DETECTION AND IMAGING IN STRONGLY BACKSCATTERING RANDOMLY LAYERED MEDIA

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

Abstract. Echoes from small reflectors buried in heavy clutter are weak and difficult to distinguish from the medium backscatter. Detection and imaging with sensor arrays in such media requires filtering out the unwanted backscatter and enhancing the echoes from the reflectors that we wish to locate. We consider a filtering and detection approach based on the singular value decomposition of the local cosine transform of the array response matrix. The algorithm is general and can be used for detection and imaging in heavy clutter, but its analysis depends on the model of the cluttered medium. This paper is concerned with the analysis of the algorithm in finely layered random media. We obtain a detailed characterization of the singular values of the transformed array response matrix and justify the systematic approach of the filtering algorithm for detecting and refining the time windows that contain the echoes that are useful in imaging.

Key words. array imaging, randomly layered media, detection, filtering, local cosine transform.

1. Introduction. We consider an inverse problem for the scalar wave equation, where the goal is to image small reflectors surrounded by heavy clutter, using an array of sensors that probes the medium with pulses and records the echoes. The clutter is due to numerous inhomogeneities that are encountered in applications such as ground penetrating radar and exploration geophysics. Heavy clutter is an issue when the cumulative scattering effect of the inhomogeneities seriously impedes the imaging process. When the coherent echoes from the reflectors, which are useful in imaging, are overwhelmed by the incoherent wave field backscattered then we are in a heavy clutter situation.

Figure 1.1 shows a schematic of the problem setup. The array $\mathcal{A}$ has N sensors that play the dual role of sources and receivers. We denote by $\vec{\mathbf{x}}_s$ and $\vec{\mathbf{x}}_r$ the location of the sources and receivers, although s and r are indexes running from 1 to N. See Appendix A for a detailed explanation of the notation. The sources probe the medium, one at a time, by sending short pulses $f(t)$ from locations $\vec{\mathbf{x}}_s$, and the receivers at $\vec{\mathbf{x}}_r$ record the echoes. The array data is the $N \times N$ response matrix

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

with entries given by the time traces of the scattered acoustic pressure $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$. We define all traces in the same time window $[0, T]$ by reseting the clock every time a source emits a pulse. The inverse problem is to estimate the compact support $\mathcal{S}^\star$ of the reflectors, given the response matrix $\mathbb{P}(t)$.

In weak clutter, the reflectors produce strong coherent echoes and we can image with the Kirchhoff migration method used routinely in radar [20] and seismic imaging [6, 17, 7]

$$\mathcal{J}(\vec{\mathbf{y}}) = \sum_{\vec{\mathbf{x}}_s, \vec{\mathbf{x}}_r \in \mathcal{A}} P\left(\tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}, \vec{\mathbf{x}}_r), \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s\right), \quad \vec{\mathbf{y}} \in \mathcal{S} \supset \mathcal{S}^\star. \tag{1.2}$$

This forms an image at points $\vec{\mathbf{y}}$ in a search domain $\mathcal{S}$ by summing the entries in the response matrix backpropagated numerically to $\vec{\mathbf{y}}$. The backpropagation is done approximately, with round-trip travel times $\tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}, \vec{\mathbf{x}}_r)$ computed, numerically, in a fictitious smooth medium. Migration methods are not useful for

[†]Computational and Applied Mathematics, Rice University, Houston, TX 77005. (rja2@rice.edu and borcea@caam.rice.edu)
[‡]Mathematics, Stanford University, Stanford CA 94305. (papanico@math.stanford.edu)
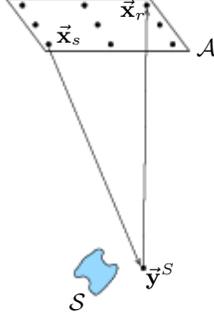[§]Applied Mathematics, University of Crete, GR-71409 Heraklion, Greece (tsogka@tem.uoc.gr)

FIG. 1.1. *General setup for array imaging.*

imaging in stronger clutter, where multiple scattering by the inhomogeneities creates long tailed traces, with incoherent arrivals (coda), observed long before and after the coherent echoes from the reflectors in $\mathcal{S}^\star$. The images are noisy, difficult to interpret and change unpredictably with the realization of the clutter.

The coherent interferometric (CINT) methods introduced and analyzed in [9, 10, 11, 13] image in clutter by backpropagating cross correlations of the traces instead of the traces themselves. The cross-correlations are over time and sensor offset windows whose size is determined by how quickly the waves decorrelate over distances and frequencies. CINT operates on the basis that the cross-correlations are rather efficient at suppressing the coda and enhancing the coherent echoes. This is true in moderate backscattering regimes, where the reflectors in $\mathcal{S}^\star$ are not further than one or two transport mean free paths [29, 30, 33] from the array. In heavy clutter, the backscattered waves dominate the coherent echoes from $\mathcal{S}^\star$ and cannot be suppressed by just taking cross-correlations of the traces. Additional filtering of the clutter effects is needed prior to the image formation with CINT or migration. The question is how to do the filtering, without a-priori information about the location of the reflectors and with no knowledge of the clutter.

The layer annihilators introduced and analyzed in [8], for imaging in randomly layered media, are examples of such filters. They distinguish the layer echoes from the coherent ones based on the dependence of their arrival times on the source-receiver offsets h. Specifically, they use a transformation between the time variable and the depth variable, called the normal move-out map, that defines the arrival time of a primary echo from a layer at depth z, where z is in one to one correspondence with the time t. The arrival times of the echoes from the compact reflectors have a different dependence on the sensor offsets h, and this is why they can be detected and emphasized by the layer annihilator filters in [8]. The one to one correspondence between the depth traveled by the waves in the medium and time exists only in one dimensional media. This is why the filters in [8] cannot be used with general, non-layered clutter.

In this paper we analyze a filtering approach that works in general cluttered media. It is based on the singular value decomposition (SVD) of the local cosine (LC) transform [25, 19] of the response matrix $\mathbb{P}(t)$. The LC transform is used to decompose the recorded traces in orthonormal bases given by smooth time windows modulated by cosine functions [25, Chapter 8]. Such orthonormal bases do not exist for smooth windows modulated by complex exponentials (like in the windowed Fourier transform), as stated by the Balian-Low theorem [25, Theorem 5.6]. The smooth windows in the LC transform avoid artificial discontinuities in the signals and large amplitude coefficients at high frequencies. Moreover, the orthonormal

2

bases lead to fast and stable reconstructions of the traces from filtered (thresholded) coefficients.

Another detection and data filtering method, based on the SVD of the Fourier transformed response matrix in a time window, is considered in [3, 2] for imaging through isotropic, strong clutter. It works with array data that are decorrelated from one receiver to the next, and uses ideas from random matrix theory to assess the medium backscatter.

Our approach (see also [14]) provides an efficient and systematic way for selecting and refining the time windows with detectable coherent echoes, independent of the correlations of the data across the array. Detection is based on the behavior of the singular values of the LC transformed $\mathbb{P}(t)$ over the frequency bandwidth and over time windows that are progressively refined.

The main result of this paper is a detailed theoretical analysis of the behavior of the singular values of the LC transformed matrix $\mathbb{P}(t)$, and a justification of the detection approach in randomly layered media. The algorithm, however, works in general clutter as shown with extensive numerical simulations in [14]. It is only the analysis of the algorithm that depends on the model of the medium. The randomly layered media considered here are of special interest because they, in fact, produce stronger backscattering than general, mostly isotropic clutter. For example, the concept of transport mean free path that quantifies the scattering effect of general clutter does not apply to randomly layered media. This is because of the wave localization phenomenon. Even small wave speed fluctuations in layered media can cause wave localization [34, 30], which means that all of the incident energy is reflected back and does not reach beyond some depth [34, 1, 22].

Our analysis does not address additive, instrument noise. We consider additive noise in the numerical study presented in [14]. In particular, we compare there the effects of strong additive noise and clutter backscatter. We observe that additive noise is much easier to mitigate than clutter effects. We also show that the detection and filtering algorithm based on the LC transform deals equally well with instrument noise and clutter backscatter. Naturally, the distribution of the singular values of the LC transformed response matrix is affected by the strength of the additive noise. If the noise is weak, the spectrum is a small perturbation of the noiseless one analyzed in this paper. The analysis does not apply to strong additive noise regimes, where the distribution of the singular values approaches that of the Wigner quarter circle law [3, 14], characteristic of random symmetric matrices with uncorrelated entries [26, 32].

The paper is organized as follows: We begin in section 2 with the formulation of the problem and the discrete LC transform used in the detection algorithm in [14]. In section 3 we present numerical results. The analysis is in sections 4-6. We end with a summary in section 7.

2. Formulation of the problem. The array imaging setup is in figure 1.1, with a small scatterer buried in a finely layered medium. We refer to appendix A for a detailed explanation of the notation used in the paper. We consider the system of coordinates with z axis orthogonal to the layers, and let the layering be confined to the half space $z < 0$. The array is on the top surface $z = 0$, and we assume for simplicity that it is linear, along the unit vector $\mathbf{e} \in \mathbb{R}^2$,

$$\vec{\mathbf{x}}_r = x_r(\mathbf{e}, 0), \quad 0 \leq x_r \leq a, \quad r = 1, 2, \ldots N, \tag{2.1}$$
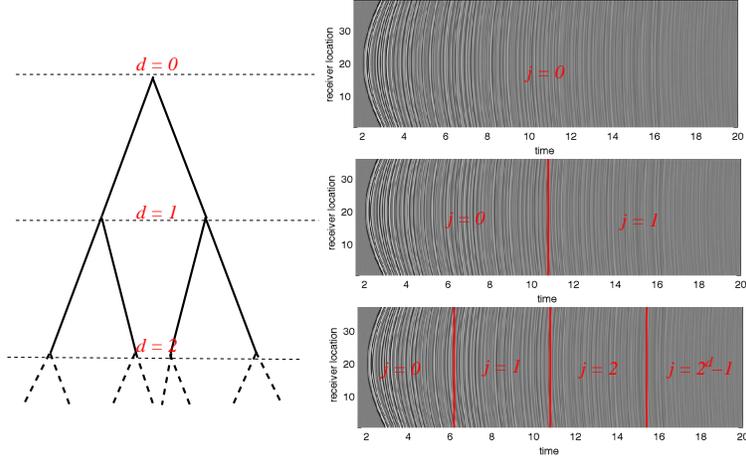
FIG. 2.1. *Illustration of the time windowing segmentations of the array data traces at different tree levels indexed by d. The schematic on the left illustrates the binary tree. On the right we show the segmentation of the data traces in the time windows indexed by j, with $j = 0, 1, \ldots, 2^{d-1}$, at tree levels denoted by d.*

where a is the array aperture. We have a reflector centered at $\vec{\mathbf{y}}^\star = (\mathbf{y}^\star, z^\star)$, with $z^\star < 0$, and we assume that its support $\mathcal{S}$ is very small, point-like.

The problem is twofold: (1) Detect the small reflector from the $N \times N$ response matrix $\mathbb{P}(t)$ dominated by the layer echoes. (2) Filter out the layer echoes so as to image mainly its support $\mathcal{S}$. We address both questions using the LC transform of $\mathbb{P}(t)$ described next.

2.1. The LC transformed response matrix. The LC transform [19, 25] on a binary tree decomposes each trace $P_{rs}(t)$ in an orthonormal basis given by smooth windows χ modulated by cosine functions. At each tree level $d \geq 0$ we have the segmentation*

$$t_j = j\Delta T_d = j\frac{T}{2^d} \tag{2.2}$$

of the time interval $[0, T]$, for $j = 0, 1, \ldots, 2^d$, as illustrated in Figure 2.1. The tree node (j, d) is associated to a space F_j^d generated by the local cosine family

$$\mathcal{F}_j^d = \left\{ \sqrt{\frac{2}{\Delta T_d}} \chi\left(\frac{t - t_j}{\Delta T_d}\right) \cos\left[w_n(t - t_j)\right] \right\}_{n \in \mathbb{N}}, \tag{2.3}$$

with frequencies

$$w_n = \frac{\pi(n + 1/2)}{\Delta T_d}, \quad n \in \mathbb{N}. \tag{2.4}$$

The union of $\mathcal{F}_j^d$ over $j = 0, 1, \ldots 2^d - 1$ gives an orthonormal basis of $L^2[0, T]$. At the next tree level $d + 1$, the spaces F_{2j}^{d+1} and F_{2j+1}^{d+1} are orthogonal, and their sum $F_{2j}^{d+1} \oplus F_{2j+1}^{d+1}$ is equal to the space F_j^d at the parent node (j, d) [25, Proposition 8.7].

The discrete cosine bases used in the numerical simulations are obtained from (2.3) by discretizing the time t at time intervals δ_T that are much smaller than the width of the pulse $f(t)$. The number $N_T = T/\delta_T$

---

*In [25], the partition is done on intervals in $\mathbb{R}$ delimited by half integer points. Here we scale the intervals by time δ_T and absorb the $1/2\delta_T$ shift of the partition points in the t variable.
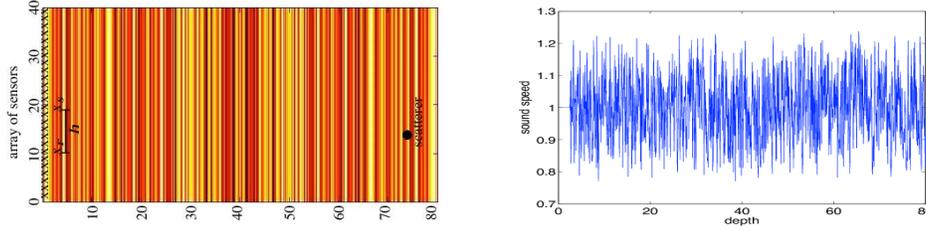
FIG. 3.1. *Left: Illustration of the setup for the numerical simulations. Right: The wave speed (in km/s) used in the simulations vs. depth $(-z)$ scaled by λ_o.*

of time samples is a power of 2, and the frequency index n in (2.4) is restricted by [25, Section 8.5.2] $n = 0, 1, \ldots, N_T/2^d - 1$. Then, the frequencies sample the same bandwidth $w_n \in (0, \pi/\delta_T)$ at all tree levels, but the sampling rate changes with d,

$$w_{n+1} - w_n = \frac{\pi}{\Delta T_d}, \qquad n = 0, 1, \ldots N_t/2^d - 2. \tag{2.5}$$

The LC transform of the response matrix at a given level $d \geq 0$ is given by

$$\widetilde{\mathbb{P}}_{rs}(t_j, w_n) = \int dt \, \mathbb{P}_{rs}(t) \sqrt{\frac{2}{\Delta T_d}} \chi\left(\frac{t - t_j}{\Delta T_d}\right) \cos\left[w_n(t - t_j)\right], \qquad r, s = 1, \ldots N. \tag{2.6}$$

It is a real and symmetric $N \times N$ matrix for all $j = 0, 1, \ldots 2^d - 1$ and $n = 0, 1, \ldots \frac{N_T}{2^d} - 1$. The detection algorithm is based on the behavior of the singular values $\{\sigma_q(t_j, w_n)\}_{q=1,\ldots N}$ of $\widetilde{\mathbb{P}}(t_j, w_n)$, across frequencies $\{w_n\}_{n=0,\ldots N_T/2^d-1}$, and in time windows indexed by t_j [14].

3. Numerical simulations. We present in this section two dimensional numerical results. We begin with the numerical setup and an illustration of the strong clutter impediment to the imaging process. Then, we show the behavior of the singular values of the LC transformed response matrix (2.6) and explain briefly the detection and imaging approach introduced in [14].

3.1. Numerical setup. The schematic of the setup is on the left in figure 3.1. The array has $N = 79$ sources and receivers. We choose the simulation parameters in a regime that is close to that encountered in exploration geophysics [34], but modified so as to articulate better the effects of the filtering algorithm. The sources in the array emit pulses $f(t)$, given by the derivative of a Gaussian, with bandwidth $2.5 - 15.5$Hz, at 6dB. The reference wavelength is $\lambda_o = 100$m calculated at frequency $\omega_o/(2\pi) = 10$Hz, and the array sensors are at distance $\lambda_o/2$ apart.

We generate the response matrix $\mathbb{P}(t)$ by solving with the finite element method described in [4, 5] the acoustic wave equation with a point source at $\vec{\mathbf{x}}_s$. The wave speed is shown in the right plot of figure 3.1. It has the mean value $c = 1$km/s and the fluctuations are generated with random Fourier series, with Gaussian correlation function and correlation length $\ell = 2$m. We have a small reflector buried in the layered medium, at depth $75\lambda_o$ and cross-range $15\lambda_o$. We model it as an acoustic soft scatterer, by setting the pressure P to zero at its boundary $\partial \mathcal{S}$. The support $\mathcal{S}$ is a disk of diameter λ_o.
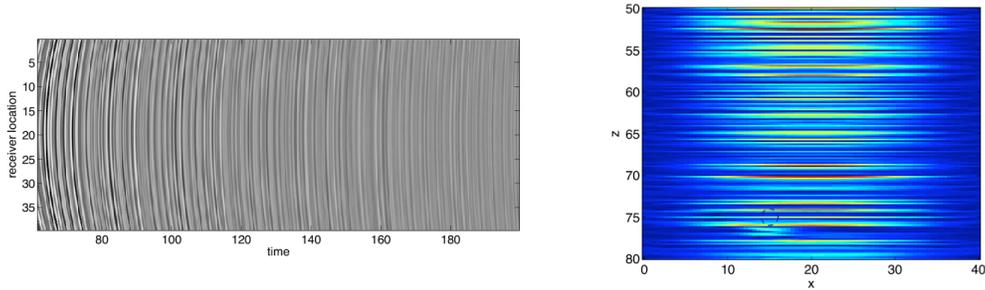
5

FIG. 3.2. *Left: The recorded time traces for the central source illumination. The abscissa is time in hundreds of ms and the ordinate is the receiver location on the array surface, in units of λ_o. Right: The Kirchhoff migration image. The abscissa is cross-range in λ_o and the ordinate is depth in λ_o. The reflector indicated with the black circle is obscured by the layers.*

The cumulative effect of the layers consists of strong backscattered waves that overwhelm the echoes from the small reflector that we wish to image. This can be seen from the time traces plotted on the left in figure 3.2, and from the Kirchhoff migration image shown on the right. The image is computed using (1.2) with round trip travel times $\tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}, \vec{\mathbf{x}}_r) = \left(|\vec{\mathbf{x}}_s - \vec{\mathbf{y}}| + |\vec{\mathbf{y}} - \vec{\mathbf{x}}_r| \right)/c$.

REMARK 3.1. *The detection approach described below does not require any knowledge of the wave speed. However, the image formation uses the smooth part $c(z)$ of the speed, which determines the round trip travel times τ of the coherent echoes. Here we suppose that we know $c(z)$ and we take it constant for simplicity. If $c(z)$ is not known, it must be estimated with additional data processing. The estimation of $c(z)$ in strongly backscattering finely layered media is considered in [8, 22, 1].*

3.2. The SVD of the LC transformed response matrix. We compute the discrete LC transform of $\mathbb{P}(t)$ with the Wavelab 850 Matlab package [21], with windows defined by the option "Sine". The traces are discretized on a uniform time mesh with $N_t = 2^{10}$ points, in the time interval $t + T_o \in [6, 20]$s, with $T_o = 6$s. We take a binary tree with maximum depth 6 (i.e. $0 \leq d \leq 6$).

There is no time windowing at root level $d = 0$, and we plot in figure 3.3 the singular values $\sigma_q(t_0, w_n)$, for $q = 1, \ldots, 10$. If the clutter were weak, the coherent echoes from $\mathcal{S}$ would have dominated the data traces[†] and we would have seen one or two large singular values separated from those associated with the clutter. However, in our case the clutter is strong, and the backscattered field obscures the coherent echoes. We obtain a cluster of singular values that fluctuate rapidly across the bandwidth.

Next, we plot in figure 3.4 the singular values $\sigma_q(t_j, w_n)$ at level $d = 3$ in the tree. Note that when we follow the pattern of $\sigma_q(t_j, w_n)$, starting with the first window at $j = 0$, the singular values remain tightly clustered, uniformly in the bandwidth, until we reach the index $j = 5$. This is the window that contains the coherent echoes from the reflector in $\mathcal{S}$, and it is distinguished from the others by one (arguably two) anomalous singular values at the lower frequencies. By anomaly we mean that at the lower frequencies the largest singular value is well separated from the rest, and its variation with the frequency differs from the others. The anomaly persists in the next window, which contains the reverberations between the small reflector and the layers, and then it disappears in the last window.

