Microquake seismic interferometry with SVD-enhanced Green's function recovery

GABRIELA MELO and ALISON MALCOLM, Massachusetts Institute of Technology

The conditions under which seismic interferometry (SI) leads to the exact Green's function (GF) are rarely met in practice. As a result, we generally recover only estimates of the true GF. This raises the questions: How good an approximation to the GF can SI give? Can we improve this estimated GF?

To recover the full GF between two receivers using SI requires that these two receivers be surrounded by a surface of sources, with both monopole and dipole sources required for accurate amplitude estimates. Because dipole sources are rarely available in practice, here we focus primarily on recovering traveltimes. Accurate estimation of these traveltimes still requires full (monopole) source coverage, however, and this is an assumption that is rarely met in practice. This results in a degradation of the quality of the recovered GF, which then needs to be carefully interpreted. In the ideal case, Snieder (2004) showed that the sources that give the main contribution to the causal and anticausal GFs are the ones located along the raypath between the two receivers, and those in the Fresnel zone around these sources. Snieder came to this conclusion using the method of stationary phase: The sources along the raypath are the sources at which the phase is stationary. Energy emanated by sources outside the Fresnel zone should cancel out, again assuming full source coverage, as they are outside the zone in which the phase is stationary.

When the coverage is not ideal, or when the source/receiver locations and raypaths are not well known, this nonstationary energy will not cancel, resulting in errors in the recovered GF. To alleviate this problem, we employ additional information than is typically used. This information comes from the collection of cross-correlated traces, one for each source for a pair of receivers, which we shall refer to as the cross-correlogram. It is by stacking the cross-correlogram in the source dimension that we obtain an interferometric GF. In general, this cross-correlogram has both stationary energy, that should contribute to the estimated GF and nonstationary energy that should not. Stationary energy in the cross-correlogram is characterized by linearity, coherency, low wavenumber, and thus nearly in-phase events along the source dimension. Nonstationary energy by contrast is characterized by nonlinearity, incoherency, high wavenumber, and out-of-phase events along the source dimension. We exploit these differences to separate the two parts of the energy in the cross-correlogram to obtain more accurate GF estimations for cases that are not ideal.

To perform this separation and extract more information from the cross-correlograms to ultimately improve the GF, we follow Melo et al. (2010) in which the singular value decomposition (SVD) (e.g., Golub and van Loan, 1996) is used to do this separation. SVD is a numerical technique commonly used in seismic data processing (e.g., Ulrych et al., 1988; Sacchi et al., 1998) to increase the signal-to-noise ratio and filter linear events. Melo et al. showed how SVD is able to enhance physical arrivals that are not properly recovered using standard stacking in SI and generally recover arrivals that would otherwise be obscured by noise. Here we further investigate the relationship between SVD and SI in the microseismic context discussed below.

To understand why SVD is able to separate stationary and nonstationary energy we must first understand the relationship between frequency and singular values. This relationship is discussed by Hansen et al. (2006) where they explain the relationship between singular values and frequency (or source-wavenumber in our case)—large singular values correspond to low frequencies and small singular values correspond to high frequencies. As they correspond to low-frequencies, large singular values are associated with events that are in phase in the cross-correlogram: stationary sources whose energy contribute to the GF. In the context of waveguides, Philippe et al. (2008) exploit the connection between singular values and

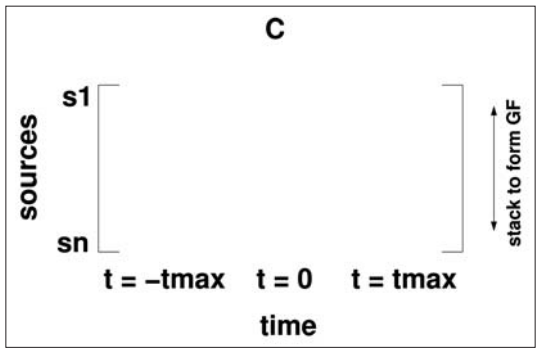

Figure 1. *Cross-correlogram matrix* ***C****. Stacking over sources gives interferometric GF.*

