Title

Microseismic source deconvolution: Wiener filter versus minimax, Fourier versus wavelets, and linear versus nonlinear

Journal

Journal of the Acoustical Society of America, 115(6)

ISSN

0001-4966

Authors

Ching, J Y
To, A C
Glaser, S D

Publication Date

2004-06-01

Peer reviewed

Microseismic source deconvolution: Wiener filter versus minimax, Fourier versus wavelets, and linear versus nonlinear

Jianye Ching,[a] Albert C. To, and Steven D. Glaser[b]
Department of Civil and Environmental Engineering, University of California at Berkeley, Berkeley, California 94720

(Received 3 August 2002; revised 13 February 2004; accepted 13 February 2004)

Deconvolution is commonly performed on microseismic signals to determine the time history of a dislocation source, usually modeled as combinations of forces or couples. This paper presents a new deconvolution method that uses a nonlinear thresholding estimator, which is based on the minimax framework and operates in the wavelet domain. Experiments were performed on a steel plate using artificially generated microseismic signals, which were recorded by high-fidelity displacement sensors at various locations. The source functions were deconvolved from the recorded signals by Wiener filters and the new method. Results were compared and show that the new method outperforms the other methods in terms of reducing noise while keeping the sharp features of the source functions. Other advantages of the nonlinear thresholding estimator include (1) its performance is close to that of a minimax estimator, (2) it is nonlinear and takes advantage of sparse representations under wavelet bases, and (3) its computation is faster than the fast Fourier transform.
 [DOI: 10.1121/1.1705658]

PACS numbers: 43.60.Pt, 43.40.Le, 43.60.Bf [JCB] Pages: 3048–3058

I. INTRODUCTION

Microseismic (MS) signals carry important information about the dislocation mechanisms within the solid that contains the source. It is of great interest to infer the characteristics of the dislocation source using the recorded MS signals, and a common way of achieving this is to deconvolve the recorded MS signals by the corresponding Green's functions. The result of the deconvolution is an estimate of the MS source function, e.g., the slip time history of the dislocation source. Using the estimated source function, it is possible to infer the dislocation mechanisms of the crack that generates the MS energy. This technique has been routinely implemented by geophysicists to invert seismic source functions using seismograms (e.g., Stump and Johnson, 1977) and by researchers in nondestructive testing (e.g., Michaels *et al.*, 1981) and fracture mechanics (e.g., Scruby *et al.*, 1985; Kim and Sachse, 1986) to determine the fracture mechanisms of cracks in solids.

A standard way of performing deconvolution is to implement time-invariant Wiener filters (e.g., Oldenburg, 1981; Bertero, 1989; Chung and Liu, 1998), which are Wiener filters with stationary assumptions. The time-invariant Wiener filters are linear estimators operating in the Fourier domain. It is intuitive to implement such time-invariant Wiener filters because (1) sinusoidal functions are eigenfunctions of convolution operators, therefore it is convenient and efficient to represent the operators in the Fourier domain. (2) Because of (1), the Fourier coefficients of stationary random processes over different frequencies are uncorrelated, and this simplifies the analyses if the MS source function is also stationary. (3) They are computationally convenient since there is only a linear inverse problem to solve, and the fast Fourier transform can be calculated rapidly.

However, time-invariant Wiener filters may perform poorly if the assumed stationary properties are not appropriate. Also, Wiener filters reject the use of a nonlinear estimator, and the Fourier basis does not necessarily provide sparse representations for MS signals and MS source functions. Last, to implement Wiener filters usually requires the knowledge of the noise power spectrum, which is another parameter to be determined.

In this paper, we introduce the nonlinear thresholding estimator (NTE) developed by Donoho and Johnstone (1992) and Johnstone and Silverman (1997). The estimator provides the following advantages: (1) It does not assume a stationary source function. Instead, the estimator is based on the minimax concept (Bickel and Doksum, 1977) from decision theory. We argue that a minimax estimator is more suitable for a deconvolution problem than time-invariant Wiener filters. (2) The NTE does not require the knowledge of the noise power spectrum. (3) The resulting algorithm is faster than the fast Fourier transform. (4) The NTE is close to a minimax estimator for a large class of functions, which no linear estimators can achieve. (5) The NTE uses sparse representations of wavelet bases, and it is particularly suitable for estimating MS source functions (MS source functions are transient and are therefore sparse under wavelet bases).

The NTE has been introduced and applied to deconvolution problems by Masuda *et al.* (1999) for synthetic data. In this paper, we apply the NTE to experimental data and discuss the following concepts that were not presented in Masuda *et al.* (1999) but are crucial to deconvolution problems: (1) the importance of sparse representations of MS signals and MS source functions, (2) the advantage of using a nonlinear filter for transient MS source functions, and (3) the advantage of a minimax estimator over Wiener filters.

[a]Electronic mail: jyching@caltech.edu
[b]Electronic mail: glaser@ce.berkeley.edu

We first define the problem of MS source deconvolution in Sec. II. Then in Sec. III, we introduce Wiener filters and time-invariant Wiener filters. In Sec. IV, we introduce the minimax framework and the NTE, and some important properties of the NTE are discussed. Section V states the key considerations regarding how to choose the estimator for deconvolution problems. Section VI presents two case studies, from which we compare the performances of four different deconvolution strategies. Sections VII and VIII contain discussions and conclusions.

II. DEFINITION OF THE PROBLEM

The MS deconvolution problem is defined as follows. Assume G is a known Green's function characterizing the host from the MS source to the sensor, $z(t)$ is stationary Gaussian random process (noise) at time t, and T is the total number of discrete data points. Given a noisy MS sensor signal $\{d(t): t=1,2,...,T\}$, where

$$d(t)=\int G(t-\tau)f(\tau)d\tau+z(t), \quad t=1,2,...,T, \tag{1}$$

our goal is to estimate the unknown MS source function f. We assume d, f, and z in the equation are all continuous-time functions in $\mathbf{L}^2$ ($\mathbf{L}^2$ is the function space that contains all functions with finite $\mathbf{L}^2$ norms). We discuss two estimators of f, Wiener filters and minimax estimator, and show that the minimax framework can lead to nonlinear estimators based on thresholding rules.

III. WIENER FILTER ESTIMATE OF AN UNKNOWN SOURCE FUNCTION

To begin with the discussions for Wiener filters (Haykin, 1991; Clarkson, 1993), let the estimated f, denoted by $\hat{f}$, be generated by passing the data d through a general linear filter h (not necessarily time-invariant):

$$\hat{f}(t)=\int_{-\infty}^{\infty} h(t,\tau)\cdot d(\tau)\cdot d\tau, \tag{2}$$

where $h(t,\tau)$ is the impulse response of the linear filter. The optimal linear filter h that minimizes the risk (also known as mean-square error)

$$E_Z(\|\hat{f}-f\|_2^2), \tag{3}$$

where $E_Z(\cdot)$ denotes expectation with respect to the probability density function (PDF) of z, and $\|\hat{f}-f\|_2^2 \equiv \int(\hat{f}(t)-f(t))^2\,dt$ is the squares of the $\mathbf{L}^2$ norm of g, is a (time-varying) Wiener filter. A Wiener filter can be found by solving the following Wiener–Hopf equation:

$$\int_{-\infty}^{\infty} R_d(\xi-t,\tau)\cdot h(t,\tau)\cdot d\tau=R_{df}(\xi-t,t) \quad \forall \xi,t, \tag{4}$$

where $R_d(t,\tau)$ is the auto-correlation function between $d(t)$ and $d(\tau)$; $R_{df}(t,\tau)$ is the cross-correlation function between $d(t)$ and $f(\tau)$.