_____

[†]See section 6 for the analysis of the SVD of the coherent part of $\widetilde{\mathbb{P}}$.
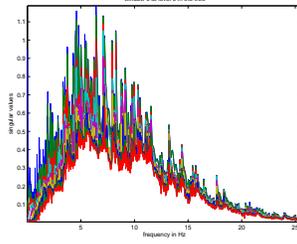
FIG. 3.3. *Singular values* $\sigma_q(t_0, w_n)$, *at the root level* $d = 0$ *in the tree. We plot the largest* 10 *of them, for all the frequency indexes* $n = 0, 1, \ldots, N_T - 1$. *The abscissa is the frequency in Hz.*
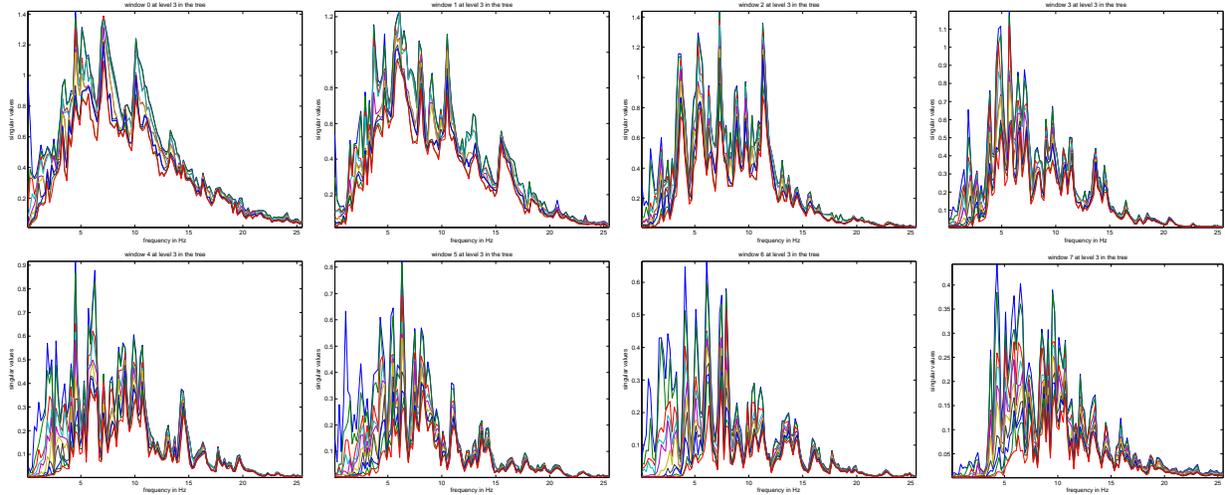

FIG. 3.4. *The singular values* $\sigma_q(t_j, w_n)$ *vs. frequency* w_n, *for* $n = 0, \ldots, N^T/2^3 - 1$, *for* $q = 1, \ldots, 10$, *at tree level* $d = 3$. *Starting from the top left corner, we take* $j = 0, 1, \ldots 2^3 - 1 = 7$.

Now, let us look at the behavior of the singular values as we progress from one tree level to the next. The bottom plots in figure 3.5 show that in the windows that contain pure clutter echoes, the singular values remain clustered uniformly over the bandwidth, and as we progress from one tree level to another. Contrast this with the top plots in figure 3.5, which show the persistent anomalous behavior of the largest singular value, at the lower frequencies, in the windows that contain the echoes from $\mathcal{S}$. Our approach uses such persistent behavior to detect and refine systematically the time windows containing the coherent echoes [14].

3.3. Time window selection. To detect the coherent echoes, we look for anomalies in the clustering of the largest singular values across the frequency band, as described here briefly. More details are in [14]. The basic idea is to introduce a "metric" that quantifies the clustering of the singular values in any given time window. With this metric, the window selection is done in a sequence of steps:

Step1. The setup: Let d be any given level in the binary tree and let $\sigma_q(t_j, w_n)$ be the singular values of the LC transformed response matrix $\widetilde{\mathbb{P}}(t_j, w_n)$, for all $j = 0, 1, \ldots 2^d - 1$ and $n = 0, 1, \ldots N_t/2^d - 1$. Choose the frequency band $\mathcal{B} \subseteq (0, \pi/\delta_T)$ and the number Q of largest singular values to be used in the detection of the time windows with coherent echoes. Let $N_{\mathcal{B}}$ be the number of frequency samples in $\mathcal{B}$. We always choose the lower part of the bandwidth in the detection, because the coherent echoes have more energy at the lower frequencies, as explained in more detail in section 6. Thus, we can index the frequencies w_n in $\mathcal{B}$
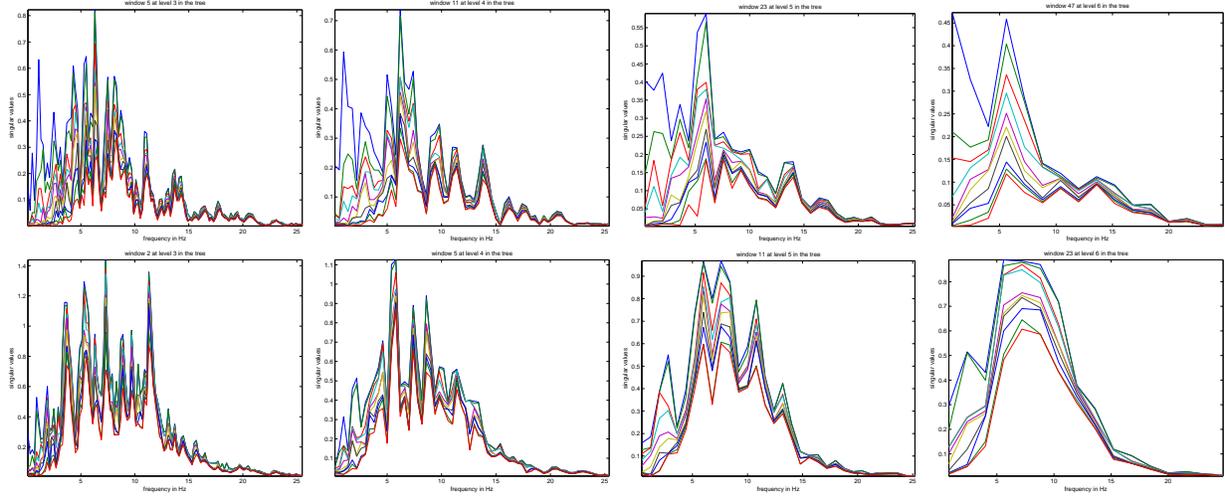
7

FIG. 3.5. *Largest: The largest* 10 *singular values in the windows that contain the coherent echoes from $\mathcal{S}$. Left to right: The 5-th window at level $d = 3$; its child, the 11-th window at level $d = 4$; its child, the 23-rd window at level $d = 5$; its child, the 47-th window at level $d = 6$. Bottom: The largest singular values in windows that contain pure clutter echoes. Left to right: The 2-nd window at level $d = 3$; its child, the 5-th window at level $d = 4$; its child, the 11-th window at level $d = 5$; its child, the 23-rd window at level $d = 6$.*

by $n = 0, 1, \ldots N_{\mathcal{B}} - 1$.

Step 2. The clustering metric: Form the matrix $\mathbb{S}(t_j) \in \mathbb{R}^{Q \times N_{\mathcal{B}}}$ with components

$$\mathbb{S}_{q,n}(t_j) = \frac{\sigma_q(t_j, w_n)}{\max_{0 \leq n' < N_{\mathcal{B}}} \sigma_q(t_j, w'_n)}, \qquad q = 1, \ldots Q, \quad n = 0, 1, \ldots N_{\mathcal{B}} - 1. \tag{3.1}$$

The rows of this matrix are the largest singular values normalized by their maximum in the bandwidth $\mathcal{B}$. Calculate the singular values $\gamma_q(t_j)$ of $\mathbb{S}(t_j) \in \mathbb{R}^{Q \times N_{\mathcal{B}}}$, for $q = 1, \ldots, \min\{Q, N_{\mathcal{B}}\}$, and define the clustering metric

$$m(t_j) = \gamma_2(t_j)/\gamma_1(t_j). \tag{3.2}$$

Step 3. The window selection: If d is the starting tree level, select the time window indexed by $t_{j^\star}$, the maximum of $m(t_j)$. Otherwise, select the time window from the two children of the previously selected window at level $d - 1$. Choose the window with the largest $m(t_j)$. Increase the tree level $d \to d + 1$ and repeat steps 1-3.

In the time windows with the largest singular values clustered uniformly in the frequency band $\mathcal{B}$, the rows of matrix $\mathbb{S}(t_j)$ are almost the same and (3.2) is small. That is to say, $\mathbb{S}(t_j)$ is almost rank one. However, when there are detectable anomalies of a few largest singular values, there is a significant second component in the row space, and (3.2) is large. This is why we use the clustering metric (3.2) to detect the windows with coherent echoes. The detection starts at some tree level d, and it continues at deeper levels $> d$ by looking at the children of the previously selected time windows.

We plot in figure 3.6 the clustering metric $m(t_j)$ for $Q = 10$ and $\mathcal{B}$ given by the lower fifth of the frequency band $(0, \pi/\delta_T)$. We choose this band because it is at the lower frequencies that we can expect to detect the coherent echoes. The medium backscatter dominates the data at the higher frequencies. If we
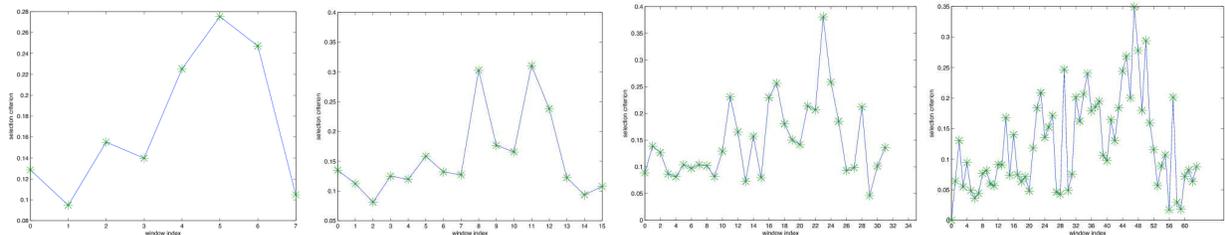
FIG. 3.6. *Plot of $m(t_j)$ as a function of the window index $j = 0, \ldots, 2^d - 1$ for tree levels $d = 3, 4, 5$ and 6.*

start the detection at tree level $d = 3$, our criterion says that we should select the window indexed by t_5. Then, at the next level $d = 4$, we must choose among the two children of this window, indexed by t_{10} and t_{11}. The second plot in figure 3.6 says that we should choose the window indexed by t_{11}. Proceeding this way, we select the window indexed by t_{23} at level $d = 5$ and the window indexed by t_{47} at level $d = 6$. These are precisely the windows considered in the top line of figure 3.5. The detection becomes ambiguous at deeper levels, and so it should because: (1) the widow support at tree level $d = 6$ is already small, comparable to the pulse width and (2) we have too few frequency samples in the bandwidth to carry on the detection.

3.4. Data filtering for imaging. Our approach filters the traces at a given tree level d, by setting to zero the LC coefficients $\widetilde{\mathbb{P}}(t_j, w_n)$ in all the windows t_j, except those where the largest singular values exhibit an anomalous behavior over the frequencies [14]. The filtered traces $\mathbb{Q}\mathbb{P}(t)$ are then reconstructed from these LC coefficients. We show in the top row of figure 3.7 the Kirchhoff migration image formed with such filtered traces, at levels $d = 3, \ldots, 6$. We also show in the bottom row the images obtained from the further filtered traces $\mathbb{Q}\widetilde{\mathbb{Q}}\mathbb{P}(t)$. The additional filter $\widetilde{\mathbb{Q}}$ amounts to projecting $\widetilde{\mathbb{P}}(t_j, w_n)$ on the space of low rank matrices with singular vectors given by the leading ones of $\widetilde{\mathbb{P}}(t_j, w_n)$, at the frequencies w_n in the lower fifth of the bandwidth. At the higher frequencies we set $\widetilde{\mathbb{P}}(t_j, w_n) = 0$.

We note that as we refine the time windows, we localize better and better the small reflector that was obscured by the layers in the image in figure 3.2.

4. Analysis setup. Our goal in the analysis is to explain qualitatively the behavior of the singular values of the LC transformed response matrix illustrated in section 3.2, which is at the core of our detection and data filtering approach. Here we give the mathematical model of the response matrix and the asymptotic regime of separation of scales used in the analysis. The SVD analysis of the LC transformed matrix is in sections 5 and 6.

4.1. Mathematical model of the LC transformed array data. The model of the array data $\mathbb{P}_{rs}(t) = P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ is based on the scalar wave equation

$$\frac{1}{V^2(\vec{\mathbf{x}})} \frac{\partial^2 P(t, \vec{\mathbf{x}}, \vec{\mathbf{x}}_s)}{\partial t^2} - \Delta P(t, \vec{\mathbf{x}}, \vec{\mathbf{x}}_s) = f(t) \frac{\partial}{\partial z} \delta(\vec{\mathbf{x}} - \vec{\mathbf{x}}_s), \quad t > 0,$$
$$P(t, \vec{\mathbf{x}}, \vec{\mathbf{x}}_s) \equiv 0, \quad t < 0, \tag{4.1}$$

with a point source at $\vec{\mathbf{x}}_s = (x_s \mathbf{e}, 0)$ emitting downward the pulse $f(t)$. The wave speed $V(\vec{\mathbf{x}})$ satisfies

$$\frac{1}{V^2(\vec{\mathbf{x}})} = \begin{cases} 1/v^2(z) + \nu(\vec{\mathbf{x}} - \vec{\mathbf{y}}^\star), & z < 0 \\ 1/c^2, & z \geq 0, \end{cases} \tag{4.2}$$
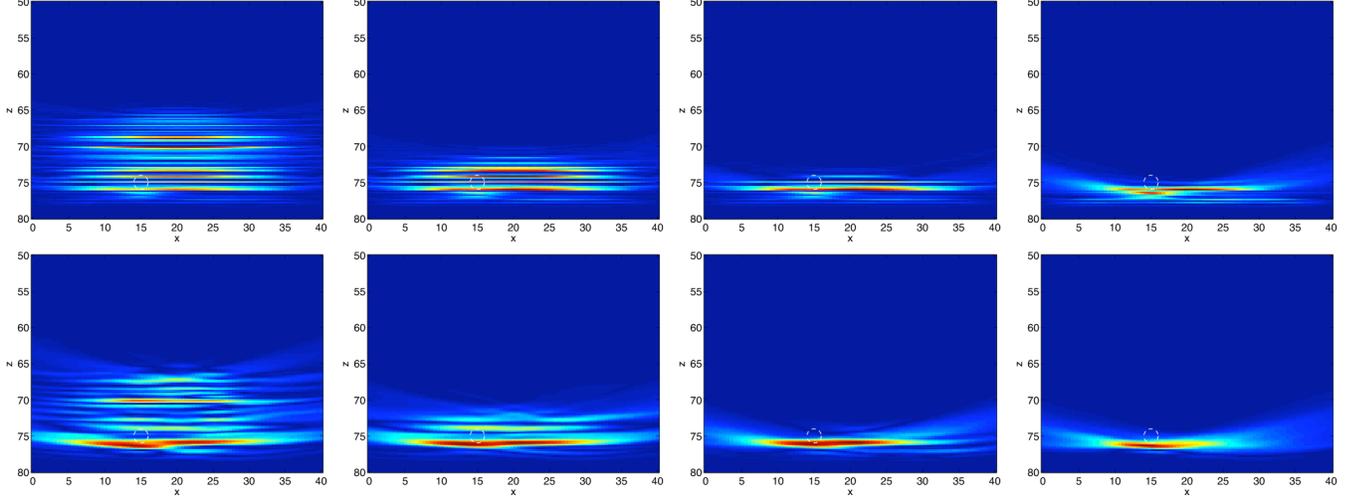
9

FIG. 3.7. *Top: Images obtained with traces reconstructed from the LC coefficients of $\mathbb{P}(t)$ in the selected windows. From left to right: $j_\star^3 = 5$, $j_\star^4 = 11$, $j_\star^5 = 23$ and $j_\star^6 = 47$. Bottom: The same plots as above, but for the traces reconstructed from the LC coefficients of $\mathbb{P}(t)$ projected on the low rank dominant subspace of matrices at he lower frequencies.*

for $\vec{\mathbf{x}} = (\mathbf{x}, z)$. Here $\nu(\vec{\mathbf{x}} - \vec{\mathbf{y}}^\star)$ models the reflectivity of the small reflector at $\vec{\mathbf{y}}^\star = (\mathbf{y}^\star, z^\star)$, and $v(z)$ is the wave speed in the layered medium. It has a smooth part c, which determines the travel times, and a rough part that scatters. We take c constant for simplicity and model the fluctuations with a random process μ,

$$\frac{1}{v^2(z)} = \frac{1}{c^2}\left[1 + \sigma\mu\left(\frac{z}{\ell}\right)\right]. \tag{4.3}$$

Here μ is a dimensionless, zero mean random function with integrable correlation function. The process is normalized so that

$$\int_{-\infty}^{\infty} dz\, E\left\{\mu(0)\mu\left(\frac{z}{\ell}\right)\right\} = \ell, \tag{4.4}$$

with ℓ the correlation length of the fluctuations, and we control its intensity

$$E\left\{\left[\sigma\mu\left(\frac{z}{\ell}\right)\right]^2\right\} = \sigma^2, \tag{4.5}$$

by adjusting the parameter σ.

4.1.1. Model of the array data. The pressure $\mathbb{P}_{rs}(t)$ recorded at the array consists of the direct arrival of the waves from $\vec{\mathbf{x}}_s$ to $\vec{\mathbf{x}}_r$, and the scattered field. We assume hereafter that the direct arrival has been removed by tapering the data for $t \leq |\vec{\mathbf{x}}_r - \vec{\mathbf{x}}_s|/c$. The scattered field observed at times

$$t < \tau^{\mathcal{C}} = \min_{r,s=1,\ldots N} \tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_r)$$

consists of the unwanted echoes $\mathcal{N}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ from the layers,

$$\mathbb{P}_{rs}(t) = \mathcal{N}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s), \qquad t < \tau^{\mathcal{C}}. \tag{4.6}$$

Around time $\tau^{\mathcal{C}}$,

$$\mathbb{P}_{rs}(t) \approx \mathcal{N}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s) + \mathcal{C}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s), \qquad t \approx \tau^{\mathcal{C}}, \tag{4.7}$$

10

where $\mathcal{C}(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)$ is the transmitted field from the source at $\vec{\mathbf{x}}_s$ to the reflector at $\vec{\mathbf{y}}^\star$ and then back to the array at $\vec{\mathbf{x}}_r$. We refer to it as the coherent field, although it is random [22, Chapter 8]. If this field were strong enough, the Kirchhoff migration imaging method would image the small reflector well, without any data filtering [12]. Here $\mathcal{C}$ is overwhelmed by the medium backscatter $\mathcal{N}$.

At later times than $\tau^{\mathcal{C}}$, the model of $\mathbb{P}_{rs}(t)$ is more complicated than (4.7), because it includes reverberations between the source and the layered medium. However, for the analysis in this paper it is sufficient to look at the two cases $t < \tau^{\mathcal{C}}$ and $t \approx \tau^{\mathcal{C}}$.

4.1.2. Model of the layer echoes. The incoherent backscattered field $\mathcal{N}(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)$ can be written as a superposition of up going plane waves

$$\mathcal{N}(t,\vec{\mathbf{x}}_r;\vec{\mathbf{x}}_s) = \frac{1}{2(2\pi)^3}\int d\omega\,\omega^2\hat{f}(\omega)\int_{K\leq 1/c} d\mathbf{K}\,\mathcal{R}(\omega,K,0)\,e^{-i\omega t+i\omega\mathbf{K}\cdot(\mathbf{x}_r-\mathbf{x}_s)}. \tag{4.8}$$

This amounts to Fourier transforming the wave equation in t and $\mathbf{x}\in\mathbb{R}^2$, and letting ω and $\omega\mathbf{K}$ be the dual variables to t and $\mathbf{x}$, respectively. We obtain a one dimensional Helmholtz equation for plane waves traveling at horizontal slowness $\mathbf{K}$ and vertical speed $c(K)=c/\sqrt{1-(cK)^2}$, where $K=|\mathbf{K}|\leq 1/c$. The evanescent waves with $K>1/c$ are neglected in (4.8).

The reflection coefficient $\mathcal{R}(\omega,K,z)$ is the ratio of the complex valued, up and down going wave amplitudes at $z\in[-L,0]$. Here $-L$ is a large enough depth that cannot influence the array data up to the time of observation. The up and down going amplitudes solve a system of first order ordinary differential equations in z, with down going amplitude at $z=0$ determined by the source excitation, and zero upgoing amplitude at $z=-L$. Although these amplitudes depend on the whole medium in $[-L,0]$, the reflection coefficient $\mathcal{R}(\omega,K,z)$ depends only on the medium below z, as if the top interval $[z,0]$ has been stripped away[22, Chapter 9]. Explicitly, $\mathcal{R}(\omega,K,z)$ is a complex valued random field satisfying the Riccati equations

$$\frac{\partial}{\partial z}\mathcal{R}(\omega,K,z) = \frac{-i\omega\sigma\mu\,(z/\ell)\,c(K)}{2c^2}\left\{e^{-2i\omega(L+z)/c(K)}-2\mathcal{R}(\omega,K,z)+e^{2i\omega(L+z)/c(K)}\mathcal{R}^2(\omega,K,z)\right\}, \quad z>-L,$$
$$\mathcal{R}(\omega,K,-L)=0. \tag{4.9}$$

In model (4.8) of the layer echoes we evaluate the reflection coefficient at the measurement surface $z=0$.

