arXiv:1607.04625v1 [math-ph] 15 Jul 2016

IMAGING IN RANDOM MEDIA WITH CONVEX OPTIMIZATION

LILIANA BORCEA AND ILKER KOCYIGIT *

Abstract. We study an inverse problem for the wave equation where localized wave sources in random scattering media are to be determined from time resolved measurements of the waves at an array of receivers. The sources are far from the array, so the measurements are affected by cumulative scattering in the medium, but they are not further than a transport mean free path, which is the length scale characteristic of the onset of wave diffusion that prohibits coherent imaging. The inversion is based on the Coherent Interferometric (CINT) imaging method which mitigates the scattering effects by introducing an appropriate smoothing operation in the image formation. This smoothing stabilizes statistically the images, at the expense of their resolution. We complement the CINT method with a convex (l_1) optimization in order to improve the source localization and obtain quantitative estimates of the source intensities. We analyze the method in a regime where scattering can be modeled by large random wavefront distortions, and quantify the accuracy of the inversion in terms of the spatial separation of individual sources or clusters of sources. The theoretical predictions are demonstrated with numerical simulations.

Key words. waves in random media, coherent interferometric imaging, l_1 optimization, mutual coherence.

1. Introduction. Waves measured by a collection of nearby sensors, called an array of receivers, carry information about their source and the medium through which they travel. We consider a typical remote sensing regime with sources of small (point-like) support, and study the inverse problem of determining them from the array measurements.

When the waves travel in a known and non-scattering (e.g. homogeneous) medium, the sources can be localized with reverse time migration [4, 5] also known as backprojection [20]. This estimates the source locations as the peaks of the image formed by superposing the array recordings delayed by travel times from the receivers to the imaging points. The accuracy of the estimates depends on the array aperture, the distance of the sources from the array, and the temporal support of the signals emitted by the sources. It may be improved under certain conditions by using l_1 optimization, which seeks to invert the linear mapping from supposedly sparse vectors of the discretized source amplitude on some mesh, to the array measurements. The fast growing literature of imaging with l_1 optimization in homogeneous media includes compressed sensing studies such as [19, 18], synthetic radar imaging studies like [1, 8], array imaging studies like [14], and the resolution study [7].

In this paper we assume that the waves travel in heterogeneous media with fluctuations of the wave speed caused by numerous inhomogeneities. The amplitude of the fluctuations is small, meaning that a single inhomogeneity is a weak scatterer. However, there are many inhomogeneities that interact with the waves on their way from the sources to the receivers, and their scattering effect accumulates. Because in applications it is impossible to know the inhomogeneities in detail, and these cannot be estimated from the array measurements as part of the inversion, the fluctuations of the wave speed are uncertain. We model this uncertainty with a random process, and thus study inversion in random media. In this stochastic framework, the actual heterogeneous medium in which the waves propagate is one realization of the random model. The data measured at the array are uncertain and the question is how to mitigate the uncertainty to get images that are robust with respect to arbitrary medium realizations, i.e., they are statistically stable.

The mitigation of uncertainty in the wave propagation model becomes important when the sources are further than a few scattering mean free paths from the array. The scattering mean free path is the length scale on which the waves randomize [23], meaning that their fluctuations from one medium realization to another are large in comparison with their coherent (statistical expectation) part. The random wave distortions registered at the array are very different from additive and uncorrelated noise assumed usually in inversion. They are more difficult to mitigate and lead to poor and unreliable source reconstructions by coherent methods like reverse time migration or standard l_1 optimization. The Coherent Interferometric (CINT) method [9, 6] is designed to deal efficiently with such random distortions, as long as there is some residual coherence in the array measurements. This holds when the sources are separated from the array by distances (ranges) that are large with respect to the scattering mean free path, but do not exceed a transport mean free path, which is the distance at which the waves forget their initial direction [23]. The

*Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043.
Email: borcea@umich.edu & ilkerk@umich.edu

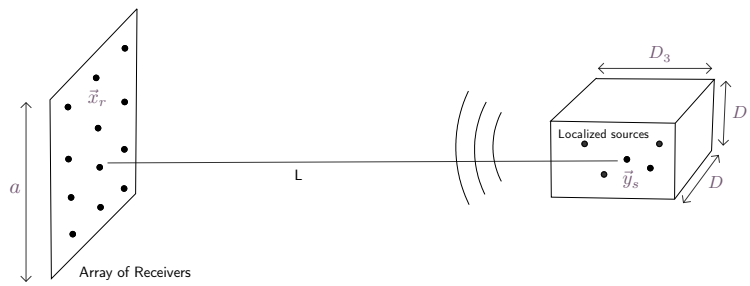

FIG. 2.1. *Schematic for the inversion setup with an array of receivers that is planar square of side a. The range direction is orthogonal to the array aperture. The sources are at ranges of order L in an imaging region $\mathfrak{D}$ modeled as a rectangular prism with size D_3 in range and D in cross-range.*

transport mean free path defines the range limit of applicability of coherent inversion methods. Beyond it only incoherent methods based on transport or diffusion equations [2] can be used.

In this paper we assume a scattering regime where the CINT method is useful. It forms images by superposing cross-correlations of the array measurements, delayed by travel times between the receivers and the imaging points. As shown in [9, 10, 6], the cross-correlations must be computed locally, in appropriate time windows, and over limited receiver offsets. This introduces a smoothing in the CINT image formation, which is essential for stabilizing statistically the images, at the expense of resolution. The larger the random distortions of the array measurements, the more smoothing is needed and the worse the resolution [9, 10]. Thus, it is natural to ask if it is possible to improve the source localization by using the prior information that the sources have small support.

We show that under generic conditions, the CINT imaging function is approximately a discrete convolution of the vector of source intensities discretized on the imaging mesh, with a blurring kernel. To reconstruct the sources we seek to undo the convolution using convex (l_1) optimization. We present an analysis of the method in a scattering regime where the random medium effects on the array measurements can be modeled by large wavefront distortions, as assumed in adaptive optics [3]. We derive from first principles the CINT blurring kernel in this regime, and state the inversion problem as an l_1 optimization. We also quantify the quality of the reconstruction with error estimates that depend on the separation of the sources, or of clusters of sources, but are independent of the source placement on or off the imaging mesh, as long as the sources are sufficiently far apart. The analysis shows that we can expect almost exact reconstructions when the sources are further apart than the CINT resolution limits. This is similar to the super-resolution results in [11], that show that one dimensional discrete convolutions can be undone by convex optimization, assuming that the minimum distance between the points in the support of the unknown vectors is $2/f_c$, where f_c is the largest "frequency" in the Fourier transform of the convolution kernel. When the sources are clustered together, the l_1 reconstruction is not guaranteed to be close to the true vector of source intensities in the point-wise sense. However, we show that its support is in the vicinity of the clusters, and its entries are related to the average source intensities there.

The paper is organized as follows: We begin in section 2 with the formulation of the inverse problem as an l_1 optimization. The analysis of the method is in sections 3 - 5, and we demonstrate its performance with numerical simulations in section 6. We end with a summary in section 7.

2. Formulation of the inverse problem. Consider the inversion setup illustrated in Figure 2.1, where N_s sources located at $\vec{\boldsymbol{y}}_s$, for $s = 1, \ldots, N_s$, emit signals $f_s(t)$ that generate sound waves recorded at a remote array of receivers placed at $\vec{\boldsymbol{x}}_r$, for $r = 1, \ldots, N_r$. For simplicity we assume that the array aperture is planar and square, of side a. This allows us to introduce a system of coordinates centered at the array, with range direction orthogonal to the array, and cross-range plane parallel to the array. In this system of coordinates we have $\vec{\boldsymbol{x}}_r = (\boldsymbol{x}_r, 0)$, with cross-range vectors $\boldsymbol{x}_r = (x_{r,1}, x_{r,2})$ satisfying $|x_{r,1}|, |x_{r,2}| \leq a/2$. The sources are at $\vec{\boldsymbol{y}}_s = (\boldsymbol{y}_s, y_{s,3})$, with range coordinates $y_{s,3}$ of order L, satisfying $L \gg a$, and two dimensional cross-range vectors $\boldsymbol{y}_s$.

In general, the signals $f_s(t)$ emitted by the sources may be pulses, chirps or even noise-like, with Fourier transforms

$$\widehat{f}_s(\omega) = \int_{-\infty}^{\infty} dt\, f_s(t) e^{i\omega t} \tag{2.1}$$

supported in the frequency interval $[\omega_o - \pi B, \omega_o + \pi B]$, where ω_o is the central frequency and B is the bandwidth. We denote the recorded waves by $p(\vec{\boldsymbol{x}}_r, t)$, and use the linearity of the wave equation to write

$$p(\vec{\boldsymbol{x}}_r, t) = \int_{\omega_o - \pi B}^{\omega_o + \pi B} \frac{d\omega}{2\pi}\, \widehat{p}(\vec{\boldsymbol{x}}_r, \omega) e^{-i\omega t}, \qquad \widehat{p}(\vec{\boldsymbol{x}}_r, \omega) = \sum_{s=1}^{N_s} \widehat{f}_s(\omega) \widehat{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s, \omega) + \widehat{n}(\vec{\boldsymbol{x}}_r, \omega), \tag{2.2}$$

for $r = 1, \ldots, N_r$. Here the Green's function $\widehat{G}$ models the propagation of time harmonic waves in the medium, and $\widehat{n}$ denotes additive and uncorrelated noise. The inverse problem is to determine the sources from these array measurements.

2.1. Imaging in homogeneous media. The Green's function in media with constant speed c_o is

$$\widehat{G}_o(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s, \omega) = \frac{\exp[i\omega\tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s)]}{4\pi|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}_s|}, \tag{2.3}$$

where

$$\tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s) = |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}_s|/c_o \tag{2.4}$$

is the travel time from the source at $\vec{\boldsymbol{y}}_s$ to the receiver at $\vec{\boldsymbol{x}}_r$. The measurements are of the form

$$p_o(\vec{\boldsymbol{x}}_r, t) = \sum_{s=1}^{N_s} \frac{f_s\left(t - \tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s)\right)}{4\pi|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}_s|} + n(\vec{\boldsymbol{x}}_r, t), \tag{2.5}$$

and in reverse time migration they are synchronized using travel time delays with respect to a presumed source location at a search point $\vec{\boldsymbol{y}}$, and then superposed to form an image

$$\mathcal{J}_o(\vec{\boldsymbol{y}}) = \sum_{r=1}^{N_r} p_o\left(\vec{\boldsymbol{x}}_r, \tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}})\right). \tag{2.6}$$

The resolution limits of the imaging function (2.6) are well known. The cross-range resolution is of order $\lambda_o L/a$, where $\lambda_o = 2\pi c_o/\omega_o$ is the central wavelength, and the range resolution is inverse proportional to the temporal support of $f_s(t)$, which determines the precision of the travel time estimation. If the signals $f_s(t)$ are pulses, their temporal support is of order $1/B$, and the range resolution is of order c_o/B. If they are chirps or other long signals that are known, then they must be compressed in time by cross-correlation (matched filtering) with the time reversed $f_s(t)$ to achieve the c_o/B range resolution [15]. If the signals are unknown and noise-like, then imaging must be based on cross-correlations of the array measurements, like in CINT.

We refer to [7] for the formulation of the inverse source problem as an l_1 optimization, and recall from there that the resolution limits $\lambda_o L/a$ and c_o/B also play a role in the successful recovery of the presumed sparse source support.

2.2. Coherent interferometric imaging. The imaging function $\mathcal{J}_o(\vec{\boldsymbol{y}})$ does not work well in random media at ranges L that exceed a few scattering mean free paths. This is because the measurements have large random distortions that are very different than the additive noise $n(\vec{\boldsymbol{x}}_r, t)$, and cannot be reduced by simply summing over the receivers as in (2.6). To mitigate these distortions we image with the CINT function

$$\mathcal{J}(\vec{\boldsymbol{y}}) = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\widetilde{\omega}\, \widehat{\phi}\left(\frac{\widetilde{\omega}}{\Omega}\right) \sum_{r=1}^{N_r} \sum_{r'=1}^{N_r} \psi\left(\frac{|\boldsymbol{x}_r - \boldsymbol{x}_{r'}|}{X(\omega)}\right) \widehat{p}(\vec{\boldsymbol{x}}_r, \omega + \widetilde{\omega}/2) \overline{\widehat{p}(\vec{\boldsymbol{x}}_{r'}, \omega - \widetilde{\omega}/2)}$$
$$\times \exp[-i(\omega + \widetilde{\omega}/2)\tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}) + i(\omega - \widetilde{\omega}/2)\tau(\vec{\boldsymbol{x}}_{r'}, \vec{\boldsymbol{y}})], \tag{2.7}$$

where $\widehat{\phi}$ and ψ are smooth window functions of dimensionless argument and support of order one, and the domain of integration is restricted by the finite bandwidth that supports the measurements,

$$\omega \pm \widetilde{\omega}/2 \in [\omega_o - \pi B, \omega_o + \pi B].$$

As in reverse time migration, the travel times are used in (2.7) to synchronize the waves due to a presumed source at the search location $\vec{\boldsymbol{y}}$. However, the image is formed by superposing cross-correlations of the array measurements $p(\vec{\boldsymbol{x}}_r, t)$, instead of the measurements themselves. The cross-correlations are convolutions of $p(\vec{\boldsymbol{x}}_r, t)$ with the time reversed $p(\vec{\boldsymbol{x}}_{r'}, t)$, for $r, r' = 1, \ldots, N_r$. The time reversal appears as complex conjugation in the frequency domain, denoted with the bar in (2.7). The time window

$$\phi\left(\Omega t\right) = \frac{1}{2\pi\Omega} \int_{-\infty}^{\infty} d\omega\, \widehat{\phi}\left(\frac{\omega}{\Omega}\right) e^{-i\omega t} \tag{2.8}$$

and spatial window $\psi(|\boldsymbol{x}|/X)$ ensure that the cross-correlations are computed locally, over receiver offsets that do not exceed the distance X, and over time offsets of order $1/\Omega$. These threshold parameters account for the fact that scattering in random media decorrelates statistically the waves at frequencies separated by more than Ω_d, the decoherence frequency, and points separated by more than X_d, the decoherence length. We refer to [10, 6] and the next section for more details. Here it suffices to recall that (2.7) is robust* when $X \lesssim X_d$ and $\Omega \lesssim \Omega_d$, and that the image $\mathcal{J}(\vec{\boldsymbol{y}})$ has a cross-range resolution of order $\lambda_o L/X$ and a range resolution of order c_o/Ω. The best focus occurs at $X \approx X_d$ and $\Omega \approx \Omega_d$, so the decoherence parameters X_d and Ω_d can be estimated with optimization, as explained in [9].

To state the inverse problem as a convex optimization, we make the simplifying assumption that the sources emit the same known pulse $f(t)$, so that

$$\widehat{f}_s(\omega) = \widehat{f}(\omega)\rho(\vec{\boldsymbol{y}}_s), \qquad s = 1, \ldots, N_s, \tag{2.9}$$

for an unknown, complex valued amplitude $\rho(\vec{\boldsymbol{y}})$. Using (2.9) in (2.2) and substituting in (2.7) we obtain

$$\mathcal{J}(\vec{\boldsymbol{y}}) \approx \sum_{s=1}^{N_s} \sum_{s'=1}^{N_s} \rho(\vec{\boldsymbol{y}}_s)\overline{\rho(\vec{\boldsymbol{y}}_{s'})}\, \kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{y}}_s, \vec{\boldsymbol{y}}_{s'}), \tag{2.10}$$

with kernel

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{y}}_s, \vec{\boldsymbol{y}}_{s'}) = &\sum_{r=1}^{N_r} \sum_{r'=1}^{N_r} \psi\left(\frac{|\boldsymbol{x}_r - \boldsymbol{x}_{r'}|}{X}\right) \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\widetilde{\omega}\, \widehat{\phi}\left(\frac{\widetilde{\omega}}{\Omega}\right) \widehat{f}\left(\omega + \widetilde{\omega}/2\right) \overline{\widehat{f}\left(\omega - \widetilde{\omega}/2\right)} \\ &\times \widehat{G}\left(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}_s, \omega + \widetilde{\omega}/2\right) \overline{\widehat{G}(\vec{\boldsymbol{x}}_{r'}, \vec{\boldsymbol{y}}_{s'}, \omega - \widetilde{\omega}/2)} \exp[-i(\omega + \widetilde{\omega}/2)\tau(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}})] + i(\omega - \widetilde{\omega}/2)\tau(\vec{\boldsymbol{x}}_{r'}, \vec{\boldsymbol{y}})],\end{aligned} \tag{2.11}$$

where the approximation is because we neglect the additive noise†.

In the analysis of the next section we take the Gaussian pulse

$$\widehat{f}(\omega) = \left(\frac{\sqrt{2\pi}}{B}\right)^{1/2} \exp\left[-\frac{(\omega - \omega_o)^2}{4B^2}\right], \tag{2.12}$$

normalized by

$$\|f\|_2 = \left[\int_{-\infty}^{\infty} dt\, |f(t)|^2\right]^{1/2} = \left[\int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, |\widehat{f}(\omega)|^2\right]^{1/2} = 1.$$

*Robustness refers to statistical stability of the image with respect to the realizations of the random medium. It means that the standard deviation of (2.7) is small with respect to its expectation near the peaks.

†Additive noise is considered in all the numerical simulations in section 6, but for simplicity we neglect it in the analysis.

This choice allows us to obtain an explicit expression of the kernel (2.11). Naturally, in practice the pulses may not be Gaussian, and the sources may emit different signals. The method described here still applies to such cases, with $\rho(\vec{\boldsymbol{y}}_s)$ replaced in (2.10) by $\|f_s\|_2$, and the substitution

$$\widehat{f}(\omega+\widetilde{\omega}/2)\overline{\widehat{f}(\omega-\widetilde{\omega}/2)} \rightsquigarrow \frac{\widehat{f}_s(\omega+\widetilde{\omega}/2)}{\|f_s\|}\frac{\overline{\widehat{f}_{s'}(\omega-\widetilde{\omega}/2)}}{\|f_{s'}\|_2}$$

in (2.11). Since f_s is unknown in general, we may only estimate the kernel κ up to unknown, constant multiplicative factors. This still allows the determination of the location of the sources, but does not give good estimates of their intensities.

2.3. The optimization problem. Let us consider a reconstruction mesh with N_z points denoted generically by $\vec{\boldsymbol{z}}$, and name $\boldsymbol{\rho}$ the column vector with entries given by the unknown $\rho(\vec{\boldsymbol{z}})$. We sample the CINT image $\mathcal{J}(\vec{\boldsymbol{y}})$ at $N_y < N_z$ points, and gather these samples in the "data" vector $\boldsymbol{d}$. It is natural to take $N_y < N_z$, because we seek to super-resolve the CINT image, which is blurred by the kernel (4.4).

At first glance it appears that we may use equation (2.10) to formulate the inversion as an optimization problem for recovering the rank one matrix $\boldsymbol{\rho}\boldsymbol{\rho}^\star$, as in [12, 13]. However, as shown in the next section, in strong random media where CINT is needed, the kernel $\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}')$ is large only when $\vec{\boldsymbol{z}}$ and $\vec{\boldsymbol{z}}'$ are nearby. In fact, for reasonable mesh sizes on which we can expect to obtain unique reconstructions, the kernel satisfies

$$\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx \begin{cases} \kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}) & \text{if } \vec{\boldsymbol{z}}'=\vec{\boldsymbol{z}}, \\ 0 & \text{if } \vec{\boldsymbol{z}}\neq\vec{\boldsymbol{z}}'. \end{cases} \tag{2.13}$$

Thus, only the diagonal entries $|\rho(\vec{\boldsymbol{z}})|^2$ of $\boldsymbol{\rho}\boldsymbol{\rho}^\star$ play a role. These are the source intensities and we denote by $\boldsymbol{u}\in\mathbb{R}^{N_z}$ the vector of unknowns formed by them. Equation (2.10) gives

$$\boldsymbol{\mathcal{M}}\,\boldsymbol{u} \approx \boldsymbol{d}, \tag{2.14}$$

where $\boldsymbol{\mathcal{M}}$ is the $N_y\times N_z$ "measurement" matrix with entries $\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}})$. We formulate the inversion as the l_1 optimization problem

$$\min_{\boldsymbol{u}\in\mathbb{R}^{N_z}} \|\boldsymbol{u}\|_1 \;\text{ such that }\; \|\boldsymbol{\mathcal{M}}\,\boldsymbol{u}-\boldsymbol{d}\|_2 \leq \text{ tolerance}, \tag{2.15}$$

where the tolerance accounts for additive noise effects and the random fluctuations of the CINT imaging function, which are small for large enough array aperture a and bandwidth B, as shown in [10, 6].

