Adaptive interferometric imaging in clutter and optimal illumination

Liliana Borcea† and George Papanicolaou‡ and Chrysoula Tsogka*

† Computational and Applied Mathematics, MS 134, Rice University, 6100 Main Street, Houston, TX 77005-1892.(`borcea@caam.rice.edu`)
‡ Department of Mathematics, Stanford University, Stanford, CA 94305. (`papanico@math.stanford.edu`)
* Department of Mathematics, University of Chicago, Chicago, IL 60637. (`tsogka@math.uchicago.edu`)

Abstract. A frequently used broadband array imaging method is Kirchhoff or travel time migration. In smooth and known media Kirchhoff migration works quite well, with range resolution proportional to the reciprocal of the bandwidth and cross range resolution that is proportional to the reciprocal of the array size. In a randomly inhomogeneous medium, Kirchhoff migration is unreliable because the images depend on the detailed scattering properties of the random medium that are not known. In [17] we introduced an imaging functional that does not depend on the detailed properties of the random medium, that is, it is statistically stable. This is the Coherent Interferometric (CINT) imaging functional, which can be viewed as a smoothed version of Kirchhoff migration. Smoothing increases the statistical stability of the image but causes blurring. In this paper we introduce an adaptive version of CINT in which there is an optimal trade-off between statistical stability and blurring. We also introduce optimal illumination schemes for achieving the best possible resolution of the images obtained with CINT.

1. Introduction

In broadband array imaging we want to determine small or distributed reflectors by sending probing signals from one or more sources at the array and recording the scattered echoes. A typical setup is shown schematically in Figure 1, where the reflectors are in the domain $\mathcal{S}$.

The source at $\mathbf{x}_s$ emits a spherical wave convolved with the signal

$$f(t) = e^{-i\omega_o t} f_B(t) \tag{1.1}$$

whose Fourier transform

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} e^{i(\omega-\omega_o)t} f_B(t) dt = \widehat{f}_B(\omega - \omega_o), \tag{1.2}$$
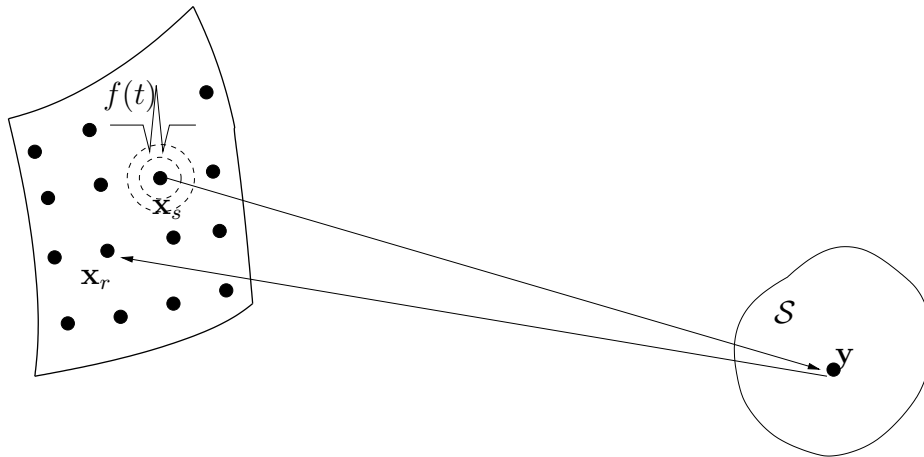
Figure 1. Schematic for imaging distributed reflectors in $\mathcal{S}$.

is supported in the frequency interval centered at ω_o with bandwidth B. The time traces of the scattered echoes $P(\mathbf{x}_r, \mathbf{x}_s, t)$ are recorded at the array receiver locations $\mathbf{x}_r$, for $r = 1, \ldots, N_r$, over some time window $t \in [t_m, t_M]$. Full array data are obtained by sending sequentially probing signals from N_s sources and recording in each case the traces $P(\mathbf{x}_r, \mathbf{x}_s, t)$, $s = 1, ..., N_s$, $r = 1, ..., N_r$. The imaging problem is to estimate from the recorded traces the support of the reflectors in $\mathcal{S}$.

When the background medium is known and is smooth or piece-wise smooth then the imaging of the reflectors can be done with Kirchhoff migration [31, 60, 30, 8, 9, 68, 49, 64, 10], as we recall in section 2. If the medium is not known then we may be able to estimate it approximately during the image formation process with a background velocity analysis [24, 66, 35]. However, this velocity estimation can only capture the smooth part of the background. It cannot get the small scale inhomogeneities that are naturally present in many important applications in geophysics, non-destructive evaluation, foliage or ground penetrating radar, etc. Since we do not know these small scale inhomogeneities we view them as clutter and we model them with spatial random processes.

Imaging with Kirchhoff migration does not work well in cluttered media when there is significant multipathing of the waves by the inhomogeneities. This is because there is considerable delay spread or coda in the recorded traces at the array, which creates noisy-looking and speckled images that are difficult to interpret. Moreover, the images change unpredictably for different realizations of the cluttered background having the same statistical properties. To make migration work in clutter, we need an efficient method for compressing the delay spread in the traces. We introduced such a method in [17, 18], called coherent interferometry (CINT), where the delay spread is reduced by cross correlating the traces over appropriate space-time windows.

In this paper we introduce an adaptive algorithm for selecting the size of the space-

time windows with a criterion that is based on both the stability and the quality of the image. We show that in cluttered media there is a trade-off between statistical stability and resolution of the image and that the adaptive algorithm that we propose selects the space-time windows so as to achieve an optimal compromise between the two.

We also consider the question of choosing optimally the waveform sent by the array in order to construct the best possible image of the reflectors. Optimal waveforms for maximizing the received signal power at the array at each frequency are studied in [53, 47, 43, 36]. This is done with the singular value decomposition of the response matrix for each frequency. Time reversal or Kirchhoff migration images constructed with this optimal illumination tend to enhance the image near the strongest reflector or near the region of highest reflectivity. Maximization of the received power in the time domain can be done with iterative time reversal [54, 27, 28]. However, the optimal waveforms for maximum received power in the time domain are narrow-band signals that are centered at a frequency which typically depends on the size of the reflectors, such as a resonant frequency. They produce strong echoes at the array, which is what is needed for detection, but they are bad for imaging because the lack of bandwidth gives poor range resolution and poor statistical stability in clutter [11, 16, 15, 20, 21]. A way to mitigate the limitations of narrow-band optimal waveforms is considered in [48].

In this paper we introduce a different method for determining optimal illumination and pulse waveforms. We use an optimization criterion that is based on the quality of the image. It is similar to what we do in adaptive coherent interferometry because it uses the coherent interferometric functional to determine the optimal illumination in essentially the same way as for smoothing clutter effects.

The paper is organized as follows: In section 2 we review the simplest form of imaging with Kirchhoff migration and present some numerical results that illustrate its lack of statistical stability in clutter. In Section 3 we consider coherent interferometric imaging and explain in section 4 how it can be viewed as a statistically smoothed version of Kirchhoff migration. The adaptive coherent interferometric approach is introduced in section 5 and its performance is illustrated with numerical simulations. In Section 6 we address the optimal illumination problem and present some numerical simulations. We end the paper with a summary and conclusions, in section 7.

2. Kirchhoff migration imaging

Imaging in known and piecewise smoothly varying environments can be done efficiently with Kirchhoff migration [31, 60, 30, 8, 9, 68, 49, 64, 10], which in its simplest form is the transformation of the array data P to an image by the functional $\mathcal{I}^{\mathrm{KM}}$

$$\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S) = \sum_{r=1}^{N_r} \sum_{s=1}^{N_s} P\left(\mathbf{x}_r, \mathbf{x}_s, \tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right). \tag{2.3}$$

Equivalently, in the frequency domain we have

$$\mathcal{I}^{\text{KM}}(\mathbf{y}^S) = \sum_{r=1}^{N_r} \sum_{s=1}^{N_s} \int_{-\infty}^{\infty} d\omega \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \exp\left(-i\omega\left[\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]\right), \quad (2.4)$$

with the Fourier transform $\widehat{P}$ defined as in (1.2). The data traces are "migrated" to a search point $\mathbf{y}^S$, which sweeps a search domain $\mathcal{D}$ that includes the reflectors. The migration is done with the travel times $\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)$ of the waves from the source at $\mathbf{x}_s$ to the search point $\mathbf{y}^S$ and then back to the receiver at $\mathbf{x}_r$. In a homogeneous medium, $\tau(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|/c_o$ with c_o the propagation speed. In smoothly varying media with propagation speed $c(\mathbf{x})$, the travel time is given by

$$\tau(\mathbf{x}, \mathbf{y}) = \min \int \frac{1}{c(\mathbf{X}(s))} ds,$$

where the minimum is over all paths $\mathbf{X}$ that start at $\mathbf{x}$ and end at $\mathbf{y}$.
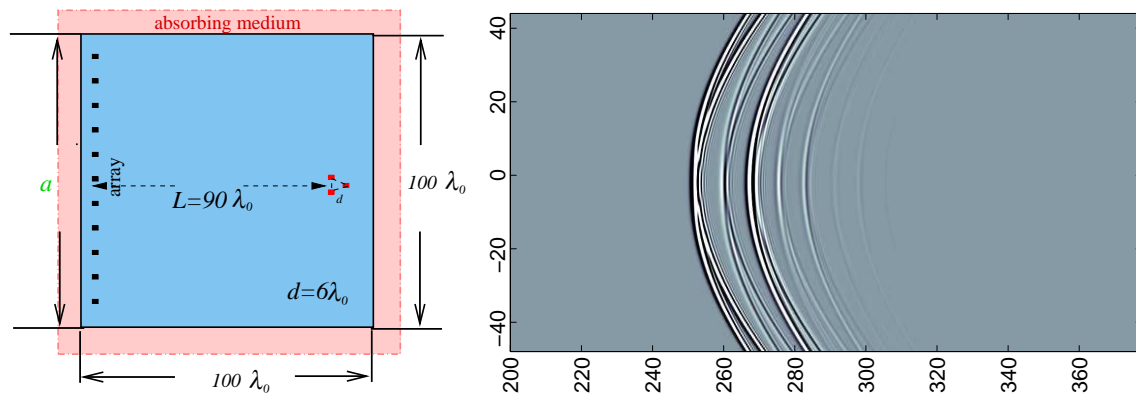
(a) The computational set-up for gathering the signals scattered by the three reflectors shown in red.

(b) Traces received in a homogeneous background, when signal $f(t)$ is sent from the central array element. The horizontal axis is time, scaled by the pulse width. The vertical axis is the receiver location, scaled by λ_o.

Figure 2. Array imaging in a homogeneous medium

Migration methods work well in smooth or piece-wise smooth media because the waves are scattered significantly only at the reflectors that we wish to image, or at the known discontinuities in the medium. Consequently, the traces are clean and the arrival times of the echoes are relatively easy to identify. We illustrate this in Figure 2-(b) for three small scatterers in a homogeneous background. The computational setup is shown in Figure 2-(a). We consider a linear array with 185 elements separated by distance $\lambda_o/2$, with an aperture of $92\lambda_o$. The range of the scatterers is $90\lambda_o$ and they are at distance $d = 6\lambda_o$ apart. The reflectors are disks of radius $\lambda_o/2$ and they are modeled as soft acoustic scatterers by setting the acoustic pressure to zero on their boundary. The pulse $f(t)$ is the derivative of a Gaussian with central frequency of 100 kHz and bandwidth $60 - 130$ kHz, measured at 6 dB. For a background speed $c_o = 3$ km/s, the central wavelength is $\lambda_o = 3$ cm. This is a typical configuration for

non-destructive testing of concrete structures with ultrasound [41]. The data traces are generated numerically by solving the acoustic wave equation as a first order velocity-pressure system with the finite element, time domain method given in [5, 6]. We model wave propagation in unbounded media by surrounding the computational domain with a perfectly matched absorbing layer [7] as shown in Figure 2-(a).

We see clearly in Figure 2-(b) the three strong hyperbolas corresponding to the direct echoes from each reflector, as well as the fainter hyperbolas coming from multiple scattering of the waves between the reflectors. The traces are zero except in the vicinity of these arrival times and the Kirchhoff migration functional (2.3) peaks at the search points that give travel times near these arrivals, along the array. The image functional $\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S)$ is shown in Figure 3 for a spatial window of 20×20 wavelengths λ_o that is centered at the configuration of the reflectors, with a spatial sampling of $\lambda_o/2$. In figure 3 and all the other figures in the paper where imaging results are displayed we use a linear color scale and the image is normalized to have maximum 1. All three reflectors are identified correctly, although there are some faint secondary peaks due to the multiple scattering between the targets that is not accounted for in the Kirchhoff migration functional.
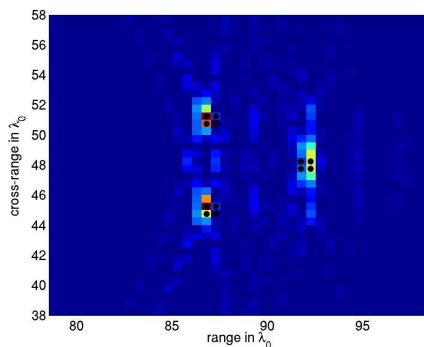

Figure 3. The Kirchhoff migration image of three reflectors in a homogeneous background as in Figure 2-(a), with the data traces of Figure 2-(b). The reflectors are indicated with black dots.

In many applications in geophysical wave propagation, ground or foliage penetrating radar, non-destructive evaluation of aging concrete structures, etc., the background is not homogeneous or smoothly varying. It is rapidly fluctuating because of inhomogeneities that can affect significantly the propagation of waves. An example of such a medium is shown in Figure 4-(a). Depending on the application, we may know the smoothly varying component of the background or we can estimate it from the data with some velocity estimation method as the image is formed [30, 24, 66, 65]. However, we usually do not know and we cannot estimate the inhomogeneities in detail, so it is natural to view them as clutter and to model them with random processes.
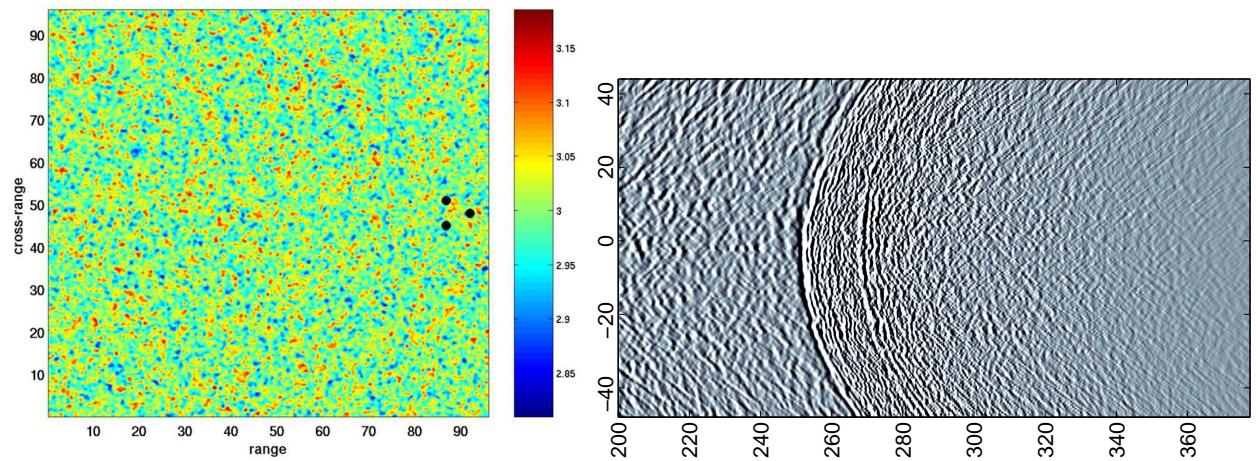
(a) Fluctuations in the sound speed $c(\mathbf{x})$ in the region between the array and the targets shown here and in Figure 2-(a) with black dots. The strength of the fluctuations of $c(\mathbf{x})$ is given in the color-bar. The axes are range and cross range in units of λ_o.

(b) Traces recorded at the array in the cluttered medium shown in Figure 4-(a). The horizontal axis is time, scaled by the pulse width. The vertical axis is the receiver location in units of the reference wavelength λ_o.

Figure 4. Array imaging in a cluttered medium.

