



Green's function estimation by seismic interferometry from limited frequency samples[☆]

Justin Jayne^{a,b}, Michael B. Wakin^{a,*}, Roel Snieder^c

^a Department of Electrical Engineering, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, United States

^b Rincon Research, United States

^c Center for Wave Phenomena, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, United States

ARTICLE INFO

Article history:

Received 16 June 2022

Revised 14 November 2022

Accepted 16 November 2022

Available online 17 November 2022

Keywords:

Compressive convolution

Cross-correlation

Random spectral sampling

Seismic interferometry

ABSTRACT

Green's function estimation is an important application of seismic interferometry but can require cross-correlating very long time series that are difficult to gather, store, and transmit in resource-constrained scenarios. We derive a compressive approach for estimating a Green's function using only a small number of random frequency samples from each signal. We bound the maximum error between this estimator and the original cross-correlation and show how this error decreases as the number of samples increases. We demonstrate the application of this technique to a numerical one-dimensional reflected wave case and to estimation of surface wave Green's functions for the western United States using USArray data. We show that the compressive approach can be extended to deconvolution as well, and we illustrate this with pressure and displacement data recorded on a volcano. We also provide guidelines for implementing the technique.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Seismic interferometry commonly refers to the cross-correlation and summation of seismograms to construct estimates of the Green's function between sensor pairs, effectively turning each pair of receivers into a virtual source-receiver pair [1–3]. Originally developed in the context of the reflection response of a horizontally layered system from recorded transmitted waves [4], the technique has been extended to multiple settings and scales [2,5–8].

Seismic interferometry can require cross-correlating very long time series: one study [7] used over one year's worth of data sampled at 0.2 Hz ($N = 10^7 - 10^8$, where N is the number of observations in the record), while another study [9] used three months' worth of data sampled at 10 Hz ($N = 10^8 - 10^9$). Gathering, storing, transmitting, and processing such lengthy time series can potentially impose heavy burdens on data processing systems. It is not difficult to imagine resource constrained scenarios in which interferometric analysis would be rendered infeasible by the need to work with large data sets. For example, both the submersible MERMAID sensor network [10] and the AGAP/GAMSEIS seismic array in

eastern Antarctica [11] represent sensor networks where storage, computation, and/or communication bandwidth might be scarce.

In this paper, we introduce a new technique called *compressive convolution*. This new approach is based on a previous construction known as the compressive matched filter [12] and uses theory from compressive sensing (CS) [13–15] to estimate either the convolution or the cross-correlation of two signals using a small number of randomly subsampled frequency components from each signal. Given the fundamental role that convolutions and correlations play in signal processing and machine learning, one can envision many possible applications of such a compressive estimator. In this paper, we focus on the application of this technique to seismic interferometry. We show that the compressive approach can reduce the number of frequencies used for Green's function estimation by orders of magnitude.

An advantage of the compressive approach is that it may be easily incorporated into existing seismic data processing schemes. The only new operation is random subsampling of the spectra at each receiver. Such random frequency samples can be computed either (i) online at each sensor from a streaming input signal, or (ii) offline at each sensor using the fast Fourier transform (FFT) on a finite vector of samples. Only these random frequency samples need be transmitted to a central node where the estimate of the Green's function is computed using a formula that we provide.

Our objective with this paper is twofold: (i) introduce compressive convolution and (ii) demonstrate its applicability to seismic

[☆] This work was supported by NSF grants CCF-1409258 and CCF-1704204 and NSF CAREER grant CCF-1149225.

* Corresponding author.

E-mail address: mwakin@mines.edu (M.B. Wakin).

interferometry. In [Section 2](#), we introduce and theoretically analyze a compressive unbiased estimator of the convolution of two arbitrary signals from random frequency samples. Our analysis bounds the error in the compressive convolution estimate in terms of the number of random frequency samples collected, and this yields insight into cases where the estimator can succeed with a number of frequency samples far lower than what would be required with classical Nyquist sampling. In [Section 3](#), we discuss how this estimator can be applied to seismic interferometry. In [Section 4](#), we apply this technique to a one-dimensional numerical simulation of the reflection response of a layered medium ([Section 4.1](#)). Next, as a more comprehensive demonstration, we estimate the Rayleigh waves propagating through the western United States based on compressive seismic interferometry of USArray seismic data ([Section 4.2](#)). Finally, we show in [Section 4.3](#) that the concept of compressive convolution can be extended to compressive deconvolution and illustrate this with pressure and displacement data recorded on a volcano. We summarize our conclusions in [Section 5](#).

2. Compressive convolution

2.1. Methodology and main results

Consider the convolution of two arbitrary real-valued signals, $v(t)$ and $w(t)$. Denote the Fourier transforms of the two signals by $\hat{v}(\omega)$ and $\hat{w}(\omega)$, respectively, and suppose there exist frequencies ω_{lo} and ω_{hi} such that $\hat{v}(\omega)$ and $\hat{w}(\omega)$ are bandlimited to the frequency interval $\Omega = [-\omega_{hi}, -\omega_{lo}] \cup [\omega_{lo}, \omega_{hi}]$ rad/s. That is, $\hat{v}(\omega) = \hat{w}(\omega) = 0$ for all $\omega \notin \Omega$. Denote the set of positive frequencies as $\Omega_+ = [\omega_{lo}, \omega_{hi}]$ rad/s.

Under these assumptions, the convolution of $v(t)$ and $w(t)$ is given by

$$\begin{aligned} x(t) &= (v * w)(t) = \frac{1}{2\pi} \int_{\Omega} (\hat{v}(\omega) \odot \hat{w}(\omega)) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \langle \hat{v}(\omega) \odot \hat{w}(\omega), e^{-i\omega t} \rangle, \quad t \in T, \end{aligned} \quad (1)$$

where $\odot$ denotes the Hadamard (elementwise) product, $\langle \cdot, \cdot \rangle$ denotes the inner product, and T is the time interval of interest.

In order to obtain a compressive estimator of $x(t)$, suppose we acquire M samples of $\hat{v}(\omega)$ and $\hat{w}(\omega)$ at positive frequencies $\omega_1, \omega_2, \dots, \omega_M$, which are drawn at random from a uniform distribution on Ω_+ , yielding vectors of frequency-domain observations $\hat{v}, \hat{w} \in \mathbb{C}^M$ with elements

$$\hat{v}[k] = \int_{-\infty}^{\infty} v(t) e^{-i\omega_k t} dt \quad \text{and} \quad \hat{w}[k] = \int_{-\infty}^{\infty} w(t) e^{-i\omega_k t} dt \quad (2)$$

for $k = 1, 2, \dots, M$. Such frequency-domain samples could be acquired either (i) by computing the integrals in [Eq. \(2\)](#) in streaming fashion as the signals $v(t)$ and $w(t)$ arrive at their sensors, or (ii) using an FFT after the signals arrive. For each value of $t \in T$, define the associated template vector

$$\mathbf{e}_t := \begin{bmatrix} e^{-i\omega_1 t} \\ e^{-i\omega_2 t} \\ \vdots \\ e^{-i\omega_M t} \end{bmatrix} \in \mathbb{C}^M. \quad (3)$$

Using the vectors $\hat{v}, \hat{w}, \mathbf{e}_t \in \mathbb{C}^M$, we define the compressive convolution estimator as follows. For each value of $t \in T$, the estimate $\tilde{x}(t)$ of $x(t)$ is defined as

$$\tilde{x}(t) := \frac{|\Omega_+|}{\pi M} \cdot \text{Re}\{\langle \hat{v} \odot \hat{w}, \mathbf{e}_t \rangle\} = \frac{|\Omega_+|}{\pi M} \cdot \text{Re}\left\{ \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e^{i\omega_k t} \right\}, \quad (4)$$

where $|\Omega_+| = \omega_{hi} - \omega_{lo}$ and $\text{Re}\{\cdot\}$ returns the real part of a possibly complex number. Because the frequencies $\omega_1, \omega_2, \dots, \omega_M$ are chosen at random, $\tilde{x}(t)$ is a random process that is a function of time. This random process $\tilde{x}(t)$ has mean function

$$\begin{aligned} \mathbb{E}[\tilde{x}(t)] &= \frac{|\Omega_+|}{\pi M} \cdot \text{Re}\left\{ \sum_{k=1}^M \mathbb{E}[\hat{v}[k] \hat{w}[k] e_i^*[k]] \right\} \\ &= 2 \cdot \text{Re}\left\{ \frac{1}{2\pi} \int_{\Omega_+} \hat{v}(\omega) \hat{w}(\omega) e^{i\omega t} d\omega \right\} = x(t), \end{aligned} \quad (5)$$

where $\cdot^*$ denotes the complex conjugate, the first equality uses the definition of expectation, the second equality uses the fact that the frequencies are drawn from a uniform distribution over Ω_+ , and the last equality follows from (1). One way to interpret the estimator (4) is that we have approximated the ideal convolution frequency-domain integral (1) as a sum of samples taken at random locations in Ω_+ . [Eq. \(5\)](#) shows that $\tilde{x}(t)$ is an unbiased estimator of $x(t)$; in short, random sampling is unbiased.

Of fundamental interest is quantifying how close the random process $\tilde{x}(t)$ is to its mean $x(t)$. In [\[16\]](#), we study a related but simpler question, characterizing the covariance of the error induced by randomly subsampling a single finite, discrete-time signal's frequency spectrum. Here, we derive a bound for the maximum deviation between the compressive convolution estimator $\tilde{x}(t)$ and its mean over all $t \in T$, a quantity denoted as $\sup_{t \in T} |\tilde{x}(t) - x(t)|$. This quantity is random, and our first theorem (the proof of which is in [Appendix A](#)) bounds its expectation.

