STATISTICAL STABILITY IN TIME REVERSAL

GEORGE PAPANICOLAOU *, LEONID RYZHIK †, AND KNUT SØLNA ‡

Abstract. When a signal is emitted from a source, recorded by an array of transducers, time reversed and re-emitted into the medium, it will refocus approximately on the source location. We analyze the refocusing resolution in a high frequency, remote sensing regime, and show that, because of multiple scattering, in an inhomogeneous or random medium it can improve beyond the diffraction limit. We also show that the back-propagated signal from a spatially localized narrow-band source is self-averaging, or statistically stable, and relate this to the self-averaging properties of functionals of the Wigner distribution in phase space. Time reversal from spatially distributed sources is self-averaging only for broad-band signals. The array of transducers operates in a remote-sensing regime so we analyze time reversal with the parabolic or paraxial wave equation.

Key words. wave propagation, random medium, Liouville-Ito equation, stochastic-flow, time reversal

AMS subject classifications. 35L05, 60H15, 35Q60

1. Introduction. In time reversal experiments a signal emitted by a localized source is recorded by an array and then re-emitted into the medium time-reversed, that is, the tail of the recorded signal is sent back first. In the absence of absorption the re-emitted signal propagates back toward the source and focuses approximately on it. This phenomenon has numerous applications in medicine, underwater acoustics and elsewhere and has been extensively studied in the literature, both from the experimental and theoretical points of view [12, 13, 14, 15, 16, 20, 24, 25, 31]. Recently time reversal has been also the subject of active mathematical research in the context of wave propagation and imaging in random media [2, 3, 4, 7, 8, 9, 32]. A schematic description of a time reversal experiment is presented in Figure 1.1.

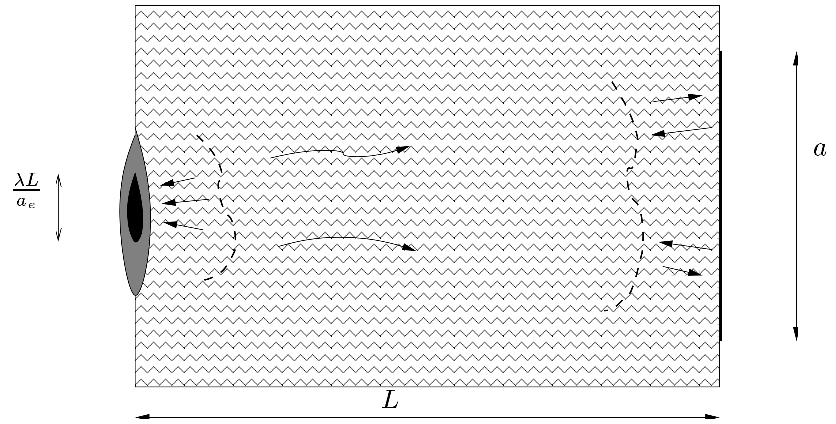

FIG. 1.1. *A pulse propagates toward a time reversal array of size a. The propagation distance L is large compared to a. The ambient medium has a randomly varying index of refraction with a typical correlation length that is small compared to a. The signal is time reversed at the array and sent back into the medium. The back propagated signal refocuses with spot size $\lambda L/a_e$, where a_e is the effective aperture of the array (Section 3.3).*

*Department of Mathematics, Stanford University, Stanford CA, 94305; papanico@math.stanford.edu,
†Department of Mathematics, University of Chicago, Chicago IL, 60637; ryzhik@math.uchicago.edu,
‡ Department of Mathematics, University of California, Irvine CA, 92697; ksolna@math.uci.edu.

For a point source in a homogeneous medium, the size of the refocused spot is approximately $\lambda L/a$, where λ is the central wavelength of the emitted signal, L is the distance between the source and the transducer array and a is the size of the array. We assume here that the array is operating in the remote-sensing regime $a \ll L$. Multiple scattering in a randomly inhomogeneous medium creates **multipathing**, which means that the transducer array can capture waves that were initially moving away from it but get scattered onto it by the inhomogeneities. As a result, the array captures a wider aperture of rays emanating from the original source and appears to be larger than its physical size. Therefore, somewhat contrary to intuition, the inhomogeneities of the medium do not destroy the refocusing but enhance its resolution. The refocused spot is now $\lambda L/a_e$, where $a_e > a$ is the **effective** size of the array in the randomly scattering medium, and depends on L. The enhancement of refocusing resolution by multipathing is called **super-resolution** [7]. The time reversed pulse is also **self averaging** and refocusing near the source is therefore **statistically stable**, which means that it does not depend on the particular realization of the random medium. There is some loss of energy in the refocused signal because of scattering away from the array but this can be overcome by amplification, up to a point.

