Communication Efficient Signal Detection for Distributed Ambient Noise Imaging

Liyan Xie, Yao Xie
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, United States
{lxie49, yao.xie}@isye.gatech.edu

Sin-Mei Wu, Fan-Chi Lin
Dept. of Geology and Geophysics
University of Utah
Salt Lake City, United States
{sinmei.wu, FanChi.Lin}@utah.edu

WenZhan Song
College of Engineering
University of Georgia
Athens, United States
wsong@uga.edu

***Abstract*—We present a signal detection scheme with low-communication overhead for the recently developed distributed ambient noise imaging (ANSI) system for real-time imaging the geophysical subsurface structure. Signal detection enables ANSI to perform adaptive data acquisition. Our method only requires each sensor to exchange a limited number of frequency samples with each other. Then each sensor directly estimates the so-called local temporal energy (LTE) using their frequency samples and received samples from another sensor, without having to form the temporal cross-correlation function. Thus, we avoid transmission of the temporal waveforms, which significantly reduce the communication overhead. Numerical examples on a real seismic dataset demonstrate the good performance of our method.**

I. INTRODUCTION

Ambient Noise Seismic Imaging (ANSI) [1], [2] is a new geophysical imaging method, which aims to capture *coherent* signals between sensors that are directly related to the geophysical structure. ANSI is a passive imaging method: it correlates ambient noises recorded at different geophone sensors, and no active sources are needed. Compared with traditional active imaging methods, such as travel-time tomography [3], where one has to rely on natural or human-induced active sources (e.g., explosion) to generate useful signals, ANSI is more environmentally sustainable.

ANSI was originally developed as an offline method using all data that have been collected. The recently developed distributed ANSI [4] is a sensor network that aims to perform *in-situ* processing of data, by allowing sensors to exchange information with each other in real-time and perform adaptive data acquisition. A key component of distributed ANSI is *signal detection*, also known as Quality Control (QC). By detecting whether or not a pair of sensors contains the coherent signal, sensors only record data when there is useful information. The signal detection for offline ANSI requires computing pairwise cross-correlation function between sensors and this is not practical for distributed system due to communication overhead.

In this paper, we present a new communication-efficient signal detection scheme for distributed ANSI. The main idea is that each sensor only communicates a limited number of frequency samples to each other, and then directly estimates the so-called local temporal energy (LTE) for signal detection. This is different from the existing centralized and offline method, which requires computing the temporal cross-correlation function and then estimate LTE. The original method is not suitable to be used in a distributed system since it requires sensors to know the complete temporal waveform of each other, and the temporal waveforms in ANSI are very long in order to accumulate enough coherence between signals. Compared with the conventional method, our method avoids directly estimating the time domain cross-correlation. Our method further reduces the communication cost by performing non-uniform downsampling within a narrow band of frequency. We test our method on a real dataset collected at Sweetwater, Texas, to show its good performance: one can achieve a similar performance detecting the coherent signal with the new scheme as the original time-domain approach.

This work is sponsored by NSF CCF-1442635.

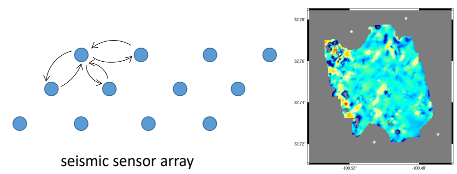


Fig. 1. Illustration of distributed ANSI system. Left: sensors communicate with each other to compute pairwise cross-correlation, to perform signal detection and then subsequent imaging. Right: an image formed by data collected by ANSI.

Notations. Given an index set Ω, denote $|\Omega|$ as the cardinality of Ω. Given a vector $\mathbf{x}$ or a matrix $\mathbf{A}$, $\mathbf{x}^\intercal$ and $\mathbf{A}^\intercal$ denote the transpose of $\mathbf{x}$ and $\mathbf{A}$. Throughout this paper, we use lower case letters to represent the signal in time domain, like $h(t)$, and use capital letters to represent the signal in frequency domain, such as $H(k)$. $|\cdot|$ denotes the modulus of a complex number. $\star$ denotes cross-correlation. X^* denotes the conjugate of a complex number X.

