Plane-wave extraction from the Stanford DAS array: adjoint and inverse

Siyuan Yuan and Biondo Biondi

ABSTRACT

We develop a beamforming algorithm (adjoint method) to extract compressional plane wave and shear plane wave components from distributed acoustic sensing (DAS) data. We also set up inverse methods regularized by the L1 and L2 norm. To understand the performance of beamforming and our inversions, we test and compare them on synthetic wavefields composed of compressional and shear waves. We consider two scenarios regarding the frequency content of the waves: impulsive plane waves and reverberant plane waves. We found that beamforming can give us a good estimation in the impulsive wave case, but it does not work well in the reverberant case. Also, beamforming is subject to crosstalk artifacts between the compressional wave and shear wave model spaces. In all our test cases, the inverse algorithm regularized by the L1 norm did the best job giving us highly focused results with no clear crosstalk artifacts.

INTRODUCTION

The Stanford DAS array has been continuously recording ground motion since September, 2016. We have cataloged more than 3,000 earthquakes in Northern California. From this rich dataset, we are interested in extracting earthquake-characterizing patterns that can be fed into machine learning algorithms to detect new earthquakes and understand the near-surface condition around our DAS array. Earthquake data tend to be dominated by plane waves, especially when their epicenters are fairly far from DAS. We aim to dig into the data and see what the dominating plane waves are; where these plane waves come from, what are their apparent speeds, and what are the compositions of compressional and shear waves. Beamforming is a common signal-processing algorithm for earthquake detection and location.

Stanford Exploration Project (SEP) has been studying beamforming since the 1980s. Cole (1995) applied beamforming to a quarry blast dataset recorded by a geophone array showing that beamforming gives clear pictures of where and at what apparent velocities the events arrive. Karrenbach and Cole (1989) extended their beamsteering method to accommodate vector data and multi-component arrays. Kostov (1989) derived algorithms to calculate the least-squares inverses of the slant-stack. Compared to traditional geophones, DAS has a much denser channel distribution. Beamforming

can potentially take advantage of these dense recordings by DAS to give us satisfactory feature extraction. Moreover, the beamforming algorithms used by Cole (1995) only work in the case of compressional waves. The Stanford DAS array records both compressional waves and shear waves (Biondi et al., 2017). Based on the fiber optic's plane-wave response theory developed by Martin et al. (2017), we generalize beamforming to deal with plane waves composed of both compressional waves and shear waves.

Beamforming is the adjoint operator of plane-wave superposition. Inverse versions of beamforming may give us more precise results than beamforming does. Here, we regularize the inverse by the L2 norm and by the L1 norm of the model respectively. Together with the generalized beamforming algorithm, we compare their plane-wave extraction performances on plane-wave synthetic data.

BEAMFORMING AND INVERSE ALGORITHMS THEORY

To set up a mathematical problem, we denote the DAS data we have by d, which has two dimensions: channel number and arrival time. To generalize beamforming to deal with both compressional waves and shear waves, we define the model space in two parts: the compressional-wave model space (m_p) and the shear-wave model space (m_s). For each of the model spaces, we are in three dimensions: azimuth (rad), apparent slowness (s/m) and interception time (s). The range of azimuth is from 0 to 2π with a spacing of 0.1047 rad (6°). The range of the apparent slowness is from 0 to 0.002 s/m (velocity range: 500 m/s to infinity). Each model point refers to the amplitude of one particular type of plane wave (compressional or shear waves) propagating with the corresponding azimuth angle and slowness reaching a reference point at an interception time τ. Martin et al. (2017) develop the fiber responses of the plane-wave propagation in both shear wave and compressional wave cases. Based on the theory, we can calculate the impulse response for each model point. Ensembling the impulse responses, we set up forward modeling operators, F_p and F_s for m_p and m_s respectively. Combining both of the contributions from the compressional waves and from shear waves on the data, we have:

$$d = \begin{bmatrix} F_p & F_s \end{bmatrix} \begin{bmatrix} m_p \\ m_s \end{bmatrix}.$$