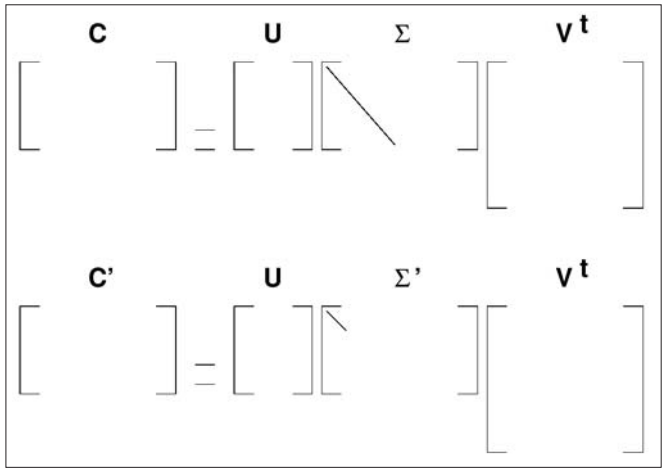

Figure 2. *Cross-correlogram matrix* ***C*** *and its lower-rank approximation* ***C'*** *obtained through SVD.*

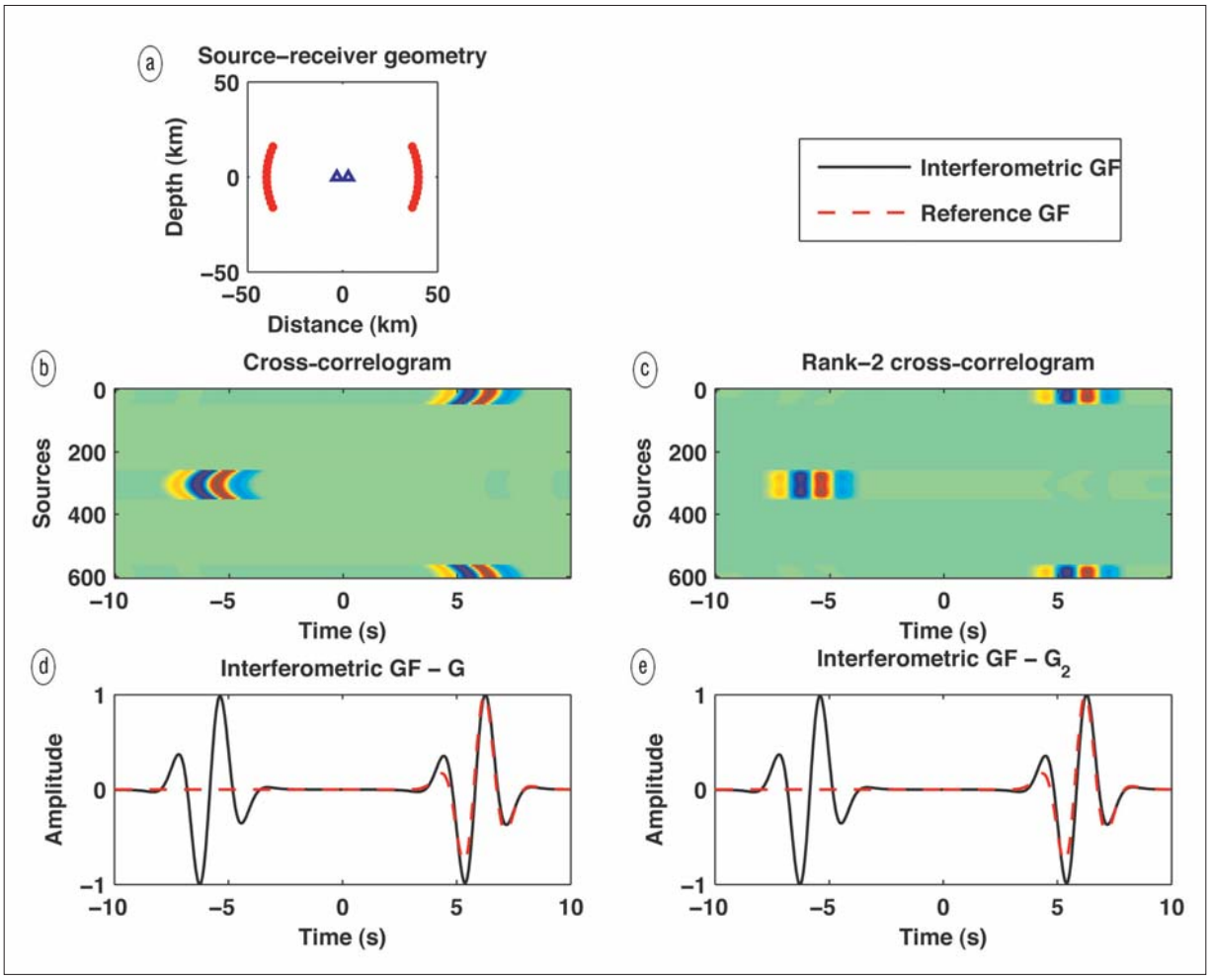

Figure 3. *(a) Source-receiver geometry with 13 evenly distributed sources (red stars)in each of the stationary zones of the receivers (blue triangles); (b) original cross-correlogram; (c) rank 2 cross-correlogram; (d) standard interferometric GF,* $\mathbf{G}$*; (e) rank 2 GF,* $\mathbf{G}_2$*. The black line corresponds to the interferometric GFs and the red line to the true GF. The GFs in (d) and (e) are similar.*