II. SIGNAL DETECTION IN ANSI

The ANSI system consists of multiple sensors (on the order of hundreds) that collect seismic data continuously. The first and foremost function for ANSI is *signal detection*. Although the data collected are eventually used to form images of the sub-surface structures (by solving the so-called Eikonal equation [5]). However, most of the time, the measurements do not contain the useful signal for imaging. Thus, signal detection (also called "quality control" by seismologists), as a first step, determines the period of continuous measurements and the subset of sensors that contain useful information for imaging. Our goal is to detect the existence of coherent signals between two seismic sensors, with the minimum data being transmitted between them.

A. Cross-correlation

We first examine the cross-correlation between a pair of sensors. Given a pair of seismic sensors, each making N observations $x_j(t)$, $j = 1, 2$, $t = 0, \cdots, N-1$. In the original form of ANSI, the cross-correlation between each pair of sensors is done in time domain [1]. The cross-correlation function $r(t)$, $t = 0, \pm 1, \pm 2, \ldots$ between $x_1(t)$ and $x_2(t)$ is given by

$$r(t) = x_1(t) \star x_2(t) \triangleq \sum_{u=0}^{N-1} x_1(u) x_2(t+u).$$

In ANSI, N is typically a very large number, because we have to rely on very long sequences to accumulate enough energy for the coherent signal.

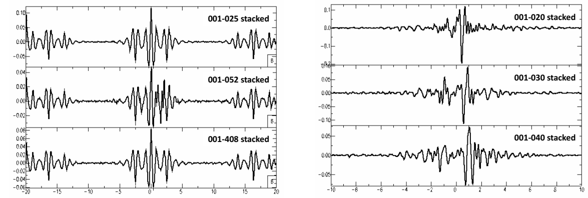
B. Composite null signal detection

The signal detection in ANSI is done through *local temporal energy* (LTE) of the cross-correlation function. which measures the energy of the cross-correlation function $r(t)$ around certain *lag time*. The lag time is determined as the expected arrival of the signal based on the station distance and the expected group velocity of the surface wave. For example, for a station pair with 1-kilometer spacing, if we expect the group velocity to be between 0.5 and 1.0 km/sec, then one can define 1 to 2 second time window as the expected signal window.

A "good" pair of sensors that have signal will have concentrated energy around certain lag, while "bad" pairs have a large amount of energy far from that lag. Fig. 2 illustrates three cases with signals (the cross-correlation function between pairs of sensors has *one* clear peak near the lag around zero), and three cases without a signal (the cross-correlation function has significant energy away from the lag 0). In the figure, the sensor indices are denoted by 001, 020, 030, 040, etc. "001-025" denotes the cross-correlation between this pair of sensors.

LTE is computed by applying time-window filtering to $r(t)$ and then compute the sum-of-squares value of the windowed signal. This can be done for various centers of the window to measure LTE at different locations, and then we may obtain a LTE profile function, as shown in Fig. 3.

The signal detection problem can be viewed as a hypothesis testing on the LTE profile function and it has a *composite null*. The alternative hypothesis is that there is a coherent waveform that is observed by two sensors (with a relative delay; because the noise sources propagate to two different sensors taking different times). Two sensors observe a common nominal



H_0: No Signal $\qquad$ H_1: Signal

Fig. 2. Cross-correlation between station pairs. *Left*: three "good" pairs that contain the signal concentrated around zero lag; *Right*: three "bad" pairs without a useful signal, where there is still a large amount of energy at large lag that is far from zero.