Let $F = \begin{bmatrix} F_p & F_s \end{bmatrix}$ and $m = \begin{bmatrix} m_p & m_s \end{bmatrix}^T$ We have:

$$d = Fm.$$

To estimate m, we have a choice of using the adjoint method (beamforming):

$$m^* = F^T d.$$

The adjoint method is computationally cheap but could give us only a rough estimation. Solving inverse problems is more expensive but could give us results with

a higher resolution. The simplest inverse problem can be set up as the following optimization problem:

$$m^{inv} = \underset{m}{\operatorname{argmin}} \|Fm - d\|_2^2.$$

Here we have a huge model space compared to the data space; we may live in the null space. For earthquake data, we expect to have dominating plane waves and in turn to have a sparse model space. Adding a regularization term can shrink our model space and alleviate the null space problem. Adding a L2 norm regularization term is a common option. The optimization problem is:

$$m_{L2}^{inv} = \underset{m}{\operatorname{argmin}} \|Fm - d\|_2^2 + \lambda \|m\|_2^2.$$

Using L1 norm regularization is, in theory, more ideal to get a sparse results shrinking most of the irrelevant model points to be zero. The optimization problem is:

$$m_{L1}^{inv} = \underset{m}{\operatorname{argmin}} \|Fm - d\|_2^2 + \lambda \|m\|_1.$$

Data modeling and inverse operators all depend on the channel layout. Figure 1 shows the layout of the Stanford DAS array. To simplify the data simulation and the setup of testing algorithms, we simplified the channel layout as shown in Figure 2. As can be seen, in the simplified map, fibers extend in two directions (shown in color blue and color red, respectively). The six lines in the map are denoted by letters from A to F. Channel numbers increase from south to north and from east to west in red lines and blue lines respectively.

In the following sections, beamforming and the inverse algorithms with regularizations shown above are examined and compared in the cases of different types of wavefields.

BEAMFORMING/INVERSE ON IMPULSIVE WAVE FIELDS

With the methodology set up in the previous section, we are ready to generate synthetic data for testing the algorithms. The simplest case is that we have only an impulse in the model space. An impulse in the model space corresponds to an impulsive wavefield received by the DAS array in the data space. For earthquake recordings, we tend to have shear waves and compressional waves coming from similar directions but with different apparent velocity. Also, the energy of the shear wave is often stronger than the compressional wave. With that in mind, when we were simulating the test data , we assumed there is a compressional plane wave propagating from the south-west to the north-east (0.7854 rad, 45°) with an apparent slowness to be 0.001 seconds/meter (1,000 meters/second). At the same time, we modeled a shear wave coming from the same direction but with a greater slowness of 0.0015 seconds/meter (667 meters/second). The two waves in the m_p and m_s are shown in Figure 3a

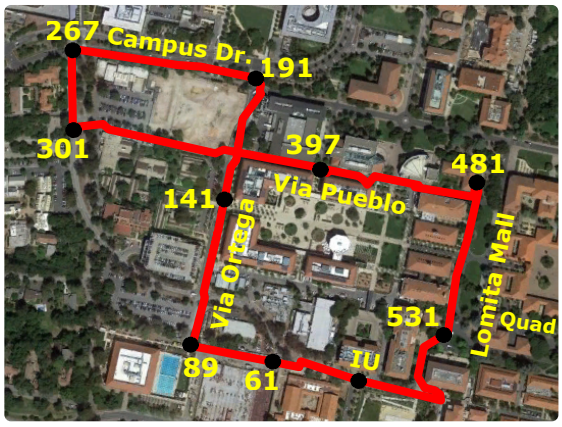


Figure 1: The layout of the fiber following telecommunications conduits overlaid on the map. The longest linear section is roughly 600 meters wide. Some deviations from straight lines had to occur due to existing conduit geometry constraints. [**NR**]