Imaging in cluttered media, in regimes with significant multiple scattering of the waves by the inhomogeneities, is a challenging problem because the traces have a lot of delay spread or coda, as illustrated in Figure 4-(b). This delay spread cannot be viewed as additive, uncorrelated noise that can be reduced by simply averaging (stacking) the traces over the array. Therefore, the Kirchhoff migration images obtained by adding the noisy traces migrated to the search points $\mathbf{y}^S$, with the travel times computed in the smooth and known part of the medium, deteriorate significantly in clutter as seen in Figure 5. The images are not only noisy but they also change unpredictably with the realizations of the clutter, that is, they are statistically unstable.
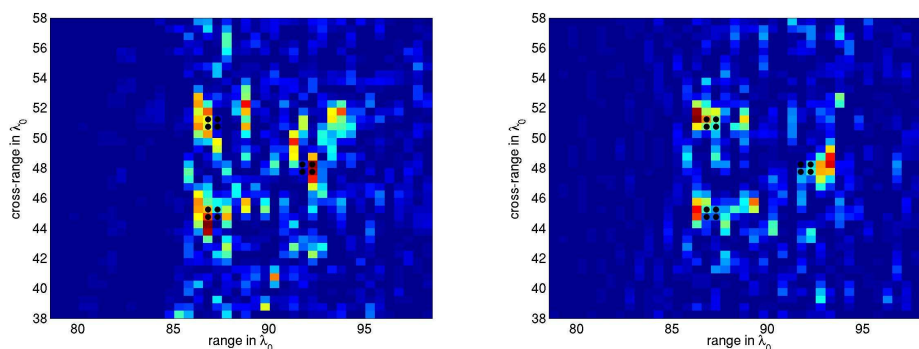

Figure 5. Kirchhoff migration images for two realizations of the clutter, with the same statistical characteristics. The data traces are as in Figure 4-(b). The support of the reflectors is indicated with black dots.

To make migration work in clutter we need an efficient technique for compressing the delay spread in the traces. We consider this in the next section.

3. Coherent interferometry

We take for simplicity the noisy traces produced with illumination from a single source located at $\mathbf{x}_s$. An effective way to reduce the delay spread in the traces is to cross correlate them and obtain the interferograms

$$
\begin{aligned}
P(\mathbf{x}_r, \mathbf{x}_s, \cdot) *_t P(\mathbf{x}_{r'}, \mathbf{x}_s, -\cdot)(t) &= \int dt' P(\mathbf{x}_r, \mathbf{x}_s, t') P(\mathbf{x}_{r'}, \mathbf{x}_s, t' - t), \\
&= \tfrac{1}{2\pi} \int d\omega \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \overline{\widehat{P}(\mathbf{x}_{r'}, \mathbf{x}_s, \omega)} e^{-i\omega t},
\end{aligned} \tag{3.5}
$$

where the bar indicates complex conjugate. These interferograms are better suited for migration than the traces themselves because by cross correlation we achieve significant cancellation of the random phases in the Fourier coefficients $\widehat{P}$. Equivalently, we reduce the unwanted effects of the clutter and we emphasize the coherent arrivals from the reflectors that we wish to image. The migration of the interferograms to the search points $\mathbf{y}^S$ is done with the travel times computed in a smooth background that we assume is known. The interferometric imaging function is

$$
\mathcal{I}^{\mathrm{INT}}(\mathbf{y}^S) = \sum_{\mathbf{x}_r, \mathbf{x}_{r'}} P(\mathbf{x}_r, \mathbf{x}_s, \cdot) *_t P(\mathbf{x}_{r'}, \mathbf{x}_s, -\cdot)|_{\tau(\mathbf{x}_r, \mathbf{y}^S) - \tau(\mathbf{x}_{r'}, \mathbf{y}^S)}. \tag{3.6}
$$

We evaluate the interferogram at the difference of the travel times because it will have a peak at that lag time when $\mathbf{y}^S$ is near a reflector.

Interferometric methods have been used before in many different contexts. For example, in geophysics, seismic traces recorded at an array of receivers are cross correlated for simulating reflection data [56, 63, 45] that can then be migrated in deterministic media to image subsurface reflectors [61]. Interferometry is also widely used in connection with matched field imaging [23, 4, 40, 34] because (3.6) takes the form of a matched field imaging functional in the frequency domain

$$
\mathcal{I}^{\mathrm{INT}}(\mathbf{y}^S) = \int_{-\infty}^{\infty} d\omega \left| \sum_{r=1}^{N_r} \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \exp\left(-i\omega \left[\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]\right) \right|^2. \tag{3.7}
$$

This matched filter form of the imaging functional has optimal noise reducing properties for data contaminated with additive, white (instrument) noise at the receivers. We use it for traces with delay spread caused by multipathing in clutter, which have much more complicated statistical structure, because we are interested in interferometry as a statistically stable imaging technique. We point the reader to our recent work in [16, 15] for a resolution study of $\mathcal{I}^{\mathrm{INT}}$ in deterministic and cluttered media. Interferometric functionals for imaging in clutter are also used in [25], in a very different way than we do with coherent interferometry, as we explain next.

The main drawback of $\mathcal{I}^{\mathrm{INT}}$ is that it provides no range resolution in clutter [16, 15], unless we have a very large array or multiple arrays that allow us to do geometric triangulation. To recover range resolution we introduced in [17, 18] a coherent interferometric imaging (CINT) functional that uses all the residual coherence in the data. While in (3.7) we treat the traces $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ as if they are uncorrelated at different frequencies, in CINT we exploit the correlations of $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ and $\widehat{P}(\mathbf{x}_{r'}, \mathbf{x}_s, \omega')$ at different frequencies and at different receivers. We can then have some range and cross range resolution even with small arrays.

There are two intrinsic and characteristic coherence parameters in the data $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$:

- the **decoherence frequency** Ω_d which is the difference in frequencies ω and ω' over which $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ and $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega')$ become uncorrelated, and

- the **decoherence length** X_d that is the difference in receiver locations $\mathbf{x}_r$ and $\mathbf{x}_{r'}$ over which $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ and $\widehat{P}(\mathbf{x}_{r'}, \mathbf{x}_s, \omega)$ become uncorrelated.

The decoherence frequency Ω_d depends on the clutter and the range L of the reflectors and when there is significant delay spread in the traces it can be much smaller than the bandwidth B. The decoherence length X_d is also determined by the clutter and the range L but it depends on the frequency as well. According to [17] it can be estimated by

$$X_d(\omega) = \frac{c_o}{\omega \kappa_d} = \frac{c_o L}{\omega a_e}, \tag{3.8}$$

where a_e is the effective aperture in time reversal [11, 16] and the dimensionless parameter κ_d quantifies the uncertainty in the direction of arrival of the echoes in clutter, as explained in detail in section 4. Note that the right side in (3.8) is the time reversal spot size in the cross range direction [32, 36, 11, 16, 51], which can be much smaller than the array aperture a in strong clutter where super-resolution of time reversal occurs.

The coherent interferometric functional is given by

$$\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d) = \int \int_{|\omega - \omega'| \leq \Omega_d} d\omega \, d\omega' \sum_{r=1}^{N_r} \sum_{\substack{r'=1 \\ |\mathbf{x}_r - \mathbf{x}_{r'}| \leq X_d\left(\frac{\omega+\omega'}{2}\right)}}^{N_r} \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \overline{\widehat{P}(\mathbf{x}_{r'}, \mathbf{x}_s, \omega')}$$

$$\exp\{-i(\omega(\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)) - \omega'(\tau(\mathbf{x}_{r'}, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)))\} \tag{3.9}$$

and it depends on the decoherence parameters Ω_d and X_d (i.e., κ_d) that are not known and must be determined from the data. Note that $\mathcal{I}^{\mathrm{CINT}}$ is equal to the square of the Kirchhoff migration functional $\mathcal{I}^{\mathrm{KM}}$ when $\Omega_d = B$ and $X_d = a$, that is, when there is no significant decoherence in the data. Note also that the matched field functional $\mathcal{I}^{\mathrm{INT}}$, defined by (3.7), can be considered as a special case of $\mathcal{I}^{\mathrm{CINT}}$, where the decoherence frequency Ω_d is zero.

The parameters Ω_d and κ_d can, in principle, be estimated directly from the data. However, this estimation is rather delicate in practice, especially in inhomogeneous random media where Ω_d and κ_d are variable. We introduce in this paper, in section 5, an adaptive algorithm for selecting these parameters using an optimization process that is based on the stability and the quality of the image that $\mathcal{I}^{\mathrm{CINT}}$ produces.

4. Coherent interferometry as statistically smoothed migration

We assume that the receivers at the array are closely spaced so we can approximate it by a continuum and we replace the sums over discrete locations $\mathbf{x}_r$ and $\mathbf{x}_{r'}$ by integrals over continuous variables $\mathbf{x}$ and $\mathbf{x}'$ on the aperture of the array. For simplicity, we take a planar array in the domain $\mathcal{A}$ and we measure the range of the reflectors in the direction orthogonal to the array and the cross range in directions parallel to it. With this convention, points $\mathbf{x}$ on the array have zero range so they can be identified with their two dimensional cross range that we also denote by $\mathbf{x}$.

Introducing the mid point and offset two dimensional vectors on the array

$$\overline{\mathbf{x}} = \frac{\mathbf{x} + \mathbf{x}'}{2}, \quad \widetilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}' \tag{4.10}$$

and the central and difference frequencies

$$\overline{\omega} = \frac{\omega + \omega'}{2}, \quad \widetilde{\omega} = \omega - \omega', \tag{4.11}$$

we rewrite the CINT functional (3.9) in the form

$$
\begin{aligned}
\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d) \sim & \int_{|\overline{\omega} - \omega_o| \leq B} d\overline{\omega} \int_{\overline{\mathbf{x}} \in \mathcal{A}} d\overline{\mathbf{x}} \int d\widetilde{\omega} \; \widehat{\Psi}(\widetilde{\omega}; \Omega_d) \int d\widetilde{\mathbf{x}} \; \widehat{\Phi}\left(\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}}; \kappa_d^{-1}\right) \\
& \widehat{P}\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} + \frac{\widetilde{\omega}}{2}\right) \overline{\widehat{P}\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} - \frac{\widetilde{\omega}}{2}\right)} \\
& \exp\left\{-i\overline{\omega}\left[\tau(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S) - \tau(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S)\right]\right\} \\
& \exp\left\{-i\widetilde{\omega}\left[\frac{\tau(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S) + \tau(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S)}{2} + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]\right\}. \tag{4.12}
\end{aligned}
$$

Here $\sim$ stands for approximate equality up to a multiplicative constant. We restrict the difference frequency $\widetilde{\omega}$ to be less than the decoherence frequency with the window function $\widehat{\Psi}(\widetilde{\omega}; \Omega_d)$, which is centered at zero and has support Ω_d. Similarly, we use the window function $\widehat{\Phi}\left(\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}}; \kappa_d^{-1}\right)$, supported in a ball of radius $\kappa_d^{-1} = \overline{\omega}X_d(\overline{\omega})/c_o$, to ensure that at central frequency $\overline{\omega}$ the offset between the receivers $|\widetilde{\mathbf{x}}|$ is less than the decoherence length $X_d(\overline{\omega})$ given by (3.8).

We have already noted that in a cluttered medium the decoherence length $X_d(\overline{\omega})$ is usually small compared to the array size. This is because it can be thought of as the

focusing spot size in time reversal. We can therefore approximate the phase in (4.12) by linearizing the travel times for small $\widetilde{\mathbf{x}}$. This gives

$$
\begin{aligned}
\tau(\overline{\mathbf{x}} + \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S) - \tau(\overline{\mathbf{x}} - \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S) &\approx \widetilde{\mathbf{x}} \cdot \nabla_{\overline{\mathbf{x}}} \tau(\overline{\mathbf{x}}, \mathbf{y}^S), \\[2mm]
\frac{\tau(\overline{\mathbf{x}} + \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S) + \tau(\overline{\mathbf{x}} - \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{y}^S)}{2} &\approx \tau(\overline{\mathbf{x}}, \mathbf{y}^S),
\end{aligned}
\tag{4.13}
$$

where both $\overline{\mathbf{x}}$ and $\widetilde{\mathbf{x}}$ are two dimensional vectors in the aperture of the array. Therefore, $\nabla_{\overline{\mathbf{x}}}$ is the tangential to the aperture gradient operator. With this phase linearization the CINT imaging functional (4.12) becomes

$$
\begin{aligned}
\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d) \sim &\int_{|\overline{\omega} - \omega_o| \le B} d\overline{\omega} \int_{\overline{\mathbf{x}} \in \mathcal{A}} d\overline{\mathbf{x}} \int d\widetilde{\omega} \, \widehat{\Psi}(\widetilde{\omega}; \Omega_d) \int d\widetilde{\mathbf{x}} \, \widehat{\Phi}\left(\frac{\overline{\omega}}{c_o} \widetilde{\mathbf{x}}; \kappa_d^{-1}\right) \\
&\widehat{P}\left(\overline{\mathbf{x}} + \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} + \tfrac{\widetilde{\omega}}{2}\right) \overline{\widehat{P}\left(\overline{\mathbf{x}} - \tfrac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} - \tfrac{\widetilde{\omega}}{2}\right)} \\
&\exp\left\{-i\overline{\omega}\widetilde{\mathbf{x}} \cdot \nabla_{\overline{\mathbf{x}}}\tau(\overline{\mathbf{x}}, \mathbf{y}^S) - i\widetilde{\omega}\left[\tau(\overline{\mathbf{x}}, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]\right\}.
\end{aligned}
\tag{4.14}
$$

From this form of the CINT functional we can give a geometrical interpretation to the parameter κ_d. We see from the exponential factor in (4.14),

$$
\exp\left\{-i\overline{\omega}\widetilde{\mathbf{x}} \cdot \nabla_{\overline{\mathbf{x}}}\tau(\overline{\mathbf{x}}, \mathbf{y}^S)\right\},
$$

that $\overline{\omega}\widetilde{\mathbf{x}}/c_o$ is the dual variable of $c_o \nabla_{\overline{\mathbf{x}}}\tau(\overline{\mathbf{x}}, \mathbf{y}^S)$, which is the direction of arrival from the search point $\mathbf{y}^S$ to the midpoint $\overline{\mathbf{x}}$ at the array when there is no clutter. Therefore, limiting the length of $\overline{\omega}\widetilde{\mathbf{x}}/c_o$ by κ_d^{-1} is equivalent to limiting the variation in the direction of arrival vector by a factor proportional to κ_d. With this interpretation, a large κ_d means that the spatial coherence window is small and there is a lot of uncertainty in the direction of arrival from the targets in clutter, as estimated by CINT.

We explain next how limiting the integrals over the space and frequency difference variables $\widetilde{\mathbf{x}}$ and $\widetilde{\omega}$ amounts to smoothing in the CINT imaging functional. Smoothing is what gives the statistical stability of the image but it also results in blurring and loss of resolution.

4.1. Statistical stability in coherent interferometric imaging

Let us begin by defining the space-frequency Wigner distribution W of the array data in the form

$$
\begin{aligned}
W\left(\overline{\mathbf{x}}, \tfrac{\overline{\omega}}{c_o}\mathbf{k}, \overline{\omega}, t\right) = &\int d\widetilde{\omega} \int d\widetilde{\mathbf{x}} \, \widehat{P}\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} + \frac{\widetilde{\omega}}{2}\right) \\
&\overline{\widehat{P}\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} - \frac{\widetilde{\omega}}{2}\right)} e^{-i\widetilde{\omega}t - i\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}} \cdot \mathbf{k}},
\end{aligned}
\tag{4.15}
$$

and introduce $\mathcal{W}(\overline{\mathbf{x}}, t, \mathbf{k})$, the integral of the Wigner distribution over the center frequency $\overline{\omega}$,

$$
\mathcal{W}(\overline{\mathbf{x}}, t, \mathbf{k}) = \int d\overline{\omega} \, W\left(\overline{\mathbf{x}}, \frac{\overline{\omega}}{c_o}\mathbf{k}, \overline{\omega}, t\right).
\tag{4.16}
$$

Here t is time and $\mathbf{k}$ is a two dimensional vector that is the dual variable of $\frac{\varpi}{c_o}\widetilde{\mathbf{x}}$, in the Fourier transform with respect to $\widetilde{\mathbf{x}}$. The vector $\mathbf{k}$ is dimensionless.