The purpose of this paper is to explore in detail the mathematical basis of pulse stabilization, beyond what was done in [7]. We want to explore in particular in what regime of parameters statistical stability is observed in time reversal. We show here that for high frequency waves in a remote sensing regime, spatially localized sources lead to statistically stable super-resolution in time reversal, even for narrow-band signals. We also show that when the source is spatially distributed, only for broad-band signals do we have statistical stability in time reversal. The regime where our analysis holds is a high frequency one, more appropriate to optical or infrared time reversal than to ultrasound, sonar or microwave radar. In this regime we can make precise what spatially localized or distributed means (see Section 3.1). The numerical simulations in [7] and [8], which are set in ultrasound or underwater sound regime, indicate that time reversal is not statistically stable for narrow-band signals even for localized sources. Only for broad-band signals is time reversal statistically stable in the regime of ultrasound experiments or sonar.

If the aperture of the transducer array is small $a/L \ll 1$, the Fresnel number $L/(ka^2)$ is of order one, and the random inhomogeneities are weak, which is often the case, we may analyze wave propagation in the paraxial or parabolic approximation [29]. The wave field is then given approximately by

$$u(t,\mathbf{x},z) = \frac{1}{2\pi}\int e^{i\omega(z/c_0-t)}\psi(z,\mathbf{x};\omega/c_0)d\omega \tag{1.1}$$

where the complex amplitude ψ satisfies the parabolic or Schrödinger equation

$$2ik\psi_z + \Delta_{\mathbf{x}}\psi + k^2(n^2-1)\psi = 0. \tag{1.2}$$

Here $\mathbf{x} = (x, y)$ are the coordinates transverse to the direction of propagation z, the wave number $k = \omega/c_0$ and $n(\mathbf{x}, z) = c_0/c(\mathbf{x}, z)$ is the random index of refraction relative to a reference speed c_0. The fluctuations of the refraction index

$$\sigma\mu(\frac{\mathbf{x}}{l},\frac{z}{l}) = n^2(\mathbf{x},z) - 1 \tag{1.3}$$

are assumed to be a stationary random field with mean zero, variance σ^2, correlation length l and normalized covariance with dimensionless arguments

$$R(\mathbf{x},z) = E\{\mu(\mathbf{x}+\mathbf{x}',z+z')\mu(\mathbf{x}',z')\}. \tag{1.4}$$

A convenient tool for the analysis of wave propagation in a random medium is the Wigner distribution [19, 28] defined by

$$W(z,\mathbf{x},\mathbf{p}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\mathbf{p}\cdot\mathbf{y}}\psi(\mathbf{x}-\frac{\mathbf{y}}{2},z)\overline{\psi(\mathbf{x}+\frac{\mathbf{y}}{2},z)}d\mathbf{y} \tag{1.5}$$

where $d = 1$ or 2 is the transverse dimension and the bar denotes complex conjugate. The Wigner distribution may be interpreted as phase space wave energy and is particularly well suited for high frequency asymptotics and random media [28]. The quantity of principal interest in time reversal, the time-reversed and back-propagated wave field, can be also expressed in terms of the Wigner distribution (see Section 3.1). The self-averaging properties of the back-propagated field are related to the self-averaging properties of functionals of the Wigner distribution in the form of integrals of W over the wave numbers $\mathbf{p}$.

In the next Section we introduce a precise scaling that corresponds to (a) high frequency, (b) long propagation distance, (c) narrow beam propagation, and (d) weak random fluctuations. In the asymptotic limit where the small parameters go to zero the Wigner distribution satisfies a stochastic partial differential equation (SPDE), a Liouville-Ito equation, that has the form

$$dW(z,\mathbf{x},\mathbf{p};k) = \left(-\frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}}W + \frac{k^2 D}{2}\Delta_{\mathbf{p}}W\right)dz - \frac{k}{2}\nabla_{\mathbf{p}}W\cdot d\mathbf{B}(\mathbf{x},z) \tag{1.6}$$

where $\mathbf{B}(\mathbf{x},z)$ is a vector-valued Brownian field with covariance

$$E\{B_i(\mathbf{x}_1,z_1)B_j(\mathbf{x}_2,z_2)\} = -\left(\frac{\partial^2 R_0((\mathbf{x}_1-\mathbf{x}_2))}{\partial x_i\partial x_j}\right) z_1\wedge z_2, \tag{1.7}$$

where $z_1\wedge z_2 = \min\{z_1,z_2\}$, and in the isotropic case

$$D = -\frac{R_0''(0)}{4},\quad R_0(\mathbf{x}) = \int_{-\infty}^{\infty} R(\mathbf{x},s)ds. \tag{1.8}$$

In Section 2.5 we analyze this SPDE in the asymptotic limit of small correlation length for $\mathbf{B}(\mathbf{x},z)$ in the transverse variables $\mathbf{x}$, and show that $W(z,\mathbf{x},\mathbf{p};k)$'s with different wave vectors $\mathbf{p}$ are uncorrelated. From this decorrelation property we deduce that for localized sources the time-reversed, back-propagated field is self-averaging, even for narrow-band signals. For distributed sources it is self-averaging only for broad-band signals. We show in detail in Section 3 how the asymptotic theory is used in time reversal. In Appendix A we introduce other scalings which lead to the same averaged SPDE but we do not analyze them in detail.

