this method, a mean VCCN is calculated by averaging the VCCNs generated with identical nuisance information. This averaged VCCN is then subtracted from the VCCN at the specific reference receiver position of interest, thereby removing the identical nuisance effects and producing the DVCC. This process is illustrated in Figure 10 using the same synthetic data from the backscattering experiment. Instead of selecting a single VCCN, such as the one corresponding to $r = 44\,\text{m}$, an averaged VCCN is computed by averaging the VCCNs from reference receiver positions spanning $r = 14\,\text{m}$ to $r = 144\,\text{m}$. This averaged VCCN is then subtracted from the VCCN for $r = 100\,\text{m}$ to generate the DVCC v_{100}. Similar to the previously discussed DVCC v_{100-44}, backscatter correlations are also evident in v_{100}. This approach has the advantage that the DVCC v_{100} does not rely on the choice of a specific baseline reference receiver. Mathematically, the averaged VCCN with the identical nuisance code $N^k_{r_j,s_j}$ can be expressed as:

$$\tfrac{1}{N}\textstyle\sum_{i=1}^{N} \texttt{Dec}\Big(\Big[P_{r_i};\, N^k_{r_j,s_j}\Big]\Big) \approx \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} c_1\, g^0_{r_i,s_j}(\delta x,\tau) + (g^k_{r_j,s_j}(0,0))^{-1} \sum_{t\neq 0} w^k(t)\, g^0_{r_j,s_j}\,(\tau - t). \tag{22}$$

  Here, the initial term represents the mean of the coherent information distributed across all reference receiver positions. Under the assumption of lateral homogeneity with the exception of the void, a majority of the reference receivers do not detect the void. Consequently, as the number of reference receivers is increased, the initial term will predominantly reflect the characteristics of the laterally homogeneous medium. The VCCN for a specific reference receiver position, denoted as r_l, is provided in Equation 19. By subtracting the averaged VCCN (Equation 22) from the VCCN corresponding to r_l, the DVCC for r_l is obtained as follows:

$$v_{r_l}(\delta x,\tau) \approx c_1\, g^0_{r_l,s_j}\,(\delta x,\tau) - \frac{1}{N}\sum_{i=1}^{N} c_1\, g^0_{r_i,s_j}\,(\delta x,\tau). \tag{23}$$

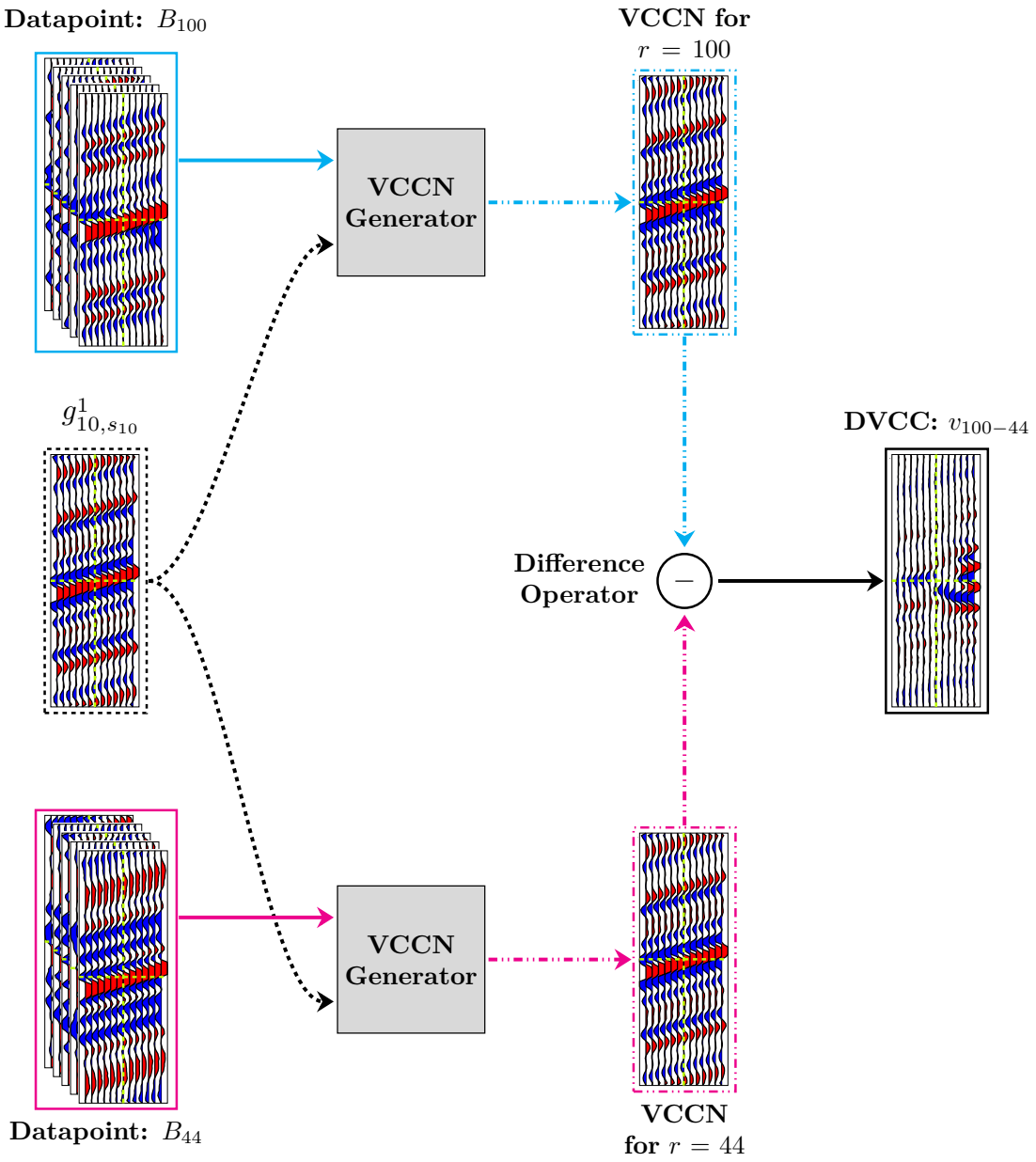

Figure 9: Illustration of the process outlined in Equation 21 to generate the DVCC utilizing synthetic data obtained from the backscattering experiment. VCCNs are calculated for reference receiver positions $r = 100\,\text{m}$ and $r = 44\,\text{m}$, applying the same nuisance code from $g^1_{10,s_{10}}$. By subtracting these VCCNs, the nuisance is eliminated, yielding the DVCC v_{100-44}. This DVCC reveals backscatter correlations, which highlight differential subsurface characteristics. The arrows indicate the flow of information. All CCNs, VCCNs, and the DVCC depicted in this plot share the identical axes as those depicted in Figure 4.

Synthetic experiments in the Figures 9 and 10 demonstrate that both approaches are effective in isolating backscatter correlations. Nonetheless, for field experiments, the second method proved more effective in extracting ringing signatures, although the cause of this difference needs further investigation. Consequently, we employed the first approach for backscattering and the second approach for ringing, to ensure the accurate extraction of void signatures, as discussed in the subsequent results section.