frequency for characterization of targets. They show that the first singular value associated with a given target is proportional to the backscattering form function of the target, and that the second singular value is proportional to the second derivative of the angular form function. Then, they use SVD to extract the backscattered frequency signature of a target in a waveguide. Here, we filter the cross-correlograms by using a lower-rank approximation, computed with SVD using the largest singular values, to enhance events that are coherent across multiple sources, thus isolating this stationary energy. In this way, we exploit the fact that stationary signal is at lower wavenumbers than non-stationary signal to separate it from nonstationary signal. We illustrate this method with synthetic results for both homogeneous and scattering media simulating a possible application with downhole receivers.

These examples are meant to illustrate the particular application we have in mind for this technique, which is the estimation of the GF between two sources in a geothermal reservoir. While most applications of SI estimate the GF between two receivers surrounded by sources, Curtis et al. (2009) show, using reciprocity, that it is also possible to use SI to estimate the GF between a pair of sources. In the microseismic context this would greatly increase the available data set as there are generally few receivers and many sources. In addition, as is well known, noise is a major issue with microseismic data. Noise-contaminated data lead to poor event locations, which creates uncertainty as to which receivers are in the Fresnel zone for a given pair of sources. Simply summing the responses from all receivers will not solve this problem because the receiver array is generally sparse. Here we show how using SVD to decompose the cross-correlogram before stacking helps to alleviate this problem. In addition to these properties, we find that SVD also allows some level of separation of the GF into different components—main arrivals (direct, singly reflected, and refracted waves), multiple scattering, and noise. To separate signal from noise directly in the GF is difficult, especially for coda waves, because the noise may have comparable amplitude and temporal frequency content to the coda. Doing this separation is important because coda waves contain information about the inhomogeneities in the medium, while noise does not. We show preliminary results illustrating that it may be possible to extract information about these different components in the cross-correlated traces before stacking them to form the GF.

Method

We now consider the cross-correlogram as a matrix, **C**, where each row is the cross-correlation of the signals at the two receivers from one source. Thus, the vertical dimension of **C** is source and the horizontal is time, as shown in Figure 1.

Next, we decompose the cross-correlogram using SVD (see Golub and van Loan for a description of SVD). The SVD decomposition of the cross-correlogram matrix is, $\mathbf{C} = \mathbf{U\Sigma V}^t$, where **U** and **V** are the left and right singular vector matrices, and **Σ** is the diagonal matrix whose elements are the singular values of **C**. Figure 2 shows how we obtain a lower-rank approximation **C'** of the cross-correlogram by selecting only the largest singular values of the SVD decomposition of **C**. Stacking the rows of **C** gives the standard interferometric GF, **G**, and stacking the rows of the approximation **C'** gives the modified interferometric GF, $\mathbf{G}_j$, where j is the rank of **C'** (the number of singular values retained). In the examples that follow, we compare these two GFs.

We now illustrate this procedure with a simple example. The model for this example is a constant velocity and density model with no reflectors, so the GF consists of the direct wave only. We examine how well we can approximate the true GF in three cases: (1) the case where there are stationary sources only; (2) nonstationary sources only; and (3) both stationary and nonstationary. In all three cases there are gaps in the source coverage and all the GF are normalized as we do not have dipole sources.

First, consider a case where there are sources only in the stationary-phase zone, Figure 3a. The energy from these sources contributes constructively to the GF. Comparing the standard (Figure 3b) and rank 2 (Figure 3c) cross-correlograms and the respective estimated GFs, **G** and $\mathbf{G}_2$, (Figure 3d and 3e), we see that the standard and the lower-rank cross-correlograms and GFs are quite similar. We use a rank 2 approximation of the cross-correlogram because there are two stationary-phase zones in the cross-correlogram and thus two signals we wish to reconstruct. In this simple example, it is obvious what the rank of the cross-correlogram approximation should be, which is not the case in general. This is a case where standard interferometry works well and the SVD technique is not necessary, although it is also not detrimental.