Next, define the convolution of $\mathcal{W}$ with Ψ and Φ, the inverse Fourier transforms of the window functions $\widehat{\Phi}$ and $\widehat{\Psi}$, respectively,

$$\mathcal{F}(\mathbf{K}, T; \Omega_d, \kappa_d) = \int dt \int d\mathbf{k}\ \mathcal{W}(\overline{\mathbf{x}}, t, \mathbf{k}) \Phi(\mathbf{K} - \mathbf{k}; \kappa_d) \Psi(T - t; \Omega_d^{-1}). \qquad (4.17)$$

We used here the fact that the support of the window functions Φ and Ψ is proportional to the reciprocal of the support of their Fourier transforms $\widehat{\Phi}$ and $\widehat{\Psi}$. Substituting (4.16) in (4.17), we have

$$\mathcal{F}(\mathbf{K}, T; \Omega_d, \kappa_d) = \int d\widetilde{\mathbf{x}} \int d\overline{\omega} \int d\widetilde{\omega}\ \widehat{P}\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} + \frac{\widetilde{\omega}}{2}\right) \overline{\widehat{P}\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} - \frac{\widetilde{\omega}}{2}\right)}$$

$$\int dt\ \Psi(T - t; \Omega_d^{-1}) e^{-i\widetilde{\omega}t} \int d\mathbf{k}\ \Phi(\mathbf{K} - \mathbf{k}; \kappa_d) e^{-i\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}} \cdot \mathbf{k}}$$

$$(4.18)$$

$$\sim \int d\widetilde{\mathbf{x}} \int d\overline{\omega} \int d\widetilde{\omega}\ \widehat{P}\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} + \frac{\widetilde{\omega}}{2}\right) \overline{\widehat{P}\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \overline{\omega} - \frac{\widetilde{\omega}}{2}\right)}$$

$$\widehat{\Psi}(\widetilde{\omega}; \Omega_d) \widehat{\Phi}\left(\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}}; \kappa_d^{-1}\right) \exp\left\{-i\widetilde{\omega}T - i\frac{\overline{\omega}}{c_o}\widetilde{\mathbf{x}} \cdot \mathbf{K}\right\}.$$

Recalling expression (4.14) of the CINT imaging functional we can write

$$\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d) \sim \int d\overline{\mathbf{x}}\ \mathcal{F}\left[c_o \nabla_{\overline{\mathbf{x}}} \tau(\overline{\mathbf{x}}, \mathbf{y}^S), \tau(\overline{\mathbf{x}}, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S); \Omega_d, \kappa_d\right]. \qquad (4.19)$$

At any given $\overline{\omega}$ and $\mathbf{k}$, the Wigner distribution (4.15) is highly fluctuating for different realizations of the clutter. However, by integrating it over $\overline{\omega}$ and by convolving it with the window functions Φ and Ψ we obtain the statistically stable CINT functional (4.19). A simpler version of this functional is discussed and analyzed in [51] for particular models of random media. Extensions to the statistical stability analysis of (4.19) are given in [14].

4.2. Resolution limits in coherent interferometry

In order to see more clearly the resolution limits implied by the smoothing, we rewrite the imaging functional (4.19) in the time domain

$$\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d) \sim \int d\overline{\mathbf{x}} \int d\widetilde{\mathbf{x}}\ \ P\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, t + \frac{\mathbf{k} \cdot \widetilde{\mathbf{x}}}{2c_o}\right) P\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, t - \frac{\mathbf{k} \cdot \widetilde{\mathbf{x}}}{2c_o}\right) \star_{\mathbf{k}}$$

$$(4.20)$$

$$\Phi(\mathbf{k}; \kappa_d)\big|_{\mathbf{k} = c_o \nabla_{\overline{\mathbf{x}}} \tau(\overline{\mathbf{x}}, \mathbf{y}^S)} \star_t \Psi(t; \Omega_d^{-1})\big|_{t = \tau(\overline{\mathbf{x}}, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)},$$

where we denote by $\star_{\mathbf{k}}$ and $\star_t$ the convolution with respect to the direction of arrival vector $\mathbf{k}$ and time t, respectively.

When there is no smoothing, the window functions Φ and Ψ are essentially delta functions and the approximations (4.13) make the right hand side in (4.20) the square of the Kirchhoff migration functional

$$\int d\overline{\mathbf{x}} \int d\widetilde{\mathbf{x}} \, P\left[\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \tau(\overline{\mathbf{x}}, \mathbf{x}_s) + \tau(\mathbf{x}_s, \mathbf{y}^S) + \frac{\nabla_{\overline{\mathbf{x}}}\tau(\overline{\mathbf{x}}, \mathbf{y}^S) \cdot \widetilde{\mathbf{x}}}{2}\right]$$

$$P\left[\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \tau(\overline{\mathbf{x}}, \mathbf{x}_s) + \tau(\mathbf{x}_s, \mathbf{y}^S) - \frac{\nabla_{\overline{\mathbf{x}}}\tau(\overline{\mathbf{x}}, \mathbf{y}^S) \cdot \widetilde{\mathbf{x}}}{2}\right]$$

$$\approx \int d\overline{\mathbf{x}} \int d\widetilde{\mathbf{x}} \, P\left[\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \tau\left(\overline{\mathbf{x}} + \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s\right) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]$$

$$P\left[\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s, \tau\left(\overline{\mathbf{x}} - \frac{\widetilde{\mathbf{x}}}{2}, \mathbf{x}_s\right) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right] \sim \left[\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S)\right]^2 .$$

$$(4.21)$$

As already noted, Kirchhoff migration is not stable in clutter so, to stabilize it we smooth it in (4.20) at the expense, however, of loss in resolution. The arrival times are blurred in (4.20) by the convolution with window Ψ whose support is proportional to $1/\Omega_d$. Therefore, the resolution in range is reduced to c_o/Ω_d. Similarly, the directions of arrival are blurred by convolution with the window function Φ whose support is proportional to κ_d. Therefore, the resolution in the cross range direction is proportional to $L\kappa_d$.

In homogeneous media the range resolution is proportional to c_o/B, where B is the bandwidth. In clutter the bandwidth is replaced by the decoherence frequency, which is usually smaller than B, so the range resolution deteriorates to c_o/Ω_d. We have recovered here, in a model independent way, the range resolution derived in [17] by analyzing specific models of the random medium.

The cross range resolution in clutter depends on κ_d, which quantifies the uncertainty in the direction of arrival of the echoes. Since κ_d^{-1} limits the length of $\overline{\omega}\widetilde{\mathbf{x}}/c_o$ in (4.12) and since $|\widetilde{\mathbf{x}}|$ is limited by $X_d(\overline{\omega})$, we have $\kappa_d = c_o/(\overline{\omega}X_d(\overline{\omega}))$ and

$$L\kappa_d = \frac{c_o L}{\overline{\omega}X_d(\overline{\omega})} = a_e,$$

where $X_d = c_o L/(\overline{\omega}a_e)$ is the time reversal spot size. We have obtained, in a model independent way, that the cross range resolution is determined by the effective aperture for the time reversal process in clutter. This is precisely the result derived in [16, 17] using specific models for the random medium.

In a homogeneous medium there is no loss of coherence, so we may take $X_d = a$ and obtain

$$L\kappa_d = L\frac{c_o}{\overline{\omega}X_d} = \frac{Lc_o}{\overline{\omega}a}.$$

This is just the Rayleigh cross range resolution formula [22]. In clutter, the array aperture is replaced by X_d, which is usually less than a, so we lose cross range resolution.

We have the implied inequalities

$$X_d(\omega) = \frac{c_o L}{\omega a_e} \leq \frac{c_o L}{\omega a} \leq \frac{c_o L}{\overline{\omega} X_d(\omega)} = a_e, \qquad (4.22)$$

which are consistent with the effective aperture a_e being greater than the physical aperture a that, in turn, is greater than the decoherence length X_d,

$$X_d \leq a \leq a_e.$$

The effective aperture a_e determines the enhanced focusing of time reversal in random media, where the multiple scattering effects are important and the array aperture a is small [11, 51, 16]. Inequality $a \leq a_e$ holds in typical imaging applications in random media, but there exist cases, such as in reflection seismology, where the array size a is larger than a_e.

   In the next section we discuss in more detail how to quantify the transition from imaging with a lot of coherence, as in Kirchhoff migration, to incoherent imaging. Imaging with the CINT functional (4.20) gives good results when there is residual coherence in the array data. This is an intermediate regime that is best characterized by the behavior of the imaging functional (4.20) itself with the decoherence parameters Ω_d and κ_d estimated by the adaptive algorithm of Section 5.

   We summarize in Table 1 the range and cross range resolution limits for deterministic and random media in the regime of significant multiple scattering. The first column in the table gives the Rayleigh resolution limits for imaging in homogeneous media with Kirchhoff migration. The last column in the table gives the resolution limits for the coherent interferometric imaging functional (4.20). Note that this column coincides with the first one when $\Omega_d = B$ and $X_d = a$ since in that case the CINT imaging function becomes the square of the Kirchhoff function, as shown in (4.21). In the middle column of the table we give the resolution limits of physical time reversal in clutter [36, 11], which is equivalent to back propagation when the medium is known exactly. It is clear that the cross range resolution is best with time reversal, which is the super-resolution phenomenon [32, 36, 11], it is intermediate with Kirchhoff migration in a homogeneous medium and it is worst when the random medium is not known, as is the case in coherent interferometry. Clutter helps in physical time reversal but it does not help in imaging.

4.3. Incoherent imaging

Imaging with the Kirchhoff migration functional (2.3) assumes that there is a lot of coherence in the array data, which means that clutter effects are not important. If clutter effects are important but there is still some coherence in the array data, then the coherent interferometric functional (3.9) works well, provided that the decoherence frequency Ω_d and the decoherence length X_d are chosen appropriately. This has been

	Deterministic	Known Random (TR)	Unknown Random (CINT)
Range	$\dfrac{c_o}{B}$	$\dfrac{c_o}{B}$	$\dfrac{c_o}{\Omega_d}$
Cross Range	$\dfrac{c_o L}{\overline{\omega} a}$	$\dfrac{c_o L}{\overline{\omega} a_e} \sim X_d$	$\dfrac{c_o L}{\overline{\omega} X_d} \sim a_e$

Table 1. Resolution limits in deterministic media and in random media, when the random medium is known, as in physical time reversal and, when the random medium is not known, as in coherent interferometry.

discussed in section 4.2, it has been summarized in Table 1 and it is illustrated with numerical simulations in section 5.

When there is a lot of clutter and therefore little coherence in the array data, imaging methods based on travel time migration or back propagation give poor resolution, even with interferometry as we see from the third column in Table 1. In terms of the observable coherence parameters Ω_d and X_d, strong clutter means that $\Omega_d << B$ and $X_d << a$. Although time reversal works well in such strong clutter and it can be used for delivering focused energy or for detection, that is, for deciding if a reflector is in the clutter or not, this is very different from imaging.

Imaging with incoherent data is usually done with large arrays or, even better, with arrays that surround the reflectors, using diffuse imaging methods [2]. These methods are based on modeling the propagation of wave energy with the diffusion equation. It is therefore important to know under what circumstances the propagation of wave energy can be modeled by diffusion. The derivation of transport and diffusion equations for waves in random media is considered in detail in [59], in the asymptotic regime of weak random fluctuations in the medium properties, long propagation distances compared to correlation lengths and wavelengths comparable to correlation lengths. An overview of the passage from waves to transport and diffusion from a more physical perspective is given in [70].

The dependence of the decoherence parameters Ω_d and X_d on the statistical properties of the random medium is carried out in [17] when backscattering can be neglected and in [19] for randomly layered media. As we have already pointed out, the advantage of using these parameters is that they can be estimated from the array data and that they also give an estimate of the resolution of the image. The diffuse transport of wave energy is, however, best assessed by estimating the associated diffusion coefficient, which in general requires more measurements in the random medium. The wave energy diffusion coefficient D can be written in the form [59, 70]

$$D = \frac{c_o l^*}{3}, \tag{4.23}$$

where c_o is the background speed of propagation and l^* is the **transport mean free path**. It is a length scale that is characteristic of the onset of wave diffusion in a random medium. If the distance between the target to be imaged and the array, the range L, is larger than l^* then wave migration or interferometric imaging methods will not work. In the numerical simulations that we presented in sections 2 and 3, the transport mean free path is estimated to be $75\lambda_o$, which is comparable to the range of the targets. Therefore, we have significant coherence in the data and coherent interferometry works well. When we contrast our numerical simulations to the time reversal experiments reported in [33], we note that in these experiments the width of the region of scatterers between the source and the time reversal array is $5-10$ transport mean free paths. This means that basically all coherence in the array data is lost and CINT imaging cannot possibly work in this environment.

5. Adaptive coherent interferometry

We have seen in the previous section that the coherent interferometric imaging functional (4.12) depends on the decoherence parameters κ_d and Ω_d, which are unknown at first. We introduce in this section an adaptive algorithm for estimating these parameters as minimizers of a functional of the image produced with $\mathcal{I}^{\mathrm{CINT}}$.

The decoherence parameters depend on (i) the statistical properties of the random medium, which are usually unknown, and (ii) the range of the reflectors that is either unknown or it can be estimated with a rough arrival time analysis. Since these parameters quantify the coherence in the data, they can be determined, in principle, by standard statistical signal processing techniques such as the variogram [57, 29, 52]. This gives quantitative estimates for κ_d and Ω_d, but the estimation is not sharp and it does not give the best image, in general. By best we mean here a stable image with the least blurring as described in the next section.

An alternative approach is to estimate κ_d and Ω_d directly from the image $\mathcal{I}^{\mathrm{CINT}}$, so that it is as stable and speckle free as possible while the blurring is kept to a minimum. We describe in section 5.1 an algorithm that quantifies this trade-off between blurring and speckle suppression and we assess its performance with numerical simulations in section 5.2.

5.1. The adaptive algorithm

The key idea behind the adaptive algorithm can be understood from Figure 6. We show three images obtained with $\mathcal{I}^{\mathrm{CINT}}$, with the same data but using three different values for the decoherence parameters: (left) no smoothing $\Omega_d = B$, $X_d = a$, which is the Kirchhoff migration image, (middle) optimal smoothing $\Omega_d = \Omega_d^\star$ and $X_d(\omega) = X_d^\star(\omega) = c_o/(\omega \kappa_d^\star)$, which is the image given by the adaptive CINT algorithm and (right) too much

smoothing $\Omega_d < \Omega_d^\star$, $X_d(\omega) < X_d^\star(\omega)$. The numerical setup is the same as in Figure 2-(a). The clutter is shown in Figure 4-(a) and the data traces are in Figure 4-(b).

From Figure 6 we draw two conclusions: (i) The left picture in Figure 6 compares poorly with the good image, shown in the middle, because of the spurious fluctuations caused by the coda in the traces. (ii) The picture on the right also compares poorly with the good image but for an entirely different reason: it is too blurry. Another conclusion emerges from repeating the experiment over many realizations of the clutter. While the picture on the left changes unpredictably (see for example Figure 5), the other two remain essentially unchanged. They are statistically stable.
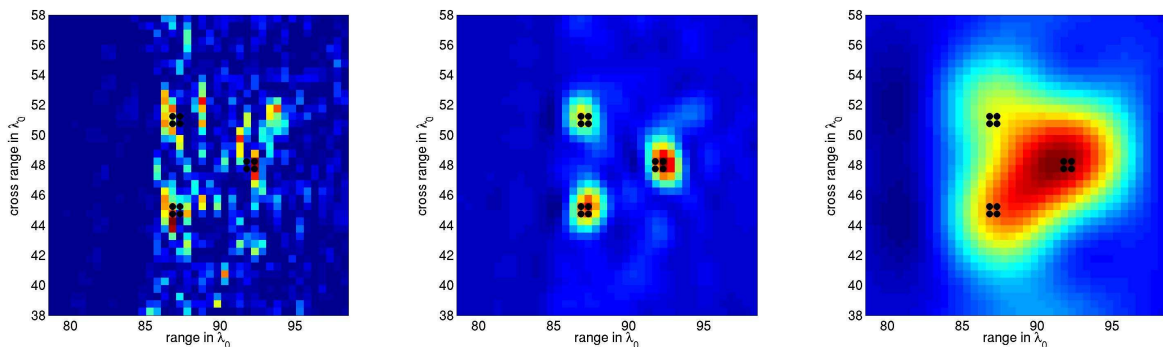
Figure 6. The coherent interferometric imaging function for three choices of smoothing parameters. The left picture is with no smoothing, the middle picture is with optimal smoothing, and the right picture with too much smoothing. The three small reflectors are indicated with black dots.