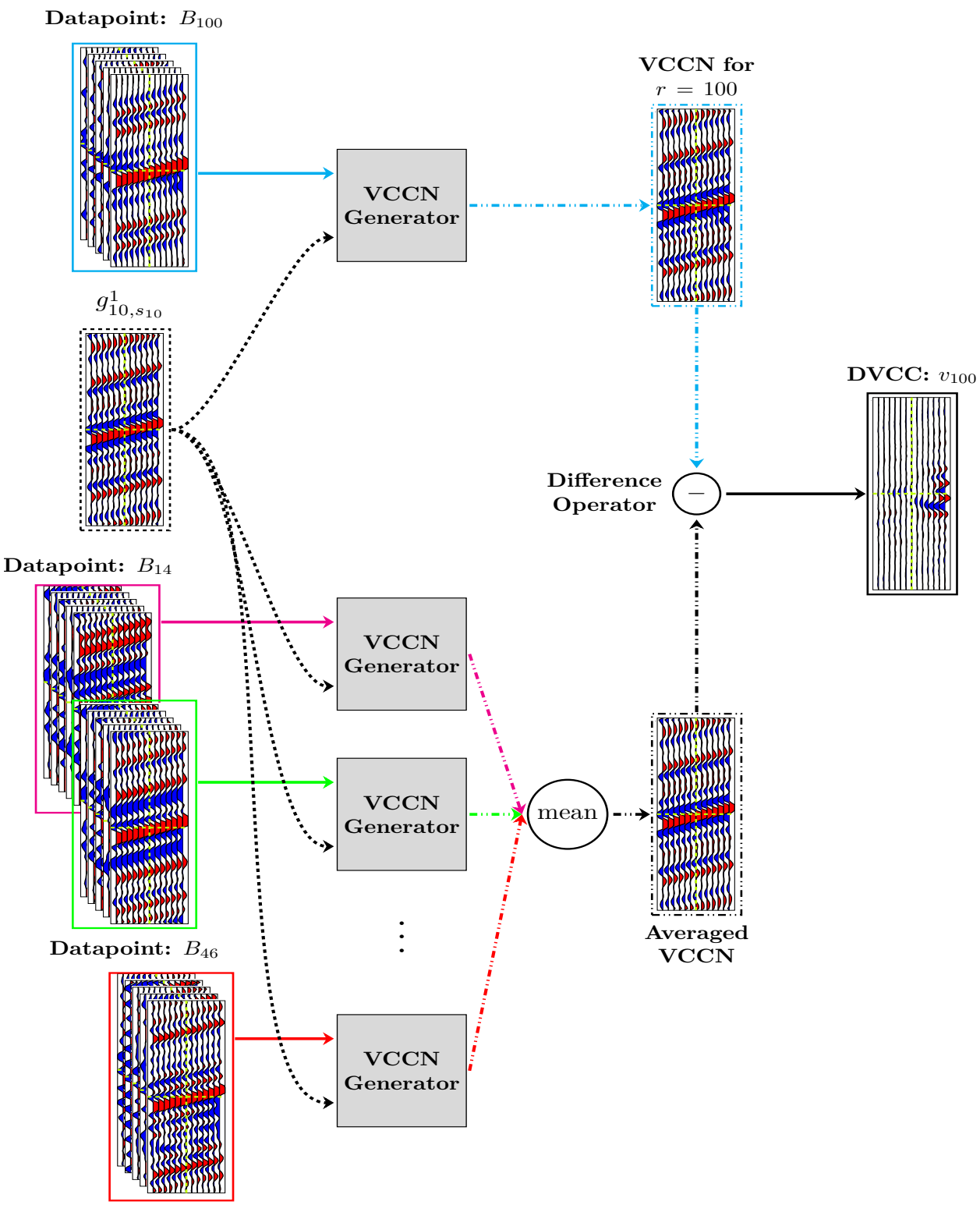

Figure 10: Illustration of the redatuming process used to compute the DVCC v_{100} by subtracting an averaged VCCN. The averaged VCCN is calculated as the mean of VCCNs generated for all reference receiver positions ($r = 14\,\text{m}$ to $r = 144\,\text{m}$). The DVCC v_{100} captures coherent subsurface information (here, backscatter correlations) specific to $r = 100\,\text{m}$ while eliminating nuisance effects. All CCNs, VCCNs, and the DVCC depicted in this plot share the identical axes as those depicted in Figure 4.

5 Results

5.1 Backscatter

Within the framework of the backscattering experiment, the outcomes for three distinct reference receivers $\in \{60,\ 100,\ 122\}$ m are delineated in Figure 4, with $r = 44$ m serving as the baseline reference receiver. These reference receivers are selected based on their proximity, the presence of a void at approximately 100 m, and the position beyond the void at 122 m. We anticipate that the difference between the stacked cross-correlated data, $\bar{g}_{44,s}$ and $\bar{g}_{60,s}$, to converge towards zero, provided that the subsurface conditions at these reference receiver locations exhibit similarity, as depicted in Fig. 2. However, due to the cross-correlation residuals, as demonstrated in Figures 4a and 4d, the discrepancy remains substantial in 2.5-minute and 5-minute scenarios, suggesting that extended durations are necessary for cross-correlations to achieve convergence. Similarly, for $\bar{g}_{100,s}$, where the receivers capture both direct and backscattered surface waves, we expect to observe backscatter correlations. However, these backscatter correlations were obscured by the cross-correlation residuals (Figures 4b and 4e). Our focus is now directed towards training SymAE with the data set $\{B_r\}$ and redatuming to generate virtual cross-correlation gathers for reference receiver positions ranging from $r = 14$ m to $r = 144$ m. Utilizing Equation 21, we generated DVCCs to monitor variations in subsurface properties between the baseline reference receiver ($r = 44$ m) and the other reference receivers $\in \{60,\ 100,\ 122\}$ m. The expectation of no variation between the reference receivers at 60 m and 44 m is validated (Figures 4g and 4j). Additionally, we verified the similarity in subsurface properties for multiple reference receiver positions up to approximately $r = 100$ m, where we expect the presence of backscattered surface waves. The DVCCs in Figures 4h and 4k distinctly depict these backscattered correlations ($_{\leftarrow,\rightarrow}g^k_{r,s}$ and $_{\rightarrow,\leftarrow}g^k_{r,s}$), demonstrating the efficacy of our method. The depth sensitivity of the method has also been evaluated. In this experiment, the dominant frequency of the backscattering was found to be 12.5 Hz, with V_s in the range of $300 - 330$ m/s at the site, resulting in a backscatter wavelength of approximately $24 - 26.4$ m. Previous studies (Sloan et al., 2016, 2021) suggest that depth sensitivity typically ranges from 0.33 to 0.5 times the wavelength, meaning the detectable depth of the void in our study would be in the range of $8 - 13.2$ m, consistent with the known depth of approximately 10 m. Moving further from the void, for $r = 122$ m, theoretically, we still expect to observe backscattered energy. However, it was not observed in this example (Figures 4i and 4l). Although the exact reason remains unknown, we attribute this to the attenuation of weaker backscattered energy — the backscattered signals were observed upto $r = 110$ m.