4.1.3. Model of the coherent echoes. The coherent echoes $\mathcal{C}(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s)$ can be modeled by

$$\mathcal{C}(t,\vec{\mathbf{x}}_r,\vec{\mathbf{x}}_s) \approx -\frac{\partial^2 P^i(t,\vec{\mathbf{y}}^\star,\vec{\mathbf{x}}_s)}{\partial t^2}\star_t G(t,\vec{\mathbf{x}}_r,\vec{\mathbf{y}}^\star), \tag{4.10}$$

where $\star_t$ denotes time convolution, $P^i(t,\vec{\mathbf{y}}^\star,\vec{\mathbf{x}}_s)$ is the "incident" pressure field impinging on the reflector at $\vec{\mathbf{y}}^\star$, and G is the causal Green's function of the wave equation in the layered medium. If there were no random fluctuations, the incident pressure field would be

$$P_o^i(t,\vec{\mathbf{y}}^\star,\vec{\mathbf{x}}_s) = -\left.\frac{\partial}{\partial z}\frac{f\left(t-|\vec{\mathbf{y}}-\vec{\mathbf{x}}_s|/c\right)}{4\pi|\vec{\mathbf{y}}-\vec{\mathbf{x}}_s|}\right|_{z=z^\star} \approx \frac{f'(t-|\vec{\mathbf{y}}^\star-\vec{\mathbf{x}}_s|/c)z^\star}{4\pi c|\vec{\mathbf{y}}^\star-\vec{\mathbf{x}}_s|^2}, \tag{4.11}$$

where we let $\vec{\mathbf{y}}^\star=(\mathbf{y}^\star,z^\star)$, and assumed a short pulse $f(t)$ to make the approximation. We would observe the pulse f' centered at travel time $|\vec{\mathbf{y}}^\star-\vec{\mathbf{x}}_s|/c$, and the amplitude change due to geometrical spreading. In the random medium we have

$$P^i(t,\vec{\mathbf{y}}^\star,\vec{\mathbf{x}}_s) \approx P_{\mathrm{ODA}}^i(t,\vec{\mathbf{y}}^\star,\vec{\mathbf{x}}_s)+\dots \tag{4.12}$$

11

with wave front P_{ODA}^i modeled by the O'Doherty Anstey (ODA) theory [27, 16, 1, 18, 31, 22]. The weaker, incoherent reverberations from the layers are denoted by "...". The ODA theory says that the transmitted field through the random medium is given by [22, 1, 27, 18, 31]

$$P_{\mathrm{ODA}}^i(t, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s) \approx \frac{\left(f \star_t \mathcal{K}_{\mathrm{ODA}}\right)'\left(t - |\vec{\mathbf{y}}^\star - \vec{\mathbf{x}}_s|/c - \delta\tau(\vec{\mathbf{y}}, \vec{\mathbf{x}}_s)\right) z^\star}{4\pi c |\vec{\mathbf{y}}^\star - \vec{\mathbf{x}}_s|^2}. \tag{4.13}$$

We have pulse spreading due to the convolution of f with the Gaussian kernel

$$\mathcal{K}_{\mathrm{ODA}}(t, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s) = \frac{\sin\theta(\vec{\mathbf{x}}_s)}{\sqrt{2\pi}t_{\mathrm{ps}}}e^{-\frac{t^2 \sin^2\theta(\vec{\mathbf{x}}_s)}{2t_{\mathrm{ps}}^2}}, \qquad \sin\theta(\vec{\mathbf{x}}_s) = \frac{|z^\star|}{|\vec{\mathbf{y}}^\star - \vec{\mathbf{x}}_s|}, \tag{4.14}$$

and a small random arrival time shift $\delta\tau(\vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)$. Here small means that $\delta\tau$ is comparable to the pulse width. The spread is proportional to t_{ps}, a parameter with units of time that depends on $z^\star$ and the correlation length ℓ, and it is more pronounced for waves propagating at shallow angles $\theta(\vec{\mathbf{x}}_s)$.

In the frequency domain, (4.13) becomes

$$P_{\mathrm{ODA}}^i(t, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s) \approx \int \frac{d\omega}{2\pi} i\omega \widehat{f}(\omega) \frac{\widehat{\mathcal{K}}_{\mathrm{ODA}}(\omega, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)z^\star}{4\pi c |\vec{\mathbf{y}}^\star - \vec{\mathbf{x}}_s|^2} e^{-i\omega[t - |\vec{\mathbf{y}}^\star - \vec{\mathbf{x}}_s|/c - \delta\tau(\vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)]}$$

$$= \int \frac{d\omega}{2\pi} \widehat{f}(\omega)\widehat{\mathcal{G}}_{\mathrm{ODA}}(\omega, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)e^{-i\omega t}, \tag{4.15}$$

where $\widehat{\mathcal{G}}_{\mathrm{ODA}}$ is like a Green's function. It gives approximately the transmitted wave field at $\vec{\mathbf{y}}^\star$ when the source at $\vec{\mathbf{x}}_s$ emits an impulse $\delta(t)$. The second factor in the convolution in (4.10), which models transmission from $\vec{\mathbf{y}}^\star$ to $\vec{\mathbf{x}}_r$ is similar to (4.15), by reciprocity. We obtain the following model of the coherent echoes

$$\mathcal{C}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s) \approx \int \frac{d\omega}{2\pi} \omega^2 \widehat{f}(\omega)\widehat{\mathcal{G}}_{\mathrm{ODA}}(\omega, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)\widehat{\mathcal{G}}_{\mathrm{ODA}}(\omega, \vec{\mathbf{x}}_r, \vec{\mathbf{y}}^\star)e^{-i\omega t}. \tag{4.16}$$

4.2. Scaling and the asymptotic regime. Our theoretical study of the spectral decomposition of the LC transformed $\mathbb{P}(t)$ is in an asymptotic regime of separation of scales that we now describe. It may be motivated by applications in exploration geophysics [34], where the waves penetrate to depths $L = 5 - 10\mathrm{km}$ that are much larger than the reference wavelength $\lambda_o \sim 100\mathrm{m}$ of the probing pulses, and the medium has strong fluctuations on a much shorter scale $\ell = 2 - 3\mathrm{m}$. Such a regime has been used in the numerical simulations in section 3.

Let L be the reference, order one length scale. This implies that the time window $[0, T]$ over which the data $\mathbb{P}(t)$ is recorded is order one, as well. To model the separation of scales, we introduce the small parameter $\varepsilon \ll 1$ given by the ratio of the pulse width and T. Specifically, we let $f^\varepsilon(t)$ be the scaled pulse

$$f^\varepsilon(t) = \varepsilon^{1/2}f\left(\frac{t}{\varepsilon}\right), \tag{4.17}$$

with Fourier transform

$$\widehat{f^\varepsilon}\left(\frac{\omega}{\varepsilon}\right) = \varepsilon^{1/2}\int dt f^\varepsilon(t)e^{\frac{i\omega t}{\varepsilon}} = \varepsilon^{3/2}\int \frac{dt}{\varepsilon}f\left(\frac{t}{\varepsilon}\right)e^{\frac{i\omega t}{\varepsilon}} = \varepsilon^{3/2}\widehat{f}(\omega). \tag{4.18}$$

Here $f(t)$ is the carrier pulse and the scaling says that $f^\varepsilon(t)$ is supported at high frequencies of order ε^{-1}. Equivalently, the reference wavelength λ_o^ε satisfies $\lambda_o^\varepsilon \sim \varepsilon L$. The correlation length is much smaller than λ_o^ε. We rename it ℓ^ε and we assume that it satisfies

$$\frac{\ell^\varepsilon}{\lambda_o^\varepsilon} \sim \frac{\lambda_o^\varepsilon}{L} \sim \varepsilon \ll 1. \tag{4.19}$$

12

The strength of the fluctuations is $\sigma \sim 1$.

It remains to specify the aperture a and the distribution of the sensors in the linear array. We take

$$x_r = r\varepsilon\Delta_x, \qquad r = 1, \ldots N, \tag{4.20}$$

with spacing $\varepsilon\Delta_x \sim \lambda_o^\varepsilon$, and let a be order one, so that the number N of sensors is large,

$$N = \frac{a}{\varepsilon\Delta_x} \sim \varepsilon^{-1} \gg 1. \tag{4.21}$$

The asymptotic regime (4.19) has been used extensively in studies of waves in randomly layered media [1, 22]. It is interesting because it considers strong fluctuations that arise in important applications. Waves penetrate to large depths in media with strong fluctuations when they interact weakly with the layers, over distances comparable to the wavelength (i.e., when $\ell^\varepsilon \ll \lambda_o^\varepsilon$). We take $\ell^\varepsilon \sim \varepsilon\lambda_o^\varepsilon$ so that over the distance $L \gg \lambda_o^\varepsilon$ the cumulative effect of the layers gives significant echoes at the array. In particular, by scaling the amplitude of the pulse with $\varepsilon^{1/2}$ in (4.18) we obtain an order one intensity of the backscattered waves [22, Section 14.3].

There are other scaling regimes that give significant backscattering. For example, the theory extends almost identically to the *weakly heterogeneous* [22, Section 18.1] regime with $\lambda_o^\varepsilon \sim \ell^\varepsilon \sim \varepsilon L$ and $\sigma \ll 1$. There is only one essential difference. The waves interact more efficiently with the fluctuations in the weakly heterogeneous regime, because $\lambda_o^\varepsilon \sim \ell^\varepsilon$, and the asymptotic results depend on the specific correlation function of the random fluctuations [22]. In our case, the waves do not see the small scales in detail because $\lambda_o^\varepsilon \gg \ell^\varepsilon$, and the fluctuations take the canonical form of *white noise* as $\varepsilon \to 0$, independent of the detailed structure of the random process μ.

4.3. Statistics of the reflection coefficient. Our analysis in section 5 is based on the statistics of the reflection coefficients

$$\mathcal{R}^\varepsilon(\omega, K, 0) = \mathcal{R}\left(\frac{\omega}{\varepsilon}, K, 0\right) \tag{4.22}$$

in the asymptotic limit $\varepsilon \to 0$, which we now summarize from [22, Section 14.3]. First, let us note from (4.9), with ω replaced by ω/ε, that $\mathcal{R}^\varepsilon(\omega, K, 0)$ satisfies a Riccati equation driven by the random process

$$\mu^\varepsilon(z) = \frac{\sigma}{\varepsilon}\mu\left(\frac{z}{\ell^\varepsilon}\right) = \frac{\sigma}{\varepsilon}\mu\left(\frac{z}{(\varepsilon/\sigma)^2 l}\right), \tag{4.23}$$

with rescaled correlation length l of order one. In the limit $\varepsilon \to 0$, we have by the central limit theorem that

$$\int_{-L}^z \mu^\varepsilon(z')dz' \to \sqrt{l}W(z), \tag{4.24}$$

where $W(z)$ is standard Brownian motion and the convergence is weak, in distribution. As we already mentioned in section 4.2, the fluctuations of the wave speed take the canonical form of white noise as $\varepsilon \to 0$, and the statistics of the reflection coefficients are analyzed using the white noise (diffusion) limit Theorem 6.5 in [22]. The relevant results for our purpose are summarized from [22, Section 14.3] in the following lemma.

13

LEMMA 4.1. *The reflection coefficients $\mathcal{R}^\varepsilon(\omega, K, 0)$ are correlated only if the frequencies and slowness moduli are close to each other, at order ε. Moreover,*

$$E\left\{\left[\mathcal{R}^\varepsilon\left(\omega + \varepsilon h/2, K + \varepsilon k/2\right)\right]^p \left[\overline{\mathcal{R}^\varepsilon\left(\omega - \varepsilon h/2, K - \varepsilon k/2\right)}\right]^q\right\} \to \delta_{pq} \int_0^\infty ds\, V_p(\omega, K, s) e^{is\left[h\left[1 - (Kc)^2\right] - \omega K k c^2\right]} \tag{4.25}$$

as $\varepsilon \to 0$, where δ_{pq} is the Kronecker delta symbol, the bar denotes complex conjugate, and

$$V_p(\omega, K, s) = \frac{2pc(K)\left[c(K)s/L_{loc}(\omega)\right]^{p-1}}{L_{loc}(\omega)\left[2 + c(K)s/L_{loc}(\omega)\right]^{p+1}}, \qquad L_{loc}(\omega) = \frac{4c^2}{\omega^2 l}. \tag{4.26}$$

Here $L_{\mathrm{loc}}(\omega)$ is the localization length [22, Section 9.2]. It coincides in layered media with the scale of exponential decay of the coherent part of the wave field, modeled with ODA [22, Section 14.2]. Specifically, the pulse spread parameter t_{ps} in (4.14) satisfies

$$\omega^2 t_{\mathrm{ps}}^2 = \frac{|z^\star|}{L_{\mathrm{loc}}(\omega)}. \tag{4.27}$$

REMARK 4.2. *We have the explicit expressions (4.25)-(4.26) of the moments of $\mathcal{R}^\varepsilon$ because we have assumed a constant background speed c. This is the simplification that we alluded at in section 4.1. The results extend to variable backgrounds, with the complication that the right hand side in (4.25) is determined by the solution of an infinite coupled system of transport equations [22, Section 14.3].*

4.4. Scaling in the LC transform. In the analysis we consider a depth d in the LC transform binary tree that gives time intervals

$$\Delta T_d = \frac{T}{2^d} = \varepsilon^{1-\gamma} \Delta_T, \qquad \gamma \in (0, 1). \tag{4.28}$$

Here Δ_T is order one, and $\gamma = 1$ means basically no time segmentation ($\Delta T_d \sim T$). The other extreme is $\gamma = 0$, where the windows are as narrow as the pulse ($\Delta T_d = \varepsilon \Delta_T$). We take $\gamma \in (0, 1)$ to ensure that the time windows are wider than the pulse, and still have time segmentation taken into account in the analysis.

Since the emitted pulse has high frequencies of order ε^{-1}, it is not difficult to show that the LC transform of the response matrix $\mathbb{P}(t)$ is supported at order ε^{-1} frequencies, as well. Therefore, we rewrite the LC coefficients (2.6) as

$$\widetilde{\mathbb{P}}_{rs}^\varepsilon(t_j, w) = \int dt\, \mathbb{P}_{rs}(t) \sqrt{\frac{2}{\varepsilon^{1-\gamma}\Delta_T}} \chi\left(\frac{t - t_j}{\varepsilon^{1-\gamma}\Delta_T}\right) \cos\left[\frac{w}{\varepsilon}(t - t_j)\right], \qquad r, s = 1, \ldots N, \tag{4.29}$$

where t_j indicates the location of the window and we drop the index n of the frequencies for simplicity of notation. The expression

$$\widetilde{\mathbb{P}}_{rs}^\varepsilon(t_j, w) = \varepsilon^{\frac{\gamma-1}{2}} \sqrt{2\Delta_T}\, \Re\left\{\int \frac{dh}{2\pi} \widehat{\mathbb{P}}_{rs}\left(\frac{w - \varepsilon^\gamma h}{\varepsilon}\right) \widehat{\chi}(h\Delta_T) e^{-\frac{i(w - \varepsilon^\gamma h)t_j}{\varepsilon}}\right\} \tag{4.30}$$

follows by direct calculation from (4.29), with $\Re$ denoting the real part.

5. Spectral analysis prior to the coherent arrivals. Prior to the coherent arrival time $\tau^{\mathcal{C}}$, the response matrix consists of pure layer echoes, and its LC transform follows from (4.6), (4.8) and (4.30). It is a real symmetric Toeplitz matrix

$$\widetilde{\mathbb{P}}_{rs}^\varepsilon(t_j, w) = \widetilde{\mathbb{D}}_{r-s}^\varepsilon(t_j, w), \tag{5.1}$$

14

defined by sequence

$$\widetilde{\mathbb{D}}_q^\varepsilon(t_j, w) \approx \frac{\varepsilon^{\frac{\gamma}{2}-1}}{(2\pi)^3} \sqrt{\frac{\Delta_T}{2}} \Re \left\{ w^2 \hat{f}(w) \int dh \, \widehat{\chi}(h\Delta_T) \int_{K \leq 1/c} d\mathbf{K} \, \mathcal{R}^\varepsilon(w - \varepsilon^\gamma h, K, 0) \times \right.$$
$$\left. \exp\left[-\frac{i(w - \varepsilon^\gamma h)t_j}{\varepsilon} + iq\Delta_x(w - \varepsilon^\gamma h)\mathbf{K} \cdot \mathbf{e}\right] \right\}, \qquad |q| \leq N - 1. \tag{5.2}$$

Since in our scaling $N \sim \varepsilon^{-1} \gg 1$, we can use the results in [24, 23, 15] to characterize its spectrum in terms of its symbol $Q_{t_j}^\varepsilon(\xi, w)$ given by

$$Q_{t_j}^\varepsilon(\xi, w) = \sum_{q=-N+1}^{N-1} e^{iq\xi} \widetilde{\mathbb{D}}_q^\varepsilon(t_j, w), \qquad \xi \in [-\pi, \pi). \tag{5.3}$$

5.1. The distribution of eigenvalues and singular values. Let us denote by $\lambda_p^{\mathbb{D}}(t_j, w)$ the eigen-values of the Toeplitz matrix (5.1), for $p = 1, \ldots, N$, and assume that they are in decreasing order. It is stated in [24, 23] that if $Q_{t_j}^\varepsilon(\cdot, w) \in L^\infty[-\pi, \pi)$, we have

$$\lim_{N \to \infty} \lambda_p^{\mathbb{D}}(t_j, w) = \sup_{\xi \in [-\pi, \pi)} Q_{t_j}^\varepsilon(\xi, w), \qquad \lim_{N \to \infty} \lambda_{N-p}^{\mathbb{D}}(t_j, w) = \inf_{\xi \in [-\pi, \pi)} Q_{t_j}^\varepsilon(\xi, w), \tag{5.4}$$

for all fixed positive integers p. In our setup N is large, but it depends on the same parameter ε as the symbol $Q_{t_j}^\varepsilon(\xi, w)$. Therefore, we write that

$$\lambda_p^{\mathbb{D}}(t_j, w) \approx \sup_{\xi \in [-\pi, \pi)} Q_{t_j}^\varepsilon(\xi, w), \qquad \lambda_{N-p}^{\mathbb{D}}(t_j, w) \approx \inf_{\xi \in [-\pi, \pi)} Q_{t_j}^\varepsilon(\xi, w), \tag{5.5}$$

for positive integers p of order one. This gives, obviously, the accumulation of the largest singular values (see also [15, Theorem 4.13])

$$\sigma_p^{\mathbb{D}}(t_j, w) \approx \|Q_{t_j}^\varepsilon(\cdot, w)\|_{L^\infty[-\pi, \pi)}, \qquad p = O(1). \tag{5.6}$$

We also have from [15, Theorem 4.5] that when $Q_{t_j}^\varepsilon(\xi, w)$ vanishes at least at one point $\xi \in [-\pi, \pi)$, which happens in our case, there is an accumulation of the singular values at zero,

$$\sigma_{N-p}^{\mathbb{D}}(t_j, w) \approx 0, \qquad p = O(1). \tag{5.7}$$

The distribution of the eigenvalues (singular values) is given approximately by Szegö's first limit theorem [15, Corollary 5.12],

$$\frac{1}{N} \sum_{p=1}^N 1_{[\alpha, \beta]}(\lambda_p^{\mathbb{D}}(t_j, w)) \approx \frac{1}{2\pi} \int_{-\pi}^\pi d\xi \, 1_{[\alpha, \beta]}\left(Q_{t_j}^\varepsilon(\xi, w)\right), \tag{5.8}$$

where $1_{[\alpha, \beta]}$ is the indicator function of the arbitrary interval $[\alpha, \beta]$ on the real line. In fact, we have [15, Theorem 5.10]

$$\frac{1}{N} \sum_{p=1}^N g(\lambda_p^{\mathbb{D}}(t_j, w)) \approx \frac{1}{2\pi} \int_{-\pi}^\pi d\xi \, g\left(Q_{t_j}^\varepsilon(\xi, w)\right), \tag{5.9}$$

for any continuous function g. We analyze next the symbol $Q_{t_j}^\varepsilon(t_j, w)$, and then use approximations (5.8) and (5.9) to study the spectrum of the LC transformed matrix.

15

5.2. Analysis of the symbol. We derive here a simpler expression of the symbol $Q_{t_j}^\varepsilon(\xi, w)$, which allows us to relate it to the random reflection coefficients $\mathcal{R}^\varepsilon$ of the waves at specific slowness moduli. We begin with the following result.

PROPOSITION 5.1. *The symbol $Q_{t_j}^\varepsilon(t_j, w)$ is given by*

$$Q_{t_j}^\varepsilon(\xi, w) \approx \frac{\varepsilon^{\frac{\gamma}{2}-1}}{(2\pi)^2} \frac{\sqrt{2\Delta_T}}{\Delta_x} \Re \left\{ w\hat{f}(w) \int dh\, \hat{\chi}(h\Delta_T) e^{-\frac{i(w-\varepsilon^\gamma h)t_j}{\varepsilon}} \int_o^{1/c} dK \mathcal{R}^\varepsilon(w - \varepsilon^\gamma h, K, 0) \times \right.$$

$$\left. \sum_{q\in\mathbb{Z}} 1_{[-1,1]} \left(\frac{2q\pi - \xi}{K\Delta_x(w - \varepsilon^\gamma h)} \right) \left[1 - \left(\frac{2q\pi - \xi}{K\Delta_x(w - \varepsilon^\gamma h)} \right)^2 \right]^{-1/2} \right\} \quad (5.10)$$

The proof is in appendix B, but the result can be understood as follows. We see from (5.2) and (5.3) that the symbol is the discrete Fourier transform of $\widetilde{\mathbb{D}}^\varepsilon$, which is itself a higher dimensional Fourier transform. Proposition 5.1 says that one Fourier transform in (5.2) is undone in the calculation of the symbol. More explicitly, the expression of $Q_{t_j}^\varepsilon$ involves the Dirichlet kernel

$$\sum_{q=-N+1}^{N-1} e^{iq\xi + iq\Delta_x(w-\varepsilon^\gamma h)\mathbf{K}\cdot\mathbf{e}} = \frac{\sin\left[\left(N - \frac{1}{2} \right) (\xi + (w - \varepsilon^\gamma h)\Delta_x \mathbf{K} \cdot \mathbf{e}) \right]}{\sin\left(\frac{\xi + (w - \varepsilon^\gamma h)\Delta_x \mathbf{K}\cdot\mathbf{e}}{2} \right)}$$

which in the limit $N \to \infty$ acts as an approximate periodic delta distribution $\delta\left[2q\pi - \xi - (w - \varepsilon^\gamma h)\Delta_x \mathbf{K} \cdot \mathbf{e} \right]$, for $q \in \mathbb{Z}$. Write then the $\mathbf{K}$ integral in (5.2) in polar coordinates (K, θ), with slowness modulus $K \in (0, 1/c)$ and $\theta \in [0, 2\pi)$, so that $\mathbf{K} \cdot \mathbf{e} = K\cos\theta$. The proposition says that we can collapse the θ integral using the asymptotic limit of the Dirichlet kernel, to obtain (5.10).

