Adaptive time-frequency detection and filtering for imaging in heavy clutter

L. Borcea and G. Papanicolaou and C. Tsogka

ADAPTIVE TIME-FREQUENCY DETECTION AND FILTERING FOR IMAGING IN HEAVY CLUTTER

L. BORCEA[†], G. PAPANICOLAOU[‡], AND C. TSOGKA[§]

Abstract. We introduce an adaptive approach for the detection of a reflector in a strongly scattering medium using a time-frequency representation of the array response matrix followed by a Singular Value Decomposition (SVD). We use the Local Cosine Transform (LCT) for the time-frequency representation and introduce a detection criterion that identifies anomalies in the top singular values, across frequencies and in different time windows, that are due to the reflector. The detection is adaptive because the time windows that contain the primary echoes from the reflector are not determined in advance. Their location and width is identified by searching through the time-frequency binary tree of the LCT. After detecting the presence of the reflector we filter the array response matrix to retain information only in the time windows that have been selected. We also project the filtered array response matrix to the subspace associated with the top singular value and then image using travel time migration. We show with extensive numerical simulations that this approach to detection and imaging works well in heavy clutter that is calibrated using random matrix theory so as to simulate regimes close to the experiments in [3]. While the detection and filtering algorithm presented here works well in general clutter it has been analyzed theoretically only for the case of randomly layered media [1].

Key words. array imaging, random media, local cosine transform, random matrix theory.

1. Introduction. Imaging with sensor arrays remote and small reflectors embedded in heterogeneous (cluttered), strongly back-scattering media is quite challenging. This is because the echoes from the reflectors are weak and overwhelmed by back-scattering from the inhomogeneities of the medium. In this paper we introduce an algorithm for detecting the presence of a reflector in a strongly scattering medium, and then imaging it after filtering the array data to minimize clutter effects.

A typical configuration for array imaging is shown in Figure 1.1. A source located at $\vec{\mathbf{x}}_s$ in an array $\mathcal{A}$ emits a broadband pulse $f(t)$ and the time traces of the scattered echoes $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ are recorded at receiver locations $\vec{\mathbf{x}}_r \in \mathcal{A}$, for $r = 1, \ldots, N$, and $t \in [0, T]$. We assume that each sensor in the array can play the dual role of source and receiver. The array response matrix

$$\mathbb{P}(t) = \{P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)\}_{r,s=1,\ldots,N}, \quad t \in [0, T], \tag{1.1}$$

is obtained by sending the same probing pulse $f(t)$ from the N sources in the array, and recording in each case the traces $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$, $s = 1, ..., N,\ r = 1, ..., N$.

Imaging reflectors in known and piecewise smoothly varying media can be done efficiently with travel time migration, which is used routinely in radar [18, 21, 23], seismic imaging [8, 19, 9], etc. In its simplest form, migration transforms the array data $\mathbb{P}(t)$ to an image $\mathcal{J}^{\rm KM}(\vec{\mathbf{y}}^s)$ at points $\vec{\mathbf{y}}^s$ in a search domain $\mathcal{D}$ that includes the reflectors and given by

$$\mathcal{J}^{\rm KM}(\vec{\mathbf{y}}^s) = \sum_{r=1}^{N_r} \sum_{s=1}^{N_s} P\left(\tau(\vec{\mathbf{x}}_r, \vec{\mathbf{y}}^s) + \tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}^s), \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s\right). \tag{1.2}$$

The data traces are "migrated" from the source at $\vec{\mathbf{x}}_s$ and the receiver $\vec{\mathbf{x}}_r$ to the search point $\vec{\mathbf{y}}^s$ with the sum travel times $\tau(\vec{\mathbf{x}}_r, \vec{\mathbf{y}}^s) + \tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}^s)$. In smoothly varying media with propagation speed $c(\vec{\mathbf{x}})$, the travel

[†]Computational and Applied Mathematics, Rice University, MS 134, Houston, TX 77005-1892. (borcea@caam.rice.edu)
[‡]Mathematics, Stanford University, Stanford CA 94305. (papanicolaou@stanford.edu)
[§]Applied Mathematics, University of Crete and IACM/FORTH, GR-71409 Heraklion, Greece (tsogka@tem.uoc.gr)

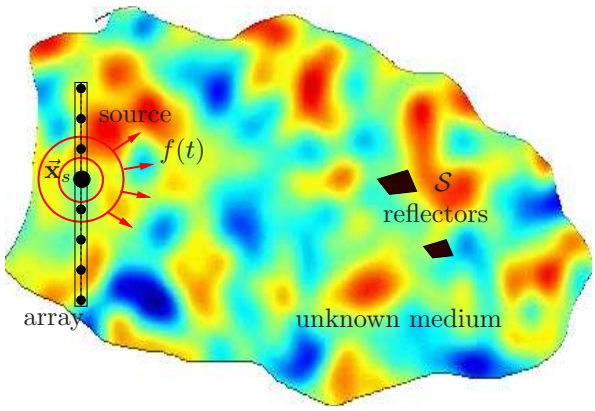

FIG. 1.1. *Typical configuration for array imaging.*

time is given by

$$\tau(\vec{\mathbf{x}},\vec{\mathbf{y}}) = \min \int c^{-1}(\mathbf{X}(u))\, du,$$

where the minimum is over all paths $\mathbf{X}(u)$, parametrized by arclength u, that start at $\vec{\mathbf{x}}$ and end at $\vec{\mathbf{y}}$. In a homogeneous medium with wavespeed c_0, the travel time is simply $\tau(\vec{\mathbf{x}},\vec{\mathbf{y}}) = |\vec{\mathbf{x}}-\vec{\mathbf{y}}|/c_0$. Peaks of the imaging function $\mathcal{J}^{\mathrm{KM}}$ give estimates of the location of the reflectors.

When there is no clutter, or when the clutter is weak, migration works well. It also works well with data corrupted by additive, uncorrelated instrument noise, if the array aperture is large, because the summation over the many sensors averages out the noise. Cumulative scattering by inhomogeneities in cluttered media has a significantly different distortion effect on the array data traces, that cannot be removed easily, by simply summing over the aperture. This is why migration is not useful for imaging in clutter. It produces images that are heavily speckled and peak at unpredictable locations.

To image in moderately strong clutter we developed the coherent interferometric (CINT) methodology [12, 14, 13, 15, 16]. CINT forms images by superposing time delayed, local cross-correlations of the array data traces $P(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)$, instead of the traces themselves. As in Kirchhoff migration, the delays are the round trip travel times computed in the smooth background medium, without the clutter. The cross-correlations are local because they are computed over suitable time windows of width $1/\Omega_d$ and over sensor offsets that do not exceed X_d. The two parameters Ω_d and X_d are the decoherence frequency and length and depend on the random medium. They are the frequency and receiver/source offsets $|\omega-\omega'|$ and $|\vec{\mathbf{x}}_r-\vec{\mathbf{x}}_{r'}|$ or $|\vec{\mathbf{x}}_s-\vec{\mathbf{x}}_{s'}|$ over which the Fourier transforms of two time traces $\widehat{P}(\omega,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)$ and $\widehat{P}(\omega',\vec{\mathbf{x}}_{r'},\vec{\mathbf{x}}_{s'})$ decorrelate due to the scattering by the clutter. In general, X_d varies with the frequency [13, 16]. If it can be approximated by a constant over the bandwidth then the CINT image has a simpler expression

$$\mathcal{J}^{\mathrm{CINT}}(\vec{\mathbf{y}}^s;\Omega_d,X_d) = \sum_{r,r'=1} \Psi\left(\frac{|\vec{\mathbf{x}}_r-\vec{\mathbf{x}}_{r'}|}{X_d}\right) \sum_{s,s'=1} \Psi\left(\frac{|\vec{\mathbf{x}}_s-\vec{\mathbf{x}}_{s'}|}{X_d}\right) \int_{-\infty}^{\infty} dt \Phi(t\Omega_d)$$
$$P(\tau(\vec{\mathbf{x}}_r,\vec{\mathbf{y}}^s)+\tau(\vec{\mathbf{x}}_s,\vec{\mathbf{y}}^s)-t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s) P(\tau(\vec{\mathbf{x}}_{r'},\vec{\mathbf{y}}^s)+\tau(\vec{\mathbf{x}}_{s'},\vec{\mathbf{y}}^s)-t,\vec{\mathbf{x}}_{r'},\vec{\mathbf{x}}_{s'}), \quad (1.3)$$

where Ψ and Φ are window functions of dimensionless arguments and order one support. This shows explicitly how the image is formed by superposing the local cross-correlations of the data traces delayed by the round

trip travel times $\tau(\vec{\mathbf{x}}_r, \vec{\mathbf{y}}^s) + \tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}^s)$ to the imaging point $\vec{\mathbf{y}}^s$. Thresholding by Ω_d and X_d over the frequency and sensor offsets is essential in CINT. It introduces a statistical smoothing in the imaging process that leads to self-averaging, statistically stable results [13, 16]. By statistical stability we mean that the images have negligible fluctuations induced by scattering in clutter. The images are essentially the same and close to their statistical mean, independent of the realization of the clutter.

Coherent interferometry works well in clutter when the range of the reflectors does not exceed a few transport mean free paths [28, 29, 31]. In heavy clutter, the coherent primary echoes from the reflectors are weak and overwhelmed by the medium back-scatter. Neither the traces nor their cross-correlations can be used directly for imaging. To extend the applicability of coherent imaging methods to such strong scattering media we introduce in this paper a preprocessing stage for the array data that detects the presence of reflectors and filters clutter back-scatter prior to imaging. The detection of the reflectors is based on the premise that they represent an anomaly in the pattern that is characteristic of signals received from a strongly backscattering random medium without a target. The image can be computed with the filtered data using either travel time migration or CINT. In this paper we use migration because it is simpler and it gives good results. In more complex reflector configurations we expect that it will be necessary to use CINT.

Concerning the clutter, we assume that the random medium fluctuations do not vary over time. This is not very restrictive as in most applications the signal gathering process at the array is so fast that the ambient medium can be assumed to be time independent. If signals are collected over different time periods, and there are time variations from one data gather to the next, then they do complicate the imaging process as they contribute to loss of coherence, which is what we seek to extract with our approach. We do not address this issue here.

In [10, 11] we considered array data filtering for imaging in randomly layered media. The filters are operators acting on the traces that annihilate echos from layers but preserve echos from compact reflectors. They are also efficient filters of fine layering effects and they improve significantly the quality of the images, as shown with a theoretical analysis and numerical simulations in [11]. However, they do not generalize to non-layered clutter.