In case 2, there are only nonstationary sources (Figure 4a). Ideally, all of this nonstationary energy should cancel but if there are gaps in the source coverage residual energy will

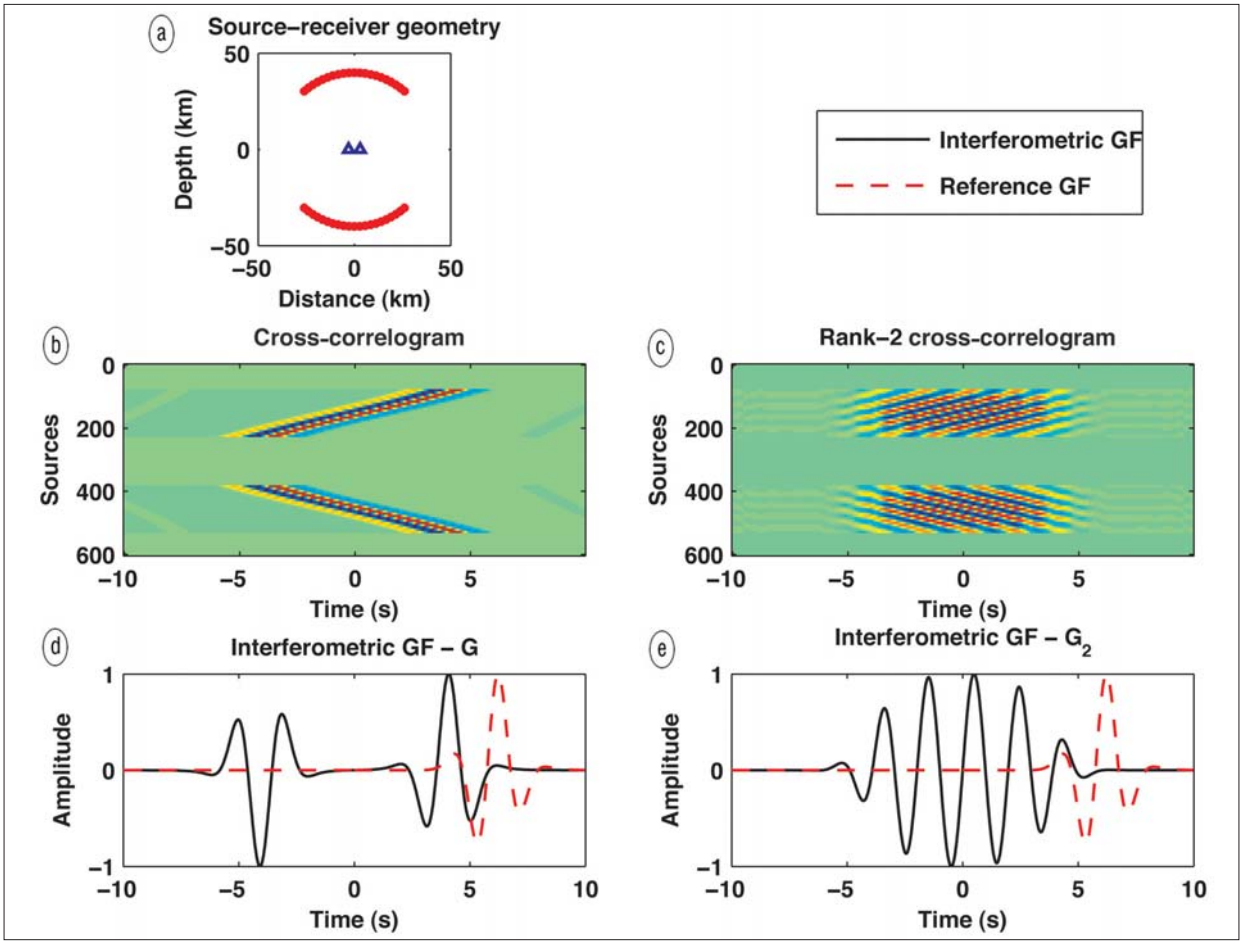

Figure 4. *(a) Source-receiver geometry with 21 evenly distributed sources (red stars) in each of the nonstationary zones of the receivers (blue triangles); (b) original cross-correlogram; (c) rank 2 cross-correlogram; (d) standard interferometric GF,* ***G****; (e) rank 2 GF,* $\mathbf{G}_2$*. The black line corresponds to the interferometric GFs and the red line to the true GF. In (e) the GF does not contain the edge effect present in (d).*