This numerical experiment shows that the two decoherence parameters Ω_d and $\kappa_d^{-1} = \frac{\omega X_d(\omega)}{c_o}$ can be obtained adaptively by using an optimization criterion that is based on the quality of the images as they are being formed. This criterion should account for the trade-off between statistical fluctuations and blurring, which is the usual bias-variance trade-off in spectral estimation of time series [55] and elsewhere. It should penalize spurious fluctuations by minimizing the gradient of the image in some appropriate norm and to control the blurring a sparsity promoting norm like L^1 should be used.

There are, of course, many criteria that meet these two objectives. In our numerical experiments we use the bounded variation norm [38] of the normalized square root of the image

$$\mathcal{J}(\mathbf{y}^S; \Omega_d, \kappa_d) = \frac{\sqrt{|\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d)|}}{\sup_{\mathbf{y}^S \in \mathcal{D}} \sqrt{|\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d)|}}. \tag{5.1}$$

We choose the square root of $\mathcal{I}^{\mathrm{CINT}}$ so that when there is no smoothing we recover the Kirchhoff migration functional. The normalization by the maximum is done so that it is the geometric features of the images that matter in the optimization process and not the magnitudes. The adaptive coherent interferometric algorithm is as follows.

Algorithm 1 *Select a search domain $\mathcal{D}$ and its sampling on some grid. For any given parameters Ω_d and κ_d, calculate $\mathcal{J}(\mathbf{y}^S; \Omega_d, \kappa_d)$ as the piecewise constant interpolation of its values at the grid points. Choose the segmentation parameters Ω_d and κ_d by minimizing*

$$\mathcal{O}(\Omega_d, \kappa_d) = \left\| \mathcal{J}(\mathbf{y}^S; \Omega_d, \kappa_d) \right\|_{L^1(\mathcal{D})} + \alpha \left\| \nabla_{\mathbf{y}^S} \mathcal{J}(\mathbf{y}^S; \Omega_d, \kappa_d) \right\|_{L^1(\mathcal{D})}, \qquad (5.2)$$

over all possible $\Omega_d \in [0, B]$ and $\kappa_d \geq 0$ and for an appropriately chosen value of the weight $\alpha > 0$.

In the objective function (5.2) we can replace L^1 by some other sparsity promoting measures, such as L^p, for $0 < p < 1$, or the Shannon entropy [62]. The parameter α controls the amount of smoothing. The larger α is, the more smoothing we do and therefore the image is blurrier. We see this in our simulations, in the right picture of Figure 6. Similarly, the smaller α is the less penalty there is on the oscillations or speckles in the image. Actually, when setting $\alpha = 0$ we found that the optimization chooses the Kirchhoff migration images, as on the left picture in Figure 6, because they have smaller L^1 norm than the good images, as in the middle of Figure 6. The parameter α should be chosen so that the two terms in (5.2) are roughly of comparable size when the objective is close to its minimum. In our numerical experiments we have found that $\alpha = 1$ works well. The choice of α does affect the results significantly and therefore it should be properly calibrated.

The optimization criterion (5.2) looks at first like any other image denoising functional used in image processing [50, 58, 26]

$$\mathcal{O}(\mathcal{I}) = \left\| \mathcal{P} - \mathcal{I} \right\|_{\mathrm{prox}} + \alpha \left\| \mathcal{I} \right\|_{\mathrm{reg}}, \qquad (5.3)$$

where $\mathcal{P}(x, y)$ is the data, in the form of a noisy image, $\mathcal{I}(x, y)$ is the denoised image to be found, $\left\| \cdot \right\|_{\mathrm{prox}}$ is a proximity norm, for example the L^2 or H^{-1} norm and $\left\| \cdot \right\|_{\mathrm{reg}}$ is a regularization norm, such as the total variation norm. However, there is a fundamental difference between (5.2) and (5.3): We do not have an image to denoise, so there is no proximity norm in (5.2). Instead, we take the L^1 norm of the image, formed during the optimization process, in order to control the blurring. Both (5.2) and (5.3) have a regularization part, the total variation of the image, but the weights α are chosen differently. In image denoising, α is usually determined with a discrepancy principle, which asks that $\left\| \mathcal{P} - \mathcal{I} \right\|_{\mathrm{prox}}$ be at the estimated noise level. In (5.2), α is chosen so that the two terms are balanced at the minimum.

Given that there are well-established image denoising techniques based on minimizing functionals like (5.3) or, by harmonic analysis methods [46], one may ask why we do not denoise the array data $P(\mathbf{x}_r, \mathbf{x}_s, t)$ directly and then do Kirchhoff migration on the denoised traces. Kirchhoff migration with denoised traces from a homogeneous medium, contaminated with additive white noise, has been done in [39] with harmonic analysis methods. We have not tried to denoise the traces in clutter for

two reasons. First, the optimization in (5.3) is computationally demanding because it is in a high dimensional space, associated with the parametrization of the traces to be denoised. In adaptive coherent interferometry we seek to estimate just two parameters: Ω_d and κ_d. Second, we do not think that denoising traces works as well as adaptive coherent interferometry because the clutter affects the traces in a complicated way and phenomena such as pulse spreading and coda cannot be modeled as additive noise, which is the implicit assumption when (5.3) is used.

5.2. Numerical results

The numerical results presented in this section are for the two dimensional setup shown in Figure 2-(a), for two configurations of the three reflectors. The first configuration is just as in Figure 2-(a) and the second one has the reflectors closer together, at distance $d = 3\lambda_o$. The fluctuations in the sound speed are generated with random Fourier series, with constant mean $c_o = 3$km/s and standard deviation 3%. We use two different correlation functions:

(i) A mono-scale, Gaussian correlation,

$$R(x_1, x_2) = R(|x_1 - x_2|) = e^{-\frac{|x_1 - x_2|^2}{2l^2}}, \tag{5.4}$$

with correlation length $l = \lambda_o/2$.

(ii) A correlation function with a range of scale sizes,

$$R(x_1, x_2) = R(|x_1 - x_2|) = \left(1 + \frac{|x_1 - x_2|}{l}\right) e^{-\frac{|x_1 - x_2|}{l}}, \tag{5.5}$$

where $l = \lambda_o$. For large spatial frequencies, the Fourier transform of (5.5) has power law behavior, as it is typical in multiscale random media, so it provides a simple model for assessing the effect of a range of scale sizes in the random medium [69].

An illustration of the Gaussian clutter is shown in Figure 4-(a) and an example of the traces in clutter is given in Figure 4-(b).

Algorithm 1 is implemented as follows. For any given parameters Ω_d and κ_d we compute the coherent interferometric image $\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \Omega_d, \kappa_d)$ by segmenting the array - bandwidth data space in rectangles $\mathcal{R}_{\overline{\mathbf{x}}, \overline{\omega}}$. The bandwidth B is divided into an integer number n_ω of intervals of size $h_\omega = B/N_\omega$, centered at frequencies $\overline{\omega}_j = (j - 1/2)h_w$, for $j = 1, \ldots, n_w$. For each $\overline{\omega}_j$, we divide the aperture a in $n_x(\overline{\omega}_j) = \mathrm{round}(2a\overline{\omega}_j \kappa_d/c_o)$ intervals of size $h_x(\overline{\omega}_j) = a/n_x(\overline{\omega}_j)$, centered at $\overline{\mathbf{x}}_q = (q - 1/2)h_x(\overline{\omega}_j)$, for $q = 1, \ldots, n_x(\overline{\omega}_j)$. This creates a nonuniform partition of the data space by rectangles $\mathcal{R}_{\overline{\mathbf{x}}_q, \overline{\omega}_j}$. We cross correlate the data in each rectangle $\mathcal{R}_{\overline{\mathbf{x}}_q, \overline{\omega}_j}$ with the data in $\mathcal{R}_{\overline{\mathbf{x}}_q, \overline{\omega}_j}$ itself and in its replicas centered at $(\overline{\mathbf{x}}_q \pm h_x(\overline{\omega}_j), \overline{\omega}_j \pm h_w)$. This makes $\Omega_d = 2B/n_w$ and $X_d(\overline{\omega}_j) = 2h_x(\overline{\omega}_j) \approx c_o/(\overline{\omega}_j \kappa_d)$. The image is computed in a square search domain $\mathcal{D}$ of size $20\lambda_o \times 20\lambda_o$ that is centered on the configuration of the reflectors and the
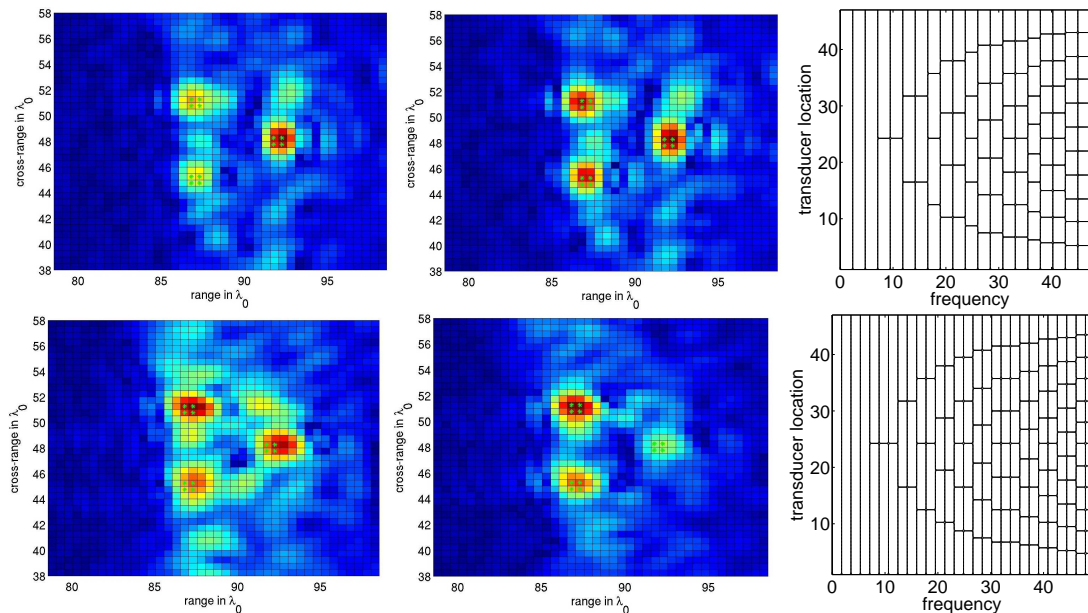
Figure 7. Images and segmentations for the cluster of three reflectors at distance $6\lambda_o$ apart. On the left the images obtained using central source illumination, in the middle a combination of three sources illumination, and on the right the segmentation shown. Top: Mono-scale, Gaussian clutter. Bottom: Multi scale random medium.

sampling is in steps of $\lambda_o/2$. We obtain the desired smoothing parameters Ω_d, κ_d by minimizing (5.2) with the matlab NOMAD optimization routines [3, 1].

We show in Figure 7 the results for the first configuration of reflectors, which are at distance $d = 6\lambda_o$ apart. First we consider a single illumination from the central element in the array and we show the optimal images and the partitioning (segmentation) of the data space in the left and right columns. The top row is for the mono-scale, Gaussian clutter and the bottom row is for the multiscale clutter. The values of the decoherence parameters can be inferred from the segmentations, as explained above. For the examples illustrated in Figure 7, we have: In the Gaussian medium $\Omega_d = B/8$ and $\kappa_d = 41.4$ and in the multiscale medium $\Omega_d = 2B/21$ and $\kappa_d = 41.4$. In theory, we expect these numbers to stay the same for different realizations of the clutter having the same statistics. However, the optimization returns slightly different results, with Ω_d ranging between $B/8 - 2B/23$ for four realizations of the Gaussian clutter and between $B/7 - 2B/21$ in the multiscale clutter. While these fluctuations cannot be predicted theoretically, because they depend on the details of the clutter realization, it is important to note that when we calculated images with the same segmentation for all four realizations of the Gaussian and multiscale media, we obtained images that are nearly as good as the optimal ones.

The middle column in Figure 7 shows images obtained with a combination of three different illuminations of the reflectors, by the central source in the array and by two

other sources at a quarter of the aperture distance above and below the array center, respectively. The combination is done as follows. First, we take a weighted average (stack) of the migrated traces

$$\widehat{m}(\mathbf{x}_r, \mathbf{y}^S, \omega) = \sum_{s=1}^{3} w_s \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) e^{-i\omega[\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)]}, \tag{5.6}$$

with some weights $\mathbf{w} = (w_1, w_2, w_3)$ to be determined. We note here that the summation is over all sources $\mathbf{x}_s$, keeping the search point $\mathbf{y}^S$ fixed. Next, we cross correlate the stacked data (5.6) and sum over the array to obtain the imaging function

$$\tilde{\mathcal{I}}(\mathbf{y}^S; \mathbf{w}, \Omega_d, \kappa_d) = \int_{|\omega - \omega_o| \leq B} d\omega \int_{\substack{|\omega' - \omega_o| \leq B \\ |\omega' - \omega| \leq \Omega_d}} d\omega'$$
$$\sum_{\substack{\mathbf{x}_r \in \mathcal{A} \\ |\mathbf{x}_r - \mathbf{x}_r'| \leq \frac{2c_o}{(\omega + \omega')\kappa_d}}} \sum_{\mathbf{x}_r' \in \mathcal{A}} \widehat{m}(\mathbf{x}_r, \mathbf{y}^S, \omega) \overline{\widehat{m}(\mathbf{x}_r', \mathbf{y}^S, \omega')}, \tag{5.7}$$

with the smoothing parameters determined by the adaptive CINT algorithm, with the central illumination. We leave the smoothing unchanged because Ω_d and κ_d should depend just on the clutter and the distance traveled by the waves, which is roughly the same for all three sources. We also choose to combine the illuminations at the level of the traces and not of the images because, as can be seen from section 4.2, the CINT images are positive nearly everywhere and so almost no error cancellation is achieved by summing them.

In (5.7) we cross correlate traces that come from three sources and they should not be further apart than the decoherence length. From the segmentations in Figure 7 we find that, depending on the frequency, the decoherence length, which is twice the size of each window as explained above, can be as small as a quarter of the aperture. However, this is just for the highest frequencies where the signal is small. For most of the bandwidth the separation of $a/2$ between the second and the third source falls within the allowed spatial offset.

We choose the weights by minimizing the bounded variation norm of the normalized square root of $\tilde{\mathcal{I}}$, subject to constraints

$$\sum_{s=1}^{3} w_s = 1, \quad w_s \geq 0, \quad s = 1, 2, 3. \tag{5.8}$$

In Figure 8 we show the results for the second configuration of the reflectors, which are at a distance $3\lambda_o$ apart. Because the imaging is done in the same clutter as in Figure 7 we keep the same segmentations of the data space. We note that the central illumination image in the Gaussian clutter is better than in the multiscale medium. However, in both cases the results improve considerably by taking the optimal combination of the three available illuminations. The reflectors are not as well resolved as in Figure 7 because they are closer together. We can improve the results further by deblurring, which we do using the level set approach described in [44].
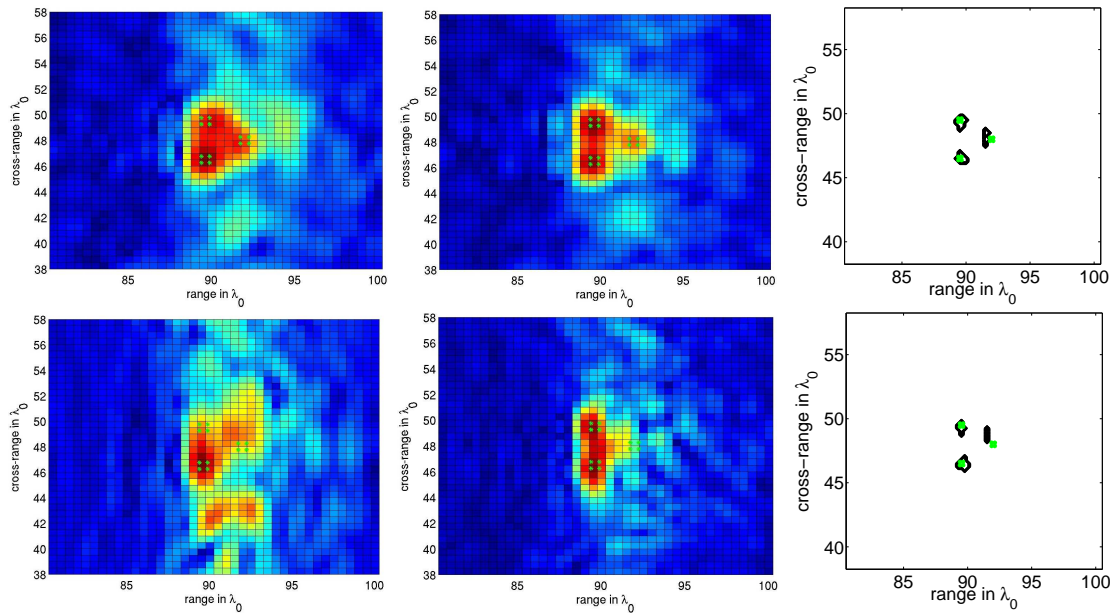
Figure 8. Images of the cluster of nearby reflectors, separated by distance $3\lambda_o$. Left: the central source illumination. Middle: illumination with a combination of three sources. Right: The deblurred images. Top: Mono-scale, Gaussian clutter. Bottom: Multi scale random medium.