The problem of imaging a reflector in a strongly scattering medium is also considered in [3]. These authors use a combination of DORT [27] with a filter that tends to preserve primary echoes from small reflectors but removes much of the multiple scattering in the array response matrix. DORT images with the array data filtered by projection on the span of singular vectors corresponding to the leading singular values of the array response matrix in the frequency domain, with entries $\widehat{P}(\omega, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$. This is done in [3] in an appropriate time window and in combination with another filter for multiple scattering effects. The most complex part in the approach of [3] is the identification of time-frequency windows for which the reflector is detectable. The identification requires the estimation of the probability distribution of the first singular value of the response matrix in the frequency domain so as to set a detection threshold.

In this paper we introduce a new and general approach for the detection of reflectors and for the filtering of heavy clutter effects. We use the singular value decomposition (SVD) of the local cosine transform (LCT) [25, 20] of the response matrix $\mathbb{P}(t)$. The LCT decomposes the traces in orthonormal bases constructed with

smooth time windows modulated by cosines [25]. The smooth windows avoid the appearance of artificial discontinuities in the transformed signals, which generate large amplitude coefficients at high frequencies. Moreover, the LCT is efficient for detection and filtering, because we can use the well-established fast algorithms for its implementation [25].

We use the LCT to decompose the data traces $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ over frequencies, locally in time. The decomposition is done on binary trees, where the time windows are halved from one level to the next. At level 0, which is the root of the tree, the local cosine transform of $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ is computed for the whole time interval, $t \in [0, T]$, and for all $r, s = 1, \ldots, N$. At level $l \geq 1$, the time interval is divided in 2^l windows of width $2^{-l}T$. The cosine transform coefficients are computed for each of these windows. The idea is to use the matrices formed with these coefficients to find, at the finest possible time scale (higher level l of the tree), the window that contains the coherent primary echoes from the reflectors. This is the adaptive part of the detection algorithm. If we knew the window, we could use a DORT-like approach to emphasize the echoes from the reflector that we wish to image. So in our approach, the difficult part is to identify the correct window.

Our window selection algorithm is based on the behavior of a group of largest singular values of the LCT of $\mathbb{P}(t)$, across frequencies. In particular, we seek to identify an anomalous pattern among the singular values as they vary over the frequencies in the bandwidth, indicating the possible presence of an object different from the inhomogeneities that make up the clutter. The idea is that, if only echoes from the cluttered background medium are present in a window, the largest singular values of the LCT in this window should all look alike and have a similar behavior across frequencies. On the other hand, we expect at least one singular value with different behavior when echoes from a detectable object are present in the time window. The detection is based on the premise that there is an anomaly or a defect in the spatial distribution of the inhomogeneities that clutter the medium and this, in turn, leads to an anomalous behavior of the singular values. The detection of the anomalous behavior is done with the singular value decomposition of the matrix whose columns are the largest singular values, across frequencies. The detection test is at present empirical. We do not have a proof that it cannot give false positive results in general clutter, where by false positive we mean that fluctuations of the singular values of response matrices from pure clutter backscatter may be mistakenly interpreted as anomalies due to a reflector. We have not seen such false positive results in our numerical simulations. The test is rigorously justified by the theory in [1], in the case of layered clutter.

After we have detected the relevant time windows, we filter the array data by zeroing the LCT coefficients in all other windows. Within the selected windows, we also project the LCT matrix on the subspace of the top singular values, that is, the ones with anomalous behavior. We then reconstruct the time traces with the inverse LCT, and image with travel time migration or with CINT. In the numerical simulations in this paper, the travel time migration images obtained from the filtered data are good so there is no need to consider CINT. However, we expect that CINT will be necessary to image in more complex cases, such as clusters of reflectors in clutter [17].

We focus in this paper on the presentation of the detection and filtering algorithm and on numerical simulations that show its effectiveness and its robustness. As already noted, a theoretical analysis of the

algorithm in the case of randomly layered media is presented in [1]. This case is particularly important in understanding imaging in heavy clutter because it can be considered to be the worst case scenario for generating back-scatter [2, 24, 29].

We formulate the imaging problem in Section 2. The setup of the numerical simulations and examples of data traces and migration images in heavy clutter are given in Section 3. The LCT is described briefly in Section 4, and more details are given in Appendix A. We next compute, in Section 5, an estimator of the probability density function for the singular values of the response matrix, and compare it to theoretical results derived using random matrix theory. The purpose of this section is to show that the numerical simulations are in a strongly scattering regime, comparable to that in the experiments in [3]. We do not use random matrix theory for any other purpose in this paper. The proposed detection and filtering algorithm is introduced in Section 6. We present numerical results for several models of clutter in Section 7. We end with a brief summary in Section 8.

2. Formulation of the array imaging problem. We assume that wave propagation is described by the scalar wave equation in an open, often unbounded domain $\Omega \subset \mathbb{R}^d$, for dimensions $d = 2, 3$,

$$\frac{1}{v^2(\vec{\mathbf{x}})}\frac{\partial^2 p(t,\vec{\mathbf{x}})}{\partial t^2} - \Delta p(t,\vec{\mathbf{x}}) = f(t)\delta(\vec{\mathbf{x}} - \vec{\mathbf{x}}_s), \quad \forall\ \vec{\mathbf{x}} \in \Omega, \quad t > 0,$$
$$p(0,\vec{\mathbf{x}}) = 0, \quad \frac{\partial p(0,\vec{\mathbf{x}})}{\partial t} = 0, \quad \forall\ \vec{\mathbf{x}} \in \Omega. \tag{2.1}$$

To fix ideas, we assume that $p(t,\vec{\mathbf{x}})$ is the acoustic pressure field at $\vec{\mathbf{x}}$ and time t, due to an excitation from a point source at $\vec{\mathbf{x}}_s$. The pulse emitted by the source is

$$f(t) = e^{-i\omega_0 t} f_{B_0}(t)$$

with Fourier transform

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t} f_{B_0}(t)dt = \hat{f}_{B_0}(\omega - \omega_0), \tag{2.2}$$

supported in the frequency interval centered at ω_0 with bandwidth B_0.

The wave speed $v(\vec{\mathbf{x}})$ is assumed to have the form

$$\frac{1}{v^2(\vec{\mathbf{x}})} = \frac{1}{c^2(\vec{\mathbf{x}})}\left(1 + \alpha\mu\left(\frac{\mathbf{x}}{\ell}, \frac{z}{\ell_z}\right) + \nu(\vec{\mathbf{x}})\right), \quad \vec{\mathbf{x}} = (\mathbf{x}, z),\ \mathbf{x} = \text{cross-range and } \ z = \text{range}. \tag{2.3}$$

Here $c(\vec{\mathbf{x}})$ is the smooth part of $v(\vec{\mathbf{x}})$, that is used in the computation of travel times for migration imaging, and $\nu(\vec{\mathbf{x}})$ is the reflectivity of the small reflector to be imaged. The clutter consists of numerous inhomogeneities that are unknown and cannot be determined from the array data as we image. There is, therefore, an uncertainty about the inhomogeneities in the medium, which we model mathematically with the random function $\mu(\frac{\mathbf{x}}{\ell}, \frac{z}{\ell_z})$. It is assumed statistically homogeneous, and with mean zero. In (2.3) the parameter α controls the strength of the fluctuations and (ℓ, ℓ_z) are the correlation lengths in the cross-range and range directions, which indicate typical sizes of the inhomogeneities. For isotropic clutter we have $\ell = \ell_z$ while in layered media ℓ is infinite and $\mu = \mu(z/\ell_z)$.

The small reflector modeled by the reflectivity $\nu(\vec{\mathbf{x}})$ is detectable in principle because it represents an anomaly or a defect in the spatial distribution of the inhomogeneities in clutter, modeled by the random

function μ. This might be due to a difference in the size and/or the material properties of the target compared to the clutter. The imaging problem is to estimate the support of $\nu(\vec{\mathbf{x}})$ from the array response matrix $\mathbb{P}(t)$. The elements of $\mathbb{P}(t)$ are given by data traces $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$, which are the solutions $p(t, \vec{\mathbf{x}}_r)$ of the wave equation (2.1) evaluated at receiver locations $\vec{\mathbf{x}}_r$. Measurement noise may be added to these solutions, when we wish to see its effects in imaging, as is done in Section 3.

3. Numerical simulations, data traces and migration images. In our numerical simulations, we generate the response matrix $\mathbb{P}(t)$ by solving the acoustic wave equation (2.1) with the finite element method described in [6, 7]. The computational domain is a rectangular set $\Omega \subset \mathbb{R}^2$, surrounded by a perfectly matched absorbing layer. Our computational setup is shown in Figure 3.1. The medium lies in the half space

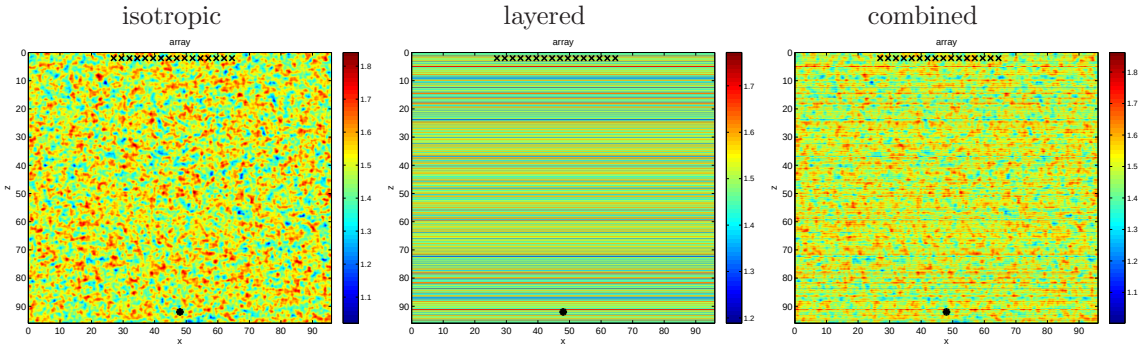

FIG. 3.1. *The setup for the numerical simulations and color plots of the fluctuating wave speed that models clutter. The color code is in kilometers per second. The horizontal axis is cross-range and the vertical is range, measured in units of λ_0. On the left we show isotropic clutter with $\ell = \ell_z = \lambda_0/4$. In the middle, we show a finely layered medium with $\ell_z = \lambda_0/50$, and on the right we show the wave speed obtained by combining the first two.*

$z > 0$ with z increasing downwards. The array is at the surface $z = 0$ and it has $N = 80$ sources and receivers. We choose the simulation parameters so as to be in a regime that is typical in ultrasonic non-destructive testing experiments [4]. The sources in the array emit pulses $f(t)$, whose Fourier transform is flat over the frequency range $B_0 = [1.5 - 4.5]$MHz*. The central wavelength is $\lambda_0 = 0.5$mm and the array sensors are at a distance $\lambda_0/2$ apart.