We prove in section 4 that the left hand side in (2.14) is approximately a discrete convolution. The l_1 optimization is useful in this context, and recovers well sources that are well separated, as expected from the results in [11]. We rediscover such results in this paper using a different analysis. We also consider cases where the sources are clustered together, and show that although we cannot expect good reconstructions in the point-wise sense, the l_1 minimizer is supported in the vicinity of the clusters, and estimates the average source intensity there.

3. Setup of the analysis. We introduce in this section the random wave speed model and the scaling assumptions which define the relations between the wavelength λ_o, the typical size ℓ of the inhomogeneities in the medium, the standard deviation σ of the fluctuations of the wave speed, the array aperture a, the range L, and the extent of the imaging region. The scaling allows us to describe the scattering effects of the random medium as large wavefront distortions. This is a simple wave propagation model that is convenient for analysis, and captures qualitatively all the important features of imaging with CINT. That is to say, equation (2.14) holds in general scattering regimes, and the CINT kernel κ has a similar form, but the mathematical expression of the decoherence length X_d and frequency Ω_d, which quantify the blurring by the kernel, are expected to change. The expressions of X_d and Ω_d in terms of λ_o,σ,ℓ and L are needed for analysis, but they are unlikely to be known in practice. This is why one should determine the decoherence parameters directly form the data or adaptively, during the CINT image formation, as in [9].

The model of the wave speed $c(\vec{\boldsymbol{x}})$ is

$$c(\vec{\boldsymbol{x}}) = c_o\left[1+\sigma\mu\left(\frac{\vec{\boldsymbol{x}}}{\ell}\right)\right]^{-1/2}, \tag{3.1}$$

where μ is a mean zero, stationary random process of dimensionless argument. We suppose that μ is bounded almost surely and denote by $\mathcal{R}$ its auto-correlation, assumed isotropic and Gaussian for convenience

$$\mathcal{R}(\vec{\boldsymbol{x}}) = \mathbb{E}\left[\mu(\vec{\boldsymbol{x}} + \vec{\boldsymbol{x}}')\mu(\vec{\boldsymbol{x}})\right] = e^{-|\vec{\boldsymbol{x}}|^2/2}. \tag{3.2}$$

Then, $\sigma \ll 1$ quantifies the small amplitude of the fluctuations of $c(\vec{\boldsymbol{x}})$, and ℓ is the correlation length, which characterizes the typical size of the inhomogeneities in the medium.

We explain in section 3.1 how far the waves should propagate in media modeled by (3.1), so that the cumulative scattering effects can be described by large wavefront distortions. We are interested in imaging with finite size arrays at long distances, so the waves propagate in the range direction, within a cone (beam) of small opening angle. This is the paraxial regime defined in section 3.2. We describe in section 3.3 the wave randomization quantified by the scattering mean free path, and the statistical decorrelation quantified by the decoherence length and frequency. We end the section with a summary of the scaling assumptions.

3.1. Wave scattering regime. We use a geometrical optics (Rytov) wave propagation model that holds in high frequency regimes with separation of scales

$$\lambda_o \ll \ell \ll L, \tag{3.3}$$

and standard deviation σ of the fluctuations satisfying

$$\sigma \ll \min\left\{\left(\frac{\ell}{L}\right)^{3/2}, \frac{\sqrt{\ell\lambda_o}}{L}\right\}. \tag{3.4}$$

It is shown in [22] that the first bound in (3.4) ensures that the rays remain straight and the variance of the amplitude of the Green's function is negligible, so we can use the same geometrical spreading factor as in the homogeneous medium. The second bound in (3.4) ensures that only first order (in σ) corrections of the travel time matter, so the propagation model is

$$\widehat{G}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) \approx \widehat{G}_o(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) \exp\left[i\omega\tau(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\frac{\sigma}{2}\int_0^1 d\vartheta\, \mu\left(\frac{(1-\vartheta)\vec{\boldsymbol{y}} + \vartheta\vec{\boldsymbol{x}}}{\ell}\right)\right]. \tag{3.5}$$

Let us write the random phase correction in (3.5) as

$$\omega\delta\tau(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) = \frac{(2\pi)^{1/4}}{2}\sigma k\sqrt{\ell|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|}\,\nu(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}), \tag{3.6}$$

where $k = \omega/c_o$ is the wavenumber and

$$\nu(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) = \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|}{\ell}}\int_0^1 d\vartheta\, \mu\left(\frac{(1-\vartheta)\vec{\boldsymbol{y}} + \vartheta\vec{\boldsymbol{x}}}{\ell}\right) \tag{3.7}$$

is defined by the integral of the fluctuations along the straight ray between $\vec{\boldsymbol{y}}$ and $\vec{\boldsymbol{x}}$. It is shown in [6, Lemma 3.1] that $\nu(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})$ converges in distribution as $\ell/|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}| \sim \ell/L \to 0$ to a Gaussian process. Its mean

$$\mathbb{E}\left[\nu(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\right] = 0, \tag{3.8}$$

and variance

$$\mathbb{E}\left[\nu^2(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\right] = \frac{|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|}{\ell\sqrt{2\pi}}\int_0^1 d\vartheta \int_0^1 d\vartheta'\, \mathcal{R}\left(\frac{(\vartheta' - \vartheta)(\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}})}{\ell}\right) \approx 1, \tag{3.9}$$

are calculated from definition (3.7) and the expression (3.2) of $\mathcal{R}$. The approximation is for ℓ/L small.

We conclude from (3.6), (3.9) and $k = O(k_o)$, with $k_o = 2\pi/\lambda_o$, that the random phase fluctuations in (3.5) have standard deviation of order $\sigma\sqrt{\ell L}/\lambda_o$. When this is small, the random medium effects are

negligible and any coherent imaging method works well. We are interested in the case of large fluctuations, so we ask that

$$\sigma \gg \frac{\lambda_o}{\sqrt{\ell L}}. \tag{3.10}$$

This is consistent with (3.4) when

$$\frac{\lambda_o/\sqrt{\ell L}}{(\ell/L)^{3/2}} = \frac{\lambda_o L}{\ell^2} \ll 1, \qquad \frac{\lambda_o/\sqrt{\ell L}}{\sqrt{\ell\lambda_o}/L} = \frac{\sqrt{\lambda_o L}}{\ell} \ll 1,$$

so we tighten our assumption (3.3) on the correlation length as

$$\sqrt{\lambda_o L} \ll \ell \ll L. \tag{3.11}$$

3.2. The paraxial regime. Suppose that the sources are contained in the search (imaging) region

$$\mathfrak{D} = [-D/2, D/2] \times [-D/2, D/2]\,[-D_3/2, D_3/2]\,, \tag{3.12}$$

which is a rectangular prism of sides D in cross-range and D_3 in range, as illustrated in Figure 2.1. When D and the array aperture a are small with respect to the range scale L, the rays connecting the sources and the receivers are contained within a cone (beam) of small opening angle, and we can use the paraxial approximation to simplify the calculations.

The paraxial regime is defined by the scaling relations

$$\lambda_o \ll D \lesssim a \ll L, \qquad D_3 \ll L, \qquad \frac{a^4}{\lambda_o L^3} \ll 1, \qquad \frac{a^2 D_3}{\lambda_o L^2} \ll 1, \tag{3.13}$$

so that for $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$ and $\vec{\boldsymbol{x}} = (\boldsymbol{x}, 0)$ we get

$$\begin{aligned} k|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}| &= k\left(y_3 + \frac{|\boldsymbol{x}|^2}{2L} - \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{L} + \frac{|\boldsymbol{y}|^2}{2L}\right) + O\left(\frac{a^4}{\lambda_o L^3}\right) + O\left(\frac{a^2 D_3}{\lambda_o L^2}\right) \\ &\approx k\left(y_3 + \frac{|\boldsymbol{x}|^2}{2L} - \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{L} + \frac{|\boldsymbol{y}|^2}{2L}\right), \end{aligned} \tag{3.14}$$

and

$$\frac{1}{4\pi|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|} = \frac{1}{4\pi L}\left[1 + O\left(\frac{D_3}{L}\right) + O\left(\frac{a^2}{L^2}\right)\right] \approx \frac{1}{4\pi L}. \tag{3.15}$$

These approximations are proved in appendix A, and the deterministic factor in (3.5) becomes

$$\widehat{G}_o(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) \approx \frac{1}{4\pi L}\exp\left[ik\left(y_3 + \frac{|\boldsymbol{x}|^2}{2L} - \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{L} + \frac{|\boldsymbol{y}|^2}{2L}\right)\right]. \tag{3.16}$$

3.3. Randomization and statistical decorrelation of the waves. Here we quantify the scattering mean free path $\mathcal{S}$, the length scale on which the waves randomize (lose coherence), and the decoherence length X_d and frequency Ω_d, which describe the statistical decorrelation of the waves due to scattering. These important scales appear in the definition of the CINT blurring kernel defined in section 4.

Proposition 3.1. *The expectation of the Green's function* (3.5) *is given by*

$$\mathbb{E}\big[\widehat{G}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega)\big] \approx \widehat{G}_o(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) e^{-\frac{|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|}{\mathcal{S}(\omega)}} \approx 0, \tag{3.17}$$

where $\mathcal{S}(\omega)$ is the scattering mean free path defined by

$$\mathcal{S}(\omega) = \frac{8}{\sqrt{2\pi}\sigma^2 k^2 \ell} \ll L. \tag{3.18}$$

This result, proved in appendix B, shows that the wavefront distortions due to scattering in the random medium do not average out. The expectation of $\widehat{G}$ is not the same as the Green's function in the homogeneous medium, but decays exponentially with the distance of propagation on the scale $\mathcal{S}(\omega)$, the scattering mean free path. The scaling assumption (3.10) and definition (3.18) give

$$\frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}{\mathcal{S}(\omega)} = O\left(\frac{L}{\mathcal{S}(\omega)}\right) = O(\sigma^2 k^2 \ell L) \gg 1,$$

which is why the expectation in (3.17) is almost zero. The standard deviation of the fluctuations is approximately

$$\text{std}[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)] \approx \sqrt{\left|\widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\right|^2 - \left|\mathbb{E}\big[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\big]\right|^2} \approx \left|\widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\right|,$$

where we used that $|\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)| \approx |\widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)|$. Thus, the random fluctuations of the waves dominate their coherent part (the expectation) at the ranges considered in our analysis,

$$\frac{\left|\mathbb{E}\big[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\big]\right|}{\text{std}[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)]} \approx e^{-\frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}{S(\omega)}} \approx 0,$$

and the wave is randomized. Reverse time migration or standard l_1 optimization methods cannot mitigate these large random distortions, as we illustrate with numerical simulations. This is why we base our inversion on the CINT method.

PROPOSITION 3.2. *Consider two points $\vec{\boldsymbol{x}} = (\boldsymbol{x}, 0)$ and $\vec{\boldsymbol{x}}' = (\boldsymbol{x}', 0)$ in the array aperture and two points $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$ and $\vec{\boldsymbol{y}}' = (\boldsymbol{y}', y_3')$ in the imaging region. Assume that the bandwidth B is small with respect to the central frequency ω_o, so that $|\omega-\omega_o|, |\omega'-\omega_o| \le \pi B \ll \omega_o$. Then, the second moments of* (3.5) *are*

$$\mathbb{E}\left[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\overline{\widehat{G}(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega')}\right] \approx \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\overline{\widehat{G}_o(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega')} e^{-\frac{|y_3-y_3'|}{\mathcal{S}} - \frac{(\omega-\omega')^2}{2\Omega_d^2} - \frac{|\boldsymbol{y}-\boldsymbol{y}'|^2 + (\boldsymbol{y}-\boldsymbol{y}')\cdot(\boldsymbol{x}-\boldsymbol{x}') + |\boldsymbol{x}-\boldsymbol{x}'|^2}{2X_d^2}}, \quad (3.19)$$

with short notation $\mathcal{S} = \mathcal{S}(\omega_o)$, and decoherence frequency Ω_d and decoherence length X_d defined by

$$\Omega_d = \frac{2\omega_o}{(2\pi)^{5/4}}\left(\frac{\lambda_o}{\sigma\sqrt{\ell L}}\right) \ll \omega_o, \qquad X_d = \sqrt{3}\ell\,\frac{\Omega_d}{\omega_o} \ll \ell. \quad (3.20)$$

Moment formula (3.19) is proved in appendix B, and the inequalities in (3.20) are due to assumption (3.10). The first exponential factor in (3.19) accounts for the randomization due to the travel time fluctuations between the two ranges. In our scaling $|y_3 - y_3'| \lesssim D_3$, and by the last inequality in (3.13), and the definition of the scattering mean free path, we have

$$\frac{|y_3-y_3'|}{\mathcal{S}(\omega_o)} = O\left(D_3\frac{\sigma^2\ell}{\lambda_o^2}\right) \ll O\left(\frac{\lambda_o L^2}{a^2}\frac{\sigma^2\ell}{\lambda_o^2}\right) \ll O\left(\frac{\ell^2}{a^2}\right), \quad (3.21)$$

where we used the bound (3.4) on σ. It is shown in [6, Section 4] that the standard deviation of the CINT image is small with respect to the expectation of its peak value (i.e., the imaging function is statistically stable) when

$$a > \ell. \quad (3.22)$$

Stability is essential for imaging to succeed, so we ask that the array aperture satisfy (3.22), and conclude from (3.21) that the second moments (3.17) simplify as[‡]

$$\mathbb{E}\left[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\overline{\widehat{G}(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega')}\right] \approx \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\overline{\widehat{G}_o(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega')} e^{-\frac{(\omega-\omega')^2}{2\Omega_d^2} - \frac{|\boldsymbol{y}-\boldsymbol{y}'|^2 + (\boldsymbol{y}-\boldsymbol{y}')\cdot(\boldsymbol{x}-\boldsymbol{x}') + |\boldsymbol{x}-\boldsymbol{x}'|^2}{2X_d^2}}, \quad (3.23)$$

[‡]Although this moment formula is derived here using the model (3.5), the result is generic and can be obtained in other scattering regimes. The only difference is the definition of X_d and Ω_d. See for example [21].

with $\widehat{G}_o$ given in (3.16).

The exponential decay in (3.23) models the statistical decorrelation of the waves due to scattering. In our context, the spatial decorrelation, modeled by the decay in $\boldsymbol{x}-\boldsymbol{x}'$ and $\boldsymbol{y}-\boldsymbol{y}'$, can be explained by the fact that rays connecting sources to far apart receivers traverse through different parts of the random medium. Because μ does not have long range correlations, the fluctuations of the travel time along such different rays are statistically uncorrelated. The waves at far apart frequencies are uncorrelated because they interact differently with the random medium. This gives the decay in $\omega-\omega'$ in equation (3.23).

Definition (2.11) of the CINT kernel involves the superposition of $\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\overline{\widehat{G}(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega')}$ over the array elements and frequencies. If the array aperture a is large with respect to X_d, the superposition stabilizes statistically because we sum many uncorrelated entries, as in the law of large numbers. This is why CINT is robust with respect to the uncertainty of the fluctuations of the wave speed, as shown in [10, 6].

3.4. Summary of the scaling assumptions. We gather here the scaling assumptions stated throughout the section, and complement them with extra assumptions on the bandwidth and the size of the imaging region. We refer to appendix C for the verification of their consistency.

The wavelength λ_o is the smallest length scale, and the range L is the largest. The assumptions (3.22) and (3.13) on the aperture are

$$\ell < a \ll (\lambda_o L^3)^{1/4}. \tag{3.24}$$

The upper bound ensures that we can use the paraxial approximation and the lower bound gives $a > \ell \gg X_d$, so that the CINT image is statistically stable.

Assumption (3.11) combined with (3.24) gives that the correlation length of the wave speed fluctuations should satisfy

$$\sqrt{\lambda_o L} \ll \ell \ll (\lambda_o L^3)^{1/4} \ll L. \tag{3.25}$$

The standard deviation σ of the fluctuations is bounded above as in (3.4), and below as in (3.10),

$$\frac{\lambda_o}{\sqrt{\ell L}} \ll \sigma \ll \frac{\sqrt{\ell\lambda_o}}{L}. \tag{3.26}$$

The cross-range and range sizes D and D_3 of the imaging region should be large with respect to the CINT resolution limits of $\lambda_o L/X$ in cross-range and c_o/Ω in range (see next section), so we can observe the image focus. We take the threshold parameters

$$X/X_d = O(1), \qquad \Omega/\min\{\Omega_d, B\} = O(1), \tag{3.27}$$

and recalling the scaling assumptions (3.13) that allow us to use the paraxial approximation, we obtain

$$\frac{c_o}{\Omega} \ll D_3 \ll \frac{\lambda_o L^2}{a^2} \quad \text{and} \quad \frac{\lambda_o L}{X} \ll D \lesssim a. \tag{3.28}$$

In general, the CINT image is statistically stable if in addition to having $a \gg X_d$, which follows from (3.20) and (3.24), the bandwidth B is larger than the decoherence frequency Ω_d. However, for the propagation model (3.5) considered in this section, where the effect of the random medium consists only of wavefront distortions and no delay spread (reverberations), the bandwidth B does not play a role in the stabilization of CINT, as shown in [6, Section 4.4.4]. Thus, we study imaging in both narrowband and broadband regimes:

The narrowband regime is defined by B satisfying

$$\omega_o\left(\frac{a}{L}\right)^2 \ll B \ll \omega_o \min\left\{1, \frac{\lambda_o L}{aX_d}\right\} = \omega_o \frac{\lambda_o L}{aX_d}. \tag{3.29}$$

As verified in Appendix C,

$$\frac{\lambda_o L}{X_d} \ll \ell < a, \qquad \frac{\lambda_o L}{aX_d} \ll \frac{\Omega_d}{\omega_o}, \tag{3.30}$$

so $B \ll \Omega_d$. This choice leads to a simpler expression of the CINT blurring kernel, but since Ω is of the order of B, the range resolution is the same as in the homogeneous medium, and cannot be improved with optimization unless the sources are very far apart in range. However, the optimization can improve the cross-range focusing.

The broadband regime is defined by

$$\Omega_d \ll B \ll \omega_o, \tag{3.31}$$

so we may seek to improve the CINT resolution in both range and cross-range. The expression of the CINT kernel is more complicated in this case, but it simplifies slightly when

$$\frac{\lambda_o}{\sqrt{\ell L}} \ll \frac{\lambda_o^{2/3}\ell^{1/6}}{L^{5/6}} \ll \sigma \ll \frac{\sqrt{\ell\lambda_o}}{L}, \qquad \frac{\ell}{a} = O(1). \tag{3.32}$$

We present the analysis that uses these conditions, which say that the fluctuations in the random medium are even stronger than in (3.26), but the correlation length is not much smaller than a. Extensions to larger apertures are possible, although the analysis is more complicated.

4. The CINT blurring kernel. Here we derive the CINT convolution model. To obtain an explicit expression of the kernel (2.11), we use the Gaussian pulse (2.12) and the Gaussian threshold windows

$$\widehat{\phi}\left(\frac{\widetilde{\omega}}{\Omega}\right) = e^{-\frac{\widetilde{\omega}^2}{2\Omega^2}}, \qquad \psi\left(\frac{|\widetilde{\boldsymbol{x}}|}{X}\right) = e^{-\frac{|\widetilde{\boldsymbol{x}}|^2}{2\Omega^2}}, \tag{4.1}$$

with X and Ω satisfying

$$X/X_d = O(1), \qquad \Omega = \begin{cases} B & \text{in narrowband regime} \\ O(\Omega_d) & \text{in broadband regime.} \end{cases} \tag{4.2}$$

As stated previously, and shown in [9], X_d and Ω_d can be estimated adaptively, by optimizing the focusing of the CINT image. This is why we can assume that X_d is known approximately. The same holds for Ω_d, if the bandwidth is big enough. The expression of the CINT kernel is simpler in the narrowband scaling (3.29), where $B \ll \Omega_d$, as shown in section 4.1, and we take $\Omega = B$. The broadband regime is discussed in section 4.2.