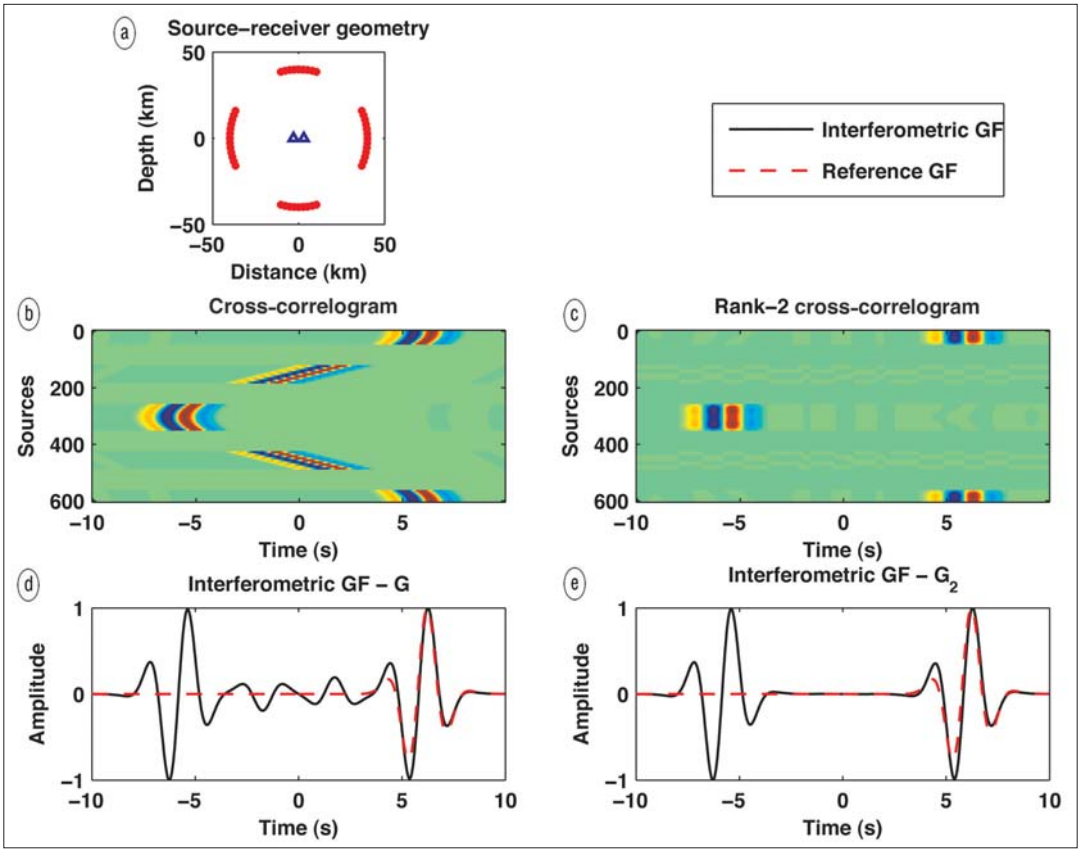

Figure 5. *(a) Source-receiver geometry with 13 and 9 evenly distributed sources (red stars) in each of the stationary and nonstationary zones of the receivers (blue triangles), respectively; (b) original cross-correlogram; (c) rank 2 cross-correlogram; (d) standard interferometric GF, **G**; (e) rank 2 GF, $\mathbf{G}_2$. The black line corresponds to the interferometric GFs and the red line to the true GF. In (e) the fluctuations are reduced and the GF is clearer than in (d).*