Theorem 1. Suppose that $|\Omega_+||T| \geq 3$. Then

$$\mathbb{E}\left[\sup_{t \in T} |\tilde{x}(t) - x(t)|\right] \leq 1.342 \sqrt{\log(2|\Omega_+||T|)} \cdot \frac{\sqrt{|\Omega_+|} \cdot \|\hat{x}(\omega)\|_2}{\sqrt{M}}. \quad (6)$$

[Theorem 1](#) characterizes the relationship between the signal bandwidth $|\Omega_+|$, convolution interval T , spectral energy $\|\hat{x}(\omega)\|_2$, number of random spectral samples M , and expected maximum deviation between $\tilde{x}(t)$ and its mean. With other parameters fixed, [Theorem 1](#) guarantees that the expected maximum deviation decreases as $1/\sqrt{M}$; as expected, collecting more random spectral samples results in more accurate estimates of $x(t)$. Examining the role of other parameters, the maximum deviation scales linearly with the spectral energy, which is natural since rescaling a signal will also rescale its estimate (and thus the error). In [Section 2.2](#), we provide more insight into the number of compressive samples M required to guarantee a certain level of accuracy in estimating $x(t)$.

Note that the requirement that $|\Omega_+||T| \geq 3$ is not a significant restriction in practice. The time-bandwidth product $|\Omega_+||T|$ (which also appears logarithmically in (6)) effectively counts the number of degrees of freedom in $x(t)$ over the observation interval T , and it is a mild assumption that $|\Omega_+||T| \geq 3$. Nevertheless, if $|\Omega_+||T| \leq 3$, both theorems hold but with larger constants.

Our second theorem (also proved in [Appendix A](#)) provides a tail bound on $\sup_{t \in T} |\tilde{x}(t) - x(t)|$, ensuring that this quantity is unlikely to achieve a value significantly larger than its expectation.

Theorem 2. Fix $\delta > 0$ and let

$$U = 4.07 \cdot \sqrt{\log(12|\Omega_+||T|/\delta)} \cdot \max\left(\frac{\sqrt{|\Omega_+|} \cdot \|\hat{x}(\omega)\|_2}{\sqrt{M}}, \sqrt{2 \log(4/\delta)} \cdot \frac{|\Omega_+| \cdot \|\hat{x}(\omega)\|_{\infty}}{M}\right). \quad (7)$$

If $|\Omega_+||T| \geq 3$, then

$$\mathbb{P}\left\{\sup_{t \in T} |\tilde{x}(t) - x(t)| > U\right\} \leq \delta. \quad (8)$$

In words, depending on which term dominates the maximum in (7), the factor U appearing in Theorem 1 scales essentially like the expectation bound on $\sup_{t \in T} |\tilde{x}(t) - x(t)|$ that was given in Theorem 1. Theorem 1 states that $\sup_{t \in T} |\tilde{x}(t) - x(t)|$ is unlikely to exceed this value U .

As detailed in Appendix A, the methodology described in this section can be generalized to complex-valued signals $v(t)$ and $w(t)$ with spectral supports on an arbitrary interval $[\omega_a, \omega_b]$: one collects samples randomly from a uniform distribution on $[\omega_a, \omega_b]$ and uses the estimator $\tilde{x}(t) = \frac{\omega_b - \omega_a}{2\pi M} \cdot (\hat{v} \odot \hat{w}, \mathbf{e}_t)$.

2.2. Insight into sample complexity

A natural question is: how large must one choose the number of random spectral samples M so that Theorems 1 and 2 provide a useful bound on the maximum deviation between $\tilde{x}(t)$ and $x(t)$? A meaningful answer to this question depends on the time- and frequency-domain properties of the true convolution signal $x(t)$. In many problems of interest—including seismic interferometry— $x(t)$ may contain a few prominent peaks, and one may wish to ensure the maximum deviation between $\tilde{x}(t)$ and $x(t)$ is small relative to the size of these peaks. From the Fourier transform definition, one can show that $|x(t)| \leq \|\tilde{x}(\omega)\|_1$ for all t . That is, the true convolution signal in the time domain has an amplitude no larger than $\|\tilde{x}(\omega)\|_1$, and strong peaks in $x(t)$ will tend to have an amplitude which is on the order of $\|\tilde{x}(\omega)\|_1$. Thus, $\|\tilde{x}(\omega)\|_1$ provides a benchmark for measuring the error in the time domain between $\tilde{x}(t)$ and $x(t)$.

With this benchmark in mind, consider the right hand side of (6), which bounds the expected maximum deviation between $\tilde{x}(t)$ and $x(t)$. Suppose that, for some small $\eta \ll 1$, we wish to ensure the right hand side of (6) is less than or equal to $\eta \cdot \|\tilde{x}(\omega)\|_1$. Rearranging terms, we see that this is ensured if

$$M \geq \left(\frac{1.342 \sqrt{\log(2|\Omega_+||T|)} \cdot \sqrt{|\Omega_+|} \cdot \|\tilde{x}(\omega)\|_2}{\eta \cdot \|\tilde{x}(\omega)\|_1} \right)^2. \quad (9)$$

This bound depends on the ratio of the L_2 and L_1 norms of $\tilde{x}(\omega)$. Roughly speaking, this ratio depends on the degree to which $\tilde{x}(\omega)$ is evenly spread across its bandwidth Ω . More precisely, from a standard inequality relating the L_2 and L_1 norms, it follows that

$$\frac{\|\tilde{x}(\omega)\|_2}{\|\tilde{x}(\omega)\|_1} \geq \frac{1}{\sqrt{|\Omega|}} = \frac{1}{\sqrt{2|\Omega_+|}},$$

where equality is achieved when $\tilde{x}(\omega)$ has constant magnitude across Ω . In such a case, (9) becomes

$$M \geq \left(\frac{1.342}{\sqrt{2}} \right)^2 \cdot \frac{\log(2|\Omega_+||T|)}{\eta^2}. \quad (10)$$

A similar analysis applies for ensuring the first term in the definition of U (see (7)) is bounded by $\eta \cdot \|\tilde{x}(\omega)\|_1$. To bound the second term in the definition of U by $\eta \cdot \|\tilde{x}(\omega)\|_1$, we require

$$M \geq \frac{4.07 \cdot \sqrt{\log(12|\Omega_+||T|/\delta)} \cdot \sqrt{2 \log(4/\delta)} \cdot |\Omega_+| \cdot \|\tilde{x}(\omega)\|_\infty}{\eta \cdot \|\tilde{x}(\omega)\|_1}. \quad (11)$$

This bound now depends on the ratio of the L_∞ and L_1 norms of $\tilde{x}(\omega)$, a ratio which again depends on the degree to which $\tilde{x}(\omega)$ is evenly spread across its bandwidth Ω . From a standard inequality, it follows that

$$\frac{\|\tilde{x}(\omega)\|_\infty}{\|\tilde{x}(\omega)\|_1} \geq \frac{1}{|\Omega|} = \frac{1}{2|\Omega_+|},$$

with equality when $\tilde{x}(\omega)$ has constant magnitude across Ω . In such a case, (11) becomes

$$M \geq \frac{2.035 \cdot \sqrt{\log(12|\Omega_+||T|/\delta)} \cdot \sqrt{2 \log(4/\delta)}}{\eta}. \quad (12)$$

To summarize the implications of (10) and (12), in cases where the spectrum of $\tilde{x}(\omega)$ has roughly constant magnitude and where $x(t)$ has a strong peak in the time domain, M need only scale logarithmically in the time-bandwidth product $|\Omega_+||T|$ to ensure the maximum deviation between $\tilde{x}(t)$ and $x(t)$ is bounded by a small multiple of the peak value of $x(t)$. Importantly, such cases arise naturally when $x(t)$ has a sinc- or wavelet-like structure in the time domain, which correspond (exactly or approximately) to a rectangular frequency spectrum across the bandwidth Ω . This logarithmic scaling is far better than the linear scaling in $|\Omega_+||T|$ that would be required by the classic Nyquist theorem for deterministic time-domain sampling. For other scenarios, where a smaller time-domain bound may be desired or where the spectrum of $\tilde{x}(\omega)$ is highly nonuniform, Theorems 1 and 2 can be interpreted correspondingly to yield insight on the required number of samples M . We direct the reader to Eftekhar et al. [12] for a similar discussion in the context of compressive matched filtering and to Snieder and Wakin [16] for more insight into the “background noise” introduced when one undersamples a signal’s frequency spectrum.

3. Compressive seismic interferometry

The operations of convolution and cross-correlation are intimately related, differing only by a time-reversal of one signal in the time-domain or, equivalently, a complex conjugation of one signal in the frequency domain. For this reason, the compressive convolution estimator can naturally be applied to estimating the cross-correlation between two wavefields, which is fundamental to seismic interferometry. For interferometric studies of surface waves one normally correlates components of the displacement to obtain the displacement Green’s function of surface waves. Strictly speaking, one should correlate the displacement with the stress to obtain the displacement Green’s function [17]. However, when the noise sources are in the far field from the receivers, one can simply correlate the displacement recorded at different receivers [18]. This has been used in observational studies with USArray [19] and with an industrial array [20] where it was shown that the correlation of noise recorded on the vertical (or radial) components recorded at different stations give the Rayleigh wave that propagates between these stations, while the correlation of the transverse components gives the Love wave.