Using Proposition 5.1, we can write a simpler expression of the symbol, as proved in appendix C. The result is due to the rapid decorrelation of the random reflection coefficients $\mathcal{R}^\varepsilon$ over slowness moduli, as summarized in section 4.3.

THEOREM 5.2. *The symbol is given by*

$$Q_{t_j}^\varepsilon(\xi, w) \approx \frac{\varepsilon^{\frac{\gamma-1}{2}}}{(2\pi)^2} \frac{\sqrt{2\Delta_T}}{\Delta_x} \Re \left\{ w\hat{f}(w) \int dh\, \hat{\chi}(h\Delta_T) e^{-\frac{i(w-\varepsilon^\gamma h)t_j}{\varepsilon}} \sum_{q\in\mathbb{Z}} 1_{[0,1/c]}(K_{q,\xi}) \times \right.$$

$$\left. \int_0^\infty dk\, R^\varepsilon(w - \varepsilon^\gamma h, K_{q,\xi}^\varepsilon(h) + \varepsilon k)\sqrt{\frac{K_{q,\xi}}{2k}} \right\}, \quad (5.11)$$

with

$$K_{q,\xi}^\varepsilon(h) = \frac{|2q\pi - \xi|}{\Delta_x(w - \varepsilon^\gamma h)} \approx K_{q,\xi}^\varepsilon(0) = K_{q,\xi}, \quad (5.12)$$

and assuming that $K_{q,\xi}$ is finite, and not of order $\le \varepsilon$. Here the approximation is in mean square sense, and therefore with high probability.

This result says that the symbol $Q_{t_j}^\varepsilon(\xi, w)$ is determined by the reflection coefficient $\mathcal{R}^\varepsilon$ of waves with slowness moduli $K \approx K_{q,\xi}^\varepsilon(h)$, the poles of the terms in (5.10). These are plane waves with slowness vectors along the direction $\mathbf{e}$ of the array. In the setup of our numerical simulations, depending on the frequency w, we have between one and three terms in the sum over q, so we may think of the symbol as being determined

16

by the reflection coefficients with slowness moduli given by (5.12) and $q = 0, \pm 1$. More explicitly, when evaluating the symbol $Q^\varepsilon_{t_j}$ at an argument $\xi \in [-\pi, \pi)$, we select backscattered plane waves traveling at different vertical speeds $c(K^\varepsilon_{q,\xi}(h))$.

REMARK 5.3. *Since $\xi \in [-\pi, \pi)$, the assumption of a finite $K_{q,\xi}$ in Theorem 5.2 is relevant only for $q = 0$. We have a very small slowness modulus $K_{q,\xi}$ when $|\xi| \leq O(\varepsilon)$, and the contribution of the term $q = 0$ to (5.10) is approximately*

$$\frac{\varepsilon^{\frac{\gamma}{2}-1}}{(2\pi)^2} \frac{\sqrt{2\Delta_T}}{\Delta_x} \Re \left\{ w\hat{f}(w) \int dh \, \hat{\chi}(h\Delta_T) e^{-\frac{i(w-\varepsilon^\gamma h)t_j}{\varepsilon}} \int_o^{1/c} dK \mathcal{R}^\varepsilon(w - \varepsilon^\gamma h, K, 0) \right\}.$$

Furthermore, this can be approximated (in mean square sense) by restricting the integral over K to a small vicinity of zero, as shown in appendix C.

5.3. Decorrelation of the symbol. It follows easily from Theorem 5.2 and Lemma 4.1 that the symbols $Q^\varepsilon_{t_j}(\xi, w)$ are correlated only if the frequencies are close to each other, at order ε. The frequency sampling in LCT is given by (2.5), and it becomes in our scaling

$$\frac{\Delta_w}{\varepsilon} = \frac{\pi}{\varepsilon^{1-\gamma}\Delta_T} \quad \rightsquigarrow \quad \Delta_w = \varepsilon^\gamma \frac{\pi}{\Delta_T} \gg O(\varepsilon). \tag{5.13}$$

Thus, the scaled frequency is sampled at rate Δ_w that is much larger than order ε, and the symbols $Q^\varepsilon_{t_j}(\xi, w)$ are decorrelated for all the frequency samples. This explains the rapid fluctuations of the singular values over the bandwidth in figures 3.4 and 3.5.

The rapid decorrelation of $\mathcal{R}^\varepsilon$ over the slowness moduli induces decorrelations of the symbol over ξ, as well. Specifically, we see from Theorem 5.2 that $Q^\varepsilon_{t_j}(\xi, w)$ and $Q^\varepsilon_{t_j}(\xi', w)$ are correlated if

$$|K_{q,\xi} - K_{q',\xi'}| \lesssim O(\varepsilon) \tag{5.14}$$

for at least one pair of indexes q, q' in the sum in (5.11). Because $\xi, \xi' \in [-\pi, \pi)$, equation (5.14) holds if

$$\text{sign}(q)\xi - \text{sign}(q')\xi' \lesssim \varepsilon, \quad \text{when } |q| = |q'|,$$
$$|\xi \mp \text{sign}(q)\pi| \lesssim \varepsilon, \quad |\xi' \pm \text{sign}(q')\pi| \lesssim \varepsilon, \quad \text{when } |q| = |q'| \pm 1, \tag{5.15}$$

where $\text{sign}(q) = 1$ for $q \geq 0$ and -1 otherwise. When only the $q = 0$ term contributes in the sum in (5.11), we have that $Q^\varepsilon_{t_j}(\xi, w)$ and $Q^\varepsilon_{t_j}(\xi', w)$ are correlated if $|\xi - \xi'| \lesssim \varepsilon$. If a few more terms appear in the sum, we also get correlations for ξ and ξ' in order ε vicinities of the ends $\pm\pi$ of the interval containing them. In any case, it is because of such rapid decorrelations of the symbol $Q^\varepsilon_{t_j}(\xi, w)$ over ξ that we expect to obtain statistically stable estimates of the distribution (5.8) of the eigenvalues over properly chosen intervals $[\alpha, \beta]$, as we discuss later.

5.4. Gaussian statistics. It is shown in [22, Section 9.3.4] that the backscattered field $\mathcal{N}$, observed around a fixed time t, converges in distribution to a Gaussian process, as $\varepsilon \to 0$. Given our representation of the symbol $Q^\varepsilon_{t_j}(\xi, w)$ in terms of $\mathcal{R}^\varepsilon$, it is not surprising that we obtain the following result.

THEOREM 5.4. *The symbol $Q^\varepsilon_{t_j}(\xi, w)$ converges in distribution to a Gaussian random variable $Q_{t_j}(\xi, w)$ as $\varepsilon \to 0$, for any fixed frequency w and $\xi \in [-\pi, \pi)$. The limit has mean zero and variance*

$$E\left\{ Q^2_{t_j}(\xi, w) \right\} = \frac{1}{8\pi} \frac{|w| |\hat{f}(w)|^2 \|\chi\|^2}{c^2 t_j \Delta_x^2} \sum_{q \in \mathbb{Z}} 1_{[0,1/c]}(K_{q,\xi}) V_1(w, K_{q,\xi}, t_j), \tag{5.16}$$
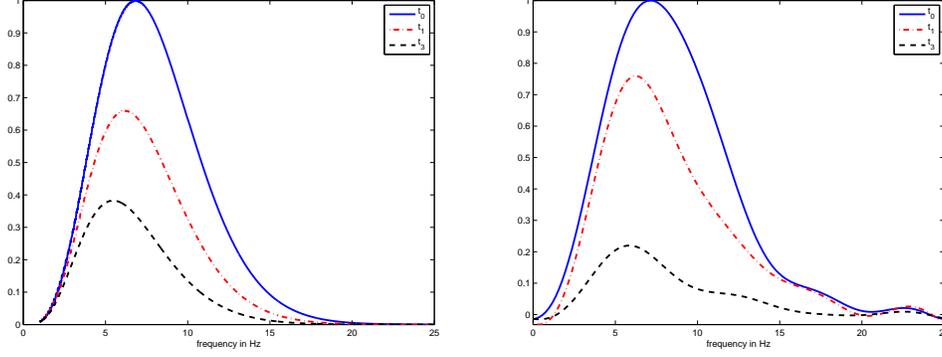
17

FIG. 5.1. *The energy of $\widetilde{\mathbb{P}}(t_j, w)$ vs. frequency w, at tree level $d = 3$ and in time windows t_j, for $j = 0, 1$ and 3. The abscissa is the frequency in Hz. On the left we show the theory prediction. On the right we show the smoothed energy computed numerically.*

where V_1 is given by (4.26) evaluated at $p = 1$, and

$$\|\chi\|^2 = \int \frac{dt}{\Delta_T} \left[\chi \left(\frac{t}{\Delta_T} \right) \right]^2 = \frac{\Delta_T}{2\pi} \int dh \, |\widehat{\chi}(h \Delta_T)|^2 \,.$$

The proof is in appendix D and it consists of showing that the moments of $Q^\varepsilon_{t_j}(\xi, w)$ converge to those of the Gaussian variable $Q_{t_j}(\xi, w)$, as $\varepsilon \to 0$.

5.5. The energy of the LC transformed matrix. Now that we know the limiting statistics of the symbol $Q^\varepsilon_{t_j}(\xi, w)$, we can estimate the energy of the LC transformed response matrix using Szegö's limit result (5.9). Specifically, we can compute the energy of $\widetilde{\mathbb{P}}^\varepsilon(t_j, w)$

$$E \left\{ \frac{1}{N} \sum_{p=1}^{N} \left[\sigma_p^{\mathbb{D}}(t_j, w) \right]^2 \right\} = E \left\{ \frac{1}{N} \|\widetilde{\mathbb{P}}^\varepsilon(t_j, w)\|_F^2 \right\} \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} d\xi \, E \left\{ \left[Q^\varepsilon_{t_j}(\xi, w) \right]^2 \right\} \approx \int_{-\pi}^{\pi} d\xi \, E \left\{ Q^2_{t_j}(\xi, w) \right\},$$

(5.17)

where $\| \cdot \|_F$ is the Frobenius norm.

We show in the left plot of figure 5.1 the theoretical prediction of the energy as a function of w, in time windows t_0, t_1 and t_3, at tree level $d = 3$. We compute it using (5.16) in the right hand side of (5.17), with the parameters defined in our numerical simulations in section 3. We plot with the solid blue line the predicted energy at t_0, normalized by its maximum. The red dash-dot line and the black dash line show the energy at times t_1 and t_3, normalized by the maximum energy at t_0. Note the shift of energy toward the lower frequencies and the overall decay as the time progresses. This is a manifestation of the wave localization phenomenon, which does not allow the waves at the higher frequencies to penetrate to large depths. When the observation time grows, we receive waves that come from deeper depths, and the energy shifts toward the lower frequencies.

In the right picture in figure 5.1 we show the numerical estimate of (5.17). We obtain it by smoothing the computed Frobenius norm as follows. We interpret $\frac{1}{N} \|\widetilde{\mathbb{P}}^\varepsilon(t_j, w)\|_F^2$ as a discrete signal and Fourier transform it in the w variable. Then, we zero all the Fourier coefficients except at the lowest three "frequencies". The numerical estimate shown in figure 5.1 is given by the inverse Fourier transform of the filtered coefficients. We note that although the numerical and theoretical estimates are not identical, the theory captures correctly
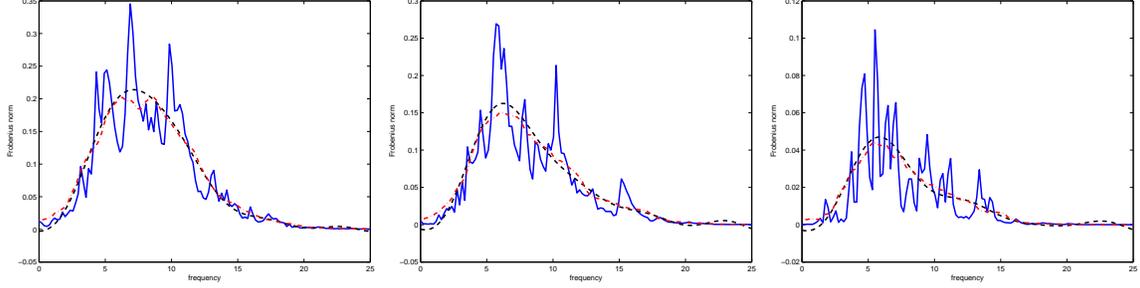
18

FIG. 5.2. *We plot with blue solid line the Frobenius norm of $\widetilde{\mathbb{P}}(t_j, w)$ at tree level $d = 3$. The black dash-dot line is the smoothed energy shown in figure 5.1. The red dash line is the Frobenius norm averaged over a sliding window with 20 frequencies. The abscissa is frequency in Hz. From left to right we take t_0, t_1 and t_3, respectively.*
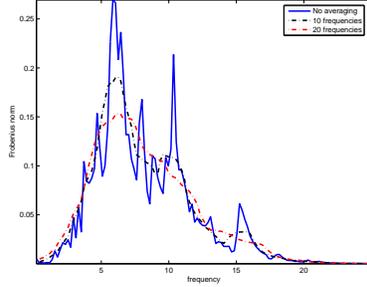

FIG. 5.3. *Illustration of the smoothing effect of averaging over a sliding window of frequencies. The blue solid line is the Frobenius norm of $\widetilde{\mathbb{P}}(t_1, w)$ at tree level $d = 3$. The black and red dashed lines are local averages of the Frobenius norm over ten and twenty frequencies, respectively.*

the behavior of the energy.

Because of the decorrelation properties of the symbol over ξ, we expect that when N is large enough, $\frac{1}{N}\|\widetilde{\mathbb{P}}^{\varepsilon}(t_j, w)\|_F^2$ approaches its statistical mean. Explicitly, the decorrelation results in section 5.3 imply that the variance of $\frac{1}{N}\|\widetilde{\mathbb{P}}^{\varepsilon}(t_j, w)\|_F^2$ is very small, of order ε. In our simulations N is not that large, and it is related to the asymptotic parameter ε. This is why we do not observe the self-averaging of the energy in figure 5.2. The computed Frobenius norm $\frac{1}{N}\|\widetilde{\mathbb{P}}^{\varepsilon}(t_j, w)\|_F^2$ is shown with the solid blue line, and the smooth numerical estimate used in figure 5.1 is shown with the black dash line. However, we do have the rapid decorrelation of the symbol over the frequencies, and when we average $\frac{1}{N}\|\widetilde{\mathbb{P}}^{\varepsilon}(t_j, w)\|_F^2$ over twenty frequencies around w, we obtain the red dash-dot line that is very similar to the smoothed black dash curve. See also figure 5.3, where we show the local averages over ten and twenty frequencies of the Frobenius norm of the LC transformed matrix in the time window indexed by t_1, at tree level $d = 3$. We see there that it is not enough to average over ten frequencies to kill all the fluctuations (see the spurious peak of the black dashed curve around the frequency 10Hz). When we average over 20 frequencies around each w, we get the smoother red dashed curve that is similar to the theoretical prediction in figure 5.1.

5.6. The distribution of singular values. The distribution of singular values is given by

$$\frac{1}{N}\sum_{p=1}^{N} 1_{[\alpha,\beta]}\left[\sigma_p^{\mathbb{D}}(t_j, w)\right] \approx \frac{1}{2\pi}\int_{-\pi}^{\pi} d\xi\, 1_{[\alpha,\beta]}\left[\left|Q_{t_j}^{\varepsilon}(\xi, w)\right|\right]. \tag{5.18}$$
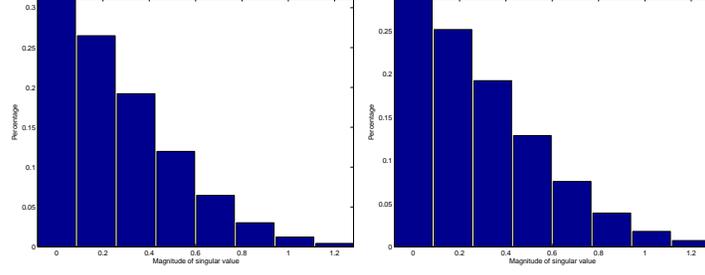
19

FIG. 5.4. *The theoretical prediction of the distribution of singular values in the first time window t_0 at tree level $d = 3$. The ordinate is the percent of singular values with magnitude in the interval indicated in the abscissa. From left to right we show the distribution at frequencies $w_{40} = 7.8Hz$, and $w_{60} = 11.7Hz$.*
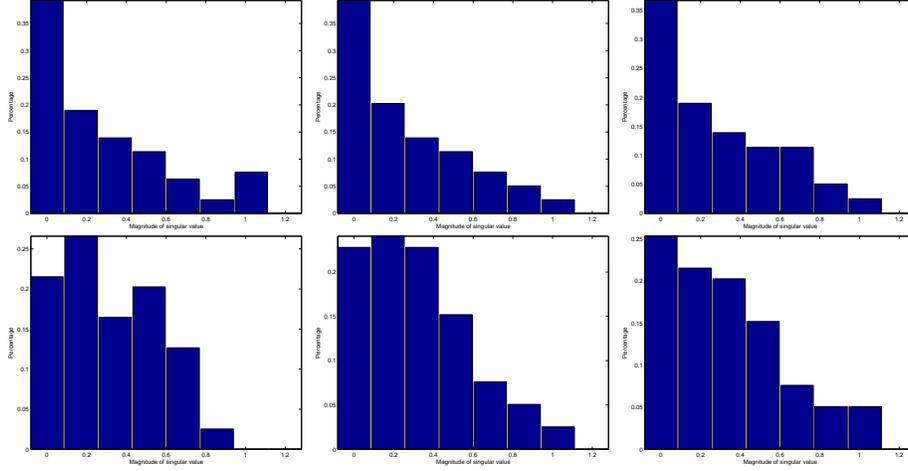

FIG. 5.5. *The numerical estimate of the distribution of singular values with magnitude in the interval indicated in the abscissa. The top row is for frequency $w_{40} = 7.8Hz$, and the bottom row for $w_{60} = 11.7Hz$. We show from left to right the distribution computed with no frequency averaging, and with ten and twenty frequency averaging, respectively.*

We show its numerical estimate in the top row in figure 5.5, for frequencies $w_{40} = 7.8$Hz and $w_{60} = 11.7$Hz. The ordinate in the plot is percent of singular values and the abscissa gives the frequency dependent intervals $[\alpha(w), \beta(w)]$ that we now describe.

If $\widetilde{\mathbb{P}}^{\varepsilon}(t_j, \omega)$ had constant energy over the frequency band, it would make sense to look at the distribution of singular values in the same interval $[\alpha, \beta]$ for all the frequencies. Here the energy varies significantly over the bandwidth, so we define relative, frequency dependent intervals as follows. First, we let $\langle \sigma_1^{\mathbb{D}}(t_j, w) \rangle$ be the largest singular value smoothed as a function of frequency. The smoothing is done the same way as in section 5.5. We take the sequence $\sigma_1^{\mathbb{D}}(t_j, w)$, Fourier transform it in w, keep its first three Fourier coefficients and inverse Fourier transform to get $\langle \sigma_1^{\mathbb{D}}(t_j, w) \rangle$. Second, we normalize the singular values

$$\sigma_p^{\mathbb{D}}(t_j, w) \rightsquigarrow \frac{\sigma_p^{\mathbb{D}}(t_j, w)}{\langle \sigma_1^{\mathbb{D}}(t_j, w) \rangle}, \qquad p = 1, \ldots, N.$$

Then, we look at the distribution of singular values in eight equally sized segments $[\alpha, \beta]$ of the interval $\left[0, \max_w \sigma_1^{\mathbb{D}}(t_j, w) / \langle \sigma_1^{\mathbb{D}}(t_j, w) \rangle \right]$.

The theoretical prediction of the distribution of singular values is in figure 5.4. We compute it using the asymptotic limit of the mean of the right hand side in (5.18). The limit is for the Gaussian process $Q_{t_j}(\xi, w)$,

as stated in Theorem 5.4.

Similar to what we said in section 5.5, we expect that the distribution stabilizes for large enough N, i.e., it approaches its statistical mean. We show in the left column in figure 5.5 the numerical estimate of the distribution at the same two frequencies as in figure 5.4. We note that at the lower frequency the distribution is qualitatively similar to the theoretical one, and smoothing by local frequency averaging is not essential. At the higher frequencies, the numerically estimated distribution is not similar to the theoretical prediction, but the results improve when averaging locally over twenty frequencies (the bottom right plot in figure 5.5).

