Efficient Seismic Forward Modeling using Simultaneous Random Sources and Sparsity

R. Neelamani, C. E. Krohn, J. R. Krebs, M. Deffenbaugh, J. E. Anderson, ExxonMobil; and J. K. Romberg, Georgia Institute of Technology*

SUMMARY

This paper proposes an approach to speed up seismic forward modeling when the Green's function of a given model is *structured* in the sense that the Green's function has a sparse representation in some known transform domain. The first step of our approach runs a forward finite-difference (FD) simulation for a duration longer than each conventional run, but with all sources activated simultaneously using different long random noise waveforms. The cumulative responses to the simultaneous sources are measured at all receivers. The second step separates the interfering Green's functions from the receiver measurements by exploiting prior knowledge of the random waveforms and the sparsity of the Green's function in a suitable domain. Simulation results demonstrate such a simultaneous source approach is indeed promising.

INTRODUCTION

One goal of seismic forward modeling is to estimate the 5-dimensional (2D receiver location * 2D source location * 1D time) Green's function for a given acoustic or elastic model. Conventionally, the Green's function for a given model is computed by running a finite-difference (FD) setup several times. During each sequential run, each source is activated with a band-limited source wavelet (e.g., a Ricker wavelet). The response to the source wavelet is computed at all receiver locations simultaneously by running FD modeling for some time, say T_{seq}. For N_s sources, the total computational effort for the *sequential* method is governed by $N_s * T_{seq}$. The complexity of FD modeling makes the sequential approach computationally expensive. In fact, it is currently infeasible to model the elastic response of a realistic 3D model. Hence more efficient modeling alternatives are desirable.

The efficiency of the conventional approach is almost independent of the complexity of the target Green's function. It is desirable to have a modeling technique whose computational efficiency scales with the complexity of the target Green's function (i.e., less computations for structured Green's function). This paper proposes a framework with this desirable property.

The first step in our proposed framework comprises running FD modeling just once, but for a longer duration $T_{sim} > T_{seq}$ than each sequential modeling run. During the entire run, *all* sources are *simultaneously* activated with different known random noise waveforms. The receivers measure the *cumulative* response to the simultaneous sources.

The second step in the proposed framework comprises separating the interfering Green's functions from the receiver measurements. This key separation step can be formulated as an ill-posed linear inverse problem. We propose to estimate the Green's function for each source-receiver pair by exploiting the *sparsity* (i.e., structure) of the Green's function in a suitable transform domain (e.g., curvelet domain (Candès et al., 2006a)) in tandem with our knowledge of the random source waveforms. Our approach draws inspiration from vibroseis acquisition in geophysics and from the recently developed *compressive sensing (CS)* field (Candès et al., 2006b; Donoho, 2006; Baraniuk, 2007).

Ikelle (2007) also advocates the use of simultaneous sources in acquisition and modeling, but employs unknown *impulsive* sources and exploits principal/independent component analysis (PCA/ICA) to estimate the unknown sources.

The total computational effort for the simultaneous source framework would be governed by T_{sim}, assuming that the computations necessary to solve the separation problem are negligible compared to FD computations. The T_{sim} necessary for satisfactory separation is expected to be controlled by the Green's function's complexity. For cases where the Green's function has an extremely sparse representation, we expect the simultaneous source method to significantly improve upon the efficiency of the sequential source method (i.e., $T_{sim} << N_s * T_{seq}$). Our encouraging experiments substantiate that the proposed modeling approach indeed holds promise.

MATHEMATICAL FORMULATION

Let $G(s_i, r_j, t_k)$ denote the Green's function for an arbitrary model with s_i and r_i denoting the source and receiver locations, and t_k denoting the discrete time-sample. Let $\Psi(s_i, t_k)$ denote the waveform emitted by a source placed at s_i. Let $R(r_j, t_k)$ denote the measurement made by a receiver placed at r_j.