We denote the Green’s function of a medium, defined as the impulse response observed at some receiver at location $\mathbf{x}_A$ from a source at location $\mathbf{x}_B$, as $G(\mathbf{x}_A, \mathbf{x}_B, t)$. Following the approach described by Wapenaar et al. [3], the Green’s function between two receivers at locations $\mathbf{x}_A$ and $\mathbf{x}_B$ can be estimated by cross-correlating and averaging an ensemble of concurrent observations $u^A(t) := \phi(\mathbf{x}_A, t)$ and $u^B(t) := \phi(\mathbf{x}_B, t)$ of a noisy, ambient wavefield $\phi(\mathbf{x}, t)$. In this approach, $u^A(t)$ and $u^B(t)$ are partitioned into L shorter segments, each of length N_{seg} samples. We denote the observations associated with each segment as $u_i^A(t)$ and $u_i^B(t)$, $i = 1, 2, \dots, L$. For example, given 100 days of observations sampled at 1 Hz, one might divide each record into $L = 100$ one-day segments, each containing $N_{\text{seg}} = 86,400$ samples. The total number of samples in each of the two records is $N_{\text{tot}} = N_{\text{seg}}L = 8.64 \times 10^6$. The Green’s function is related to the observations via the equation $[G(\mathbf{x}_A, \mathbf{x}_B, t) + G(\mathbf{x}_A, \mathbf{x}_B, -t)] * \rho_{SS}(t) = \langle \langle u^A(t) * u^B(-t) \rangle \rangle$, where we use $\langle \langle \cdot \rangle \rangle$ to denote ensemble averaging, and where $\rho_{SS}(t)$ denotes the autocorrelation of the noise source [2,21]. Consequently, a conventional seismic interferometry estimator using L

segments of data is given by

$$\begin{aligned}\Gamma_{AB}(t) &:= \frac{1}{L} \sum_{i=1}^L u_i^A(t) * u_i^B(-t) \\ &\approx [G(\mathbf{x}_A, \mathbf{x}_B, t) + G(\mathbf{x}_A, \mathbf{x}_B, -t)] * \rho_{SS}(t).\end{aligned}\quad (13)$$

A compressive estimator of $\Gamma_{AB}(t)$ can be formed by applying Eq. (4) to Eq. (13), allowing $u_i^A(t)$ to play the role of $v(t)$ and $u_i^B(-t)$ to play the role of $w(t)$ in each segment $i = 1, 2, \dots, L$. In particular, assume the data is real-valued with positive bandwidth Ω_+ . In time segment $T_i = [T_{i,\text{start}}, T_{i,\text{end}}]$, collect spectral samples at randomly chosen frequencies $\omega_{i,1}, \omega_{i,2}, \dots, \omega_{i,M} \in \Omega_+$:

$$\hat{u}_i^A[k] := \int_{-\infty}^{\infty} u_i^A(t) e^{-i\omega_{i,k}t} dt = \int_{T_{i,\text{start}}}^{T_{i,\text{end}}} u_i^A(t) e^{-i\omega_{i,k}t} dt, \quad (14)$$

$$\hat{u}_i^{*B}[k] := \int_{-\infty}^{\infty} u_i^B(-t) e^{-i\omega_{i,k}t} dt = \int_{-T_{i,\text{end}}}^{-T_{i,\text{start}}} u_i^B(-t) e^{-i\omega_{i,k}t} dt, \quad (15)$$

for $k = 1, 2, \dots, M$. Combining all of these spectral samples, we obtain the compressive estimator

$$\tilde{\Gamma}_{AB}(t) := \frac{|\Omega_+|}{\pi ML} \sum_{i=1}^L \text{Re} \left\{ \sum_{k=1}^M \hat{u}_i^A[k] \hat{u}_i^{*B}[k] e^{i\omega_{i,k}t} \right\}. \quad (16)$$

4. Examples

4.1. One-dimensional reflected wave interferometry

Following the example of Claerbout [4], our first demonstration is a simulation of one-dimensional reflected wave interferometry. This example parallels a simplified exploration geophysics scenario in which the reflection response $R(t)$, the response of the system to an impulsive source located at the surface, may reveal the depth and reflection/transmission coefficients of subsurface boundaries.

Claerbout's insight was in showing that $R(t)$ is related to $T(t)$, the transmission response of the system to a deep impulsive source, and that one may obtain a modulated form of $R(t)$ by observing the response $u(t)$ of the system to some deep source signal $S(t)$:

$$(R(t) + R(-t)) * \rho_{SS}(t) = \rho_{SS}(t) - \rho_{uu}(t), \quad (17)$$

where $\rho_{SS}(t)$ is the autocorrelation of $S(t)$ and $\rho_{uu}(t)$ is the autocorrelation of $u(t)$.

We simulate a reflected wave interferometry scenario with a single reflective boundary located 1500 m below the surface and a noisy source located 3000 m below the surface. In this scenario, we use a reflection coefficient $r = 0.9$, transmission coefficient $t = 0.44$, and seismic wave velocity of 1500 m/s. The source signal is simulated as bandlimited Gaussian noise with a central frequency of 30 Hz, bandwidth of 10 Hz, and duration of 300 s. The sampling frequency, f_{samp} , of the simulated receiver is set to 1000 Hz over the 300 s duration of the simulation, so the simulated received signals are of length $N = 3 \times 10^5$ samples.

Because the spectrum of the source is essentially supported on $\Omega = [-2\pi 30, -2\pi 10] \cup [2\pi 10, 2\pi 30]$ rad/s, we restrict our analysis to using only frequencies in Ω . The source and received signal autocorrelation functions in Eq. (17) can be approximated using

$$\rho_{SS}(t) \approx \frac{1}{2\pi} \int_{\Omega} (\hat{S}(\omega) \odot \hat{S}^*(\omega)) e^{i\omega t} d\omega, \quad (18)$$

$$\rho_{uu}(t) \approx \frac{1}{2\pi} \int_{\Omega} (\hat{u}(\omega) \odot \hat{u}^*(\omega)) e^{i\omega t} d\omega. \quad (19)$$

Using only frequencies in the band Ω , we can reduce N from 3×10^5 real-valued time-domain samples to 6×10^3 complex-valued

frequency-domain samples. (Because the signals are real-valued, it suffices to keep the spectral samples only on the interval $\Omega_+ = [2\pi 10, 2\pi 30]$ rad/s.) Based on these deterministic bandlimited samples, the results of a non-compressive analysis are shown in Fig. 1. The ideal reflection response $R(t)$ is shown in Fig. 1(a) and the autocorrelation function of the source $\rho_{SS}(t)$ is shown in Fig. 1(b). The causal parts of the left- and right-hand sides of Eq. (17) are shown in Fig. 1(c) and (d), respectively.

We then repeated this experiment using compressive convolution to calculate estimates of the autocorrelation of $u(t)$, denoted $\tilde{\rho}_{uu}(t)$, using between $M = 6 \times 10^1$ and $M = 3 \times 10^3$ complex-valued spectral subsamples of $u(t)$. All samples were collected uniformly at random from the interval $\Omega_+ = [2\pi 10, 2\pi 30]$ rad/s. The resulting compression ratios M/N range from 0.01 to 0.5, where N is the number of deterministic samples on Ω_+ described in the previous paragraph. The estimated Green's functions are shown in Fig. 2. While the compressive estimates grow gradually less accurate as the subsampling ratio decreases, it is clear that the shape of the modulated reflection response (Fig. 1(c)), is generally recognizable, even when $\frac{M}{N} \leq 0.01$.

We can quantify this global preservation of the locations of the peaks of $\tilde{\rho}_{uu}(t)$ using Theorem 1:

$$\mathbb{E} \left[\sup_{t \in T} |\tilde{\rho}_{uu}(t) - \rho_{uu}(t)| \right] \leq 1.342 \sqrt{\log(2|\Omega_+||T|)} \cdot \frac{\sqrt{|\Omega_+|} \cdot \|\hat{\rho}_{uu}(\omega)\|_2}{\sqrt{M}}. \quad (20)$$

Across 100 independent experiments for each value of M , Fig. 3 shows the empirical mean value of $\sup_{t \in T} |\tilde{\rho}_{uu}(t) - \rho_{uu}(t)|$ as a function of M . In agreement with Theorem 1, the mean maximum absolute error is approximately proportional to $1/\sqrt{M}$ (the dashed line on Fig. 3 is $C/\sqrt{M}$, where C is a scaling constant chosen for the sake of illustration).

As a coda to this example, the results provide a useful guideline for choosing M . If one has an estimate of the order of magnitude of the features of $\tilde{\rho}_{uu}(t)$ that are of interest, one should choose M such that the right-hand side of Eq. (20) is approximately equal to that magnitude. This ensures that the “noise floor” from the estimation errors does not overwhelm the features of interest. See Section 2.2 for further discussion.

4.2. USArray seismic data analysis

Our next example involves testing the compressive approach to seismic interferometry on real-world seismic data. Inspired by the work of Lin et al. [7], we also estimated surface wave velocity in the western USA by cross-correlating observations of ocean-generated ambient seismic noise made with the EarthScope/USArray Transportable Array. We cross-correlated observations made at each station (spread across several western states) with those made at Station R06C, located southeast of Lake Tahoe. Cross-correlation yields an estimate of the travel time between the elements of the virtual source-receiver pair. Velocity was calculated using the great circle distance between stations.

We used observations of the 24 s Rayleigh wave signal measured at a sampling frequency of 1 Hz between October 2004 and November 2007 (available at http://ds.iris.edu/gmap/#network=_US-TA&planet=earth). Following the data processing scheme described by Bensen et al. [22], we bandpass filtered the observations using a Butterworth filter with a passband of 0.033–0.055 Hz (i.e., periods between 18 and 30 s), partitioned the record, and removed the mean before cross-correlating and stacking. Using a non-compressive approach, we were able to replicate travel times that [7] estimated between the R06C station and other stations in the western United States whose records overlap with that of R06C for at least 365 consecutive days.

