Abstract. The recovery of geological reflection coefficients from seismic data includes a deconvolution operation. The sparse spike deconvolution algorithm used in seismic inversion is computed with an $l_1$ minimization. Although this procedure was developed in 1973, there is no mathematical model that explains the efficiency of this approach for seismic data. Using recent results on sparse signal representations in redundant dictionaries, this paper proves that reflectivities that are sums of Diracs sufficiently far away can indeed be recovered with an $l_1$ penalized deconvolution. Numerical examples on seismic signals illustrate these results.

1. Seismic Inversion

Measuring the reflectivity of the underground is necessary for mineral and oil exploration. Seismic techniques compute this reflectivity by sending pressure wave in the underground and by recording the reflected pressure waves on the surface, as a function of time and spatial position. Seismic inversion includes different steps such as migration and stacking to invert the wave propagation equation. At a given position of the surface, in a first approximation one can relate the resulting seismic data $Y$ to the underground reflectivity $R$ through a convolution equation, as a function of the depth variable $z$. The convolution kernel is a “wavelet” $\psi$ which depends upon the pressure wave sent in the underground. A bounded noise $W$ that incorporates all the errors of this linear model is added to the convolution equation:

$$Y = \psi * R + W.$$  

To invert the convolution equation (1), geophysicists model the reflectivity $R$ as a sum of Diracs that are reflectivity coefficients between different homogeneous layers:

$$R = \sum_{i \in S} a_i \delta_i$$

Each Dirac $\delta_i$ is located at a depth $i$ which is a junction of two homogeneous geological layers.

Clearbout and Muir [4] proposed in 1973 to use a $l_1$ minimization to recover $R$. Santosa and Symes [12] implemented this idea in 1986 with an $l_1$ relaxed minimization. The resulting sparse spike deconvolution algorithm defines the solution as:

$$R = \arg \min_f \frac{1}{2} \|Y - \psi * f\|^2 + \gamma \|f\|_1.$$

Daubechies, Defrise and De Moll [5], Chambolle [3] and Figueiredo and Nowak [8] proposed in 2003 a new iterative algorithm to solve this $l_1$ minimization problem. All numerical results in this paper are computed with this algorithm.

Although sparse spike deconvolutions are often used in seismic data processing, there has been very little mathematical analysis of the estimation error depending upon the properties of the underground reflectivity. In the following we introduce a minimum scale parameter on the reflectivity model. The minimum scale $\Delta$ of $R = \sum_{i \in S} a_i \delta_i$ is the minimum distance between two reflectivity Diracs in $R$:

$$\Delta = \min_{(i,j) \in S^2} |i - j|.$$

The minimum scale of $R$ depends only on its support $S$ so this notion applies to a set $S$ as well. Using recent results on sparse signal representations in redundant dictionaries, we shall prove that for any wavelet $\psi$ a sparse spike deconvolution can compute the exact support of a reflectivity if its minimum scale is sufficiently large. This result is valid for general $l_1$ penalized deconvolution algorithms.

2. Deconvolution Without Noise

To simplify the mathematical analysis, the deconvolution problem without noise is studied first. We want to find conditions under which $R$ can be exactly recovered from:

$$Y = \psi * R$$

with an $l_1$ minimization:

$$R_0 = \arg \min \|f\|_1, \quad \text{with } \psi * f = Y.$$
The wavelet $\psi$ is a band-pass filter which removes the lowest frequencies as well as high frequencies.

**Figure 1.** At the top is a seismic image with a vertical axis corresponding to depth. A grey point corresponds to a zero. Below are shown three results of a sparse spike deconvolution using a parameter $\gamma$ in (2) that decreases from top to bottom. Observe that the supports of these solutions increase as $\gamma$ decreases.