In practice, a time-invariant Wiener filter is often used, which requires the assumption that both f and d are weakly stationary processes. Under this assumption, Eq. (4) becomes

$$\int_{-\infty}^{\infty} R_d(\xi-\tau)\cdot h(\tau)\cdot d\tau=R_{df}(\xi) \quad \forall \xi, \tag{5}$$

where $R_d(\xi)$ and $R_{fd}(\xi)$ are the auto-correlation function of d and the cross-correlation function between f and d, respectively, with lag ξ; h is now a time-invariant filter, and $h(\tau)$ is its impulse response. With the assumption that f and z in Eq. (1) are uncorrelated, Eq. (5) can be written in the frequency (Fourier) domain as

$$(|G(\omega)|^2\cdot|F(\omega)|^2+P_Z(\omega))\cdot H(\omega)=G(\omega)^*\cdot|F(\omega)|^2, \tag{6}$$

where $G(\omega)$ is the Fourier transform of the Green's function G; $H(\omega)$ and $F(\omega)$ are the Fourier transforms of h and f; G^* denotes the complex conjugate of G; and $P_Z(\omega)$ is the power spectrum of z. It follows that

$$H(\omega)=\frac{G(\omega)^*}{|G(\omega)|^2+P_Z(\omega)/|F(\omega)|^2}. \tag{7}$$

In principle, $|F(\omega)|^2$ is not completely known *a priori*, so a key step in designing a Wiener filter is to choose a reasonable $|F(\omega)|^2$. Different choices of $|F(\omega)|^2$ lead to different designs of Wiener filters. For instance, Oldenburg (1981) chooses $|F(\omega)|^2=1/\tan\theta$, where θ is a resolution factor that varies from 0 to 2π; Bertero (1989) chooses $|F(\omega)|^2=\mu\cdot|D(\omega)|^2/P_Z(\omega)$, where μ is usually chosen as the average power of the noise spectrum; Chung and Liu (1998) choose $|F(\omega)|^2=|D(\omega)|^2/|G(\omega)|^2$. All of these algorithms lead to low-pass filters, i.e., estimators with time-invariant moving-average windows.

IV. MINIMAX ESTIMATE OF AN UNKNOWN SOURCE FUNCTION

A. Minimax rule

Under the minimax framework, it is assumed that f is smooth in a certain sense, i.e., f is within some known function set $\mathbf{H}$. For instance, if f is a solution of some second-order differential equation, f is within $\mathbf{H}=\mathbf{C}^2$, the function space of twice differentiable functions. Formally, the minimax estimate of f over $\mathbf{H}$ is the one that ''minimizes'' the ''maximum'' risk over $\mathbf{H}$. In other words,

$$\begin{aligned}\hat{f}_H^M(d)&=\arg\inf_{\hat{f}(\cdot)} R^M(\hat{f}(d),H)\\&=\arg\inf_{\hat{f}(\cdot)}\left[\sup_{f\in H} R(\hat{f}(d),f)\right]\\&=\arg\inf_{\hat{f}(\cdot)}\left[\sup_{f\in H} E_Z\left\|\hat{f}(d)-f\right\|_2^2\right],\end{aligned} \tag{8}$$

where $\hat{f}_H^M(\cdot)$ denotes the minimax estimator, which is an operator mapping $d\in\mathbf{L}^2$ to another function in $\mathbf{L}^2$, and $R^M(\hat{f}(d),H)$ denotes the maximum risk over the function set $\mathbf{H}$. That is, the minimax estimator $\hat{f}_H^M(\cdot)$ is the operator $\hat{f}(\cdot)$ that minimizes the maximum risk $R^M(\hat{f}(d),H)$. The minimized risk is called the minimax risk, i.e.,

$$R^{Min\ max}(d,H)=\inf_{\hat{f}(\cdot)} R^M(\hat{f}(d),H). \tag{9}$$

It can be seen from Eqs. (8) and (9) that

$$E_Z\|\hat{f}_H^M(d)-f\|_2^2=R(\hat{f}_H^M(d),f)<R^{Min\ max}(d,H) \quad \forall f\in H, \tag{10}$$

where the definition of $R(\hat{f}(d),f)$ is identical to the one in Eq. (3). That is, the resulting risk $R(\hat{f}_H^M(d),f)$ for every f in **H** is guaranteed to be less than the fixed number $R^{Min\ max}(d,H)$, which is minimized in the minimax framework. In the case that **H** is sufficiently large, the minimax framework allows us to find an estimate of f without applying strong subjective constraints to f, e.g., the stationary properties assumed by time-invariant Wiener filters.

Unfortunately, even for the simple deconvolution problem, the minimax estimator cannot be found for any arbitrary **H**. Nevertheless, Donoho and Johnstone (1994) showed that it is possible to construct estimators, the so-called nonlinear thresholding estimators, which have maximum risks that are close to the minimax risk [Eq. (9)] over all Besov function classes (Triebel, 1983), i.e.,

$$\sup_{f\in B_{p,q}^s} R(\hat{f}^{NTE}(d),f)\sim R^{Min\ max}(d,B_{p,q}^s), \tag{11}$$

where $\hat{f}^{NTE}(\cdot)$ is the NTE estimator; $B_{p,q}^s$ is a Besov function space defined as follows:

$$B_{p,q}^s=\left\{f:\left(\int\left(\frac{w_{r,p}(f;h)}{h^s}\right)^q\frac{dh}{h}\right)^{1/q}<\infty\right\},$$

$$w_{r,p}(f;h)=\left\|\sum_{k=0}^{r}\binom{r}{k}(-1)^k f(t+kh)\right\|_{L^p}. \tag{12}$$

Since the NTEs operate under wavelet bases, we discuss wavelet bases in the next section.

B. Wavelet bases and sparse representations

Traditionally, deconvolution problems are usually manipulated in the Fourier domain, e.g., a low-pass filter is a linear estimator operating in the Fourier domain. However, the Fourier basis can seldom provide a sparse representation for f unless f is a periodic function. In this section, we introduce wavelet bases, which are more suitable to represent transient functions. In MS signal analysis, wavelets have been applied to material classification (Qi, 2000) and fracture mode classification and determination of phase and group velocities (Takemoto *et al.*, 2000).

A wavelet basis (Daubechies, 1992; Meyer, 1992; Mallat, 1998) is an orthonormal basis that decomposes the $\mathbf{L}^2$ function space into a series of function spaces containing functions with different degrees of oscillations:

$$L^2=\bigoplus_{j=-\infty}^{\infty} W_j, \tag{13}$$

where $\oplus$ is the direct sum of linear spaces. W_j with a small j contains functions with slow oscillations; W_j with a large j contains functions with rapid oscillations. Wavelet basis functions include father wavelets $\{\varphi_{m,k}:k=1,2,...,2^m\}$, which span $V_m=\oplus_{j=-\infty}^{m-1}W_j$, and mother wavelets $\{\psi_{j,k}:k=1,2,...,2^j\}$, which span W_j (Mallat, 1998). We denote the wavelet coefficients of a function $\xi\in\mathbf{L}^2$ by $\alpha_{m,k}(\xi)=\langle\varphi_{m,k},\xi\rangle$ and $\beta_{j,k}(\xi)=\langle\psi_{j,k},\xi\rangle$, where $\langle g,h\rangle$ is the inner product between function g and h. For a sampled function of sample size $T=2^J$, the decomposition of f is

$$f=\sum_{k=1}^{2^m}\alpha_{m,k}(f)\cdot\varphi_{m,k}+\sum_{j=m}^{J-1}\sum_{k=1}^{2^j}\beta_{j,k}(f)\cdot\psi_{j,k}. \tag{14}$$

An appealing feature of a wavelet basis is that many useful functions have sparse representations under wavelet bases. The sparseness is due to the following two facts: (1) with a careful selection of mother wavelets, the functions can be made orthogonal to polynomials less than a certain degree (Mallat, 1998), and (2) mother wavelets have compact support (or nearly compact support) regions. Consequently, transient signals which are usually piecewise smooth functions are sparse in the wavelet domain since many wavelet coefficients $\beta_{j,k}$ in the smooth regions that are close to polynomials are close to zeros. In addition, a nonsmooth function whose region of support does not intersect with the support region of a mother wavelet is orthogonal to that mother wavelet; hence, a small support region also helps supply sparseness.