In the sequential source scenario, with a single source placed at s_i, the receiver measures

$$R(r_j, t_k) = G(s_i, r_j, t_k) \circledast \Psi(s_i, t_k) = \sum_{\ell} G(s_i, r_j, t_\ell)\, \Psi(s_i, t_{k-\ell}),$$

where $\circledast$ denotes convolution, and $t_k \in [0, T_{seq}]$. Since $\Psi(s_i, t_k)$ is typically chosen to be a band-limited impulse function, the receivers directly measure the band-limited Green's function.

In the simultaneous source scenario, since sources at several s_i are activated simultaneously, the receiver measures the linear superposition of all the Green's functions convolved with the respective source waveforms;

$$\begin{aligned} R(r_j, t_k) &= \sum_i G(s_i, r_j, t_k) \circledast \Psi(s_i, t_k) \\ &= \sum_{i,\ell} G(s_i, r_j, t_\ell) \Psi(s_i, t_{k-\ell}), \end{aligned} \qquad (1)$$

with $t_k \in [0, T_{sim}]$. In this case, the Green's functions $G(s_i, r_j, t_k)$ need to be separated from the measurements $R(r_j, t_k)$.

Green's function separation from Equation (1) is clearly an ill-posed linear inverse problem because the number of measurements is less than the number of inversion variables. This problem naturally raises three fundamental questions.
1) Measurement: What source waveforms Ψ's would facilitate better separation?
2) Efficiency: What is the minimum number of samples (T_{sim}) needed to accurately estimate the Green's functions?
3) Algorithm: What algorithm can be used to efficiently separate the Green's function?

COMPRESSIVE SENSING (CS) OVERVIEW

The recently developed *compressive sensing* (CS) field (Candès et al. (2006b); Donoho (2006); Baraniuk (2007); also see www.dsp.rice.edu/cs) contains theoretical insights that we will find useful to rigorously answer the three questions raised in the previous section. This section reviews the relevant CS results. CS provides a general framework to sense and reconstruct signals that enjoy sparse representations. In geophysics, the use of sparsity to solve ill-posed linear problems is common. However, as is well-known, the use of sparsity can yield unstable results. CS' strength and contribution is that it identifies a coherent measurement-to-reconstruction framework that facilitates robust exploitation of sparsity.

Assume that we wish to sense a discrete-time signal g that is M samples long and that we are allowed to measure the projections λ_k of g onto *measurement* vectors ψ_k of our choice.

$$\lambda_k = \langle g, \psi_k \rangle, \tag{2}$$

where $\langle .,. \rangle$ denote inner products.

CS rigorously addresses the three questions (mentioned above) that arise when one tries to reconstruct a signal g from linear measurements λ_k's. Clearly, for arbitrary g, standard linear algebra dictates that at least M measurements would be necessary to recover g exactly. However, if g is known to be *sparse* in an orthonormal transform domain $\mathscr{H}$, then we can use the CS methodology to accurately reconstruct g with measurements that are far fewer than M. A signal g is said to be sparse in $\mathscr{H}$ if most of its energy is captured by a few big coefficients of its transform-domain representation $\mathscr{H}(g)$.

Measurement: For sensing sparse signals, CS recommends choosing random noise functions as the ψ_k in Equation (2). Such vectors provide robust reconstruction of signals that are sparsely represented by *any* orthonormal transform $\mathscr{H}$.

Efficiency: If $\mathscr{H}(g)$ contains exactly K non-zero coefficients (perfectly sparse), then CS proves that it is possible to recover g *exactly* from approximately $K \log M$ measurements with random ψ_k's. In other words, the number of necessary measurements is predominantly controlled not by the size of the signal but by the signal's complexity (sparsity). Note that at least K measurements are necessary to recover any K-sparse signal, *even if the location of the K non-zero $\mathscr{H}(g)$ coefficients are known*. Thus CS guarantees that with random measurement vectors, the number of measurements necessary to reconstruct g is near-optimal (up to $\log M$ factors). The recovery is also stable when ψ_k are corrupted with noise and when g is only approximately sparse.