5.1.1 2D synthetic experiment validation

We performed a set of elastic simulations by employing a Ricker wavelet within a constant velocity medium containing a void. In order to create a noise data set to validate field observations, the simulation data were convolved with vehicular noise similar to that depicted in Figure 1b. The simulations were performed in the finite-difference-time domain (FDTD) using Geophyinv.jl package. True cross-correlated Green's functions (Figures 11a–11c) were also generated for an impulsive source using a Ricker wavelet. The results of this synthetic data set exhibited a high degree of congruence with the field results and were further evaluated by calculating the mean squared error (MSE) between the normalized true cross-correlated Green functions and the results derived from linear stacking (Figures 11d–11i) and redatuming (Figures 11j–11l) across the 2.5-minute and 5-minute intervals. In contrast to the field experiment, we detected significant backscattered energy in the DVCCs (Figures 11i and 11l) for $r = 122$ m due to the non-attenuating medium used in our simulations.

5.2 Ringing

In addition to backscattering, the ringing pattern constitutes another distinct signature of the void. However, ringing patterns can be misinterpreted due to cross-correlation residuals, as discussed in the theory section (refer to Figure 6). By leveraging SymAE's capability to extract coherent information and eliminate cross-correlation residuals, we trained SymAE with the data set $\{Y_r\}$ and conducted redatuming to isolate the distinct ringing pattern. DVCCs (Equation 23), derived by subtracting the averaged VCCN, for states $\in \{60,\ 104,\ 120\}$ m with the nuisance information from the reference position $r = 44$ m are illustrated in Figures 12d–12f. In the absence of cross-correlation residuals, no ringing pattern is discernible for $r = 60$ m and $r = 104$ m, as depicted in Figures 12d and 12e, respectively. However, Figure 12f demonstrates that the ringing pattern is consistent for $r = 120$ m. It should be noted that the CCNs in the data set $\{Y_r\}$ contain all positive and negative time lags, but only results for higher positive lags are shown to highlight that the ringing pattern is a continuous wave train, best observed at higher lags. These observations strongly corroborate our hypothesis that cross-correlation residuals can lead to misinterpretations of the ringing pattern using linear stacking.

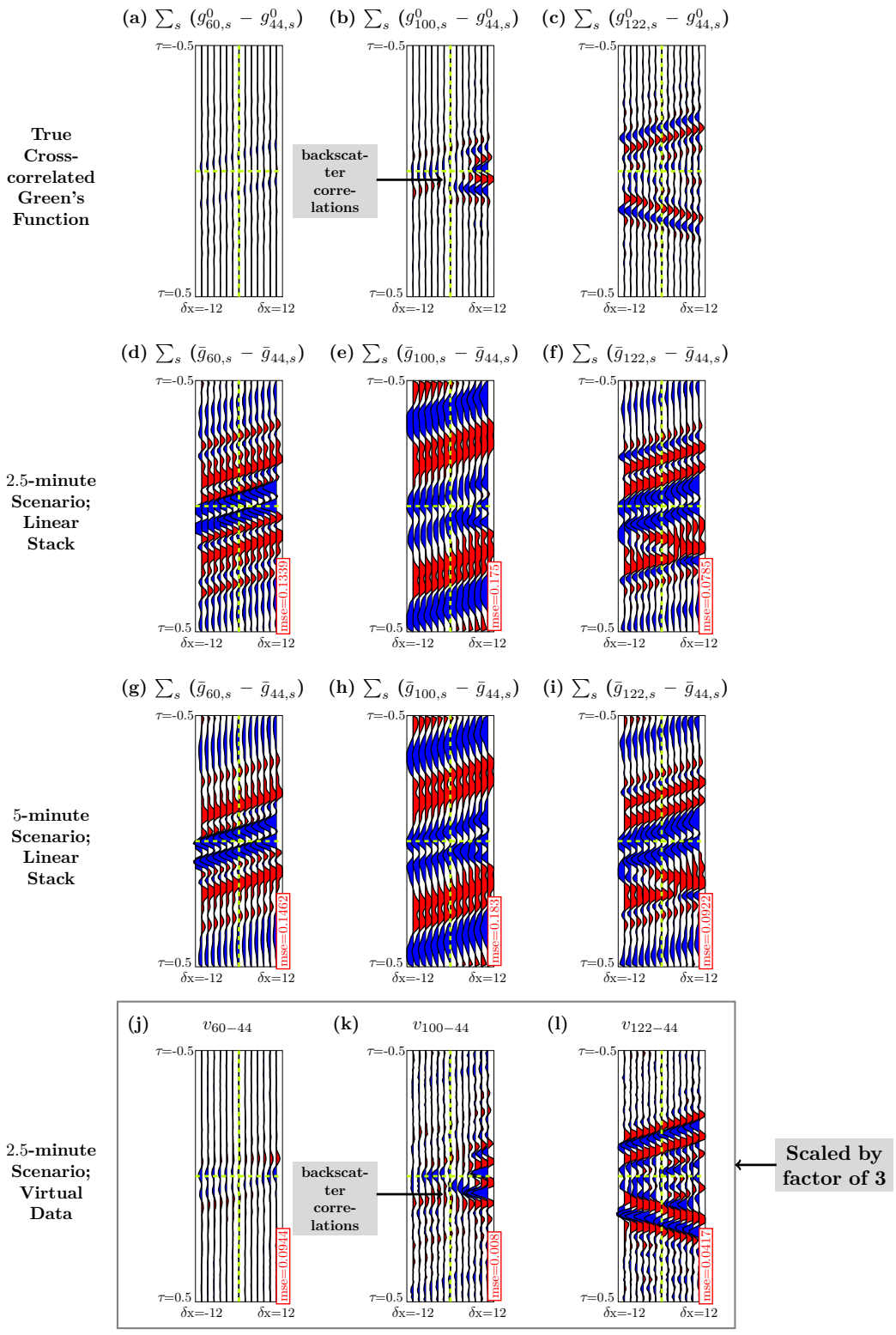

Figure 11: Similar to Figure 4, but depicting results from a synthetic experiment. Lower mean squared error (MSE) values (red boxes) indicate lower cross-correlation residuals — thanks to SymAE redatuming. The true cross-correlated Green's functions were resampled for MSE calculation purposes. During MSE computation, all stacked CCNs, DVCCs, and the true correlated Green's functions were subject to normalization.