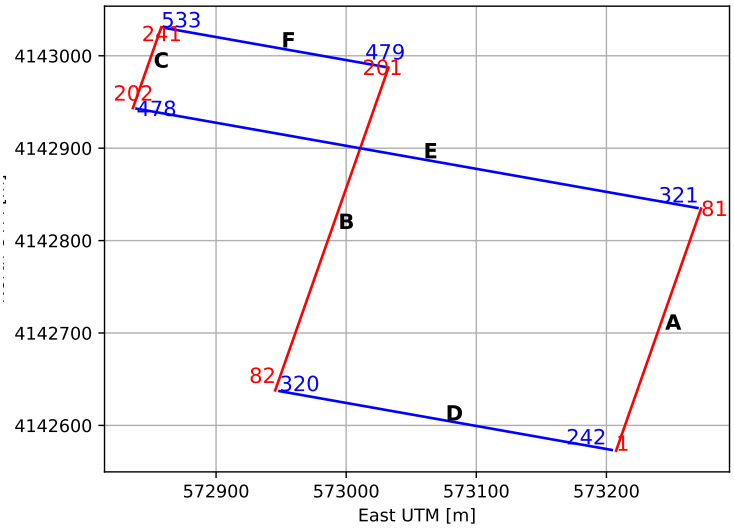


Figure 2: The layout of the simplified DAS channel map for data simulation and algorithms testing purposes. The fibers are extended into two directions colored in blue and red respectively. 534 channels are distributed with a spacing of 4 meters along each line in the plot. The six lines in the map are denoted by letters from A to F. The channel numbers at the ends of the six lines are labeled. [**ER**]

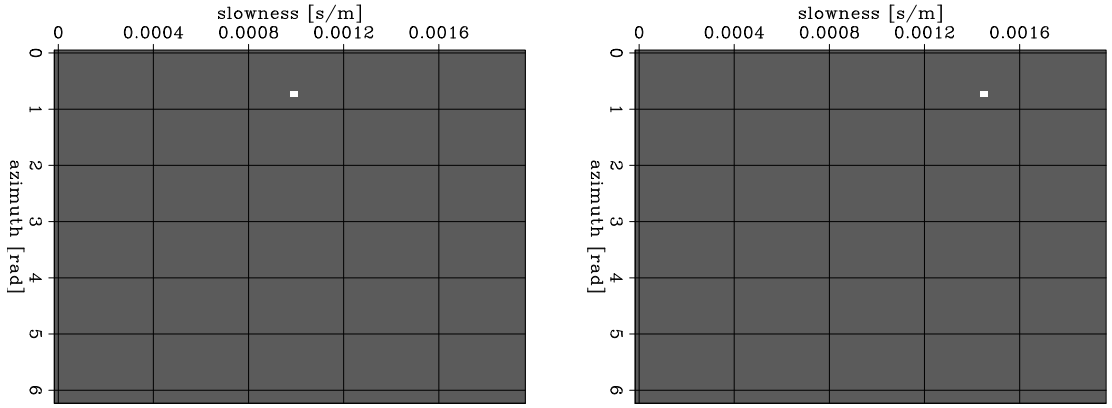


Figure 3: (Left) Compressional-wave model space.The impulse has an azimuth angle of 0.7854 rad (propagating from south-west to north-east) and has a slowness of 0.001 s/m (1000 m/s); (Right) Shear-wave model space. The azimuth angle is the same as the compressional wave's azimuth angle. The slowness is 0.0015 s/m (667 m/s). The amplitude of the shear wave impulse is three times stronger than that of the compressional wave. [**ER**]

and Figure 3b respectively. The amplitude of the shear-wave impulse is three times stronger than that of the compressional wave.

Applying the forwarding operator F to the impulses shown in Figure 3a and Figure 3b, we obtain the data with the compressional wave and the shear wave interfering with each other shown in Figure 4. The compressional wave data is shown as the six light grey oblique lines, while the shear wave is shown as three black lines from channel number 0 to 230 and three white lines after channel 230. Whiter signal means more positive amplitude. Because the amplitude of the shear wave is three times higher than the compressional wave, the intersections of the two types of wave are more toward to the shear wave signal. Thus, recovering the compressional wave can be a challenge.