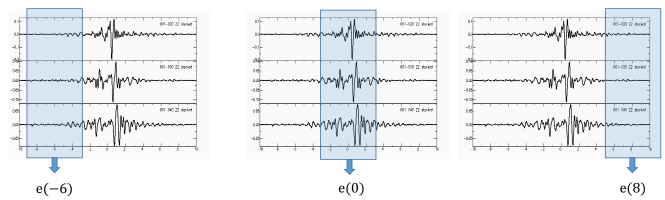


Fig. 3. Time-window filtering used to estimate $e(t_0)$, the LTE of cross-correlation function within a specific time window centered at t_0.

waveform (with a relative delay τ, the "lag" in seismology) buried in noise

$$\mathsf{H}_1 : \begin{cases} x_1(t) = \rho_1 s(t) + n_1(t), \\ x_2(t) = \rho_2 s(t-\tau) + n_2(t), \end{cases} \tag{1}$$

where ρ_1 and ρ_2 are the magnitudes observed by two sensors, $s(t)$ is an unknown deterministic unit-energy signal, $\int s^2(t)dt = 1$, $n_1(t)$ and $n_2(t)$ are zero-mean noises that are independent of each other. Thus their cross-correlation function will have *one* clear peak around certain lag above the noise floor when the two waveforms are properly aligned. The null hypothesis of "no signal" can be viewed as when either $r(t)$ consists of only noise, or there is significant energy outside of the vicinity of the lag.

Based on this, a signal is detected, if the LTE around the lag exceeds a threshold $b > 0$ and there is no significant LTE far from the lag (below another threshold $b' > 0$)

$$\{e_0 > b\} \quad \text{and} \quad \{e_i < b'\}, i = 1 \ldots, M, \tag{2}$$

$e_0 = \sum_{t \in \mathcal{S}_0} |r(t)|^2$, $e_i = \sum_{t \in \mathcal{S}_i} |r(t)|^2$ where $\mathcal{S}_0$ is a pre-specified window around the lag and $\mathcal{S}_i$, $i = 1, \ldots, M$ defines the two sides of M regions that are away from the lag.

Hence, to perform signal detection in ANSI, one need to estimate LTEs $\{e_0, e_1, \ldots, e_M\}$. Existing signal detection in ANSI is done in the time domain, which first needs to estimate the cross-correlation function $r(t)$. Next, we present a new approach, which avoids estimating $r(t)$ but rather directly estimate LTEs using limited frequency samples to lower the communication cost.

III. Communication Efficient Algorithm

In the distributed setting, we would like sensors to communicate as little data as possible to achieve the detection goal. We thus let sensors to transmit the sub-sampled frequency components of their signals, and our goal is to perform signal detection with as limited communication as possible. The new detection scheme consists of three steps: (1) Each sensor computes the Fourier transform (FFT) of their signals and performs band-pass (BP) filtering; (2) Sensors communicate a subset of non-uniform samples of their frequency components; (3) Each sensor multiplies its frequency samples with the received samples from another sensor to recover the LTE, and to decide whether or not there is a coherent signal between them.

A. Cross-correlation in frequency domain

It is well known that the cross-correlation in the time domain is equivalent to multiplication in the frequency domain. Define the Fourier transform of signals $x_j(t)$ as

$$X_j(k) = \sum_{n=0}^{N-1} x_j(n) e^{-i\frac{2\pi nk}{N}},\ k = 0, \ldots, N-1,$$

where i is the imaginary number. Let $R(k)$ be the Fourier transform of $r(t)$, by convolution theorem we have

$$R(k) = X_1^*(k) X_2(k), \quad k = 0, 1, \ldots, N-1. \tag{3}$$

To improve computational efficiency, we can use Fast Fourier Transform (FFT). The complexity of FFT is $O(N \log(N))$, while the complexity of directly computing cross-correlation by convolution is $O(N^2)$.

B. Communication efficient scheme

Typically geophysicists are only interested in a pre-specified narrow band of frequencies that are linked to particular imaging depth into the ground [1]. Hence, instead of transmitting the raw signal waveform, each sensor can only transmit a few frequency components within the targeted frequency band. To further reduce communication cost, one can perform non-uniform downsampling at each sensor, and only transmit a subset of frequency components in that band. As a consequence of non-uniform downsampling, one can no longer perform inverse FFT to go back to the time domain.