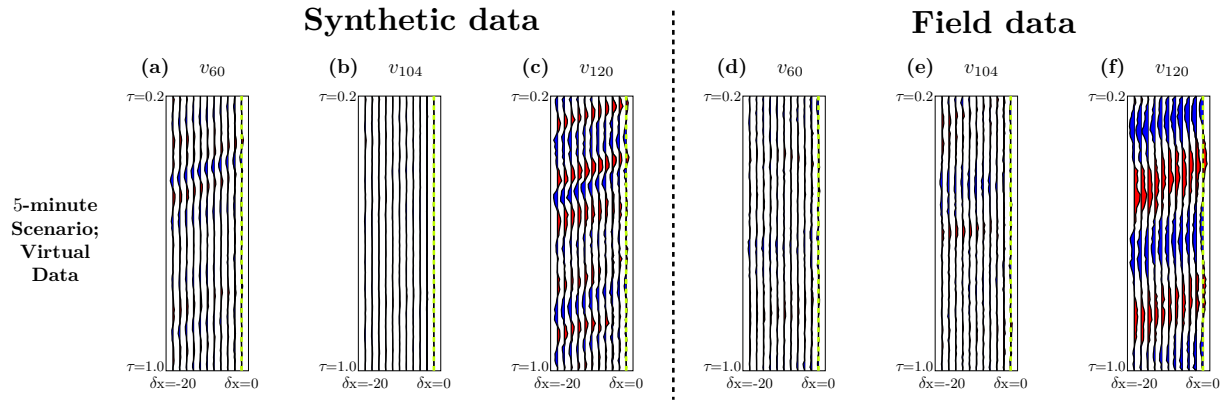

Figure 12: Ringing patterns as a void signature: DVCCs, derived by subtracting the averaged VCCN from those at $r \in \{60,\ 104,\ 120\}$ m subsequent to SymAE training and redatuming: (a) and (b) exhibit the absence of ringing patterns for $r = 60$ m and $r = 104$ m, and (c) demonstrates the persistence of the ringing pattern for $r = 120$ m for the synthetic data. Panels (d)-(f) present the same analysis for the field data. DVCCs are scaled by a factor of 2 compared to stacked CCNs in Figure 6.

5.2.1 2D synthetic experiment validation

We generated a data set by convolving the measured vehicular noise with the data obtained in the ringing simulations discussed in Figure 5. The data were organized according to the method described to generate datapoints Y_r to extract ringing effects. SymAE was trained using this synthetic data set. It was anticipated that only the reference receiver at $r = 120$ m, being closest to the void when the source was above it, would exhibit a distinct ringing pattern similar to Figure 5o. However, in Figures 6a– 6f, similar patterns were observed in the 2.5-minute and 5-minute scenarios, making it difficult to detect the ringing pattern. This also suggests that the ringing pattern and cross-correlation residuals can be visualizing similar. Then, we analyzed the DVCCs (Equation 23), derived by subtracting the averaged VCCN, for states $\in \{60,\ 104,\ 120\}$ m with the nuisance information from the reference position $r = 44$ m. These DVCCs, shown in Figures 12a–12c, clearly demonstrated that SymAE was able to remove the cross-correlation residuals for $r = 60$ and $r = 104$ m and the ringing pattern is intact for $r = 120$ m. Moreover, a brief discussion on 3D synthetic simulations is provided in Appendix C.

6 Discussion

By leveraging a neural network-based nonlinear stacking approach, SymAE addresses the limitations of traditional linear stacking, significantly enhancing the S/N and revealing distinct void signatures, specifically backscattering and ringing patterns. Backscattering phenomena have been extensively studied, and their characteristics are well-documented in prior research. The relationship between backscattered wavelengths and void depth is particularly useful, allowing us to estimate the depth of the void at site 1 to be in the range of $8 - 13.2$ m, consistent with its known depth of approximately 10 m. This makes backscattering a reliable indicator for void detection and depth estimation. However, synthetic experiments suggest that backscattering signals are not unique to voids; similar effects can occur around high-velocity anomalies such as boulders and concrete foundations. These anomalies, if present, may create false positives. In this study, no such anomalies were detected at site 1, allowing a more straightforward interpretation of the backscatter signals. The phenomenon of ringing remains less well-understood due to the variety of potential contributing mechanisms. These include circumferential wave propagation along void boundaries (Korneev, 2009), internal acoustic modes (Rubin et al., 2014), and resonant emissions from energy trapped within the void (Schneider et al., 2017). This study discusses the ringing pattern as the resonant emission of trapped seismic energy. Additionally, we observed that ringing is evident primarily when the source is positioned close to or directly above the void, though this observation may be site specific. The ringing of the void depends on its nature, size, depth, and material properties. Our synthetic experiments for high-velocity anomalies such as boulders or concrete foundations indicate an absence of ringing. In contrast, ringing is observed for low-velocity air-filled or partially-filled voids. Thus, ringing can serve as a distinctive signature of low-velocity voids. Field data did not definitively reveal how the diameter of the void influences the detectability of backscattering and ringing, although synthetic tests, using low-frequency Ricker wavelet, suggest a minimum diameter of roughly 1.5 m is required for these signatures to be observed.

SymAE is particularly effective in identifying these void signatures by deriving the coherent subsurface details from short 5-minute intervals of seismic noise. Employing cross-correlated data facilitated the training of SymAE, as obtaining coherent information from unprocessed (raw) noise posed difficulties. In our study, SymAE successfully processed data over a survey length of 148 m in a single setting, with the capacity to handle larger data sets, allowing a wider area coverage per survey. SymAE's training process is efficient, typically taking $5 - 10$ minutes for a 148 m data set, although it requires a sufficient volume of datapoints to ensure robust learning and avoid overfitting. Furthermore, SymAE is a train-on-the-fly technique instead of relying on pre-training, which permits flexibility and the ability to train in any new environment.

7 Conclusion

In this paper, we introduced, 1. a vehicle source that generates lower frequencies to augment the depth of geophysical investigations, and 2. a novel application of a neural network architecture,named SymAE, to seismic noise, specifically designed to effectively separate coherent subsurface information from nuisance source effects. SymAE exhibits a pronounced enhancement in mitigating cross-correlation residuals, which present substantial challenges in the processing of short-duration vehicle seismic noise. Our methodology addresses the inherent limitations of conventional linear stacking by leveraging the non-linear nature of neural networks. SymAE successfully extracted both void signatures in approximately 80% of the investigated void locations. Combining SymAE with redatuming techniques, to eliminate nuisance effects, has advanced the detection and analysis of subsurface voids. This underscores SymAE's potential to significantly enhance S/N in geophysical data processing, particularly with shorter noise intervals, to detect subsurface features in urban environments. The ability of SymAE in mitigating cross-correlation residuals, explored in this study, has potential applications across diverse fields such as environmental monitoring, groundwater surveillance, and civil engineering. Future studies might concentrate on applying SymAE to short-duration ambient noise data in urban settings for rapid and precise near-surface assessments.

A MASW experiment

At site 1, we executed comprehensive MASW surveys utilizing a 20 kg sledgehammer as the seismic source. The data acquisition adhered to the ideal parameters noted in Park et al. (2002) and Park (2003). A 24-channel setup with vertical geophones spaced 2 m apart collected the data. To enhance the S/N, three hammer shots were stacked at each spot. Figure A.1 illustrates the dispersion image and the 1D inversion outcome for the stacked shot gather when the 12^{th} receiver was positioned over the void (i.e., at $x = 100$ m). We expected to observe a low-velocity contrast in the inversion due to the characteristics of an air-filled void, yet no such contrast appeared. Furthermore, the dispersion image lacked low-frequency components, underscoring the necessity of deploying a low-frequency source for successful void detection.

B Choice of δx

The selection of δx differs between the backscattering and ringing experiments, as it is determined by the nature of the void signatures being extracted. In the backscattering experiment, δx ranges from negative to positive values because the backscattered signals often exhibit a hyperbolic pattern, necessitating a broad range. In contrast, for the ringing experiment, δx is limited to negative values (i.e., decreasing x in Figure 2), as the ringing effect is predominantly observed in receivers located to the left of the reference receiver. Moreover, δx must be chosen such that the flattened dimension n (outlined in Table 1) of the flattened CCN stays adequately low, ideally maintaining the vector dimension under 4000. This limitation is critical to prevent the overfitting of fully connected feedforward networks used in the SymAE model.