If $R = \sum_{i \in S} a_i \delta_i$ then

$$Y(z) = \psi \ast R(z) = \sum_{i \in S} a_i \psi(z - i).$$  \hspace{1cm} (5)

If the support of $\psi$ is included in an interval of width $\Delta$ then if the minimum scale of $R$ is larger than $\Delta$ the wavelet components $\psi(z - i)$ do not overlap, as illustrated by figure 2. It thus seems clear that the $a_i$ as well the positions $i$ can be recovered from $Y$ and that the $l_1$ minimization recovers $R_0 = R$. Our goal is to extend this property to filters $\psi$ that are not compactly supported and find the largest possible lower bound on the minimum scale to reconstruct signals with an $l_1$ minimization.

To use recent results on sparse signal representations in redundant dictionaries, we introduce a dictionary constructed by translating the wavelet $\psi$ at all locations. It is a matrix whose column vectors are

$$D = [g_i = \psi \ast \delta_i \text{ for } 1 \leq i \leq N].$$
Viewing $D$ as a matrix, one can rewrite $l_1$ minimization:
\begin{equation}
R_0 = \arg \min_f \|f\|_1, \quad \text{with } Df = Y.
\end{equation}
Gribonval and Nielsen in [11] give a condition for a vector $R$ to be recovered by (6). This condition relates the support of $R$ to the kernel of $D$ without using the more classical coherence measure of $D$. Tropp [13] and Fuchs [9, 10] have refined these conditions to guarantee the reconstruction of $R$ with the $l_1$ minimization (6).

Using an approach similar to Fuchs and Tropp, the following lemma introduces a weaker condition called Weak Exact Recovery Coefficient (WERC), which is sufficient to guarantee the exact reconstruction of $R$ in our deconvolution problem.

**Lemma 1.** Let $S \subset \{1, \ldots, N\}$, we define
\[
WERC(S) = \frac{\sup_{g \in S} \sum_{k \in S} |\langle g_k, g_i \rangle|}{1 - \sup_{i \in S} \sum_{k \in S, k \neq i} |\langle g_k, g_i \rangle|}
\]
If $WERC(S) < 1$ and if the support of $R$ is included in $S$ then $R$ is the only solution of (6).

This lemma is proved in [7]. Using this WERC condition the following theorem proves that any signal whose minimum scale is larger than a $\Delta$ large enough is uniquely recovered by an $l_1$ minimization.

**Lemma 2.** If $\psi$ is a filter whose Fourier transform is $C^1$ then there exists $\Delta$ and $\Lambda$ such that any set $S$ whose minimum scale is above $\Delta$ satisfies $WERC(S) \leq \Lambda < 1$.

This lemma is also proved in [7]. As a consequence of Lemma 1 and Lemma 2, it results that any signal $R$ whose minimum scale is larger than $\Delta$ is recovered by the $l_1$ minimization (6) from the convolved data $Y = DR = R \ast \psi$.

3. Sparse Spike Deconvolution With Noise

Let us now come back to the original noisy deconvolution problem (1), with a convolution operator that is rewritten using the dictionary matrix $D$:
\[
Y = DR + W.
\]
We shall prove that a sparse spike deconvolution with an $l_1$ minimization can recover a signal whose support is the same as the support of $R$. An approximation of $R$ is then computed by finding the minimum norm signal which is solution of 3, and whose support is equal to the calculated support. This is performed with alternate projections on this support as in the algorithm of Candès and Tao [2].

A sparse spike deconvolution computes
\begin{equation}
R_0 = \arg \min_f \frac{1}{2} \|Df - Y\|^2 + \gamma \|f\|_1, \quad \gamma > 0.
\end{equation}
Fuchs [9, 10] and Tropp [13] have introduced two conditions which guaranty that the support of $R_0$ is the same as the support of $R$. Using their work, the following theorem proves that the support of signals having a sufficiently large minimum scale is recovered by (7) if $\gamma$ is well chosen.