6. Detection of the coherent echoes. The LC transform of the response matrix in time windows with $t_j \approx \tau^{\mathcal{C}}$ follows from (4.7) and (4.30),

$$\widetilde{\mathbb{P}}^{\varepsilon}_{r,s}(t_j, w) \approx \widetilde{\mathbb{D}}^{\varepsilon}_{r-s}(t_j, w) + \widetilde{\mathcal{C}}^{\varepsilon}_{r,s}(t_j, w), \tag{6.1}$$

where $\widetilde{\mathcal{C}}^{\varepsilon}$ is the LC transform of the coherent matrix (4.16). Because our time windows are much broader than the pulse, we can write

$$\widetilde{\mathcal{C}}^{\varepsilon}_{r,s}(t_j, w) = \int dt\, \mathcal{C}_{rs}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s) \sqrt{\frac{2}{\varepsilon^{1-\gamma}\Delta_T}} \chi\left(\frac{t - t_j}{\varepsilon^{1-\gamma}\Delta_T}\right) \cos\left[\frac{w}{\varepsilon}(t - t_j)\right]$$

$$\approx \sqrt{\frac{2}{\varepsilon^{1-\gamma}\Delta_T}} \chi(0) \int dt\, \mathcal{C}_{rs}(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s) \cos\left[\frac{w}{\varepsilon}(t - t_j)\right].$$

Here we assume that the coherent arrivals are well contained in the window χ, to extend the integral to the entire real line. This is consistent with our assumption in (5.13) that the time window is much larger than the pulse width. Thus, we have approximately

$$\widetilde{\mathcal{C}}^{\varepsilon}_{r,s}(t_j, w) \approx \sqrt{\frac{2}{\varepsilon^{1-\gamma}\Delta_T}} \chi(0) \Re\left\{ e^{-i\frac{w}{\varepsilon}t_j} \widehat{\mathcal{C}}\left(\frac{w}{\varepsilon}, \vec{\mathbf{x}}_s, \vec{\mathbf{x}}_r\right) \right\}$$

$$\approx \varepsilon^{\frac{\gamma}{2}-1} \sqrt{\frac{2}{\Delta_T}} \chi(0) \Re\left\{ w^2 \widehat{f}(w) e^{-i\frac{w}{\varepsilon}t_j} \widehat{\mathcal{G}}^{\varepsilon}_{\mathrm{ODA}}(w, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s) \widehat{\mathcal{G}}^{\varepsilon}_{\mathrm{ODA}}(w, \vec{\mathbf{x}}_r, \vec{\mathbf{y}}^\star) \right\}, \tag{6.2}$$

from model (4.16) with $\widehat{\mathcal{G}}^{\varepsilon}_{\mathrm{ODA}}(w, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s) = \widehat{\mathcal{G}}_{\mathrm{ODA}}(w/\varepsilon, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_s)$.

To state the result that justifies the detection of the coherent echoes, we denote by $\lambda_p(t_j, w)$ and $\lambda_p^{\mathcal{C}}(t_j, w)$ the eigenvalues of $\widetilde{\mathbb{P}}^{\varepsilon}(t_j, w)$ and $\widetilde{\mathcal{C}}^{\varepsilon}(t_j, w)$, and suppose that they are in decreasing order. Recall from section 5.1 that $\lambda_p^{\mathbb{D}}(t_j, w)$ are the eigenvalues of the Toeplitz matrix $\widetilde{\mathbb{D}}^{\varepsilon}_{r-s}(t_j, w)$ containing the layer echoes. Since t_j and w are fixed here, we drop them from the arguments of the eigenvalues for simplicity of notation. We have the following result proved in appendix E.

THEOREM 6.1. *The rank of the coherent matrix $\widetilde{\mathcal{C}}^{\varepsilon}(t_j, w)$ is at most two. Depending on the sign of its nonzero eigenvalues, we have the following interlacing relations:*
Case 1: *When $\lambda_N^{\mathcal{C}} \leq 0 < \lambda_1^{\mathcal{C}}$,*

$$\lambda_1^{\mathbb{D}} \lesssim \lambda_1 \leq \lambda_1^{\mathbb{D}} + \lambda_1^{\mathcal{C}} \qquad and \qquad \lambda_N^{\mathbb{D}} + \lambda_N^{\mathcal{C}} \leq \lambda_N \lesssim \lambda_N^{\mathbb{D}}. \tag{6.3}$$

Case 2: *When $\lambda_1^{\mathcal{C}} \geq \lambda_2^{\mathcal{C}} \geq 0$,*

$$\lambda_1^{\mathbb{D}} \leq \lambda_1 \leq \lambda_1^{\mathbb{D}} + \lambda_1^{\mathcal{C}} \qquad and \qquad \lambda_1^{\mathbb{D}} \lesssim \lambda_2 \leq \lambda_1^{\mathbb{D}} + \lambda_2^{\mathcal{C}}. \tag{6.4}$$

21

Case 3: When $\lambda_N^{\mathcal{C}} \le \lambda_{N-1}^{\mathcal{C}} \le 0$,

$$\lambda_N^{\mathbb{D}} + \lambda_N^{\mathcal{C}} \le \lambda_N \le \lambda_N^{\mathbb{D}} \qquad and \qquad \lambda_N^{\mathbb{D}} + \lambda_{N-1}^{\mathcal{C}} \le \lambda_{N-1} \lesssim \lambda_N^{\mathbb{D}}. \tag{6.5}$$

In all three cases, the largest in magnitude of the remaining eigenvalues cannot be distinguished from those of the matrix $\mathbb{D}^\varepsilon(t_j, w)$ *of pure layer echoes.*

The interlacing relations in this theorem say that we can hope to detect the coherent echoes when the spectral norm of $\widetilde{\mathcal{C}}^\varepsilon(t_j, w)$ is large enough

$$\|\widetilde{\mathcal{C}}^\varepsilon(t_j, w)\| = \max\left\{|\lambda_1^{\mathcal{C}}|, |\lambda_N^{\mathcal{C}}|\right\}. \tag{6.6}$$

By (6.2), the amplitudes of the entries in $\widetilde{\mathcal{C}}^\varepsilon(t_j, w)$ are related to the amplitudes of $\widehat{\mathcal{G}}_{\mathrm{ODA}}^\varepsilon$, which decay exponentially with w^2, as explained in section 4.1.3. The decay means that the coherent waves lose energy to the incoherent ones, backscattered by the layers. The spectral norm (6.6) is very small at the high frequencies, and relations (6.3)-(6.5) say that the eigenvalues of $\widetilde{\mathbb{P}}^\varepsilon(t_j, w)$ cannot be distinguished from those of the Toeplitz matrix $\widetilde{\mathbb{D}}^\varepsilon(t_j, w)$ of pure layer echoes. It is only at the low frequencies, where (6.6) is large enough, that we get a significant perturbation of the eigenvalues, as seen in the top row plots in figure 3.5.

7. Summary. Sensor array imaging of remote reflectors embedded in heterogeneous (cluttered), strongly scattering media is difficult because the useful coherent echoes are overwhelmed by the medium backscatter. Coherent imaging in such environments can work only if we pre-process the data with filters that tend to suppress the clutter backscatter and enhance the coherent arrivals. The question is how to design such filters when we have no prior information about the location of the reflectors and the scattering medium. The only implicit assumption is that the reflectors that we wish to find have different scattering properties than the clutter, so that the question of imaging them makes sense.

In this paper we analyze in detail a new detection and filtering approach. It requires the array response matrix $\mathbb{P}(t)$ obtained by emitting pulses from the array, one source at a time, and recording the echoes at the receivers over a time window $t \in (0, T]$. The entries (traces) in this matrix are dominated by the "noise-like" medium backscatter and the detection of the weak but coherent echoes embedded in them is based on a spectral analysis of the local cosine (LC) transform $\widetilde{\mathbb{P}}(t_j, w_n)$ of $\mathbb{P}(t)$. We use the LC transform to decompose the traces in orthonormal bases given by smooth time windows indexed by t_j and modulated by cosine functions that oscillate at frequency samples w_n in the bandwidth. The wider the time windows, the finer the frequency sampling.

Our approach is a systematic method for selecting the time windows that contain detectable coherent echoes, based on the behavior of the singular values of $\widetilde{\mathbb{P}}(t_j, w_n)$ over the frequencies w_n and in progressively refined time windows. We use the LC transform on binary trees, so that the time refinement consists of splitting each window in two equal parts. The key observation is that in the time windows that contain pure backscatter from clutter, the largest singular values are clustered together and have a similar behavior across the frequency band. It is only in the time windows that contain detectable coherent echoes that the largest singular values exhibit an anomalous behavior, especially at the lower frequencies. Our method identifies the time windows of interest by detecting anomalies in the behavior of the largest singular values of $\widetilde{\mathbb{P}}(t_j, w_n)$.

Once such windows are identified at a given time segmentation (level in the binary tree), we refine them by studying the spectrum of the LC transformed matrix of responses in the two sub-windows corresponding to the children nodes at the next level in the tree. Proceeding this way we have a systematic selection of smaller and smaller time windows that contain the coherent echoes that are useful in imaging.

The filtering of the data involves three steps: (1) Setting to zero the LC coefficients in all the windows except the selected ones, at the deepest level in the tree (i.e., the finest time segmentation). (2) Projecting the LC transformed response matrix to the subspace of low rank matrices with singular vectors corresponding to the largest anomalous singular values. This projection is done in the lower frequency sub-band where such anomalies can be detected. The LC coefficients are set to zero in the remainder of the bandwidth. (3) The inverse LC transform of the filtered LC coefficients gives the filtered data.

The detection and filtering algorithm considered in this paper is general in the sense that it applies to many different types of cluttered media. We refer to [14], where the algorithm is presented in more detail and results are presented for various types of clutter. The focus of this paper is on the analysis of the algorithm, which depends on the model of the clutter. We consider finely layered media for two reasons: (1) The layered media are among the most strongly backscattering ones. For example, phenomena such as wave localization occur even when the wave fluctuations due to layering are weak [34, 1, 22]. (2) The LC transformed response matrix of layer backscatter is Toeplitz and symmetric, and we can relate the singular values to its symbol. We show here how the symbol is related to the reflection coefficient of the layered medium (the kernel of $\mathbb{P}(t)$) and then use the theory of waves in randomly layered media [1, 22] to obtain a detailed analysis of the spectrum of $\widetilde{\mathbb{P}}(t_j, w_n)$ in the time windows that contain only layer echoes. In the windows that contain coherent echoes, $\widetilde{\mathbb{P}}(t_j, w_n)$ is a low rank perturbation of the Toeplitz one, and we can bound the largest singular values away from those due to clutter, providing therefore justification for our detection and filtering approach. The detection is successful when the time window is narrow enough so the coherent wave energy is not completely overwhelmed by the backscatter (i.e., the coherent echoes are detectable at the lower frequencies), and yet wider than the pulse width.

The analysis of detection and filtering in general (not layered) clutter is left for another publication.

Appendix A. Explanation of notation. We denote vectors in $\mathbb{R}^3$ by bold letters topped with an arrow, and vectors in $\mathbb{R}^2$ by bold letters. We use hats to denote Fourier transforms with respect to time. The real part is denoted by $\Re$, the imaginary part by $\Im$ and the complex conjugate by an overbar.

• **Sensor locations:** The locations of the sources in the array are $\vec{\mathbf{x}}_s$, with index s indicating a source. The receiver locations are denoted by $\vec{\mathbf{x}}_r$. Because all N sensors in the array $\mathcal{A}$ play the dual role of sources

and receivers, indices s and r take values $1, 2, \ldots, N$.

- **The system of coordinates:** The range axis z originates at the array and it is orthogonal to the layers. The array is linear, along the unit vector $\mathbf{e}$ at the surface $z = 0$. Thus, the receiver locations are $\vec{\mathbf{x}}_r = x_r(\mathbf{e}, 0)$, with arclength $x_r \in [0, a]$ and a the array aperture. Similarly, the sources are at $\vec{\mathbf{x}}_s = x_s(\mathbf{e}, 0)$, where $x_s \in [0, a]$.

- **The reflector and the image:** The reflector is at location $\vec{\mathbf{y}}^\star = (\mathbf{y}^\star, z^\star)$, with range $z^\star < 0$ and cross-range $\mathbf{y}^\star \in \mathbb{R}^2$. When we image, we search for it in a domain $\mathcal{S}$ that contains $\vec{\mathbf{y}}^\star$. The search points are denoted by $\vec{\mathbf{y}} = (\mathbf{y}, z)$, where we distinguish again between the range z and cross-range $\mathbf{y}$ coordinates. The imaging function is $\mathcal{J}(\vec{\mathbf{y}})$. We form it by migrating the array data using travel times $\tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}, \vec{\mathbf{x}}_r)$ between the source at $\vec{\mathbf{x}}_s$, the search point $\vec{\mathbf{y}}$ and then the receiver at $\vec{\mathbf{x}}_r$.

- **The array data:** The entries of the array response matrix $\mathbb{P}(t) \in \mathbb{R}^{N \times N}$ are denoted by $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$. The letters P and $\mathbb{P}$ stand for pressure.

- **The model of the time traces:** In section 4.1 we denote the mathematical model of the time traces $P(t, \vec{\mathbf{x}}_r, \vec{\mathbf{x}}_s)$ by $\mathcal{N}$ or $\mathcal{N} + \mathcal{C}$, depending on the time of observation. Letter $\mathcal{C}$ stands for the coherent part of the array data, observed around travel time $\tau(\vec{\mathbf{x}}_s, \vec{\mathbf{y}}^\star, \vec{\mathbf{x}}_r)$. Letter $\mathcal{N}$ stands for the "noise", the incoherent backscattered waves by the random layers. The coherent part is modeled by the O'Doherty Anstey theory via the kernel (Green's function) denoted by $\widehat{\mathcal{G}}_{\text{ODA}}^{\varepsilon}$. The incoherent part is modeled as a superposition of upgoing plane waves, with amplitude modeled by the random reflection coefficient $\mathcal{R}$. We denote the frequency by ω and let $\mathbf{K}$ be the two dimensional horizontal slowness vector of the plane waves. The vertical wave speed of the plane waves is denoted by $c(K)$, where $K = |\mathbf{K}|$.

- **The random layering and asymptotic regime:** We model the layering with the mean zero random process $\sigma\mu(z/\ell)$, with correlation length ℓ and standard deviation σ. The analysis is in a regime of separation of scales modeled by the small parameter ε, as explained in section 4.2. We use superscripts ε to indicate the dependence on $\varepsilon \ll 1$, and in the analysis we take the limit $\varepsilon \to 0$.

- **The LC transform:** We denote by $d = 0, 1, \ldots$ the depth in the binary tree, and by ΔT_d the width of the time windows. In the analysis we relate ΔT_d to the asymptotic parameter $\varepsilon \ll 1$ as $\Delta T_d = \varepsilon^{\gamma} \Delta_T$, with $\gamma \in (0, 1)$ and Δ_T a reference, order one time interval. For a fixed tree level d, the array response matrix $\mathbb{P}(t)$ is decomposed in a cosine basis, in the time window $\chi\left(\frac{t - t_j}{\Delta T_d}\right)$ with location indexed by $t_j = j\Delta T_d$, with $j = 0, 1, \ldots 2^d - 1$. The discrete frequencies are denoted by w_n. The LC transformed matrix is denoted by $\widetilde{\mathbb{P}}(t_j, w_n) \in \mathbb{R}^{N \times N}$. We use tilde to denote the LC coefficients. Recall that hats stand for Fourier transforms.

- **The spectral decomposition of the LC transformed matrix:** We denote the eigenvalues and singular values of $\widetilde{\mathbb{P}}(t_j, w_n)$ by $\lambda_q(t_j, w_n)$ and $\sigma_q(t_j, w_n)$, for $q = 1, \ldots N$.

Prior to the arrival of the coherent echoes, $\widetilde{\mathbb{P}}(t_j, w_n)$ is a symmetric and real Toeplitz matrix, defined by the sequence $\widetilde{\mathbb{D}}_q(t_j, w_n)$, as explained in section 5. Its eigenvalues are denoted by $\lambda_q^{\mathbb{D}}(t_j, w_n)$ and the singular values by $\sigma_q^{\mathbb{D}}(t_j, w_n)$. We use the Szegö theory to relate them to the symbol $Q_{t_j}^{\varepsilon}(\xi, w)$ of the Toeplitz matrix. This symbol is defined as the discrete Fourier transform of $\widetilde{\mathbb{D}}_q(t_j, w_n)$, in the index q. The symbol is indexed by t_j, which determines the time window, the frequency w_n, and ξ, the dual of q in the Fourier transform. The eigenvalues and singular values of the coherent part of the LC transformed response matrix are denoted

by $\lambda_q^{\mathcal{C}}(t_j, w_n)$ and $\sigma_q^{\mathcal{C}}(t_j, w_n)$, respectively.

• **The time window selection:** Our criterium for selecting automatically the time windows indexed by t_j, at a fixed tree level d, is based on a metric $m(t_j)$. To calculate this metric, we form the matrix $\mathbb{S}(t_j)$ with rows defined by the values of the largest Q singular values $\sigma_q(t_j, w_n)$ of the LC transformed matrix $\widetilde{\mathbb{P}}(t_j, w_n)$. The rows are indexed by $q = 1, \ldots, Q$ and the columns by the index n of the frequencies. The metric $m(t_j)$ is defined as the ratio $\gamma_2(t_j)/\gamma_1(t_j)$ of the second and first largest singular values of $\mathbb{S}(t_j)$.

Appendix B. Proof of Proposition 5.1. Let us write the symbol (5.3) in the form

$$Q_{t_j}^\epsilon(\xi, w) = \frac{\epsilon^{\gamma/2-1}}{(2\pi)^3} \sqrt{\frac{\Delta_T}{2}} \Re \left\{ w^2 \widehat{f}(w) \int_{\mathbb{R}} dh \, \widehat{\chi}(h\Delta_T) \, e^{-i\frac{(w-\epsilon^\gamma h)t_j}{\epsilon}} \mathcal{I}^\epsilon(\xi, w) \right\}, \tag{B.1}$$

where

$$\mathcal{I}^\epsilon(\xi, w) = \sum_{q=-N+1}^{N-1} e^{iq\xi} \int_{K \leq 1/c} d\mathbf{K} \, \mathcal{R}^\epsilon\left(w - \epsilon^\gamma h, K, 0\right) \, e^{iq\Delta_x(w-\epsilon^\gamma h)\mathbf{K}\cdot\mathbf{e}}. \tag{B.2}$$

Proposition 5.1 follows from the following lemma and the integrability of $|\widehat{\chi}|$.

LEMMA B.1. *We have*

$$\mathcal{I}^\epsilon(\xi, w) = \mathcal{J}^\varepsilon(\xi, w) + \mathcal{E}^\varepsilon(\xi, w), \tag{B.3}$$

where

$$\mathcal{J}^\varepsilon(\xi, w) = \frac{4\pi}{\Delta_x(w-\epsilon^\gamma h)} \int_0^{1/c} dK \, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0) \sum_{q \in \mathbb{Z}} \frac{1_{[-1,1]}\left(\frac{2q\pi-\xi}{K\Delta_x(w-\epsilon^\gamma h)}\right)}{\left[1 - \left(\frac{2q\pi-\xi}{K\Delta_x(w-\epsilon^\gamma h)}\right)^2\right]^{1/2}} \tag{B.4}$$

and the residual $\mathcal{E}^\varepsilon(\xi, w)$ converges uniformly to zero as $\varepsilon \to 0$.

Proof: To simplify the notation, let

$$\phi(u) = \frac{1}{\sqrt{1-u^2}}, \quad \text{and} \quad \Delta^\epsilon(w) := \Delta_x(w - \epsilon^\gamma h),$$

and introduce the Dirichlet kernel

$$\mathcal{D}_N(u) := \sum_{q=-N+1}^{N-1} e^{-iqu} = \frac{\sin\left[(N-1/2)u\right]}{\sin(u/2)}.$$

In polar coordinates (K, θ), defined so that $\mathbf{K} \cdot \mathbf{e} = K\cos(\theta)$, we have

$$\mathcal{I}^\epsilon(\xi, w) = \int_0^{1/c} dK \, K \, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0) \int_0^{2\pi} d\theta \, \mathcal{D}_N\left(\xi + \Delta^\epsilon(w)K\cos\theta\right). \tag{B.5}$$

For the inner integral we write

$$\int_0^{2\pi} d\theta \, \mathcal{D}_N\left(\xi + \Delta^\epsilon(w)K\cos\theta\right) = 2\int_{-1}^1 du \, \mathcal{D}_N\left(\xi + \Delta^\epsilon(w)K \, u\right)\phi(u)$$

$$= 2\int_{-1}^1 du \, \mathcal{D}_N\left(\xi + \Delta^\epsilon(w)K \, u\right)\left(\phi_s(u) + \phi_r(u)\right)$$

$$= \frac{2}{\Delta^\epsilon(w)K} \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta \, \mathcal{D}_N\left(\xi + \zeta\right)\left[\phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right) + \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right)\right],$$

25

where ϕ_s is a nonnegative smooth function bounded above by ϕ and satisfying

$$0 \leq \phi_s(u) \leq \phi(-1 + \delta) = \phi(1 - \delta) \sim \delta^{-1/2}, \qquad u \in [-1, 1],$$

for a fixed $\delta > 0$. The function $\phi_r = \phi - \phi_s \geq 0$ is the non smooth remainder. Thus, the integral splits naturally in smooth and remainder parts,

$$\mathcal{I}^\epsilon(\xi, w) = \mathcal{I}^\epsilon_{\phi_s}(\xi, w) + \mathcal{I}^\epsilon_{\phi_r}(\xi, w).$$

To minimize the computations, we design ϕ_s to have the following properties:

(1) $\phi_s = \phi$ in the interval $[-1 + \delta, 1 - \delta]$ with $\phi_s(-1) = \phi_s(1) = 0$.