We propose to use the following scheme. Each sensor computes the discrete Fourier transform of their local signal $x_j(t)$. Then performs narrow band-pass filtering to keep only a range of frequency components that we are interested in. Then each sensor performs downsampleing to further reduce the frequency samples in that range. Let

$$\Omega = \{f_1, \cdots, f_m\} \subset \{0, 1, \cdots, N-1\}$$

denote the indices of the downsampled frequency components, which is the same across all sensors. Finally, sensor j only transmits the subset of frequency samples $\{X_j(f_k)\}$, where $f_k \in \Omega$.

Assume $m = |\Omega|$ frequency samples are selected. Then this reduces the amount of data to be transmitted from $\mathcal{O}(N)$ to $\mathcal{O}(m)$. Usually, in ANSI, $N \gg m$, because we need to observe a long enough noise sequence to observe coherent signal; however, the frequency band we are interested is usually narrow. For instance, in our real data example, $N = 15,000$, and m ranges from 120 to 1200 (with downsampling), which corresponds to 99.2% $\sim$ 92% reduction in the amount of data to be communicated. This reduction is for each sensor. In a large sensor network, this reduction will significantly reduce the total communication overhead.

C. Estimate local temporal energy from frequency samples

Once receiving the frequency samples from other sensors, each sensor needs to perform signal detection from estimated cross-correlation. However, since each sensor only has a set of incomplete frequency components, each sensor only knows a partial cross-correlation function by multiplying received frequency components with its own. It only knows $\{R(f_k)\}$, for $f_k \in \Omega$. Hence, one cannot perform the conventional inverse FFT to recover the time-domain $r(t)$ from the incomplete frequency components. Thus, one cannot directly estimate the local temporal energy using $r(t)$.

Now we present a method to estimate the local temporal energy from $\{R(f_k)\}$, $f_k \in \Omega$. Define a time-window function as $g(t - t_0)$, where t_0 is the center. For instance, $g(t)$ is a rectangular function with unit area. We consider the Gaussian filter since its Fourier transform is also a Gaussian function and smooth:

$$g(t) = \sqrt{\frac{a}{\pi}} e^{-at^2}, \tag{4}$$

where $1/2a$ is the variance of the Gaussian filter and we set the window length to be $2/\sqrt{2a}$, which means we only measure the LTE in a narrow time interval $[t_0 - 1/\sqrt{2a}, t_0 + 1/\sqrt{2a}]$.

The windowed temporal cross-correlation function within time window centered at t_0 is given by

$$h(t) = r(t) g(t - t_0), \quad t = 0, 1, \ldots, N-1. \tag{5}$$

The LTE within this time window can be estimated from (5) as

$$\hat{e}(t_0) = \sum_{t=0}^{N-1} h^2(t), \quad t_0 = 0, \pm 1, \ldots.$$

So far, we have found an expression for the LTE in the time domain. Next, we will use the time-frequency duality to find an expression in terms of frequency samples. For each time window centering at t_0, let $G(k)$ denotes the Fourier transform of $g(t)$, thus $G(k)e^{-i2\pi k t_0/N}$ is the Fourier transform of $g(t - t_0)$. Let $R(k)$ be the Fourier transform of $r(t)$, due to time-frequency duality, we have

$$H(k) = R(k) \star (G^*(k) e^{i2\pi k t_0/N}). \tag{6}$$

Since Fourier transform is an isometry by Plancherel theorem, we can estimate LTE, equivalently, as

$$\hat{e}(t_0) = \sum_{k=0}^{N-1} |H(k)|^2. \tag{7}$$

Hence now we only need to estimate $H(k)$. Below, we will use the relation (6), and assume that $H(k)$'s are smooth, which is reasonable since it is resulted from filtering by a window function.