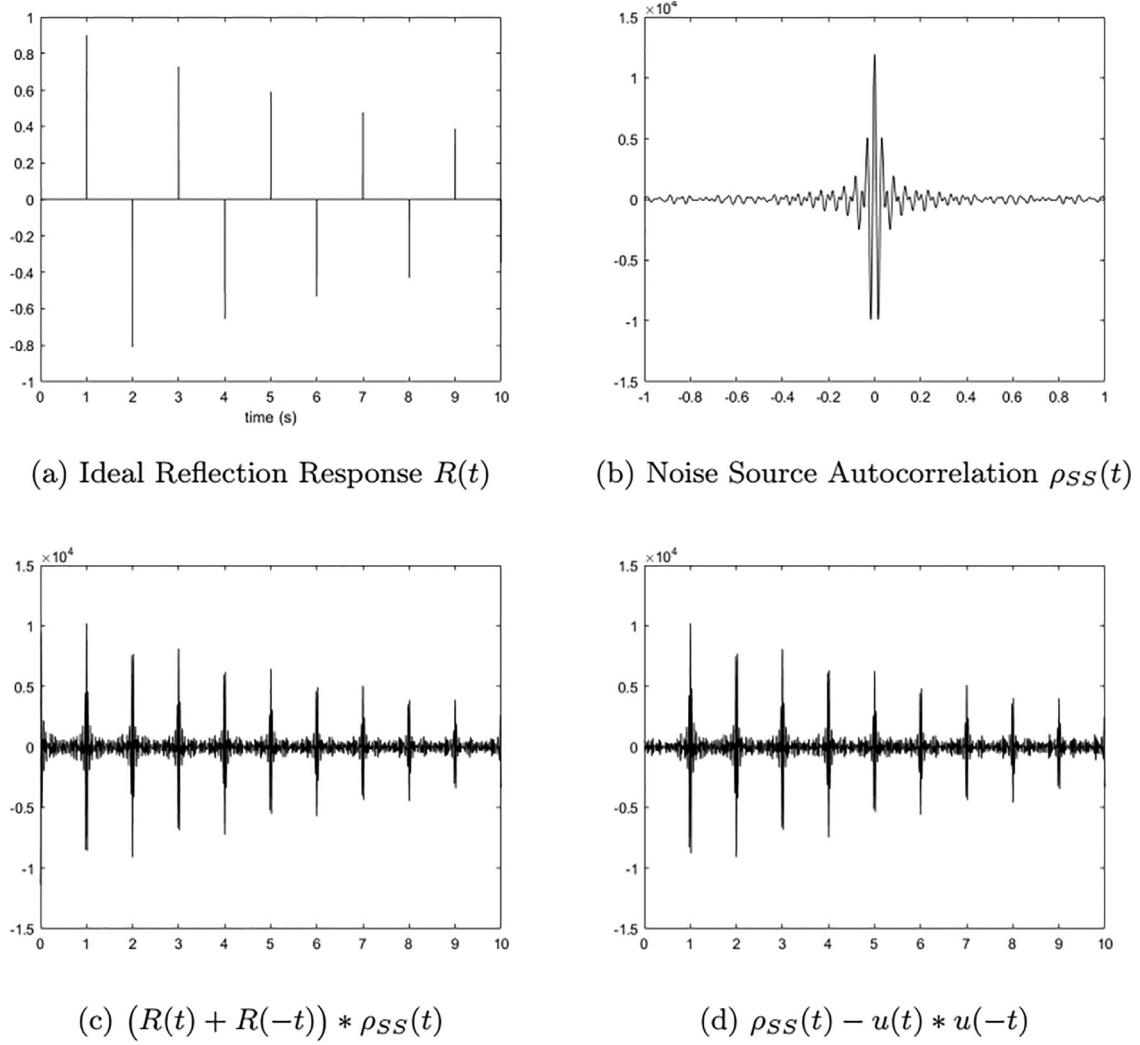


Fig. 1. Non-compressive example and illustration of terms appearing in Eq. (17).

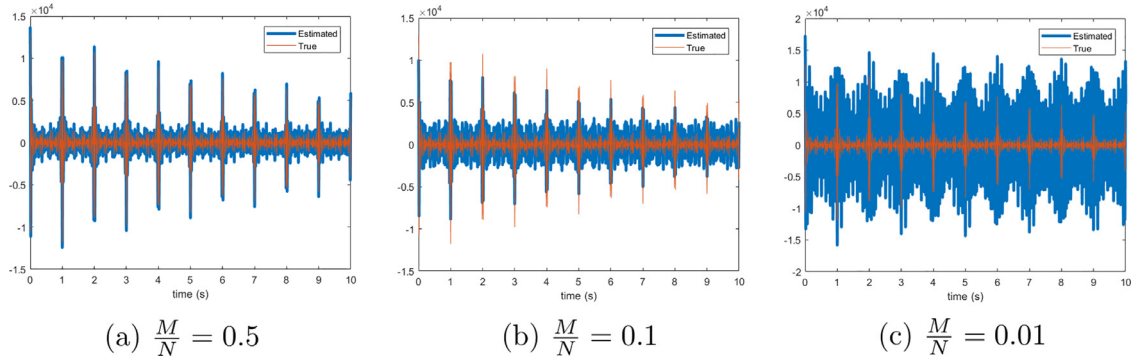


Fig. 2. Dependence of compressive estimate $\tilde{\rho}_{uu}(t)$ on subsampling ratio M/N . $\tilde{\rho}_{uu}(t)$ is shown in blue; $\rho_{uu}(t)$ is shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Consider the case of the virtual source-receiver pair formed by stations R06C and R04C. Station R04C is located 133 km west of R06C in Lone, CA. The records from these two stations overlap for the 643 days between August 6, 2005 and May 28, 2007. At a sampling rate of 1 Hz, $N_{\text{tot}} = 5.6 \times 10^7$. We set $N_{\text{seg}} = 86,400$ and $L = 643$. The non-compressive cross-correlation based on Eq. (13) is shown in Fig. 4(a). The peak value of $\Gamma_{AB}(t)$, corresponding to the time-of-arrival, occurs at 47 s, corresponding to a seismic velocity of 2.8 km/s. This is in accord with the value obtained by

Lin et al. [7]. Fig. 4(b)–(d) show the corresponding compressive estimates using compression ratios (defined as M/N_{seg}) of 0.01, 0.001, and 0.0001, respectively. While all the analyses result in the same estimate of seismic velocity, the peak is not as distinct at more aggressive compression ratios.

To quantify the effectiveness of the compressive approach, we propose the signal-to-noise ratio (SNR) between the compressive and conventional estimator as a metric. To justify this choice, we performed a set of experiments using the R06C-R04C pair: for a

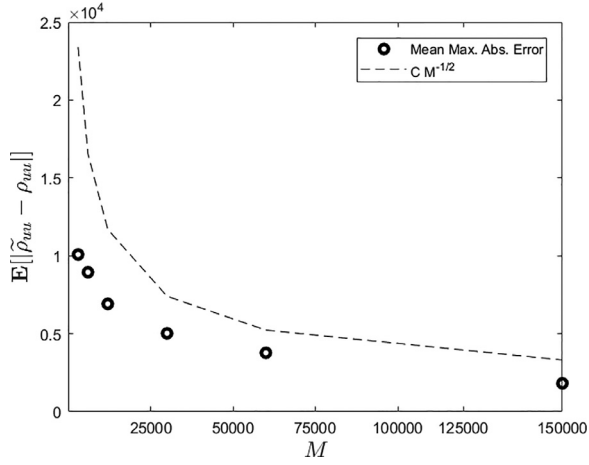


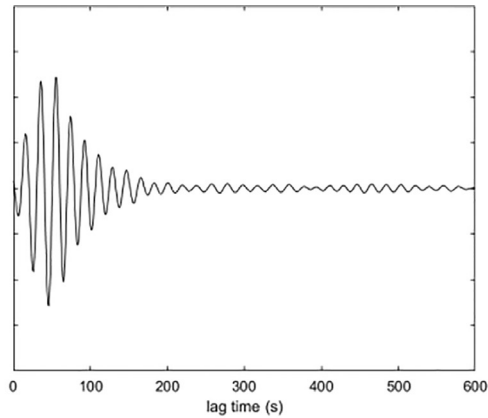
Fig. 3. Simulated mean maximum absolute error $\mathbb{E}[\sup_{t \in T} |\tilde{\rho}_{uu}(t) - \rho_{uu}(t)|]$ (shown as circles) as a function of M . The dashed line corresponds to $C/\sqrt{M}$.

given compression ratio, we randomly generated a set of M frequencies, sampled the spectra at those frequencies, compressively cross-correlated the observations, and calculated the SNR. We performed 100 simulations for each value of M for compression ratios of $1/100, 1/200, \dots, 1/10,000$. Boxplots of SNR as a function

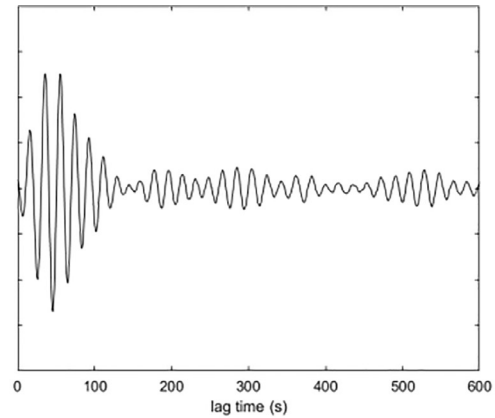
of compression ratio for these experiments are shown in Fig. 5(a). As expected, SNR gets worse as M/N gets smaller.