Algorithm: CS also identifies computationally tractable algorithms (convex programs) that can accurately reconstruct g from the measurements λ_k. A perfectly sparse g can be recovered exactly from approximately $K \log M$ λ_k's by solving the following linear program.

$$\begin{aligned} \widehat{g} &= \arg\min_g \|\mathscr{H}(g)\|_1 \text{ such that} \\ \lambda_k &= \langle \widehat{g}, \psi_k \rangle, \text{for all } k. \end{aligned} \tag{3}$$

$\|\mathscr{H}(g)\|_1$ denotes the ℓ_1 norm of $\mathscr{H}(g)$. In words, Equation (3) dictates that we should seek a signal that not only satisfies the measurements, but also has a sparse representation in the domain $\mathscr{H}$.

EFFICIENT SEISMIC FORWARD MODELING

In geophysics, the vibroseis data acquisition community pioneered the use of simultaneous sources, which are activated with long waveforms, as an efficient alternative to improve the signal-to-noise ratio (SNR) of land data (Allen et al., 1998; Krohn and Johnson, 2006). Seismic forward modeling is closely linked to acquisition because the same equation explains both processes. The vibroseis method provides us with the seed inspiration to adopt a simultaneous source approach to forward modeling. To fine-tune our framework and to justify our answers to the three questions raised earlier, we invoke CS.

Measurement: We propose to activate all sources during modeling with different long random noise functions. The receiver measurements in Equation (1) at each t_k are given by the projection of the Green's function onto a collection of source waveforms; the projection takes place along the time and source axes. From a CS perspective, each measurement vector in our setup is a time-shifted version of the collection of source waveforms. That is, the $\{\Psi(s_i, t_{k-\ell})\}_k$ in Equation (1) is analogous to the ψ_k in Equation (2). The use of random noise source waveforms maximizes the "randomness" of our measurement vectors, as prescribed by CS.

Efficiency: The length of the source waveforms, which in turn controls the forward modeling run time T_{sim}, would be controlled by the sparsity of $\mathscr{H}\left(G(s_i, r_j, t_k)\right)$, with more complex $G(s_i, r_j, t_k)$ needing larger T_{sim}. In the best-case scenario, given a single receiver measurement at r_j with T_{sim} samples, we could hope to recover approximately T_{sim} largest components of a common-receiver gather (i.e., r_j is fixed) from $G(s_i, r_j, t_k)$ in an arbitrary orthonormal transform domain $\mathscr{H}$; we ignore any log terms for brevity and convenience.

However, the measurement vectors $\{\Psi(s_i, t_{k-\ell})\}_k$ in the proposed modeling setup are not independent random noise realizations; the convolutional constraint placed by the modeling setup prevents us from making them so. Therefore, we cannot prove the efficiency of our framework currently by directly applying known CS results. Our ongoing investigation indicates that for some special transform domains, we may be able to provide some guarantees about our framework's efficiency (Bajwa et al., 2007; Romberg, 2007).

Since the run time T_{sim} is expected to be controlled by $\mathscr{H}\left(G(s_i,r_j,t_k)\right)$'s sparsity, the choice of an appropriate $\mathscr{H}$ is crucial to maximize the simultaneous source approach's efficiency. We expect that high dimensional extensions of the curvelet transform (Candès et al., 2006a)* would make suitable choices for $\mathscr{H}$ because they would provide sparse representations for typical 5D Green's functions.

Algorithm: In contrast to field acquisition, the source waveforms in forward modeling are known precisely. If the sparsity-promoting transform $\mathscr{H}$ is also known, then the Green's functions can be estimated from the receiver measurements by solving

$$\widehat{G} = \arg\min_G \|\mathscr{H}(G)\|_1 \quad \text{such that Equation (1) holds for all } k. \tag{4}$$

If the receiver measurements are noisy (e.g., corrupted with computational dispersion noise), we would seek a Green's function with a sparse representation that satisfies the receiver measurements only up to a predetermined error level.