Since we only know a subset of frequency components for $R(k)$, to use (6) and (7), we need to deal with missing frequency samples. Similar to [6], we use total variation (TV) norm minimization problem. Define $\tilde{G}(k) = G^*(k)e^{i2\pi k t_0/N}$, and formalize vectors $\boldsymbol{H}$, $\boldsymbol{R}$ and the convolution matrix $\mathbf{G}$,

$$\mathbf{G} = \begin{bmatrix} \tilde{G}(0) & \tilde{G}(-1) & \cdots & \tilde{G}(1-N) \\ \tilde{G}(1) & \tilde{G}(0) & \cdots & \tilde{G}(2-N) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{G}(N-1) & \tilde{G}(N-2) & \cdots & \tilde{G}(0) \end{bmatrix} \in \mathbb{R}^{N \times N},$$

$$\mathbf{H} = [H(0), \cdots, H(N-1)]^\intercal \in \mathbb{R}^{N \times 1},$$

$$\mathbf{R} = [R(0), \cdots, R(N-1)]^\intercal \in \mathbb{R}^{N \times 1}.$$

Then to estimate LTE centered at t_0, equation (6) which corresponds to

$$\mathbf{H} = \mathbf{G}\mathbf{R}. \tag{8}$$

Once we can estimate $\widehat{\mathbf{H}}$ that satisfies (8), we can use (7) to estimate the LTE at each window location t_0.

With downsampling, we only have samples $X_j(f_1), \ldots, X_j(f_m)$, $j = 1, 2$, from which we only know $R(f_1), \cdots, R(f_m)$, which only corresponds to a subset of entries of the $\mathbf{R}$ vector. We consider two approaches to estimate $\widehat{\mathbf{H}}$. The first simple approach is to fill zeros for the missing entries of $\mathbf{R}$, and solve $\mathbf{H}$ using (8). The second approach considers a regularized solution, satisfying the data constraint.Define

$$\mathbf{R}_\Omega = [R(f_1), \cdots, R(f_m)]^\intercal \in \mathbb{R}^{m \times 1}.$$

let $\mathbf{G}_\Omega \in \mathbb{R}^{N \times m}$ denotes the matrix consisting of all columns of $\mathbf{G}$ at indices from Ω. Then the zero-filling solution should satisfy $\mathbf{H} = \mathbf{G}_\Omega \mathbf{R}_\Omega$, or $\mathbf{G}_\Omega^\intercal \mathbf{H} = \mathbf{G}_\Omega^\intercal \mathbf{G}_\Omega \mathbf{R}_\Omega$. This creates m linear constraints on the variables. We then find the solution that has the smallest total variation norm:

$$\begin{aligned} \underset{\mathbf{H}}{\text{minimize}} \quad & \mathrm{TV}(\mathbf{H}), \\ \text{subject to} \quad & \mathbf{G}_\Omega^\intercal \mathbf{H} = \mathbf{G}_\Omega^\intercal \mathbf{G}_\Omega \mathbf{R}_\Omega. \end{aligned} \tag{9}$$

Here $\mathrm{TV}(\mathbf{H}) = \sum_i |H(i+1) - H(i)|$ is the total variation (TV) norm of $\mathbf{H}$, which can be viewed as the ℓ_1 norm of its derivative [7]. The motivation for using TV norm is that we would like to find a piece-wise constant solution. As one can imagine, the "true" energy function should have a set of non-zero values in a window centered about the true "lag", and be nearly zero outside of that window. The convex optimization problem here can be solved using standard optimization softwares such as CVX [8].

IV. Numerical results on real-data

We test our method on a real dataset. The dataset consists of 37 days of records (from 2016/09/16 to 2016/10/22) at Sweetwater, Texas. There are seven sensors in this dataset, with sensor indices given by 001, 020, 030, 040, 025, 052 and 408. We form six pairs by letting sensor 001 to communicate with the remaining six sensors. The sampling rate of each sensor is 50 Hz, which means there are 50 temporal samples in a second.

For data in one day, we divide samples into 288 blocks; each block contains a 5-minute length data since the lag time is less than 5 minutes by our prior knowledge. For each pair of sensors, we compute the cross-correlation function using the 5-minute data, and then the final results are averaged over 288 blocks. The resulted average temporal cross-correlation function are shown in Fig. 2. Sensors 020-001, 030-001 and 040-001 are "good" pairs with a coherent signal, while 025-001, 052-001 and 408-001 are "bad" pairs with no coherent signal. We use band-pass filtering to keep only 1-3 Hz content. Below, the time window function is Gaussian (4) with width 1.2 seconds.