Typically, the receivers are separated by distances of order λ_o, so that they behave collectively as an array. Since $\lambda_o \ll a$, we have $N_r = O\left(a^2/\lambda_o^2\right) \gg 1$, and we can approximate the sums in (2.11) by integrals

$$\sum_{r=1}^{N} \rightsquigarrow \frac{N_r}{a^2}\int_{\mathcal{A}} d\boldsymbol{x},$$

where $\mathcal{A}$ denotes the array aperture, the square of side a. To avoid specifying the finite aperture in the integrals, and to simplify the calculations, we use a Gaussian apodization factor

$$\psi_{\mathcal{A}}(\boldsymbol{x}) = \exp\left[-\frac{|\boldsymbol{x}|^2}{2(a/6)^2}\right], \tag{4.3}$$

which is negligible outside the disk of radius $a/2$.

4.1. The CINT kernel in the narrowband regime. The calculation of the kernel (2.11) is in appendix D, and we state the results in the next proposition.

PROPOSITION 4.1. *Let $\vec{\boldsymbol{z}} = (\boldsymbol{z}, z_3)$, $\vec{\boldsymbol{z}}' = (\boldsymbol{z}', z_3')$ and $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$ be three points in the imaging region and define the center and difference vectors*

$$\frac{\vec{\boldsymbol{z}} + \vec{\boldsymbol{z}}'}{2} = (\overline{\boldsymbol{z}}, \overline{z}_3), \qquad \vec{\boldsymbol{z}} - \vec{\boldsymbol{z}}' = (\widetilde{\boldsymbol{z}}, \widetilde{z}_3).$$

Under the assumptions (3.24)-(3.29)*, the CINT kernel* (2.11) *is approximated by*

$$\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx C \exp\left[-\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2R^2} - \frac{(\overline{z}_3-y_3)^2}{2R_3^2}\right] \mathfrak{M}(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'), \tag{4.4}$$

where C is a constant, and

$$R = \frac{L}{k_o X_e}, \qquad R_3 = \frac{c_o}{\Omega_e}, \tag{4.5}$$

with X_e and Ω_e defined by

$$\frac{1}{X_e^2} = \frac{1}{X_d^2} + \frac{1}{X^2} + \frac{1}{4(a/6)^2}, \qquad \frac{1}{\Omega_e^2} = \frac{1}{\Omega_d^2} + \frac{1}{\Omega^2} + \frac{1}{4B^2}. \tag{4.6}$$

The factor $\mathfrak{M}(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}')$ is complex, with absolute value

$$|\mathfrak{M}(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}')| = \exp\left[-\frac{\widetilde{z}_3^2}{2\widetilde{R}^2} - \frac{|\widetilde{\boldsymbol{z}}|^2}{2}\left(\frac{1}{\gamma X_d^2} + \frac{1}{\widetilde{R}^2}\right)\right], \tag{4.7}$$

where

$$\frac{1}{\gamma} = 1 - \frac{X_e^2}{4X_d^2} > \frac{3}{4}, \quad \widetilde{R}_3 = \frac{c_o}{B}, \quad \widetilde{R} = 6\sqrt{2}\frac{L}{k_o a}. \tag{4.8}$$

The parameter R defined in (4.5) is the CINT cross-range resolution limit, the length scale of exponential decay of the kernel $\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}')$ with $\overline{\boldsymbol{z}}-\boldsymbol{y}$. Definitions (4.6), (3.20) and assumption (4.2) give that

$$X_e = O(X_d) \ll a,$$

so the resolution is worse than in homogeneous media,

$$R \gg \frac{L}{k_o a}. \tag{4.9}$$

This is due to the smoothing needed to stabilize statistically the image [9]. The goal of the convex optimization (2.15) is to overcome this blurring and localize better the sources in cross-range.

The parameter R_3 is the CINT range resolution limit. Because we are in the narrowband regime, we obtain from definition (4.6) and (4.2) that $\Omega_e \approx B$, and therefore R_3 is similar to the range resolution in homogeneous media,

$$R_3 = \frac{c_o}{\Omega_e} \approx \frac{c_o}{B}. \tag{4.10}$$

The results obtained in [7] for imaging with l_1 optimization in homogeneous media show that it is not possible to improve the c_o/B range resolution, unless the sources are very far apart in range. We cannot expect to do better in random media, so we do not seek any super-resolution in range, in the narrowband regime.

Note that by the first inequality in (3.30) we have $\widetilde{R} \ll X_d$, so the kernel decays with the offsets $\widetilde{\boldsymbol{z}}$ and $\widetilde{z}$ on the length scales $\widetilde{R}$ and $\widetilde{R}_3$. These scales are, up to a constant of order one, the resolution limits of imaging in homogeneous media.

4.2. The CINT kernel in the broadband regime. The expression of the CINT kernel is stated in the next proposition, proved in appendix D.

PROPOSITION 4.2. *Consider the same points and notation as in Proposition 4.1. Under the assumptions* (3.24)-(3.28) *and* (3.31)-(3.32)*, the CINT kernel is given by*

$$\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx \frac{C}{\sqrt{1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}}} \exp\left[-\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2R^2} - \frac{\left(\overline{z}_3-y_3+\frac{|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2}{2L}\right)^2}{2R_3^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]}\right] \mathfrak{M}(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'), \tag{4.11}$$

where C is a constant,

$$\theta = \frac{6\omega_o X_e}{\Omega_e a}, \tag{4.12}$$

and $\mathfrak{M}(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}')$ is a complex multiplicative factor with absolute value

$$|\mathfrak{M}(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}')| = \exp\left\{ -\frac{\widetilde{z}_3^2}{2\widetilde{R}_3^2} - \frac{|\widetilde{\boldsymbol{z}}|^2}{2}\left[\frac{1}{\gamma X_d^2} + \frac{1}{\widetilde{R}^2}\right] + \frac{\left|\frac{(\overline{\boldsymbol{z}}-\boldsymbol{y})}{R}\cdot\frac{\widetilde{\boldsymbol{z}}}{\widetilde{R}}\right|^2}{4\theta^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]}\right\} \tag{4.13}$$

Because $X_e = O(X_d)$ and $\Omega_e = \Omega_d$, we obtain from definitions (4.12) and (3.20) that

$$\theta = O\left(\frac{\ell}{a}\right) = O(1), \tag{4.14}$$

where we used the assumption (3.32) on the aperture.[§] The first term in the exponential in (4.11) gives the focusing in cross-range, which is the same as in the narrowband case: $|\overline{\boldsymbol{z}} - \boldsymbol{y}| = O(R)$. This means that the denominators in (4.11) are order one,

$$1 + \frac{|\overline{\boldsymbol{z}} - \boldsymbol{y}|^2}{2\theta^2 R^2} = O(1).$$

The second term in the exponential in (4.11) gives the focusing in range. In our setting we have by the paraxial approximation that

$$\overline{z}_3 - y_3 + \frac{|\overline{\boldsymbol{z}}|^2 - |\boldsymbol{y}|^2}{2L} \approx |(\overline{\boldsymbol{z}}, \overline{z}_3)| - |(\boldsymbol{y}, y_3)|,$$

so CINT estimates the distance from the center of the array to $(\overline{\boldsymbol{z}}, \overline{z}_3)$ with resolution of order R_3. Since $\Omega_e \ll B$, this resolution is worse than in homogeneous media

$$R_3 = \frac{c_o}{\Omega_e} = O\left(\frac{c_o}{B}\right) = O(\widetilde{R}_3),$$

so in the broadband regime it makes sense to seek an improvement of both the range and cross-range resolution with optimization.

Equations (4.11) and (4.13) show that the kernel decays with the offset $\widetilde{z}_3$ on the same scale $\widetilde{R}_3$ as before. To see the decay with the offset $\widetilde{\boldsymbol{z}}$, we note that the last two terms in (4.13) satisfy

$$\begin{aligned}\frac{|\widetilde{\boldsymbol{z}}|^2}{2}\left[\frac{1}{\gamma X_d^2} + \frac{1}{\widetilde{R}^2}\right] - \frac{\left|\frac{(\overline{\boldsymbol{z}}-\boldsymbol{y})}{R}\cdot\frac{\widetilde{\boldsymbol{z}}}{\widetilde{R}}\right|^2}{4\theta^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]} &= \frac{|\widetilde{\boldsymbol{z}}|^2}{2R^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]}\left[1 + \frac{\widetilde{R}^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]}{\gamma X_d^2}\right] + \frac{\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{R^2}\frac{|\widetilde{\boldsymbol{z}}|^2}{\widetilde{R}^2} - \left|\frac{(\overline{\boldsymbol{z}}-\boldsymbol{y})}{R}\cdot\frac{\widetilde{\boldsymbol{z}}}{\widetilde{R}}\right|^2}{4\theta^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]} \\ &\gtrsim \frac{|\widetilde{\boldsymbol{z}}|^2}{2R^2\left[1+\frac{|\overline{\boldsymbol{z}}-\boldsymbol{y}|^2}{2\theta^2R^2}\right]},\end{aligned}$$

where we used that $\widetilde{R} \ll X_d$, as explained in the previous section. This shows that the kernel decays with the cross-range offsets on the same scale $\widetilde{R}$ as before.

[§]In the narrowband case the aperture may be much larger than ℓ, as in (3.24). It is only in the broadband case that we take $\ell = O(a)$ to simplify the expression of the CINT kernel.

4.3. The approximate convolution model. Let us discretize the imaging region $\mathfrak{D}$ defined in (3.12) on a mesh with size $\vec{\boldsymbol{h}} = (h, h, h_3)$. In principle, the steps h and h_3 may be chosen arbitrarily small, to avoid discretization error due to sources being off the mesh. However, we know from [7] and the analysis below and the numerical simulations that we cannot expect reconstructions at scales that are finer than the resolution limits in homogeneous media. This motivates us to formulate the inversion using the assumption that the sources are further apart than $3\widetilde{R}$ in cross-range and $3\widetilde{R}_3$ in range. This leads to a simpler optimization problem because by Propositions 4.1 and 4.2 we have

$$|\mathfrak{M}(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}')| \leq \exp\left(-\frac{9}{2}\right) \ll 1, \quad \text{if } |\boldsymbol{z} - \boldsymbol{z}'| \geq 3\widetilde{R} \text{ or } |z_3 - z_3'| \geq 3\widetilde{R}_3,$$

and we may work only with the diagonal part of the CINT kernel.

We obtain the linear system of equations (2.14), with vector $\boldsymbol{u}$ of components $|\rho(\vec{\boldsymbol{z}})|^2$ at the N_z mesh points in $\mathfrak{D}$. The "data" vector $\boldsymbol{d}$ consists of the samples of the CINT image at $N_y < N_z$ equidistant points in $\mathfrak{D}$, and in the narrowband regime the $N_y \times N_z$ matrix $\boldsymbol{\mathcal{M}}$ has entries

$$m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}} = C \exp\left[-\frac{|\boldsymbol{z} - \boldsymbol{y}|^2}{2R^2} - \frac{(\overline{z}_3 - y_3)^2}{2R_3^2}\right], \qquad B \ll \Omega_d, \tag{4.15}$$

with constant C. This depends only on $\vec{\boldsymbol{y}} - \vec{\boldsymbol{z}}$, so we have a convolution as stated in section 2.3. In the broadband regime, the entries of $\boldsymbol{\mathcal{M}}$ are

$$m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}} = \frac{C}{\sqrt{1 + \frac{|\boldsymbol{z}-\boldsymbol{y}|^2}{2\theta^2 R^2}}} \exp\left[-\frac{|\boldsymbol{z} - \boldsymbol{y}|^2}{2R^2} - \frac{|z_3 - y_3 + \frac{|\boldsymbol{z}-\boldsymbol{y}|^2}{2L} + \frac{\boldsymbol{y}\cdot(\boldsymbol{z}-\boldsymbol{y})}{L}|^2}{2R_3^2\left[1 + \frac{|\boldsymbol{z}-\boldsymbol{y}|^2}{2\theta^2 R^2}\right]}\right], \qquad B \gg \Omega_d, \tag{4.16}$$

with constant C. Were it not for the last term in (4.16), we would have a convolution. This term is large only at points $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$ with $\boldsymbol{y}$ near the boundary of the imaging region ($|\boldsymbol{y}| \sim D < a$), because by definition (4.12) and the assumption $\theta = O(1)$ we get

$$\left|\frac{\boldsymbol{y}\cdot(\boldsymbol{z} - \boldsymbol{y})/L}{R_3}\right| = O\left(\frac{|\boldsymbol{y}|}{L}\frac{R}{R_3}\right) = O\left(\frac{|\boldsymbol{y}|}{a\theta}\right) = O\left(\frac{|\boldsymbol{y}|}{a}\right).$$

For points with $|\boldsymbol{y}| \ll D < a$ the right hand side in (4.16) is approximately a function of $\vec{\boldsymbol{y}} - \vec{\boldsymbol{z}}$, corresponding to a convolution model.

5. Resolution analysis. In this section we analyze the reconstruction of the vector $\boldsymbol{u}$ of source intensities using the convex optimization formulation described in section 2.3. To simplify the analysis, we treat the approximation in (2.14) as an equality, and study the l_1 optimization

$$\min_{\boldsymbol{u}\in\mathbb{R}^{N_z}} \|\boldsymbol{u}\|_1 \quad \text{such that} \quad \boldsymbol{\mathcal{M}}\,\boldsymbol{u} = \boldsymbol{d}. \tag{5.1}$$

This neglects additive noise and random fluctuations of the CINT function, which are small in our scaling. It also implies that the sources are on the reconstruction mesh, so that the equality constraint in (5.1) holds for the true discretized source intensity. Naturally, in practice the sources may lie anywhere in $\mathfrak{D}$, and noise and distortions due to the random medium play a role. This is why we use the more robust formulation (2.15) in the numerical simulations in section 6.

We expect from the study [11] of deconvolution using l_1 optimization that the solution of (5.1) should be a good approximation of the unknown vector of intensities if the sources are well separated. We show in this section that this is indeed the case. We also consider the case of clusters of nearby sources, and show that the l_1 solution is useful when the clusters are well separated. The analysis is built on our recent results in [7].

5.1. Definitions. Let $\boldsymbol{Y} = \{\vec{\boldsymbol{y}}_s, \quad s = 1, \ldots, N_s\}$ be the set that supports the unknown, point-like sources in $\mathfrak{D}$. We quantify the spatial separation between them using the following definition:

DEFINITION 5.1. *The points in $\boldsymbol{Y}$ are separated by at least $\vec{\boldsymbol{H}} = (H, H, H_3)$, if the intersection of $\boldsymbol{Y}$ with any rectangular prism of sides less then H in cross-range and H_3 in range consists of at most one point.*

For example, if the sources are all in the same cross-range plane, and the minimum distance between any two of them is H_{min}, we may take $H = H_{min}$ and $H_3 = D_3$.

We search the sources on a mesh with N_z points denoted generically by $\vec{\boldsymbol{z}}$. The mesh discretizes $\mathfrak{D}$, and we call it $\mathfrak{D}_z$. For simplicity we let $\boldsymbol{Y} \subset \mathfrak{D}_z$. To any $\vec{\boldsymbol{z}} \in \mathfrak{D}_z$, we associate the column vector $\boldsymbol{m}_{\vec{\boldsymbol{z}}} \in \mathbb{R}^{N_y}$ of the matrix $\boldsymbol{\mathcal{M}}$. Its entries are given in (4.15) in the narrowband regime and by (4.16) in the broadband regime, for $N_y < N_z$ points $\vec{\boldsymbol{y}}$ at which we sample the CINT image.

DEFINITION 5.2. *We quantify the interaction between two presumed sources at $\vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}' \in \mathfrak{D}_{\vec{\boldsymbol{z}}}$ using the cross-correlation of the associated column vectors in $\boldsymbol{\mathcal{M}}$,*

$$\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} = \frac{|\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'} \rangle|}{\|\boldsymbol{m}_{\vec{\boldsymbol{z}}}\|_2 \|\boldsymbol{m}_{\vec{\boldsymbol{z}}'}\|_2}. \tag{5.2}$$

Here $\langle \cdot, \cdot \rangle$ is the Euclidian inner product and $\| \cdot \|_2$ is the Euclidian norm.

Note that (5.2) is symmetric and non-negative, and attains its maximum at $\vec{\boldsymbol{z}}' = \vec{\boldsymbol{z}}$, where $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}} = 1$. We will show below that (5.2) decreases as the points $\vec{\boldsymbol{z}}$ and $\vec{\boldsymbol{z}}'$ grow apart. This motivates the next definition which uses $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}$ to measure the distance between $\vec{\boldsymbol{z}}$ and $\vec{\boldsymbol{z}}'$.

DEFINITION 5.3. *We define the semimetric $\Delta : \mathfrak{D}_z \times \mathfrak{D}_z \to [0, 1]$ by*

$$\Delta(\vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') = 1 - \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, \qquad \forall \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}' \in \mathfrak{D}_z, \tag{5.3}$$

and let

$$\mathscr{B}_r(\vec{\boldsymbol{z}}) = \{\vec{\boldsymbol{z}}' \in \mathfrak{D}_z \quad s.t. \quad \Delta(\vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') < r\} \tag{5.4}$$

be the open balls defined by Δ.

We will show that $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}$ decays as $\|\vec{\boldsymbol{z}} - \vec{\boldsymbol{z}}'\|_2$ grows. Thus, we say that points $\vec{\boldsymbol{z}}'$ outside $\mathscr{B}_r(\vec{\boldsymbol{z}})$ have a weaker interaction with $\vec{\boldsymbol{z}}$ than points inside $\mathscr{B}_r(\vec{\boldsymbol{z}})$. Moreover, we may relate intuitively $\Delta(\vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}')$ to the Euclidian distance $\|\vec{\boldsymbol{z}} - \vec{\boldsymbol{z}}'\|_2$.

DEFINITION 5.4. *We define the interaction coefficient of the set $\boldsymbol{Y}$ of source locations by*

$$\mathcal{I}(\boldsymbol{Y}) = \max_{\vec{\boldsymbol{z}} \in \mathfrak{D}_z} \sum_{\vec{\boldsymbol{y}} \in \boldsymbol{Y} \setminus \mathscr{N}(\vec{\boldsymbol{z}})} \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}}, \tag{5.5}$$

where $\mathscr{N}(\vec{\boldsymbol{z}})$ is the closest point to $\vec{\boldsymbol{z}}$ in $\boldsymbol{Y}$, as measured with the semimetric Δ.

In general more than one point may be closest to $\vec{\boldsymbol{z}}$. If this is so, we let $\mathscr{N}(\vec{\boldsymbol{z}})$ be any such point.

5.2. Results. The results stated here describe the relation between the reconstruction $\boldsymbol{u}_\star$, the solution of the convex optimization problem (5.1), and the true unknown vector $\boldsymbol{u}$ of source intensities. The next theorem shows that $\boldsymbol{u}_\star$ is essentially supported in the set $\boldsymbol{Y}$, when the points there are well separated.