C. Nonlinear thresholding estimators

Equation (1) can be rewritten using an operator form as the following equation:

$$d(t)=(G\cdot f)(t)+z(t), \quad t=1,2,...,T, \tag{15}$$

where $(G\cdot)$ denotes the corresponding convolution operator. Then the following equation is equivalent to Eq. (15):

$$(G^{-1}\cdot d)(t)=f(t)+(G^{-1}\cdot z)(t), \quad t=1,2,...,T, \tag{16}$$

where G^{-1} denotes the inverse operation of G. The philosophy of the NTE is to compare the magnitudes of $G^{-1}\cdot d$ and $G^{-1}\cdot z$ under a wavelet basis. If the wavelet coefficient $\beta_{j,k}(G^{-1}\cdot d)$ is much larger than $\beta_{j,k}(G^{-1}\cdot z)$, $\beta_{j,k}(G^{-1}\cdot d)$ is considered to contain a significant amount of information about f and is therefore kept; otherwise, $\beta_{j,k}(G^{-1}\cdot d)$ is discarded and set to zero. Formally, the two NTEs proposed by Johnstone and Silverman (1997), called hard and soft thresholding estimators, are described by the following algorithms:

(1) Compute $y=G^{-1}\cdot d$ (or pseudo-inverse if G^{-1} does not exist) and find $\alpha(y)_{m,k}$ and $\beta(y)_{j,k}$.

(2) Find the estimate of the standard deviation of $\beta_{j,k}(G^{-1}\cdot z)$ for scale j, denoted by $\hat{\sigma}_j$, by the median value of $\{|\beta_{j,k}(y)|:k=1,...,2^j\}$ divided by 0.6745 (Donoho and Johnstone, 1992) and compute the level-dependent thresholding level $\lambda_j=\hat{\sigma}_j\sqrt{2\log(T)}$.

Discussion: Due to the fact that f is sparse under wavelet bases, the set $\{|\beta_{j,k}(f)|:k=1,...,2^j\}$ contains mostly numbers that are close to zero. Therefore, most of the numbers contained in the set $\{|\beta_{j,k}(y)|:k=1,...,2^j\}$ are close to those in the set $\{|\beta_{j,k}(G^{-1}z)|:k=1,...,2^j\}$. Consequently, the median value of $\{|\beta_{j,k}(y)|:k=1,...,2^j\}$ should be close to $\{|\beta_{j,k}(G^{-1}z)|:k=1,...,2^j\}$.

Also, since mother wavelets are almost eigenfunctions for many convolution operators (Meyer, 1992), so the stationary Gaussian noise $G^{-1}\cdot z$ becomes almost independent under wavelet bases. Therefore, the median value of $\{|\beta_{j,k}(G^{-1}z)|:k=1,...,2^j\}$ divided by 0.6745 is a plausible estimate of its standard deviation. The constant $\sqrt{2\log(T)}$ is chosen because $\beta_{j,k}(G^{-1}z)$ can be greater than λ_j with only an insignificant probability.

(3) Threshold $\beta_{j,k}(y)$ uses the hard thresholding, i.e.,

$$\tilde{\beta}_{j,k}(y)=\begin{cases}\beta_{j,k}(y) & |\beta_{j,k}(y)|>\lambda_j,\\ 0 \quad \text{otherwise,}\end{cases} \qquad \forall j,k, \tag{17}$$

or uses the soft thresholding, i.e.,

$$\tilde{\beta}_{j,k}(y)=\begin{cases}(|\beta_{j,k}(y)|-\lambda_j)\operatorname{sgn}(\beta_{j,k}(y)) & |\beta_{j,k}(y)|>\lambda_j,\\ 0 \quad \text{otherwise,}\end{cases} \qquad \forall j,k. \tag{18}$$

(4) Find the NTE of f, denoted by $\hat{f}^{NTE}$, by the inverse wavelet transform

$$\hat{f}^{NTE}=\sum_{k=1}^{2^m}\alpha_{m,k}(y)\phi_{m,k}+\sum_{j=m}^{J-1}\sum_{k=1}^{2^j}\tilde{\beta}_{j,k}(y)\varphi_{j,k}. \tag{19}$$

Equations (17) and (18) indicate that the NTEs depend on the data d in nonlinear ways. The thresholding estimator described above does not require any knowledge about the power spectrum of the noise z: all we need to do is to estimate $\hat{\sigma}_j$ for each scale. The thresholding estimate $\hat{f}^{NTE}$ is nearly minimax over every Besov function space, while no linear estimator can be nearly-minimax over every Besov function space (Johnstone and Silverman, 1997). This uniform near-minimaxity over all the Besov spaces has the following advantage: although in principle we need to know which Besov space [i.e., the values of p, q, and s in Eq. (11)] that f belongs to before we implement the NTEs, in practice we can ignore this step because it is very unlikely that f does not belong to any of the Besov function classes.

V. KEY CONSIDERATIONS

We have introduced Wiener filters and NTEs in the last few sections. We now summarize the key considerations regarding the original deconvolution problem:

A. Wiener filters versus minimax

Which framework should we choose, Wiener filters or minimax? For designing a time-invariant Wiener filter, assigning a reasonable $|F(\omega)|^2$ [Eq. (7)] is crucial. Unfortunately, for the deconvolution problem it is usually difficult to gain knowledge about $|F(\omega)|^2$ *a priori*. Also time-invariant Wiener filters assume that f is stationary; however, it is hardly the case that the MS source function is stationary. MS source functions are usually transient, hence nonstationary.

On the other hand, the minimax framework does not assume f to be stationary. Also, under the minimax framework, it is guaranteed that the resulting risk $R(\hat{f}_H^M(d),f)$ is smaller than $R^{Min\ max}(d,H)$ for every f in $\mathbf{H}$ [see Eq. (10)]. We have seen that the NTEs perform close to a minimax estimator over any Besov function class; this makes the NTEs even more attractive.

B. Linear versus nonlinear

In general, nonlinear estimators can be assumed superior to linear ones since linear estimators are special cases of nonlinear ones. Another way of justifying the superiority is through the following example. Consider the following function estimation problem:

$$d(t)=f(t)+e(t), \quad t=1,2,...,T, \tag{20}$$

where e is a white Gaussian process. The goal is to estimate f based on the data $\{d(t):t=1,2,...,T\}$. Now let us restrict ourselves with a keep-or-kill estimator in the Fourier domain: the estimation procedure is to first compute the Fourier coefficients of d, keep $K(\ll T)$ Fourier coefficients of d and set others to zeros, and the estimate of f is computed as the inverse Fourier transform of the manipulated Fourier coefficients of d.

Clearly, the best strategy is to keep the K largest Fourier coefficients if the goal is to minimize the estimation risk $R(\hat{f}(d),f)$, and this strategy corresponds to a nonlinear estimator since the decision of keeping or killing any coefficient depends on the Fourier amplitude of d. A linear estimator prohibits the dependence of the estimator on the data, so the corresponding K Fourier coefficients have to be prechosen. Therefore, all linear estimators perform worse than the best nonlinear estimator unless the prechosen K coefficients coincide with the K largest coefficients.

Wiener filters are linear estimators while the NTEs are nonlinear. In fact, the NTEs behave similarly as the best nonlinear estimator mentioned in the last paragraph: the NTEs tend to keep wavelet coefficients with large amplitudes and kill those with small amplitudes.