However, the quantity of interest in seismic interferometry usually is estimated time-of-arrival. We also measured the effectiveness of the compressive estimator by calculating the error in time-of-arrival (where time-of-arrival is the lag time at which the cross-correlation curve is maximum) with respect to the non-compressive estimate. We denote the error as e_{toa} . For our experiments, an absolute time-of-arrival error of 3 s corresponds to a relative error of 5%, a threshold we deem significant. We use this threshold to produce Fig. 5(b), which shows the probability that $|e_{\text{toa}}| > 3$ s as a function of the compression ratio. We observe that the probability that the compressive estimate is significantly different from the conventional estimate increases as M/N decreases. Fig. 5(c) then plots the probability of a significant error in time-of-arrival as a function of SNR. We observe that SNR is in fact a strong predictor of the probability of a significant time-of-arrival error, with high SNR values corresponding to low error probability. We will show how this can be used in the following discussion.

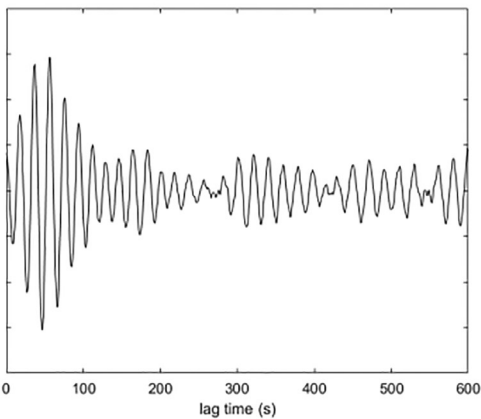
Finally, we consider the question of how, in scenarios such as this, one might optimize the choices of L and M for a given total number of compressive samples LM . Using our previous observation that the probability of error varies inversely with SNR, one should choose L and M such that SNR is maximized. To guide this choice, denote the observations made during the i th segment



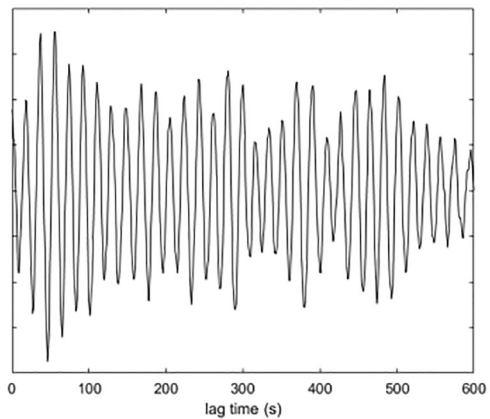
(a) Non-Compressive



(b) Compressive: $M/N_{\text{seg}} = 0.01$



(c) Compressive: $M/N_{\text{seg}} = 0.001$



(d) Compressive: $M/N_{\text{seg}} = 0.0001$

Fig. 4. Cross-correlation of overlapping records from stations R06C and R04C.

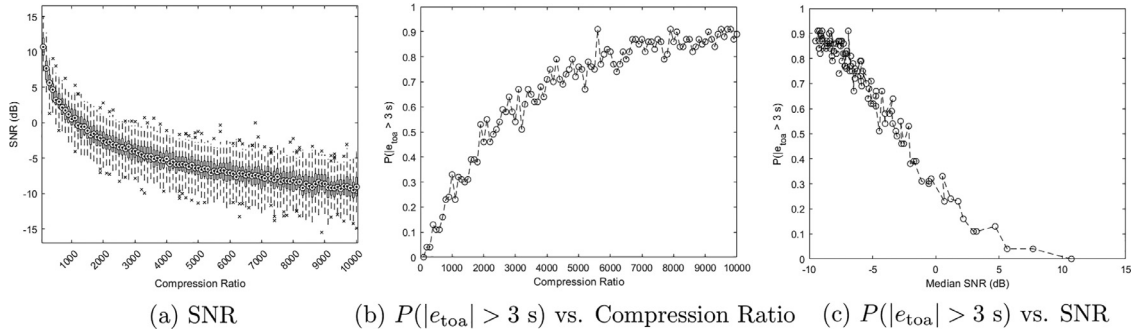


Fig. 5. Monte Carlo results: SNR and time-of-arrival error.

as $u_i^A(t)$ and $u_i^B(t)$, $i = 1, 2, \dots, L$ with the corresponding cross-correlation

$$x_i(t) = u_i^A(t) * u_i^B(-t), \quad i = 1, 2, \dots, L. \quad (21)$$

The conventional Green's function estimator from Eq. (13) becomes

$$\Gamma_{AB}(t) = \frac{1}{L} \sum_{i=1}^L x_i(t). \quad (22)$$

Because of noise in the original data and randomness in the source signal, each cross-correlation $x_i(t)$ is itself a random process, serving as an estimate of the Green's function $g(t)$, which we use as shorthand for the right hand side of Eq. (13). Suppose that $\mathbb{E}[x_i(t)] = g(t)$ for each segment i , and suppose $\sigma^2(t) = \text{Var}[x_i(t)]$ is the variance of this estimate. Then it follows that $\mathbb{E}[\Gamma_{AB}(t)] = g(t)$ as well, with $\text{Var}[\Gamma_{AB}(t)] = \frac{\sigma^2(t)}{L}$ if each segment provides an independent estimate.

Now, let $\tilde{x}_i(t)$ denote the compressive estimator of the cross-correlation of the i th segment, and denote the compressive estimator of the Green's function from Eq. (16) as

$$\tilde{\Gamma}_{AB}(t) := \frac{1}{L} \sum_{i=1}^L \tilde{x}_i(t). \quad (23)$$

Each $\tilde{x}_i(t)$ serves as an estimate of the respective cross-correlation $x_i(t)$: conditioning on $x_i(t)$, we have $\mathbb{E}[\tilde{x}_i(t)] = x_i(t)$, and in Appendix B we derive the fact that $\text{Var}[\tilde{x}_i(t)] \leq \frac{B}{M}$, where B is an upper bound on $\frac{|\Omega_+| \cdot \|\tilde{x}_i(\omega)\|_2^2}{2\pi^2}$ for all i .

Accounting for both the randomness in $x_i(t)$ and in the compressive estimator, we conclude that $\mathbb{E}[\tilde{\Gamma}_{AB}(t)] = g(t)$. Assuming independence across segments and modeling the deviations in $x_i(t)$ and $\tilde{x}_i(t)$ as independent, we also have

$$\text{Var}[\tilde{\Gamma}_{AB}(t)] \leq \frac{\sigma^2(t)}{L} + \frac{B}{LM}. \quad (24)$$

Assuming the length of each segment is fixed, then in order to maximize the SNR of $\tilde{\Gamma}_{AB}(t)$, we should minimize its variance. If the total compressive sample budget LM is fixed, then the variance in Eq. (24) can be minimized by making L (the number of segments) as large as possible, and taking only a small number M of compressive samples in each segment.

This observation is validated by the experiment shown in Fig. 6, which illustrates how the tradeoff between L and M affects SNR if the number of compressive samples (LM) is held constant. That plot shows the SNR with respect to $g(t)$, which is based on an uncompressed estimate obtained with all 643 days of data.

4.3. Deconvolution

As a final demonstration of the effectiveness and flexibility of random spectral subsampling, we present an example involving deconvolution of two signals from random samples of their frequency spectra.

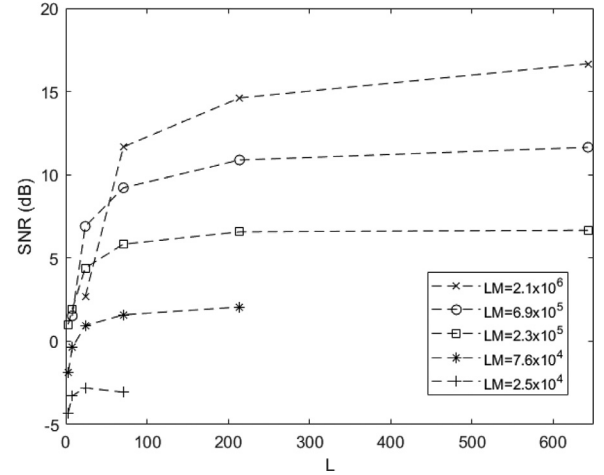
Fig. 6. Signal-to-Noise Ratio vs. number of one-day segments L . On each curve, the total number of compressive samples LM is approximately constant.

Fig. 7 shows 1000 s air pressure and displacement signals recorded at the Arenal Volcano in northern Costa Rica in 1997, along with their frequency (amplitude) spectra. As detailed in Snieder and Hagerty [23], deconvolving the pressure from the displacement gives an estimate of the displacement associated with a single pressure pulse from the volcano. Since volcanoes are very inhomogeneous, wave propagation is complicated, and the deconvolved waveforms consist of a long wave-train of scattered waves [23]. Deconvolution is a process similar to convolution, with the difference being that it corresponds to division in the frequency domain instead of multiplication. In our notation, letting $w(t)$ denote the air pressure signal and $v(t)$ denote the displacement signal, the goal in Snieder and Hagerty [23] is to compute the deconvolved signal $x(t)$ whose Fourier transform equals

$$\hat{x}(\omega) = \frac{\hat{v}(\omega) \odot \hat{w}^*(\omega)}{|\hat{w}(\omega)|^2 + \epsilon}, \quad (25)$$

where ϵ is a small regularization parameter. (One may contrast this with the relation $\hat{x}(\omega) = \hat{v}(\omega) \odot \hat{w}(\omega)$ in the convolution (1).)