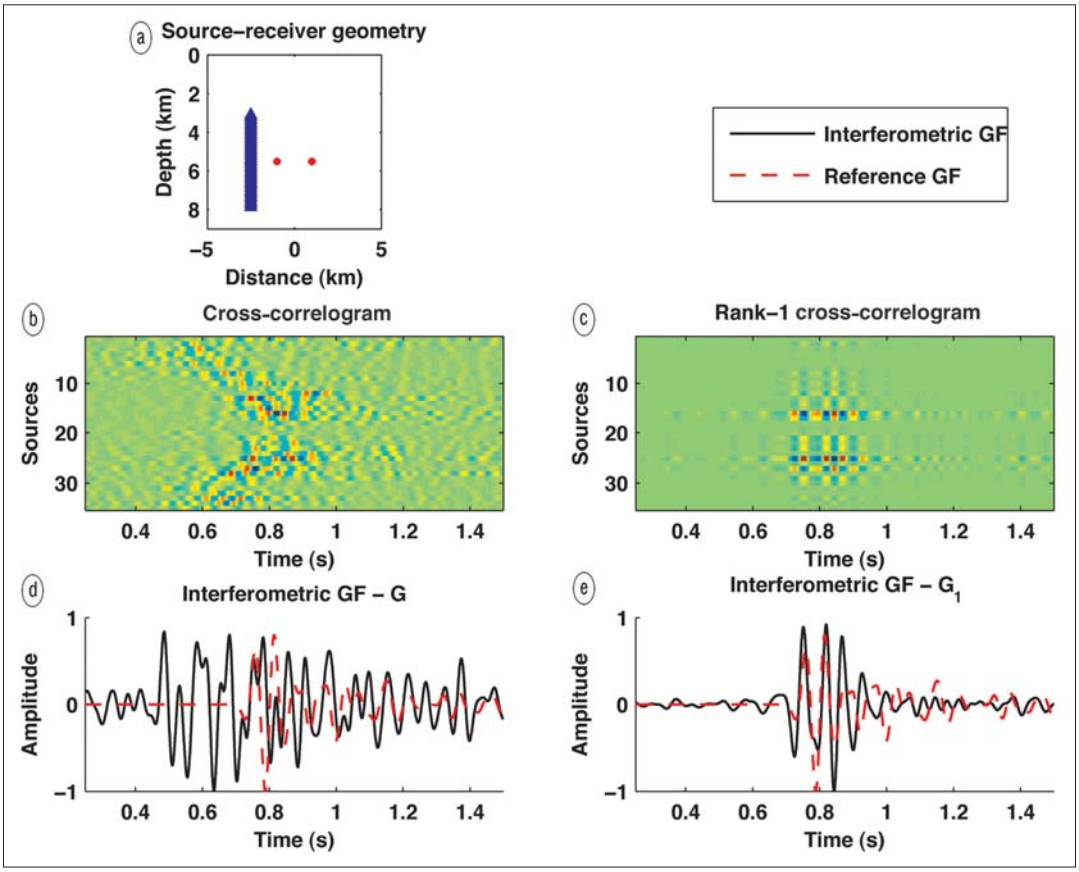

Figure 6. *(a) Source-receiver geometry = one borehole with 35 receivers (blue triangles) and two microquakes (red stars); (b) original cross-correlogram; (c) rank 1 cross-correlogram; (d) standard interferometric GF, **G**; (e) rank 1 GF, $\mathbf{G}_1$. The black line corresponds to the interferometric GFs and the red line to the true GF. In (e) the fluctuations are reduced and the GF is clearer than in (d).*

remain because of edge effects. As is clear in Figure 4d, **G** is not a good estimate of the correct GF, but it appears as if it contains a physical arrival. While **G** (Figure 4d) contains two nonphysical arrivals due to edge effects, $\mathbf{G}_2$ (Figure 4e) does not. The rank 2 cross-correlogram (Figure 4c) in this case does not enhance any linearity and does not even resemble the original cross-correlogram (Figure 4b). The rank 2 cross-correlogram and $\mathbf{G}_2$ thus act as a diagnostic of nonphysical arrivals.

Case 3 mixes the two previous cases. Figure 5a shows sources uniformly distributed in each stationary zone and each nonstationary zone, but with gaps in between. The cross-correlogram (Figures 5b and 5c) thus has energy contributing to the GF and energy that should cancel out completely but, because of the gaps, does not. The rank 2 approximation filters the pseudo-noise caused by the imperfect cancellation of nonstationary energy, and $\mathbf{G}_2$ is more accurate then **G** as seen in Figures 5d and 5e.

Examples

This example approximately mimics an idealized source/receiver geometry of a downhole monitoring of microseismic activity in a geothermal reservoir. We use a single borehole with 35 receivers and estimate the GF between two microquakes as shown in both Figure 6a and Figure 7a. The reference and interferometric GFs shown here are all normalized. The idea is to obtain the interferometric GF between a pair of sources (microquakes) instead of a pair of receivers, similar to what is done in Curtis et al. (2009) for a larger-scale problem. The medium is weakly scattering with a constant background velocity and density. We study two cases: first, we do the cross-correlations in a clean data set (Figure 6) and second in a noisy data set (Figure 7). The additive noise and random scattering we use here are realizations of a Gaussian random field with prescribed correlation lengths along given directions. We focus our observations on two things: the phase of the direct wave and the energy in the coda. We added enough noise to completely obscure the direct wave and distort the waveform of the coda wave in the interferometric GF and show how SVD improves both of these measures.