We show three examples of fluctuating wave speeds of strongly back-scattering media in Figure 3.1. In all cases, the smooth part of the speed is constant $c(\vec{\mathbf{x}}) = c_0 = 1.5$km/s, and the fluctuations are generated with random Fourier series. On the left the fluctuations are isotropic $\mu_i(\vec{\mathbf{x}}) = \mu\left(\frac{x}{\ell}, \frac{z}{\ell}\right)$, with correlation

$$E\{\mu_i(\vec{\mathbf{x}}_1)\mu_i(\vec{\mathbf{x}}_2)\} = R(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2) = \left(1 + \frac{|\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2|}{\ell}\right) e^{-\frac{|\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2|}{\ell}}, \qquad \ell = \lambda_0/4,$$

and standard deviation $\alpha_i = 0.12$. In the middle, the medium is layered, with fluctuations $\mu_l(\vec{\mathbf{x}}) = \mu\left(\frac{z}{\ell_z}\right)$ for $\vec{\mathbf{x}} = (x, z)$, satisfying

$$E\{\mu_l(z_1)\mu_l(z_2)\} = \left(1 + \frac{|z_1 - z_2|}{\ell_z}\right) e^{-\frac{|z_1 - z_2|}{\ell_z}}, \qquad \ell_z = \lambda_0/50,$$

and the standard deviation is $\alpha_l = 0.08$. On the right, the fluctuations are given by the combination

$$\mu_c(\vec{\mathbf{x}}) = \frac{1}{\sqrt{2}}\left(\mu_i(\vec{\mathbf{x}}) + \mu_l(\vec{\mathbf{x}})\right),$$

*In the finely layered medium the pulse is the derivative of a Gaussian with central frequency 3MHz.

with standard deviation $\alpha_c = 0.17$.

The small reflector that we wish to image is located at range $90\lambda_0$ and cross-range $48\lambda_0$. We model it as an acoustic soft scatterer shaped as a disk of radius $\lambda_0/2$, by setting the pressure p to zero on its boundary.

For comparison, we also consider imaging the same reflector in a homogeneous medium (no clutter), using data with additive white noise. We take a very low signal to noise ratio, SNR $= -10$ dB, corresponding to a noise standard deviation that is ten times larger than the power of the signal.

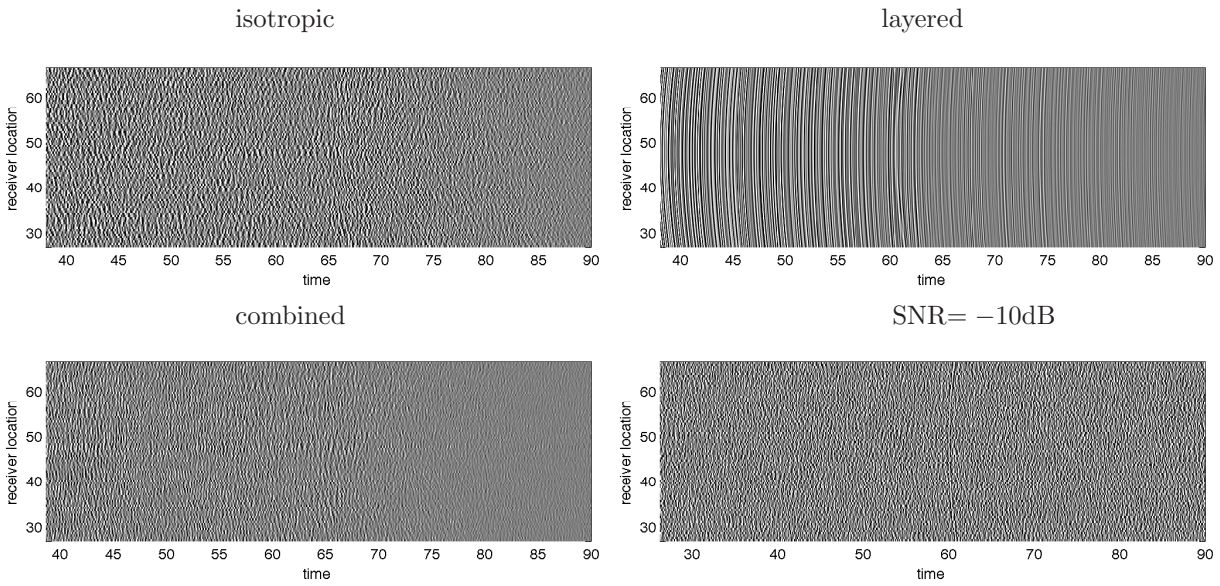

FIG. 3.2. *Data traces recorded at the array for the central source illumination. Lengths are scaled in units of λ_0 and time is measured in μs.*

In Figure 3.2 we show the traces recorded at the array for the central source illumination. They look very noisy and we cannot distinguish any coherent echoes from the reflector that we wish to image, due to the strong multiple scattering in the medium. As we will see in Section 6, the coherent arrivals are still present in the traces and we can enhance them by looking at the data not in the entire time window, but rather in a local time-frequency plane.

Using the Kirchhoff migration function (1.2) we obtain the images shown in Figure 3.3. The images in the isotropic and combined media are noisy and fail to estimate the location of the reflector. The image in the layered medium is better, in the sense that it has a maximum at the true reflector location, although it peaks at layers above it, as well. That the image has a peak at the reflector may be due to the symmetry of the setup. See for example the migration image shown on the top left in Figure 7.10, where the reflector is shifted in cross-range with respect to the center of the array. There is no peak in that image at the location of the reflector. All we can see are layers. The rightmost image in Figure 3.3 is for the additive noise to the array data $P_0(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ in the homogeneous medium. The noise is strong, but the image is excellent. This

is not surprising and is in fact expected from the expression of the Kirchhoff migration function

$$\mathcal{J}^{\text{KM}}(\vec{\mathbf{y}}^s) = \sum_{r=1}^{N_r}\sum_{s=1}^{N_s} P_0\left(\tau(\vec{\mathbf{x}}_r,\vec{\mathbf{y}}^s)+\tau(\vec{\mathbf{x}}_s,\vec{\mathbf{y}}^s),\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s\right) + \sum_{r=1}^{N_r}\sum_{s=1}^{N_s}\eta_{rs}(\tau(\vec{\mathbf{x}}_r,\vec{\mathbf{y}}^s)+\tau(\vec{\mathbf{x}}_s,\vec{\mathbf{y}}^s)). \tag{3.1}$$

Here $\eta_{rs}(t)$ are the uncorrelated entries of the additive white noise matrix, and the travel time is given by $\tau(\vec{\mathbf{x}}_s,\vec{\mathbf{y}}^s) = |\vec{\mathbf{x}}-\vec{\mathbf{y}}|/c_0$. Due to the summation over the sources and receivers, the noise term is effectively zero.

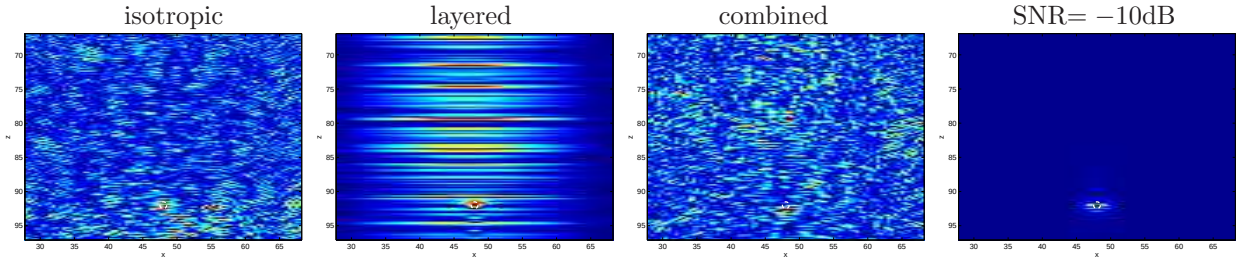

FIG. 3.3. *Migration images obtained using (1.2) with array data shown in Figure 3.2. In the imaging function we use all sources at the array while in Figure 3.2 we only show the traces for the central source. In the images, length is in units of* λ_0.

We note that although the first and the last images in Figure 3.3 are very different, the corresponding traces in Figure 3.2 look quite similar. We cannot see by inspection any difference between the traces coming from a scattering medium and those coming from a homogeneous medium but have a lot of additive measurement noise. There is, however, an important difference: additive measurement noise is white, *i.e.*, uncorrelated and, as we see from the last image at Figure 3.3, it is averaged out by the Kirchhoff migration function. The clutter back scatter is more difficult to deal with, and it cannot be removed by simply summing over the array sensors, as we see from the first three images in Figure 3.3.

Our aim in this paper is to introduce a detection and filtering method that gives recognizable images of the reflector in heavy clutter, despite the very bad images shown in Figure 3.3.

4. The local cosine transform of the response matrix. We will use extensively a local time-frequency decomposition of the data that is based on the discrete Local Cosine Transform (LCT) on a binary tree [20, 25]. We show in Figure 4.1 an illustration of the binary tree used in the time-frequency decomposition of the array data. Level $l = 0$ corresponds to one window that contains the data traces over the entire duration of the recording. At level $l = 1$, the data are segmented in two time windows indexed by $j = 0$ and $j = 1$. At level $l = 2$, the data are segmented in four time windows indexed by $j = 0, 1, 2$ and 3, etc. In Appendix A we give a more detailed description of LCT bases.