C Synthetic experiment: 3D modeling of void

In addition to conducting synthetic experiments in 2D, where the void was modeled as a circular anomaly within an elastic and homogeneous medium, we extended our analysis to a three-dimensional (3D) framework. In this 3D experiment, we modeled the void as a cuboid within a 3D acoustic homogeneous medium. This shift to acoustic rather than elastic modeling was necessitated by computational limitations, specifically the memory and time constraints associated with elastic modeling in 3D. The synthetic model comprises a 3D medium with dimensions 250 m x 300 m x 100 m in the x-, y-, and z- directions, respectively. The cuboidal void, measuring 2 m x 300 m x 2 m, is located at a depth of 10 m along the z-axis and extends parallel to the y-axis. The x-position of the void is 100 m. The acquisition

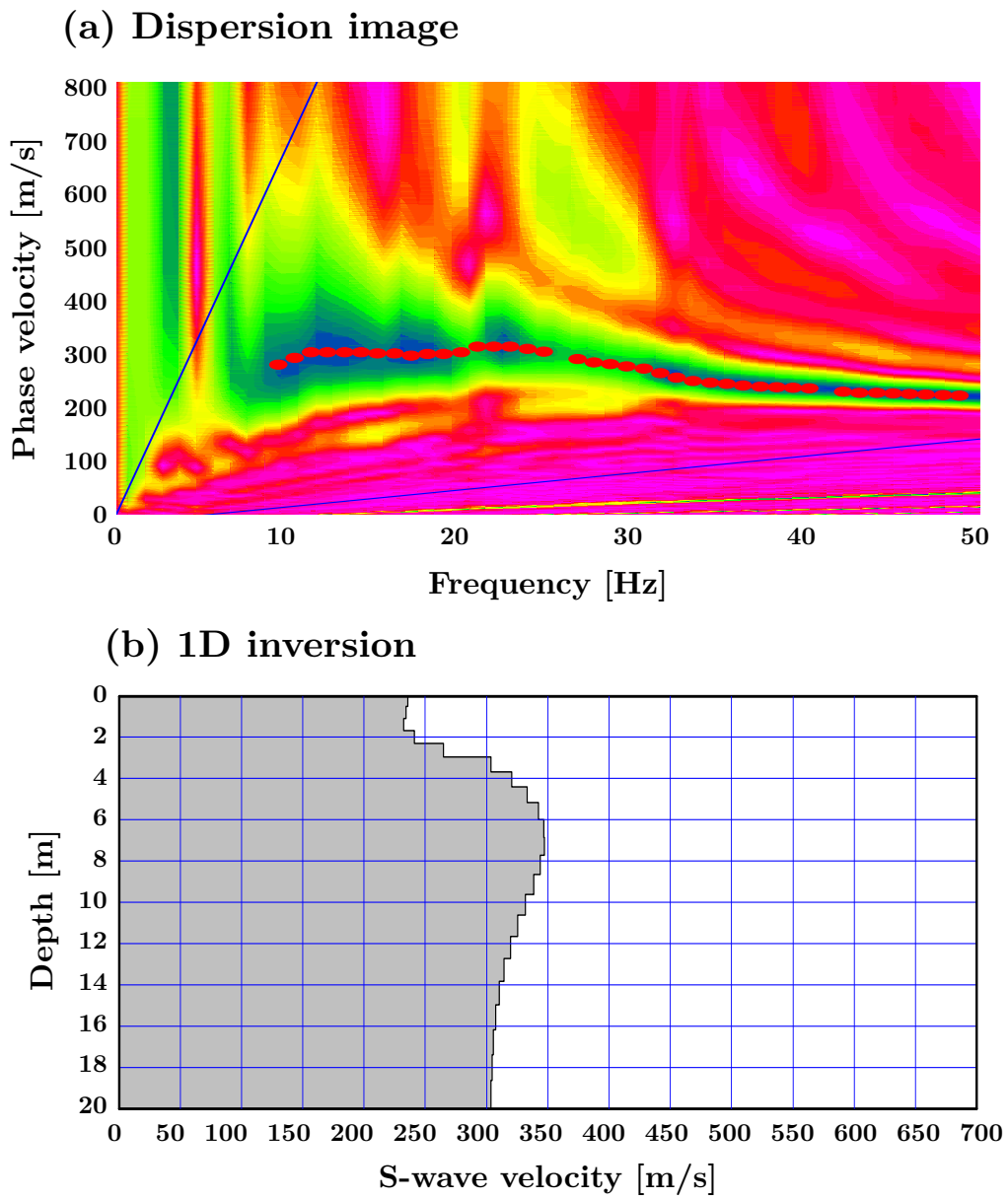

Figure A.1: (a): Dispersion image of the stack of three gathers with the dispersion curve picked using an auto-picking function. The plot clearly shows the absence of low frequencies. (b): 1D inversion of the same data, highlighting the inability of MASW to detect the low-velocity void at 10 m depth.

system is oriented parallel to the x-axis. Modeling was performed using the finite-difference-time domain (FDTD) method implemented in GeoPhyInv.jl, with a Ricker wavelet used as the source. Given memory limitations, a cuboidal representation was selected for the void, as modeling it as a cylinder would have necessitated a finer point spacing to achieve accuracy, which was beyond our computational capabilities. Figures C.1 and C.2 display snapshots of the wavefield in the x-y and x-z planes, illustrating wave propagation within the homogeneous medium both with and without the cuboidal void, as well as the differences in the resulting wavefields. These figures demonstrate the void's ringing effect due to the wave interactions. Figure C.3 shows the raw shot gathers (without cross-correlations) for the homogeneous medium, both with and without the cuboidal void, and the differences in these wavefields. Our results indicate that the backscatter and ringing signatures observed in the 3D acoustic modeling are qualitatively similar to those noted in the 2D elastic experiments. This consistency across dimensional scales justifies our use of 2D synthetic