(2) ϕ'_s is continuously differentiable with $\phi'_s(-1) = \phi'_s(1) = 0$.

To obtain such ϕ_s just pick $g \in C^1[-1, 1]$ such that

$$\int_{-1}^1 du \ g(u) = 0, \quad \text{and} \quad g = \phi' \quad \text{in} \ [-1 + \delta, 1 - \delta] \quad \text{with} \quad g(-1) = g(1) = 0.$$

Then, define $\phi_s(u) := \int_{-1}^u du' \ g(u')$ for $u \in [-1, 1]$. Clearly g can be chosen such that ϕ_s is nonnegative.

Step 1. Convergence of the smooth part. It is well known that the Dirichlet kernel converges in distribution to the 2π-periodic Dirac measure. Therefore

$$\frac{1}{2\pi} \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta \ \mathcal{D}_N(\xi + \zeta) \phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right) = \sum_{q \in \mathbb{Z}} 1_{[-1,1]}\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right) \phi_s\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right) + \mathcal{E}^\varepsilon_s(\xi, w),$$

with remainder $\mathcal{E}^\varepsilon_s(\xi, w)$ converging uniformly to 0 as $N \to \infty$. In our case $N \sim \varepsilon^{-1}$ so the limit is as $\varepsilon \to 0$. The proof of this fact is similar to the proofs of uniform convergence of Fourier series. Let

$$a_q = \frac{1}{2\pi} \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta \ e^{-iq\zeta} \phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right)$$

be the Fourier coefficients, and use the properties of ϕ_s and integration by parts to obtain

$$|a_q| = \left| \frac{1}{2\pi} \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta \ e^{-iq\zeta} \phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right) \right| \leq \frac{1}{2\pi q^2 \Delta^\epsilon(w)K} \int_{-1}^1 du \ |\phi''_s(u)| \leq \frac{C}{q^2 \delta^{5/2} \Delta^\epsilon(w)K}.$$

Here C is an order one constant and we used the estimate

$$\int_{-1}^1 du \ |\phi''_s(u)| = \text{var}\{\phi'_s\} \leq O(\delta^{-5/2}),$$

that follows from the construction of ϕ_s. Explicitly, the construction ensures that $\phi_s(u) \leq O(\delta^{-1/2})$ and that the first and second derivatives of ϕ_s do not exceed $O(\delta^{-3/2})$ and $O(\delta^{-5/2})$, respectively. Therefore,

$$\left| \frac{1}{2\pi} \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta \ \mathcal{D}_N(\xi + \zeta) \phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right) - \sum_{q \in \mathbb{Z}} 1_{[-1,1]}\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right) \phi_s\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right) \right|$$

$$= \left| \sum_{q=-N+1}^{N-1} a_q e^{-iq\xi} - \sum_{q=-\infty}^{\infty} a_q e^{-iq\xi} \right| \leq \sum_{|q| \geq N-1} |a_q| \leq \frac{C'}{N \delta^{5/2} \Delta^\varepsilon(w)K}, \tag{B.6}$$

for C' another order one constant.

26

Step 2. Estimate of the non smooth remainder. Using that $|\mathcal{R}^\epsilon| \leq 1$, we have

$$\left|\mathcal{I}^\epsilon_{\phi_r}(\xi,w)\right| = \frac{2}{\Delta^\epsilon(w)} \left|\int_0^{1/c} dK\, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0) \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta\, \mathcal{D}_N(\xi + \zeta)\, \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right)\right|$$

$$\leq \frac{2}{\Delta^\epsilon(w)} \int_0^{1/c} dK \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta\, |\mathcal{D}_N(\xi + \zeta)|\, \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right)$$

$$= \frac{2}{\Delta^\epsilon(w)} \int_{-\Delta^\epsilon(w)/c}^{\Delta^\epsilon(w)/c} d\zeta\, |\mathcal{D}_N(\xi + \zeta)| \int_{|\zeta|/\Delta^\epsilon(w)}^{1/c} dK\, \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right).$$

Now $0 \leq \phi_r(u) \leq \frac{1}{\sqrt{1-u^2}}\left[1_{[1-\delta,1]}(u) + 1_{[-1,-1+\delta]}(u)\right]$, and therefore

$$\int_{|\zeta|/\Delta^\epsilon(w)}^{1/c} dK\, \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right) = 2 \int_{|u|/\Delta^\epsilon(w)}^{1/c} dK\, \phi_r\left(\frac{|\zeta|}{\Delta^\epsilon(w)K}\right) = \frac{2|\zeta|}{\Delta^\varepsilon(w)} \int_{1-\delta}^1 \frac{du}{u^2} \phi_r(u)$$

$$\leq \frac{2|\zeta|}{\Delta^\varepsilon(w)} \int_{1-\delta}^1 \frac{du}{u^2\sqrt{1-u^2}} = \frac{C|\zeta|\sqrt{\delta}}{\Delta^\varepsilon(w)},$$

with C an order one constant. The estimate of the remainder becomes

$$\left|\mathcal{I}^\epsilon_{\phi_r}(\xi,w)\right| \leq \frac{2}{\Delta^\epsilon(w)} \int_{-\Delta^\epsilon(w)/c}^{\Delta^\epsilon(w)/c} d\zeta\, |\mathcal{D}_N(\xi + \zeta)| \int_{|\zeta|/\Delta^\epsilon(w)}^{1/c} dK\, \phi_r\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right)$$

$$\leq \frac{C'\sqrt{\delta}}{\Delta^\varepsilon(w)} \int_{-\Delta^\epsilon(w)/c}^{\Delta^\epsilon(w)/c} d\zeta\, |\mathcal{D}_N(\xi + \zeta)| \leq \frac{C''\sqrt{\delta}\ln(N)}{\Delta^\varepsilon(w)}, \tag{B.7}$$

with C' and C'' order one constants. Here we used that the L^1 norm of the Dirichlet kernel diverges as $\ln(N)$ in the limit $N \to \infty$, i.e., as $\varepsilon \to 0$.

Step 3. The final estimate. The triangle inequality gives

$$|\mathcal{I}^\epsilon(\xi,w) - \mathcal{J}^\varepsilon(\xi,w)| \leq \left|\mathcal{I}^\epsilon_{\phi_s}(\xi,w) - \mathcal{J}^\varepsilon(\xi,w)\right| + \left|\mathcal{I}^\epsilon_{\phi_r}(\xi,w)\right|$$

$$\leq \left|\widetilde{\mathcal{I}}^\epsilon_{\phi_s,\delta}(\xi,w) - \widetilde{\mathcal{J}}^\varepsilon_\delta(\xi,w)\right| + \left|\mathcal{I}^\epsilon_{\phi_s,\delta}(\xi,w)\right| + |\mathcal{J}^\epsilon_\delta(\xi,w)| + \left|\mathcal{I}^\epsilon_{\phi_r}(\xi,w)\right|,$$

where $\mathcal{I}^\epsilon_{\phi_s}(\xi,w) = \mathcal{I}^\epsilon_{\phi_s,\delta}(\xi,w) + \widetilde{\mathcal{I}}^\epsilon_{\phi_s,\delta}(\xi,w)$, and

$$\mathcal{I}^\epsilon_{\phi_s,\delta}(\xi,w) = \frac{2}{\Delta^\epsilon(w)} \int_0^\delta dK\, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0) \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta\, \mathcal{D}_N(\xi + \zeta)\, \phi_s\left(\frac{\zeta}{\Delta^\epsilon(w)K}\right).$$

Similarly, $\mathcal{J}^\epsilon = \mathcal{J}^\epsilon_\delta + \widetilde{\mathcal{J}}^\epsilon_\delta$, where

$$\mathcal{J}^\epsilon_\delta(\xi,w) = \frac{4\pi}{\Delta^\varepsilon(w)} \int_0^\delta dK\, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0) \sum_{q\in\mathbb{Z}} 1_{[-1,1]}\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right) \phi_s\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right).$$

We have

$$\left|\mathcal{I}^\epsilon_{\phi_s,\delta}(\xi,w)\right| \leq \frac{C}{\sqrt{\delta}\Delta^\varepsilon(w)} \int_0^\delta dK \int_{-\Delta^\epsilon(w)K}^{\Delta^\epsilon(w)K} d\zeta\, |\mathcal{D}_N(\xi + \zeta)| \leq \frac{C'''\sqrt{\delta}\ln(N)}{\Delta^\varepsilon(w)} \tag{B.8}$$

because $|\phi_s| \leq O(\delta^{-1/2})$ and $|\mathcal{R}^\varepsilon| \leq 1$. Moreover, when estimating $\mathcal{J}^\varepsilon_\delta(\xi,w)$, we note that only the term $q = 0$ may contribute in the sum, and even then only if $|\xi|/\Delta^\varepsilon(w) \leq O(\delta)$. Otherwise, no slowness magnitude $K \in (0,\delta)$ is in the support of the indicator function $1_{[-1,1]}\left(\frac{2\pi q - \xi}{\Delta^\epsilon(w)K}\right)$, for $\xi \in [-\pi,\pi)$. Thus, we have

$$|\mathcal{J}^\varepsilon_\delta(\xi,w)| \leq \frac{4\pi}{\Delta^\varepsilon(w)} \int_{|\xi|/\Delta^\varepsilon(w)}^\delta dK \left[1 - \frac{\xi^2}{(K\Delta^\epsilon(w))^2}\right]^{-1/2} = \frac{4\pi}{\Delta^\varepsilon(w)} \left[\delta^2 - \frac{\xi^2}{(\Delta^\varepsilon(w))^2}\right]^{1/2} \leq \frac{C^{iv}\delta}{\Delta^\varepsilon(w)}. \tag{B.9}$$

The constants C''' and C^{iv} in the estimates (B.8) and (B.9) are order one. Finally, (B.6) gives

$$\left|\widetilde{\mathcal{I}}^{\epsilon}_{\phi_s,\delta}(\xi,w) - \widetilde{\mathcal{J}}^{\varepsilon}(\xi,w)\right| \leq \frac{4\pi}{\Delta^{\epsilon}(w)} \int_{\delta}^{1/c} dK \left| \frac{1}{2\pi} \int_{-\Delta^{\epsilon}(w)K}^{\Delta^{\epsilon}(w)K} d\zeta \, \mathcal{D}_N\left(\xi+\zeta\right)\phi_s\left(\frac{\zeta}{\Delta^{\epsilon}(w)K}\right) - \right.$$

$$\left. \sum_{q\in\mathbb{Z}} 1_{[-1,1]}\left(\frac{2\pi q - \xi}{\Delta^{\epsilon}(w)K}\right)\phi_s\left(\frac{2\pi q - \xi}{\Delta^{\epsilon}(w)K}\right) \right| \leq \frac{C'}{N\delta^{5/2}\left(\Delta^{\varepsilon}(w)\right)^2}\ln\left(\frac{1}{c\delta}\right). \text{(B.10)}$$

Putting (B.7)-(B.10) together,

$$|\mathcal{I}^{\epsilon}(\xi,w) - \mathcal{J}^{\varepsilon}(\xi,w)| \leq \frac{C'}{N\delta^{5/2}\left(\Delta^{\varepsilon}(w)\right)^2}\ln\left(\frac{1}{c\delta}\right) + (C''+C''')\frac{\sqrt{\delta}\ln(N)}{\Delta^{\varepsilon}(w)},$$

and letting $\delta = N^{-1/3} \sim \varepsilon^{1/3}$, we get

$$|\mathcal{I}^{\epsilon}(\xi,w) - \mathcal{J}^{\varepsilon}(\xi,w)| = |\mathcal{E}^{\epsilon}(\xi,w)| \leq \widetilde{C}\frac{\ln(N)}{N^{1/6}\left(\Delta^{\varepsilon}(w)\right)^2}, \tag{B.11}$$

for $\widetilde{C}$ yet another order one constant. Finally note that since $\Delta^{\varepsilon}(w) \approx w\Delta_x$, and there is a w^2 factor in (B.1), the bound on the residual becomes

$$w^2 |\mathcal{I}^{\epsilon}(\xi,w) - \mathcal{J}^{\varepsilon}(\xi,w)| \leq \frac{\widetilde{C}\ln(N)}{N^{1/6}\Delta_x^2}$$

and tends uniformly to zero as $N \sim \varepsilon^{-1} \to \infty$. $\quad\square$

Appendix C. Proof of Theorem 5.2. We assume for simplicity that $\widehat{\chi}$ is smooth and of compact support. The proof is divided in four steps.

Step 1. The set up. Suppose for the moment that $\{K_{q,\xi}\} \cap [0,1/c] \subseteq (0,1/c)$, and note that this implies that $\xi \neq 0$. The case $K_{q,\xi} = 0$ is considered at the end. Fix $\epsilon > 0$ and $\delta > 0$ such that $\epsilon \ll \delta \ll K_{q,\xi}$ and recall that

$$Q^{\epsilon}_{t_j}(\xi,w) = \frac{\epsilon^{\gamma/2-1}}{(2\pi)^2}\frac{\sqrt{2\Delta_T}}{\Delta_x}\Re\left\{ w\widehat{f}(w) \, e^{-\frac{iwt_j}{\epsilon}} \int dh \, \widehat{\chi}(\Delta_T h) \, e^{iht_j/\epsilon^{1-\gamma}}\sum_{q\in\mathbb{Z}}\int_0^{1/c} dK \right.$$

$$\left. \mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma}h, K, 0\right)\phi\left(\frac{K^{\epsilon}_{q,\xi}(h)}{K}\right) \right\},$$

where we have introduced $\phi(u) := \frac{1_{[0,1]}(u)}{\sqrt{1-u^2}}$ to simplify notation. To deal with the singularity at $K = K^{\epsilon}_{q,\xi}(h)$, we decompose the inner integral in two parts. The first one is for K that lie δ-close to $K^{\epsilon}_{q,\xi}(h)$, namely in the interval $I^{\epsilon}_{q,\xi}(h) = [K^{\epsilon}_{q,\xi}(h), K^{\epsilon}_{q,\xi}(h) + \delta] \cap (0,1/c)$. The second part is for the complement of $I^{\epsilon}_{q,\xi}(h)$ in $(0,1/c)$, denoted $\widetilde{I}^{\epsilon}_{q,\xi}(h)$. We have

$$\int_0^{1/c} dK \, \mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma}h, K, 0\right)\phi\left(\frac{K^{\epsilon}_{q,\xi}(h)}{K}\right) = \mathcal{T}_{q,\xi}(h) + \widetilde{\mathcal{T}}_{q,\xi}(h), \tag{C.1}$$

where

$$\mathcal{T}_{q,\xi}(h) = \int_{I^{\epsilon}_{q,\xi}(h)} dK \, \mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma}h, K, 0\right)\phi\left(\frac{K^{\epsilon}_{q,\xi}(h)}{K}\right), \tag{C.2}$$

$$\widetilde{\mathcal{T}}_{q,\xi}(h) = \int_{\widetilde{I}^{\epsilon}_{q,\xi}(h)} dK \, \mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma}h, K, 0\right)\phi\left(\frac{K^{\epsilon}_{q,\xi}(h)}{K}\right). \tag{C.3}$$

28

To evaluate (C.2), we magnify the interval $I_{q,\xi}^\epsilon(h)$ by performing the change of variables

$$K = K_{q,\xi}^\epsilon(h) + \epsilon k, \qquad 0 \le k \le \delta/\epsilon$$

and using the approximation

$$\phi\left(\frac{K_{q,\xi}^\epsilon(h)}{K_{q,\xi}^\epsilon(h) + \epsilon k}\right) \approx \sqrt{\frac{K_{q,\xi}}{2\epsilon k}}. \tag{C.4}$$

Therefore,

$$\mathcal{T}_{q,\xi}(h) \approx \varepsilon \int_0^{\delta/\epsilon} dk\, \mathcal{R}^\epsilon\left(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0\right) \sqrt{\frac{K_{q,\xi}}{2\varepsilon k}}$$

$$\approx \sqrt{\frac{\epsilon}{2}} \int_0^\infty dk\, \mathcal{R}^\epsilon\left(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0\right) \sqrt{\frac{K_{q,\xi}}{k}}. \tag{C.5}$$

We made the second asymptotic equivalence because $\delta \gg \varepsilon$. Then, the total contribution of these terms is

$$\sum_{q\in\mathbb{Z}} \mathcal{T}_{q,\xi}(h) \approx \sqrt{\frac{\epsilon}{2}} \sum_{q\in\mathbb{Z}} \mathbf{1}_{[0,1/c]}\left(K_{q,\xi}\right) \int_0^\infty dk\, \mathcal{R}^\epsilon\left(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0\right) \sqrt{\frac{K_{q,\xi}}{k}}.$$

This is the leading term of $Q_{t_j}^\epsilon(\xi, w)$, consisting of the total contribution of the vicinities about the singularities ± 1 of the ϕ–kernel. The remainder is determined by the sum of the terms $\widetilde{\mathcal{T}}_{q,\xi}(h)$.

Step 2. Estimate of the intensity for the leading term. Let us denote by $L_q^\epsilon(\xi, w)$ the leading term containing $\mathcal{T}_{q,\xi}(h)$. It has mean zero and intensity

$$\mathrm{E}\left\{|L_q^\epsilon(\xi, w)|^2\right\} \approx \frac{\epsilon^{\gamma-1}}{2(2\pi)^4} \frac{\Delta_T}{\Delta_x^2} K_{q,\xi}\, w^2 |\widehat{f}(w)|^2 \int\int dh\, dh'\, \widehat{\chi}(\Delta_T\, h)\widehat{\chi}(\Delta_T\, h')\, e^{i(h-h')t_j/\epsilon^{1-\gamma}}$$

$$\int\int \frac{dk\, dk'}{\sqrt{k\, k'}} \mathrm{E}\left\{\mathcal{R}^\epsilon(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0)\overline{\mathcal{R}^\epsilon(w - \epsilon^\gamma h', K_{q,\xi}^\epsilon(h') + \epsilon k', 0)}\right\},$$

because

$$\mathrm{E}\left\{\mathcal{R}^\epsilon(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0)\mathcal{R}^\epsilon(w - \epsilon^\gamma h', K_{q,\xi}^\epsilon(h') + \epsilon k', 0)\right\} =$$

$$\mathrm{E}\left\{\mathcal{R}^\epsilon(w - \epsilon^\gamma h, K_{q,\xi}^\epsilon(h) + \epsilon k, 0)\overline{\mathcal{R}^\epsilon(-w + \epsilon^\gamma h', K_{q,\xi}^\epsilon(h') + \epsilon k', 0)}\right\} \approx 0$$

by the decorrelation of the reflection coefficients over frequency intervals that are larger than $O(\varepsilon)$. Now change variables

$$h \to h - \epsilon^{1-\gamma}\frac{\tilde{h}}{2}\,, \quad h' \to h + \epsilon^{1-\gamma}\frac{\tilde{h}}{2},$$

$$k \to k - \frac{\tilde{k}}{2}\,, \quad k' \to k + \frac{\tilde{k}}{2},$$

and use the compact support of $\widehat{\chi}$ to obtain $|\epsilon\tilde{h}| = |\epsilon^\gamma(h - h')| \ll 1$, and therefore

$$K_{q,\xi}^\epsilon\left(h + \epsilon^{1-\gamma}\frac{\tilde{h}}{2}\right) = \left|\frac{2\pi q - \xi}{\Delta_x(w - \epsilon^\gamma h - \epsilon\frac{\tilde{h}}{2})}\right| \approx K_{q,\xi}^\epsilon(h) + \epsilon\frac{K_{q,\xi}}{2w}\,\tilde{h}.$$

29

We obtain

$$\mathrm{E}\left\{|L_q^\epsilon(\xi,w)|^2\right\} \approx \frac{K_{q,\xi}}{2(2\pi)^4}\frac{\Delta_T}{\Delta_x^2}\,w^2|\widehat{f}(w)|^2\int dh\,\,|\widehat{\chi}(\Delta_T\,h)|^2\int d\tilde{h}\,\,e^{-i\tilde{h}t_j}\int_0^\infty dk\int_{-2k}^{2k}d\tilde{k}\,\frac{\mathcal{S}^\epsilon(\xi,w,h,\tilde{h},k,\tilde{k})}{k\sqrt{1-(\tilde{k}/2k)^2}}$$

with

$$\mathcal{S}^\epsilon(\xi,w,h,\tilde{h},k,\tilde{k}) = \mathrm{E}\left\{\mathcal{R}^\epsilon\left(w-\epsilon^\gamma h+\epsilon\tilde{h}/2,K_{q,\xi}^\epsilon(h)+\epsilon k+\frac{\epsilon}{2}\left(\tilde{k}-K_{q,\xi}\tilde{h}/w\right),0\right)\right.$$
$$\left.\overline{\mathcal{R}^\epsilon\left(w-\epsilon^\gamma h-\epsilon\tilde{h}/2,K_{q,\xi}^\epsilon(h)+\epsilon k-\frac{\epsilon}{2}\left(\tilde{k}-K_{q,\xi}\tilde{h}/w\right),0\right)}\right\}.$$