**Theorem 1.** Suppose that $\psi$ has a Fourier transform which is $C^1$. There exists $\Delta$ and $\Lambda$ such that if $R$ has a minimum scale larger than $\Delta$, $WERC(S) < \Lambda$ and if
\begin{equation}
\gamma \geq \frac{\|W\|_\infty \|\psi\|_1}{(1 - \Lambda)}
\end{equation}
then the solution (7) has a support included in the support of $R$.

Moreover if
\[
\min_{z, R(z) \neq 0} |R(z)| > 5\|W\|_\infty \|\psi\|_1 (1 - \Lambda)
\]
then there exists $\gamma$ such that if $R$ has a minimum scale larger than $\Delta$ then the solution (7) has a support equal to support of $R$.

This theorem is proved in [7]. It shows that if the regularization parameter $\gamma$ is large then the solution has a support included in the support of the original signal. Moreover, if the original signal has spikes of sufficiently large amplitude then its support can be exactly recovered by a sparse spike deconvolution.

4. Minimum Scale Bound

Lemma 1 shows that the lower bound $\Delta$ on the signal minimum scale can be calculated with the WERC, which depends on the dictionary $D$ and hence of the wavelet $\psi$. We give a numerical method to compute such a lower bound for a given $\psi$. Since
\[
\langle g_i, g_j \rangle = \langle \psi \ast \delta_i, \psi \ast \delta_j \rangle = \psi \ast \tilde{\psi}(|i - j|)
\]
with $\tilde{\psi}(z) = \psi(-z)$. Let us introduce
\[
\forall k, \phi(k) = \max_{|j| \geq k} \psi \ast \tilde{\psi}(j).
\]
Let $\Delta_0$ be the minimum scale of $S$. One can verify that
\[
\alpha(S) = \sup_{i \in S} \sum_{k \in S, k \neq i} |\langle g_k, g_i \rangle| \leq 2 \sum_k \phi(k \Delta_0)
\]
and
\[
\sup_{j \in S} \sum_{k \in S} |\langle g_k, g_j \rangle| \leq \max_{j \in \Delta_0} (\phi(j) + \phi(\Delta_0 - j)) + \alpha(S).
\]
Using these two equations, we obtain an upper bound of \( WERC(S) \) as a function of the minimum scale \( \Delta_0 \). Using lemma 1 we derive a lower bound \( \Delta \) on the minimum scale of a signal \( R \) to recover it with an \( l_1 \) minimization. This uses the \( ERC \) criteria introduced by Tropp, in [13]. Tropp proves that if a set \( S \) satisfies \( ERC(S) < 0 \) then there exists a signal supported in \( S \) and whose support can’t be recovered by (7). If the minimum scale of \( S \) is \( \Delta_1 \), this proves that the constant \( \Delta \) in Theorem 1 is larger than \( \Delta_1 \). We thus get a lower bound on \( \Delta \).

For a filter \( \psi \) such that \( \psi(\omega) = \cos^2(\omega) \) for \( \omega \in [-\pi/2, \pi/2] \) and \( \psi(\omega) = 0 \) for \( |\omega| \in ]\pi/2, \pi] \), the lower bound is equal to the upper bound and the optimal \( \Delta \) in Theorem 1 is 5.

If \( \psi(\omega) \) is dilated by a factor 5/6, the upper and lower bounds remain equal and \( \Delta = 6 \). If it is dilated by 2/3 the upper and lower bounds are also equal and \( \Delta = 8 \). If it is dilated by 1/2 the calculated upper bound is 12 and the calculated lower bound is 11. The optimal value for \( \Delta \) in Theorem 1 may thus be 11 or 12.

This shows that the minimum scale of a sparse signal gives a precise criteria to evaluate if it can be reconstructed with a sparse spike deconvolution using an \( l_1 \) penalization. These results were applied to seismic signal processing but remain valid for other deconvolution problems.

References


[6] D. L. Donoho. For most underdetermined systems of equations, the minimal \( l_1 \) norm near-solution approximate the sparsest near-solution. 2004.