C. Fourier versus wavelets

Which orthonormal basis should we choose, Fourier or wavelets? A transient signal f is usually sparse in the wavelet domain. However, f is usually not sparse under the Fourier domain unless f is periodic or stationary. Therefore, the wavelet domain is usually a better platform for function estimation than the Fourier domain for transient f since in the wavelet domain, it is easier to distinguish the patterns of the signal f and the noise z (z is never sparse under any orthonormal basis). MS source functions are usually transient, i.e., nonperiodic and nonstationary, and so are not sparse in the Fourier domain, but are sparse under the wavelet domain. Time-invariant Wiener filters operate in the Fourier domain while the NTEs operate in the wavelet domain.

VI. CASE STUDIES

A. Experimental setup and data processing

In order to demonstrate the utility of our proposed method, several experiments are conducted on a 600 $\times$600$\times$50 mm steel plate (see Fig. 1). Artificial sources are used to generate MS signals, which are recorded by wide-

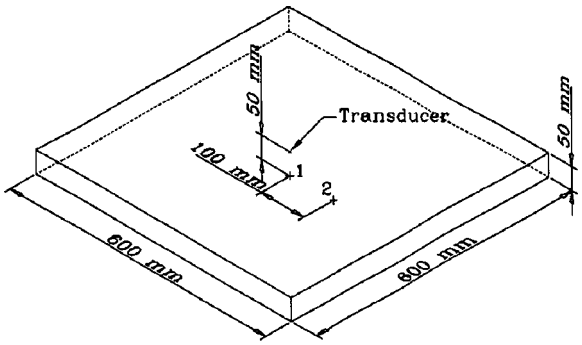

FIG. 1. The geometry of the steel plate (in mm). The location of the sensor is indicated, and numbers 1 and 2 in the figure indicate the two locations where the artificial sources were applied.

band, high-fidelity displacement sensors (Glaser *et al.*, 1998). Using the sensor signals, our goal is to estimate the time histories of the artificial MS source functions. Three classical artificial vertical force sources (Breckenridge *et al.*, 1990) are applied on the surface of the steel plate. They are (1) the fracture of a glass capillary of 0.4-mm diameter (called the capillary break), (2) the fracture of a pencil lead of 0.3-mm diameter (called the pencil lead break), and (3) the impact of a 3.16-mm-diam stainless steel ball dropped vertically (called the ball drop). Figure 2 shows the highly accurate source time functions of a 0.4-mm capillary break, a 0.3-mm pencil lead break and a 3.175-mm-diam steel ball drop obtained by deconvolution from signals measured by capacitive sensors (Breckenridge *et al.*, 1990). All these sources are well documented and can be reproduced easily and consistently.

The capillary break is generated by loading a rod vertically on a glass capillary placed flat on the plate until the capillary shatters suddenly. Immediately before the capillary breaks, the plate is displaced locally by the vertical force exerted by the rod through the capillary. Once the capillary breaks, the surface displacement recovers suddenly. This sudden rebound generates a steplike function that has a short rise time [Fig. 2(a)]. The pencil lead break is generated by loading vertically the tip of a mechanical lead pencil that is held at some angle relative to the plate until the lead tip breaks. The mechanism of the pencil lead break is similar to that of the capillary break except that when the lead tip breaks, the fracture energy displaces the broken lead tip further into the plate before the local displacement on the plate rebounds. Therefore, the source function of the pencil lead break has a distinct ''dip'' at the wavefront before the rebound occurs with a steplike function as shown in Fig. 2(b). The ball drop source is the most reproducible and has a bell-shaped source function [Fig. 2(c)].

Because of the wideband steplike source function, the capillary break signal is used to derive the empirical Green's function from which the pencil lead break and the ball drop sources are calculated. In the calculation, the rise time of the capillary break is neglected and its source signal is treated as a perfect step function. Therefore, the numerical derivative (the central difference method) of the capillary break signal is treated as a perfect impulse response and is used as the empirical Green's function. The technique has been performed by Michaels *et al.* (1981), who have shown excellent results.

A Glaser-type broadband piezoelectric sensor (Glaser *et al.*, 1998) (bandwidth 12 kHz to 1 MHz) is placed at the center of the upper side of the steel plate (see Fig. 1). The sources are applied at two different surface locations of the steel plate (see Fig. 1): location 1 (epicentral through plate), on the bottom side of the plate directly below the sensor, and location 2 (off-epicentral through plate), on the bottom side

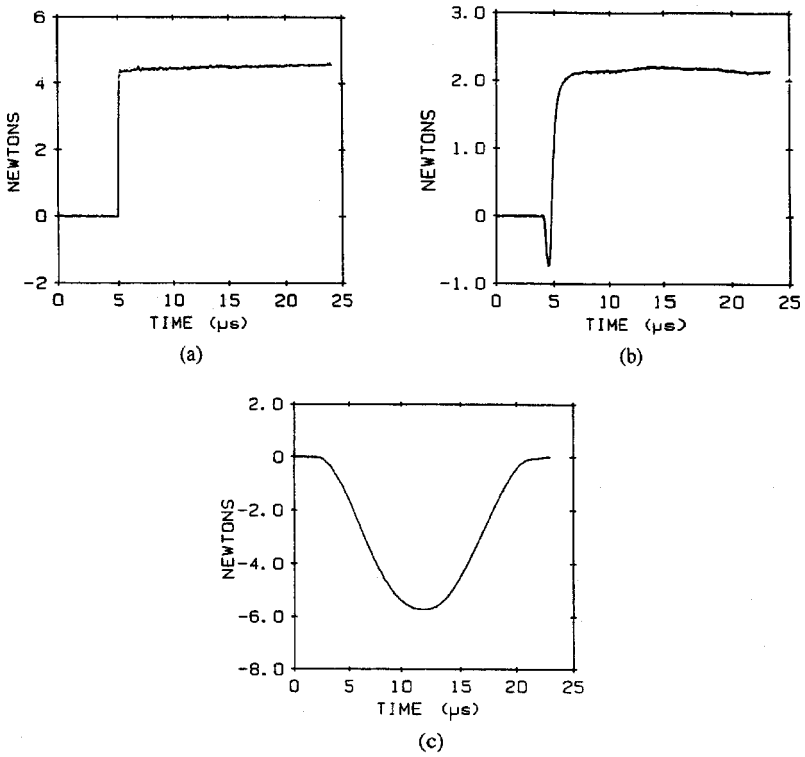

FIG. 2. (a) The glass capillary break source function; (b) The pencil lead break source function; (c) The ball drop source function; (from Breckenridge *et al.*, 1990).

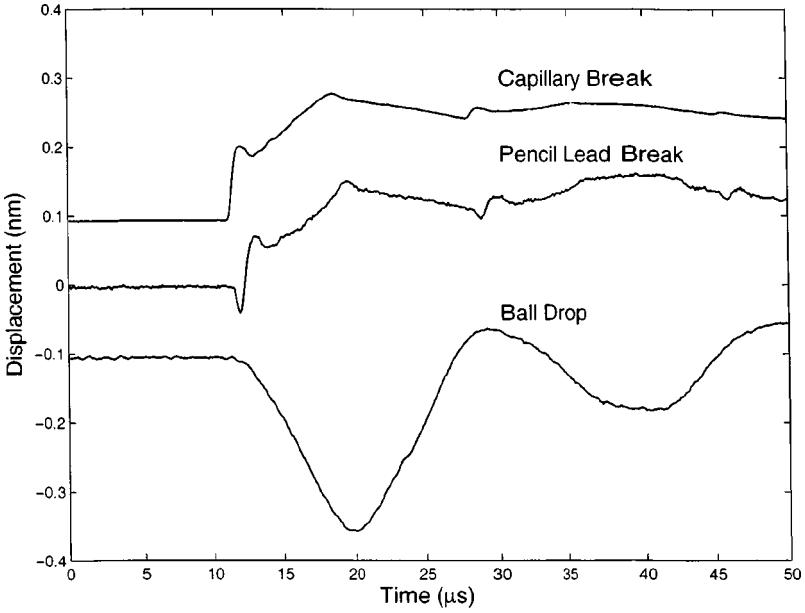

FIG. 3. Sensor signals for the three artificial sources applied at location 1. The pencil lead break and ball drop signals are added with white Gaussian noise of constant amplitude.

at a horizontal distance of 100 mm from the sensor. The transient responses (MS signals) are digitally sampled at a 0.1-μs interval with 14-bit resolution.