The moment formula (4.25) gives

$$\lim_{\epsilon\to 0}\mathcal{S}^\epsilon(\xi,w,h,\tilde{h},k,\tilde{k}) = \int_0^\infty du\,V_1(w,K_{q,\xi},u)\,e^{i\tilde{h}u(1-(K_{q,\xi}c)^2)}e^{-iwuK_{q,\xi}(\tilde{k}-K_{q,\xi}\tilde{h}/w)c^2}$$
$$= \int_0^\infty du\,V_1(w,K_{q,\xi},u)\,e^{i\tilde{h}u}e^{-iwuK_{q,\xi}\tilde{k}c^2},$$

so that

$$\lim_{\epsilon\to 0}\mathrm{E}\left\{|L_q^\epsilon(\xi,w)|^2\right\} = \frac{K_{q,\xi}}{2(2\pi)^4}\frac{\Delta_T}{\Delta_x^2}w^2|\widehat{f}(w)|^2\int dh\,\,|\widehat{\chi}(\Delta_T\,h)|^2\int_0^\infty du\,V_1(w,K_{q,\xi},u)$$
$$\int_{-\infty}^\infty d\tilde{h}\,e^{i\tilde{h}(u-t_j)}\int_0^\infty dk\int_{-2k}^{2k}d\tilde{k}\frac{e^{-iwuK_{q,\xi}\tilde{k}c^2}}{k\sqrt{1-(\tilde{k}/2k)^2}}.$$

The inner integral is given by

$$\int_{-2k}^{2k}d\tilde{k}\,\frac{e^{-iwuK_{q,\xi}\tilde{k}c^2}}{k\sqrt{1-(\tilde{k}/2k)^2}} = 2\int_{-1}^1 d\zeta\,\frac{e^{-i2wuK_{q,\xi}\zeta kc^2}}{\sqrt{1-\zeta^2}} = 2\pi J_0\left(2wuK_{q,\xi}kc^2\right),$$

where J_0 is the 0–Bessel function of the first kind, satisfying

$$\int_0^\infty dk\,J_0(2wuK_{q,\xi}kc^2) = \frac{1}{2wuK_{q,\xi}c^2}.$$

The integral in $\tilde{h}$ is

$$\int_{-\infty}^\infty d\tilde{h}\,e^{i\tilde{h}(s-t_j)} = 2\pi\delta(s-t_j),$$

and the intensity becomes

$$\lim_{\epsilon\to 0}\mathrm{E}\left\{|L_q^\epsilon(\xi,w)|^2\right\} = \frac{1}{4(2\pi)^2}\frac{\Delta_T}{c^2t_j\Delta_x^2}\,w\,|\widehat{f}(w)|^2V_1(w,K_{q,\xi},t_j)\int dh\,\,|\widehat{\chi}(\Delta_t h)|^2. \tag{C.6}$$

Finally, let us note that the terms $L_q^\epsilon(\xi,w)$ are uncorrelated. Indeed, the expression $\mathrm{E}\{L_q^\epsilon(\xi,w)\overline{L_{q'}^\epsilon(\xi,w)}\}$ involves terms

$$\mathrm{E}\left\{\mathcal{R}^\epsilon\left(w,K_{q,\xi}^\epsilon(h),0\right)\overline{\mathcal{R}^\epsilon\left(w,K_{q',\xi}^\epsilon(h),0\right)}\right\} \approx 0,$$

by the decorrelation properties of the reflection coefficients and by $\left|K_{q,\xi}^\epsilon(h)-K_{q',\xi}^\epsilon(h)\right| = \mathrm{O}(1)$ for $q\neq q'$. Thus, we can write

$$E\left\{|\sum_q L_q^\varepsilon(\xi,w)|^2\right\} \approx \sum_q E\left\{|L_q^\varepsilon(\xi,w)|^2\right\}. \tag{C.7}$$

30

Step 3. Estimate on the error term. It follows from *Step 1* that the error term consists of the sum of the integrals over $\widetilde{I}^\epsilon_{q,\xi}(h)$, i.e., containing terms $\widetilde{\mathcal{T}}_{q,\xi}(h)$. Now we prove that the intensity of these terms vanishes in the asymptotic limit $\epsilon \to 0$. First compute,

$$\mathcal{E}^\epsilon(\xi, w) = \frac{\epsilon^{\gamma/2-1}}{(2\pi)^2} \frac{\sqrt{2\Delta_T}}{\Delta_x} \Re \left\{ w\, \widehat{f}(w)\, e^{-i\frac{w}{\epsilon}t_j} \int dh\, \widehat{\chi}(\Delta_T h)\, e^{i\frac{h}{\epsilon^{1-\gamma}}t_j} \sum_{q\in\mathbb{Z}} \widetilde{\mathcal{T}}_{q,\xi}(h) \right\}$$

Note that $\phi\left(K_{q,\xi}/K\right)$ is smooth in $\widetilde{I}^\epsilon_{q,\xi}(h)$, because we are far away from the singularity at $K = K^\epsilon_{q,\xi}(h)$. Therefore,

$$\left| \phi\left(\frac{K^\epsilon_{q,\xi}(h)}{K}\right) - \phi\left(\frac{K_{q,\xi}}{K}\right) \right| \leq \left\| D\phi\left(\frac{K_{q,\xi}}{K}\right) \right\|_\infty |\epsilon^\gamma h| = \max_{\{|y-w|\leq|\epsilon^\gamma h|\}} \left\{ \frac{K\, K_{q,\xi}(y)^2}{y\left(K^2 - K_{q,\xi}(y)^2\right)^{3/2}} \right\} |\epsilon^\gamma h|$$

$$\leq \frac{K\sqrt{K_{q,\xi}}}{w\delta^{3/2}} |\epsilon^\gamma h| = \mathrm{O}\left(\frac{\epsilon^\gamma}{\delta^{3/2}}\right),$$

where $K_{q,\xi}(y) = \left| \frac{2\pi r - \xi}{\Delta_x y} \right|$. For the last inequality, we have chosen $|\epsilon^\gamma h| \ll \delta$, so that $K \geq K^\epsilon_{q,\xi}(y) + 2^{-3/2}\delta$ for all $y \in \{|y-w| \leq |\epsilon^\gamma h|\}$. This is possible as long as $K \in \widetilde{I}^\epsilon_{q,\xi}(h)$, because the support of $\phi\left(K^\epsilon_{q,\xi}(h)/K\right)$ does not intersect the interval $[0, K^\epsilon_{q,\xi}(h))$. We have now obtained that

$$\mathcal{E}^\epsilon(\xi, w) = \frac{\epsilon^{\gamma/2-1}}{(2\pi)^2} \frac{\sqrt{2\Delta_T}}{\Delta_x} \Re \left\{ w\, \widehat{f}(w)\, e^{-i\frac{w}{\epsilon}t_j} \int dh\, \widehat{\chi}(\Delta_T h)\, e^{i\frac{h}{\epsilon^\gamma}t_j} \right.$$

$$\left. \sum_{q\in\mathbb{Z}} \int_{\widetilde{I}^\epsilon_{q,\xi}(h)} dK\, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0)\, \phi\left(\frac{K_{q,\xi}}{K}\right) \right\} + \mathrm{O}\left(\frac{\epsilon^\gamma}{\delta^{3/2}}\right),$$

and consequently,

$$\mathrm{E}\left\{\mathcal{E}^\epsilon(\xi, w)^2\right\} = \frac{\epsilon^{\gamma-2}}{2(2\pi)^4} \frac{\Delta_T}{\Delta_x^2}\, w^2 |\widehat{f}(w)|^2 \int dh\, \widehat{\chi}(\Delta_T h) \int dh'\, \overline{\widehat{\chi}(\Delta_T h')}\, e^{i\frac{(h-h')}{\epsilon^{1-\gamma}}t_j}$$

$$\sum_{q,q'\in\mathbb{Z}} \int_{\widetilde{I}^\epsilon_{q,\xi}(h)} dK \int_{\widetilde{I}^\epsilon_{q',\xi}(h')} dK'\, \mathrm{E}\left\{\mathcal{R}^\epsilon(w - \epsilon^\gamma h, K, 0)\, \overline{\mathcal{R}^\epsilon(w - \epsilon^\gamma h', K', 0)}\right\}$$

$$\phi\left(\frac{K_{q,\xi}}{K}\right) \phi\left(\frac{K_{q',\xi}}{K'}\right) + \mathrm{O}\left(\frac{\epsilon^\gamma}{\delta^{3/2}}\right).$$

Here we have neglected as before two terms, using the rapid decorrelation in the frequency variable of the reflexion coefficients. With the change of variables

$$h \to h - \epsilon^{1-\gamma}\frac{\tilde{h}}{2}, \quad h' \to h + \epsilon^{1-\gamma}\frac{\tilde{h}}{2}$$

$$K \to K + \epsilon\frac{\tilde{k}}{2}, \quad K' \to K - \epsilon\frac{\tilde{k}}{2},$$

we get

$$\mathrm{E}\left\{\mathcal{E}^\epsilon(\xi, w)^2\right\} = \frac{1}{2(2\pi)^4} \frac{\Delta_T}{\Delta_x^2}\, w^2 |\widehat{f}(w)|^2 \int dh\, |\widehat{\chi}(\Delta_T h)|^2 \sum_{q,q'\in\mathbb{Z}} \int_{\widetilde{I}^\epsilon_{q,\xi}(h)} dK\, \phi\left(\frac{K_{q,\xi}}{K}\right) \phi\left(\frac{K_{q',\xi}}{K}\right)$$

$$\int d\tilde{h}\, e^{-i\tilde{h}t_j} \int d\tilde{k}\, S^\epsilon(w, h, \tilde{h}, K, \tilde{k}) + \mathrm{O}\left(\frac{\epsilon^\gamma}{\delta^{3/2}}\right),$$

31

where

$$S^\epsilon(w, h, \tilde{h}, K, \tilde{k}) = \mathrm{E}\left\{ \mathcal{R}^\epsilon\left(w - \epsilon^\gamma h + \epsilon\frac{\tilde{h}}{2}, K + \epsilon\frac{\tilde{k}}{2}, 0\right) \overline{\mathcal{R}^\epsilon\left(w - \epsilon^\gamma h - \epsilon\frac{\tilde{h}}{2}, K - \epsilon\frac{\tilde{k}}{2}, 0\right)}\right\}.$$

$$\to \int_0^\infty ds\, V_1(w, K, s) e^{i\tilde{h}s(1 - (Kc)^2)} e^{-iwsK\tilde{k}c^2}, \qquad \text{as } \varepsilon \to 0.$$

Here the convergence is in L^∞ weak–$\star$ sense and pointwise. Thus, as $\epsilon \to 0$, we have the distributional limit (pointwise for $t_j \neq 0$),

$$\int \frac{d\tilde{h}}{2\pi} \int \frac{d\tilde{k}}{2\pi}\, S^\epsilon(w, h, \tilde{h}, K, \tilde{k})\, e^{-i\tilde{h}t_j} \to \int_0^\infty ds\, V_1(w, K, s)\delta\left[s\left(1 - (Kc)^2\right) - t_j\right] \delta\left(wsKc^2\right)$$

$$= \frac{V_1(w, K, 0)}{wKc^2}\, \delta(t_j).$$

We can now write

$$\lim_{\epsilon \to 0} \mathrm{E}\{\mathcal{E}^\epsilon(\xi, w)^2\} = \frac{1}{2(2\pi)^4}\frac{\Delta_T}{(c\Delta_x)^2}\, w|\widehat{f}(w)|^2 \int dh\, |\widehat{\chi}(\Delta_T h)|^2$$

$$\sum_{q, q' \in \mathbb{Z}} \int_{K > \max\{K_{q,\xi} + \delta, K_{q',\xi}\}} \frac{dK}{K}\, \phi\left(\frac{K_{q,\xi}}{K}\right)\, \phi\left(\frac{K_{q',\xi}}{K}\right) V_1(w, K, 0)\, \delta(t_j).$$

For the region of integration we have used that the integrand vanishes in $[0, \max\{K_{q,\xi}, K_{q',\xi}\}]$. Note also that the integrand in K is bounded for all q and q', by our choice of parameter $\delta > 0$. Thus, for $t_j \neq 0$,

$$\mathrm{E}\{\mathcal{E}^\epsilon(\xi, w)^2\} \to 0 \quad \text{pointwise as } \varepsilon \to 0.$$

REMARK C.1. *The proof above assumes that $K_{q,\xi} \gg O(\varepsilon)$ (recall approximation (C.4)). The case $K_{q,\xi} \leq O(\varepsilon)$ that arises when $q = 0$ and $|\xi| \leq O(\varepsilon)$, can be analyzed separately, using similar arguments. Alternatively, to determine the intensity of the contribution from such small slowness moduli, we can take $\varepsilon \ll |\xi| \ll 1$ and evaluate the intensity of the leading term as in* Step 2 *(equation (C.6) with $q = 0$). Then, using the continuity of this intensity in ξ, we can take the limit $|\xi| \to 0$.* $\quad\square$

Appendix D. Proof of Theorem 5.4. In this section we prove that the symbol $Q^\epsilon_{t_j}(\xi, w)$ converges in the asymptotic limit $\epsilon \to 0$ to Gaussian distributed process in the variable ξ. To this end, we show that the moments of the process agree, in the limit, with those of a Gaussian variable. Recall that the symbol is given by

$$Q^\epsilon_{t_j}(\xi, w) = \sum_{q=-N+1}^{N-1} e^{iq\xi}\widetilde{\mathbb{D}}^\epsilon_q(t_j, w), \qquad \xi \in [-\pi, \pi).$$

The coefficients of this trigonometric series are given by equation (5.2). It was proved in Theorem 5.2 that the symbol can be written as

$$Q^\epsilon_{t_j}(\xi, w) \approx \widetilde{Q}^\epsilon_{t_j}(\xi, w) + \overline{\widetilde{Q}^\epsilon}_{t_j}(\xi, w).$$

where $\widetilde{Q}$ is uniformly equivalent as $\epsilon \to 0$ to

$$\widetilde{Q}^\epsilon_{t_j}(\xi, w) \approx \frac{\epsilon^{\frac{\gamma-1}{2}}}{2(2\pi)^2}\frac{\sqrt{2\Delta_T}}{\Delta_x} w\widehat{f}(w) \int dh\, \widehat{\chi}(h\Delta_T)\, e^{-i\frac{t_j}{\epsilon}(w - \epsilon^\gamma h)} \sum_{q \in \mathbb{Z}} \mathbf{1}_{[0, 1/c]}(K_{q,\xi})$$

$$\int_0^\infty dk\, \mathcal{R}^\epsilon(w - \epsilon^\gamma h, K^\epsilon_{q,\xi}(h) + \epsilon k, 0)\sqrt{\frac{K_{q,\xi}}{2k}}.$$

32

With this expression in mind, we compute

$$
\mathrm{E}\left\{Q_{t_j}^\epsilon(\xi,w)^n\right\} \approx \mathrm{E}\left\{\left(\widetilde{Q}_{t_j}^\epsilon(\xi,w) + \overline{\widetilde{Q}_{t_j}^\epsilon(\xi,w)}\right)^n\right\}
$$

$$
= \sum_{p=0}^n \binom{n}{p} \mathrm{E}\left\{\widetilde{Q}_{t_j}^\epsilon(\xi,w)^p \,\overline{\widetilde{Q}_{t_j}^\epsilon}(\xi,w)^{n-p}\right\} = \sum_{p=0}^n \binom{n}{p} \mathrm{M}_{p,n-p}^\epsilon,
$$

where the coefficients $\mathrm{M}_{p,q}$ have the asymptotic expression

$$
\mathrm{M}_{p,q}^\epsilon = \left(\frac{\epsilon^{\frac{\gamma-1}{2}}}{2(2\pi)^2}\frac{\sqrt{2\Delta_T}}{\Delta_x} w\widehat{f}(w)\right)^{p+q} \int d\mathbf{h}\int d\mathbf{h}' \prod_{l=1}^p \widehat{\chi}(\Delta_T h_l) \prod_{m=1}^q \overline{\widehat{\chi}(\Delta_T h_m')} e^{-i\frac{t_j}{\epsilon}\left(\sum_l(w-\epsilon^\gamma h_l)-\sum_m(w-\epsilon^\gamma h_m')\right)}
$$

$$
\int d\mathbf{k}\int d\mathbf{k}' \prod_{l=1}^p \sum_{q_l} \mathbf{1}_{[0,1/c]}(K_{q_l,\xi})\sqrt{\frac{K_{q_l,\xi}}{2k_l}} \prod_{m=1}^q \sum_{q_m'} \mathbf{1}_{[0,1/c]}(K_{q_m',\xi})\sqrt{\frac{K_{q_m',\xi}}{2k_m'}}
$$

$$
E\left\{\prod_{l=1}^p \mathcal{R}^\epsilon(w-\epsilon^\gamma h_l, K_{q_l,\xi}^\epsilon(h_l)+\epsilon k_l,0)\prod_{m=1}^q \overline{\mathcal{R}^\epsilon(w-\epsilon^\gamma h_m', K_{q_m',\xi}^\epsilon(h_m')+\epsilon k_m',0)}\right\}.
$$

The integration is performed in the product measure

$$
d\mathbf{h}\,d\mathbf{k}\,d\mathbf{h}'\,d\mathbf{k}' = \prod_{l=1}^p dk_l\,dh_l \prod_{m=1}^q dk_m'\,dh_m'.
$$

It is well known that when $p \neq q$ the expected value inside $\mathrm{M}_{p,q}^\epsilon$ vanishes uniformly in ϵ, therefore these terms do not contribute to the asymptotic limit. Note this is the case for $M_{p,n-p}^\varepsilon$, whenever n is odd. When $p = q$ it is possible to use the symmetry in the integration $\{h_m'\}$ to write

$$
\mathrm{M}_{p,p}^\epsilon = p!\left(\frac{\epsilon^{\frac{\gamma-1}{2}}}{2(2\pi)^2}\frac{\sqrt{2\Delta_T}}{\Delta_x} w\widehat{f}(w)\right)^{2p} \int d\mathbf{h}\int_{\{\mathbf{h}'\nearrow\}} \prod_{l=1}^p \widehat{\chi}(h_l\Delta_T)\overline{\widehat{\chi}(h_l'\Delta_T)}\, e^{-it_j\sum_l \epsilon^{1-\gamma}(h_l'-h_l)}
$$

$$
\int d\mathbf{k}\int d\mathbf{k}' \prod_{l=1}^p \sum_{q_l}\mathbf{1}_{[0,1/c]}(K_{q_l,\xi})\sqrt{\frac{K_{q_l,\xi}}{2k_l}}\sum_{q_l'}\mathbf{1}_{[0,1/c]}(K_{q_l',\xi})\sqrt{\frac{K_{q_l',\xi}}{2k_l'}}
$$

$$
\mathrm{E}\left[\prod_{l=1}^p \mathcal{R}^\epsilon(w-\epsilon^\gamma h_l, K_{q_l,\xi}^\epsilon(h_l)+\epsilon k_l)\overline{\mathcal{R}^\epsilon(w-\epsilon^\gamma h_l', K_{q_l',\xi}^\epsilon(h_l')+\epsilon k_l')}\right],
$$

The notation $\{h'\nearrow\}$ means that the integration is done in the set $\{h_1' \leq h_2' \leq \cdots \leq h_p'\}$.