THEOREM 5.5. *Suppose that the source locations in $\boldsymbol{Y}$ are separated by at least $\vec{\boldsymbol{H}} = (H, H, H_3)$ in the sense of Definition 5.1, where*

$$\frac{H}{R} = \alpha, \quad \frac{H_3}{R_3} = \alpha_3, \tag{5.6}$$

for $\alpha, \alpha_3 > 1$. Take $r \in (0, 1)$ small enough so that the balls $\mathscr{B}_r(\vec{\boldsymbol{y}}_s)$ centered at $\vec{\boldsymbol{y}}_s \in \boldsymbol{Y}$, for $s = 1, \ldots, N_s$, are disjoint. Let $\boldsymbol{u}_\star$ be the l_1 minimizer in (5.1), and decompose it as $\boldsymbol{u}_\star = \boldsymbol{u}_\star^{(i)} + \boldsymbol{u}_\star^{(o)}$, where $\boldsymbol{u}_\star^{(i)}$ is supported

in the union $\bigcup_{s=1}^{N_s} \mathscr{B}_r(\vec{\boldsymbol{y}}_s)$ *of balls centered at the points in* $\boldsymbol{Y}$*, and* $\boldsymbol{u}_\star^{(o)}$ *is supported in the complement of this union. Then, there exists a constant* C *that is independent of* α *and* α_3*, such that*

$$\|\boldsymbol{u}_\star^{(o)}\|_1 \leq \frac{C}{r}\mathfrak{F}(\alpha,\alpha_3)\|\boldsymbol{u}_\star\|_1, \tag{5.7}$$

where $\mathfrak{F}(\alpha,\alpha_3)$ *is function that decays with* α *and* α_3*. In the narrowband case it is given by*

$$\mathfrak{F}(\alpha,\alpha_3) = \frac{\exp\left[-\left(\frac{\min\{\alpha,\alpha_3\}}{4}\right)^2\right]}{\alpha^2\alpha_3}, \tag{5.8}$$

for arbitrary $\alpha,\alpha_3 > 1$*. In the broadband case* $\alpha_3 > 8\alpha/\theta$*, and*

$$\mathfrak{F}(\alpha,\alpha_3) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\alpha}{4}\right)^2\right] + \exp[-\alpha]}{\alpha^3}. \tag{5.9}$$

Note that the scales of separation between the sources are the resolution parameters R and R_3 of CINT. The parameters α and α_3 in the separation assumption may be any non-negative real numbers, but the statement of the theorem is useful only when the coefficient in front of $\|\boldsymbol{u}_\star\|_1$ in (5.7) is smaller than one. This happens for large enough α and α_3. The larger the separation between the sources, the smaller the right hand side in (5.7) is, and the better the concentration of the support of $\boldsymbol{u}_\star$ near the points in $\boldsymbol{Y}$. The next corollary gives an estimate of the error of the reconstruction.

COROLLARY 5.6. *Let* $\boldsymbol{u} \in \mathbb{R}^{N_z}$ *be the vector of true source intensities, and use the same assumptions and notation as in Theorem 5.5. Denote the entries of the* l_1 *minimizer* $\boldsymbol{u}_\star$ *by* $u_\star(\vec{\boldsymbol{z}})$*, where* $\vec{\boldsymbol{z}}$ *are the* N_z *points on the mesh* $\mathfrak{D}_z$*. Define the effective reconstructed source intensity vector* $\overline{\boldsymbol{u}}_\star \in \mathbb{R}^{N_z}$*, with entries*

$$\overline{u}_\star(\vec{\boldsymbol{z}}) = \begin{cases} \sum_{\vec{\boldsymbol{z}}' \in \mathscr{B}_r(\vec{\boldsymbol{z}})} u_\star(\vec{\boldsymbol{z}}')\,\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, & \text{if } \vec{\boldsymbol{z}} \in \boldsymbol{Y}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.10}$$

Then, we have the following estimate of the relative error

$$\frac{\|\boldsymbol{u} - \overline{\boldsymbol{u}}_\star\|_1}{\|\boldsymbol{u}\|_1} \leq \frac{C}{r}\mathfrak{F}(\alpha,\alpha_3), \tag{5.11}$$

with constant C *independent of* α *and* α_3*.*

This result says that when the sources are far apart, the effective intensity vector $\overline{\boldsymbol{u}}_\star$ is close to the true solution $\boldsymbol{u}$. By definition, the support of $\overline{\boldsymbol{u}}_\star$ is at the source points in $\boldsymbol{Y}$. Its entries $\overline{u}_\star(\vec{\boldsymbol{z}})$ at $\vec{\boldsymbol{z}} \in \boldsymbol{Y}$ are weighted averages of the entries of $\boldsymbol{u}_\star$ at points $\vec{\boldsymbol{z}}' \in \mathscr{B}_r(\vec{\boldsymbol{z}})$, with weights $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}$. When r is small, these weights are close to one, so $\overline{u}_\star(\vec{\boldsymbol{z}})$ is approximately the sum of the entries of $\boldsymbol{u}_\star$ supported in the ball $\mathscr{B}_r(\vec{\boldsymbol{z}})$.

Theorem 5.5 and its corollary are not useful when the sources are clustered together. The next result deals with this case, when the clusters are well separated.

THEOREM 5.7. *Let* $\epsilon \in (0,1)$ *be such that there exists a subset* $\boldsymbol{Y}_\epsilon$ *of* $\boldsymbol{Y}$*, satisfying*

$$\boldsymbol{Y} \subset \bigcup_{\vec{\boldsymbol{z}} \in \boldsymbol{Y}_\epsilon} \mathscr{B}_\epsilon(\vec{\boldsymbol{z}}), \qquad \mathscr{B}_\epsilon(\vec{\boldsymbol{z}}) \bigcap \mathscr{B}_\epsilon(\vec{\boldsymbol{z}}') = \emptyset, \quad \forall\, \vec{\boldsymbol{z}},\vec{\boldsymbol{z}}' \in \boldsymbol{Y}_\epsilon, \quad \vec{\boldsymbol{z}} \neq \vec{\boldsymbol{z}}'. \tag{5.12}$$

Suppose that the points in $\boldsymbol{Y}_\epsilon$ *are separated by at least least* $\vec{\boldsymbol{H}} = (H,H,H_3)$ *in the sense of Definition 5.1, where* H *and* H_3 *satisfy* (5.6)*, for some* $\alpha,\alpha_3 > 1$*. Let* r *satisfy* $\epsilon < r < 1$*, and decompose the* l_1 *minimizer* $\boldsymbol{u}_\star$ *in* (5.1) *as* $\boldsymbol{u}_\star = \boldsymbol{u}_\star^{(i)} + \boldsymbol{u}_\star^{(o)}$*, where* $\boldsymbol{u}_\star^{(i)}$ *is supported in the union* $\bigcup_{\vec{\boldsymbol{z}} \in \boldsymbol{Y}_\epsilon} \mathscr{B}_r(\vec{\boldsymbol{z}})$ *of balls centered at the points in* $\boldsymbol{Y}_\epsilon$*, and* $\boldsymbol{u}_\star^{(o)}$ *is supported in the complement of this union. There exists a constant* C *that is independent of* α *and* α_3*, such that*

$$\|\boldsymbol{u}_\star^{(o)}\|_1 \leq \frac{C}{r}\mathfrak{F}(\alpha,\alpha_3)\|\boldsymbol{u}_\star\|_1 + \frac{\epsilon}{r}\|\boldsymbol{u}\|_1, \tag{5.13}$$

where $\boldsymbol{u}$ is the vector of true source intensities.

Equation (5.12) says that we can cover the sources with disjoint balls of radius ϵ, centered at the points in $\boldsymbol{Y}_\epsilon$. Thus, we call $\boldsymbol{Y}_\epsilon$ the effective support of the sources, and ϵ the radius of the clusters. The statement of the theorem is that when the clusters are well separated, and they have small radius, the l_1 minimizer will be supported in their vicinity. As expected, (5.13) converges to (5.7) in the limit $\epsilon \to 0$.

5.3. Proofs. We use [7, Theorem 4.1 and Corollary 4.2] which state that for the decomposition of the l_1 minimizer $\boldsymbol{u}_\star$ as in Theorem 5.5, and for the effective reconstructed source intensity vector $\overline{\boldsymbol{u}}_\star$ defined in (5.10), we have

$$\|\boldsymbol{u}_\star^{(o)}\|_1 \leq \frac{2\mathcal{I}(\boldsymbol{Y})}{r}\|\boldsymbol{u}_\star\|_1 \quad \text{and} \quad \frac{\|\boldsymbol{u}-\overline{\boldsymbol{u}}_\star\|_1}{\|\boldsymbol{u}\|_1} \leq \frac{2\mathcal{I}(\boldsymbol{Y})}{r}. \tag{5.14}$$

To determine the interaction coefficient $\mathcal{I}(\boldsymbol{Y})$ of the sources, we estimate first the cross-correlations $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}$:

LEMMA 5.8. *The cross-correlations $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}$ defined in* (5.2) *satisfy*

$$\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \approx \exp\left[-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2} - \frac{(z_3-z_3')^2}{4R_3^2}\right], \tag{5.15}$$

in the narrowband regime and

$$\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \leq C\exp\left[-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2} - \frac{\theta\left|z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}\right|}{R_3}\right], \tag{5.16}$$

in the broadband regime, for all $\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}' \in \mathfrak{D}_z$, with $\vec{\boldsymbol{z}} = (\boldsymbol{z}, z_3)$ and $\vec{\boldsymbol{z}}' = (\boldsymbol{z}', z_3')$. The constant C in (5.16) *is given by*

$$C = \frac{3e^{-\theta^2}}{2erfc(\sqrt{2}\,\theta)}, \tag{5.17}$$

with θ defined in (4.12) *of order one.*

The proof is in Appendix E. The next lemma, proved in sections 5.3.1 and 5.3.2, gives the estimate of $\mathcal{I}(\boldsymbol{Y})$, that combined with (5.14) proves Theorem 5.5 and Corollary 5.6.

LEMMA 5.9. *Suppose that the points in $\boldsymbol{Y}$ are separated by at least $\vec{\boldsymbol{H}} = (H, H, H_3)$ in the sense of Definition 5.1, where H and H_3 are as in* (5.6)*, for some $\alpha, \alpha_3 > 1$. The interaction coefficient satisfies*

$$\mathcal{I}(\boldsymbol{Y}) \leq C\mathfrak{F}(\alpha, \alpha_3) \tag{5.18}$$

for a constant C that is independent of α and α_3 and $\mathfrak{F}(\alpha, \alpha_3)$ as defined in Theorem 5.5.

To prove Theorem 5.7, we use [7, Theorem 4.4] which states that

$$\|\boldsymbol{u}_\star^{(o)}\|_1 \leq \frac{2\mathcal{I}(\boldsymbol{Y}_\epsilon)}{r}\|\boldsymbol{u}_\star\|_1 + \frac{\|\boldsymbol{u}\| - \|\overline{\boldsymbol{u}}\|_1}{r}, \tag{5.19}$$

for $\overline{\boldsymbol{u}}$ defined by

$$\overline{\boldsymbol{u}}(\vec{\boldsymbol{z}}) = \begin{cases} \sum_{\vec{\boldsymbol{z}}' \in \mathscr{B}_\epsilon(\vec{\boldsymbol{z}}) \cap \boldsymbol{Y}} u(\vec{\boldsymbol{z}}')\,\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, & \text{if } \vec{\boldsymbol{z}} \in \boldsymbol{Y}_\epsilon, \\ 0, & \text{otherwise.} \end{cases} \tag{5.20}$$

The interaction coefficient of the effective support $\boldsymbol{Y}_\epsilon$ is as in Lemma 5.9, so it remains to estimate the last term in (5.19). Let us define the set

$$S_{\vec{\boldsymbol{z}},\epsilon} = \{\vec{\boldsymbol{z}}' \in \boldsymbol{Y} \ \text{ s.t. } \ \vec{\boldsymbol{z}}' \in \mathscr{B}_\epsilon(\vec{\boldsymbol{z}})\},$$

so that with definition (5.20) we can write

$$\|\overline{\boldsymbol{u}}\|_1 = \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}_\epsilon} \left| \sum_{\vec{\boldsymbol{z}}'\in S_{\vec{\boldsymbol{z}},\epsilon}} u(\vec{\boldsymbol{z}}')\, \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \right| = \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}_\epsilon} \sum_{\vec{\boldsymbol{z}}'\in S_{\vec{\boldsymbol{z}},\epsilon}} |u(\vec{\boldsymbol{z}}')|\, \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'},$$

where we used that by definition $\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \geq 0$ and $u(\vec{\boldsymbol{z}}') = |\rho(\vec{\boldsymbol{z}}')|^2 \geq 0$. Since the norm of the vector of the true source intensities is given by

$$\|\boldsymbol{u}\|_1 = \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}} |u(\vec{\boldsymbol{z}})| = \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}_\epsilon} \sum_{\vec{\boldsymbol{z}}'\in S_{\vec{\boldsymbol{z}}\,\epsilon}} |u(\vec{\boldsymbol{z}}')|,$$

we obtain that

$$\|\boldsymbol{u}\|_1 - \|\overline{\boldsymbol{u}}\|_1 = \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}_\epsilon} \sum_{\vec{\boldsymbol{z}}'\in S_{\vec{\boldsymbol{z}},\epsilon}} |u(\vec{\boldsymbol{z}}')|\,(1 - \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}) < \epsilon \sum_{\vec{\boldsymbol{z}}\in \boldsymbol{Y}_\epsilon} \sum_{\vec{\boldsymbol{z}}'\in S_{\vec{\boldsymbol{z}},\epsilon}} |u(\vec{\boldsymbol{z}}')| = \epsilon \|\boldsymbol{u}\|_1. \tag{5.21}$$

The inequality is because

$$\Delta(\vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') = 1 - \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} < \epsilon, \quad \forall\, \vec{\boldsymbol{z}}' \in \mathscr{B}_\epsilon(\vec{\boldsymbol{z}}).$$

The statement of Theorem 5.7 follows by substitution of (5.21) in (5.19), and using the estimate in Lemma 5.9, with $\boldsymbol{Y}$ replaced by $\boldsymbol{Y}_\epsilon$. $\square$.

5.3.1. Proof of Lemma 5.9 in the narrowband regime. Recall Definition 5.4 of $\mathcal{I}(\boldsymbol{Y})$, and let $\vec{\boldsymbol{z}} \in \mathfrak{D}_z$ be the maximizer of the sum in (5.5), so that

$$\mathcal{I}(\boldsymbol{Y}) = \sum_{\vec{\boldsymbol{y}}\in \boldsymbol{Y}\setminus \mathscr{N}(\vec{\boldsymbol{z}})} \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}}. \tag{5.22}$$

We denote the components of $\vec{\boldsymbol{z}}$ by z_j, for $j = 1, 2, 3$, and conclude from the source separation assumption in the lemma that the set

$$\mathscr{S}_{\vec{\boldsymbol{z}}} = \left\{ \vec{\boldsymbol{z}}' = (z_1', z_2', z_3') \in \mathfrak{D} \ \text{ s.t. } \ |z_j - z_j'| < H,\ j = 1, 2, \ \ |z_3 - z_3'| < H_3 \right\} \tag{5.23}$$

contains at most one point in $\boldsymbol{Y}$. This may be $\mathscr{N}(\vec{\boldsymbol{z}})$, the closest point in $\boldsymbol{Y}$ to $\vec{\boldsymbol{z}}$ with respect to the semimetric Δ, satisfying

$$\mathcal{I}_{\vec{\boldsymbol{z}},\mathscr{N}(\vec{\boldsymbol{z}})} \geq \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}}. \qquad \forall\, \vec{\boldsymbol{y}} \in \boldsymbol{Y}. \tag{5.24}$$

Alternatively, $\mathscr{S}_{\vec{\boldsymbol{z}}}$ may be empty or contain another point in $\boldsymbol{Y}$. In either case, we obtain from equations (5.22) and (5.24) that

$$\mathcal{I}(\boldsymbol{Y}) \leq \sum_{\vec{\boldsymbol{y}}\in \boldsymbol{Y}\setminus \mathscr{S}_{\vec{\boldsymbol{z}}}} \mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}},$$

and from the bound in Lemma 5.8,

$$\mathcal{I}(\boldsymbol{Y}) \leq C \sum_{\vec{\boldsymbol{y}}\in \boldsymbol{Y}\setminus \mathscr{S}_{\vec{\boldsymbol{z}}}} \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}}, \ \text{ for } \ \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}} = \exp\left[-\frac{|\boldsymbol{z} - \boldsymbol{y}|^2}{8R^2} - \frac{(z_3 - y_3)^2}{R_3} \right]. \tag{5.25}$$

Using again the source separation assumption in the lemma, we conclude that for any $\vec{\boldsymbol{y}} \in \boldsymbol{Y}$, we can define a set $\mathscr{H}_{\vec{\boldsymbol{y}}}$, in the form of a rectangular prism of sides $H/2$ in cross-range and $H_3/2$ in range, satisfying

$$\vec{\boldsymbol{y}} \in \mathscr{H}_{\vec{\boldsymbol{y}}} \ \text{ and } \ \mathscr{H}_{\vec{\boldsymbol{y}}} \bigcap \mathscr{H}_{\vec{\boldsymbol{y}}'} = \emptyset, \qquad \forall\, \vec{\boldsymbol{y}} \neq \vec{\boldsymbol{y}}' \in \boldsymbol{Y}. \tag{5.26}$$

There are many such sets, but we make our choice so that $\vec{\boldsymbol{y}}$ is the furthermost point to $\vec{\boldsymbol{z}}$ in $\mathscr{H}_{\vec{\boldsymbol{y}}}$, satisfying

$$\mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}} \leq \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, \qquad \forall\, \vec{\boldsymbol{z}}' \in \mathscr{H}_{\vec{\boldsymbol{y}}}. \tag{5.27}$$

This allows us to write

$$\mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}} \leq \frac{8}{H^2 H_3} \int_{\mathscr{H}_{\vec{\boldsymbol{y}}}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, \qquad \forall\, \vec{\boldsymbol{y}} \in \boldsymbol{Y}, \tag{5.28}$$

and obtain from (5.25) that

$$\mathcal{I}(\boldsymbol{Y}) \leq \frac{8C^2}{H^2 H_3} \sum_{\boldsymbol{y}\in \boldsymbol{Y}\setminus \mathscr{S}_{\vec{\boldsymbol{z}}}} \int_{\mathscr{H}_{\vec{\boldsymbol{y}}}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \leq \frac{8C^2}{H^2 H_3} \int_{\mathbb{R}^3\setminus \mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'}, \tag{5.29}$$

with $\mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}$ defined as in (5.23), with half the values of H and H_3,

$$\mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}} = \left\{ \vec{\boldsymbol{z}}' = (z_1', z_2', z_3') \in \mathfrak{D} \ \text{ s.t. } \ |z_j - z_j'| < H/2,\ j = 1,2, \ \ |z_3 - z_3'| < H_3/2 \right\}. \tag{5.30}$$

The last inequality in (5.29) is because the integrand is positive, the sets $\mathscr{H}_{\vec{\boldsymbol{y}}}$ are disjoint, and

$$\bigcup_{\vec{\boldsymbol{y}}\in \boldsymbol{Y}\setminus \mathscr{S}_{\vec{\boldsymbol{z}}}} \mathscr{H}_{\vec{\boldsymbol{y}}} \subset \mathbb{R}^3 \setminus \mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}.$$

We estimate the integral in (5.29) by decomposing the set $\mathscr{S}^c_{\vec{\boldsymbol{z}},\frac{1}{2}} = \mathbb{R}^3 \setminus \mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}$ in three components denoted by $\mathscr{C}_{\vec{\boldsymbol{z}},j}$, where

$$\mathscr{C}_{\vec{\boldsymbol{z}},j} = \left\{ \vec{\boldsymbol{z}}' \in \mathbb{R}^3 \ \text{ s.t. } \ |z_j - z_j'| \geq H/2 \right\}, \quad j = 1,2,$$

and

$$\mathscr{C}_{\vec{\boldsymbol{z}},3} = \left\{ \vec{\boldsymbol{z}}' \in \mathbb{R}^3 \ \text{ s.t. } \ |z_j - z_j'| < H/2,\ |z_3 - z_3'| \geq H_3/2 \right\}.$$