In the first example (Figure 6), even though the data are noise free, there is enough nonstationary energy in the cross-correlogram from receivers outside the Fresnel zone (Figure 6b) to create high-amplitude fluctuations that hide the direct

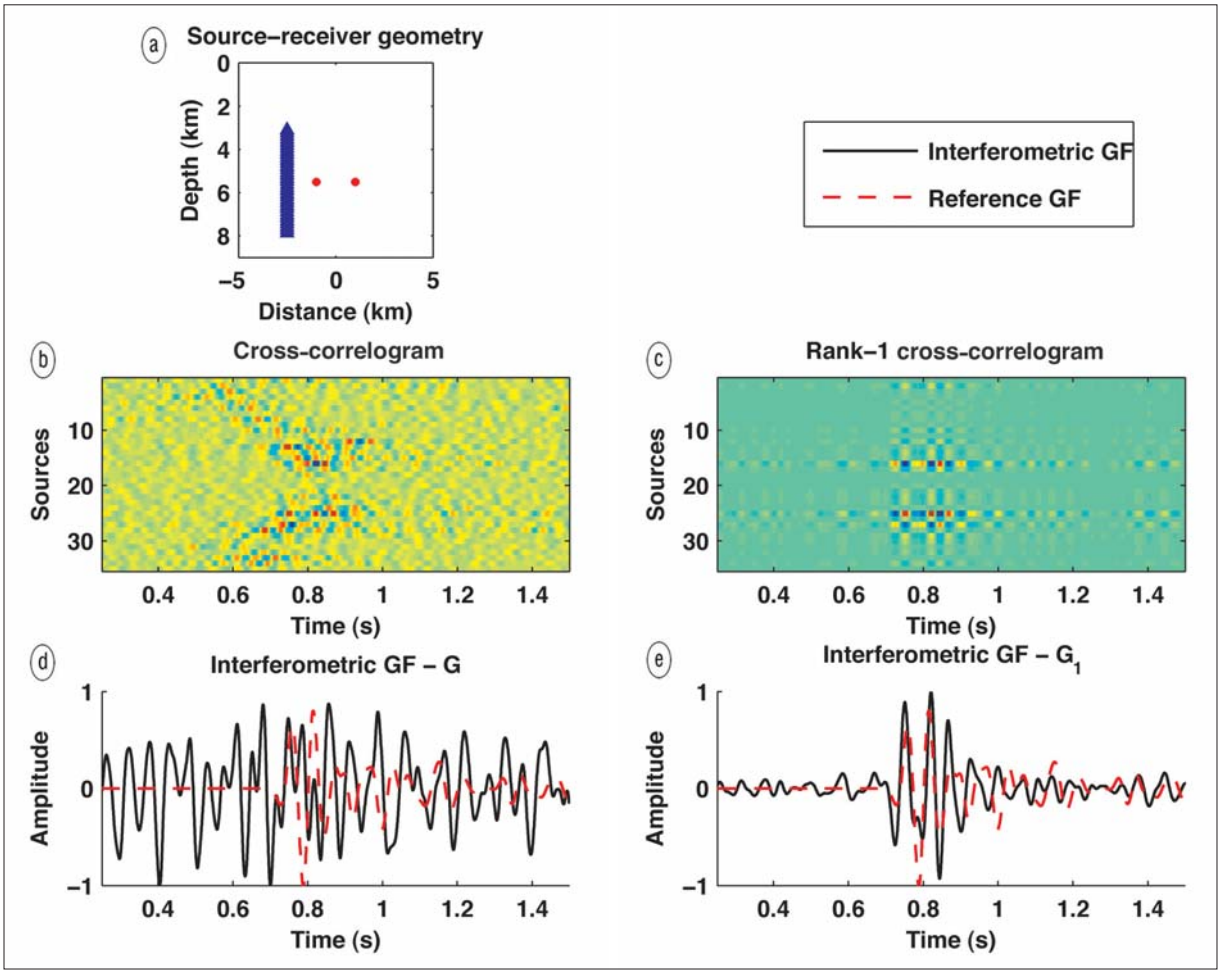

Figure 7. *(a) Source-receiver geometry = one borehole with 35 receivers (blue triangles) and two microquakes (red stars); (b) original cross-correlogram; (c) rank 1 cross-correlogram; (d) standard interferometric GF,* ***G****; (e) rank 1 GF,* $\mathbf{G}_1$*. The black line corresponds to the interferometric GFs and the red line to the true GF. In (e) the random noise is reduced to the level of the coda and the GF is clearer than in (d).*

wave. Here we chose the rank 1 approximation (Figure 6c) because the GF consists of only one direct wave and thus there is only one zone of stationary-phase energy in the cross-correlogram. In Figure 6e, the fluctuations were damped in $\mathbf{G}_1$ as compared with $\mathbf{G}$ in Figure 6d, leading to a much clearer GF. Cross-correlation of both GFs, $\mathbf{G}$ and $\mathbf{G}_1$, with the reference GF peaks at 0.2835 and 0.0020 s (time-sampling interval is 0.0005 s), respectively, shows a significantly more accurate phase estimate with $\mathbf{G}_1$ than with $\mathbf{G}$. We use the L2-norm (square-root of sum of squares) of the coda waves as a measure of the energy in the coda. By comparing the reference and the interferometric coda waves, we find relative errors of 98% for the coda wave in $\mathbf{G}$ and only 12% for $\mathbf{G}_1$, demonstrating that $\mathbf{G}_1$ is a better approximation of the true GF than $\mathbf{G}$, in this norm.