Given the recorded MS signals and the Green's functions, i.e., d and G in Eq. (1), we adapt four deconvolution methods to estimate the unknown MS source functions—f in Eq. (1). The four methods are (1) the nonlinear hard thresholding estimator, (2) the maximum likelihood estimator, i.e., the direct deconvolution estimate $G^{-1}d$, which maximizes the likelihood function, and two time-invariant Wiener filters, (3) Eq. (7) with $|F(\omega)|^2=|D(\omega)|^2/|G(\omega)|^2$, and (4) Eq. (7) with $|F(\omega)|^2=\text{avg}(|D(\omega)|^2)/|G(\omega)|^2$, where $|D(\omega)|^2$ is the power spectrum of the data estimated by the smoothed periodogram of the data (Marple and Lawrence, 1987). Both Wiener filters are intuitive because one expects the prior PDF of the unknown source function to be close to that of the data. These two Wiener filters also coincide with the deconvolution filters used by Chung and Liu (1998) and Oldenburg (1981), respectively. The nonlinear hard thresholding estimator is implemented using the procedure presented in Sec. VIII with the Daubechies wavelet (Mallat, 1998) of order 12.

As discussed before, the numerical derivative of the capillary break signal is used as the empirical Green's function at each source location (locations 1 and 2 in Fig. 1), and the pencil lead break and the ball drop source functions are estimated for each location. Artificial white Gaussian noise of constant variance is added to the recorded pencil lead break and ball drop signals to study the effects of noise (noise was not added to the capillary break signals since they were used as empirical Green's functions). The amplitude of the artificial Gaussian noise is much larger than usual experimental noise, so the noise power spectrum, which is needed for the two Wiener filters, can be computed *a priori*. Note that the noise power spectrum is not necessary for the NTE and the maximum likelihood estimator.

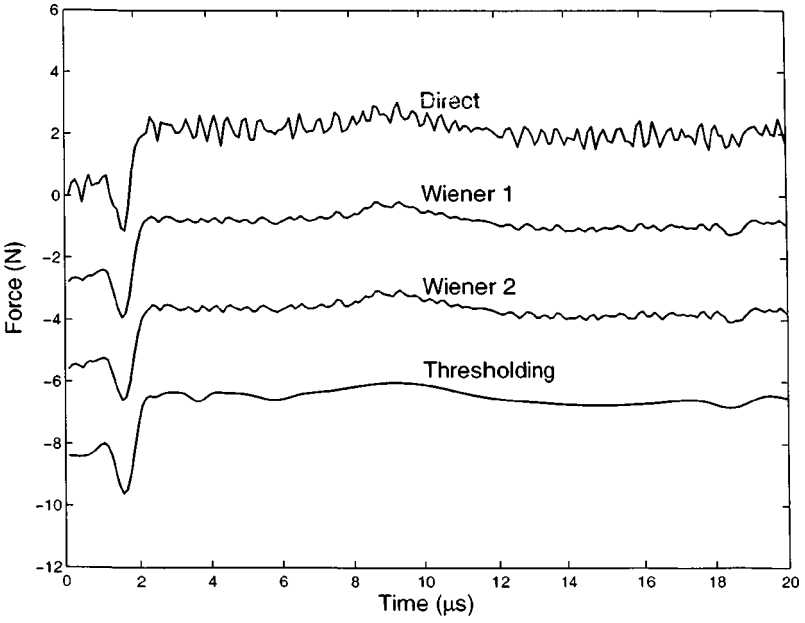

FIG. 4. The estimated pencil lead source functions using the four deconvolution methods (location 1).

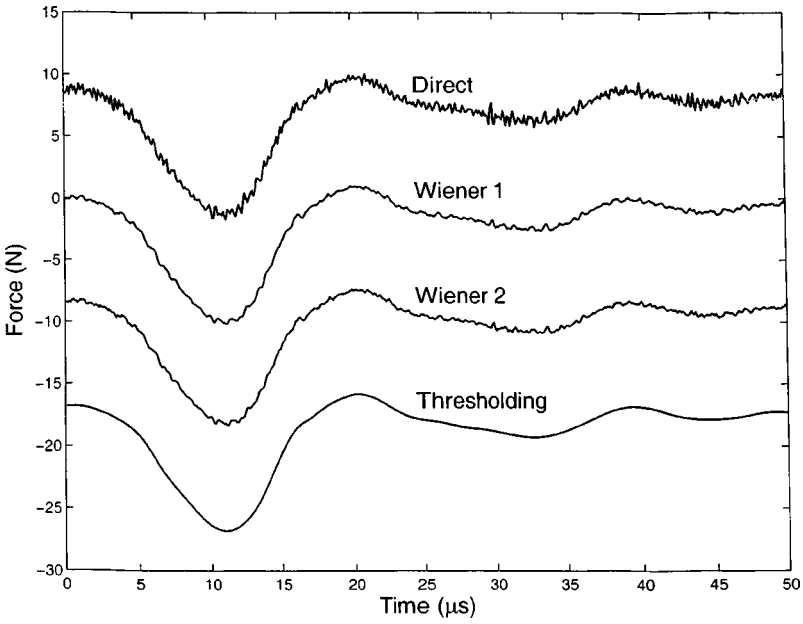

FIG. 5. The estimated ball drop source functions using the four deconvolution methods (location 1).

B. Analysis of results at location 1

The sensor signals for the three artificial sources applied at location 1 (epicentral) are shown in Fig. 3. Gaussian noise of 5% signal power is added to the pencil lead break and ball drop signals to simulate noisy experimental conditions. The numerical derivative of the capillary break signal is taken as the empirical Green's function. Given the empirical Green's function, the pencil lead and ball drop source functions are estimated using the four deconvolution methods. (The estimated pencil lead and ball drop source functions are shown in Figs. 9 and 10, respectively.) The four estimates are similar except that the direct deconvolution estimate is more noisy. Note that the four estimators preserve the sharp downward jump at the beginning of the pencil lead break source function.

All the waveforms of our estimated source time functions (Figs. 4 and 5) compare favorably to those estimated by Breckenridge *et al.* (1990) shown in Figs. 2(b) and 2(c). Table I summarizes the rise times and peak forces of our estimated source time functions and those of Breckenridge *et al.* (1990). The values of the rise time and peak force of the pencil lead break source time functions are very similar. For the ball drop, the rise time and peak force for various estimates (Table I) are different because the diameter of the ball and the height of the drop are different among the estimates. In general, the larger the diameter of the ball and the higher the ball is dropped, the longer the rise time and the larger the peak force. Nonetheless, the waveforms of the ball drop estimates are very similar.

The sources applied at location 1 generate sensor signals of large amplitudes, so the signal-to-noise ratios are high. Therefore, the direct deconvolution estimates are not dramatically noisy, and the two Wiener filters and the NTE can only improve the results with limited degrees by filtering out the high-frequency noise from the direct deconvolution estimates. It is also found that the two Wiener filters and the NTE perform similarly in the location 1 cases.