We have

$$\begin{aligned}
\frac{8}{H^2 H_3} \int_{\mathscr{C}_{\vec{\boldsymbol{z}},1}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} &= \frac{8}{H^2 H_3} \int_{|z_1'-z_1|\geq H/2} dz_1'\, e^{-\frac{(z_1'-z_1)^2}{4R^2}} \int_{-\infty}^{\infty} dz_2'\, e^{-\frac{(z_2'-z_2)^2}{4R^2}} \int_{-\infty}^{\infty} dz_3'\, e^{-\frac{(z_3'-z_3)^2}{4R_3^2}} \\
&= \frac{32\pi R R_3}{H^2 H_3} 2 \int_{H/2}^{\infty} dt\, e^{-\frac{t^2}{4R^2}} \\
&= \frac{64\sqrt{\pi}}{\alpha^2 \alpha_3} \operatorname{erfc}\left(\frac{\alpha}{4}\right),
\end{aligned} \tag{5.31}$$

where we evaluated the integrals over z_3' and z_2' in the second line, and used (5.6) in the last line. The integral over $\mathscr{C}_{\vec{\boldsymbol{z}},2}$ is the same, so it remains to estimate

$$\begin{aligned}
\frac{8}{H^2 H_3} \int_{\mathscr{C}_{\vec{\boldsymbol{z}},3}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} = \frac{8}{H^2 H_3} \int_{|z_1'-z_1|<H/2} dz_1'\, e^{-\frac{(z_1'-z_1)^2}{4R^2}} \int_{|z_2'-z_2|<H/2} dz_2'\, e^{-\frac{(z_2'-z_2)^2}{4R^2}} \\
\times \int_{|z_3'-z_3|\geq H_3/2} dz_3'\, e^{-\frac{(z_3'-z_3)^2}{4R_3^2}}.
\end{aligned} \tag{5.32}$$

We bound the integrals over z_1' and z_2' by those of the real line, and rewrite the integral over z_3' in terms of the complementary error function, to obtain

$$\frac{8}{H^2 H_3} \int_{\mathscr{C}_{\vec{\boldsymbol{z}},3}} d\vec{\boldsymbol{z}}'\, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \leq \frac{64\sqrt{\pi}}{\alpha^2 \alpha_3} \operatorname{erfc}\left(\frac{\alpha_3}{4}\right). \tag{5.33}$$

The statement of Lemma 5.9 follows from (5.29), with right hand side given by the sum of the integrals over $\mathscr{C}_{\vec{\boldsymbol{z}},1}$ and $\mathscr{C}_{\vec{\boldsymbol{z}},2}$ estimated in (5.31), and over $\mathscr{C}_{\vec{\boldsymbol{z}},3}$, estimated in (5.33). □

5.3.2. Proof of Lemma 5.9 in the broadband regime. We obtain from Lemma 5.8, the same way as above, and with the same notation, that

$$\mathcal{I}(\boldsymbol{Y}) \leq C \sum_{\vec{\boldsymbol{y}} \in \boldsymbol{Y} \setminus \mathscr{S}_{\vec{\boldsymbol{z}}}} \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}}, \quad \text{for} \quad \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{y}}} = \exp\left[-\frac{|\boldsymbol{z}-\boldsymbol{y}|^2}{8R^2} - \frac{\theta \left| z_3 - y_3 + \frac{|\boldsymbol{z}|^2 - |\boldsymbol{y}|^2}{2L} \right|}{R_3} \right]. \tag{5.34}$$

We also define as before, using the source separation assumption in the lemma, the set $\mathscr{H}_{\vec{\boldsymbol{y}}}$, satisfying (5.26) and (5.27). This leads us to the bound (5.29), with the set $\mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}$ defined in (5.30).

We estimate the integral in (5.29) by decomposing the set $\mathbb{R}^3 \setminus \mathscr{S}_{\vec{\boldsymbol{z}},\frac{1}{2}}$ in three parts $S_{\vec{\boldsymbol{z}},j}$, where

$$S_{\vec{\boldsymbol{z}},j} = \left\{ \vec{\boldsymbol{z}}' \in \mathbb{R}^3 \ \text{ s.t. } \ |z_j - z_j'| \geq H/2 \right\}, \quad j = 1, 2,$$

and

$$S_{\vec{\boldsymbol{z}},3} = \left\{ \vec{\boldsymbol{z}}' \in \mathbb{R}^3 \ \text{ s.t. } \ |z_j - z_j'| < H/2, \ |z_3 - z_3'| \geq H_3/2 \right\}.$$

We have

$$\begin{aligned}
\frac{8}{H^2 H_3} \int_{S_{\vec{\boldsymbol{z}},1}} d\vec{\boldsymbol{z}}' \, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} &= \frac{8}{H^2 H_3} \int_{|z_1' - z_1| \geq H/2} dz_1' \, e^{-\frac{(z_1'-z_1)^2}{8R^2}} \int_{-\infty}^{\infty} dz_2' \, e^{-\frac{(z_2'-z_2)^2}{8R^2}} \int_{-\infty}^{\infty} dz_3' \, e^{-\frac{\theta}{R_3}\left| z_3' - z_3 + \frac{|\boldsymbol{z}|^2 - |\boldsymbol{z}'|^2}{2L} \right|} \\
&= \frac{128\pi R^2 R_3}{H^2 H_3 \theta} \operatorname{erfc}\left(\frac{H}{4\sqrt{2}R} \right) \\
&\leq \frac{128\pi}{\alpha^2 \alpha_3 \theta} e^{-\left(\frac{h}{4\sqrt{2}R}\right)^2} \\
&\leq \frac{16\pi}{\alpha^3} e^{-\left(\frac{\alpha}{4\sqrt{2}}\right)^2},
\end{aligned} \tag{5.35}$$

where we evaluated the integrals over z_3' and z_2' in the second line. The first inequality is because the complementary error function satisfies $\operatorname{erfc}(x) \leq \exp(-x^2)$, and the second inequality is by the assumption on α_3. The integral over $S_{\vec{\boldsymbol{z}},2}$ is the same, so it remains to estimate

$$\begin{aligned}
\frac{8}{H^2 H_3} \int_{S_{\vec{\boldsymbol{z}},3}} d\vec{\boldsymbol{z}}' \, \mathscr{E}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \approx \frac{8}{H^2 H_3} &\int_{|z_1' - z_1| < H/2} dz_1' \, e^{-\frac{(z_1'-z_1)^2}{8R^2}} \int_{|z_2' - z_2| < H/2} dz_2' \, e^{-\frac{(z_2'-z_2)^2}{8R^2}} \\
&\times \int_{|z_3' - z_3| \geq H_3/2} dz_3' \, e^{-\frac{\theta}{R_3}\left| z_3' - z_3 + \frac{|\boldsymbol{z}'|^2 - |\boldsymbol{z}|^2}{2L} \right|}.
\end{aligned} \tag{5.36}$$

Because $|z_j' - z_j| < H/2$ for $j = 1, 2$, and therefore $|\boldsymbol{z}' - \boldsymbol{z}| < H/\sqrt{2}$, we have

$$\theta \left| \frac{|\boldsymbol{z}'|^2 - |\boldsymbol{z}|^2}{2LR_3} \right| \leq \frac{\theta |\boldsymbol{z}' - \boldsymbol{z}| |\boldsymbol{z}' + \boldsymbol{z}|}{2LR_3} \leq \frac{\theta H D}{2LR_3} = \frac{\theta \alpha D R}{2LR_3} = \frac{3D\alpha}{a} \lesssim 3\alpha. \tag{5.37}$$

Here we used that $|\boldsymbol{z}' + \boldsymbol{z}| \leq D\sqrt{2}$ for all $\boldsymbol{z}, \boldsymbol{z}' \in [-D/2, D/2] \times [-D/2, D/2]$, and substituted $H = R\alpha$ and definitions (4.5) and (4.12). The last inequality is by assumption (3.28). With this bound we can estimate the integral over z_3' as follows

$$\begin{aligned}
\int_{|z_3' - z_3| \geq H_3/2} dz_3' \, e^{-\frac{\theta}{R_3}\left| z_3' - z_3 + \frac{|\boldsymbol{z}'|^2 - |\boldsymbol{z}|^2}{2L} \right|} &= \frac{R_3}{\theta} \left[\int_{\frac{\theta H_3}{2R_3}}^{\infty} dt \, e^{-\left| -t + \frac{\theta(|\boldsymbol{z}'|^2 - |\boldsymbol{z}|^2)}{2LR_3} \right|} + \int_{\frac{\theta H_3}{2R_3}}^{\infty} dt \, e^{-\left| t + \frac{\theta(|\boldsymbol{z}'|^2 - |\boldsymbol{z}|^2)}{2LR_3} \right|} \right] \\
&\leq \frac{2R_3}{\theta} \int_{\frac{\theta H_3}{2R_3}}^{\infty} dt \, e^{-(t - 3\alpha)} \\
&= \frac{2R_3}{\theta} e^{-\left(\frac{\theta \alpha_3}{2} - 3\alpha\right)} \int_0^{\infty} dt e^{-t} \\
&\leq \frac{2R_3}{\theta} e^{-\alpha},
\end{aligned} \tag{5.38}$$

where the last inequality is by the assumption $H_3/R_3 = \alpha_3 \geq 8\alpha/\theta$. Substituting in (5.36) we get

$$\frac{8}{H^2 H_3}\int_{S_{\vec{z},3}} d\vec{z}'\, \mathscr{E}_{\vec{z},\vec{z}'} \leq \frac{16R_3}{H^2 H_3\theta} e^{-\alpha}\int_{|z_1'-z_1|<H/2} dz_1'\, e^{-\frac{(z_1'-z_1)^2}{8R^2}} \int_{|z_2'-z_2|<H/2} dz_2'\, e^{-\frac{(z_2'-z_2)^2}{8R^2}}$$
$$\leq \frac{128\pi R^2 R_3}{H^2 H_3\theta} e^{-\alpha} = \frac{128\pi}{\alpha^2\alpha_3\theta} \leq \frac{16\pi}{\alpha^3} e^{-\alpha}. \tag{5.39}$$

Here we bounded each Gaussian integral by $2\sqrt{2\pi}R$, which is the integral over the real line, and used again the assumption on α_3.

The statement of Lemma 5.9 follows from (5.29), with right hand side given by the sum of the integrals over $S_{\vec{z},1}$ and $S_{\vec{z},2}$ estimated in (5.35), and over $S_{\vec{z},3}$, estimated in (5.39). □

6. Numerical simulations. We present here numerical simulations obtained with the wave propagation model described in equation (3.5), with random travel time fluctuations computed by the line integrals in (3.6), in one realization of the random process μ. We generate it numerically using random Fourier series [17], for the Gaussian autocorrelation (3.2). All the length scales are normalized by ℓ in the simulations, and are chosen to satisfy marginally the assumptions in section 3.4. Specifically, we take $\lambda_o = 1.75\cdot 10^{-4}\,\ell$ and $L = 800\,\ell$, so that

$$\sqrt{\lambda_o L} = 0.12 < \ell < 9.72\,\ell = (\lambda_o L^3)^{1/4},$$

and the aperture is $a = 16\ell$. This is slightly larger than the bound in (3.24), but we also have the apodization (4.3). We verify that

$$\frac{\lambda_o}{\sqrt{\ell L}} = 6.17\cdot 10^{-7} \ll \frac{\sqrt{\lambda_o \ell}}{L} = 5.22\cdot 10^{-6} \ll \left(\frac{\ell}{L}\right)^{3/2} = 4.42\cdot 10^{-5},$$

and we take the strength of the fluctuations $\sigma = 1.5\cdot 10^{-6}$. With this choice we obtain

$$\frac{\Omega_d}{\omega_o} = 0.083 \quad \text{and} \quad X_d = 0.14\ell.$$

We show results in two dimensions, for a linear array and a narrowband regime with bandwidth $B = 0.0032\omega_o$. Since in this regime we can only expect improvements in the cross-range localization of the sources, we focus attention at a given range, and display cross-range sections of the images.

The migration and CINT images are calculated as in equations (2.6) and (2.7) from the data contaminated with 5% additive, uncorrelated, Gaussian noise. The thresholding parameters in the CINT image formation are $\Omega = B/2$ and $X = X_d/2$, and the sources are off the reconstruction mesh. The mesh size is $H = \lambda_o L/(6X)$, unless stated otherwise The optimization formulation is

$$\min_{\boldsymbol{u}} \|\boldsymbol{u}\|_1 \quad \text{such that} \quad \|\boldsymbol{\mathcal{M}}\,\boldsymbol{u} - \boldsymbol{d}\|_2 \leq \delta, \tag{6.1}$$

with tolerance $\delta = 0.05\|\boldsymbol{d}\|_2$. The entries of matrix $\boldsymbol{\mathcal{M}}$ are as defined in (4.15), with constant C given in appendix D. For comparison, we also present the results of a direct application of L_1 optimization to the array data, without using the CINT image formation, as in [7]. We call this method "direct l_1 optimization" and refer to [7, Appendix A] for details. We also refer to (6.1) as "l_1 optimization" and solve it with the software package [16].

We begin with the results in Figure 6.1, for two sources that are at about $2\lambda_o L/X$ apart. The exact source locations are indicated on the abscissa in the plots, where the units are in $\lambda_o L/X$. We display in the top row the CINT and migration images, and in the bottom row the solutions of the l_1 optimization and the direct l_1 optimization. Both migration and direct l_1 optimization give spurious peaks, due to the random medium. To confirm this, we show in Figure 6.2 the results of the direct l_1 optimization for the same sources in the homogeneous medium, where the reconstruction is very good. The CINT image shown in the top left plot is blurry, and it cannot distinguish the two sources. The l_1 optimization improves the

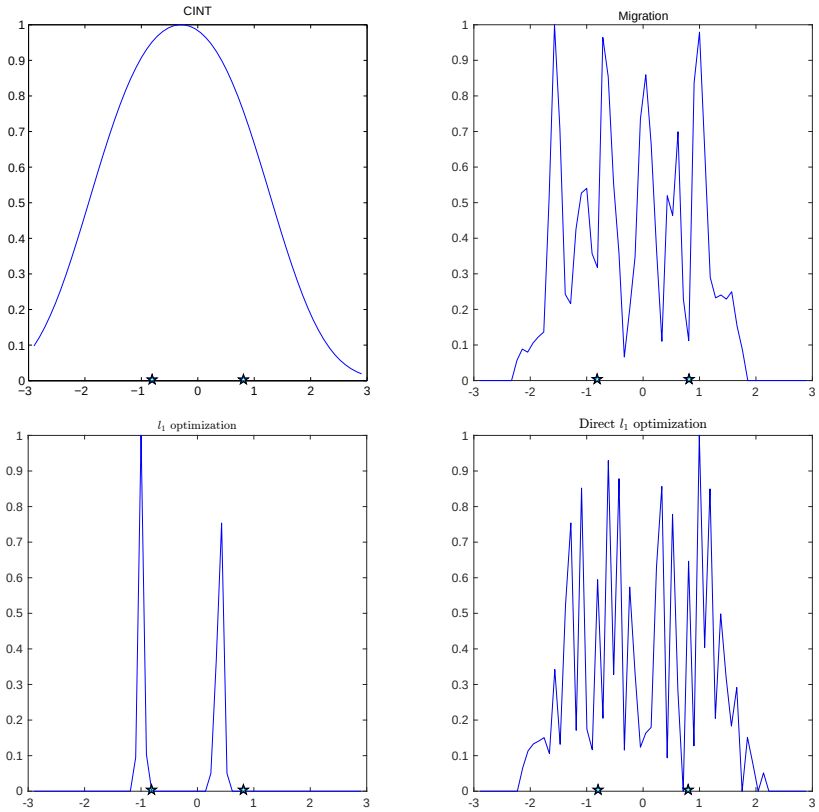

FIG. 6.1. *Results in one realization of the random medium. The abscissa is the cross-range in units $\lambda_o L/X$, and the source locations are shown with the stars. We display the CINT and migration images in the top row, and the l_1 and direct l_1 optimization results in the bottom row.*

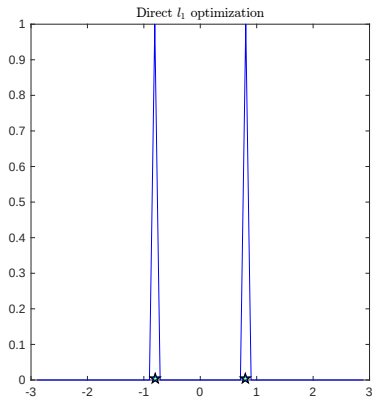

FIG. 6.2. *Result of direct l_1 optimization in the homogeneous medium. The abscissa is the cross-range in units $\lambda_o L/X$, and the source locations are shown with the stars.*

result, although there is a small shift in the estimate of the source locations. This shift changes from one realization to another, and it is due to the small random fluctuations of the CINT image.

To illustrate the robustness of the methods to different realizations of the random medium, we display in Figure 6.3 the histograms of the number of peaks found by each method at a given cross-range location. We define filtered peaks as local maxima whose values are above 33% of the maximum of the image. The height of the histograms varies among the plots in Figure 6.3 because each method finds a different number

of peaks. On the average, the migration images have 9.5 peaks, the direct l_1 method finds 9.65 peaks, the CINT image has 1.01 peaks and the l_1 optimization finds 2.04 peaks. While both migration and direct l_1 find many spurious peaks, that are far from the source locations, the l_1 optimization separates well the two sources and almost always peaks at their true locations.

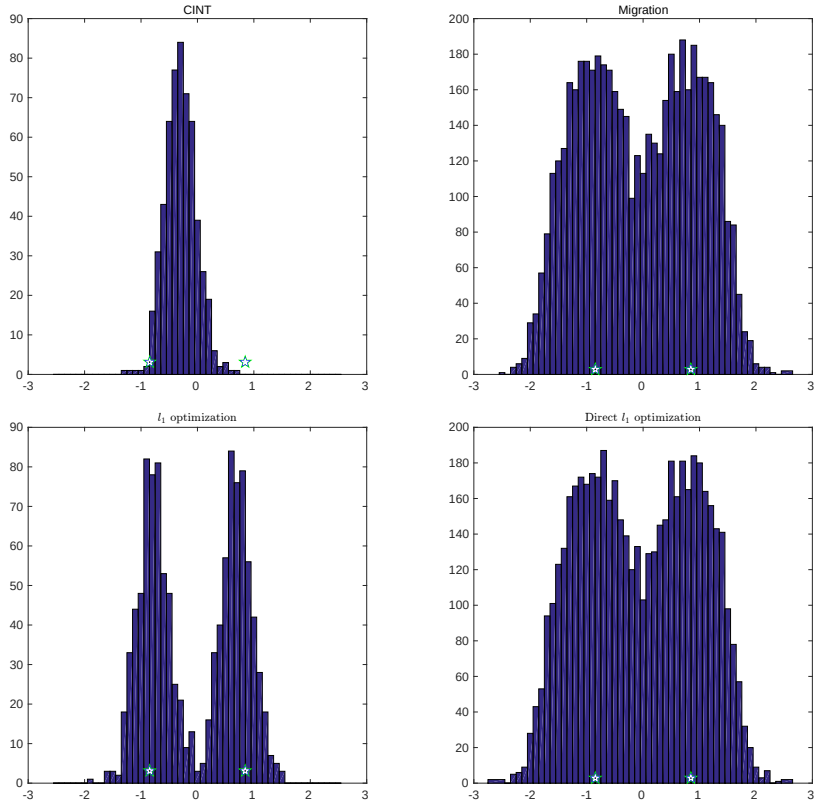

FIG. 6.3. *Histogram of the number of peaks obtained in* 500 *realizations of the random medium. The abscissa is the cross-range in units* $\lambda_o L/X$*, and the source locations are shown with the stars. We display the results for the CINT and migration images in the top row, and the* l_1 *and direct* l_1 *optimization results in the bottom row. The heights of the histograms are different because each method finds a different number of peaks. On the average, the number of peaks found in simulations are Migration: 9.5, CINT: 1.01,* l_1 *optimization: 2.04, Direct* l_1 *optimization: 9.65.*

We display in Figure 6.4 the effect of the mesh size H on the l_1 optimization results. Because we have 5 sources in this simulation, that are closer apart than $\lambda_o L/X$, we do not expect a nearly exact reconstruction with the l_1 optimization. Thus, we display in addition to the actual reconstructions the aggregated values recovered in the intervals of length $r = \lambda_o L/(4X)$, centered at the true source locations. We observe that the results improve as we decrease the mesh size from $H = \lambda_o L/X$ to $\lambda_o L/(8X)$. This is due to the fact that the sources are off the grid, and the discretization error decreases as we reduce H.