We describe very briefly the preliminary results of the deblurring, shown on the right in Figure 8. The blurring kernel is taken as a Gaussian

$$\mathcal{K}(\mathbf{y}^S, \mathbf{y}) = \beta e^{-\frac{(y_s-y)^2}{2s_x^2} - \frac{(z_s-z)^2}{2s_z^2}}, \quad \text{for } \mathbf{y} = (y, z), \ \mathbf{y}^S = (y_s, z_s), \quad (5.9)$$

with parameters β, s_x and s_z that depend on the random medium and on the range L. These parameters are related to the decoherence parameters but here they are estimated by fitting the kernel to the image obtained with the adaptive CINT algorithm of the target behind in Figure 7. The form (5.9) of the blurring kernel can be justified analytically for some specific models for the clutter as we do in [17]. Although the numerical simulations are in a different regime than the one analyzed in [17], we find that the results are qualitatively similar to those predicted in [17], which is why we use the same type of blurring kernel.

6. Optimal illumination and waveform design

We have already seen in section 5.2 that we can improve significantly the quality of the coherent interferometric images by appropriate averaging (stacking) over the different illuminations. In this section we explore further the question of optimal illumination and introduce a new variational approach for its solution. We separate the optimal illumination problem from the estimation of the optimal segmentation discussed in section 5 because the later is determined by the clutter. While the clutter is a fixed but

unknown entity in the inverse problem, the optimal illumination question is part of the data acquisition framework that we can change in order to improve the results.

As discussed in the introduction, the traditional approach to the optimal illumination problem has focused on maximizing the power of the scattered echoes recorded at the receivers [53, 47, 43, 36]. This is done with the singular value decomposition of the response matrix for each frequency. Kirchhoff migration images constructed with this optimal illumination enhance the image near the strongest reflector. In the time domain, maximization of the received power can be done with iterative time reversal [54, 27, 28]. The resulting optimal waveforms are, however, narrow-band signals, which are bad for imaging because the lack of bandwidth gives poor range resolution and poor statistical stability in clutter [11, 16, 15, 20, 21].

In this section we propose a different approach. We choose the illumination pattern and the waveform by using an optimality criterion that is based on the image itself, as we do in adaptive coherent interferometry.

6.1. Formulation of the optimal illumination problem

The idea behind our formulation of the optimal illumination problem is quite similar to that of section 5.1. Since it is the image quality that we wish to improve, we should choose the illumination using an optimization criterion that measures the quality of the image. How we measure this depends on the problem. For example, when imaging small reflectors in clutter the measure can be as in section 5.1, where we take the bounded variation norm of the square root of the normalized images. Here, we take a simpler choice, the L^1 norm of the CINT image, normalized by its maximal value. We choose this norm because it is easier to analyze and we discard the norm of $\nabla \mathcal{I}^{\mathrm{CINT}}$ because we suppose that spurious fluctuations such as those on the left in Figure 6 have been eliminated by the prior estimation of the statistical smoothing parameters Ω_d and κ_d. Of course, other objective functions could be used for other problems, but this is not the point that we wish to make here. Instead, let us fix the L^1 norm as our measure of the quality of the image and concentrate next on how to find the illuminations that optimize it.

Assuming that we have N_s sources at locations $\mathbf{x}_s$, $s = 1, \ldots, N_s$, we illuminate the reflectors with one source at a time and we gather the scattered echoes $P(\mathbf{x}_r, \mathbf{x}_s, t)$ at receivers $\mathbf{x}_r$, $r = 1, \ldots, N_r$. If we could send a flat pulse in the frequency domain $[\omega_o - B, \omega_o + B]$, we would have the impulse response matrix $\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ over the bandwidth of interest. Then, the echoes due to an arbitrary waveform $g(\mathbf{x}_s, t)$ sent from $\mathbf{x}_s$, would be given by

$$P(\mathbf{x}_r, \mathbf{x}_s, t) = \int_{|\omega - \omega_o| \leq B} d\omega \widehat{g}(\mathbf{x}_s, \omega) \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) e^{-i\omega t}. \tag{6.1}$$

Alternatively, we could collect data $P(\mathbf{x}_r, \mathbf{x}_s, t)$ for all the waveforms in the optimization

process, which in most situations is quite expensive.

6.1.1. Optimal illumination for detection Before formulating the new optimal illumination problem that is based on the quality of the image, we recall the form of the optimal illumination problem for maximum received power at the array and, therefore, for detecting the presence of reflectors in the medium. If the sources at $\mathbf{x}_s$ emit simultaneously signals $g(\mathbf{x}_s, t)$, then the echo received at $\mathbf{x}_r$ is given in the frequency domain by

$$\sum_{s=1}^{N_s} \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{g}(\mathbf{x}_s, \omega) \tag{6.2}$$

and the total power received at the array is

$$P_{tot} = \int_{|\omega - \omega_o| \leq B} d\omega \sum_{r=1}^{N_r} \left| \sum_{s=1}^{N_s} \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{g}(\mathbf{x}_s, \omega) \right|^2. \tag{6.3}$$

The best signals $\widehat{g}(\mathbf{x}_s, \omega)$ for detecting reflectors in the medium maximize P_{tot}, subject to the normalization constraint

$$\int_{|\omega - \omega_o| \leq B} d\omega \sum_{s=1}^{N_s} |\widehat{g}(\mathbf{x}_s, \omega)|^2 = 1 \tag{6.4}$$

of the power emitted from the sources. We denote the maximum power by

$$\mathcal{P} = \max_{\widehat{g}} P_{tot}.$$

Let us use the short notation $\widehat{\Pi}(\omega)$ for the $N_r \times N_s$ impulse response matrix with entries $\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ and let $\lambda^{(M)}(\omega)$ and $\widehat{g}^{(M)}(\omega)$ be the largest eigenvalue and corresponding normalized eigenvector of the hermitian, positive definite matrix $\widehat{\Pi}^\star(\omega)\widehat{\Pi}(\omega)$, where the superscript $\star$ denotes the adjoint. Let also $\omega_{\max}$ be the (resonant) frequency, assumed unique, at which $\lambda^{(M)}(\omega)$ is maximum. If for $\delta > 0$ and small we define

$$\widehat{g}(\mathbf{x}_s, \omega) = \begin{cases} \frac{1}{\sqrt{2\delta}} \widehat{g}_s^{(M)}(\omega), & |\omega_{\max} - \omega| \leq \delta, \\ 0, & \text{otherwise}, \end{cases} \tag{6.5}$$

where $\widehat{g}_s^{(M)}(\omega)$ is the s-th component of the eigenvector $\widehat{g}^{(M)}(\omega)$, it is easily seen that the maximizer of P_{tot} is given by (6.5) in the limit $\delta \to 0$ and that $\mathcal{P} = \lambda^{(M)}(\omega_{\max})$. This illumination is clearly a narrow-band signal and it is the one that arises with iterative time reversal [27, 54]. The result is optimal for detection, but the lack of bandwidth is bad for migration imaging, as we illustrate in Figure 11.

If the illumination is chosen to equal the eigenvector corresponding to the largest eigenvalue

$$\widehat{g}(\mathbf{x}_s, \omega) = \frac{1}{\sqrt{2B}} \widehat{g}_s^{(M)}(\omega), \quad |\omega - \omega_o| \leq B,$$

then the received power is

$$\frac{1}{2B} \int_{|\omega-\omega_o|\leq B} d\omega \lambda^{(M)}(\omega) \leq \mathcal{P}.$$

For well separated targets, the largest eigenvalue of $\widehat{\Pi}^\star(\omega)\widehat{\Pi}(\omega)$ is associated with echoes received from the most strongly reflecting one, across the bandwidth. The Kirchhoff migration image with this illumination is sharply peaked at the most strongly reflecting target, as is seen from the numerical simulations shown in Figure 12. This is the DORT method of selective focusing in time reversal [53, 36], which is implemented with the singular value decomposition of the response matrix $\widehat{\Pi}(\omega)$. This approach was extended to statistically stable imaging in random media in [20, 21].

6.1.2. Optimal illumination for imaging Illuminating with the source waveforms $\widehat{g}_s^{(M)}(\omega)$, which are the components of the leading eigenvector of matrix $\widehat{\Pi}^\star\widehat{\Pi}$, is an efficient way of delivering energy or beamforming onto the target that sends back the strongest echo. Such selective beamforming is not used in coherent interferometric imaging, or in classic migration, because these are imaging methods that rely on cancellation of phases by back propagation. When we have data from N_s sources we combine them in the CINT imaging function in the form

$$\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \widehat{f}) = \int_{|\omega-\omega_o|\leq B} d\omega \int_{\substack{|\omega'-\omega_o|\leq B \\ |\omega'-\omega|\leq \Omega_d}} d\omega' \sum_{r=1}^{N_r} \sum_{\substack{r'=1 \\ |\mathbf{x}_r-\mathbf{x}_{r'}|\leq X_d\left(\frac{\omega+\omega'}{2}\right)}}^{N_r}$$

$$e^{-i\omega\tau(\mathbf{x}_r,\mathbf{y}^S)} \sum_{s=1}^{N_r} \widehat{\Pi}(\mathbf{x}_r,\mathbf{x}_s,\omega)\widehat{f}(\mathbf{x}_s,\omega)e^{-i\omega\tau(\mathbf{x}_s,\mathbf{y}^S)} \tag{6.6}$$

$$e^{i\omega'\tau(\mathbf{x}_{r'},\mathbf{y}^S)} \sum_{\substack{s'=1 \\ |\mathbf{x}_s-\mathbf{x}'_s|\leq X_d\left(\frac{\omega+\omega'}{2}\right)}}^{N_s} \overline{\widehat{\Pi}(\mathbf{x}_{r'},\mathbf{x}_{s'},\omega')\widehat{f}(\mathbf{x}'_s,\omega')e^{-i\omega'\tau(\mathbf{x}_{s'},\mathbf{y}^S)}},$$

so that the source functions have the form

$$\widehat{g}(\mathbf{x}_s,\omega) = \widehat{f}(\mathbf{x}_s,\omega)e^{-i\omega\tau(\mathbf{x}_s,\mathbf{y}^S)}. \tag{6.7}$$

The functions $\widehat{f}(\mathbf{x}_s,\omega)$ need only play the role of weights here because the $\widehat{g}(\mathbf{x}_s,\omega)$ beamform automatically, in the smooth and known background medium, to the search point $\mathbf{y}^S$. We will see in the next section (below (6.17)) that the source functions $\widehat{g}(\mathbf{x}_s,\omega)$ have the form of a singular vector for an isolated point scatterer at $\mathbf{y}^S$.

In the case of one source, (6.6) reduces to the CINT functional (3.9). For many sources, we cross correlate traces for sources and receivers that are not more than the decoherence length $X_d(\omega) = c_o/(\omega\kappa_d)$ apart. The segmentation parameters Ω_d and κ_d are assumed to have been determined prior to addressing the optimal illumination

question, so we suppress them from the arguments of the imaging function which is a quadratic form in $\widehat{f}(\mathbf{x}_s, \omega)$.

We wish now to determine the illumination functions $\widehat{f}(\mathbf{x}_s, \omega)$ by minimizing the objective function

$$\mathcal{O}(\widehat{f}) = \|\mathcal{J}(\mathbf{y}^S; \widehat{f})\|_{L^1(\mathcal{D})}, \tag{6.8}$$

where $\mathcal{J}(\mathbf{y}^S; \widehat{f})$ is the CINT image $\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \widehat{f})$ normalized to maximum value one. We shall optimize over illuminations of the form

$$\widehat{f}(\mathbf{x}_s, \omega) = w_s \widehat{f}_B(\omega - \omega_o), \quad s = 1, ..., N_s, \tag{6.9}$$

which correspond to sending the same pulse $f(t) = e^{-i\omega_o t} \widehat{f}_B(t)$ from all the sources in the array. This is a simplifying assumption that is motivated partially by an analytic result which shows that the product form (6.9) is optimal for imaging single point scatterers in homogeneous media [13]. This does not hold in general, so (6.9) leads to a suboptimal illumination problem, which is, however, numerically robust and less costly to implement. This is discussed further in the Section 6.3. The minimization is performed subject to the following constraints:

(i) We view $\mathbf{w} = (w_1, \ldots, w_{N_s})$ as a vector of weights assigned to the sources at $\mathbf{x}_s$, after collecting the data traces $\widehat{f}_B(\omega - \omega_o)\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega)$. This means that

$$\sum_{s=1}^{N_s} w_s = 1, \quad w_s \geq 0, \quad s = 1, \ldots, N_s. \tag{6.10}$$

(ii) The support of $\widehat{f}_B(\omega - \omega_o)$ is restricted to the fixed frequency band $[\omega_o - B, \omega_o + B]$ and we ask that

$$\widehat{f}_B(\omega - \omega_o) \geq 0, \quad \text{for all } \omega \in [\omega_o - B, \omega_o + B]. \tag{6.11}$$

$$\int_{|\omega - \omega_o| \leq B} d\omega \widehat{f}_B(\omega - \omega_o) = 1. \tag{6.12}$$

The constraint $\widehat{f}_B(\omega - \omega_o) \geq 0$ is a natural one in clutter because Ω_d is usually small relative to the scale of variation of the broadband waveforms that we seek, which means that $\mathcal{I}^{\mathrm{CINT}}$ involves essentially absolute values squared of $\widehat{f}_B(\omega - \omega_o)$. In the absence of clutter, Ω_d is replaced by B, so $\mathcal{I}^{\mathrm{CINT}}$ is like the square of the Kirchhoff migration function and the image is determined by $\widehat{f}_B(\omega - \omega_o)$ and not its absolute value. However, after writing $(\mathcal{I}^{\mathrm{KM}})^2$ as a double integral over the frequencies and after integrating over $\mathbf{y}^S \in \mathcal{D}$ to get the objective function (6.8), we see that the contribution from frequencies that are not equal is small. Thus, in both smooth and cluttered media the objective function depends approximately on $|\widehat{f}_B(\omega - \omega_o)|$ and this motivates our restriction of f_B to the class of correlation functions, which have nonnegative Fourier transforms. This leads to a significant simplification in the optimization problem because we can normalize the pulse, in the L^1 sense, with a linear equality constraint. Other

normalizations, such as L^2, that have the physical meaning of fixing the power delivered by each source can be considered as well, but at the expense of complicating the numerical optimization.

Finally we also impose a quadratic constraint on the waveforms with a lower bound $\mathcal{N}$ on the received power at the array

$$\sum_{r=1}^{N_r} \int_{|\omega - \omega_o| \leq B} d\omega \left| \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \right|^2 \geq \mathcal{N}, \quad \text{for all } s = 1, \ldots, N_s. \tag{6.13}$$

Here we take $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{f}_B(\omega - \omega_o)$. This constraint is a very important feature of our imaging algorithm and is essential for applications where instrument or environmental noise at the array is an issue. Clearly the level $\mathcal{N}$ must be less than the maximum received total power at the array, $0 < \mathcal{N} < \mathcal{P}$. Setting the level $\mathcal{N}$ close to $\mathcal{P}$ is appropriate when the signals received at the array are weak and instrument noise affects them significantly. In this case, the images obtained with the optimal waveforms are similar to those given by the optimal power illumination from (6.3) and (6.4). When there is no clutter, this is clearly seen in the numerical simulations shown in Figures 13, 14 and 16 that we describe in the next section. When the level $\mathcal{N}$ is set significantly below $\mathcal{P}$, the optimal waveforms give images that have much better resolution than those produced with optimal received power waveforms. The ratio $\mathcal{P}/\mathcal{N}$ plays the role of a signal to noise ratio so we define a proxy SNR by

$$SNR = \frac{\mathcal{P}}{\mathcal{N}}. \tag{6.14}$$

6.2. Optimal illumination and the singular value decomposition

In order to understand better the scope of the optimal illumination problem in section 6.1.2, we consider its analytical form for the case of M point scatterers at locations $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_M$ in a homogeneous medium. We assume for simplicity an equal number $N > M$ of sources and receivers collocated at points $\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N$ in the array.