Beamforming results

Appying beamforming to the data, we recover the compressional wave model and the shear wave model shown in Figure 5a and Figure 5b respectively. Compared with the true model points shown in Figure 3a and Figure 3b, we see that we have bright spots focused on the correct positions, which means that both compressional wave and shear wave are recovered fairly well. But the downside is that we see a clear crosstalk problem in the compressional model space shown in color black in Figure 5a at around the location where the shear wave impulse is. And the compressional wave is recovered with a lower signal to noise ratio when compared to the shear wave.

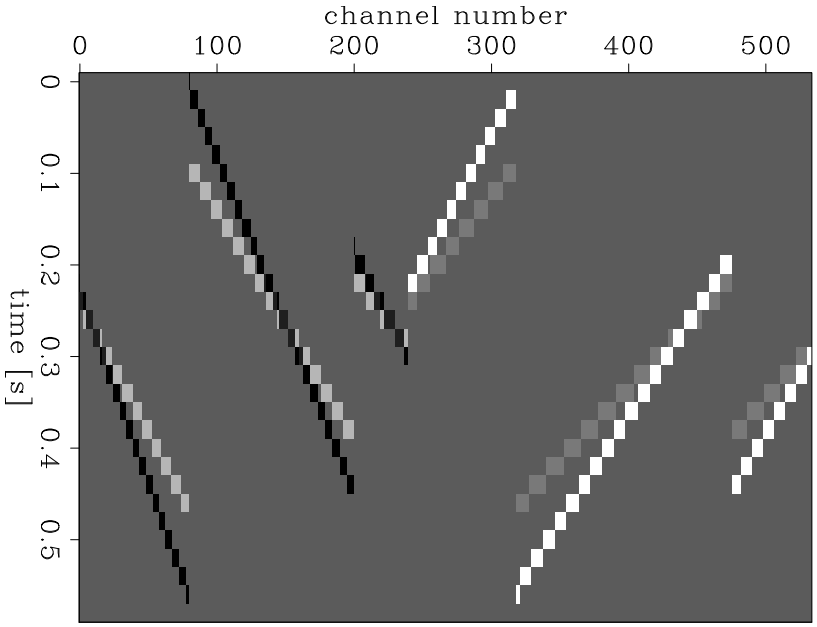


Figure 4: Data simulated using the model shown in Figure 3a and Figure 3b.The compressional wave data is shown as the six light grey oblique lines, while the shear wave is shown as three black lines from channel number 0 to 230 and three white lines after channel 230. Whiter signal means the amplitude is more positive. The channel numbers correspond to the numbers labeled in Figure 2. [**ER**]