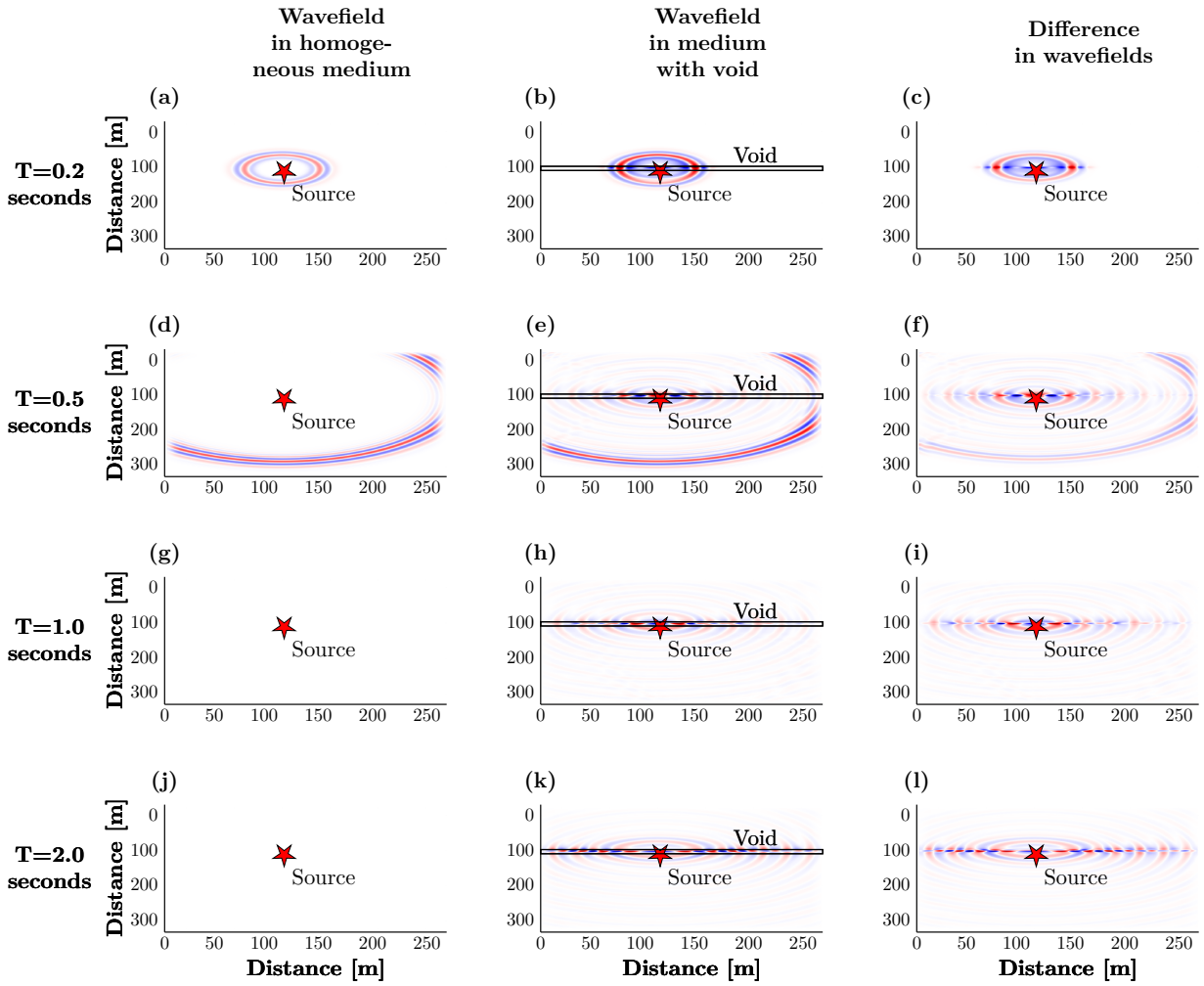

Figure C.1: Seismic wavefield snapshots, analogous to Figure 5, in the x-y plane along the $z = 10\,\text{m}$ slice of a 3D simulation. Panels (a)–(l) show temporal snapshots of seismic wavefields in both a homogeneous medium and a medium containing a void, along with the difference between the two wavefields.

data in the primary analysis and discussion of results, while 3D modeling provides additional insights into the impact of the void's shape and dimensionality on scattering behaviors.

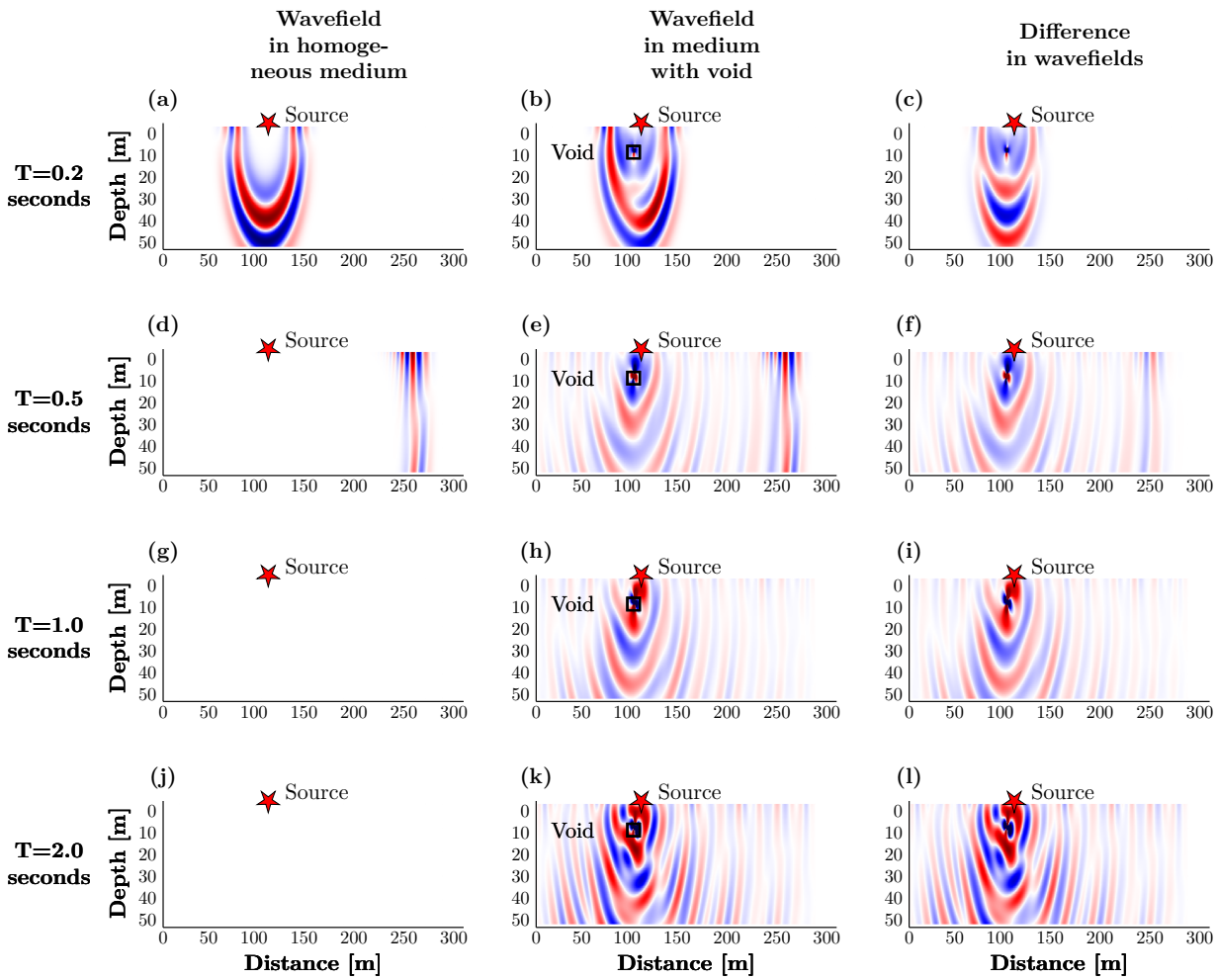

Figure C.2: Seismic wavefield snapshots, similar to Figure C.1, but in the x-z plane along the slice $y = 100\,\text{m}$ of a 3D simulation.