Some measure of goodness of fit can be obtained by convolving each of the four estimated source functions with the empirical Green's function. These fits should be close to the actual recorded pencil lead or ball drop signals. The fits by the four estimators together with the recorded signals are shown in Figs. 6 and 7 for pencil lead and ball drop sources, respectively. The direct deconvolution estimate always give excellent fits since the estimate is simply $G^{-1}d$ and must return d when convolved with the Green's function. However, this perfect fit is not desirable since the direct deconvolution estimate does not only fit the signal but also the noise. The other three estimates all tend to fit the signals instead of the noise. But notice that the fits of the second Wiener filter estimates cannot adapt the sharp features of the signals. The fits of the first Wiener filter estimates can adapt those sharp features, but the estimates themselves are more noisy. Compared to the other estimates, the nonlinear thresholding estimates are noise-free while being able to adapt the sharp features of the signals.

TABLE I. Summary of properties of source time functions.

Sources	Rise time (μs)	Peak force (N)
Pencil lead break 0.3 mm (from Breckenridge *et al.*, 1990)	2.50	2.05
Pencil lead break 0.3 mm (at location 1 based on the new method)	2.15	1.84
Pencil lead break 0.3 mm (at location 2 based on the new method)	2.31	2.22
Ball drop (diameter 3.18 mm and drop height 2.6 mm) (from Breckenridge *et al.*, 1990)	9.50	5.72
Ball drop (diameter 3.16 mm and drop height 5.1 mm) (at location 1 based on the new method)	10.30	10.05
Ball drop (diameter 3.16 mm and drop height 3.9 mm) (at location 2 based on the new method)	9.00	8.68

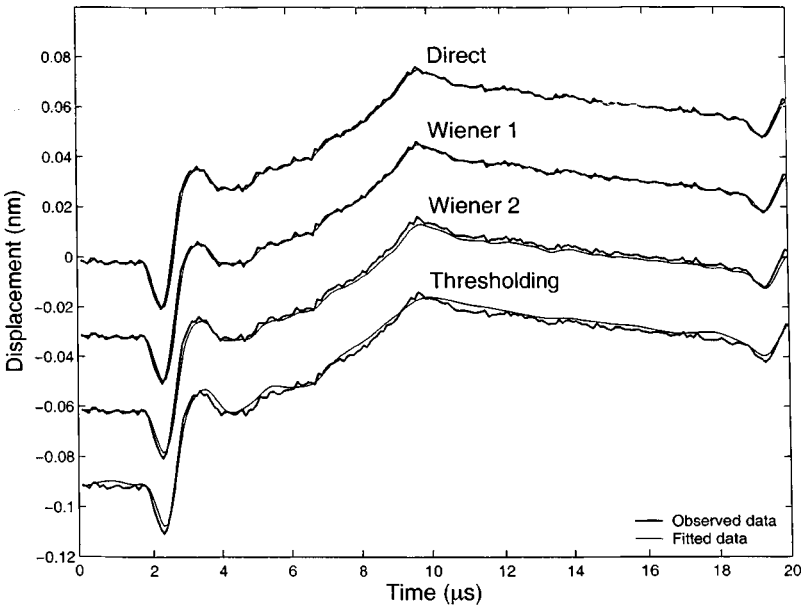

FIG. 6. The fits of the four estimates for the pencil lead break data (location 1). The recorded pencil lead break signal added with Gaussian noise is plotted as heavy lines.

C. Analysis of results at location 2

The sensor signals for the three artificial sources applied at location 2 (off-epicentral) are shown in Fig. 8. Gaussian noise of the same amplitude as the epicentral case is added to the pencil lead break and ball drop signals to simulate noisy experimental conditions. Since the off-epicentral signals are weaker, the corresponding signal-to-noise ratios are lower (the Gaussian noise is roughly 10% of the signal power). The numerical derivative of the capillary break signal is again taken as the empirical Green's function. The estimated pencil lead and ball drop source functions using the four deconvolution methods are shown in Figs. 9 and 10. Note that the direct deconvolution estimator is quite noisy because of the low signal-to-noise ratios, and the two Wiener filters and the NTE significantly improve the results by removing the high-frequency noise from the direct deconvolution estimates. Also note that the estimated source functions are not as sharp as those for location 1, and the Wiener filters and NTE tend to suppress sharp features of their estimates as they eliminate noise.

There is a clear difference between the two Wiener filters in Figs. 9 and 10: the first Wiener filter results in sharper estimates than the second one does, also the amplitudes of the first Wiener filter estimate are larger than those of the second one. On the other hand, the estimated ball drop source function of the first Wiener filter contains more high-frequency oscillations than that of the second one. This phenomenon is due to the bias-variance tradeoff: the first Wiener filter is less biased while its variance is larger, but the converse is true for the second one. Since the Wiener filters are linear, the second Wiener filter reduces the noise level by smoothing its estimates uniformly over time. However, this time-invariant smoothing strategy tends to shrink the overall amplitudes of the estimates as well, not just the noise ampli-

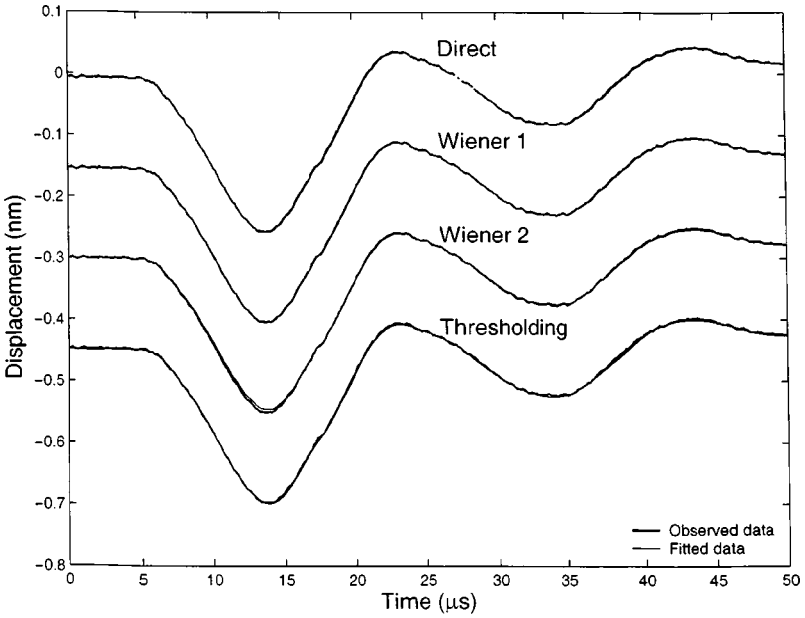

FIG. 7. The fits of the four estimates for the ball drop data (location 1). The recorded ball drop signal added with Gaussian noise is plotted as heavy lines.

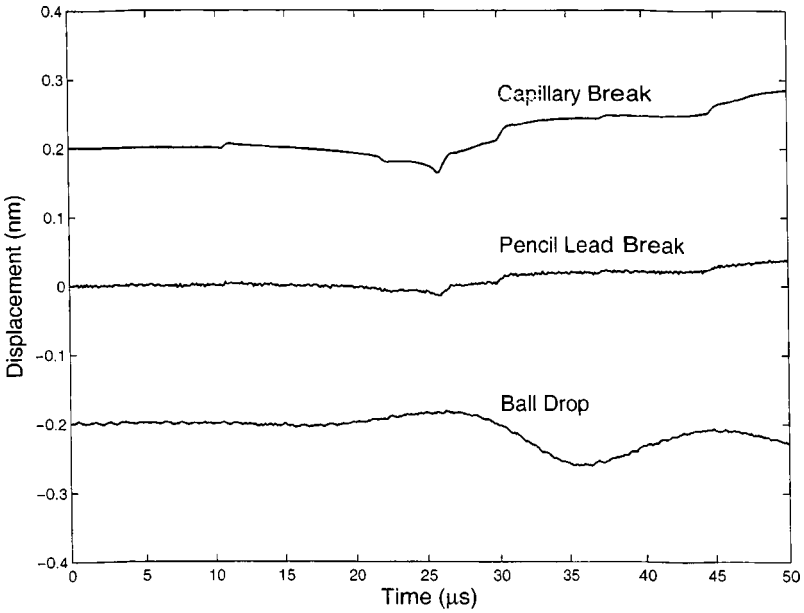

FIG. 8. Sensor signals for the three artificial sources applied at location 2. The pencil lead break and ball drop signals are added with white Gaussian noise of constant amplitude.

tude. Moreover, if the unknown function f is highly nonstationary, this linear time-invariant smoothing strategy can result in large bias. Both pencil lead break and ball drop source functions can be considered highly nonstationary. Therefore, large bias is found in the second Wiener filter for both the pencil lead break and ball drop source functions. This bias-variance tradeoff is not significant for the cases of location 1 since the signal-to-noise ratios are high.