To make this example more realistic, we add weakly correlated noise to the data in Figure 6. In Figure 7d, $\mathbf{G}$ appears strongly contaminated by noise and neither the direct arrival nor the coda wave are visible. In Figure 7e, fluctuations and random noise are strongly attenuated in $\mathbf{G}_1$, revealing not only the direct arrival but also reducing the noise close to the coda-wave level. The phase differences between $\mathbf{G}$ and $\mathbf{G}_1$ and the reference GF are 0.2850 and 0.0015 s, respectively. The relative errors in the L2-norm of the coda are 148% for $\mathbf{G}$ and 11% for $\mathbf{G}_1$. We see that SVD eliminates most of the noise in the coda wave, as well as the fluctuations before the direct wave, demonstrating its stability with respect to noise.

For the noisy case, we performed tests for a variety of receiver apertures, noise levels, and spacing between receivers. We find that the absolute improvement obtained through SVD varies from case to case but the phase of direct wave and the coda energy, in general, are closer to correct with SVD than without. This noise attenuation is particularly important in microseismic studies as the data are typically quite noisy. Stability with respect to aperture is also important because errors in the location of microquakes can be significant.

Discussions, conclusions, and future work

The accurate estimation of the GF with source coverage that is not ideal remains a significant problem in SI. We have shown how using SVD to approximate cross-correlograms before stacking is a promising approach to alleviate this problem. In general, for the SVD technique to work, there must be more stationary energy than nonstationary energy in the cross-correlogram, although this requirement can be relaxed somewhat through normalization of the traces in the cross-correlogram. How much more energy is necessary and how much noise can be accommodated are subjects of ongoing research.

We are also continuing to investigate which properties of coda waves can be accurately inferred from the GF obtained through SVD; the preliminary results shown here indicate that such properties are better recovered with SVD than stacking alone. Separating real signal from noise would lead to a coda that truly reflects the scattering characteristics of the medium thus allowing for the use of coda waves to retrieve information about the scattering strength and through this about fracture characteristics in a reservoir. TLE

References

Curtis, A., H. Nicolson, D. Halliday, J. Trampert, and B. Baptie, 2009, Virtual seismometers in the subsurface of the earth from seismic interferometry: Nature Geoscience, **2**, no. 10, 700–704, doi:10.1038/ngeo615.

Golub, G. and C. van Loan, 1996, Matrix computations: Johns Hopkins University Press.

Hansen, P. C., M. E. Kilmer, and R. Kjeldsen, 2006, Exploiting residual information in the parameter choice for discrete ill-posed problems: BIT, **46**, no. 1, 41–59, doi:10.1007/s10543-006-0042-7.

Melo, G., A. Malcolm, D. Mikessel, and K. van Wijk, 2010, Using svd for improved interferometric Green's functions: 80th Annual International Meeting, SEG, Expanded Abstracts, 3986–3990, doi:10.1190/1.351368.

Philippe, F. D., C. Prada, J. de Rosny, D. Clorennec, J.-G. Minonzio, and M. Fink, 2008, Characterization of an elastic target in a shallow water waveguide by decomposition of the time-reversal operator: The Journal of the Acoustical Society of America, **124**, 779–787, doi:10.1121/1.2939131.

Sacchi, M. D., T. J. Ulrych, and C. Magnuson, 1998, Eigenimage analysis of common offset sections: Signal-to-noise enhancement and pre-stack data compression.

Snieder, R., 2004, Extracting the Green's function from correlation of coda waves: A derivation based on stationary phase: Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, **69**, no. 4, 046610, doi:10.1103/PhysRevE.69.046610.

Ulrych, T. J., S. Freire, and P. Siston, 1988, Eigenimage processing of seismic sections: 58th Annual International Meeting, SEG, Expanded Abstracts, 1261–1265, doi:10.1190/1.1892508.

Acknowledgments: We thank Oleg Poliannikov, Pierre Gouedard, Bongani Mashele, Huajian Yao, and Michael Fehler for great input during our discussions. This work is supported by grants from the Department of Energy (DE-FG36-08GO18197) and Chevron, and the founding members consortium at Earth Resources Laboratory (ERL).

Corresponding author: gmelo@mit.edu