Let us denote by

$$\widehat{g}_o(\omega, \mathbf{y}_j) = [\widehat{G}_o(\mathbf{x}_p, \mathbf{y}_j, \omega)]_{p=1...N} \tag{6.15}$$

the vector of Fourier coefficients of the received signals at the array, due to a delta function impulse sent from a point source at $\mathbf{y}_j$. Here $\widehat{G}_o$ is the free space Green's function

$$\widehat{G}_o(\mathbf{x}, \mathbf{y}, \omega) = \frac{e^{i\omega\tau(\mathbf{x},\mathbf{y})}}{4\pi|\mathbf{x} - \mathbf{y}|}. \tag{6.16}$$

If we ignore multiple scattering between the point targets, the $N \times N$ impulse response matrix is

$$\widehat{\Pi}(\omega) = [\widehat{\Pi}(\mathbf{x}_p, \mathbf{x}_q, \omega)]_{p,q=1...N} = \sum_{j=1}^{M} \xi_j(\omega) \widehat{g}_o(\omega, \mathbf{y}_j) \widehat{g}_o(\omega, \mathbf{y}_j)^T, \tag{6.17}$$

where the superscript T stands for transpose and $\xi_j(\omega)$ is the scattering amplitude of the j-th scatterer. In the case of well separated scatterers the vectors $\widehat{g}_o(\omega, \mathbf{y}_j)$ are approximately orthogonal and expression (6.17) is the singular value decomposition of $\widehat{\Pi}(\omega)$. Its singular values are given by

$$\sigma_j(\omega) = |\xi_j(\omega)| \|\widehat{g}_o(\omega, \mathbf{y}_j)\|^2$$

and the right singular vectors by

$$\widehat{u}(\omega, \mathbf{y}_j) = \overline{\widehat{g}_o(\omega, \mathbf{y}_j)} / \|\widehat{g}_o(\omega, \mathbf{y}_j)\|.$$

From (6.16) we see that $\|\widehat{g}_o(\omega, \mathbf{y}_j)\|$ does not depend on ω. Therefore the right singular vectors have approximately the form

$$\widehat{u}_s(\omega, \mathbf{y}_j) = \tilde{w}_s e^{-i\omega\tau(\mathbf{x}_s, \mathbf{y}_j)} \tag{6.18}$$

for some weights $\tilde{w}_s$. Up to allowing for more general frequency dependence, this is the form that the illumination functions have in (6.7), with the product form (6.9) for $\widehat{f}(\mathbf{x}_s, \omega)$ and with $\mathbf{y}_j$ replaced by the search point $\mathbf{y}^S$.

Suppose now that we assign uniform weights $w_s = 1$ (not normalized) to each source in the array and that we image by migrating the data to search points $\mathbf{y}^S$, with the full Green's function. The resulting imaging functional is

$$\mathcal{I}(\mathbf{y}^S; \widehat{f}_B) = \int d\omega \widehat{f}_B(\omega - \omega_o) \overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{\Pi}(\omega) \overline{\widehat{g}_o(\omega, \mathbf{y}^S)} \sim \mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S; \widehat{f}_B) \tag{6.19}$$

and it behaves roughly like the Kirchhoff migration function, which uses just travel times [67]. If we insert (6.17) into the expression of $\mathcal{I}(\mathbf{y}^S)$ we obtain

$$\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S; \widehat{f}_B) \sim \int d\omega \widehat{f}_B(\omega - \omega_o) \sum_{j=1}^{M} \xi_j(\omega) \left[\overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{g}_o(\omega, \mathbf{y}_j) \right]^2. \tag{6.20}$$

The terms $\overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{g}_o(\omega, \mathbf{y}_j)$ in the integral have small absolute value unless the search point $\mathbf{y}^S$ is close to a scatterer $\mathbf{y}_j$. In fact, $|\overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{g}_o(\omega, \mathbf{y}_j)|$ is the point spread function for time reversal imaging of a point source at $\mathbf{y}_j$, at frequency ω and in the homogeneous medium. When the range L is large compared to the array aperture a and the wavelength is small, the support of this point spread function is of order $\lambda_o L/a$, in cross range. This means that $\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S)$ peaks for search points $\mathbf{y}^S$ within a cross range distance of order $\lambda_o L/a$ from some target location $\mathbf{y}_j$ and that it is small elsewhere in the search domain. This is for the usual Kirchhoff migration, without any optimization of the weights.

If we weight differently the sources in the migration imaging function we get

$$\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S; \mathcal{W}, \widehat{f}_B) \sim \int d\omega \widehat{f}_B(\omega - \omega_o) \sum_{j=1}^{M} \xi_j(\omega) \overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{g}_o(\omega, \mathbf{y}_j) \widehat{g}_o(\omega, \mathbf{y}_j)^T \mathcal{W} \overline{\widehat{g}_o(\omega, \mathbf{y}^S)}, \tag{6.21}$$

where $\mathcal{W}$ is the $N \times N$ diagonal matrix with entries equal to the weights w_s, $s = 1, ..., N$. The square of this functional is the same as $\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S)$, for Ω_d replaced by B and $X_d = a$,

$$\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \mathcal{W}, \widehat{f}_B) \sim \left[\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S; \mathcal{W}, \widehat{f}_B) \right]^2. \tag{6.22}$$

The $L^1(\mathcal{D})$ norm of $\mathcal{I}^{\mathrm{KM}}$ is given by

$$\|\mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \mathcal{W}, \widehat{f}_B)\|_{L^1(\mathcal{D})} \sim \|\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S; \mathcal{W}, \widehat{f}_B)\|_{L^2(\mathcal{D})}^2 \approx \int d\omega |\widehat{f}_B(\omega - \omega_o)|^2$$

$$\sum_{j=1}^{M} |\xi_j(\omega)|^2 \int_{\mathcal{D}} d\mathbf{y}^S |\overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \widehat{g}_o(\omega, \mathbf{y}_j)|^2 |\overline{\widehat{g}_o(\omega, \mathbf{y}^S)}^T \mathcal{W} \widehat{g}_o(\omega, \mathbf{y}_j)|^2 \tag{6.23}$$

and this determines our objective function (6.8), up to a normalization. Here we used the fact that when integrating with respect to $\mathbf{y}^S$ the square of $\mathcal{I}^{\mathrm{KM}}(\mathbf{y}^S)$, written as a double integral over two frequencies, the cross terms in the double sum are small, as are contributions from frequencies that are not nearly equal.

To complete the minimization problem under constraints (6.10), (6.11), (6.12) and (6.13), let us write next the form of the power constraint (6.13), when $\widehat{\Pi}$ is given by (6.17). Using the assumed orthogonality of $\widehat{g}_o(\omega, \mathbf{y}_j)$, for $j = 1, ..., M$, we find

$$\int d\omega |\widehat{f}_B(\omega - \omega_o)|^2 \sum_{j=1}^{M} |\xi_j(\omega)|^2 \|\widehat{g}_o(\omega, \mathbf{y}_j)\|^2 \mathrm{diag}(\overline{\widehat{g}_o(\omega, \mathbf{y}_j)} \widehat{g}_o(\omega, \mathbf{y}_j)^T) \geq \mathcal{N} I, \tag{6.24}$$

where I is the identity matrix and notation $\mathrm{diag}(A)$ stands for a diagonal matrix, with entries A_{ii}, for $i = 1, \ldots, N$ and for an arbitrary $A \in \mathbb{R}^{N \times N}$. Finally, we note that (6.15) implies that terms $\|\widehat{g}_o(\omega, \mathbf{y}_j)\|^2 \mathrm{diag}(\overline{\widehat{g}_o(\omega, \mathbf{y}_j)} \widehat{g}_o(\omega, \mathbf{y}_j)^T)$ in (6.24) are independent of frequency and we let $c_{js} > 0$ denote their s-element. Inequalities (6.24) become

$$\int d\omega |\widehat{f}_B(\omega - \omega_o)|^2 \sum_{j=1}^{M} |\xi_j(\omega)|^2 c_{js} \geq \mathcal{N} , \quad s = 1, ..., N. \tag{6.25}$$

We can now see the structure of the minimization problem for $\mathcal{O}(\mathcal{W}, \widehat{f})$, under the constraints (6.10),(6.11),(6.12) and (6.25). When the objective does have approximately the form (6.23), then the positivity constraint (6.11) is not a restriction since only the absolute value of $\widehat{f}_B$ appears in (6.23) and (6.25). The interpretation of w_s as weights of the sources can be understood by noting that in (6.23), the steering of the array data to the search points $\mathbf{y}^S \in \mathcal{D}$ is done by the migration vector $\widehat{g}_o(\omega, \mathbf{y}^S)$, which has all the relevant phase information. We have therefore, in this case of well separated point targets, a robust optimization problem for optimal illumination without any essential loss in generality. The results of the numerical simulations in section 6.5 support very well this conclusion.

The main advantage of the optimal illumination problem that we have introduced here is that it is robust in random media. This is because it uses the CINT functional, which is statistically stable, and because the constraints that we have used, (6.11) in particular, are compatible with the smoothing that adaptive CINT introduces.

6.3. The implementation of the optimization problem

Let us discretize the frequency band on a uniform mesh, with N_f points and spacing $\Delta\omega$. We denote by $\widehat{\mathbf{f}}_{\mathbf{B}}$ the vector of components $\widehat{f}_B(\omega_j - \omega_o)$, for $\omega_j = (j-1)\Delta\omega + \omega_o - B$ and $j = 1, \ldots, N_f$. The vector of weights is $\mathbf{w} = (w_1, \ldots, w_{N_s})$. We wish to solve the following optimization problem.

Problem 1 *Minimize objective function*

$$\mathcal{O}(\mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}}) = \left\| \mathcal{J}(\mathbf{y}^S; \mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}}) \right\|_{L^1(\mathcal{D})}, \tag{6.26}$$

given by the $L^1(\mathcal{D})$ norm of the normalized image

$$\mathcal{J}(\mathbf{y}^S; \mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}}) = \frac{\mathcal{I}^{CINT}(\mathbf{y}^S; \mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}})}{\max_{\mathbf{y}^S \in \mathcal{D}} \mathcal{I}^{CINT}(\mathbf{y}^S; \mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}})}, \tag{6.27}$$

over the weights $\mathbf{w}$ and the vector $\widehat{\mathbf{f}}_{\mathbf{B}}$ of the Fourier coefficients of the pulse, subject to constraints (6.10) and

$$\widehat{f}_B(\omega_j - \omega_o) \geq 0, \quad j = 1, \ldots, N_f, \tag{6.28}$$

$$\Delta\omega \sum_{j=1}^{N_f} \widehat{f}_B(\omega_j - \omega_o) = 1, \tag{6.29}$$

$$w_s^2 \left[\Delta\omega \sum_{j=1}^{N_f} \sum_{r=1}^{N_r} \left| \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega_j) \right|^2 - \mathcal{N} \right] \geq 0, \quad s = 1, \ldots, N_s, \tag{6.30}$$

Note that we multiply the inequalities in (6.30) by the weights to ensure that only sources that participate in the imaging, that is, have nonzero weights, restrict the feasibility region of the waveform.

This optimization can be quite expensive. However, our numerical experiments indicate that we can take a simpler, iterative approach that is based on the following two sub-optimal problems:

Problem 2 *Fix $\widehat{\mathbf{f}}_{\mathbf{B}}$ and minimize $\mathcal{O}(\mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}})$ over the weights $\mathbf{w}$, with constraints (6.10).*

Problem 3 *Fix the weights $\mathbf{w}$ to those determined in Problem 2 and minimize $\mathcal{O}(\mathbf{w}, \widehat{\mathbf{f}}_{\mathbf{B}})$ over $\widehat{\mathbf{f}}_{\mathbf{B}}$, subject to constraints (6.28) - (6.29). Then, return to Problem 2 to update the weights and iterate.*

While Problems 2 and 3 are sub-optimal, they can be solved with smaller computational effort than Problem 1 and, at least for the numerical simulations presented in section 6.4, we can achieve numerical convergence after about three or four iterations, with results that are identical to those of Problem 1, which we found to be significantly more time consuming.

In the case of infinite SNR at the array ($\mathcal{N} = 0$), Problems 2 and 3 are easier to solve, because of the simple, linear constraints. For a finite SNR, the quadratic
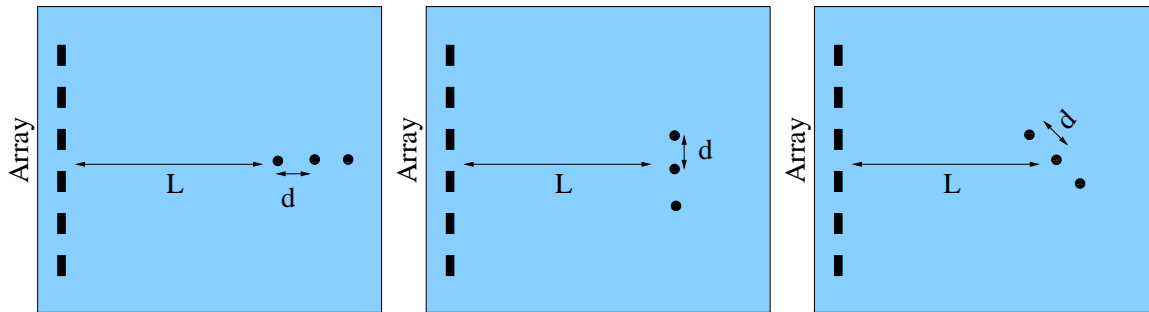
Figure 9. The three array-target configurations.

constraints (6.30) complicate Problem 3 and, in fact, we can have no feasible solution if the level $\mathcal{N}$ is set higher than the maximal power $\mathcal{P}$, which means very low SNR. In the numerical simulations shown in section 6.4, we choose $\mathcal{N}$ as a fraction of the maximal power at the array, so we always have a feasible solution.

6.4. The numerical setup

We illustrate the performance of our optimal illumination approach with numerical simulations for three small targets in a homogeneous medium. It is computationally very expensive to both generate and process the data needed for solving the optimal illumination problem in clutter. The results of these simulations will be presented in a forthcoming paper [12].

We take the three configurations of the targets shown in Figure 9. The array consists of $N_s = 51$ elements, playing the dual role of sources and receivers, placed uniformly in the aperture $a = 11.02\lambda_o$, where λ_o is the central wavelength for the frequency band $0.1 - 1.3$ MHz, at propagation speed $c_o = 1.5$ km/s. The range of the targets is $L = 12.79\lambda_o$ and it is measured with respect to the middle target. The distance between the targets is $d = 1.1\lambda_o$. The background is homogeneous, with sound speed $c_o = 1.5$ km/s and the targets are modeled as identical point scatterers, with scattering amplitude given by a Gaussian function of frequency, centered at $\nu_o = 0.7$ MHz and of standard deviation 0.4 MHz. The impulse response matrix $\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is computed with the Foldy-Lax approach [37, 42, 20] described briefly in the next section.

6.4.1. *The Foldy-Lax model* The scattered field due to M isotropic point scatterers can be modeled by the solution of a system of linear equations in the frequency domain. This is the simplest form of the Foldy-Lax multiple scattering models [37, 42]. The impulse response matrix $\widehat{\Pi}$ is given by

$$\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{m=1}^{M} \widehat{\xi}_m(\omega)\widehat{G}_o(\mathbf{x}_r, \mathbf{y}_m, \omega)\widehat{\mathcal{G}}(\mathbf{y}_m, \mathbf{x}_s, \omega),$$

$$\widehat{\mathcal{G}}(\mathbf{y}_m, \mathbf{x}_s, \omega) = \widehat{G}_o(\mathbf{y}_m, \mathbf{x}_s, \omega) + \sum_{m' \neq m} \widehat{\xi}_{m'}(\omega) \widehat{G}_o(\mathbf{y}_m, \mathbf{y}_{m'}, \omega) \widehat{\mathcal{G}}(\mathbf{y}_{m'}, \mathbf{x}_s, \omega) \ , \quad m = 1, ..., M$$

where $\widehat{\xi}_m(\omega)$ model the frequency dependent scattering amplitudes of the scatterers. As noted above, in our numerical simulations, $M = 3$ and all the scattering amplitudes are the same and equal to a Gaussian function of the frequency.