At any level $l \geq 0$ in the binary tree we compute the real, symmetric $N \times N$ matrices of coefficients

$$\hat{\mathbb{P}}^l(t_j^l,\omega_n^l) = \left\{\hat{P}^l(t_j^l,\omega_n^l,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)\right\}_{r,s=1,\ldots,N}, \tag{4.1}$$

with

$$\hat{P}^l(t_j^l,\omega_n^l,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s) = \int dt\, P(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)\sqrt{\frac{2}{\Delta t_l}}\chi\left(\frac{t-t_j}{\Delta t_l}\right)\cos[\omega_n^l(t-t_j)]. \tag{4.2}$$

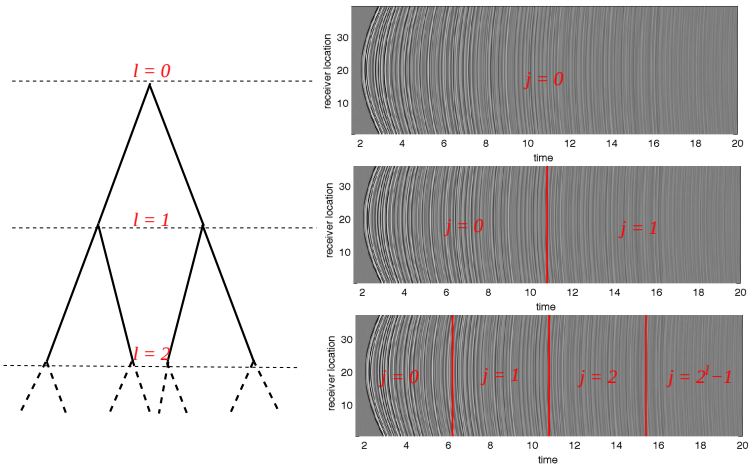

FIG. 4.1. *Illustration of the time windowing segmentations of the array data traces for the layered medium at different tree levels indexed by l. The schematic on the left illustrates the binary tree. On the right we show the segmentation of the data traces in the time windows indexed by j, with $j = 0, 1, \ldots, 2^{l-1}$, at tree levels denoted by l.*

Here t_j^l are the mesh points that define the width and location of the time windows at level l,

$$t_j^l = j\Delta t_l = \frac{jT}{2^l}, \quad j = 0, 1, \ldots, 2^l. \tag{4.3}$$

Each (j, l) node is associated with the subspace

$$\mathcal{F}_j^l = \text{span}\left\{ \sqrt{\frac{2}{\Delta t_l}} \chi\left(\frac{t - t_j^l}{\Delta t_l}\right) \cos[\omega_n^l (t - t_j^l)] \right\}_{n \in \mathbb{N}} \tag{4.4}$$

where

$$\omega_n^l = \frac{\pi\,(n + 1/2)}{\Delta t_l}, \quad n \in \mathbb{N}, \tag{4.5}$$

are the frequencies associated with the decomposition in the smooth windows χ.

The union over $j = 0, 1, \ldots, 2^l - 1$ of the bases associated with $\mathcal{F}_j^l$ gives an orthonormal basis of $L^2[0, T]$. At the next tree level $l + 1$, the spaces $\mathcal{F}_{2j}^{l+1}$ and $\mathcal{F}_{2j+1}^{l+1}$ are orthogonal, and their sum $\mathcal{F}_{2j}^{l+1} \oplus \mathcal{F}_{2j+1}^{l+1}$ is equal to the space $\mathcal{F}_j^l$ at the parent node (j, l) in the tree [25].

The discrete cosine bases used in the numerical simulations are obtained from (4.4) by discretizing the time interval $[0, T]$ with $N_T = 2^m$ time samples. In this case, the frequencies (4.5) sample the *same* bandwidth $(0, \pi N_T/T)$, in steps $\pi/\Delta t_l$, that increase with the tree level l.

5. Singular value distribution of the response matrix. In this section we compute an estimator of the probability density function of the singular values of the numerically simulated array response matrices. The purpose of this estimation is to show that our numerical simulations are calibrated to be in strong scattering regimes that are comparable to those of the experiments in [4].

It is well known that the distribution of the eigenvalues of $N \times N$ real symmetric matrices with independent, identically distributed entries, with mean zero and variance $1/N$, converges as $N \to \infty$ to the Wigner

semi circle law [26, 30] whose density is

$$\rho_{\mathrm{SC}}(\sigma) = \begin{cases} \dfrac{1}{2\pi}\sqrt{4-\sigma^2} & \text{for } -2 \le \sigma \le 2 \\ 0 & \text{otherwise,} \end{cases} \tag{5.1}$$

The limit distribution of the singular values[†], the absolute value of the eigenvalues, is described by the quarter circle law, whose density is,

$$\rho_{\mathrm{QC}}(\sigma) = \begin{cases} \dfrac{1}{\pi}\sqrt{4-\sigma^2} & \text{for } 0 \le \sigma \le 2 \\ 0 & \text{otherwise,} \end{cases} \tag{5.2}$$

It is also well known that, as $N \to \infty$, the extremal eigenvalues of real symmetric matrices with independent, identically distributed entries, with mean zero and variance $1/N$, converge to 2 and -2 [26, 30]. This is why the density (5.2) is supported in the interval $\sigma \in [0, 2]$.

The results reported in [4] suggest that, in strong clutter, and when there is no correlation over the receiver offsets, the distribution of singular values of array response matrices is in good agreement with the quarter circle law.

In Section 5.1 we consider the estimation of the correlation coefficient of the numerically simulated array response matrix. Then, we estimate in Section 5.2 the distribution of the singular values and compare it with the quarter circle law and the results in the experiments in [4]. We note that when the correlation over the receivers is negligible, then the distribution of the singular values is in very good agreement with the quarter circle law. It deviates from it when there is correlation over the receivers.

5.1. Correlation coefficient of the data. Given the LCT of the array response matrix (4.1) at a fixed level l, we compute the normalized correlation over the receiver offsets as follows

$$\Gamma_m^l = \frac{\sum_{j,n,r,s} \hat{P}^l(t_j^l, \omega_n^l, \vec{\mathbf{x}}_{r+m}, \vec{\mathbf{x}}_s)\hat{P}^l(t_j^l, \omega_n^l, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)}{\sum_{j,n,r,s}\left(\hat{P}^l(t_j^l, \omega_n^l, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)\right)^2}. \tag{5.3}$$

Here the sums are over the time window indices $j = 0, 1, \ldots, 2^l - 1$, frequency indices $n = 0, 1, \ldots, N_T/2^l - 1$, and the sources and receivers $s, r = 1, \ldots, N$. The integer m is the distance between the two receivers in units of array pitch (sensor separation) which is $\lambda_0/2$ in our case.

The correlation coefficients Γ_m^l for our numerically simulated response matrices are shown as functions of m in Figure 5.1, for tree level $l = 5$. Similar behavior is observed for other tree levels. As expected, in the layered case the response matrix exhibits strong correlations due to the invariance of the medium in the cross-range direction. In all the other cases, the entries of the response matrix are basically un-correlated, because Γ_m is very small for $m \neq 0$.

[†]The distribution of the singular values is the non-decreasing function of σ, given by the fraction of the N number of singular values of the $N \times N$ matrix, that are less or equal than σ. The limit distribution is absolutely continuous, with density given by (5.2).

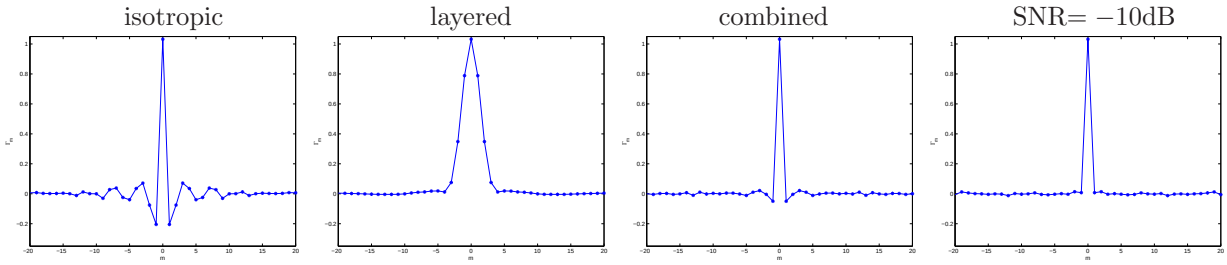

FIG. 5.1. *Normalized correlation coefficients* Γ_m^l *for our data, at tree level* $l = 5$.

5.2. The probability density function of the singular values. For each l and t_j^l, we do the SVD of $\hat{\mathbb{P}}^l(t_j^l, \omega_n^l)$, frequency by frequency and obtain the singular values $\sigma_q^{l,j}(\omega_n^l)$, for $q = 1, \ldots, N$. We then normalize them as

$$\widetilde{\sigma}_q^{l,j}(\omega_n^l) = \frac{\sigma_q^{l,j}(\omega_n^l)}{\sqrt{\frac{1}{N}\sum_{p=1}^{N}(\sigma_p^{l,j}(\omega_n^l))^2}}. \tag{5.4}$$

This normalization is consistent with the hypothesis in the quarter circle law that the variance of the entries of the matrix is $1/N$.

The estimate of the distribution of the singular values is given by the histogram $\mathcal{H}(\sigma)$, computed from (5.4) as follows. The bins of the histogram are the intervals $[i\ w, (i+1)w]$, with w the width of the bin and i non-negative integers. The value of $\mathcal{H}(\sigma)$ is the number of normalized singular values $\widetilde{\sigma}_q^{l,j}(\omega_n^l)$ contained in the same bin as σ, for all 2^l time window indices $j = 0, 1, \ldots, 2^l - 1$, the $N_T/2^l$ frequency indices $n = 0, 1, \ldots, N_T/2^l - 1$, and the N singular value indices $q = 1, \ldots, N$. The probability density function of the singular values is estimated as

$$\rho_{Data}(\sigma) = \frac{\mathcal{H}(\sigma)}{N_T N w}. \tag{5.5}$$

Note that the total number of singular values involved in the computation of the histogram is independent of the tree level l, and equal to $N_T N$.

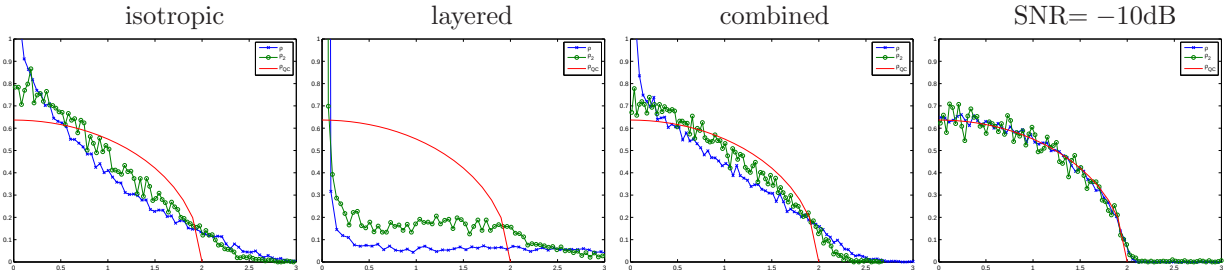

FIG. 5.2. *Estimator* $\rho_{Data}(\sigma)$ *of the probability density function of the singular values and comparison with the quarter circle law* $\rho_{QC}(\sigma)$ *(red line). The blue line corresponds to the estimator computed from the whole array response matrix. The green line is the estimator computed from the reduced matrix, obtained by keeping only the odd indices in the rows and columns.*

The estimator $\rho_{Data}(\sigma)$ for different numerical data sets is plotted in Figure 5.2. In these plots we used one hundred bins to compute the histogram. The width w of each bin is therefore 1% of the maximum value

of $\widetilde{\sigma}_q^{l,j}$. It compares quite well with $\rho_{\mathrm{QC}}(\sigma)$ in the case of the additive white noise. The estimator $\rho_{Data}(\sigma)$ is also relatively close to the quarter circle law in the combined medium, after removing every other entry in the rows and columns of the array response matrix. The estimator is further from $\rho_{\mathrm{QC}}(\sigma)$ in the isotropic medium, and it is very different in the finely layered medium. In the latter case, the medium invariance in the cross-range direction produces a Toeplitz array response matrix $\mathbb{P}(t)$, with persistent correlations along columns/rows. We have done the detailed analysis of the singular value distribution in randomly layered media in [1].