Next, let us change variables

$$
h_l \to h_l - \frac{\epsilon^{1-\gamma}}{2}\widetilde{h}_l \quad\text{and}\quad h_l' \to h_l + \frac{\epsilon^{1-\gamma}}{2}\widetilde{h}_l
$$

$$
k_l \to k_l - \frac{\widetilde{k}_l}{2} \quad\text{and}\quad k_l' \to k_l + \frac{\widetilde{k}_l}{2}
$$

and use the approximations

$$
K_{q_l,\xi}^\epsilon\left(h_l - \epsilon^{1-\gamma}\frac{\widetilde{h}_l}{2}\right) \approx K_{q_l,\xi}^\epsilon(h_l) + \epsilon\frac{K_{q_l,\xi}}{2w}\widetilde{h}_l
$$

$$
K_{q_l',\xi}^\epsilon\left(h_l + \epsilon^{1-\gamma}\frac{\widetilde{h}_l}{2}\right) \approx K_{q_l',\xi}^\epsilon(h_l) - \epsilon\frac{K_{q_l',\xi}}{2w}\widetilde{h}_l.
$$

33

We have

$$M_{p,p}^{\epsilon} \approx p! \left(\frac{1}{2(2\pi)^2} \frac{\sqrt{\Delta_T}}{\Delta_x} w \widehat{f}(w) \right)^{2p} \int d\mathbf{h} \prod_{l=1}^{p} |\widehat{\chi}(h_l \Delta_T)|^2 \int d\widetilde{\mathbf{h}} e^{-i\frac{t_j}{2} \sum_l \widetilde{h}_l}$$

$$\prod_{l=1}^{p} \int_0^{\infty} dk_l \int_{-2k_l}^{2k_l} d\widetilde{k}_l \sum_{q_l} \mathbf{1}_{[0,1/c]}(K_{q_l,\xi}) \frac{K_{q_l,\xi}}{k_l \sqrt{1 - (\widetilde{k}_l/2k_l)^2}} \, S_p^{\epsilon}(\xi, w, h_l, \widetilde{h}_l, k_l, \widetilde{k}_l),$$

where we used that for fixed $(h_1, h_2, \cdots, h_p)$ the integration set becomes

$$\{h_l' : h_l + \frac{\epsilon^{1-\gamma}}{2} \widetilde{h}_l \nearrow\} \to \mathbb{R}^p \text{ as } \epsilon \to 0,$$

and we let

$$S_p^{\epsilon}(\xi, w, h_l, \widetilde{h}_l, k_l, \widetilde{k}_l) = \mathrm{E}\left\{ \prod_{l=1}^{p} \mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma} h_l + \epsilon \widetilde{h}_l/2, K_{q_l,\xi}^{\epsilon}(h_l) + \epsilon k_l + \frac{\epsilon}{2}\left(\widetilde{k}_l - K_{q_l,\xi} \widetilde{h}_l/w \right), 0 \right) \right.$$

$$\left. \overline{\mathcal{R}^{\epsilon}\left(w - \epsilon^{\gamma} h_l - \epsilon \widetilde{h}_l/2, K_{q_l,\xi}^{\epsilon}(h_l) + \epsilon k_l - \frac{\epsilon}{2}\left(\widetilde{k}_l - K_{q_l,\xi} \widetilde{h}_l/w \right), 0 \right)} \right\}.$$

The multi frequency moments are [22, Sections 9.2.4, 14.3]

$$\lim_{\epsilon \to 0} S_p^{\epsilon}(\xi, w, h_l, \widetilde{h}_l, k_l, \widetilde{k}_l) = \prod_{l=1}^{p} \int_0^{\infty} ds \, V_1(w, K_{q_l,\xi}, s) \, e^{i\widetilde{h}_l s} e^{-iws K_{q_l,\xi} \widetilde{k}_l c^2},$$

and we obtain after computations similar to those in *Step 2* in Appendix C that

$$\lim_{\epsilon \to 0} M_{p,p}^{\epsilon} = p! \prod_{l=1}^{p} \left(\frac{1}{2(2\pi)^2} \frac{\sqrt{\Delta_T}}{\Delta_x} w \widehat{f}(w) \right)^2 \int_{-\infty}^{\infty} dh_l \, |\widehat{\chi}(\Delta_T \, h_l)|^2 \int_0^{\infty} ds_l \, V_1(w, K_{q_l,\xi}, s_l)$$

$$\int_{-\infty}^{\infty} d\widetilde{h}_l \, e^{i\widetilde{h}_l(s_l - t_j)} \sum_{q_l} \mathbf{1}_{[0,1/c]}(K_{q_l,\xi}) \int_0^{\infty} dk_l \int_{-2k_l}^{2k_l} d\widetilde{k}_l \frac{e^{-iws_l K_{q_l,\xi} \widetilde{k}_l c^2}}{k_l \sqrt{1 - (\widetilde{k}_l/2k_l)^2}}$$

$$= \frac{p!}{2^p} \prod_{l=1}^{p} \frac{1}{4(2\pi)^2} \frac{\Delta_T}{c^2 t_j \Delta_x^2} w |\widehat{f}(w)|^2 \int dh_l \, |\widehat{\chi}(\Delta_T \, h_l)|^2 \sum_{q_l} \mathbf{1}_{[0,1/c]}(K_{q_l,\xi}) V_1(w, K_{q_l,\xi}, s_l)$$

$$= \frac{p!}{2^p} \left(\lim_{\epsilon \to 0} E[Q_{t_j}^{\epsilon}(\xi, w)^2] \right)^p.$$

We have now proved that

$$\lim_{\epsilon \to 0} \mathrm{E}\left[Q_{t_j}^{\epsilon}(\xi, w)^{2p+1} \right] \to 0$$

$$\lim_{\epsilon \to 0} \mathrm{E}\left[Q_{t_j}^{\epsilon}(\xi, w)^{2p} \right] \to \binom{2p}{p} \frac{p!}{2^p} \left(\lim_{\epsilon \to 0} E[Q_{t_j}^{\epsilon}(\xi, w)^2] \right)^p,$$

which is in agreement with the moment relations of a Gaussian process. $\square$

Appendix E. Proof of Theorem 6.1. Let us begin by writing the coherent part in (6.2) as

$$\widetilde{\mathcal{C}}^{\varepsilon}(t_j, w) = \Re\left\{ (\alpha + i\beta) (\mathbf{C} + i\mathbf{S}) (\mathbf{C} + i\mathbf{S})^T \right\}, \tag{E.1}$$

where $\mathbf{C}$ and $\mathbf{S}$ are vectors in $\mathbb{R}^N$ with components given by the real and imaginary parts of $\widehat{\mathcal{G}}_{\mathrm{ODA}}^{\varepsilon}$,

$$C_r = \Re\{\widehat{\mathcal{G}}_{\mathrm{ODA}}^{\varepsilon}(w, \vec{\mathbf{y}}^{\star}, \vec{\mathbf{x}}_r)\}, \qquad S_r = \Im\{\widehat{\mathcal{G}}_{\mathrm{ODA}}^{\varepsilon}(w, \vec{\mathbf{y}}^{\star}, \vec{\mathbf{x}}_r)\}, \qquad r = 1, \ldots N, \tag{E.2}$$

34

and

$$\alpha + i\beta = \varepsilon^{\frac{\gamma}{2}-1}\sqrt{\frac{2}{\Delta_T}}\chi(0)w^2\widehat{f}(w)e^{-iwt_j}.\tag{E.3}$$

Hereafter we suppose that t_j and w are fixed, and we drop them from the arguments. We have

$$\widetilde{\mathcal{C}}^\varepsilon = \alpha\left(\mathbf{CC}^T - \mathbf{SS}^T\right) - \beta\left(\mathbf{CC}^T + \mathbf{SS}^T\right).\tag{E.4}$$

The proof of Theorem 6.1 follows immediately from the eigenvalue bounds [28, Theorem 10.3.1]

$$\lambda_{i+j-1} \le \lambda_i^{\mathbb{D}} + \lambda_j^{\mathcal{C}},\tag{E.5}$$

$$\lambda_{N+1-(i+j-1)} \ge \lambda_{N+1-i}^{\mathbb{D}} + \lambda_{N+1-j}^{\mathcal{C}},\tag{E.6}$$

for all integers i, j satisfying $1 \le i + j - 1 \le N$, and the following lemma.

LEMMA E.1. *The rank of the LC transformed coherent matrix $\widetilde{\mathcal{C}}_{r,s}^\varepsilon$ is at most two. In the most likely case that $\alpha \ne 0$, only one eigenvalue is positive,*

$$\lambda_N^{\mathcal{C}} \le \lambda_{N-1}^{\mathcal{C}} = \ldots = \lambda_2^{\mathcal{C}} = 0 < \lambda_1^{\mathcal{C}}.\tag{E.7}$$

When $\alpha = 0$, the nonzero eigenvalues have the same sign.

Indeed, say that $\alpha \ne 0$, so that (E.7) holds. Then take $i = j = 1$ in (E.5) and $i = N - 1, j = 2$ in (E.6) to obtain

$$\lambda_2^{\mathbb{D}} = \lambda_2^{\mathbb{D}} + \lambda_{N-1}^{\mathcal{C}} \le \lambda_1 \le \lambda_1^{\mathbb{D}} + \lambda_1^{\mathcal{C}}.$$

Relation (6.3) follows from (5.5), which says that $\lambda_p^{\mathbb{D}} \approx \lambda_1^{\mathbb{D}}$ for $p \ll N$. The other relations are obtained in a similar way. For example, taking $i = 1$, $j = 2$ in (E.5) and $i = N - 2$, $j = 2$ in (E.6), and using (E.7) we get

$$\lambda_2^{\mathbb{D}} \approx \lambda_3^{\mathbb{D}} = \lambda_3^{\mathbb{D}} + \lambda_{N-1}^{\mathcal{C}} \le \lambda_2 \le \lambda_1^{\mathbb{D}} + \lambda_2^{\mathcal{C}} = \lambda_1^{\mathbb{D}} \approx \lambda_2^{\mathbb{D}}$$

and so on. This proves Case 1 in Theorem 6.1.

When $\alpha = 0$ and

$$\lambda_N^{\mathcal{C}} = \lambda_{N-1}^{\mathcal{C}} = \ldots = \lambda_3^{\mathcal{C}} = 0 \le \lambda_2^{\mathcal{C}} < \lambda_1^{\mathcal{C}},\tag{E.8}$$

we take $i = j = 1$ in (E.5) and $i = N, j = 1$ in (E.6) to obtain

$$\lambda_1^{\mathbb{D}} = \lambda_1^{\mathbb{D}} + \lambda_N^{\mathcal{C}} \le \lambda_1 \le \lambda_1^{\mathbb{D}} + \lambda_1^{\mathcal{C}}.$$

Similarly, choices $i = 1, j = 2$ in (E.5) and $i = N - 1, j = 1$ in (E.6) give

$$\lambda_1^{\mathbb{D}} \approx \lambda_2^{\mathbb{D}} = \lambda_2^{\mathbb{D}} + \lambda_N^{\mathcal{C}} \le \lambda_2 \le \lambda_1^{\mathbb{D}} + \lambda_2^{\mathcal{C}}.$$

Furthermore, $i = N - 2, j = 3$ in (E.5) and $i = j = 1$ in (E.6) give

$$\lambda_N^{\mathbb{D}} = \lambda_N^{\mathbb{D}} + \lambda_N^{\mathcal{C}} \le \lambda_N \le \lambda_{N-2}^{\mathbb{D}} + \lambda_3^{\mathcal{C}} = \lambda_{N-2}^{\mathbb{D}} \approx \lambda_N^{\mathbb{D}},$$

and so on. This proves Case 2 of Theorem 6.1. Case 3 follows similarly.

35

Proof of Lemma E.1. Assuming that $\alpha \neq 0$, rewrite (E.4) as

$$\widetilde{\mathcal{C}}^\varepsilon = \alpha \mathbb{M}, \qquad \mathbb{M} = \left(1 + \frac{\beta^2}{\alpha^2}\right) \mathbf{C}\mathbf{C}^T - \mathbf{V}\mathbf{V}^T, \qquad \text{where} \quad \mathbf{V} = \mathbf{S} + \frac{\beta}{\alpha}\mathbf{C}. \tag{E.9}$$

This identity can be checked with a straightforward calculation. Now, there are two cases to consider for characterizing the spectrum of the real, symmetric matrix $\mathbb{M}$.

Case (i): Vectors $\mathbf{C}$ and $\mathbf{V}$ are orthogonal. Then, $\mathbb{M}$ has rank two, with $\mathbf{C}$ and $\mathbf{V}$ the eigenvectors corresponding to the nonzero eigenvalues

$$\mathbb{M}\mathbf{C} = \lambda_1 \mathbf{C}, \qquad \lambda_1 = 1 + \frac{\beta^2}{\alpha^2}\|\mathbf{C}\|^2, \qquad \mathbb{M}\mathbf{V} = \lambda_N \mathbf{V}, \qquad \lambda_N = -\|\mathbf{V}\|^2. \tag{E.10}$$

Case (ii): Vectors $\mathbf{C}$ and $\mathbf{V}$ are not orthogonal. Here we construct an orthonormal basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$ of $\mathbb{R}^N$, so that $\text{span}\{\mathbf{q}_1, \mathbf{q}_2\} = \text{span}\{\mathbf{C}, \mathbf{V}\}$, by taking

$$\mathbf{q}_1 = \frac{\mathbf{C}}{\|\mathbf{C}\|}, \qquad \mathbf{q}_2 = \frac{\mathbf{V} - \left(\mathbf{q}_1^T \mathbf{V}\right)\mathbf{q}_1}{\|\mathbf{V} - \left(\mathbf{q}_1^T \mathbf{V}\right)\mathbf{q}_1\|}, \tag{E.11}$$

Let $\mathbb{Q}$ be the orthogonal matrix in $\mathbb{R}^{N \times N}$ with columns $\mathbf{q}_j$, for $j = 1, \ldots, N$. Since

$$\mathbf{C}\mathbf{C}^T = \|\mathbf{C}\|^2 \mathbf{q}_1 \mathbf{q}_1^T = \mathbb{Q}\,\text{diag}\left(\|\mathbf{C}\|^2, 0, \ldots, 0\right)\mathbb{Q}^T, \tag{E.12}$$

we obtain from definition (E.9) of $\mathbb{M}$ that

$$\mathbb{M} = \mathbb{Q}\left[\text{diag}\left((1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2, 0, \ldots, 0\right) - \left(\mathbb{Q}^T \mathbf{V}\right)\left(\mathbb{Q}^T \mathbf{V}\right)^T\right]\mathbb{Q}^T. \tag{E.13}$$

That is to say, $\mathbb{M}$ is related via a similarity transformation to matrix

$$\mathbb{U} = \text{diag}\left((1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2, 0, \ldots, 0\right) - \left(\mathbb{Q}^T \mathbf{V}\right)\left(\mathbb{Q}^T \mathbf{V}\right)^T. \tag{E.14}$$

But by our choice of the basis, $\mathbb{U}$ has the following block structure

$$\mathbb{U} = \begin{pmatrix} \widetilde{\mathbb{U}} & 0 \\ 0 & 0 \end{pmatrix}, \qquad \widetilde{\mathbb{U}} = \begin{pmatrix} (1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2 & 0 \\ 0 & 0 \end{pmatrix} - \widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^T, \qquad \widetilde{\mathbf{V}} = (\mathbf{q}_1, \mathbf{q}_2)^T \mathbf{V}, \tag{E.15}$$

so the nonzero eigenvalues of $\mathbb{M}$ are the eigenvalues of $\widetilde{\mathbb{U}} \in \mathbb{R}^{2 \times 2}$.

Let $\widetilde{\lambda}_1$ and $\widetilde{\lambda}_2$ be the eigenvalues of $\mathbb{M}$. We obtain by direct calculation that

$$\widetilde{\lambda}_{1,2} = \frac{1}{2}\left\{(1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2 - \|\widetilde{\mathbf{V}}\|^2 \pm \sqrt{\left[(1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2 - \|\widetilde{\mathbf{V}}\|^2\right]^2 + 4\widetilde{V}_2^2(1 + \beta^2/\alpha^2)\|\mathbf{C}\|^2}\right\}, \tag{E.16}$$

and therefore $\widetilde{\lambda}_2 \leq 0 < \widetilde{\lambda}_1$. This proves the case $\alpha \neq 0$.

When $\alpha = 0$, we obtain from (E.4) that

$$\widetilde{\mathcal{C}}^\varepsilon = -\beta\left(\mathbf{C}\mathbf{C}^T + \mathbf{S}\mathbf{S}^T\right). \tag{E.17}$$

When $\mathbf{C}$ and $\mathbf{S}$ are orthogonal, then $\widetilde{\mathcal{C}}^\varepsilon$ has two nonzero eigenvalues, given by $-\beta\|\mathbf{C}\|^2$ and $-\beta\|\mathbf{S}\|^2$. If $\mathbf{C}$ and $\mathbf{S}$ are not orthogonal, we proceed as above and construct an orthonormal basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$ of $\mathbb{R}^N$, so that $\mathbf{C} = \|\mathbf{C}\|\mathbf{q}_1$ and $\mathbf{S} \in \text{span}\{\mathbf{q}_1, \mathbf{q}_2\}$. Then, we obtain the similarity transformation

$$\widetilde{\mathcal{C}}^\varepsilon = -\beta\mathbb{Q}\,\mathbb{U}\,\mathbb{Q}^T, \tag{E.18}$$

where $\mathbb{Q}$ is the orthogonal matrix in $\mathbb{R}^{N \times N}$ with columns $\mathbf{q}_j$ and

$$\mathbb{U} = \begin{pmatrix} \widetilde{\mathbb{U}} & 0 \\ 0 & 0 \end{pmatrix}, \qquad \widetilde{\mathbb{U}} = \begin{pmatrix} \|\mathbf{C}\|^2 & 0 \\ 0 & 0 \end{pmatrix} + \widetilde{\mathbf{S}}\widetilde{\mathbf{S}}^T, \qquad \widetilde{\mathbf{S}} = (\mathbf{q}_1, \mathbf{q}_2)^T \mathbf{S}. \tag{E.19}$$

Now we can compute the eigenvalues of the 2×2 matrix $\widetilde{\mathbb{U}}$

$$\widetilde{\lambda}_{1,2} = \frac{1}{2}\left[\|\mathbf{C}\|^2 + \|\mathbf{S}\|^2 \pm \sqrt{(\|\mathbf{C}\|^2 + \|\mathbf{S}\|^2)^2 - 4\|\mathbf{C}\|^2 \widetilde{S}_2^2}\right] \tag{E.20}$$

and conclude easily that they are nonnegative. The nonzero eigenvalues of $\widetilde{\mathcal{C}}^{\varepsilon}$ are equal to $-\beta\widetilde{\lambda}_{1,2}$, and they have the same sign. $\quad\square$

REFERENCES

[1] M. Asch, W. Kohler, G. Papanicolaou, M. Postel, and B. White, *Frequency content of randomly scattered signals*, SIAM Review, 33 (1991), pp. 519–625.
[2] A. Aubry and A. Derode, *Detection and imaging in a random medium: A matrix method to overcome multiple scattering and aberration*, Journal of Applied Physics, 106 (2009), pp. 044903–044903.
[3] ———, *Random Matrix Theory Applied to Acoustic Backscattering and Imaging In Complex Media*, Physical review letters, 102 (2009), p. 84301.
[4] E. Bécache, P. Joly, and C. Tsogka, *Etude d'un nouvel élément fini mixte permettant la condensation de masse*, C. R. Acad. Sci. Paris Sér. I Math., 324 (1997), pp. 1281–1286.
[5] ———, *An analysis of new mixed finite elements for the approximation of wave propagation problems*, SIAM J. Numer. Anal., 37 (2000), pp. 1053–1084.
[6] Biondo Biondi, *3D Seismic Imaging*, no. 14 in Investigations in Geophysics, Society of Exploration Geophysicists, Tulsa, 2006.
[7] N. Bleistein, J.K. Cohen, and J.W. Stockwell Jr., *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer, New York, 2001.
[8] L. Borcea, F. González del Cueto, G. Papanicolaou, and C. Tsogka, *Filtering random layering effects in imaging*, SIAM Multiscale Modeling and Simulation, 8 (2010), pp. 751–781.
[9] L. Borcea, G. Papanicolaou, and C. Tsogka, *Interferometric array imaging in clutter*, Inverse Problems, 21 (2005), pp. 1419–1460.
[10] ———, *Adaptive interferometric imaging in clutter and optimal illumination*, Inverse Problems, 22 (2006), pp. 1405–1436.
[11] ———, *Coherent interferometric imaging in clutter*, Geophyics, 71 (2006), pp. 1165–1175. Submitted to SIAM Multiscale Analysis.
[12] ———, *Coherent interferometry in finely layered random media*, SIAM Multiscale Modeling and Simulation, 5 (2006), pp. 62–83.
[13] ———, *Asymptotics for the space-time wigner transform with applications to imaging*, in Stochastic differential equations: theory and applications, Peter H. Baxendale and Sergey V. Lototsky, eds., vol. 2 (In honor of Prof. Boris L. Rozovskii) of Interdisciplinary Mathematical Sciences, World Scientific, 2007, pp. 91–112.
[14] ———, *Adaptive time-frequency detection and filtering for imaging in heavy clutter*. submitted to SIAM Imaging Science, 2010.
[15] A. Böttcher and B. Silbermann, *Introduction to large truncated Toeplitz matrices*, Springer Verlag, 1999.
[16] R. Burridge, G. Papanicolaou, and B. White, *One dimensional wave propagation in a highly discointinuous medium*, Wave Motion, 10 (1988), pp. 19–44.
[17] J. F. Claerbout, *Fundamentals of geophysical data processing : with applications to petroleum prospecting*, CA : Blackwell Scientific Publications, Palo Alto, 1985.
[18] J. F. Clouet and J. P. Fouque, *Spreading of a pulse travelling in random media*, Annals of Applied Probability, Vol.4, No.4, (1994).
[19] R. R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier a fenêtre*, C.R. Acad. Sci., (1991), pp. 259–261.
[20] J.C. Curlander and R.N. McDonough, *Synthetic Aperture Radar*, Wiley, New York, 1991.
[21] D. Donoho, A. Maleki, and M. Shahram, *Wavelab 850*. downnload from http://www-stat.stanford.edu/ wavelab/.
[22] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna, *Wave Propagation and Time Reversal in Randomly Layered Media*, Springer, April 2007.
[23] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications, University of Calif*, Press, Berkeley and Los Angeles, 20 (1958), pp. 22–26.
[24] M. Kac, WL Murdock, and G. Szegö, *On the eigenvalues of certain Hermitian forms*, J. Rat. Mech. Anal, 2 (1953), pp. 767–800.
[25] S. Mallat, *A wavelet tour of signal processing*, Academic Press, second ed., 1999.
[26] V. Marcenko and L. Pastur, *Distributions of eigenvalues for some sets of random matrices*, Math. USSR-Sbornik, 1 (1967), pp. 457–483.
[27] R.F. O'Doherty and N. A. Anstey, *Reflections on amplitudes*, Geophysical Prospecting, 19 (1971), pp. 430–458.

[28] B.N. Parlett, *The symmetric eigenvalue problem*, Society for Industrial Mathematics, 1998.

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

[30] P. Sheng, *Introduction to wave scattering, localization and mesoscopic phenomena*, Academic Press, 1995.

[31] K. Sølna and G. Papanicolaou, *Ray theory for a locally layered random medium*, Waves Random Media, (2000).

[32] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*, Commun. Inf. Theory, 1 (2004), pp. 1–182.

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

[34] B. White, P. Sheng, and B. Nair, *Localization and backscattering spectrum of seismic waves in stratified lithology*, Geophysics, 55 (1990), pp. 1158–1165.