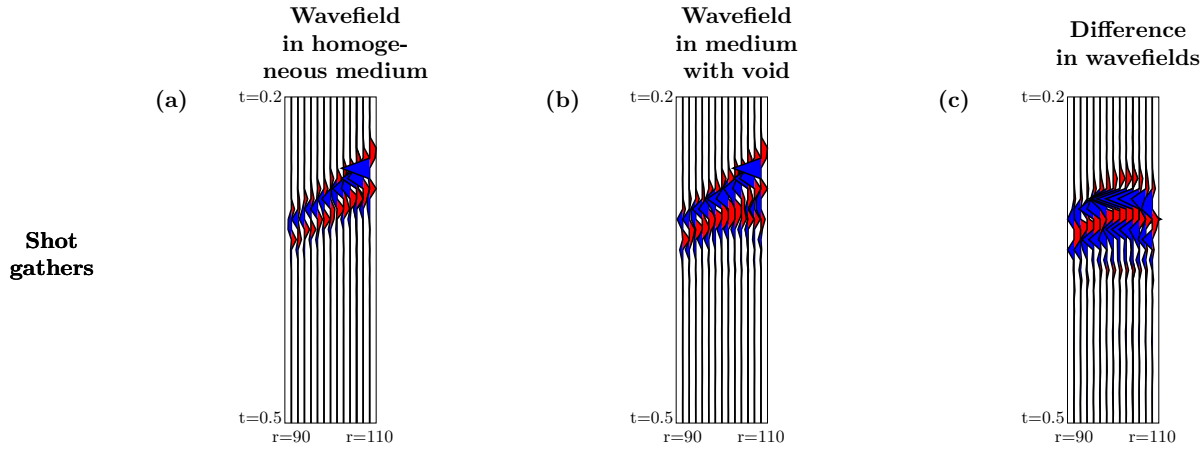

Figure C.3: Raw shot gathers recorded during the 3D simulation: Panels (a)–(c) display shot gathers for the homogeneous medium, both with and without the cuboidal void, along with the differences in these wavefields. The void is located at $x = 100\,\text{m}$, and the gathers are plotted for receivers ranging from $r = 90\,\text{m}$ to $r = 110\,\text{m}$. The backscatters appear as a hyperbola in panel (c).

References

Behm, M., G. M. Leahy, and R. Snieder, 2014, Retrieval of local surface wave velocities from traffic noise–an example from the La Barge basin (Wyoming): Geophysical Prospecting, **62**, 223–243.

Belfer, I., I. Bruner, S. Keydar, A. Kravtsov, and E. Landa, 1998, Detection of shallow objects using refracted and diffracted seismic waves: Journal of Applied Geophysics, **38**, 155–168.

Bharadwaj, P., M. Li, and L. Demanet, 2020, Symae: An autoencoder with embedded physical symmetries for passive time-lapse monitoring, *in* SEG Technical Program Expanded Abstracts 2020: Society of Exploration Geophysicists, 1586–1590.

Bharadwaj, P., M. Li, and L. Demanet, 2022, Redatuming physical systems using symmetric autoencoders: Physical Review Research, **4**, 023118.

Bharadwaj, P., M. Sharma, I. Lohan, and A. P. Singh, 2024, Learning earthquake sources using symmetric autoencoders: arXiv:2304.02404.

Binley, A., S. S. Hubbard, J. A. Huisman, A. Revil, D. A. Robinson, K. Singha, and L. D. Slater, 2015, The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales: Water Resources Research, **51**, 3837–3866.

Branham, K. L., and D. Steeples, 1988, Cavity detection using high-resolution seismic reflection methods: Mining Engineering, **40**, 115–119.

Campillo, M., and A. Paul, 2003, Long-range correlations in the diffuse seismic coda: Science, **299**, 547–549.

Cheng, F., J. Xia, M. Behm, Y. Hu, and J. Pang, 2019, Automated data selection in the Tau–p domain: Application to passive surface wave imaging: Surveys in Geophysics, **40**, 1211–1228.

Cheng, F., J. Xia, Y. Xu, Z. Xu, and Y. Pan, 2015, A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves: Journal of Applied Geophysics, **117**, 126–135.

Clements, T., and M. A. Denolle, 2018, Tracking groundwater levels using the ambient seismic field: Geophysical Research Letters, **45**, 6459–6465.

Cybenko, G., 1989, Approximation by superpositions of a sigmoidal function: Mathematics of Control, Signals and Systems, **2**, 303–314.

De Ridder, S., B. Biondi, and R. Clapp, 2014, Time-lapse seismic noise correlation tomography at Valhall: Geophysical Research Letters, **41**, 6116–6122.

Dobecki, T., 1988, A rapid seismic technique for detecting subsurface voids and unmapped mine workings: 1st EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems, European Association of Geoscientists & Engineers, cp–214.

Fang, G., Y. E. Li, Y. Zhao, and E. R. Martin, 2020, Urban near-surface seismic monitoring using distributed acoustic sensing: Geophysical Research Letters, **47**, e2019GL086115.

Gouedard, P., L. Stehly, F. Brenguier, M. Campillo, Y. C. De Verdière, E. Larose, L. Margerin, P. Roux, F. J. Sánchez-Sesma, N. Shapiro, et al., 2008, Cross-correlation of random fields: Mathematical approach and applications: Geophysical Prospecting, **56**, 375–393.

Gunasekera, R. C., G. R. Foulger, and B. R. Julian, 2003, Reservoir depletion at the Geysers geothermal area, California, shown by four-dimensional seismic tomography: Journal of Geophysical Research: Solid Earth, **108**, 2134.

Gutiérrez, F., M. Parise, J. De Waele, and H. Jourde, 2014, A review on natural and human-induced geohazards and impacts in karst: Earth-Science Reviews, **138**, 61–88.

Hanafy, S. M., A. AlTheyab, and G. T. Schuster, 2015, Controlled noise seismology, *in* SEG Technical Program Expanded Abstracts 2015: Society of Exploration Geophysicists, 5102–5106.

Hornik, K., M. Stinchcombe, and H. White, 1989, Multilayer feedforward networks are universal approximators: Neural Networks, **2**, 359–366.

Ivanov, J., R. D. Miller, D. Z. Feigenbaum, and S. L. Peterie, 2017, Detecting subsurface objects and identifying voids possibilities using the backscatter analysis of surface wave (BASW) method: International Conference on Engineering Geophysics, Al Ain, United Arab Emirates, 9-12 October 2017, Society of Exploration Geophysicists, 116–119.

Korneev, V., 2009, Resonant seismic emission of subsurface objects: Geophysics, **74**, T47–T53.

Landa, E., and S. Keydar, 1998, Seismic monitoring of diffraction images for detection of local heterogeneities: Geophysics, **63**, 1093–1100.

Larose, E., A. Derode, M. Campillo, and M. Fink, 2004, Imaging from one-bit correlations of wideband diffuse wave fields: Journal of Applied Physics, **95**, 8393–8399.

Li, J., and S. Hanafy, 2016, Skeletonized inversion of surface wave: Active source versus controlled noise comparison: Interpretation, **4**, SH11–SH19.

Louie, J. N., 2001, Faster, better: shear-wave velocity to 100 meters depth from refraction microtremor arrays: Bulletin of the Seismological Society of America, **91**, 347–364.