RESULTS

We evaluated the efficacy of simultaneous source modeling by using a 3D model comprising diffractors and dipping reflectors. We placed 64 sources and 128 receivers along orthogonal lines (see Figure 1(a)). The desired Green's function is a 3D volume (1D source * 1D receiver * 1D time); see Figure 2(a). We assume that the Green's function amplitudes are negligible (assumed zero) after 1 second.

To simulate a simultaneous source modeling setup, each shot-gather from the desired Green's function was convolved with a different band-limited random waveform that was $T_{sim} = 8$ seconds long, and then stacked (in accordance with Equation (1)). Figure 1(b) illustrates some sample random waveforms that were employed during the experiment. The Fourier spectrum of each waveform is flat, but the phase of the Fourier components were chosen at random. Figure 1(c) displays the measurements made by the 128 receivers.

Figure 2(b) shows the Green's function estimated by our method from the receiver measurements. The estimate not only matches the receiver measurements up to 0.01 percent in energy terms (see Equation (4)) but also possesses a sparse 3D curvelet representation. The difference plot Figure 2(c) confirms that our method captures most of the coherent features in the actual Green's function reliably (SNR $\approx$ 7.5dB).

For reference, Figure 2(d) shows the Green's function estimated by a cross-correlation approach. Each shot-gather slice from the 3D volume is extracted by cross-correlating the receiver measurements with the known random source waveform,

$$\widehat{G}(s_n,r_j,t_k) = R(r_j,t_k) \circledast \Psi(s_n,-t_k). \tag{5}$$

Such an approach would be feasible if the cross-terms $\Psi(s_i,t_k) \circledast \Psi(s_n,-t_k)$, $i \neq n$, that result upon expanding Equation (5) are negligible. It is well-known that the cross-correlation between two random functions is zero. However, the cross-correlations go to zero only in the statistical sense (i.e., averaged over several different noise realizations). In the simultaneous source setup, the cross-term cannot be neglected, as illustrated by poor result in Figure 2(d).

For the chosen model, a sequential approach would take 64 seconds to measure Figure 2(a) because each shot-gather in the Green's function would be obtained by running the modeling for $T_{seq} = 1$ second. Since $T_{sim} = 8$ seconds, the modeling effort for the simultaneous source approach would be 8 times lesser than the sequential approach (assuming that the separation step takes negligible computations).

DISCUSSION

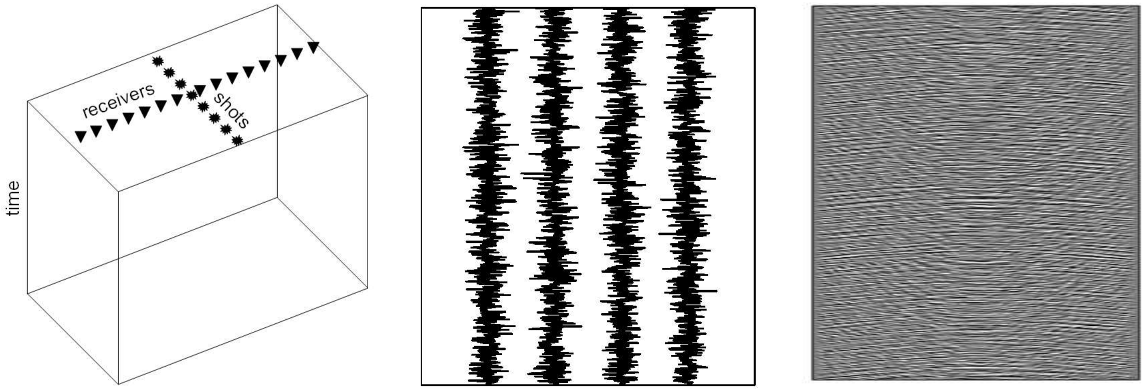
This paper proposes a novel simultaneous source approach that speeds up seismic forward modeling when the target Green's function is structured (i.e., sparse in some known domain). The total modeling time for the method is controlled by the complexity of the Green's function. Thus, more structured Green's functions (e.g., in simple models) can be generated with less modeling effort. Based on recent advances in the CS field, we believe that our approach may enjoy desirable theoretical properties; this an area of ongoing investigation. Our experimental results confirm that our approach is indeed promising.