We note that the results regarding the distribution of singular values from the numerical simulations, shown in Figure 5.2, compare well with similar results obtained experimentally and shown in [5, Section 3]

6. The detection and filtering algorithm. In this section we formulate and describe the detection and filtering algorithm that we use for imaging in heavy clutter. We use two basic computational tools: The discrete LCT on a binary tree, as presented in [20, 25], and the singular value decomposition.

• **Inputs** The algorithm takes two inputs:

The $N \times N$ response matrix $\mathbb{P}(t)$, for time $t \in [T_o, T_f]$ sampled on a mesh with $N_T = 2^m$ points, where m is a non negative integer.

The maximum level D in the binary tree used in the discrete LCT. Note that D must be an integer between 0 and m, although we should choose it so that we have enough samples of the signal in each time window.

• **Data processing:**

1. LCT each trace in $\mathbb{P}(t)$. At any level l in the binary tree, with $0 \leq l \leq D$, we obtain the real $N \times N$ matrices of coefficients $\hat{\mathbb{P}}^l(t_j^l, \omega_n^l)$ given by (4.1).

2. For each l and t_j^l, do the SVD of $\hat{\mathbb{P}}^l(t_j^l, \omega_n^l)$, frequency by frequency. Denote by $\sigma_q^{l,j}(\omega_n^l)$ the singular values, for $q = 1, \ldots N$.

3. Choose the frequency band $B \subseteq (0, \pi N_T/T)$ and the number Q of top singular values, to be used in the selection of the time windows, and the data filtering process.

• **Detection:**

4. Use the singular values computed at step 2 to form the matrices

$$\mathbb{S}^{l,j} = \left\{ \tilde{\sigma}_q^{l,j}(\omega_n^l) \right\}_{1 \leq q \leq Q,\, n \in \mathcal{N}^l}, \quad \text{where } \tilde{\sigma}_q^{l,j}(\omega_n^l) = \frac{\sigma_q^{l,j}(\omega_n^l)}{\max\limits_{n'} \sigma_q^{l,j}(\omega_{n'_l})}. \tag{6.1}$$

The columns in $\mathbb{S}^{l,j}$ are the first $Q \leq N$ singular values normalized to maximum value one. The matrices are formed over the set of frequency indices

$$\mathcal{N}^l = \left\{ n = 0, 1, \ldots, N_T/2^l - 1, \;\; \text{s.t. } \omega_n^l \in B \right\}, \tag{6.2}$$

with cardinality $|\mathcal{N}^l|$.

5. Do the SVD of matrices $\mathbb{S}^{l,j}$, for $j = 0, \ldots, 2^l - 1$ and $l = 0, \ldots, D$. Calculate

$$\lambda^{l,j} = \gamma_2^{l,j} / \gamma_1^{l,j}, \tag{6.3}$$

where $\gamma_q^{l,j}$, for $1 \leq q \leq \min\{Q, |\mathcal{N}^l|\}$, are the singular values of $\mathbb{S}^{l,j}$.

6. Select the time window of interest as follows:

For $l = 0 : D$

If $\lambda^{l,j}$ has a maximum above a predetermined threshold, let l_0 be this l and stop.

Next, initialize $j_\star^{l_0} = \arg\max_j \lambda^{l_0,j}$.

For $l = l_0 + 1 : D$

Compute $j_\star^l = \underset{j \in \{2j_\star^{l-1}, 2j_\star^{l-1}+1\}}{\operatorname{argmax}} \lambda^{l,j}$.

If $j_\star^l$ is a maximum of $\lambda^{l,j}$ above a predetermined threshold, for j in the vicinity of $j_\star^l$, continue.

Otherwise, set $l = l - 1$ and stop.

Select the window at $t^l_{j_\star^l}$.

• **Data filtering:**

7. In the selected time window, define the filter $\mathcal{F}^{j_\star^l}$, that sets to zero the LCT coefficients in the windows that have not been selected, at level l,

$$\mathcal{F}^{j_\star^l} \hat{\mathbb{P}}^l(t_j^l, \omega_n^l) = 0 \quad \text{for } j = 0, 1, \ldots, 2^l - 1, \quad j \neq j_\star^l \quad \text{and } n = 0, 1, \ldots N_T/2^l - 1. \tag{6.4}$$

Additional filtering is done by the filter $\mathcal{Q}$, which projects $\mathcal{F}^{j_\star^l} \hat{\mathbb{P}}^l(t_j^l, \omega_n^l)$ on the subspace of low rank matrices with singular vectors corresponding to the "anomalous" top singular values. The projection is done for frequency indices n in the set $\mathcal{N}^l$. All other coefficients are set to zero.

• **Output:**

8. The algorithm outputs the filtered response matrix in the time domain. The filtered traces are given by the inverse LCT of the entries in $\mathcal{Q}\mathcal{F}^{j_\star^l} \hat{\mathbb{P}}^l(t^l_{j_\star}, \omega_n^l)$, at the selected tree level l. We denote the output by $\mathcal{Q}\mathcal{F}^{j_\star^l} \mathbb{P}(t)$.

The detection steps 4-6 are the key part of the algorithm because they are designed to recognize an anomaly in the pattern of the singular values. As we shall see in the next section, the top singular values follow a similar pattern, that is, they are clustered across frequencies in the windows that contain echoes from clutter only. To test for such clustering, we compare, in steps 4 and 5, the normalized singular values across the selected frequency band. We normalize the singular values because their actual magnitude should not enter the pattern recognition. The main idea in step 5 is that because of the clustering of the singular values, across frequencies, the matrices $\mathbb{S}^{l,j}$ have almost rank one in the windows with pure clutter echoes. This is so if we choose Q appropriately, in order to include all of the large, strongly clustered singular values. In the numerical simulations of section 7, we work with $Q = 10$ singular values out of $N = 80$. However, Q will depend on the data set and it must be chosen after a preliminary inspection of the behavior of the singular values.

We expect an anomaly in the pattern of the singular values in the windows that contain detectable coherent echoes. When this occurs, the column space of the matrices $\mathbb{S}^{l,j}$ has a significant second dimension and therefore the second singular value of $\mathbb{S}^{l,j}$ is large. This explains the window selection criterion in step 6. The thresholds that are needed in this step must be determined by a preliminary exploration of the data. Note that the selection begins at the first level of the tree where a clear maximum of $\lambda^{l,j}$ occurs, and that the windows are progressively refined until we reach the highest level where the selection is unambiguous.

In general, we should work with a band B of low frequencies, where the clutter effects are weaker. In the numerical simulations of section 7, we take B to be the first half of the frequency band, that is, we choose to work with $B = [1.5, 3]$MHz instead of the whole bandwidth for our signal, which is $[1.5, 4.5]$MHz. This choice is expected to be dependent on the data set, and it should be made after a preliminary inspection of the behavior of the singular values.

The filtering step 7 in the algorithm removes the contribution of all the windows except the selected ones. It also projects the matrix of coefficients, in the selected windows, on the subspace of low rank matrices with singular vectors corresponding to the "anomalous" singular values. By "anomalous" singular values we mean those that have a different behavior than the other top singular values. These singular values are expected to be due to the coherent echoes, per our selection criterion in steps 5 - 6. The projection is done with the truncated SVD, as follows. Let

$$\mathcal{F}^{j^l_\star}\hat{\mathbb{P}}^l(t^l_{j^l_\star}, \omega^l_n) = \mathbb{U}(t^l_{j^l_\star}, \omega^l_n)\Sigma(t^l_{j^l_\star}, \omega^l_n)\mathbb{U}^T(t^l_{j^l_\star}, \omega^l_n), \quad n \in \mathcal{N}^l, \tag{6.5}$$

be the SVD of the matrix of coefficients in the selected window. The singular vectors are in the orthogonal matrices $\mathbb{U}(t^l_{j^l_\star}, \omega^l_n)$, and $\Sigma(t^l_{j^l_\star}, \omega^l_n)$ is the diagonal matrix of singular values. The filtered matrix $\mathcal{Q}\mathcal{F}^{j^l_\star}\hat{\mathbb{P}}^l(t_{j^l_\star}, \omega^l_n)$ is obtained by setting to zero in $\Sigma(t^l_{j^l_\star}, \omega^l_n)$ the entries that are not "anomalous" singular values.

An important feature of the detection and filtering algorithm is that it applies to general clutter. The calibration that determines the number of top singular values used in the detection, the frequency sub-band, and the detection threshold is influenced by the type of clutter, but we do not need to know it in advance. The calibration can be carried out directly from the data.

7. Numerical simulations with the detection and filtering algorithm. Our implementation of the discrete LCT is using the Wavelab 850 MATLAB package [22], with windows option "Sine" and a binary tree of depth $D = 5$. The traces in $\mathbb{P}(t)$ are discretized on a uniform time mesh with $N_T = 2^{11}$ points, in the time interval $[T_o, T_f] = [33, 90]\mu$s.

7.1. Results for the combined medium. Let us begin at the root of the tree, where $l = 0$. If the small reflector that we wish to image was in a homogeneous medium or in weak clutter, we would observe one or two large singular values of the matrices $\hat{\mathbb{P}}^0(t^0_0, \omega^0_n)$ [1, Theorem 6.1]. However, the combined medium is strongly back-scattering and, as shown in Figure 7.1, we cannot observe anything useful at root level $l = 0$. We have tight clustering of the top ten singular values across the frequencies. Because the energy of the clutter back-scatter is dominant in the large time interval $[T_o, T_f]$, it is difficult to detect the presence of the small reflector. We should look at higher levels in the tree, to get more information.

We plot in Figure 7.2 the singular values $\sigma^{4,j}_q(\omega^4_n)$, in six windows at level $l = 4$, and for $q = 1, \ldots, 10$.

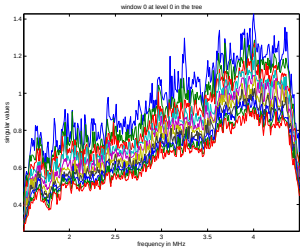

FIG. 7.1. *Combined medium. Singular values $\sigma_q^{0,0}(\omega_n^0)$, at the root level $l = 0$ in the tree. We plot the first $Q = 10$ of them as a function of frequency. The abscissa is frequency in MHz.*

Starting with the window indexed by $j = 5$, we note that the singular values remain tightly clustered, uniformly in the bandwidth, until we reach the index $j = 7$. This is the window that contains the coherent echoes from the reflector, and it is distinguished from the others by an anomalous singular value (arguably two) at the lower frequencies. The anomaly is also present in the next window, which contains the interactions between the reflector and the random medium, and then it disappears in the last two windows. This behavior persists for all three cluttered media considered in our simulations. It has also been analyzed and explained theoretically in [1], for randomly layered media.