For each 5-minute data, for each LTE centered at t_0, the computation time for zero-filling least-square is less than 1 second, and for the total variation norm minimization is 7.68 second on average (when the down-sampling rate is 0.8). Different LTE centers can be done in parallel. Thus our algorithm can be run on-the-fly.

A. Recovery of local temporal energy

We first compare the LTE profile function that are (i) obtained directly from time-domain cross-correlation function $r(t)$, as the conventional method did; (ii) "zero-filing" least-square solution obtained from solving (8), and (iii) TV minimization solution from solving (9). We test the methods on three pairs of "good" sensors that contain a coherent signal, and three pairs of "bad" sensors that do not contain a coherent signal. The results are presented in Fig. 4. For "good" pairs (left column of the figure) and top two "bad" pairs, both communication efficient methods have reasonable performance, and the TV minimization approach seems to recover the LTE function better. For some "bad" pairs with low SNR (such as 001-408), it is harder to recover the original LTE, however, the detection results are still consistent with the original method, since the the recovered LTE indeed has a high energy out of the "lag" near 0.

B. SNR versus downsampling rate

To quantify performance, we define a measure SNR $= H_s/H_n$ as illustrated in Fig. 5. Here H_s is the mean LTE in time windows near the true lag (which is 0 in our cases) and H_n is the mean LTE in time windows outside of 0.

We further study how the downsampling rate affects the performance of the proposed method. From results shown in Fig. 6, we can see that as we reduce the downsampling rate, SNR decreases and it becomes more difficult to tell apart "good" and "bad" pairs due to more information losses. The "zero-filling" method fails to tell apart "good" and "bad" pairs even when the downsampling rate is as high as around 0.8, but the TV-min method works well when the downsampling rate is higher than 0.4.

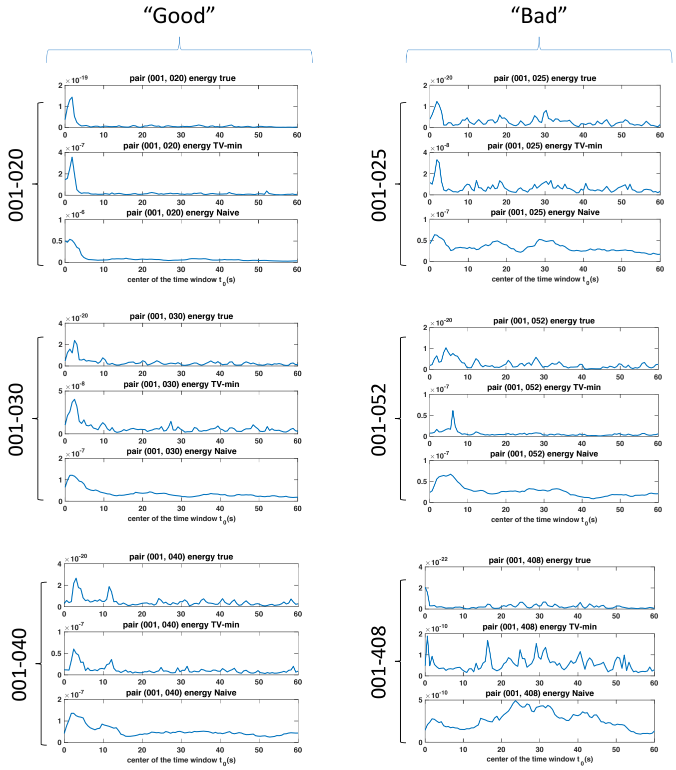


Fig. 4. The local temporal energy (LTE) at downsampling rate 0.8, for the six pairs of cross-correlation functions in Fig. 2. For each panel: *Upper*: “true” LTE directly computed from $r(t)$. *Middle*: LTE from TV minimization. *Lower*: LTE from “zero-filling”.