Due to the similarity between convolution and deconvolution, it is straightforward to adapt our compressive estimator (4) to estimate the deconvolution of $v(t)$ and $w(t)$ from random samples of their frequency spectra. In particular, to implement the deconvolution described by (25), one can replace the product $\hat{v}[k] \hat{w}[k]$ appearing in (4) with the ratio $\frac{\hat{v}[k] \hat{w}^*[k]}{|\hat{w}[k]|^2 + \epsilon}$. To demonstrate this, we randomly select 33.3% of the spectral samples of $v(t)$ and $w(t)$ over the bandwidth $\Omega_+ = [0, 20\pi]$ rad/s and implement the compressive deconvolution estimator. The result, plotted as the blue curve in Fig. 8 over the time scale of interest -6 s to 10 s, closely matches the deconvolved signal obtained from the full pressure and velocity recordings, presented as the red curve in Fig. 8.

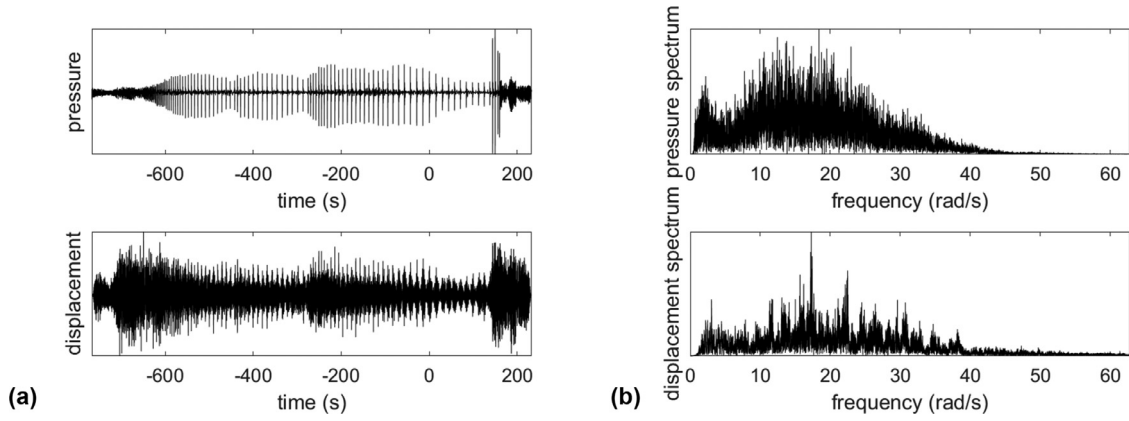


Fig. 7. Air pressure and displacement recorded at the Arenal Volcano in Costa Rica as used in the study of Snieder and Hagerty [23]. (a) Time series. (b) Magnitude of frequency spectra.

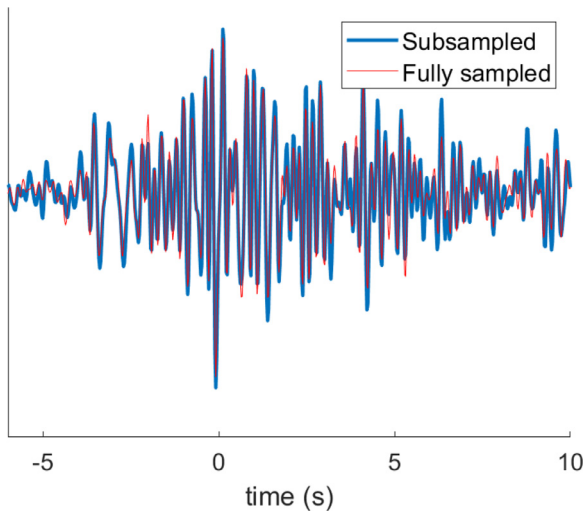


Fig. 8. Deconvolution of the Arenal pressure and displacement recordings obtained from a deconvolution using all frequencies (red curve) and from frequencies randomly subsampled with a factor 3 (blue curve). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Conclusion

We present a new method for estimating the cross-correlation of two arbitrary signals from random samples of their frequency spectra, and we demonstrate the applicability of this method to seismic interferometry. The compressive cross-correlation estimator approximates conventional cross-correlation using Eq. (4). In Theorems 1 and 2, we provide uniform bounds on the deviation between the true cross-correlation function and its compressive estimate. These theorems reveal how the maximum deviation decreases as a function of the number M of compressive samples, and give guidance on how to choose M as a function of other problem parameters. Rather than affecting the features of interest in the cross-correlation function, estimation errors tend to manifest as a “noise floor” whose magnitude increases with M . Thus, as discussed in Section 2.2, the number of samples M should be chosen so that the right-hand side of Theorem 1 is sufficiently smaller than $\sup_{t \in T} |x(t)|$.

The compressive cross-correlation estimator can be applied to seismic interferometry with data collected over multiple time segments by using Eq. (16). For a given segment length, this technique works best if we maximize the number of segments L

and then minimize the number of samples M drawn from each segment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The USArray data used in our experiments was obtained from http://ds.iris.edu/gmap/#network=_US-TA&planet=earth.

Appendix A. Proofs of Theorems 1 and 2

A1. Conversion to complex baseband problem

Recall that we assume the signals $v(t)$ and $w(t)$ are real-valued and bandlimited to the frequency interval $\Omega = [-\omega_{hi}, -\omega_{lo}] \cup [\omega_{lo}, \omega_{hi}]$ rad/s. It follows immediately that $x(t) = (v * w)(t)$ is also real-valued and bandlimited to the same interval. Recall also that our frequency-domain samples are collected from the interval $\Omega_+ = [\omega_{lo}, \omega_{hi}]$ rad/s.

To begin our proof, in this section we argue that our problem of interest—bounding the quantity $|\tilde{x}(t) - x(t)|$ —can be solved by considering a different and simpler to analyze scenario, where all signals are complex-valued and baseband, and the frequency-domain samples are collected from the baseband frequencies. We construct this argument in two steps.

1. First, define the complex version of $v(t)$ using only its positive frequency content:

$$v_+(t) = \frac{1}{2\pi} \int_{\Omega_+} \widehat{v}(\omega) e^{i\omega t} d\omega$$

and note that $v(t) = 2\text{Re}\{v_+(t)\}$. Similarly define $w_+(t)$ and $x_+(t)$, and note that $x_+(t) = (v_+ * w_+)(t)$. Let $\widehat{v}_+ = \widehat{v} \in \mathbb{C}^M$ and $\widehat{w}_+ = \widehat{w} \in \mathbb{C}^M$, and note that $\widehat{v}_+$ and $\widehat{w}_+$ correspond to vectors of random frequency-domain samples of $v_+(t)$ and $w_+(t)$, respectively, with the frequencies drawn from the full spectral support of these signals, which is Ω_+ . Now define the estimator

$$\tilde{x}_+(t) = \frac{|\Omega_+|}{2\pi M} \cdot \langle \widehat{v}_+ \odot \widehat{w}_+, \mathbf{e}_t \rangle$$

and note that $\mathbb{E}\tilde{x}_+(t) = x_+(t)$ and $\tilde{x}(t) = 2\text{Re}\{\tilde{x}_+(t)\}$. Now, we have

$$\begin{aligned}
|\tilde{x}(t) - x(t)| &= |2\text{Re}\{\tilde{x}_+(t)\} - 2\text{Re}\{x_+(t)\}| \\
&= 2|\text{Re}\{\tilde{x}_+(t) - x_+(t)\}| \\
&\leq 2|\tilde{x}_+(t) - x_+(t)|.
\end{aligned} \tag{A.1}$$

Therefore, to study the real-valued bandpass problem, it suffices to study the problem where all signals are complex-valued and bandpass, and the frequency-domain samples are collected from the bandpass frequencies.

2. As a second step, we relate the complex-valued bandpass problem to a complex-valued baseband problem. Define ω_c to be the center frequency in Ω_+ : $\omega_c = \frac{\omega_{lo} + \omega_{hi}}{2}$. Now define the signals $v_c(t) = v_+(t)e^{-i\omega_c t}$ and $w_c(t) = w_+(t)e^{-i\omega_c t}$. Both signals are baseband, with frequency-domain support on $\Omega_c = [-\omega_h, \omega_h]$, where $\omega_h = \frac{\omega_{hi} - \omega_{lo}}{2}$. Define the convolution $x_c(t) = (v_c * w_c)(t)$ and note that $x_c(t) = x_+(t)e^{-i\omega_c t}$. Let $\hat{v}_c = \hat{v}_+ \in \mathbb{C}^M$ and $\hat{w}_c = \hat{w}_+ \in \mathbb{C}^M$, and note that $\hat{v}_c$ and $\hat{w}_c$ correspond to vectors of random frequency-domain samples of $v_c(t)$ and $w_c(t)$, respectively, with the frequencies drawn from the full spectral support of these signals, which is Ω_c . For each $k = 1, 2, \dots, M$, define the baseband frequencies $\underline{\omega}_k := \omega_k - \omega_c$; we can view $\underline{\omega}_1, \underline{\omega}_2, \dots, \underline{\omega}_M$ as frequencies drawn at random from the uniform distribution on Ω_c . The underline notation emphasizes that they are baseband frequencies rather than bandpass frequencies. Similarly, define the associated baseband template vector

$$\underline{e}_t := \begin{bmatrix} e^{-i\underline{\omega}_1 t} \\ e^{-i\underline{\omega}_2 t} \\ \vdots \\ e^{-i\underline{\omega}_M t} \end{bmatrix} \in \mathbb{C}^M. \tag{A.2}$$