On the other hand, the NTE simultaneously eliminates noise and preserves the amplitudes and sharp changes of the estimates. Note that the NTE also preserves the downward jump at the beginning of the pencil lead break source function, while the jump is smoothed out by the second Wiener filter. As discussed previously, this excellent performance is due to the implementation of the minimax framework of the NTE, the nonlinearity of the NTE, and sparse representations of transient functions under wavelet bases.

Similar to those at location 1, the waveforms of our estimated pencil lead break and ball drop source time functions at location 2 compare favorably with those obtained by Breckenridge *et al.* (1990) (Fig. 2). The rise times and peak forces of the sources at location 2 are summarized in Table I and are similar to those obtained at location 1. The fits by the four estimators together with the recorded signals are shown in Figs. 11 and 12 for pencil lead and ball drop sources, respectively. They are all close to the recorded signals.

VII. DISCUSSION

Our results show that the NTE outperforms the two time-invariant Wiener filters in the experimental case studies. The success of the NTE is because wavelet bases provide sparse representations for the transient MS source functions, the estimator is nonlinear, and it is based on the minimax

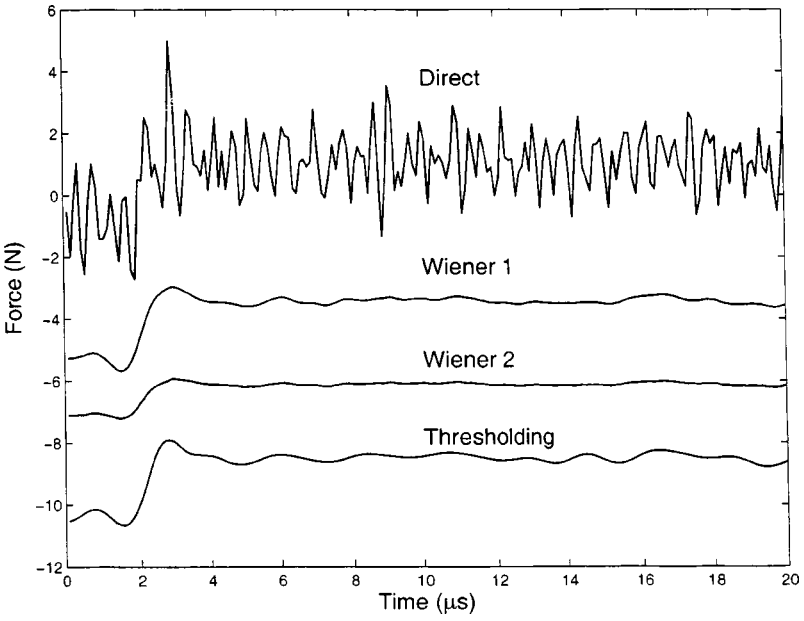

FIG. 9. The estimated pencil lead source functions using the four deconvolution methods (location 2).

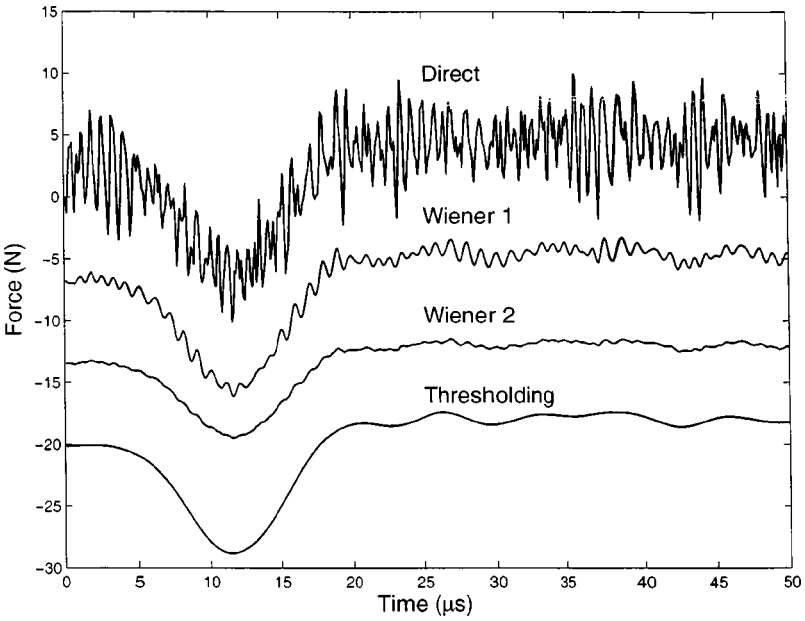

FIG. 10. The estimated ball drop source functions using the four deconvolution methods (location 2).

framework. On the other hand, Wiener filters perform worse probably because the assumed stationary properties are not appropriate for MS source functions.

Some previous research has found that it is possible to finely tune a time-invariant Wiener filter to get better estimation results, e.g., Oldenburg (1981) developed a time-invariant Wiener filter with a fine-tuning parameter θ. The second Wiener filter used in the case studies is actually a special case of the Wiener filter having θ equals 45°. It is possible to adjust θ to get a better estimate than our second Wiener filter estimate, but this leads to an *ad hoc* parameter fitting procedure while it still excludes the use of a nonlinear estimator. In contrast, the NTE does not have such a free parameter.

VIII. CONCLUSIONS

Two types of estimators, Wiener filters and minimax estimators, for microseismic deconvolution problems are described and discussed. Time-invariant Wiener filters are linear filters operating in the Fourier domain. These estimators assume stationary properties for the unknown microseismic source function. They are usually not optimal for deconvolution problems due to the following reasons: (1) microseismic source functions are usually transient and nonstationary, hence the assumed stationary properties are not appropriate. (2) Microseismic source functions do not necessarily have sparse representations under the Fourier basis. (3) Linear estimators are usually not the best estimators to use.

A nonlinear thresholding estimator is presented in this paper, which operates in the wavelet domain. This estimator is usually more suitable than Wiener filters for deconvolution problems because (1) wavelet bases provide sparse representations transient microseismic source functions, (2) the estimator is nearly minimax over all Besov function classes; therefore, in practice, prior information about the microseis-

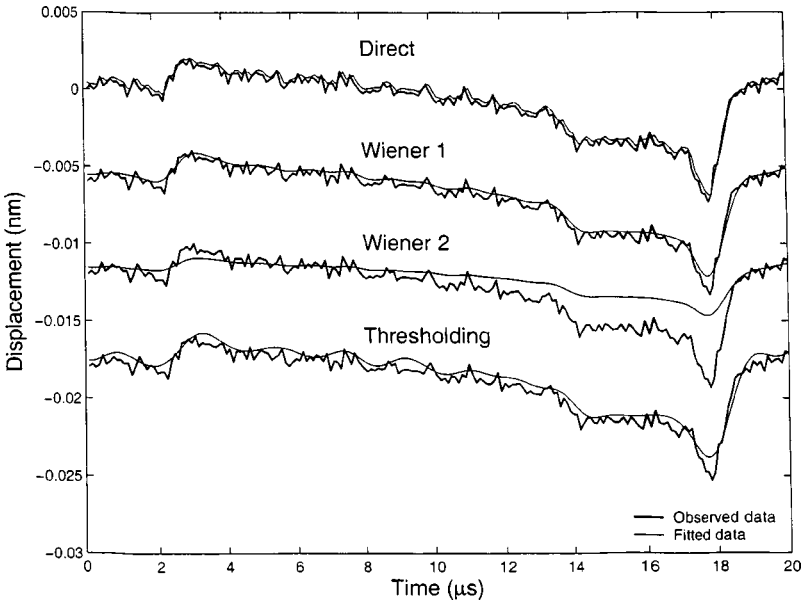

FIG. 11. The fits of the four estimates for the pencil lead break data (location 2). The recorded pencil lead break signal added with Gaussian noise is plotted as heavy lines.