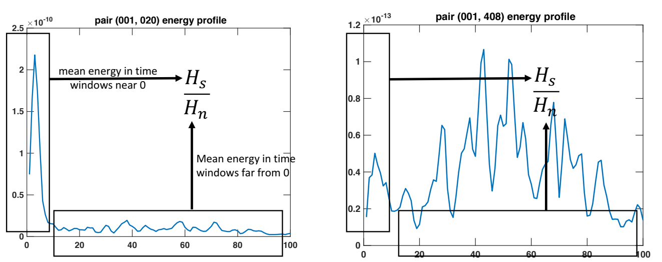
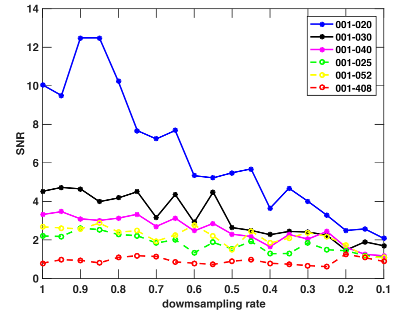


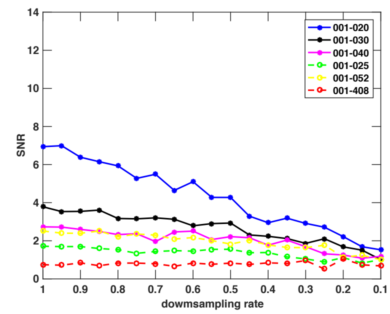
Fig. 5. Definition of SNR $= H_s/H_n$. Left: SNR is high for “good” pairs. Right: SNR is low for “bad” pairs.

V. Summary and Discussion

In this paper, we present a communication-efficient signal detection scheme for Ambient Noise Seismic Imaging systems. Detecting whether a pair of sensors contains the coherent signal is an essential part of ANSI. Since detection is based on cross-correlation function, sensors need to communicate with each other. Without a communication-efficient scheme, sensors transmit raw signals, which can be extremely inefficient. We exploited the time-frequency duality and presented a scheme where each sensor only sends a limited number of frequency components. Furthermore, since the detection is based on the *local temporal energy* (LTE), we do not aim to recover the cross-correlation function, but rather present a novel approach to directly estimate LTE from limited frequency samples. We tested two recovery methods based on simple “zero-filling” and total variation norm minimization on real data. Both methods perform reasonably well, with the TV minimization approach give higher SNR gain in certain cases. Ongoing work includes developing real-time signal detection and decentralized detection with limited frequency samples.



(a): SNR versus downsampling rate for TV-min method



(b): SNR versus dowbsampling rate for Naive method

Fig. 6. Performance of the communication efficient method. The metric SNR $= H_s/H_n$ over different downsampling rates.

References

[1] Fan-Chi Lin, Morgan P Moschetti, and Michael H Ritzwoller, “Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps,” *Geophysical Journal International*, vol. 173, no. 1, pp. 281–298, 2008.

[2] GD Bensen, MH Ritzwoller, MP Barmin, AL Levshin, F Lin, MP Moschetti, NM Shapiro, and Yanyan Yang, “Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements,” *Geophysical Journal International*, vol. 169, no. 3, pp. 1239–1260, 2007.

[3] JA Hole, “Nonlinear high-resolution three-dimensional seismic travel time tomography,” *Journal of Geophysical Research: Solid Earth*, vol. 97, no. B5, pp. 6553–6562, 1992.

[4] Maria Valero, Jose Clemente, Goutham Kamath, Yao Xie, Fan-Chi Lin, and WenZhan Song, “Real-time ambient noise subsurface imaging in distributed sensor networks,” in *Smart Computing (SMARTCOMP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–8.

[5] Fan-Chi Lin, Michael H Ritzwoller, and Roel Snieder, “Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array,” *Geophysical Journal International*, vol. 177, no. 3, pp. 1091–1110, 2009.

[6] Yonina C Eldar, “Analog compressed sensing,” in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 2949–2952.

[7] Leonid I Rudin, Stanley Osher, and Emad Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[8] Michael Grant, Stephen Boyd, and Yinyu Ye, “Cvx: Matlab software for disciplined convex programming,” 2008.