6.5. Numerical simulations

We consider first, in Figures 11 and 12, the optimal illumination for detection, discussed in section 6.1.1. We show in Figure 11 the Kirchhoff migration images with the illumination given by the monochromatic singular vector $\widehat{f}^{(M)}(\omega_{max})$, corresponding to the maximal eigenvalue $\lambda^M(\omega)$ of the response matrix, at the resonant frequency ω_{max}. This $\omega_{\max}$ depends on the configuration of the scatterers, as we can see from Figure 10, where we display the first three singular values of the response matrix as functions of frequency, for the three different configurations. It is clear from Figure 11 that although this method gives the maximum received power at the array, it does not provide good images.
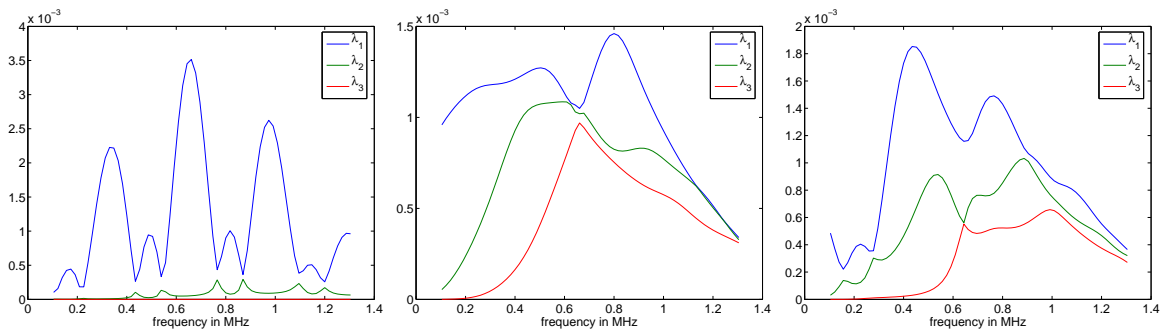
Figure 10. The first three singular values of the response matrix as a function of frequency (in MHz) for the three scatterer configurations. In the first column we show the results for configuration 1, in the second for configuration 2 and in the third for configuration 3.

The second method discussed in section 6.1.1 uses as illuminating signal the singular vector corresponding to the maximal singular value of the response matrix, frequency by frequency in the bandwidth. Since the targets in our three configurations are not well separated, the resulting images are not sharply peaked at a single target. We see from Figure 12 that for the first configuration, where the rank of the response matrix is deficient (1 instead of 3 as expected), all three targets are imaged. For the second configuration, where the singular value curves cross at 0.6MHz, the image has a stronger peak at central target but the other two are also imaged. It is only in the third configuration that the image is peaked at the closest target to the array. While this illumination images better than that for optimal detection, it is not optimal for
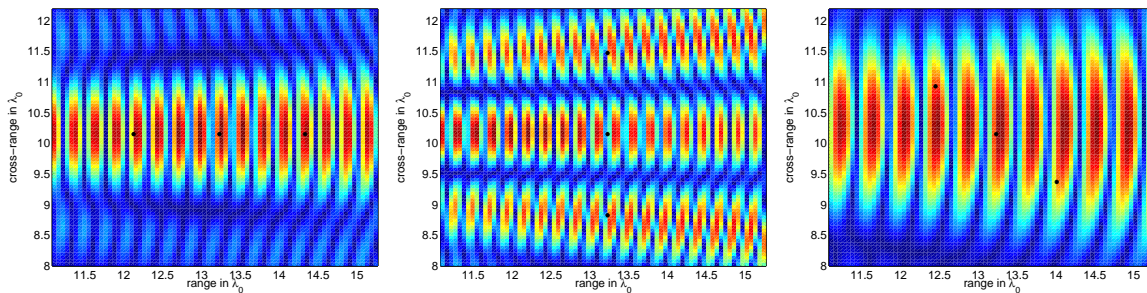
Figure 11. Kirchhoff migration images obtained by illuminating the scatterers with the singular vector corresponding to the maximal eigenvalue $\lambda^M(\omega)$ of the response matrix at frequency ω_{max} for which $\lambda^M(\omega)$ is maximum. In the first column we show the results for configuration 1, in the second for configuration 2 and in the third for configuration 3.

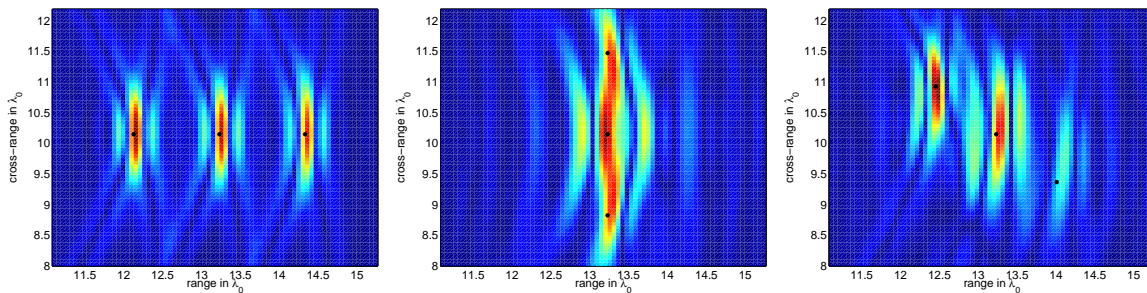imaging, as can be seen by comparing the results in Figure 12 with those in Figures 13-16.

Figure 12. Kirchhoff migration images obtained by illuminating the targets with the singular vector corresponding to the maximal eigenvalue of the response matrix. In the first column we show the results for configuration 1, in the second for configuration 2 and in the third for configuration 3.

We now turn to the numerical simulations using the optimal illumination approach for imaging, described in section 6.3. Since we are in a homogeneous medium, we calculate the image (6.6) for smoothing parameters $\Omega_d = B$ and $X_d(\omega) = a$. We also took $O(\mathbf{w}, \widehat{\mathbf{f}_B})$ as the bounded variation norm of the normalized image in order to be consistent with what is done in section 5. However, since there is no clutter here, the L^1 norm of the image gives almost the same, slightly sharper results.

We implement Problems 2 and 3 by restricting the frequency band to $0.1 - 1.3$ MHz, and we discretize $\widehat{f}_B(\omega - \omega_o)$ by dividing the bandwidth in ten sub-bands of width $\Delta\nu = 0.12$ MHz and by setting $\widehat{f}_B$ constant in each sub-band. Similarly, we group the sources in blocks of three elements and we distribute the weights uniformly in each group. This way we are left with seventeen blocks of weights and ten sub-band values of $\widehat{f}_B$ to determine by iterating between Problems 2 and 3. In all cases we obtain convergence of the iteration in three or four steps and for the first target configuration

we verified that the solution coincides with that given by the computationally expensive Problem 1. We also tested the solution by starting with different initial waveforms in the optimization and we found that the results were always the same.
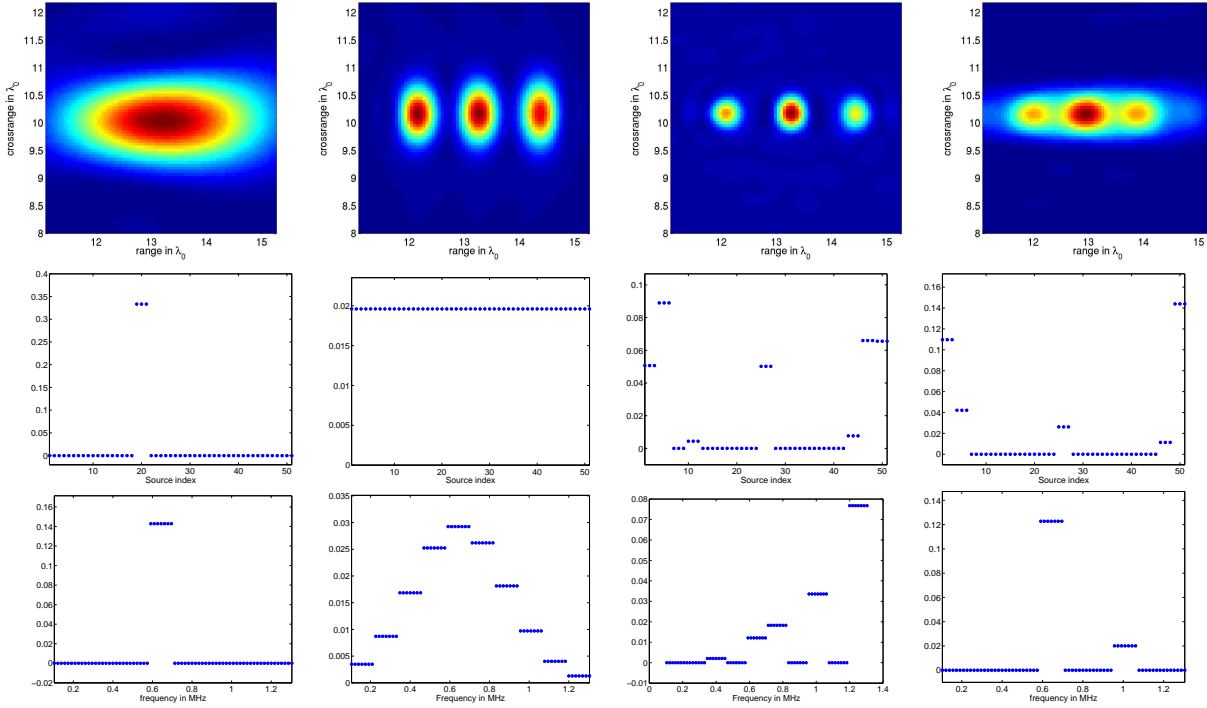

Figure 13. Images with optimal waveform illumination for configuration 1. First column: the image, weights and pulse given by the maximal received power criterion. Second column: the initial image and the initial values of the weights and pulse. Third column: the optimal image, weights and pulse for infinite SNR. Fourth column: the optimal image, weights and pulse for low SNR.

The results are shown in Figures 13 - 16. We show four images in each figure, from left to right: (i) the image given by the maximum received power criterion and constraints (6.10), (6.28), (6.29); (ii) the image at the beginning of the optimization process; (iii) the optimal image for infinite SNR ($\mathcal{N} = 0$) and (iv) the optimal image for low SNR, with $\mathcal{N}$ set to 75% of the maximal power achieved at the array, for any source. The weights are shown in the second row, below the corresponding image, and the Fourier coefficients $\widehat{f}_B$ are shown in the third row.

In these numerical simulations the maximal power can be achieved by illuminating the reflectors from a central block of sources and by sending a narrow-band signal near the resonant frequency of the reflectors. The corresponding images are very blurry. For a fare comparison of these results with the results obtained by illuminating the scatterers with the singular vector $\widehat{f}^{(M)}(\omega_{max})$, we show in Figure 15 Kirchhoff migration images obtained by illuminating the scatterers with the singular vector $\widehat{f}^{(M)}(\omega_{max})$ for all frequencies in a sub-band of width $\Delta\nu = 0.12$ MHz centered at ω_{max} (which corresponds to using $\widehat{f}^{(M)}(\omega_{max})$ constant in one sub-band). Note that for the first
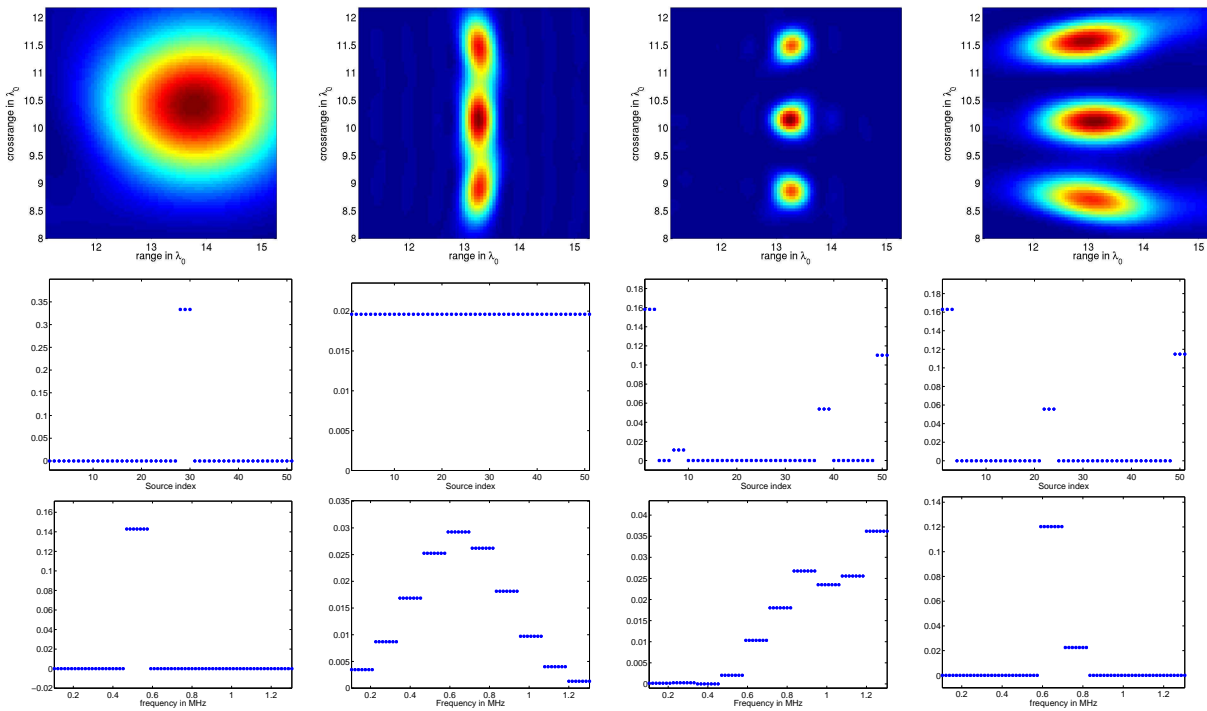
Figure 14. Images with optimal waveform illumination for configuration 2. First column: the image, weights and pulse given by the maximal power criterion. Second column: the initial image and the initial values of the weights and pulse. Third column: the optimal image, weights and pulse for infinite SNR. Fourth column: the optimal image, weights and pulse for low SNR.

and third configurations the two algorithms give very similar results. For the second configuration the image in Figure 15 is better than the one given by our optimization algorithm for maximal power (see Figure 14). For all the configurations, the initial images are better than the maximal power ones because of the broadband initial pulse.
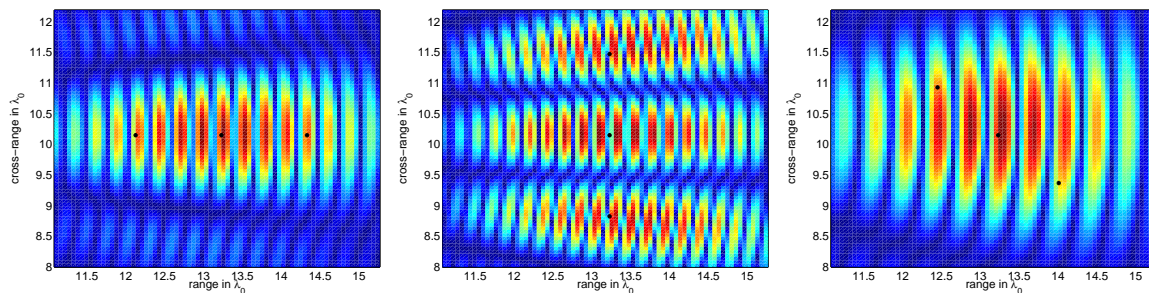

Figure 15. Kirchhoff migration images obtained by illuminating the scatterers with the singular vector $\widehat{f}^{(M)}(\omega_{max})$ for all frequencies in a sub-band of width $\Delta\nu = 0.12$ MHz centered at ω_{max}. In the first column we show the results for configuration 1, in the second for configuration 2 and in the third for configuration 3.