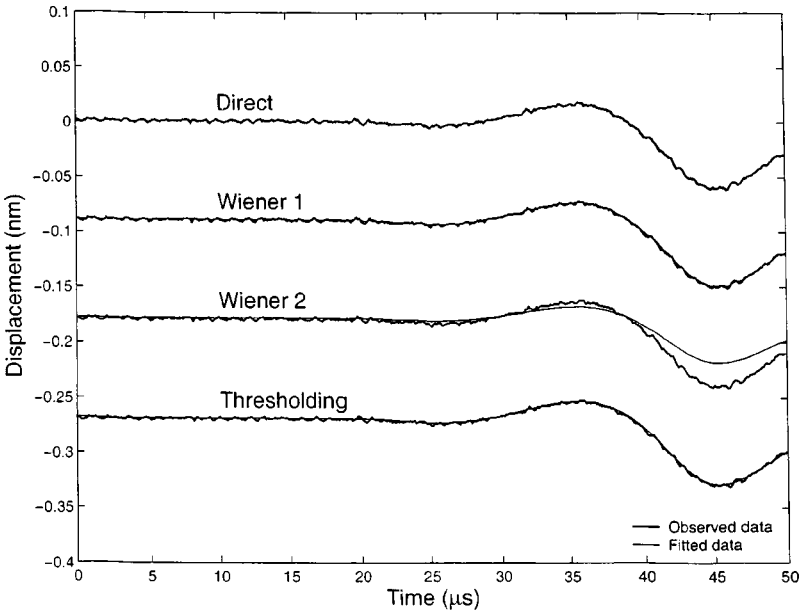

FIG. 12. The fits of the four estimates for the ball drop data (location 2). The recorded ball drop signal added with Gaussian noise is plotted as heavy lines.

mic source function is not needed, and (3) the estimator is nonlinear.

Two time-invariant Wiener filters and the nonlinear thresholding estimator are compared using several case studies of experimental microseismic data. From the results of the case studies, we conclude that the nonlinear thresholding estimator outperforms the Wiener filters. The nonlinear thresholding estimator is able to achieve two goals simultaneously: (1) effectively remove estimation noise and (2) preserve the sharp features in the source functions. The two Wiener filters fail to achieve the two goals simultaneously in some of the case studies.

ACKNOWLEDGMENTS

We would like to acknowledge the valuable discussions and opinions from Dr. Jenher Jeng for the wavelet theory and the concepts of nonparametric statistics. We also would like to acknowledge the valuable opinions and comments from the reviewers.

Bertero, M. (**1989**). ''Linear inverse and ill-posed problems,'' in *Advances in Electronics and Electron Physics*, edited by P. W. Hawkes (Academic, New York), pp. 1–120.

Bickel, P. J., and Doksum, K. A. (**1977**). *Mathematical Statistics: Basic Ideas and Selected Topics* (Holden-Day, San Francisco).

Breckenridge, F. R., Proctor, T. M., Hsu, N. N., Fick, S. E., and Eitzen, D. G. (**1990**). ''Transient sources for acoustic emission work,'' in *Progress in Acoustic Emission, V, Proc. 10th International Acoustic Emission Symposium*, edited by K. Yamagushi, H. Takakashi, and H. Niitsuma (Japanese Society for Non-Destructive Inspection, Tokyo), pp. 20–37.

Chung, C., and Liu, Q. H. (**1998**). ''Inversion of source-time functions using borehole array sonic waveforms,'' J. Acoust. Soc. Am. **103**, 3163–3168.

Clarkson, P. M. (**1993**). *Optimal and Adaptive Signal Processing* (CRC, Boca Raton, FL).

Daubechies, I. (**1992**). *Ten Lectures on Wavelets* (SIAM, Philadelphia).

Donoho, D. L., and Johnstone, I. M. (**1992**). ''Minimax estimation via wavelet shrinkage,'' Technical Report, Department of Statistics, Stanford University.

Donoho, D. L., and Johnstone, I. M. (**1994**). ''Ideal spatial adaptation by wavelet shrinkage,'' Biometrika **81**, 425–455.

Glaser, S. D., Weiss, G., and Johnson, L. R. (**1998**). ''Body waves recorded inside an elastic half space by an embedded, wideband velocity sensor,'' J. Acoust. Soc. Am. **104**, 1404–1412.

Haykin, S. (**1991**). *Adaptive Filter Theory*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ).

Johnstone, I. M., and Silverman, B. W. (**1997**). ''Wavelet threshold estimators for data with correlated noise,'' J. R. Stat. Soc. Ser. B. Methodol. **59**, 319–351.

Kim, K. Y., and Sachse, W. (**1986**). ''Characteristics of a microseismic source from a thermal crack in glass,'' Int. J. Fract. **31**, 211–231.

Mallat, S. G. (**1998**). *A Wavelet Tour of Signal Processing* (Academic, San Diego).

Marple, S. L., and Lawrence, S. (**1987**). *Digital Spectral Analysis with Applications* (Prentice-Hall, Englewood Cliffs, NJ).

Masuda, A., Yamamoto, S., and Sone, A. (**1999**). ''Deconvolution of time series using wavelet transform (Introduction of wavelet packets with soft-thresholding and generation of optimized wavelets),'' JSME Int. J., Ser. C **42**, 188–194.

Meyer, Y. (**1992**). *Wavelets and Operators* (Cambridge U.P., London).

Michaels, J. E., Michaels, T. E., and Sachse, W. (**1981**). ''Applications of deconvolution to microseismic signal analysis,'' Mater. Eval. **39**, 1032–1036.

Oldenburg, O. W. (**1981**). ''A comprehensive solution to the linear deconvolution problem,'' Geophys. J. R. Astron. Soc. **65**, 331–357.

Qi, G. (**2000**). ''Wavelet-based AE characterization of composite materials,'' NDT&E Int. **33**, 133–144.

Scruby, C. B., Baldwin, G. R., and Stacey, K. A. (**1985**). ''Characterisation of fatigue crack extension by quantitative microseismic,'' Int. J. Fract. **28**, 201–222.

Stump, B. W., and Johnson, L. R. (**1977**). ''The determination of source properties by the linear inversion of seismograms,'' Bull. Seismol. Soc. Am. **67**, 1489–1502.

Takemoto, M., Nishino, H., and Ono, K. (**2000**). ''Wavelet transform: applications to AE signal analysis,'' in *Acoustic Emission—beyond the millennium*, edited by T. Kishi, M. Ohtsu, and S. Yuyama (Elsevier, New York), pp. 35–56.

Triebel, H. (**1983**). *Theory of Function Spaces* (Birkhäuser Verlag, Boston).

Zukas, J., Nicholas, T., Swift, H., Greszczuk, L., and Curran, D. (**1982**). *Impact Dynamics* (Wiley, New York).