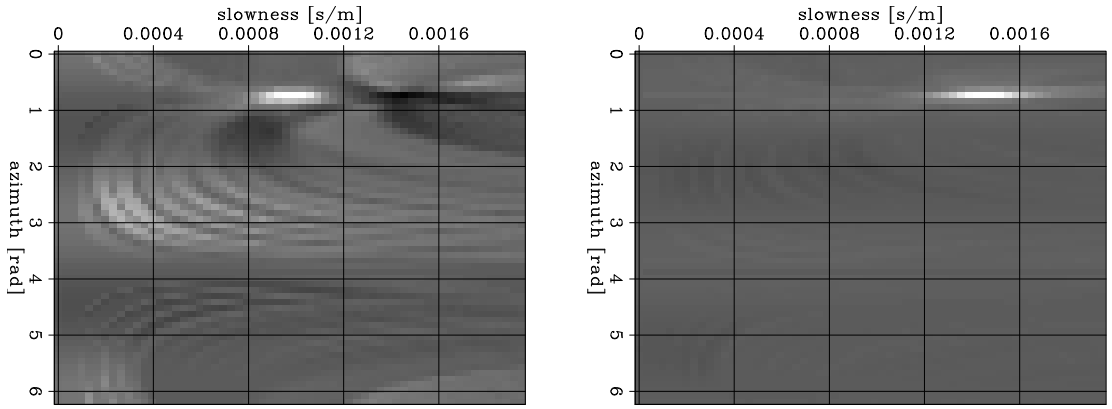
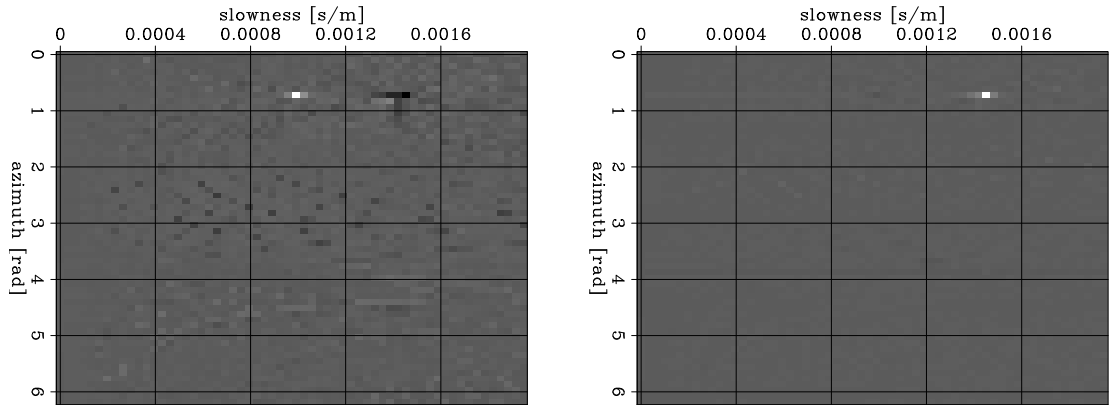


Figure 5: (Left) Compressional-wave model space recovered by beamforming. We see the crosstalk artifact in color black around the same position of the bright spot shown on the right panel. (Right) Shear-wave model space recovered by beamforming. Shear-wave is recovered well with no obvious crosstalk and a higher signal to noise ratio. [**ER**]

Inverse with L2 norm regularized results

Applying the inverse with a L2 norm regularization term to the data shown in Figure 4, we recover the compressional wave model and the shear wave model shown in Figure 6a and Figure 6b respectively. We see that inverse does give us a much better focused results compared to beamforming. But we still have the crosstalk contamination shown in Figure 6a. The signal to noise ratio of the shear wave is higher than that of the compressional wave.



(a) (b)

Figure 6: (Left) Compressional-wave model space recovered by the inverse with a L2 norm regularization. We see the crosstalk artifact in color black around the same position of the bright spot shown on the right panel. (Right) Shear-wave model space recovered by the inverse with a L2 norm regularization. Shear-wave is recovered well with no obvious crosstalk and a higher signal to noise ratio. [**ER**]

Inverse with L1 norm regularized results

Applying the inverse with a L1 norm regularization term to the data shown in Figure 4, we recover the compressional wave model and the shear wave model shown in Figure 7a and Figure 7b respectively. We see that with a L1 norm term, inverse highly recovers both the compressional wave and the shear wave without clear crosstalk problem.

BEAMFORMING/INVERSE ON REVERBERANT WAVE FIELDS

We have seen in DAS recordings that we always have reverberant waves. Here we simulate sinusoid waves with a frequency of 4 Hz. Suppose that we have waves coming with the same azimuth and the same apparent velocity as in the impulsive case (Figure 3a and Figure 3b). But instead of having only one impulse at a fixed