The last illustration, in Figure 6.5, shows the effect of the source separation on the quality of the reconstructions. As expected from the results in section 5, the reconstruction is better when the sources are further apart.

7. Summary. We studied receiver array imaging of remote localized sources in random media, using convex optimization. The scattering regime is defined by precise scaling assumptions, and leads to large random wavefront distortions of the waves measured at the array. Conventional imaging methods like reverse time migration, also known as backprojection [4, 20], or standard l_1 optimization [14, 19], cannot deal with such distortions and produce poor and unreliable results. We base our imaging on the coherent interferometric (CINT) method [9] which mitigates random media effects like wavefront distortions at the expense of image resolution. The goal of the convex optimization is to remove the blur in the CINT images

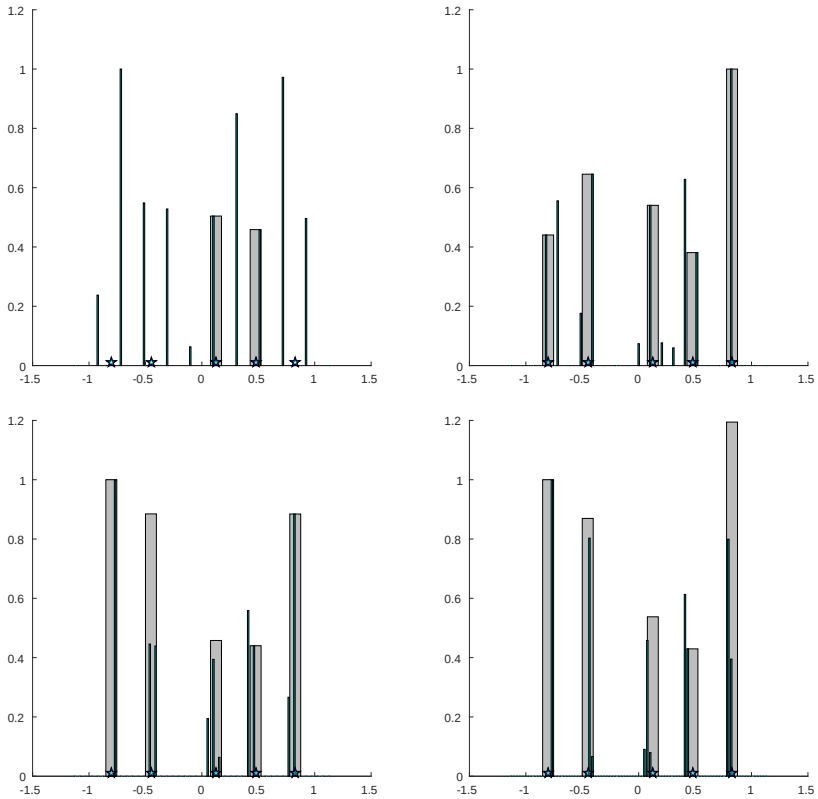

FIG. 6.4. *We display l_1 optimization results in one realization of the random medium, for 5 sources with locations indicated by the stars in the abscissa. The units of the abscissa are in $\lambda_o L/X$. The dark thin bars show the reconstruction and the light gray bars give the aggregated values of the reconstruction in the intervals of length $r = L\lambda_o/(4X)$, centered at the source locations. The mesh size is, clockwise, starting from the top left, $H = 1, 1/2, 1/8$ and $1/4$ of $\lambda_o L/X$. The results improve as we refine the mesh.*

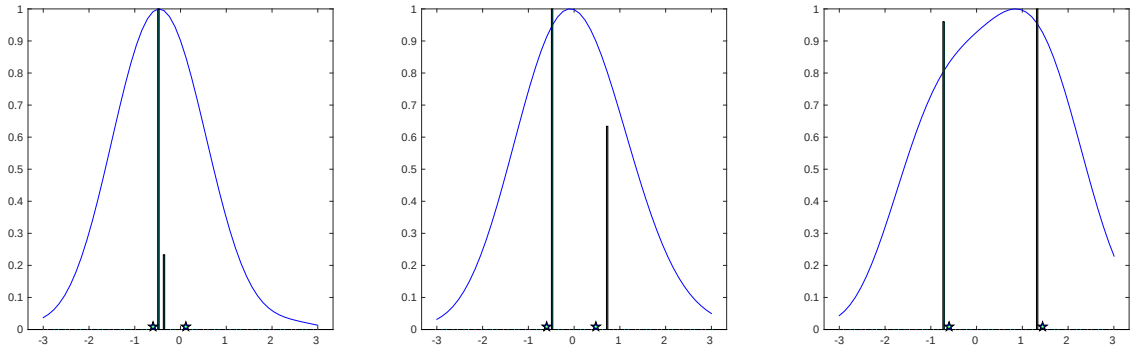

FIG. 6.5. *We display l_1 optimization results in one realization of the random medium, for 2 sources separated by, from left to right, 3/4, 1 and 2 $\lambda_o L/X$. The units of the abscissa are in $\lambda_o L/X$.*

and thus improve the source localization. We show with a detailed analysis that under generic conditions the CINT imaging function is approximately a convolution of a blurring kernel with the discretized unknown source intensity on the imaging mesh. The kernel has a generic expression, that depends on the known CINT resolution limits obtained in [9, 10, 6] for various wave propagation models. The optimization seeks to undo this convolution. The analysis and numerical simulations show that it gives very good estimates of the source locations when they are sufficiently far apart. This is in agreement with the results in [11].

We also show that when the sources are clustered together, the estimates are not close to the true locations pointwise, but they are supported in their vicinity.

Acknowledgments. We gratefully acknowledge support from ONR Grant N00014-14-1-0077 and the AFOSR Grant FA9550-15-1-0118.

Appendix A. The paraxial approximation. Equation (3.15) follows easily from the Taylor expansion

$$|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| = \sqrt{y_3^2+|\boldsymbol{x}-\boldsymbol{y}|^2} = L\left[1+O\left(\frac{D_3}{L}\right)+O\left(\frac{a^2}{L^2}\right)\right],$$

where we used that $y_3 = L+O(D_3)$, $|\boldsymbol{x}| = O(a)$ and $|\boldsymbol{y}| = O(D) \le O(a)$. For the phase we obtain similarly

$$k_o|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| = k_o\sqrt{y_3^2+|\boldsymbol{y}-\boldsymbol{x}|^2} = k_o\left(y_3+\frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{2y_3}\right)+O\left(\frac{k_o|\boldsymbol{x}-\boldsymbol{y}|^4}{y_3^3}\right),$$

where $k_o = \omega_o/c_o = 2\pi/\lambda_o$. Expanding the square in the right hand side and using that $y_3 = L+O(D_3)$,

$$k_o|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| = k_o\left(y_3+\frac{|\boldsymbol{x}|^2-2\boldsymbol{x}\cdot\boldsymbol{y}+|\boldsymbol{y}|^2}{2L}\right)+O\left(\frac{a^4}{\lambda_o L^3}\right)+O\left(\frac{a^2D_3}{\lambda_o L^2}\right)+O\left(\frac{aDD_3}{\lambda_o L^2}\right)+O\left(\frac{D^2D_3}{\lambda_o L^2}\right).$$

The last two terms in the remainder are dominated by the previous ones, because $D < a$, and approximation (3.14) follows.

Appendix B. Derivation of the statistical moments. We begin with the second moments of the Gaussian process (3.7), with Gaussian autocorrelation (3.2), and then prove Propositions 3.1 and 3.2.

Consider points $\vec{\boldsymbol{y}}$ and $\vec{\boldsymbol{y}}'$ in the imaging region, so that $|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| \ge |\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|$, and write

$$\begin{aligned}\mathbb{E}\left[\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\nu(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}')\right] &= \frac{\sqrt{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'||\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}}{\sqrt{2\pi}\ell}\iint_0^1 d\vartheta d\vartheta'\, e^{-\frac{1}{2\ell^2}\left|(\vartheta'-\vartheta)(\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}})+\vartheta'(\vec{\boldsymbol{x}}'-\vec{\boldsymbol{x}})+(1-\vartheta')(\vec{\boldsymbol{y}}'-\vec{\boldsymbol{y}})\right|^2}\\ &= \sqrt{\frac{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}}\int_0^1 d\vartheta'\int_{-(1-\vartheta')|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|/\ell}^{\vartheta'|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|/\ell}\frac{d\widetilde{\vartheta}}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left|\widetilde{\vartheta}\frac{(\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}})}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}+\vartheta'\frac{(\vec{\boldsymbol{x}}'-\vec{\boldsymbol{x}})}{\ell}+(1-\vartheta')\frac{(\vec{\boldsymbol{y}}'-\vec{\boldsymbol{y}})}{\ell}\right|^2},\end{aligned}$$

where we changed variables

$$(\vartheta,\vartheta') \rightsquigarrow (\widetilde{\vartheta},\vartheta'),\quad \widetilde{\vartheta} = (\vartheta'-\vartheta)|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|/\ell.$$

Let $\vec{\boldsymbol{m}} = (\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}})/|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|$ and note that since $|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| \sim L \gg \ell$ and the Gaussian is negligible for $|\widetilde{\vartheta}| > 3$, we can extend the $\widetilde{\vartheta}$ integral to the real line and obtain

$$\mathbb{E}\left[\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\nu(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}')\right] \approx \sqrt{\frac{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}}\int_0^1 d\vartheta\, e^{-\frac{1}{2\ell^2}\left|(I-\vec{\boldsymbol{m}}\vec{\boldsymbol{m}}^T)\left[\vartheta(\vec{\boldsymbol{x}}'-\vec{\boldsymbol{x}})+(1-\vartheta)(\vec{\boldsymbol{y}}'-\vec{\boldsymbol{y}})\right]\right|^2}. \tag{B.1}$$

In the system of coordinates centered at the array, we calculate

$$\vec{\boldsymbol{m}} = \frac{(\boldsymbol{x},0)-(\boldsymbol{y},L+y_3)}{\sqrt{(L+y_3)^2+|\boldsymbol{x}-\boldsymbol{y}|^2}} = \left[-\vec{\boldsymbol{e}}_3+\frac{(\boldsymbol{x}-\boldsymbol{y},0)}{L+y_3}\right]\left[1+O\left(\frac{a^2}{L^2}\right)\right],$$

with $\vec{\boldsymbol{e}}_3$ the unit vector in the range direction, and obtain

$$|\vec{\boldsymbol{m}}+\vec{\boldsymbol{e}}_3| = O\left(\frac{a}{L}\right) \ll 1.$$

The projections on the plane orthogonal to $\vec{\boldsymbol{m}}$ are

$$(I-\vec{\boldsymbol{m}}\vec{\boldsymbol{m}}^T)\frac{(\vec{\boldsymbol{x}}'-\vec{\boldsymbol{x}})}{\ell} = \frac{(\boldsymbol{x}'-\boldsymbol{x},0)}{\ell}+O\left(\frac{a^2}{\ell L}\right), \tag{B.2}$$

$$(I-\vec{\boldsymbol{m}}\vec{\boldsymbol{m}}^T)\frac{(\vec{\boldsymbol{y}}'-\vec{\boldsymbol{y}})}{\ell} = \frac{(\boldsymbol{y}'-\boldsymbol{y},0)}{\ell}+O\left(\frac{aD}{\ell L}\right)+O\left(\frac{aD_3}{\ell L}\right), \tag{B.3}$$

with negligible residuals by assumptions (3.24)-(3.25) and (3.28), which give

$$\frac{a^2}{\ell L} \ll \frac{\sqrt{\lambda_o L}}{\ell} \ll 1, \qquad \frac{aD_3}{\ell L} \ll \frac{\lambda_o L}{\ell a} < \frac{\lambda_o L}{\ell^2} \ll 1.$$

Thus, we can approximate (B.1) as

$$\mathbb{E}\left[\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\nu(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}')\right] \approx \sqrt{\frac{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}} \int_0^1 d\vartheta\, e^{-\frac{1}{2\ell^2}\left|\vartheta(\boldsymbol{x}'-\boldsymbol{x})+(1-\vartheta)(\boldsymbol{y}'-\boldsymbol{y})\right|^2}. \tag{B.4}$$

This expression can be simplified further for small offsets satisfying $|\boldsymbol{y}'-\boldsymbol{y}| \ll \ell$ and $|\boldsymbol{x}'-\boldsymbol{x}| \ll \ell$, by expanding the exponential and then integrating in ϑ,

$$\mathbb{E}\left[\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\nu(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}')\right] \approx \sqrt{\frac{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}} \left[1 - \frac{|\boldsymbol{y}'-\boldsymbol{y}|^2 + (\boldsymbol{y}'-\boldsymbol{y})\cdot(\boldsymbol{x}'-\boldsymbol{x}) + |\boldsymbol{x}'-\boldsymbol{x}|^2}{6\ell^2}\right]. \tag{B.5}$$

Note that the small receiver offset condition is consistent with $|\boldsymbol{x}'-\boldsymbol{x}| \ll X \sim X_d \ll \ell$ used in the CINT image formation.

The proof of Proposition 3.1 follows easily from (B.5) and definitions (3.5), (3.7). Because the process $\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})$ is Gaussian, we have

$$\begin{aligned}
\mathbb{E}\left[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\right] &\approx \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\mathbb{E}\left[\exp\left(i\frac{(2\pi)^{1/4}}{2}k\sigma\sqrt{\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}\nu(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\right)\right] \\
&\approx \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\exp\left\{-\frac{\sqrt{2\pi}}{8}k^2\sigma^2\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|\mathbb{E}\left[\nu^2(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\right]\right\} \\
&= \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega)\exp\left[-\frac{\sqrt{2\pi}}{8}k^2\sigma^2\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|\right].
\end{aligned} \tag{B.6}$$

This is equation (3.17) with scattering mean free path $\mathcal{S}(\omega)$ defined as in (3.18). □

To prove Proposition 3.2 we use again definitions (3.5) and (3.7), the Gaussianity of the process ν, and result (B.1) to write

$$\mathbb{E}\left[\widehat{G}(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega+\widetilde{\omega}/2)\overline{\widehat{G}(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega-\widetilde{\omega}/2)}\right] \approx \widehat{G}_o(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}},\omega+\widetilde{\omega}/2)\overline{\widehat{G}_o(\vec{\boldsymbol{x}}',\vec{\boldsymbol{y}}',\omega-\widetilde{\omega}/2)}\,\mathcal{E}, \tag{B.7}$$

with

$$\begin{aligned}
\mathcal{E} = \exp\Bigg\{&-\frac{\sqrt{2\pi}}{8}\left(\frac{\omega+\widetilde{\omega}/2}{c_o}\right)^2\sigma^2\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| - \frac{\sqrt{2\pi}}{8}\left(\frac{\omega-\widetilde{\omega}/2}{c_o}\right)^2\sigma^2\ell|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'| \\
&+\frac{\sqrt{2\pi}(\omega^2-\widetilde{\omega}^2/4)}{4c_o^2}\sigma^2\ell|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|\int_0^1 d\vartheta\, e^{-\frac{1}{2\ell^2}\left|\vartheta(\boldsymbol{x}'-\boldsymbol{x})+(1-\vartheta)(\boldsymbol{y}'-\boldsymbol{y})\right|^2}\Bigg\}.
\end{aligned} \tag{B.8}$$

Note that

$$\begin{aligned}
\frac{\sqrt{2\pi}}{8}\left(\frac{\omega+\widetilde{\omega}/2}{c_o}\right)^2\sigma^2\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| &= \frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}{\mathcal{S}(\omega+\widetilde{\omega}/2)} \gg 1, \\
\frac{\sqrt{2\pi}}{8}\left(\frac{\omega-\widetilde{\omega}/2}{c_o}\right)^2\sigma^2\ell|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'| &= \frac{|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{\mathcal{S}(\omega-\widetilde{\omega}/2)} \gg 1,
\end{aligned}$$

so $\mathcal{E}$ is exponentially small unless the last term in (B.8) compensates the first two. This happens only when $|\boldsymbol{y}-\boldsymbol{y}'| \ll \ell$. For other offsets the integral over ϑ is small and, consequently, $\mathcal{E} \approx 0$.

When $|\boldsymbol{y}-\boldsymbol{y}'| \ll \ell$, we can approximate the ϑ integral as in (B.5), and obtain

$$\mathcal{E} \approx \exp\left\{-\frac{\sqrt{2\pi}}{8}\left(\frac{\omega+\widetilde{\omega}/2}{c_o}\right)^2 \sigma^2\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| - \frac{\sqrt{2\pi}}{8}\left(\frac{\omega-\widetilde{\omega}/2}{c_o}\right)^2 \sigma^2\ell|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'| \right.$$
$$\left. + \frac{\sqrt{2\pi}(\omega^2-\widetilde{\omega}^2/4)}{4c_o^2}\sigma^2\ell|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|\left[1-\frac{|\boldsymbol{y}'-\boldsymbol{y}|^2+(\boldsymbol{y}'-\boldsymbol{y})\cdot(\boldsymbol{x}'-\boldsymbol{x})+|\boldsymbol{x}'-\boldsymbol{x}|^2}{6\ell^2}\right]\right\}. \tag{B.9}$$

Rearranging the terms and using definition (3.18) of the scattering mean free path we get

$$\mathcal{E} \approx \exp\left\{-\frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|-|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{\mathcal{S}(\omega)} - \frac{\sqrt{2\pi}\sigma^2\ell\widetilde{\omega}^2}{8c_o^2}\left(\frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|+3|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{4}\right) - \frac{\sqrt{2\pi}\sigma^2\ell\omega\widetilde{\omega}}{8c_o^2}(|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|-|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|)\right.$$
$$\left. - \frac{\sqrt{2\pi}(\omega^2-\widetilde{\omega}^2/4)\sigma^2|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{24c_o^2\ell}\left(|\boldsymbol{y}'-\boldsymbol{y}|^2+(\boldsymbol{y}'-\boldsymbol{y})\cdot(\boldsymbol{x}'-\boldsymbol{x})+|\boldsymbol{x}'-\boldsymbol{x}|^2\right)\right\}. \tag{B.10}$$

Recall that $|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}| \geq |\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|$, and conclude from the decay of the first term that $\mathcal{E}$ is large if

$$|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|-|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}| = O(\mathcal{S}(\omega)).$$

Substituting in (B.10), and using the scales

$$\Omega_d = \frac{2c_o}{(2\pi)^{1/4}\sigma\sqrt{\ell|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|}}, \quad X(\omega) = \frac{\sqrt{3}\ell\Omega_d}{\omega}, \tag{B.11}$$

we get

$$\mathcal{E} \approx \exp\left\{-\frac{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|-|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'|}{\mathcal{S}(\omega)} - \frac{\widetilde{\omega}^2}{2\Omega_d^2}\left[1+O\left(\frac{\mathcal{S}(\omega)}{L}\right)\right] + O\left(\frac{\sigma^2\ell\omega\widetilde{\omega}\mathcal{S}(\omega)}{c_o^2}\right)\right.$$
$$\left. - \frac{(|\boldsymbol{y}'-\boldsymbol{y}|^2+(\boldsymbol{y}'-\boldsymbol{y})\cdot(\boldsymbol{x}'-\boldsymbol{x})+|\boldsymbol{x}'-\boldsymbol{x}|^2)}{2X(\omega)^2}\left[1+O\left(\frac{\mathcal{S}(\omega)}{L}\right)+O\left(\frac{\Omega_d^2}{\omega^2}\right)\right]\right\}. \tag{B.12}$$

This equation simplifies because $\mathcal{S}(\omega) \ll L$ and $\Omega_d \ll \omega = O(\omega_o)$. Moreover, (B.12) is large only when $|\widetilde{\omega}| \lesssim \Omega_d$, so we estimate using definitions (3.18) and (B.11) and the assumption (3.26) that

$$\frac{\sigma^2\ell\omega\widetilde{\omega}\mathcal{S}}{c_o^2} \lesssim O\left(\frac{\sigma^2\ell\omega_o\Omega_d\mathcal{S}}{c_o^2}\right) = O\left(\frac{\lambda_o}{\sigma\sqrt{\ell L}}\right) \ll 1.$$

Similarly, $\mathcal{E}$ is large for cross-range offsets of at most $O(X)$, so we can write

$$|\vec{\boldsymbol{x}}-\vec{\boldsymbol{y}}|-|\vec{\boldsymbol{x}}'-\vec{\boldsymbol{y}}'| \approx y_3-y_3'+\frac{|\boldsymbol{x}|^2-|\boldsymbol{x}'|^2}{2L}+\frac{|\boldsymbol{y}|^2-|\boldsymbol{y}'|^2}{2L}-\frac{\boldsymbol{x}\cdot\boldsymbol{y}-\boldsymbol{x}'\cdot\boldsymbol{y}'}{L} = y_3-y_3'+O\left(\frac{aX}{L}\right).$$

We also have $X \sim X_d$, and

$$\frac{aX_d}{\mathcal{S}L} = O\left(\frac{aX_d}{L}\frac{\sigma^2\ell}{\lambda_o^2}\right) = O\left(\frac{a\sigma\ell^{3/2}}{\lambda_o L^{3/2}}\right) \ll O\left[\frac{\sigma\ell^{3/2}}{(\lambda_o L)^{3/4}}\right] \ll O\left(\frac{\ell^2}{\lambda_o^{1/4}L^{1+3/4}}\right) \ll O\left[\left(\frac{\lambda_o}{L}\right)^{1/4}\right] \ll 1,$$

where the two equalities are by definition (3.18) and (3.20) of $\mathcal{S}$ and X_d, the first inequality is by (3.24), the second by (3.26), the third by (3.25), and the last by (3.3). Gathering the results we arrive at the approximation

$$\mathcal{E} \approx \exp\left[-\frac{|y_3-y_3'|}{\mathcal{S}(\omega)} - \frac{\widetilde{\omega}^2}{2\Omega_d^2} - \frac{|\boldsymbol{y}'-\boldsymbol{y}|^2+(\boldsymbol{y}'-\boldsymbol{y})\cdot(\boldsymbol{x}'-\boldsymbol{x})+|\boldsymbol{x}'-\boldsymbol{x}|^2}{2X^2(\omega)}\right]. \tag{B.13}$$

The statement of Proposition 3.2 follows from this equation and $|\omega - \omega_o| \lesssim \pi B \ll \omega_o$. $\square$

Appendix C. Consistency of scaling. The scaling (3.26) is consistent because

$$\frac{(\ell/L)^{3/2}}{\sqrt{\ell\lambda_o}/L} = \frac{\ell}{\sqrt{\lambda_o L}} \gg 1 \quad \text{and} \quad \frac{\sqrt{\ell\lambda_o}/L}{\lambda_o/\sqrt{\ell L}} = \frac{\ell}{\sqrt{\lambda_o L}} \gg 1,$$

where we used (3.25).