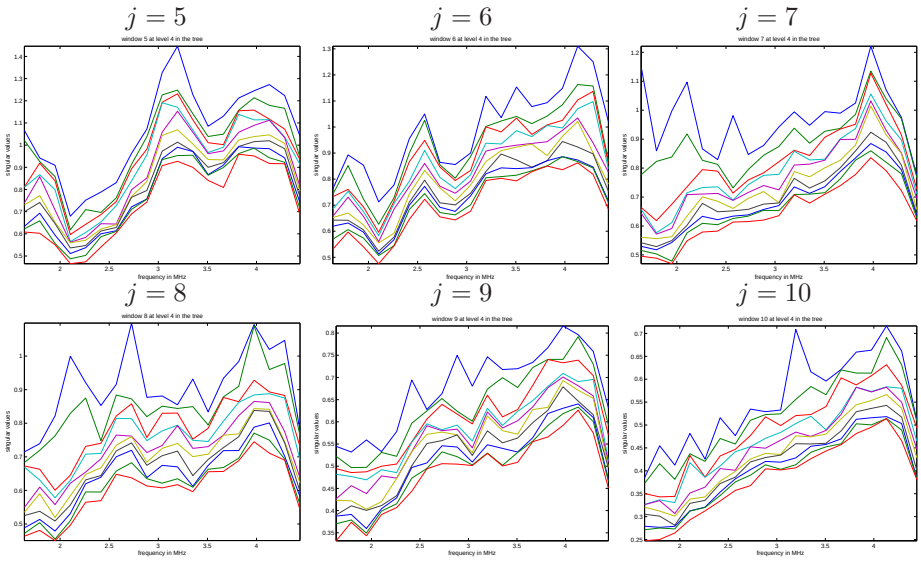

FIG. 7.2. *Combined medium. The singular values $\sigma_q^{4,j}(\omega_n^4)$ as functions of frequency, for $q = 1, \ldots, 10$. Starting from the top left corner, we take the time windows indexed by $j = 5, 6, \ldots 10$. Each eigenvalue is plotted with a different color.*

In Figure 7.3 we illustrate the behavior of the singular values as we progress upward in the tree, from one level to the next. The bottom plots show that, in the time windows that contain only the medium back-scatter, the singular values remain clustered uniformly over the bandwidth. Moreover, for levels $l = 4$ and 5, the top plots in Figure 7.3 show that there is a persistent anomalous behavior of one singular value (arguably two), at the lower frequencies, in the windows that contain the echoes from the small reflector that we wish to image. It is because of this anomaly in the pattern of the singular values that we can detect the time windows of interest, with steps 5-6 in the algorithm.

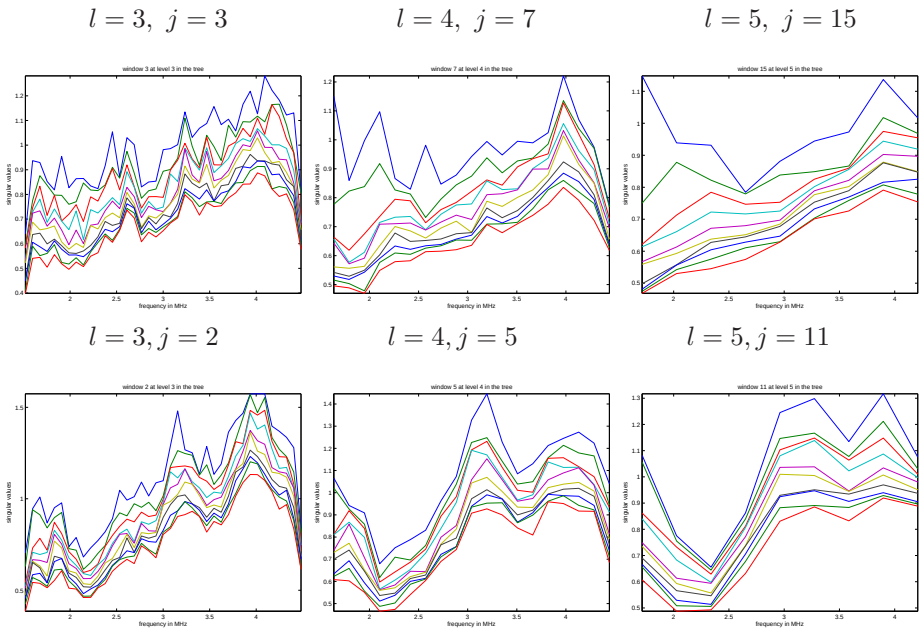

FIG. 7.3. *Combined medium. Top: The top* 10 *singular values in the windows that contain the coherent echoes from the small reflector. Left to right: The 3-rd window at level* $l = 3$*; its child, the 7-th window at level* $l = 4$*; its child, the* 15*-th window at level* $l = 5$*. Bottom: The top singular values in windows that contain pure clutter echoes. Left to right: The 2-nd window at level* $l = 3$*; its child, the* 5*-th window at level* $l = 4$*; its child, the 11-th window at level* $l = 5$*.*

After inspecting the plots in Figure 7.3, we conclude that we should have made the window selection by using only the lower half of the frequency band. This is what we call the bandwidth B in step 3 of the algorithm. We plot in Figure 7.4 the selection criterion $\lambda^{l,j}$, as a function of the window index j, for levels $l = 3, \ldots, 5$. Note that it is difficult to detect, by inspection, an anomalous behaviour of a singular value at the lower tree level $l = 3$, in the time window indexed by $j = 3$, which is where the coherent echoes lie. The algorithm detects this window, just barely, because $\lambda^{3,j}$ has a slight peak at $j = 3$. The detection is unambiguous at the higher levels $l = 4$ and 5, as seen from the middle and right pictures in Figure 7.4.

Following step 6 of the algorithm, we select at level 3 the time window indexed by $j_\star^3 = 3$. Moving upward to level $l = 4$, step 6 of the algorithm says to choose $j_\star^4$ as

$$j_\star^4 = \underset{j \in \{6,7\}}{\operatorname{argmax}} \lambda^{4,j}, \quad \text{where} \quad 2j_\star^3 = 6.$$

We see from Figure 7.4 that $j_\star^4 = 7$ and since it is a clear local maximum of $\lambda^{4,j}$, we accept it and move to level $l = 5$. Proceeding as above,

$$j_\star^5 = \underset{j \in \{14,15\}}{\operatorname{argmax}} \lambda^{5,j}, \quad \text{where} \quad 2j_\star^4 = 14.$$

Figure 7.4 shows that $j_\star^5 = 15$ and we accept it, as it is a clear local maximum of $\lambda^{5,j}$. The detection stops here because we have reached the highest level in the tree.

The next step in the algorithm is the data filtering done at tree level $l = 5$, where the selected time window lies. We set to zero the coefficients $\hat{\mathbb{P}}^5(t_j^5, \omega_n^5)$ for $j \in \{0, \ldots, 2^5 - 1\} \setminus \{15\}$, and we project the matrix $\hat{\mathbb{P}}^5(t_{15}^5, \omega_n^5)$ for $\omega_n \in B$, on the subspace of rank one matrices with singular vectors corresponding to

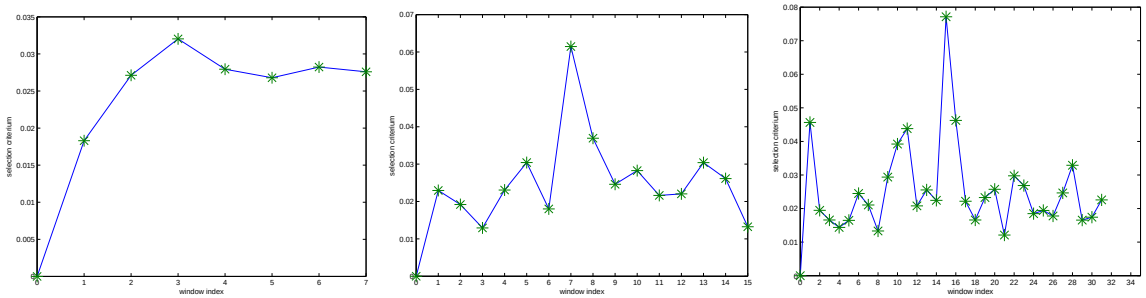

FIG. 7.4. *Combined medium. Plot of* $\lambda^{l,j}$ *as a function of the window index* $j = 0, \ldots, 2^l - 1$ *for tree levels* $l = 3, 4, 5$.

the anomalous top singular value. The resulting traces obtained with the inverse LCT, after the filtering, are shown in Figure 7.5.

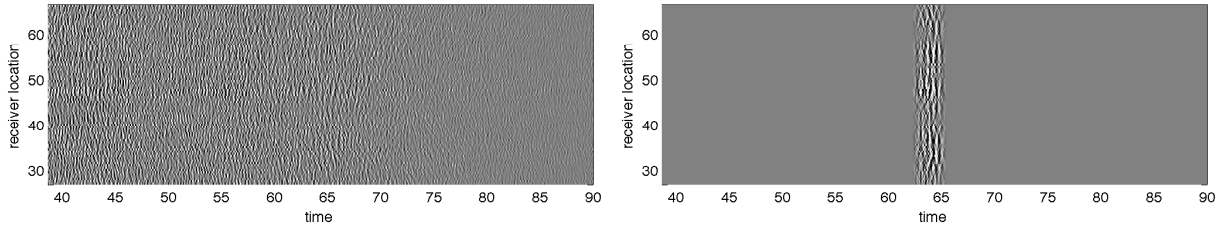

FIG. 7.5. *Combined medium. Original and filtered traces produced by our algorithm. The abscissa is time in* μs.