This paper ignored the computational expense of solving the separation problem associated with the proposed approach. For extremely large problems (e.g.,, when a large number of sources are simultaneously activated with very long random waveforms), the computational cost to solve the separation problem using conventional optimization techniques may not be negligible, thereby adversely affecting the overall efficiency of the proposed approach. The issue of balancing the efficiency obtained by running multiple random sources simultaneously with the computations require to solve the associated separation problem needs further investigation.

This paper also ignores our method's ability to efficiently use all available processors during an FD simulation. An FD simulation for one source location typically employs an expanding computational grid around the source location. For parallel algorithms, this may leave some processors idle until the wave-field reaches their domain-decomposed region. In contrast, the proposed method would have a larger zone of active wave propagation at early simulation times, thereby utilizing the processors more efficiently.

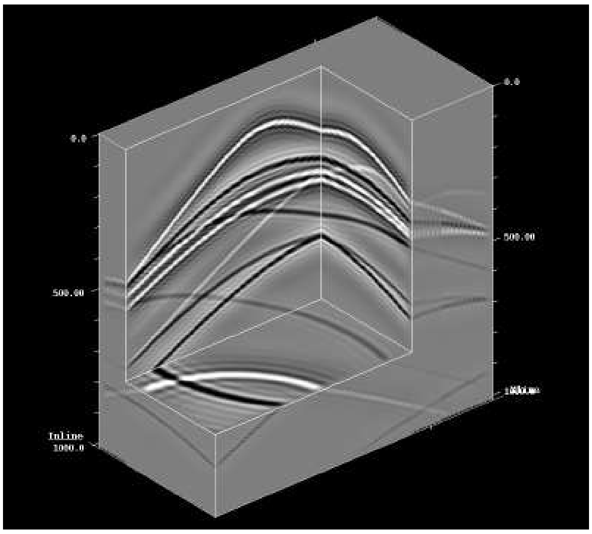
We have also obtained substantial uplift in the quality of real vibroseis data by employing this paper's sparsity-based separation algorithm (Neelamani and Krohn, 2008). The application of the proposed framework to further improve acquisition efficiency is an interesting topic for future investigation.

* Though the curvelet transform is not an orthonormal transform (its a tight frame), it enjoys several desirable properties of orthonormal transforms.

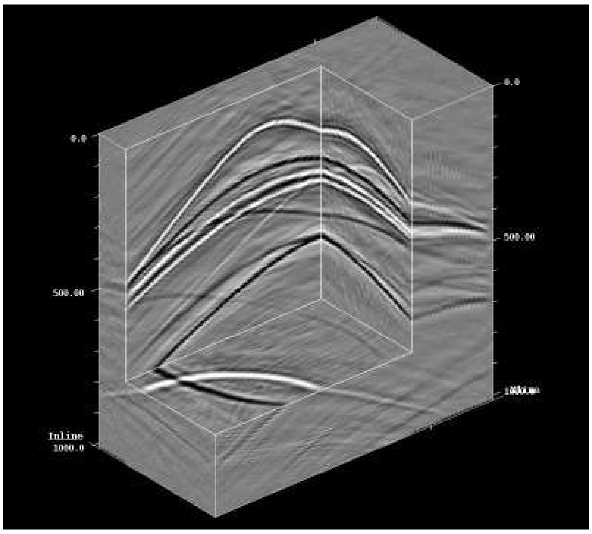


(a) (b) (c)