The scaling (3.29) is consistent because by assumption (3.24) and (3.20),

$$\frac{(a/L)^2}{\lambda_o L/(aX_d)} = \frac{a^4}{\lambda_o L^3}\frac{X_d}{a} \ll \frac{\ell}{a} < 1.$$

We also have

$$\frac{\lambda_o L/(aX_d)}{\Omega_d/\omega_o} = O\left(\frac{\lambda_0 L}{a\ell(\Omega_d/\omega_o)^2}\right) = O\left(\frac{\sigma^2 L^2}{a\lambda_o}\right) \ll \frac{\ell}{a} < 1,$$

so (3.29) implies $B \ll \Omega_d$.

Definition (3.20) and assumptions (3.25), (3.26) give

$$\frac{\lambda_o L}{X_d} = O\left(\frac{\lambda_o L}{\ell}\frac{\omega_o}{\Omega_d}\right) = O\left(\sigma\frac{L^{3/2}}{\ell^{1/2}}\right) \ll \sqrt{\lambda_o L} \ll \ell.$$

This verifies that the assumptions (3.28) on D are consistent.

The assumptions (3.28) on D_3 are consistent in the narrowband regime, when (3.29) holds, because by (3.27) we have $\Omega/B = O(1)$ and

$$\frac{c_o/B}{\lambda_o L^2/a^2} = O\left(\frac{\omega_o}{B}\frac{a^2}{L^2}\right) \ll 1.$$

In the broadband regime $\Omega = \Omega_d$ and assumptions (3.28) on D_3 are consistent because

$$\frac{c_o/\Omega_d}{\lambda_o L^2/a^2} = O\left(\frac{\sigma\sqrt{\ell L}}{\lambda_o L^2/a^2}\right) \ll \frac{\ell}{L} \ll 1,$$

by definition (3.20) and assumptions (3.24) and (3.26). Moreover, assumption (3.32) is consistent because

$$\frac{\lambda_o^{2/3}\ell^{1/6}/L^{5/6}}{\lambda_o/\sqrt{\ell L}} = \left(\frac{\ell}{\sqrt{\lambda_o L}}\right)^{2/3} \gg 1, \qquad \frac{\lambda_o^{2/3}\ell^{1/6}/L^{5/6}}{\sqrt{\lambda_o \ell}/L} = \left(\frac{\sqrt{\lambda_o L}}{\ell}\right)^{1/3} \ll 1,$$

by assumption (3.25).

Appendix D. Derivation of the CINT kernel. We begin with the expression (2.11) of the CINT kernel, and use the Gaussian pulse, thresholding windows and apodization to obtain

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') \approx & \frac{N_r^2\sqrt{2\pi}}{a^4 B}\int d\boldsymbol{x}\, e^{-\frac{|\boldsymbol{x}|^2}{(a/6)^2}}\int d\widetilde{\boldsymbol{x}}\, e^{-\frac{|\widetilde{\boldsymbol{x}}|^2}{2X^2}-\frac{|\widetilde{\boldsymbol{x}}|^2}{4(a/6)^2}}\int d\omega\, e^{-\frac{(\omega-\omega_o)^2}{2B^2}}\int d\widetilde{\omega}\, e^{-\frac{\widetilde{\omega}^2}{2\Omega^2}-\frac{\widetilde{\omega}^2}{8B^2}} \\ & \times \widehat{G}\left((\boldsymbol{x}+\widetilde{\boldsymbol{x}}/2, 0), \vec{\boldsymbol{z}}, \omega+\widetilde{\omega}/2\right)\overline{\widehat{G}\left((\boldsymbol{x}-\widetilde{\boldsymbol{x}}/2, 0), \vec{\boldsymbol{z}}', \omega-\widetilde{\omega}/2\right)} \\ & \times \exp\left[-i\frac{(\omega+\widetilde{\omega}/2)}{c_o}\left|(\boldsymbol{x}+\widetilde{\boldsymbol{x}}/2, 0)-\vec{\boldsymbol{y}}\right| + i\frac{(\omega-\widetilde{\omega}/2)}{c_o}\left|(\boldsymbol{x}-\widetilde{\boldsymbol{x}}/2, 0)-\vec{\boldsymbol{y}}\right|\right], \end{aligned} \tag{D.1}$$

for points $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$, $\vec{\boldsymbol{z}} = (\boldsymbol{z}, z_3)$ and $\vec{\boldsymbol{z}}' = (\boldsymbol{z}', z_3')$ in the imaging region. The integration over $\boldsymbol{x}$ and $\widetilde{\boldsymbol{x}}$ extends to the whole plane $\mathbb{R}^2$, and those over ω and $\widetilde{\omega}$ to the real line, with the aperture and bandwidth

restrictions ensured by the Gaussians. Because CINT is statistically stable in our scaling, we may approximate the right hand side of (D.2) by its expectation. Using (3.23) and the paraxial approximation (3.16), we obtain

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx & \frac{N_r^2\sqrt{2\pi}}{a^4B(4\pi L)^2}e^{-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{2X_d^2}}\int d\boldsymbol{x}\, e^{-\frac{|\boldsymbol{x}|^2}{(a/6)^2}}\int d\widetilde{\boldsymbol{x}}\, e^{-\frac{|\widetilde{\boldsymbol{x}}|^2}{2X_e^2}-\frac{\widetilde{\boldsymbol{x}}\cdot(\boldsymbol{z}-\boldsymbol{z}')}{2X_d^2}}\int d\omega\, e^{-\frac{(\omega-\omega_o)^2}{2B^2}}\int d\widetilde{\omega}\, e^{-\frac{\widetilde{\omega}^2}{2\Omega_e^2}}\\ & \times\exp\left\{i\frac{\widetilde{\omega}}{c_o}\left[\frac{z_3+z_3'}{2}-y_3+\frac{1}{2L}\left(\frac{|\boldsymbol{z}|^2+|\boldsymbol{z}'|^2}{2}-|\boldsymbol{y}|^2\right)-\frac{\boldsymbol{x}}{L}\cdot\left(\frac{\boldsymbol{z}+\boldsymbol{z}'}{2}-\boldsymbol{y}\right)-\frac{\widetilde{\boldsymbol{x}}}{4L}\cdot(\boldsymbol{z}-\boldsymbol{z}')\right]\right\}\\ & \times\exp\left\{i\frac{\omega}{c_o}\left[z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}-\frac{\boldsymbol{x}\cdot(\boldsymbol{z}-\boldsymbol{z}')}{L}-\frac{\widetilde{\boldsymbol{x}}}{L}\cdot\left(\frac{\boldsymbol{z}+\boldsymbol{z}'}{2}-\boldsymbol{y}\right)\right]\right\},\end{aligned}\tag{D.2}$$

with $X_e = O(X_d)$ and $\Omega_e = O(B)$ defined in (4.6). Note that the last term in the second line of (D.2) is negligible, because by assumptions (3.24), (3.29) and definition (3.20),

$$\frac{\widetilde{\omega}\widetilde{\boldsymbol{x}}\cdot(\boldsymbol{z}-\boldsymbol{z}')}{4Lc_o} = O\left(\frac{BX_d^2}{c_oL}\right) = O\left(\frac{B}{\omega_o}\frac{aX_d}{\lambda_oL}\frac{X_d}{a}\right) \ll 1,$$

in the narrowband case. Moreover, in the broadband case

$$\frac{\widetilde{\omega}\widetilde{\boldsymbol{x}}\cdot(\boldsymbol{z}-\boldsymbol{z}')}{4Lc_o} = O\left(\frac{\Omega_dX_d^2}{c_oL}\right) = O\left(\frac{\ell^2}{\lambda_oL}\frac{\Omega_d^3}{\omega_o^3}\right) = O\left(\frac{\lambda_o^2\ell^{1/2}}{\sigma^3L^{5/2}}\right) \ll 1,$$

where the first equalities are by definitions (3.20), and the bound is by (3.32). Let us introduce the center and difference coordinates

$$\overline{\boldsymbol{z}} = \frac{\boldsymbol{z}+\boldsymbol{z}'}{2}, \qquad \widetilde{\boldsymbol{z}} = \boldsymbol{z}-\boldsymbol{z}', \qquad \overline{z}_3 = \frac{z_3+z_3'}{2}, \qquad \widetilde{z}_3 = z_3-z_3',\tag{D.3}$$

and note that

$$\frac{\widetilde{\omega}}{c_oL}\left(\frac{|\boldsymbol{z}|^2+|\boldsymbol{z}'|^2}{2}-|\boldsymbol{y}|^2\right) = \frac{\widetilde{\omega}}{c_o}\frac{(|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2)}{L}+\frac{\widetilde{\omega}|\widetilde{\boldsymbol{z}}|^2}{4c_oL} \approx \frac{\widetilde{\omega}}{c_o}\frac{(|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2)}{L},$$

because

$$\frac{\widetilde{\omega}}{c_o}\frac{|\widetilde{\boldsymbol{x}}|^2}{L} = O\left(\frac{\Omega_dX_d^2}{c_oL}\right) \ll 1.$$

The kernel (D.2) simplifies as

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx & \frac{N_r^2\sqrt{2\pi}}{a^4B(4\pi L)^2}e^{-\frac{|\widetilde{\boldsymbol{z}}|^2}{2X_d^2}}\int d\widetilde{\omega}\,\exp\left[-\frac{\widetilde{\omega}^2}{2\Omega_e^2}+i\widetilde{\omega}\left(\frac{\overline{z}_3-y_3}{c_o}+\frac{|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2}{2Lc_o}\right)\right]\\ & \times\int d\omega\,\exp\left[-\frac{(\omega-\omega_o)^2}{2B^2}+i\omega\left(\frac{\widetilde{z}_3}{c_o}+\frac{\overline{\boldsymbol{z}}\cdot\widetilde{\boldsymbol{z}}}{Lc_o}\right)\right]\\ & \times\int d\widetilde{\boldsymbol{x}}\,\exp\left[-\frac{|\widetilde{\boldsymbol{x}}|^2}{2X_e^2}-\widetilde{\boldsymbol{x}}\cdot\left(\frac{\widetilde{\boldsymbol{z}}}{2X_d^2}+\frac{i\omega(\overline{\boldsymbol{z}}-\boldsymbol{y})}{Lc_o}\right)\right]\\ & \times\int d\boldsymbol{x}\,\exp\left[-\frac{|\boldsymbol{x}|^2}{(a/6)^2}-i\boldsymbol{x}\cdot\left(\frac{\widetilde{\omega}(\overline{\boldsymbol{z}}-\boldsymbol{y})}{Lc_o}+\frac{\omega\widetilde{\boldsymbol{z}}}{Lc_o}\right)\right],\end{aligned}\tag{D.4}$$

with Ω_e and X_e defined in (4.6), and after evaluating the last two integrals, we get

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx & \frac{N_r^2\sqrt{2\pi}X_e^2}{288Ba^2L^2}e^{-\frac{|\widetilde{\boldsymbol{z}}|^2}{2X_d^2}}\int d\widetilde{\omega}\,\exp\left[-\frac{\widetilde{\omega}^2}{2\Omega_e^2}+i\widetilde{\omega}\left(\frac{\overline{z}_3-y_3}{c_o}+\frac{|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2}{2Lc_o}\right)\right]\\ & \times\int d\omega\,\exp\left[-\frac{(\omega-\omega_o)^2}{2B^2}+i\omega\left(\frac{\widetilde{z}_3}{c_o}+\frac{\overline{\boldsymbol{z}}\cdot\widetilde{\boldsymbol{z}}}{Lc_o}\right)\right]\\ & \times\exp\left[-\frac{X_e^2}{2}\left|\frac{\omega(\overline{\boldsymbol{z}}-\boldsymbol{y})}{Lc_o}-\frac{i\widetilde{\boldsymbol{z}}}{2X_d^2}\right|^2-\frac{1}{2}\left(\frac{a}{6\sqrt{2}}\right)^2\left|\frac{\widetilde{\omega}(\overline{\boldsymbol{z}}-\boldsymbol{y})}{Lc_o}+\frac{\omega\widetilde{\boldsymbol{z}}}{Lc_o}\right|^2\right].\end{aligned}\tag{D.5}$$

Let us change variables $w = \omega - \omega_o$ and use the notation

$$\overline{\boldsymbol{\zeta}} = \frac{\overline{\boldsymbol{z}} - \boldsymbol{y}}{L/(k_o X_e)}, \qquad \widetilde{\boldsymbol{\zeta}} = \frac{\widetilde{\boldsymbol{z}}}{6\sqrt{2}L/(k_o a)}, \qquad \beta = \frac{\overline{z}_3 - y_3}{c_o/\Omega_e} + \frac{|\overline{\boldsymbol{z}}|^2 - |\boldsymbol{y}|^2}{2Lc_o/\Omega_e}, \qquad \theta = \frac{\omega_o X_e}{\Omega_e(a/6)}. \tag{D.6}$$

Define also

$$\frac{1}{\gamma} = 1 - \frac{X_e^2}{4X_d^2} > \frac{3}{4}, \tag{D.7}$$

with the inequality implied by the definition of X_e. Substituting in (D.5) we get

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') \approx & \frac{N_r^2\sqrt{2\pi}X_e^2}{288Ba^2L^2} \exp\left[-\frac{|\widetilde{\boldsymbol{z}}|^2}{2\gamma X_d^2} - \frac{|\overline{\boldsymbol{\zeta}}|^2}{2} - \frac{|\widetilde{\boldsymbol{\zeta}}|^2}{2} + ik_o\Big(\widetilde{z}_3 + \frac{\widetilde{\boldsymbol{z}}\cdot\overline{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}} - \boldsymbol{y})}{2X_d^2L}\Big)\right] \\ & \times \int dw \exp\left[-\frac{w^2}{2}\Big(\frac{1}{B^2} + \frac{|\overline{\boldsymbol{\zeta}}|^2 + |\widetilde{\boldsymbol{\zeta}}|^2}{\omega_o^2}\Big) + \frac{iw}{c_o}\Big(\widetilde{z}_3 + \frac{\overline{\boldsymbol{z}}\cdot\widetilde{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}} - \boldsymbol{y})}{2X_d^2L}\Big) - \frac{w(|\overline{\boldsymbol{\zeta}}|^2 + |\widetilde{\boldsymbol{\zeta}}|^2)}{\omega_o}\right] \\ & \times \int d\widetilde{\omega} \exp\left[-\frac{\widetilde{\omega}^2}{2\Omega_e^2}\Big(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2}\Big) + \frac{i\widetilde{\omega}}{\Omega_e}\Big(\beta + \frac{i(\omega_o + w)\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\sqrt{2}\omega_o\theta}\Big)\right],\end{aligned} \tag{D.8}$$

and integrate next over $\widetilde{\omega}$. We obtain after rearranging the terms that

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') \approx & \frac{\pi N_r^2 X_e^2 \Omega_e}{144Ba^2L^2\sqrt{1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2}}} \exp\left[-\frac{|\widetilde{\boldsymbol{z}}|^2}{2\gamma X_d^2} - \frac{\Delta^2}{2} - \frac{\beta^2}{2(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}\right] \\ & \times \exp\left[ik_o\Big(\widetilde{z}_3 + \frac{\widetilde{\boldsymbol{z}}\cdot\overline{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}} - \boldsymbol{y})}{2X_d^2L}\Big) - \frac{i\beta\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\sqrt{2}\theta(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}\right] \\ & \times \int dw \exp\left[-\frac{w^2}{2B^2}\Big(1 + \frac{\Delta^2B^2}{\omega_o^2}\Big) + \frac{iw\eta}{B} - \frac{w\Delta^2}{\omega_o}\right],\end{aligned} \tag{D.9}$$

with notation

$$\eta = \frac{B}{c_o}\left[\widetilde{z}_3 + \frac{\overline{\boldsymbol{z}}\cdot\widetilde{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}} - \boldsymbol{y})}{2X_d^2L} - \frac{\beta c_o\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\sqrt{2}\omega_o\theta(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}\right], \tag{D.10}$$

$$\Delta^2 = |\overline{\boldsymbol{\zeta}}|^2 + |\widetilde{\boldsymbol{\zeta}}|^2 - \frac{|\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}|^2}{2\theta^2(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}. \tag{D.11}$$

Note that (D.11) is non-negative because

$$|\widetilde{\boldsymbol{\zeta}}|^2 - \frac{|\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}|^2}{2\theta^2(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} = \frac{|\widetilde{\boldsymbol{\zeta}}|^2}{(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} + \frac{(|\overline{\boldsymbol{\zeta}}|^2|\widetilde{\boldsymbol{\zeta}}|^2 - |\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}|^2)}{(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} \geq \frac{|\widetilde{\boldsymbol{\zeta}}|^2}{(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}.$$

Now we integrate in w in equation (D.9), and obtain

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}}, \vec{\boldsymbol{z}}') \approx & \frac{\sqrt{2}\pi^{3/2}N_r^2X_e^2\Omega_e}{144a^2L^2\sqrt{1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2}}\sqrt{1 + \frac{\Delta^2B^2}{\omega_o^2}}} \exp\left[-\frac{|\widetilde{\boldsymbol{z}}|^2}{2\gamma X_d^2} - \frac{\beta^2}{2(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} - \frac{\eta^2}{2(1 + \frac{B^2\Delta^2}{\omega_o^2})} - \frac{\Delta^2}{2(1 + \frac{B^2\Delta^2}{\omega_o^2})}\right] \\ & \times \exp\left[ik_o\Big(\widetilde{z}_3 + \frac{\widetilde{\boldsymbol{z}}\cdot\overline{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}} - \boldsymbol{y})}{2X_d^2L}\Big) - \frac{i\beta\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\sqrt{2}\theta(1 + \frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} - \frac{i\eta B\Delta^2}{\omega_o(1 + \frac{B^2\Delta^2}{\omega_o^2})}\right].\end{aligned} \tag{D.12}$$