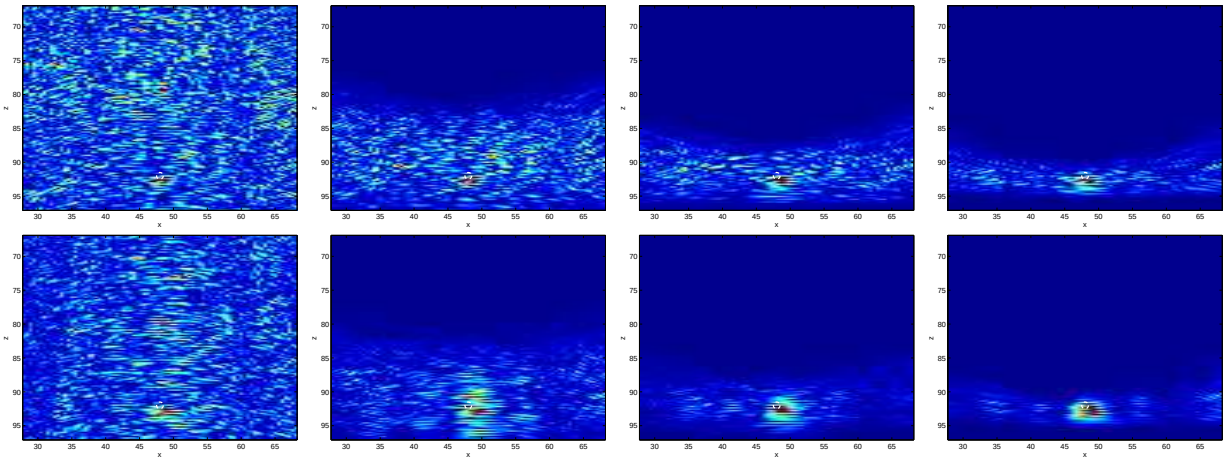

FIG. 7.6. *Combined medium. Top: Migration images obtained with the data reconstructed from the filtered by* $\mathcal{F}^{j^l_\star}$ *LCT coefficients in the windows selected in the iteration at step 6 of the algorithm. From left to right:* $j^3_\star = 3$, $j^4_\star = 7$ *and* $j^5_\star = 15$. *Bottom: The same plots for the data reconstructed after applying also the projection filter* $\mathcal{Q}$. *The abscissa is cross-range in* λ_0 *and the ordinate is depth in* λ_0. *The small scatterer is indicated with a white circle.*

The migration images with the traces reconstructed from the selected windows $j^l_\star$ at levels $l = 3, \ldots, 5$ are in Figure 7.6. We show in the bottom row the images obtained with the filtered data, per step 7 in the algorithm. The images in the top row are obtained by skipping the subspace projection $\mathcal{Q}$ at step 7. The results in Figure 7.6 illustrate the effect of the two filtering operators $\mathcal{F}^{j^l_\star}$ and $\mathcal{Q}$ on the images. As l

increases, the filter $\mathcal{F}^{j_\star^l}$ narrows the image in range around the correct location, while the filter $\mathcal{Q}$ focuses the image in cross-range. Although we observe a beneficial effect of $\mathcal{Q}$ at all tree levels, it is the combination of the two filters $\mathcal{F}^{j_\star^5}$ and $\mathcal{Q}$ that produces the best image, shown at the bottom right in Figure 7.6. This image is comparable with the one in Figure 7.7, obtained in the ideal case of the reflector in an homogeneous medium, at infinite SNR.

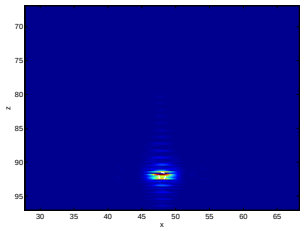

FIG. 7.7. *Migration image for the reflector embedded in a homogeneous medium. In the migration we use the same bandwidth B as the one used in our filtering algorithm. The abscissa is cross-range in λ_0 and the ordinate is depth in λ_0. The small scatterer is indicated with a white circle.*

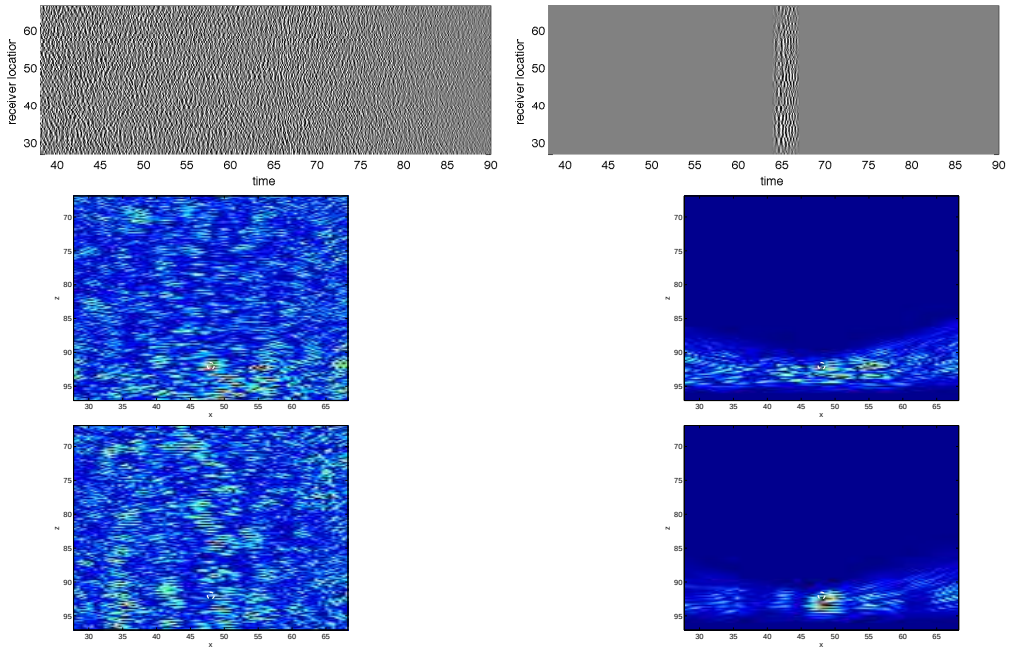

FIG. 7.8. *Isotropic medium. Top: Original and filtered traces produced by our algorithm. The abscissa is time in μs. Bottom: Migration images with filtered data per step 7 of the algorithm, applied to windows $j_\star^0 = 0$ on the left and $j_\star^5 = 16$ on the right. Middle: The same plots, except that the filtered data is obtained by skipping the subspace projection $\mathcal{Q}$ at step 7. The abscissa is cross-range in λ_0 and the ordinate is depth in λ_0. The small scatterer is indicated with a white circle.*

7.2. Results for isotropic and layered clutter. We show in Figure 7.8 the results in isotropic clutter. The array data traces are on the top left, and on the top right we show the filtered traces produced by our algorithm. The selected time window is indexed by $j_\star^5 = 16$, at tree level $l = 5$. The middle left migration image is obtained with the raw traces. It is noisy and has several maxima that are not near the reflector location. The middle right image is obtained by applying the subspace projection filter $\mathcal{Q}$ in the

entire time window, at the root of the tree. We note that this projection does not improve the image at all. The migration image with the filtered data obtained with step 7 of our algorithm, in the selected time window at t_{16}^5, is shown at the bottom right in Figure 7.8. The image obtained from the partially filtered data by skipping the projection $\mathcal{Q}$ is on the right in the middle row. We observe, as before, that it is the combination of the two filters $\mathcal{F}^{j_\star^l}$ and $\mathcal{Q}$ that produces the best images.

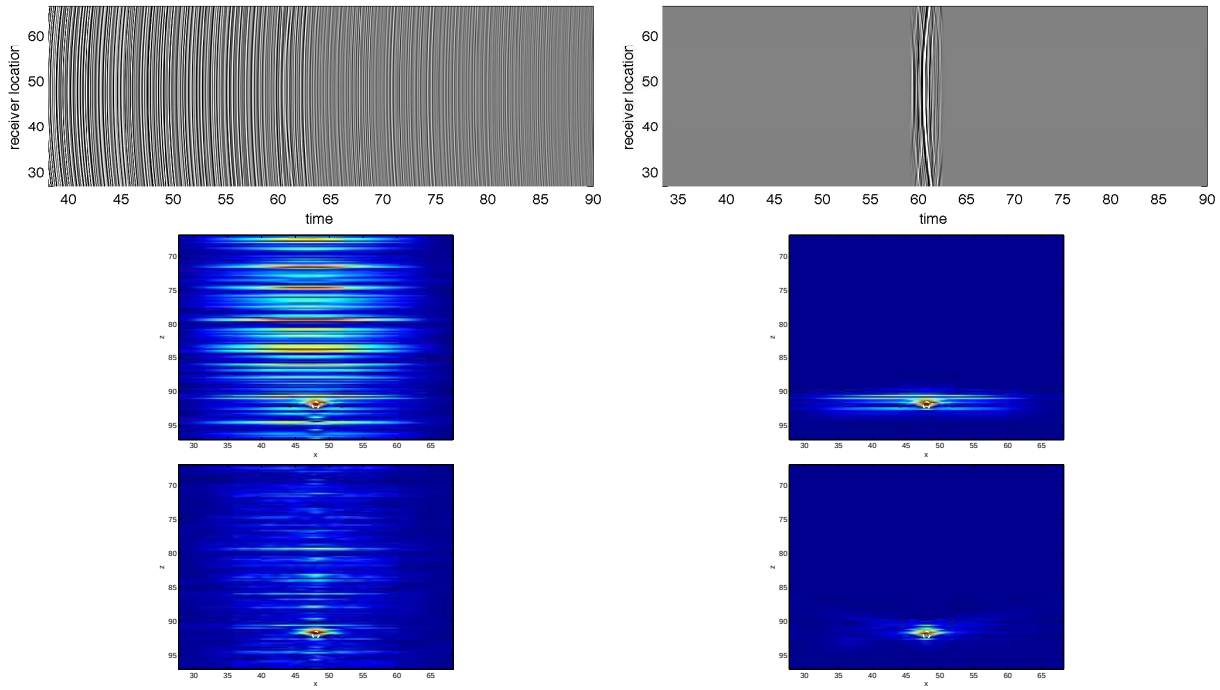

FIG. 7.9. *Layered medium. Top: Original and filtered traces produced by our algorithm. The abscissa is time in μs. Bottom: Migration images with filtered data per step 7 of the algorithm, applied to windows $j_\star^0 = 0$ on the left and $j_\star^5 = 15$ on the right. Middle: The same plots, except that the filtered data is obtained by skipping the subspace projection $\mathcal{Q}$ at step 7. The abscissa is cross-range in λ_0 and the ordinate is depth in λ_0. The small scatterer is indicated with a white circle.*

The results for the layered medium are shown in Figure 7.9. The layout is the same as in Figure 7.8. What is surprising in this figure is the beneficial effect of the subspace projection filter $\mathcal{Q}$ in the entire time window, at the root of the tree, as seen from the bottom left image. This is due to the symmetry of the setup, and does not persist when we displace the reflector in cross-range, as illustrated in Figure 7.10. It is the latter figure that illustrates, as before, the beneficial effect of the filtering of the data on the image.