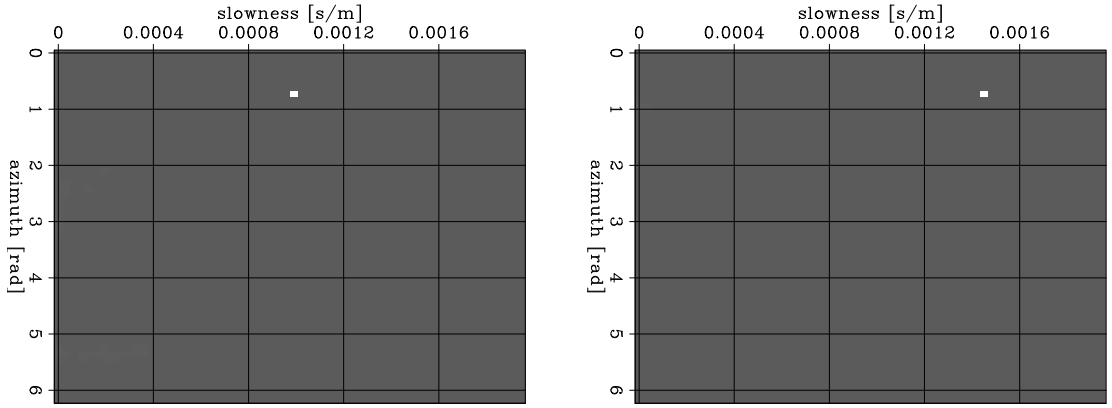


Figure 7: (Left) Compressional-wave model space recovered by the inverse with a L1 norm regularization. (Right) Shear-wave model space recovered by the inverse with a L1 norm regularization. Both the compressional wave and the shear wave are recovered well. [**ER**]

interception time in the impulsive wave case, we record sinusoid amplitudes with a frequency to be 4 Hz. We can simulate a reverberant compressional wave and shear wave interfering with each other shown in Figure 8. The data are more similar to the shear wave simulation since the shear wave has a higher amplitude with respect to the compressional wave.

Beamforming results

Figure 9a and Figure 9b show the beamforming results for compressional and shear waves respectively. We see that the shear wave is recovered fairly well, while the resolution is worse than the impulsive wave case. The poor resolution makes sense considering that beamforming is known to be bad at dealing with data with compact spectrum. From Figure 9a, we see that the compressional wave is not recovered at all. This could be because in the reverberant case, the compressional wave is much more contaminated by the shear wave.

Inverse with L2 norm regularized results

Applying the inverse with a L2 norm regularization term to the data shown in Figure 8, we recover the shear wave model shown in Figure 10b. Better than beamforming, the inverse does recover the compressional wave, although there are high amplitude noise.

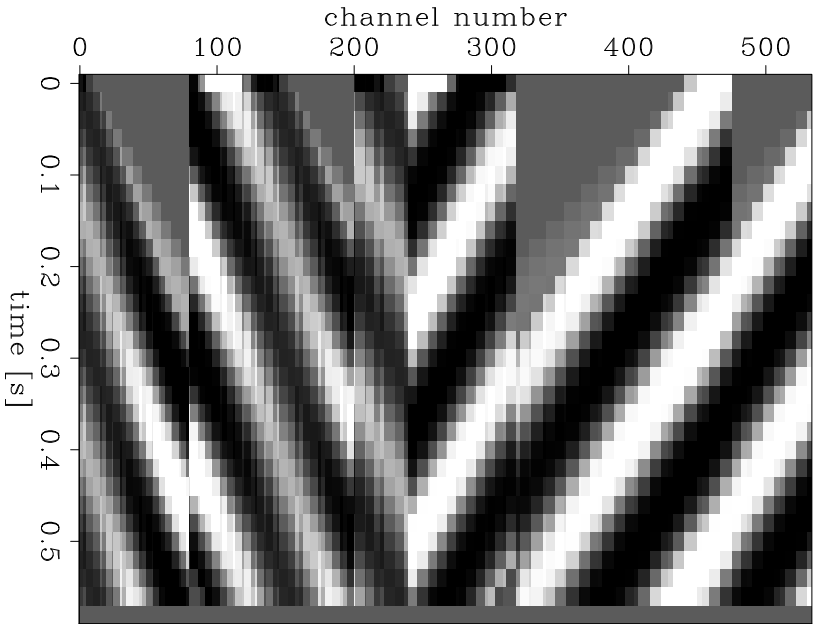


Figure 8: Data simulated using the model shown in Figure 3a and Figure 3b. The amplitudes of the impulse are sinusoid functions of the interception time with a frequency to be 4 Hz. The data is more toward to the shear wave simulation since the shear wave has a higher amplitude with respect to the compressional wave. The channel numbers correspond to the numbers labeled in Figure 2. [**ER**]