This expression simplifies, because the exponential decay in Δ^2 ensures that the kernel is large only when $\Delta = O(1)$. Since $B/\omega_o \ll 1$ by assumption, we see that

$$\frac{B^2\Delta^2}{\omega_o^2} \ll 1. \tag{D.13}$$

By the same reasoning, the kernel is large when $|\eta| = O(1)$, but then

$$\left|\frac{\eta B\Delta^2}{\omega_o(1+\frac{B^2\Delta^2}{\omega_o^2})}\right| \approx |\eta|\Delta^2\frac{B}{\omega_o} = O\left(\frac{B}{\omega_o}\right) \ll 1. \tag{D.14}$$

We also have from $\Delta = O(1)$ and definition (D.11) that both $|\overline{\boldsymbol{\zeta}}|$ and $|\widetilde{\boldsymbol{\zeta}}|$ are $O(1)$. Then, using the definition of $\widetilde{\boldsymbol{\zeta}}$ in (D.6), we can estimate

$$\frac{B}{c_o}\left|\frac{\overline{\boldsymbol{z}}\cdot\widetilde{\boldsymbol{z}}}{L}\right| \le O\left(\frac{B}{c_o}\frac{\lambda_o L/aa}{L}\right) = O\left(\frac{B}{\omega_o}\right) \ll 1, \tag{D.15}$$

and

$$\left|\frac{\beta B\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\omega_o\theta}\right| = O\left(\frac{B}{\omega_o\theta}\right) = O\left(\frac{B}{\omega_o}\right) \ll 1. \tag{D.16}$$

Here we used that θ^{-1} is at most of order 1, which follows from its definition in (D.6). Indeed, in the broadband case we obtain from assumption (3.32) that

$$\theta^{-1} = \frac{a\Omega_e}{6\omega_o X_e} = O\left(\frac{a\Omega_d}{\omega_o X_d}\right) = O\left(\frac{a}{\ell}\right) = O(1). \tag{D.17}$$

In the narrowband case we obtain

$$\theta^{-1} = \frac{a\Omega_e}{6\omega_o X_e} = O\left(\frac{aB}{\omega_o X_d}\right) \ll \frac{\lambda_o L}{X_d^2} = O\left(\frac{\lambda_o L}{\ell^2(\Omega_d/\omega_o)^2}\right) = O\left(\frac{\sigma^2 L^2}{\ell\lambda_o}\right) \ll 1, \tag{D.18}$$

where the second equality is because $\Omega_e = O(B)$, the first inequality is by assumption (3.29), the following equalities are by the definitions (3.20) of X_d and Ω_d and the last inequality is by assumption (3.26). We also have the estimate

$$\begin{aligned}\frac{B}{c_o}\left|\frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}}-\boldsymbol{y})}{X_d^2 L}\right| &= O\left(\frac{B}{c_o L}\frac{L}{k_o a}\frac{L}{k_o X_d}\right) = O\left(\frac{B}{\omega_o}\frac{\lambda_o L}{aX_d}\right) = O\left(\frac{B}{\omega_o}\frac{\lambda_o L}{a\ell\Omega_d/\omega_o}\right)\\ &= O\left(\frac{B}{\omega_o}\frac{\sigma L^{3/2}}{a\ell^{1/2}}\right) \ll \frac{B}{\omega_o}\frac{\sqrt{\lambda_o L}}{a} \ll 1,\end{aligned} \tag{D.19}$$

where we used again definitions (3.20) and the assumption (3.26). Note that estimates (D.15)-(D.19) and definition (D.10) imply that

$$\eta \approx \frac{B\widetilde{z}_3}{c_o}. \tag{D.20}$$

Substituting all the results in (D.12), we get the kernel

$$\begin{aligned}\kappa(\vec{\boldsymbol{y}},\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}') \approx &\frac{\sqrt{2}\pi^{3/2}N_r^2 X_e^2\Omega_e}{144a^2L^2\sqrt{1+\frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2}}}\exp\left[-\frac{|\widetilde{\boldsymbol{z}}|^2}{2\gamma X_d^2} - \frac{\beta^2}{2(1+\frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})} - \frac{\widetilde{z}_3^2}{2(c_o/B)^2} - \frac{\Delta^2}{2}\right]\\ &\times\exp\left[ik_o\Big(\widetilde{z}_3 + \frac{\widetilde{\boldsymbol{z}}\cdot\overline{\boldsymbol{z}}}{L} + \frac{X_e^2\widetilde{\boldsymbol{z}}\cdot(\overline{\boldsymbol{z}}-\boldsymbol{y})}{2X_d^2 L}\Big) - \frac{i\beta\overline{\boldsymbol{\zeta}}\cdot\widetilde{\boldsymbol{\zeta}}}{\sqrt{2}\theta(1+\frac{|\overline{\boldsymbol{\zeta}}|^2}{2\theta^2})}\right].\end{aligned} \tag{D.21}$$

The statement of Proposition 4.2 follows.

In the narrowband case we know from (D.18) that $\theta \gg 1$. We also have

$$\frac{\Omega_e}{c_o}\left|\frac{(|\overline{\boldsymbol{z}}|^2-|\boldsymbol{y}|^2)}{L}\right| = O\left(\frac{B|(\overline{\boldsymbol{z}}-\boldsymbol{y})\cdot(\overline{\boldsymbol{z}}+\boldsymbol{y})|}{c_o L}\right) = O\left(\frac{B}{c_o}\frac{L/(k_o X_e)a}{L}\right) = O\left(\frac{\lambda_o L}{X_d^2}\right) = O\left(\frac{\sigma^2 L^2}{\lambda_o \ell}\right) \ll 1, \tag{D.22}$$

where we used that $\Omega_e = O(B)$, $X_e = O(X_d)$ and $|\overline{\boldsymbol{\zeta}}| = O(1)$ i.e., $|\overline{\boldsymbol{z}}-\boldsymbol{y}| = O(L/(k_o X_d)$. This estimate follows from definition (3.20) and assumptions (3.26) and (3.29). The expression (4.4) in Proposition 4.1 follows from (D.21) and definition (D.6) of β. □

Appendix E. Proof of Lemma 5.8. We estimate the inner product of the columns of the matrix using the continuum approximation

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle = \sum_{\vec{\boldsymbol{y}}\in\mathfrak{D}_z} m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}}\, m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}'} \approx \frac{1}{h^2 h_3}\int_{-\frac{D}{2}}^{\frac{D}{2}} dy_1 \int_{-\frac{D}{2}}^{\frac{D}{2}} dy_2 \int_{-\frac{D_3}{2}}^{\frac{D_3}{2}} m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}}\, m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}'}, \tag{E.1}$$

where $\vec{\boldsymbol{y}} = (\boldsymbol{y}, y_3)$, with $\boldsymbol{y} = (y_1, y_2)$, and h and h_3 are the mesh sizes in cross-range and range.

E.1. The narrowband regime. With the expression (4.15) of $m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}}$ we get

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx \frac{C^2}{h^2 h_3}\int_{\mathbb{R}^2} d\boldsymbol{y}\, \exp\left[-\frac{(|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2)}{2R^2}\right]\int_{-\infty}^{\infty} dy_3\, \exp\left[-\frac{(y_3-z_3)^2}{2R_3^2}-\frac{(y_3-z_3')^2}{2R_3^2}\right], \tag{E.2}$$

where we used the paraxial approximation and extended the integrals to the whole space with negligible error[¶], due to the Gaussians. Evaluating the integrals,

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx \frac{C^2\pi^{3/2}R^2 R_3}{h^2 h_3}\exp\left[-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}-\frac{(z_3-z_3')^2}{4R_3^2}\right]. \tag{E.3}$$

Obviously, the maximum of (E.3) is attained at $\vec{\boldsymbol{z}} = \vec{\boldsymbol{z}}'$, where

$$\|\boldsymbol{m}_{\vec{\boldsymbol{z}}}\|_2^2 = \langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}}\rangle \approx \frac{C^2\pi^{3/2}R^2 R_3}{h^2 h_3}.$$

The result in Lemma 5.8 follows. □

E.2. The broadband regime. With the expression (4.16) of $m_{\vec{\boldsymbol{y}},\vec{\boldsymbol{z}}}$ we get

$$\begin{aligned}\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx \frac{C^2}{h^2 h_3}\int_{\mathbb{R}^2} d\boldsymbol{y}\, &\frac{\exp\left[-\frac{(|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2)}{2R^2}\right]}{\left[1+\frac{|\boldsymbol{y}-\boldsymbol{z}|^2}{2(\theta R)^2}\right]^{1/2}\left[1+\frac{|\boldsymbol{y}-\boldsymbol{z}'|^2}{2(\theta R)^2}\right]^{1/2}}\\ &\times\int_{-\infty}^{\infty} dy_3\, \exp\left[-\frac{\left(y_3-z_3+\frac{|\boldsymbol{y}|^2-|\boldsymbol{z}|^2}{2L}\right)^2}{2R_3^2\left(1+\frac{|\boldsymbol{y}-\boldsymbol{z}|^2}{2\theta^2R^2}\right)}-\frac{\left(y_3-z_3'+\frac{|\boldsymbol{y}|^2-|\boldsymbol{z}'|^2}{2L}\right)^2}{2R_3^2\left(1+\frac{|\boldsymbol{y}-\boldsymbol{z}'|^2}{2\theta^2R^2}\right)}\right],\end{aligned} \tag{E.4}$$

where we used the paraxial approximation and extended the integrals to the whole space with negligible error[‖], due to the Gaussians. Evaluating the integral over y_3 and renaming the constant, we obtain

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}}, \boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx CR_3\int_{\mathbb{R}^2} d\boldsymbol{y}\, \frac{\exp\left[-\frac{(|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2)}{2R^2}-\frac{\left(z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}\right)^2}{4R_3^2\left(1+\frac{|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2}{4\theta^2R^2}\right]}\right]}{\left(1+\frac{|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2}{4\theta^2R^2}\right)^{1/2}}. \tag{E.5}$$

[¶] This is assuming that $\vec{\boldsymbol{z}}$ and $\vec{\boldsymbol{z}}'$ (i.e., the sources) are not near the edge of the imaging region.

[‖] This is assuming that $\vec{\boldsymbol{z}}$ and $\vec{\boldsymbol{z}}'$ (i.e., the sources) are not near the edge of the imaging region.

Note that

$$\frac{|\boldsymbol{y}-\boldsymbol{z}|^2+|\boldsymbol{y}-\boldsymbol{z}'|^2}{2}=|\boldsymbol{y}-\overline{\boldsymbol{z}}|^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4},$$

where $\overline{\boldsymbol{z}}=(\boldsymbol{z}+\boldsymbol{z}')/2$. Substituting in (E.3) and changing variables as $\boldsymbol{v}=(\boldsymbol{y}-\overline{\boldsymbol{z}})/R$, we get

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}},\boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx CR^2R_3e^{-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}\int_{\mathbb{R}^2}d\boldsymbol{v}\,\frac{\exp\left[-|\boldsymbol{v}|^2-\frac{\left(z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}\right)^2}{4R_3^2\left(1+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8\theta^2R^2}+\frac{|\boldsymbol{v}|^2}{2\theta^2}\right)}\right]}{\left(1+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8\theta^2R^2}+\frac{|\boldsymbol{v}|^2}{2\theta^2}\right)^{1/2}}.$$

The integrand depends only on $v=|\boldsymbol{v}|$, so we can write the integral in polar coordinates and obtain

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}},\boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx 2\pi CR^2R_3e^{-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}\int_0^\infty dv\,v\,\frac{\exp\left[-v^2-\frac{\left(z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}\right)^2}{4R_3^2\left(1+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8\theta^2R^2}+\frac{v^2}{2\theta^2}\right)}\right]}{\left(1+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8\theta^2R^2}+\frac{v^2}{2\theta^2}\right)^{1/2}}. \tag{E.6}$$

To eliminate the algebraic factors, let us change coordinates again

$$t=\theta\left[2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4\theta^2R^2}+\frac{v^2}{\theta^2}\right]^{1/2}\quad\text{s.t.}\quad dt=\frac{1}{\theta}\left[2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4\theta^2R^2}+\frac{v^2}{\theta^2}\right]^{-1/2}vdv.$$

Equation (E.6) becomes

$$\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}},\boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx 2\pi\theta CR^2R_3e^{2\theta^2}\int_{\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}}^\infty dt\;\exp\left[-\frac{\left(z_3-z_3'+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2L}\right)^2}{2R_3^2t^2}-t^2\right],$$

with integral over t evaluated below, in terms of the complementary error function

$$\begin{aligned}\langle \boldsymbol{m}_{\vec{\boldsymbol{z}}},\boldsymbol{m}_{\vec{\boldsymbol{z}}'}\rangle \approx \frac{\pi^{3/2}}{4}\theta CR^2R_3e^{2\theta^2}\Bigg\{&e^{-\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}-\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right]\\&+e^{\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}+\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right]\Bigg\}.\end{aligned} \tag{E.7}$$

We are interested in the cross-correlation defined in (5.2). The norm $\|\boldsymbol{m}_{\vec{\boldsymbol{z}}}\|_2$ is obtained by letting $\vec{\boldsymbol{z}}=\vec{\boldsymbol{z}}'$ in (E.7), and the result is

$$\begin{aligned}\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \approx \frac{1}{2\mathrm{erfc}(\sqrt{2}\theta)}\Bigg\{&e^{-\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}-\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right]\\&+e^{\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}+\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right]\Bigg\}.\end{aligned} \tag{E.8}$$

We can bound the right hand side using the elementary inequality $\mathrm{erfc}(t) \le e^{-t^2}$, for all $t \ge 0$. This gives

$$
\begin{aligned}
e^{\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}&\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}+\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right] \\
&\le \exp\left[-2\theta^2-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}-\frac{\theta^2\left(\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right)^2}{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}\right] \\
&\le \exp\left[-\theta^2-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}-\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|\right], \qquad \text{(E.9)}
\end{aligned}
$$

where the last inequality is because

$$
\begin{aligned}
\left[2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}+\frac{\theta^2\left(\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right)^2}{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}\right]-\left[\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}+\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|\right]= \\
\left[\sqrt{\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}}-\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{2\sqrt{\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}}}\right]^2 \ge 0.
\end{aligned}
$$

For the other term in (E.8) the bound is the same when the argument of the complementary error function is non-negative. If the argument is negative, then using that $\mathrm{erfc}(t) \le 2$ for all t, we get

$$
e^{-\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}-\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right] \le 2e^{-\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|} \qquad \text{(E.10)}
$$

But since in this case

$$
\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right| > 2\sqrt{2}\left(\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}\right),
$$

we can write

$$
(\sqrt{2}-1)\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right| > 2\sqrt{2}(\sqrt{2}-1)\left(\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}\right) > \theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}.
$$

Substituting in (E.10) we get that

$$
\begin{aligned}
e^{-\sqrt{2}\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}\mathrm{erfc}&\left[\sqrt{2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}}-\frac{\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|}{\sqrt{2\left(2\theta^2+\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{4R^2}\right)}}\right] \le \\
&2\exp\left[-\theta^2-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}-\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|\right], \qquad \text{(E.11)}
\end{aligned}
$$

and using this result and (E.9) in (E.8) we get

$$
\mathcal{I}_{\vec{\boldsymbol{z}},\vec{\boldsymbol{z}}'} \le \frac{3e^{-\theta^2}}{2\mathrm{erfc}(\sqrt{2}\theta)}\exp\left[-\frac{|\boldsymbol{z}-\boldsymbol{z}'|^2}{8R^2}-\theta\left|\frac{z_3-z_3'}{R_3}+\frac{|\boldsymbol{z}|^2-|\boldsymbol{z}'|^2}{2LR_3}\right|\right]. \qquad \text{(E.12)}
$$

This is the result in Lemma 5.8. □

REFERENCES

[1] Laura Anitori, Ali Maleki, Matern Otten, Richard G Baraniuk, and Peter Hoogeboom, *Design and analysis of compressed sensing radar detectors*, Signal Processing, IEEE Transactions on, 61 (2013), pp. 813–827.

[2] Simon R Arridge and John C Schotland, *Optical tomography: forward and inverse problems*, Inverse Problems, 25 (2009), p. 123010.

[3] Jacques M Beckers, *Adaptive optics for astronomy-principles, performance, and applications*, Annual review of astronomy and astrophysics, 31 (1993), pp. 13–62.

[4] B. Biondi, *3D seismic imaging*, Society of Exploration Geophysicists, 2006.

[5] N. Bleistein, J. K. Cohen, and J.J.W. Stockwell, *Mathematics of multidimensional seismic imaging, migration, and inversion*, vol. 13, Springer, 2001.

[6] Liliana Borcea, Josselin Garnier, George Papanicolaou, and Chrysoula Tsogka, *Enhanced statistical stability in coherent interferometric imaging*, Inverse problems, 27 (2011), p. 085004.

[7] Liliana Borcea and Ilker Kocyigit, *Resolution analysis of imaging with l1 optimization*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 3015–3050.

[8] Liliana Borcea, Miguel Moscoso, George Papanicolaou, and Chrysoula Tsogka, *Synthetic aperture imaging of direction-and frequency-dependent reflectivities*, SIAM Journal on Imaging Sciences, 9 (2016), pp. 52–81.

[9] Liliana Borcea, George Papanicolaou, and Chrysoula Tsogka, *Adaptive interferometric imaging in clutter and optimal illumination*, Inverse Problems, 22 (2006), p. 1405.

[10] ———, *Asymptotics for the space-time wigner transform with applications to imaging*, Stochastic Differential Equations: Theory and Applications (in Honor of Prof. Boris L. Rozovskii), Interdiscip. Math. Sci, 2 (2007), pp. 91–112.

[11] Emmanuel J Candès and Carlos Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics, 67 (2014), pp. 906–956.

[12] Emmanuel J Candes, Thomas Strohmer, and Vladislav Voroninski, *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, 66 (2013), pp. 1241–1274.

[13] Anwei Chai, Miguel Moscoso, and George Papanicolaou, *Array imaging using intensity-only measurements*, Inverse Problems, 27 (2010), p. 015005.

[14] A. Chai, M. Moscoso, and G. Papanicolaou, *Robust imaging of localized scatterers using the singular value decomposition and l1 minimization*, Inverse Problems, 29 (2013), p. 025016.

[15] John C. Curlander and Robert N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*, Wiley-Interscience, 1991.

[16] CVX Research, *Cvx: matlab software for disciplined convex programming, version 2.0.* http://cvxr.com/cvx`http://cvxr.com/cvx`, August 2012.

[17] Luc Devroye, *Nonuniform random variate generation*, Handbooks in operations research and management science, 13 (2006), pp. 83–121.

[18] A. C. Fannjiang, *Compressive inverse scattering: I. High-frequency SIMO/MISO and MIMO measurements*, Inverse Problems, 26 (2010), p. 035008.

[19] A. C. Fannjiang, T. Strohmer, and P. Yan, *Compressed remote sensing of sparse objects*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 595–618.

[20] David C Munson, James Dennis O'Brien, and W Kenneth Jenkins, *A tomographic formulation of spotlight-mode synthetic aperture radar*, Proceedings of the IEEE, 71 (1983), pp. 917–925.

[21] George Papanicolaou, Lenya Ryzhik, and Knut Sølna, *Self-averaging from lateral diversity in the Itô-Schrödinger equation*, Multiscale Modeling & Simulation, 6 (2007), pp. 468–492.

[22] SM Rytov, Yu A Kravtsov, and VI Tatarskii, *Principle of statistical radiophysics iv: Wave propagation through random media. chapter 4*, 1989.

[23] MCW van van Rossum and Th M Nieuwenhuizen, *Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion*, Reviews of Modern Physics, 71 (1999), p. 313.