Finally, define the estimator

$$\tilde{x}_c(t) = \frac{|\Omega_c|}{2\pi M} \cdot \langle \hat{v}_c \odot \hat{w}_c, \underline{e}_t \rangle \tag{A.3}$$

and note that $\mathbb{E}\tilde{x}_c(t) = x_c(t)$. Now, we have

$$\begin{aligned}
|\tilde{x}_+(t) - x_+(t)| &= \left| \frac{|\Omega_+|}{2\pi M} \cdot \langle \hat{v} \odot \hat{w}, e_t \rangle - (v_+ * w_+)(t) \right| \\
&= \left| e^{-i\omega_c t} \left(\frac{|\Omega_+|}{2\pi M} \cdot \langle \hat{v} \odot \hat{w}, e_t \rangle - (v_+ * w_+)(t) \right) \right| \\
&= \left| \frac{|\Omega_c|}{2\pi M} \cdot \langle \hat{v}_c \odot \hat{w}_c, \underline{e}_t \rangle - (v_c * w_c)(t) \right| \\
&= |\tilde{x}_c(t) - x_c(t)|.
\end{aligned} \tag{A.4}$$

We see that to study the real-valued bandpass problem, it suffices to study the problem where all signals are complex-valued and baseband, and the frequency-domain samples are collected from the baseband frequencies. That is, moving forward we will focus our analysis on the estimator $\tilde{x}_c(t)$ defined in (A.3), where $\underline{\omega}_1, \underline{\omega}_2, \dots, \underline{\omega}_M$ as frequencies drawn at random from the uniform distribution on Ω_c . From our analysis above, we know that

$$\mathbb{E} \left[\sup_{t \in T} |\tilde{x}(t) - x(t)| \right] \leq 2\mathbb{E} \left[\sup_{t \in T} |\tilde{x}_c(t) - x_c(t)| \right] \tag{A.5}$$

and that

$$\mathbb{P} \left\{ \sup_{t \in T} |\tilde{x}(t) - x(t)| > U \right\} \leq \mathbb{P} \left\{ \sup_{t \in T} |\tilde{x}_c(t) - x_c(t)| > U/2 \right\}. \tag{A.6}$$

A2. Expectation and tail bounds

In this section, we focus on bounding

$$\mathbb{E} \left[\sup_{t \in T} |\tilde{x}_c(t) - x_c(t)| \right] \text{ and } \mathbb{P} \left\{ \sup_{t \in T} |\tilde{x}_c(t) - x_c(t)| > U/2 \right\}.$$

After doing so, we can use (A.5) and (A.6) to complete the proofs of Theorems 1 and 2, respectively.

Our proofs in this section follow closely the analysis of the compressive matched filter by Eftekhari et al. [12], extending their analysis from autocorrelation to cross-correlation. Similar to Eftekhari et al. [12], we make the following definitions for notational convenience:

$$M_0 = M_0(x_c) := \sup_{\underline{\omega} \in \Omega_c} |\hat{x}_c(\underline{\omega})| = \|\hat{x}_c\|_\infty$$

and

$$M_1 = M_1(x_c, M, \Omega_c) := \sqrt{\frac{M}{|\Omega_c|}} \|\hat{x}_c\|_2.$$

Now, define

$$\xi(t) := \langle \hat{v}_c \odot \hat{w}_c, \underline{e}_t \rangle = \sum_{k=1}^M \hat{v}_c[k] \hat{w}_c[k] \underline{e}_t^*[k] \tag{A.7}$$

and note that

$$\tilde{x}_c(t) = \frac{|\Omega_c|}{2\pi M} \xi(t). \tag{A.8}$$

We are interested in bounding $|\tilde{x}_c(t) - x_c(t)|$, which can be expressed as

$$|\tilde{x}_c(t) - x_c(t)| = |\tilde{x}_c(t) - \mathbb{E}[\tilde{x}_c(t)]| = \frac{|\Omega_c|}{2\pi M} |\xi(t) - \mathbb{E}[\xi(t)]|. \tag{A.9}$$

Therefore, it suffices to bound $|\xi(t) - \mathbb{E}[\xi(t)]|$. Now define the centered process

$$\psi(t) := \xi(t) - \mathbb{E}[\xi(t)] = \sum_{k=1}^M \hat{x}_c(\underline{\omega}_k) e^{i\underline{\omega}_k t} - \frac{2\pi M}{|\Omega_c|} x_c(t). \tag{A.10}$$

Let $\psi'(t)$ denote an independent copy of $\psi(t)$ generated from an independent set of samples $\underline{\omega}'_1, \underline{\omega}'_2, \dots, \underline{\omega}'_M$ drawn uniformly from Ω_c , and define

$$\zeta(t) := \psi(t) - \psi'(t) = \sum_{k=1}^M \hat{x}_c(\underline{\omega}_k) e^{i\underline{\omega}_k t} - \hat{x}_c(\underline{\omega}'_k) e^{i\underline{\omega}'_k t}. \tag{A.11}$$

Since each term in (A.11) is a symmetric random variable, $\zeta(t)$ has the same distribution as

$$\zeta'(t) := \sum_{k=1}^M \epsilon_k \left(\hat{x}_c(\underline{\omega}_k) e^{i\underline{\omega}_k t} - \hat{x}_c(\underline{\omega}'_k) e^{i\underline{\omega}'_k t} \right),$$

where $\epsilon_1, \epsilon_2, \dots, \epsilon_M$ is a sequence of independent Rademacher (± 1) random variables.

Our goal is to bound $\sup_t |\psi(t)|$, which equals $\sup_t |\xi(t) - \mathbb{E}[\xi(t)]|$. The expectation of $\sup_t |\psi(t)|$ can be controlled through the corresponding expectation of $\sup_t |\zeta'(t)|$. In particular,

$$\mathbb{E} \sup_t |\psi(t)| \leq \mathbb{E} \sup_t |\zeta'(t)|, \tag{A.12}$$

which is proved in Lemma 9 of Eftekhari et al. [12]. We may then bound the quantity $\mathbb{E} \sup_t |\zeta'(t)|$ using a chaining argument outlined in [12, Section IV] for a similar random process. Omitting the intermediate details, we arrive at the conclusion that, assuming $|\Omega_c||T| \geq 3$,

$$\mathbb{E} \sup_t |\zeta'(t)| \leq M_1 \left(4.25 \sqrt{\log(2|\Omega_c||T|)} + 2.28 \right) \tag{A.13}$$

$$\leq 5.96 M_1 \sqrt{\log(2|\Omega_c||T|)}. \tag{A.14}$$

Combining (A.5), (A.9), (A.10), (A.12), and (A.14), and using the fact that $\|\hat{x}\|_2 = \sqrt{2}\|\hat{x}_c\|_2$, we complete the proof of Theorem 1.

To establish a tail bound, we bound the deviation of $\sup_t |\psi(t)|$ from its expectation through the corresponding deviation of $\sup_t |\zeta'(t)|$. In particular, in Lemma 10 of Eftekhari et al. [12], it is proved that for any $\lambda \geq 0$,

$$\mathbb{P} \left\{ \sup_t |\psi(t)| > 2\mathbb{E} \sup_t |\psi(t)| + \lambda \right\} \leq 2\mathbb{P} \left\{ \sup_t |\zeta'(t)| > \lambda \right\}. \quad (\text{A.15})$$

We may bound $\mathbb{P} \left\{ \sup_t |\zeta'(t)| > \lambda \right\}$ using a combination of a chaining argument and an application of Bernstein's inequality, similar to the steps outlined in Eftekhari et al. [12, Section IV] for a similar random process. Omitting the intermediate details, we arrive at the conclusion that

$$\mathbb{P} \left\{ \sup_{t \in T} |\zeta'(t)| > u \right\} \leq \frac{\delta}{2}$$

where

$$u := 7.11 \max \left(M_1, M_0 \sqrt{\log(4/\delta)} \right) \sqrt{\log(12|\Omega_c||T|/\delta)}.$$

Thus, using (A.13) and (A.15), we have

$$\begin{aligned} & \mathbb{P} \left\{ \sup_t |\psi(t)| > 8.5M_1 \sqrt{\log(2|\Omega_c||T|)} + 4.56M_1 + u \right\} \\ & \leq \mathbb{P} \left\{ \sup_t |\psi(t)| > 2\mathbb{E} \sup_t |\psi(t)| + u \right\} \\ & \leq 2\mathbb{P} \left\{ \sup_t |\zeta'(t)| > u \right\} \\ & \leq \delta. \end{aligned} \quad (\text{A.16})$$

Note that

$$\begin{aligned} & 8.5M_1 \sqrt{\log(2|\Omega_c||T|)} + 4.56M_1 + u \\ & \leq \max \left(M_1, M_0 \sqrt{\log(4/\delta)} \right) \\ & \quad \cdot \left(8.5 \sqrt{\log(2|\Omega_c||T|)} + 4.56 + 7.11 \sqrt{\log(12|\Omega_c||T|/\delta)} \right) \\ & \leq \max \left(M_1, M_0 \sqrt{\log(4/\delta)} \right) \cdot \left(15.61 \sqrt{\log(12|\Omega_c||T|/\delta)} + 4.56 \right) \\ & \leq \max \left(M_1, M_0 \sqrt{\log(4/\delta)} \right) \\ & \quad \cdot \left(15.61 \sqrt{\log(12|\Omega_c||T|/\delta)} + \frac{4.56}{\sqrt{\log 36}} \sqrt{\log(12|\Omega_c||T|/\delta)} \right) \\ & \leq 18.02 \max \left(M_1, M_0 \sqrt{\log(4/\delta)} \right) \sqrt{\log(12|\Omega_c||T|/\delta)}, \end{aligned} \quad (\text{A.17})$$

where the third inequality uses the assumption that $|\Omega_c||T| \geq 3$. Combining (A.6), (A.9), (A.10), (A.16), and (A.17), and using the fact that $\|\hat{x}\|_2 = \sqrt{2}\|\hat{x}_c\|_2$, we complete the proof of Theorem 2.