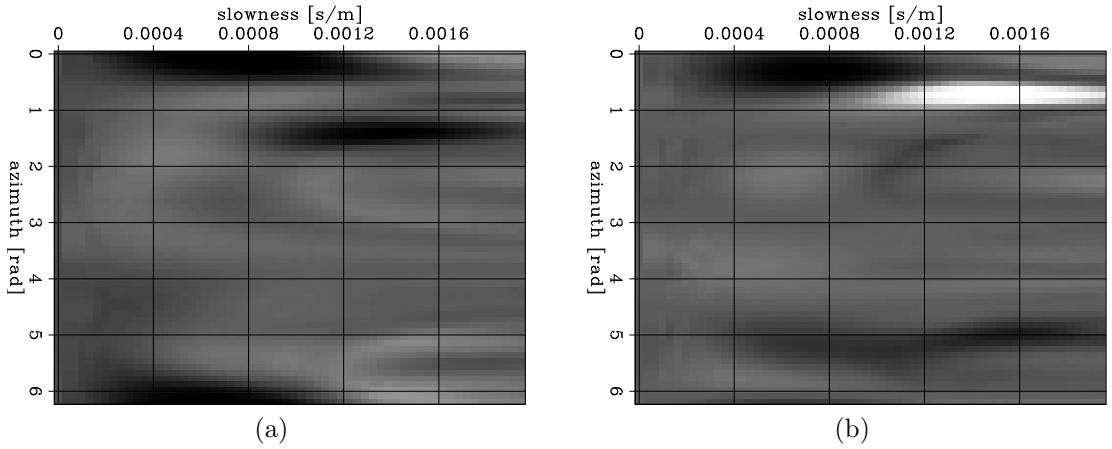


Figure 9: Reverberant wave beamforming results(Left) Recovered compressional-wave model space. The compressional wave is not recovered. (Right) Recovered shear-wave model space. The shear wave is recovered with a poor resolution compared to the impulsive wave case. [**ER**]

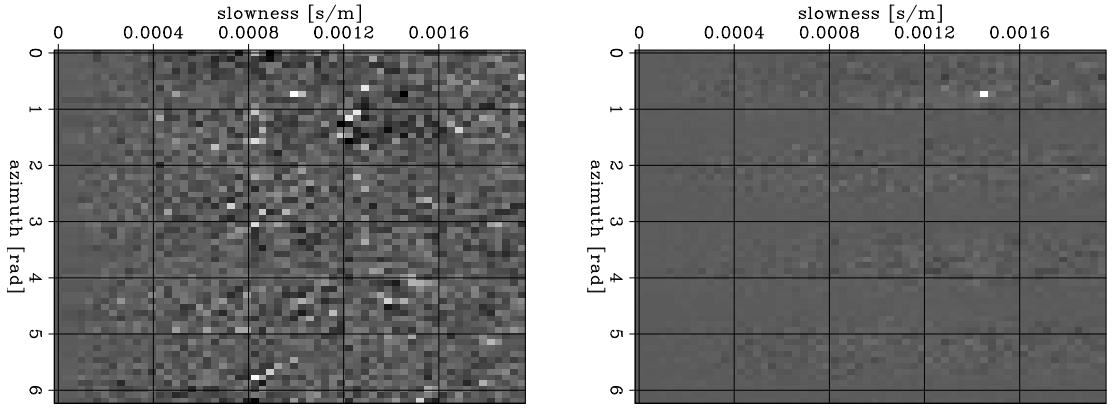


Figure 10: Reverberant wave L2 norm inverse results. (Left) Recovered compressional-wave model space. The compressional wave is recovered with a high noise level. (Right) Recovered shear-wave model space. The shear wave is recovered well. [**ER**]

Inverse with L1 norm regularized results

Applying the inverse with a L1 norm regularization term to the data shown in Figure 8, we recover the compressional wave model and shear wave model shown in Figure 11a and Figure 11b respectively. We see that with a L1 norm term, the inverse highly recovers both of the compressional wave and the shear wave without obvious crosstalk.