The optimal images for the infinite SNR case are significantly sharper than the initial ones and the optimization tends to put large weights on the sources at the opposite

ends of the array. These are the sources that provide the best illumination of the target configuration. The optimal pulses are also interesting because they display a general increasing trend with the frequency, which is what gives better image resolution [13]. The images become blurrier, however, as we lower the SNR. This is expected because for $\mathcal{N}$ set close to the maximum received power $\mathcal{P}$, we recover the results shown in the first column of Figures 13-16.
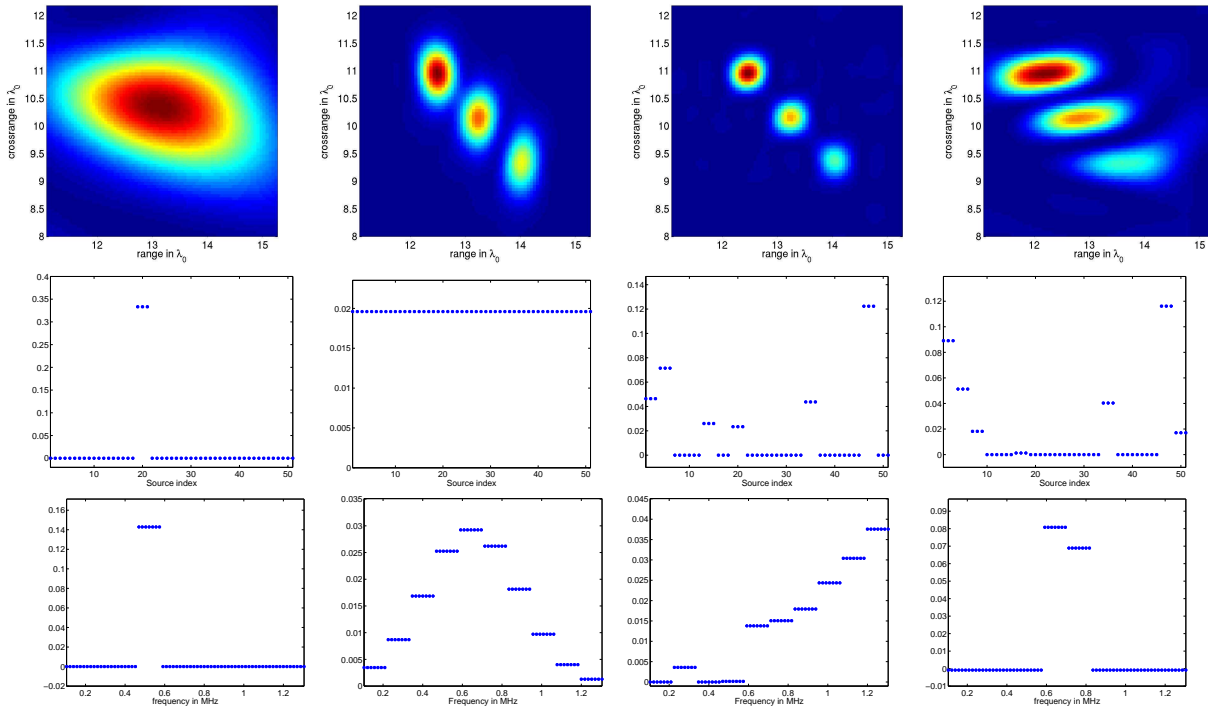
Figure 16. Images with optimal waveform illumination for configuration 3. First column: the image, weights and pulse given by the maximal power criterion. Second column: the initial image and the initial values of the weights and pulse. Third column: the optimal image, weights and pulse for infinite SNR. Fourth column: the optimal image, weights and pulse for low SNR.

7. Summary and conclusions

In this paper we consider array imaging methods that perform well in cluttered media. We also consider optimal illumination schemes that improve the quality of the images.

We introduce first the coherent interferometric functional (CINT) in section 3. It is a travel time migration imaging functional of the array data, that is robust in cluttered media. This is because migration or back propagation is done on local cross correlations of the data traces and not on the traces themselves, which is what Kirchhoff migration does. The cross correlations must be computed locally in space and time so as to enhance information in the traces and reduce the delay spread due to clutter. The

essential complexity of the CINT imaging method is in the determination of the space-time, or space-frequency localization parameters, which are the decoherence parameters introduced in section 3. They depend on the random medium and they are therefore not known. We introduce in section 5 an adaptive method for determining optimally the decoherence parameters. Optimality is based on the quality of the image that CINT itself produces. We want to minimize the speckles in the image that are due to clutter but we also want to limit the amount of blurring that speckle smoothing introduces. This is the trade-off that adaptive CINT quantifies and out of which the decoherence parameters are optimally estimated. In section 5.2 we present the results of extensive numerical simulations that assess the performance of adaptive CINT.

We showed in [17] that the decoherence parameters are not only needed for implementing CINT but they also determine the resolution of the image that CINT produces. This was done in [17] for a class of models of random media for which asymptotic analysis can be used, along with the paraxial approximation. We carried out a similar resolution analysis for CINT in randomly layered media in [19]. In section 4 we show in a self consistent, model independent way, how the decoherence parameters determine the resolution of the images that CINT produces. We do this by expressing CINT as a smoothed Wigner function of the data that has a particular form, from which the resolution limits of the image can be determined. This is done in section 4.2.

In section 6 we consider the problem of how to illuminate the object to be imaged, so that the images that CINT produces have the best resolution. We first present in section 6.1.1 the well-known optimal illumination algorithm for detection, which is the illumination that gives array data with the highest received power. This is also the illumination that results from iterative time reversal. Then, we introduce in section 6.1.2 a new optimal illumination algorithm that is based on the quality of the image that CINT produces, for a given signal to noise ratio. In section 6.5 we show the results of numerical simulations using this new optimal illumination algorithm. Generating data for realistic optimal illumination simulations in clutter is very demanding computationally so in this paper we only show results for media without clutter. When the signal to noise ratio is high, then optimal illumination gives much better images than the optimal illumination for maximum received power at the array. The optimal waveforms that give the best images have an intricate form that is very different from that of the waveforms for maximum received power.

Acknowledgments

References

[1] M. Abramson. *Pattern Search Algorithms for Mixed Variable General Constrained Optimization Problems.* PhD thesis, Rice University, 2002. Software available at: http://en.afit.edu/ENC/Faculty/MAbramson/abramson.html.

[2] S.R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:41–93, 1999.

[3] C. Audet and J. Dennis. A pattern search filter method for nonlinear programming without derivatives. Technical report, Rice University, http://www.caam.rice.edu, 2000, Revised November 2003.

[4] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky. An overview of matched field methods in ocean acoustics. *IEEE J. Ocean Eng.*, 18:401–424, 1993.

[5] 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.*, 324:1281–1286, 1997.

[6] E. Bécache, P. Joly, and C. Tsogka. An analysis of new mixed finite elements for the approximation of wave propagation problems. *SIAM J. Numer. Anal.*, 37:1053–1084, 2000.

[7] J.P. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Comp. Physics.*, 114:185–200, 1994.

[8] G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *J. Math. Phys.*, 26(1):99–108, 1985.

[9] G. Beylkin and R. Burridge. Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion*, 12(1):15–52, 1990.

[10] N. Bleistein, J.K. Cohen, and J.W. Stockwell Jr. *Mathematics of multidimensional seismic imaging, migration, and inversion.* Springer, New York, 2001.

[11] P. Blomgren, G. Papanicolaou, and H. Zhao. Super-resolution in time-reversal acoustics. *Journal of the Acoustical Society of America*, 111:238–248, 2002.

[12] L. Borcea, G. Papanicolaou, and C. Tsogka. Adaptive optimal illumination for imaging in clutter. preprint.

[13] L. Borcea, G. Papanicolaou, and C. Tsogka. Optimal waveform design for active array imaging. preprint.

[14] L. Borcea, G. Papanicolaou, and C. Tsogka. Statistical stability of coherent interferometric imaging in random media. preprint.

[15] L. Borcea, G. Papanicolaou, and C. Tsogka. A resolution study for imaging and time reversal in random media. *Contemporary Math*, 333:63–77, 2003.

[16] L. Borcea, G. Papanicolaou, and C. Tsogka. Theory and applications of time reversal and interferometric imaging. *Inverse Problems*, 19:S134–164, 2003.

[17] L. Borcea, G. Papanicolaou, and C. Tsogka. Interferometric array imaging in clutter. *Inverse Problems*, 21(4):1419–1460, 2005.

[18] L. Borcea, G. Papanicolaou, and C. Tsogka. Coherent interferometric imaging. 2006. to appear in Geophysics.

[19] L. Borcea, G. Papanicolaou, and C. Tsogka. Coherent interferometry in finely layered random media. 2006. to appear in the SIAM Multiscale Modeling and Simulation.

[20] L. Borcea, G. Papanicolaou, C. Tsogka, and J. Berryman. Imaging and time reversal in random media. *Inverse Problems*, 18:1247–1279, 2002.

[21] L. Borcea, G. Papanicolaou, C. Tsogka, and J. Berryman. Statistically stable ultrasonic imaging in random media. *J. Acoust. Soc. Am.*, 112:1509–1522, 2002.

[22] M. Born and E. Wolf. *Principles of optics.* Academic Press, New York, 1970.

[23] H. P. Bucker. Use of calculated sound field and matched-field detection to locate sound sources in shallow water. *J. Acoust. Soc. Am.*, 59:368–373, 1976.

[24] J. Carazzone and W. Symes. Velocity inversion by differential semblance optimization. *Geophysics*, 56, 1991.

[25] T. K. Chan, Y. Kuga, and A. Ishimaru. Experimental studies on circular sar imaging in clutter using angular correlation function technique. *IEEE Trans. Geoscience and Remote Sensing*, 37:2192–2197, 1999.

[26] T.F. Chan and J. Shen. *Image Processing and Analysis*. SIAM, 2005.

[27] M. Cheney, D. Isaacson, and M. Lassas. Optimal acoustic measurements. *SIAM J. Appl. Math.*, 61(5):1628–1647, 2001.

[28] E. Cherkaeva and A. C. Tripp. 67th Annual Meeting of Society of Exploration Geophysicists. In *SEG97 Expanded abstracts*, pages 92–95, Tulsa, OK, 1997.

[29] G. Christakos. *Random Field Models in Earth Sciences*. Dover Publications Inc., 2005.

[30] J. F. Claerbout. *Fundamentals of geophysical data processing : with applications to petroleum prospecting*. CA : Blackwell Scientific Publications, Palo Alto, 1985.

[31] J. F. Claerbout and S. M. Doherty. Downward continuation of moveout-corrected seismograms. *Geophysics*, 37(5):741–768, 1972.

[32] A. Derode, P. Roux, and M. Fink. Robust acoustic time reversal with high-order multiple scattering. *Physical Review Letters*, 75:4206–4209, 1995.

[33] A. Derode, A. Tourin, and M. Fink. Limits of time-reversal focusing through multiple scattering: Long range correlation. *J. Acoust. Soc. Am.*, 107:2987–2998, 2000.

[34] G. L. D'Spain, J. J. Murray, W. S. Hodkiss, N. O. Booth, and P. W. Schey. Mirages in shallow water matched field processing. *J. Acoust. Soc. Am.*, 105:3245–3265, 1999.

[35] E. Dussaud. *Velocity analysis in the presence of uncertainty*. PhD thesis, Rice University, 2006.

[36] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, and M. Tanter. Time-reversed acoustics. *Rep. Prog. Phys.*, 63:1933–1994, 2000.

[37] L. L. Foldy. The multiple scattering of waves. i. general theory of isotropic scattering by randomly distributed scatterers. *Phys. Rev.*, 67:107–119, 1945.

[38] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, Basel, 1984.

[39] Gilles Hennenfent and Felix J. Herrmann. Seismic denoising with unstructured curvelets.

[40] J. L. Krolik. Matched field minimum-variance beamforming in random ocean channel. *J. Acoust. Soc. Am.*, 92:1408–1419, 1992.

[41] K.J. Langenberg, R. Marklein, K. Mayer, T. Krylov, P. Ampha, M. Krause, and D. Streicher. Wavefield inversion in nondestructive testing. In I. M. Pinto, V. Galdi, and L. B. Felsen, editors, *Electromagnetics in a Complex World - Challenges and Perspectives*, pages 277–285. Springer, 2003.

[42] M. Lax. Multiple scattering of waves. *Rev. Modern Phys.*, 23:287–310, 1951.

[43] F. Lin, A. I. Nachman, and R. C. Waag. Quantitative imaging using time-domain eigenfunction method. *J. Acoust. Soc. Am.*, 108:899–911, 2000.

[44] A. Litman, D. Lesselier, and F. Santosa. Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set. *Inverse Problems*, 14:685–706, 1998.

[45] O.I. Lobkis and R.L. Weaver. On the emergence of the Green's function in the correlations of a diffuse field. *Journal of the Acoustical Society of America*, 110(6):3011–3017, 2001.

[46] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[47] T. D. Mast, A. I. Nachman, and R. C. Waag. Focusing and imaging using eigenfunctions of the scattering operator. *J. Acoust. Soc. Amer.*, 102:715–725, 1997.

[48] G. Montaldo, M. Tanter, and M. Fink. Revisting iterative time reversal processing: Application to detection of multiple targets. *J. Acoust. Soc. Am.*, 115(2):776–784, 2004.

[49] C. J. Nolan and W. Symes. Global solution of a linearized inverse problem for the wave equation. *Comm. PDE*, 22(5-6):919–952, 1997.

[50] S. Osher, A. Solé, and L. Vese. Image decomposition and restoration using total variation

minimization and the h^{-1} norm. *SIAM Multiscale Modeling Simulations*, 1(3):349–370, 2003.

[51] G. Papanicolaou, L. Ryzhik, and K. Solna. Statistical stability in time reversal. *SIAM J. Applied Mathematics*, 64(4):1133–1155, 2004.

[52] G. Papanicolaou and K. Solna. Wavelet based estimation of kolmogorov turbulence. In P. Doukhan, G. Oppenmeim, and M. S. Taqqu, editors, *Long-range Dependence: Theory and Applications*, pages 473–505. Birkhäuser, 2003.

[53] C. Prada and M. Fink. Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media. *Wave Motion*, 20:151–163, 1994.

[54] C. Prada, J. L. Thomas, and M. Fink. The iterative time-reversal process: analysis of the convergence. *J. Acoust. Soc. Am.*, 97:62–71, 1995.

[55] M. B. Priestley. *Spectral Analysis and Time Series.* Probability and Mathematical Statistics. Academic Press, 2004.

[56] J. Rickett and J.F. Claerbout. Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring. *The leading edge*, 18:957–960, 1999.

[57] B. Ripley. *Spatial Statistics.* Wiley, 1981.

[58] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.

[59] L. Ryzhik, G. Papanicolaou, and J. B. Keller. Transport equations for elastic and other waves in random media. *Wave Motion*, 24:327–370, 1996.

[60] P. Schultz and J. F. Claerbout. Velocity estimation and downward-continuation by wavefront synthesis. *Geophysics*, 43(4):691–714, 1978.

[61] G. T. Schuster, J. Yu, J. Sheng, and J. Rickett. Interferometric daylight seismic imaging. *Geophysics Journal International*, 157:838–852, 2004.

[62] C.E. Shannon and W. Weaver. *The mathematical theory of communication.* University of Illinois Press, 1963.

[63] R. Snieder, A. Gret, H. Douma, and J. Scales. Coda wave interferometry for estimating nonlinear behavior in seismic velocity. *Science*, 295:2253–2255, 2002.

[64] C. Stolk and M. V. deHoop. Microlocal analysis of seismic inverse scattering in anisotropic elastic media. *Comm. Pure Appl. Math.*, 55(3):261–301, 2002.

[65] C.C. Stolk and W. Symes. Smooth objective functionals for seismic velocity inversion. *Inverse Problems*, 19(1):73–89, 2003.

[66] W. Symes. A differential semblance criterion for inversion of multioffset seismic reflection data. *J. Geophys. Res.*, 88, 1993.

[67] W. Symes. Lecture notes in seismic imaging. Mathematical Geophysics Summer School, Stanford, available at www.trip.caam.rice.edu, 1998.

[68] A. P. E. ten Kroode, D. J. Smit, and A. R. Verdel. A microlocal analysis of migration. *Wave Motion*, 28:149–172, 1998.

[69] B. J. Uscinski. Analytical solution of the fourth-moment equation and interpretation as a set of phase screens. *Journal of the Optical Society of America*, 2:2077–2091, 1985.

[70] M. C. W. van Rossum and Th. M. Nieuwenhuizen. Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion. *Reviews of Modern Physics*, 71:313–371, 1999.