8. Summary. We have introduced an algorithm for the detection of a reflector in a strongly scattering medium and for filtering of the clutter back-scatter in array imaging. The algorithm is based on a time-frequency representation of the array response matrix obtained with the local cosine transform, followed by a singular value decomposition. Detection is done by identifying anomalies in the group of top singular values, across frequencies and in different time windows. The anomalies are due to the reflector and are identified with another singular value decomposition. The detection is adaptive because the time windows that contain the primary echoes from the reflector are not determined in advance. Their location and width is identified by searching through the time-frequency binary tree of the local cosine transform. Filtering

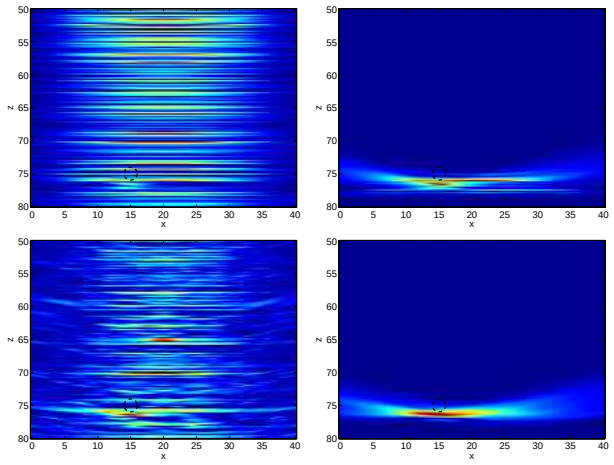

FIG. 7.10. *Layered medium and reflector at different location, indicated with a black circle. Bottom: Migration images with filtered data per step 7 of the algorithm, applied to windows $j_\star^0 = 0$ on the left and $j_\star^5 = 15$ on the right. Top: The same plots, except that the filtered data is obtained by skipping the subspace projection $\mathcal{Q}$ at step 7. The abscissa is cross-range in λ_0 and the ordinate is depth in λ_0.*

retains only information in the time windows that have been selected in the detection stage. We have shown with extensive numerical simulations that this algorithm works well in heavy clutter that is calibrated using random matrix theory as a guide. The algorithm works in general clutter, but has been analyzed theoretically only for the case of randomly layered media in [1].

Acknowledgments. The work of L. Borcea was partially supported by the Office of Naval Research, grant N00014-09-1-0290, and by the National Science Foundation, grant DMS-0907746. The work of G. Papanicolaou was partially supported by AFOSR grant FA9550-08-1-0089. The work of C. Tsogka was partially supported by the European Research Council Stanting Grant, GA 239959.

Appendix A. The local cosine transform. We summarize from [25] a few facts on the discrete LCT transform on binary trees, which are relevant to the algorithm in section 6 and the numerical implementation used in section 7.

The discrete local cosine bases on binary trees segment the time domain $[T_o, T_f]$ in dyadic intervals. Explicitly, let us define at any level $l \geq 0$ in the tree, the time segmentation

$$t_j^l = \left(\alpha_j^l + 1/2\right) dt + T_o, \tag{A.1}$$

where α_j^l are the half integer values [25, section 8.5.2]

$$\alpha_j^l = jN_T 2^{-l} - 1/2, \quad j = 0, 1, \ldots, 2^l, \tag{A.2}$$

and $N_T = 2^m$, with $m > l$, is the number of time samples in $[T_o, T_f]$, so that $dt = T/N_T$, with $T = T_f - T_o$. We obtain 2^l time intervals of length $\Delta t_l = T/2^l$, covered with the overlapping time windows[‡]

$$\psi_j^l(t) = g_j^l\left(\frac{t - T_o}{dt}\right). \tag{A.3}$$

[‡]In the paper we let $\chi\left(\frac{t-t_j^l}{\Delta t_l}\right) = \psi_j^l(t)$.

Here g_j^l are functions supported in $[\alpha_j^l - \eta, \alpha_{j+1}^l + \eta]$, given by

$$g_j^l(s) = \begin{cases} \beta\left(\frac{s-\alpha_j^l}{\eta}\right) & \text{if } \ s \in [\alpha_j^l - \eta, \alpha_j^l + \eta] \\ 1 & \text{if } \ s \in [\alpha_j^l + \eta, \alpha_{j+1}^l - \eta] \\ \beta\left(\frac{\alpha_{j+1}^l - s}{\eta}\right) & \text{if } \ s \in [\alpha_{j+1}^l - \eta, \alpha_{j+1}^l + \eta] \\ 0 & \text{otherwise.} \end{cases} \tag{A.4}$$

Since we must restrict the support of the time windows to $[T_o, T_f]$, we modify the definition (A.3) for indices $j = 0$ and $j = 2^l - 1$ as follows,

$$\begin{aligned} \psi_0^l(t) &= 1 \quad \text{if } t \in [T_o, T_o + \eta dt] \\ \psi_{2^l-1}^l(t) &= 1 \quad \text{if } t \in [T_f - \eta dt, T_f]. \end{aligned} \tag{A.5}$$

The smoothness of the windows depends on the regularity of the function β which enters its definition (A.4), and the overlap η. To have orthonormal cosine bases, the function β must satisfy the identity

$$\beta^2(s) + \beta^2(-s) = 1, \quad s \in [-1, 1], \tag{A.6}$$

and the boundary conditions

$$\begin{aligned} \beta(s) &= 0 \quad \text{if } s < -1 \\ \beta(s) &= 1 \quad \text{if } s > 1. \end{aligned} \tag{A.7}$$

In our numerical implementation we use the option "Sine" in the Wavelab 850 Matlab package [22], which corresponds to

$$\beta(s) = \sin\left[\frac{\pi(1+s)}{4}\right]. \tag{A.8}$$

This gives continuous, but not differentiable windows $\psi_j^l(t)$. Smoother windows are given by the iterated sine functions, as shown in [25, page 362].

The overlap η is uniform in l and it is defined by equation

$$2\eta = 2^{-D}, \tag{A.9}$$

in terms of the deepest level D in the tree [25, page 370]. Note that a large η means smoother windows, at the expense of time resolution, due to the bound constraint $l \leq D = -2\log_2(2\eta)$.

REFERENCES

[1] R. Alonso, L. Borcea, G. Papanicolaou, and C. Tsogka, *Detection and imaging in strongly backscattering randomly layered media*. submitted to Inverse Problems.

[2] M. Asch, W. Kohler, G. Papanicolaou, M. Postel, and B. White, *Frequency content of randomly scattered signals*, SIAM Review, 33 (1991), pp. 519–625.

[3] A. Aubry and A. Derode, *Detection and imaging in a random medium: A matrix method to overcome multiple scattering and aberration*, J.APPL.PHYS., 106 (2009), p. 044903.

[4] ———, *Random matrix theory applied to acoustic backscattering and imaging in complex media*, Phys. Rev. Lett., 102 (2009), p. 084301.

[5] ———, *Singular value distribution of the propagation matrix in random scattering media*, Waves in Random and Complex Media, 20 (2010), pp. 333–363.
[6] E. Bécache, P. Joly, and C. Tsogka, *Etude d'un nouvel élément fini mixte permettant la condensation de masse*, C. R. Acad. Sci. Paris Sér. I Math., 324 (1997), pp. 1281–1286.
[7] ———, *An analysis of new mixed finite elements for the approximation of wave propagation problems*, SIAM J. Numer. Anal., 37 (2000), pp. 1053–1084.
[8] Biondo Biondi, *3D Seismic Imaging*, no. 14 in Investigations in Geophysics, Society of Exploration Geophysicists, Tulsa, 2006.
[9] N. Bleistein, J.K. Cohen, and J.W. Stockwell Jr., *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer, New York, 2001.
[10] L. Borcea, F. González del Cueto, G. Papanicolaou, and C. Tsogka, *Filtering deterministic layer effects for imaging*, SIAM Multiscale Modeling and Simulations, 7 (2009), pp. 1267–1301.
[11] ———, *Filtering random layering effects in imaging*, SIAM Multiscal Modeling and Simulations, 8 (2010), pp. 751–781.
[12] L. Borcea, G. Papanicolaou, and C. Tsogka, *Interferometric array imaging in clutter*, Inverse Problems, 21 (2005), pp. 1419–1460.
[13] ———, *Adaptive interferometric imaging in clutter and optimal illumination*, Inverse Problems, 22 (2006), pp. 1405–1436.
[14] ———, *Coherent interferometric imaging*, Geophysics, 71 (2006), pp. SI165–SI175.
[15] ———, *Coherent interferometry in finely layered random media*, SIAM J. on Multiscale Model. Simul., 5 (2006), pp. 62–83.
[16] ———, *Asymptotics for the space-time Wigner transform with applications to imaging*, in Stochastic Differential Equations: Theory and Applications. Volume in Honor of Professor Boris L. Rozovskii, P. H. Baxendale and S. V. Lototsky, eds., vol. 2 of Interdisciplinary Mathematical Sciences, World Scientific, 2007, pp. 91–112.
[17] ———, *Optimal illumination and waveform design for imaging in random media*, JASA, 122 (2007), pp. 3507–3518.
[18] B. Borden, *Mathematical problems in radar inverse scattering*, Inverse Problems, 19 (2002), pp. R1–R28.
[19] J. F. Claerbout, *Fundamentals of geophysical data processing : with applications to petroleum prospecting*, CA : Blackwell Scientific Publications, Palo Alto, 1985.
[20] R. R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier a fenêtre*, C.R. Acad. Sci., (1991), pp. 259–261.
[21] J.C. Curlander and R.N. McDonough, *Synthetic Aperture Radar*, Wiley, New York, 1991.
[22] D. Donoho, A. Maleki, and M. Shahram, *Wavelab 850*. `http://www-stat.stanford.edu/~wavelab/`.
[23] C. Elachi, *Spaceborne Radar Remote Sensing: Applications and Techniques*, IEEE, New York, 1987.
[24] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna, *Wave Propagation and Time Reversal in Randomly Layered Media*, Springer, April 2007.
[25] S. Mallat, *A wavelet tour of signal processing*, Academic Press, second ed., 1999.
[26] V. Marcenko and L. Pastur, *Distributions of eigenvalues for some sets of random matrices*, Math. USSR-Sbornik, 1 (1967), pp. 457–483.
[27] C. Prada and M. Fink, *Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media*, Wave Motion, 20 (1994), pp. 151–163.
[28] L. Ryzhik, G. Papanicolaou, and J. B. Keller, *Transport equations for elastic and other waves in random media*, Wave Motion, 24 (1996), pp. 327–370.
[29] P. Sheng, *Introduction to wave scattering, localization and mesoscopic phenomena*, Academic Press, 1995.
[30] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*, Commun. Inf. Theory, 1 (2004), pp. 1–182.
[31] M. C. W. van Rossum and Th. M. Nieuwenhuizen, *Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion*, Reviews of Modern Physics, 71 (1999), pp. 313–371.