Appendix B. Variance of the estimator $\tilde{x}(t)$

In this section, we derive the variance of $\tilde{x}(t)$ at any given $t \in T$. Again, we use similar arguments to [12]. Using the compressive convolution estimator (4) and the definition of variance, we have

$$\begin{aligned} \text{Var}\{\tilde{x}(t)\} &= \text{Var} \left\{ \frac{|\Omega_+|}{\pi M} \cdot \text{Re} \left[\sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right] \right\} \\ &\leq \text{Var} \left\{ \frac{|\Omega_+|}{\pi M} \cdot \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right\} \\ &= \mathbb{E} \left[\left| \frac{|\Omega_+|}{\pi M} \cdot \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right|^2 \right] \end{aligned}$$

$$\begin{aligned} &= \mathbb{E} \left[\left| \frac{|\Omega_+|}{\pi M} \cdot \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right|^2 \right] - 4|x_+(t)|^2. \end{aligned} \quad (\text{B.1})$$

Concerning the first term in (B.1), we have

$$\begin{aligned} & \mathbb{E} \left[\left| \frac{|\Omega_+|}{\pi M} \cdot \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right|^2 \right] \\ &= \frac{|\Omega_+|^2}{\pi^2 M^2} \mathbb{E} \left[\left| \sum_{k=1}^M \hat{v}[k] \hat{w}[k] e_t^*[k] \right|^2 \right] \\ &= \frac{|\Omega_+|^2}{\pi^2 M^2} \mathbb{E} \left[\left| \sum_{k=1}^M \hat{x}(\omega_k) e^{i\omega_k t} \right|^2 \right] \\ &= \frac{|\Omega_+|^2}{\pi^2 M^2} \sum_{k_1=1}^M \sum_{k_2=1}^M \mathbb{E} [\hat{x}(\omega_{k_1}) \hat{x}(\omega_{k_2}) e^{i(\omega_{k_1} - \omega_{k_2})t}] \\ &= \frac{|\Omega_+|^2}{\pi^2 M^2} \sum_{k_1=1}^M \mathbb{E} [|\hat{x}(\omega_{k_1})|^2] \\ &\quad + \frac{|\Omega_+|^2}{\pi^2 M^2} \sum_{k_1=1}^M \sum_{k_1 \neq k_2}^M \mathbb{E}_{\omega_{k_1}} [\hat{x}(\omega_{k_1}) e^{i\omega_{k_1} t}] \cdot \mathbb{E}_{\omega_{k_2}} [\hat{x}(\omega_{k_2}) e^{-i\omega_{k_2} t}] \\ &= \frac{|\Omega_+|^2}{\pi^2 M^2} \sum_{k_1=1}^M \frac{1}{|\Omega_+|} \int_{\Omega_+} |\hat{x}(\omega)|^2 d\omega \\ &\quad + \frac{|\Omega_+|^2}{\pi^2 M^2} \sum_{k_1=1}^M \sum_{k_1 \neq k_2}^M \frac{1}{|\Omega_+|^2} \left(\int_{\Omega_+} \hat{x}(\omega) e^{i\omega t} d\omega \right) \cdot \left(\int_{\Omega_+} \hat{x}(\omega) e^{i\omega t} d\omega \right)^* \\ &\leq \frac{|\Omega_+|^2}{\pi^2 M^2} \left(\frac{M}{2|\Omega_+|} \|\hat{x}(\omega)\|_2^2 + \frac{M(M-1)4\pi^2}{|\Omega_+|^2} |x_+(t)|^2 \right) \\ &= \frac{|\Omega_+| \cdot \|\hat{x}(\omega)\|_2^2}{2\pi^2 M} + 4 \frac{M-1}{M} |x_+(t)|^2. \end{aligned} \quad (\text{B.2})$$

Returning to (B.1), we have

$$\begin{aligned} \text{Var}\{\tilde{x}(t)\} &\leq \frac{|\Omega_+| \cdot \|\hat{x}(\omega)\|_2^2}{2\pi^2 M} + 4 \frac{M-1}{M} |x_+(t)|^2 - 4|x_+(t)|^2 \\ &= \frac{|\Omega_+| \cdot \|\hat{x}(\omega)\|_2^2}{2\pi^2 M} - \frac{4}{M} |x_+(t)|^2 \\ &\leq \frac{|\Omega_+| \cdot \|\hat{x}(\omega)\|_2^2}{2\pi^2 M}. \end{aligned} \quad (\text{B.3})$$

References

- [1] R. Snieder, Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E* 69 (2004) 046610, doi:10.1103/PhysRevE.69.046610.
- [2] E. Larose, L. Margerin, A. Derode, B. van Tiggelen, M. Campillo, N. Shapiro, A. Paul, L. Stehly, M. Tanter, Correlation of random wavefields: an interdisciplinary review, *Geophysics* 71 (2006) S111–S121.
- [3] K. Wapenaar, D. Draganov, R. Snieder, X. Campman, A. Verdel, Tutorial on seismic interferometry: part 1—Basic principles and applications, *Geophysics* 75 (5) (2010) 75A195–75A209, doi:10.1190/1.3457445.
- [4] J.F. Claerbout, Synthesis of a layered medium from its acoustic transmission response, *Geophysics* 33 (2) (1968) 264, doi:10.1190/1.1439927.
- [5] R.L. Weaver, O.I. Lobkis, Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Phys. Rev. Lett.* 87 (2001) 134301, doi:10.1103/PhysRevLett.87.134301.
- [6] A. Curtis, P. Gerstoft, H. Sato, R. Snieder, K. Wapenaar, Seismic interferometry – turning noise into signal, *Lead. Edge* 25 (2006) 1082–1092.
- [7] F.-C. Lin, M.P. Moschetti, M.H. Ritzwoller, Surface wave tomography of the Western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.* 173 (1) (2008) 281–298, doi:10.1111/j.1365-246X.2008.03720.x.

- [8] R. Snieder, E. Larose, Extracting earth's elastic wave response from noise measurements, *Ann. Rev. Earth Planet. Sci.* 41 (2013) 183–206. <http://arjournals.annualreviews.org/eprint/Yba7vHbEPazD9Ryqy38M/full/10.1146/annurev-earth-050212-123936>
- [9] N. Nakata, J.P. Chang, J.F. Lawrence, P. Boué, Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry, *J. Geophys. Res.* 120 (2) (2015) 1159–1173, doi:10.1002/2015JB011870.
- [10] A. Sukhovich, S. Bonnieux, Y. Hello, J.-O. Irissou, F.J. Simons, G. Nolet, Seismic monitoring in the oceans by autonomous floats, *Nat. Commun.* 6 (2015). <https://www.nature.com/articles/ncomms9027>
- [11] A.C. Lough, D.A. Wiens, A. Nyblade, Reactivation of ancient Antarctic rift zones by intraplate seismicity, *Nat. Geosci.* 11 (2018) 515–519. <https://www.nature.com/articles/s41561-018-0140-6>
- [12] A. Eftekhar, J. Romberg, M.B. Wakin, Matched filtering from limited frequency samples, *IEEE Trans. Inf. Theory* 59 (6) (2013) 3475–3496, doi:10.1109/TIT.2013.2243495.
- [13] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2) (2006) 489–509.
- [14] D.L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory* 52 (4) (2006) 1289–1306.
- [15] E.J. Candes, M.B. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.* 25 (2) (2008) 21–30, doi:10.1109/MSP.2007.914731.
- [16] R. Snieder, M.B. Wakin, When randomness helps in undersampling, *SIAM Rev.* 64 (4) (2022) 1062–1080.
- [17] D.-J. van Manen, A. Curtis, J.O. Robertsson, Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity, *Geophysics* 71 (4) (2006) SI47–SI60, doi:10.1190/1.2213218.
- [18] D. Halliday, A. Curtis, Seismic interferometry, surface waves and source distribution, *Geophys. J. Int.* 175 (3) (2008) 1067–1087, doi:10.1111/j.1365-246X.2008.03918.x.
- [19] F.-C. Lin, M.P. Moschetti, M.H. Ritzwoller, Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.* 173 (1) (2008) 281–298, doi:10.1111/j.1365-246X.2008.03720.x.
- [20] M. Behm, R. Snieder, G. Leahy, Retrieval of local surface wave velocities from traffic noise - an example from the LaBarge basin (Wyoming), *Geophys. Prosp.* 62 (2014) 223–243. <http://inside.mines.edu/~rsnieder/Behm14Retrieval.pdf>.
- [21] R. Snieder, K. Wapenaar, U. Wegler, Unified Green's function retrieval by cross-correlation; connection with energy principles, *Phys. Rev. E* 75 (2007) 036103.
- [22] G.D. Bensen, M.H. Ritzwoller, M.P. Barmin, A.L. Levshin, F. Lin, M.P. Moschetti, N.M. Shapiro, Y. Yang, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.* 169 (3) (2007) 1239–1260, doi:10.1111/j.1365-246X.2007.03374.x.
- [23] R. Snieder, M. Hagerty, Monitoring change in volcanic interiors using coda wave interferometry: application to Arenal Volcano, Costa Rica, *Geophys. Res. Lett.* 31 (9) (2004).