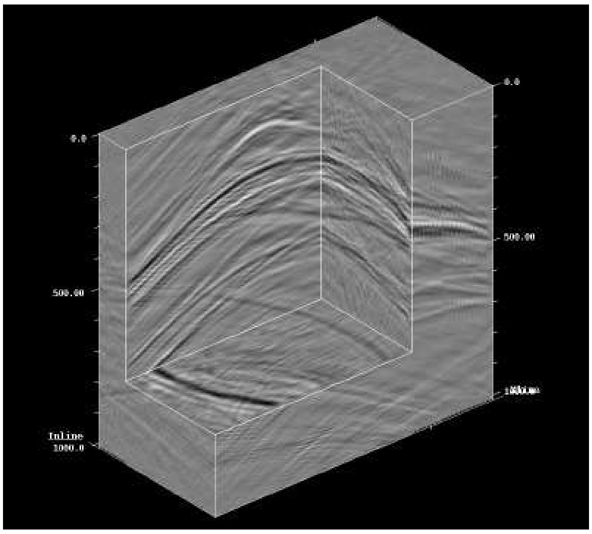
Figure 1: (a) Shot and receiver geometry. (b) Sample simultaneous sweeps. (c) Receiver measurements.



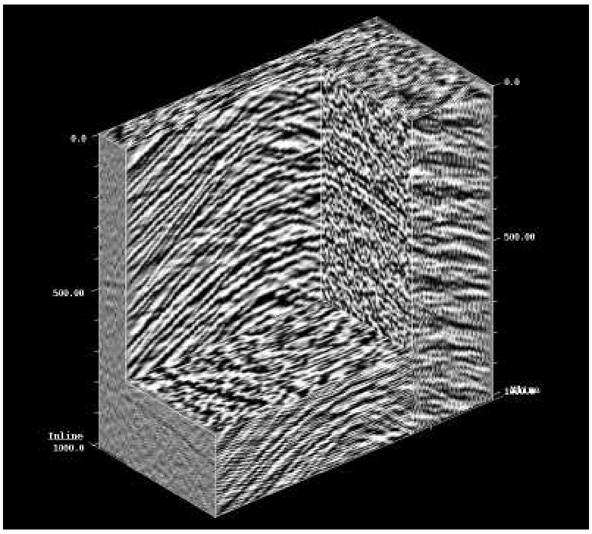
(a) Desired Green's function.



(b) Simultaneous source modeling estimate (8x faster).



(c) Difference ((a) minus (b))



(d) Cross-correlation estimate.

Figure 2: Simulation results.

REFERENCES

Allen, K. P., M. L. Johnson, and J. S. May, 1998, High Fidelity Vibratory Seismic (HFVS) Method for acquiring seismic data: 68th Ann. Internat. Mtg, 140–143, Soc. of Expl. Geophys.

Bajwa, W. U., J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, 2007, Toeplitz-structured compressed sensing matrices: IEEE/SP 14th Workshop on Statistical Signal Processing (SSP), 26–29.

Baraniuk, R., July 2007, Compressive sensing [lecture notes]: Signal Processing Magazine, IEEE, **24**, 118–121.

Candès, E., L. Demanet, D. Donoho, and L. Ying, 2006a, Fast discrete curvelet transforms: SIAM Multiscale Model. Simul., **5**, 861–899.

Candès, E. J., J. K. Romberg, and T. Tao, 2006b, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223.

Donoho, D., 2006, Compressed sensing: Information Theory, IEEE Transactions on, **52**, 1289–1306.

Ikelle, L. T., 2007, Coding and decoding: Seismic data modeling, acquisition, and processing: 77th Ann. Internat. Mtg, 66–70, Soc. of Expl. Geophys.

Krohn, C. E. and M. L. Johnson, 2006, HFVS^{tm}: Enhanced data quality through technology integration: Geophysics, **71**, E13–E23.

Neelamani, R. and C. E. Krohn, 2008, Simultaneous sourcing without compromise: Presented at the 70th EAGE Conference and Exhibition, Rome.

Romberg, J. K., 2007, Sensing by random convolution: Presented at the IEEE 2nd International Workshop on Computational Advances in Multi- Sensor Adaptive Processing.