Lumley, D. E., 2001, Time-lapse seismic reservoir monitoring: Geophysics, **66**, 50–53.

Mayne, W. H., 1962, Common reflection point horizontal data stacking techniques: Geophysics, **27**, 927–938.
Mi, B., and J. Xia, 2021, Multisource multichannel analysis of surface waves for near-surface characterization in urban areas: Sixth International Conference on Engineering Geophysics, Virtual, 25–28 October 2021, Society of Exploration Geophysicists, 122–125.
Morton, S. L., J. Ivanov, S. L. Peterie, R. D. Miller, and A. J. Livers-Douglas, 2021, Passive multichannel analysis of surface waves using 1D and 2D receiver arrays: Geophysics, **86**, EN63–EN75.
Olivier, G., F. Brenguier, M. Campillo, R. Lynch, and P. Roux, 2015, Body-wave reconstruction from ambient seismic noise correlations in an underground mine: Geophysics, **80**, KS11–KS25.
Park, C., 2003, Multichannel analysis of surface waves (MASW)-an overview: Journal of the Korean Geophysical Society, **6**, 99–105.
Park, C., R. Miller, D. Laflen, C. Neb, J. Ivanov, B. Bennett, and R. Huggins, 2004, Imaging dispersion curves of passive surface waves, *in* SEG Technical Program Expanded Abstracts 2004: Society of Exploration Geophysicists, 1357–1360.
Park, C., R. Miller, and H. Miura, 2002, Optimum field parameters of an masw survey [exp. abs.]: SEG-J, Tokyo, 22–23.
Park, C., R. Miller, N. Rydén, J. Xia, and J. Ivanov, 2005, Combined use of active and passive surface waves: Journal of Environmental and Engineering Geophysics, **10**, 323–334.
Park, C. B., R. D. Miller, and J. Xia, 1999, Multichannel analysis of surface waves: Geophysics, **64**, 800–808.
Peterie, S. L., and R. D. Miller, 2015, Near-surface scattering phenomena and implications for tunnel detection: Interpretation, **3**, SF43–SF54.
Romero Jr, A. E., T. V. McEvilly, and E. L. Majer, 1997, 3D microearthquake attenuation tomography at the Northwest Geysers geothermal region, California: Geophysics, **62**, 149–167.
Rubin, S., V. Shtivelman, S. Keydar, and A. Lev, 2014, Seismic ringing effect in the shallow subsurface: Near Surface Geophysics, **12**, 687–696.
Rückemann, C. P., 2012, Comparison of stacking methods regarding processing and computing of geoscientific depth data: Proceedings of the Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2012), 35–40.
Sabra, K. G., P. Roux, and W. Kuperman, 2005, Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function: The Journal of the Acoustical Society of America, **118**, 3524–3531.
Schimmel, M., E. Stutzmann, and J. Gallart, 2011, Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale: Geophysical Journal International, **184**, 494–506.
Schneider, F., S. Esterhazy, I. Perugia, and G. Bokelmann, 2017, Seismic resonances of spherical acoustic cavities: Geophysical Prospecting, **65**, 1–24.
Schwenk, J. T., S. D. Sloan, J. Ivanov, and R. D. Miller, 2016, Surface-wave methods for anomaly detection: Geophysics, **81**, EN29–EN42.
Schwenk, J. T., S. D. Sloan, R. D. Miller, and J. Ivanov, 2014, Correlation of the backscatter analysis of surface waves method (BASW) for anomaly detection, *in* SEG Technical Program Expanded Abstracts 2014: Society of Exploration Geophysicists, 2029–2035.
Sloan, S., H. Cudney, R. Miller, J. Ivanov, and M. Moran, 2016, Depth estimation of voids using backscattered surface waves, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 2362–2366.
Sloan, S. D., and D. Z. Feigenbaum, 2020, Examples of seismic resonance from air-filled voids, *in* SEG Technical Program Expanded Abstracts 2020: Society of Exploration Geophysicists, 1955–1959.
Sloan, S. D., and D. Z. Feigenbaum, 2024, Field examples of void-related resonance in shallow seismic data: Journal of Environmental and Engineering Geophysics, **29**, 37–42.
Sloan, S. D., D. Z. Feigenbaum, and R. H. Stevens, 2021, Backscattered surface waves and the relationship between wavelength and signature response: Sixth International Conference on Engineering Geophysics, Virtual, 25–28 October 2021, Society of Exploration Geophysicists, 48–51.
Sloan, S. D., J. J. Nolan, S. W. Broadfoot, J. R. McKenna, and O. M. Metheny, 2013, Using near-surface seismic refraction tomography and multichannel analysis of surface waves to detect shallow tunnels: A feasibility study: Journal of Applied Geophysics, **99**, 60–65.
Sloan, S. D., S. L. Peterie, R. D. Miller, J. Ivanov, J. T. Schwenk, and J. R. McKenna, 2015, Detecting clandestine tunnels using near-surface seismic techniques: Geophysics, **80**, EN127–EN135.
Snieder, R., 2004, Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase: Physical Review E, **69**, 046610.
Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, 2014, Dropout: a simple way to prevent neural networks from overfitting: The Journal of Machine Learning Research, **15**, 1929–1958.
Steeples, D. W., 2005, Near-surface geophysics: 75 years of progress: The Leading Edge, **24**, s82–s85.
Wang, Y., M. Khorrami, K. T. Tran, and D. Horhota, 2023, Application of ambient noise tomography for deep void detection: Journal of Applied Geophysics, **209**, 104922.

Wapenaar, K., 2004, Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation: Physical Review Letters, **93**, 254301.
Weaver, R. L., 2005, Information from seismic noise: Science, **307**, 1568–1569.
Weaver, R. L., and O. I. Lobkis, 2001, Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies: Physical Review Letters, **87**, 134301.
Weaver, R. L., and O. I. Lobkis, 2005, Fluctuations in diffuse field–field correlations and the emergence of the Green's function in open systems: The Journal of the Acoustical Society of America, **117**, 3432–3439.
Xie, J., Y. Yang, and Y. Luo, 2020, Improving cross-correlations of ambient noise using an rms-ratio selection stacking method: Geophysical Journal International, **222**, 989–1002.
Youssef, A. M., Y. A. Zabramwi, F. Gutiérrez, A. M. Bahamil, Z. A. Otaibi, and A. J. Zahrani, 2020, Sinkholes induced by uncontrolled groundwater withdrawal for agriculture in arid Saudi Arabia: Integration of remote-sensing and geophysical (ERT) techniques: Journal of Arid Environments, **177**, 104132.
Zaheer, M., S. Kottur, S. Ravanbhakhsh, B. Póczos, R. Salakhutdinov, and A. J. Smola, 2017, Deep sets: Advances in Neural Information Processing Systems, **30**, 3392–3402.
Zhang, H., B. Mi, Y. Liu, C. Xi, and K. L. Kouadio, 2021, A pitfall of applying one-bit normalization in passive surface-wave imaging from ultra-short roadside noise: Journal of Applied Geophysics, **187**, 104285.
Zhang, K., H. Li, X. Wang, and K. Wang, 2020, Retrieval of shallow S-wave profiles from seismic reflection surveying and traffic-induced noise: Geophysics, **85**, EN105–EN117.