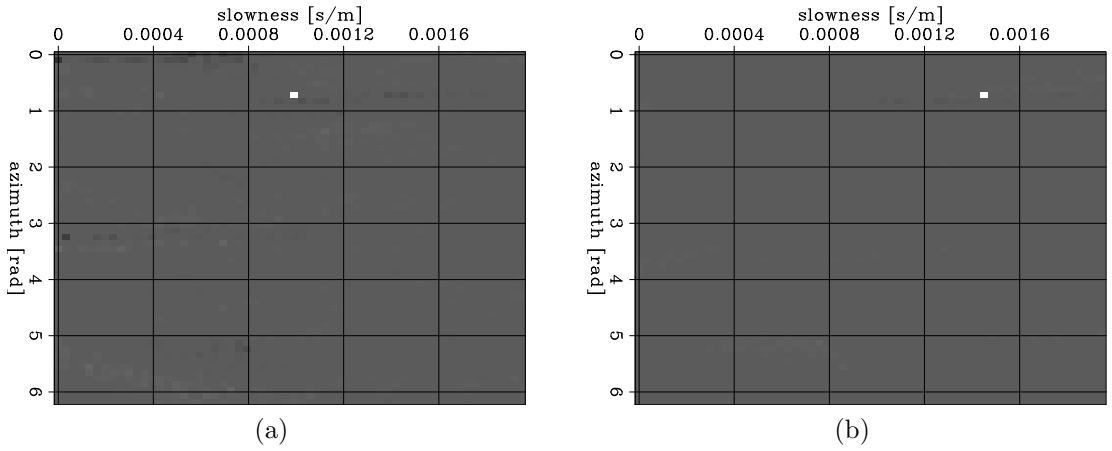


Figure 11: Reverberant wave L1 norm inverse results. (Left) Recovered compressional-wave model space. (Right) Recovered shear-wave model space. Both of the compressional wave and the shear wave are recovered without any crosstalk. [**ER**]

CONCLUSION AND FUTURE WORK

We found that no matter whether we have impulsive waves or reverberant waves, the inverse algorithms regularized by the L1 norm do the best job giving us highly focused results. In the impulsive case, both beamforming and the inverse algorithms regulerized by the L2 norm focused fairly well. But they are subject to the crosstalk problem. In the reverberant case, beamforming and L2 norm inverse fail to recover the compressional wave. This may be due to the fact that in the reverberant case, the compressional wave is overwhelmed by the shear wave. We can learn that beamforming and L2 norm regularization are insensitive to compressional waves when the compressional wave has less energy compared to that of shear waves.

The future work would be testing these algorithms on real earthquake data to see their ability to tell us where the earthquakes come, at what apparant velocities, and with what compositions of the compressional and shear waves. We can then use these information as features of template matching or machine learning algorithms to detect new earthquakes.

ACKNOWLEDGEMENT

We would like to thank OptaSense Ltd for making the DAS recording equipment available. We thank our colleagues at Stanford for helpful discussions and insights, particular Robert Clapp, Eileen Martin and Yinbin Ma.

REFERENCES

Biondi, B., E. Martin, S. Cole, and M. Karrenbach, 2017, Earthquakes analysis using data recorded by the Stanford DAS Array: SEP-Report, **168**, 11–26.

Cole, S. P., 1995, Passive seismic and drill-bit experiments using 2-d arrays: SEP-Report, **86**.

Karrenbach, M. and S. P. Cole, 1989, Beam steering using 3-component data: SEP-Report, **61**.

Kostov, C., 1989, Finite-aperture slant-stack transforms: SEP-Report, **61**.

Martin, E. R., B. Biondi, G. Fabien-Ouellet, and R. G. Clapp, 2017, Sensitivity analysis of distributed acoustic sensing arrays : SEP-Report, **170**, 57–80.