Throughout the paper we define the Fourier transform by

$$\hat{f}(\mathbf{k}) = \int d\mathbf{x}e^{-i\mathbf{k}\cdot\mathbf{x}}f(\mathbf{x})$$

so that

$$f(\mathbf{x}) = \int \frac{d\mathbf{k}}{(2\pi)^d}e^{i\mathbf{k}\cdot\mathbf{x}}\hat{f}(\mathbf{k}).$$

G. Papanicolaou was supported in part by grants AFOSR F49620-01-1-0465, NSF DMS-9971972 and ONR N00014-02-1-0088, L. Ryzhik by NSF grant DMS-9971742, an Alfred P. Sloan Fellowship and ONR grant N00014-02-1-0089. K. Solna by NSF grant DMS-0093992 and ONR grant N00014-02-1-0090.

2. Scaling and asymptotics.

2.1. The rescaled problem. To carry out the asymptotic analysis we begin by rewriting the Schrödinger equation (1.2) in dimensionless form. Let L_z and $L_{\mathbf{x}}$ be characteristic length scales in the propagation direction, as, for example, the distance L between the source and the transducer array for L_z and the array size a for L_x. We introduce a dimensionless wave number $k' = k/k_0$ with $k_0 = \omega_0/c_0$ and ω_0 a central frequency. We rescale $\mathbf{x}$ and z by $\mathbf{x} = L_{\mathbf{x}}\mathbf{x}'$, $z = L_z z'$ and rewrite (1.2) in the new coordinates dropping primes:

$$2ik\frac{\partial\psi}{\partial z} + \frac{L_z}{k_0 L_{\mathbf{x}}^2}\Delta\psi + k^2 k_0 L_z \sigma\mu\left(\frac{\mathbf{x}L_{\mathbf{x}}}{l}, \frac{zL_z}{l}\right)\psi = 0. \tag{2.1}$$

The physical parameters that characterize the propagation problem are: (a) the central wave number k_0, (b) the strength of the fluctuations σ, and (c) the correlation length l. We introduce now three dimensionless variables

$$\delta = \frac{l}{L_{\mathbf{x}}}, \quad \varepsilon = \frac{l}{L_z}, \quad \gamma = \frac{1}{k_0 l} \tag{2.2}$$

which are the reciprocals of the **transverse scale** relative to correlation length, the reciprocal of the **propagation distance** relative to correlation length, and the central **wave length** relative to the correlation length. We will assume that the dimensionless parameters γ, σ, ε and δ are small

$$\gamma \ll 1; \quad \sigma \ll 1; \quad \delta \ll 1; \quad \varepsilon \ll 1\ . \tag{2.3}$$

This is a regime of parameters where super-resolution phenomena can be observed.

To make the scaling more precise we introduce the Fresnel number

$$\theta = \frac{L_z}{k_0 L_{\mathbf{x}}^2} = \gamma\frac{\delta^2}{\varepsilon}. \tag{2.4}$$

We can then rewrite the Schrödinger equation (2.1) in the form

$$2ik\theta\psi_z + \theta^2\Delta_{\mathbf{x}}\psi + \frac{k^2\delta}{\varepsilon^{1/2}}\mu(\frac{\mathbf{x}}{\delta}, \frac{z}{\varepsilon})\psi = 0. \tag{2.5}$$

provided that we relate ε to σ and δ by

$$\varepsilon = \sigma^{2/3}\delta^{2/3}. \tag{2.6}$$

One way that the asymptotic regime (2.3) can be realized is with the ordering

$$\theta \ll \varepsilon \ll \delta \ll 1\ , \tag{2.7}$$

and $\gamma \ll \sigma^{4/3}\delta^{-2/3}$, corresponding to the high-frequency limit. We see from the scaled Schrödinger equation (2.5) that this regime can be given the following interpretation. We have first a **high frequency** limit $\theta \to 0$, then a **white noise** limit $\varepsilon \to 0$, and then a **broad beam** limit $\delta \to 0$. We will analyze in detail and interpret these limits in the following Sections. Another scaling in which (2.3) is realized is $\varepsilon \ll \theta \ll \delta \ll 1$. This is a regime in which the white noise limit is carried out first, then the high frequency limit and then the broad beam limit. We do not analyze this case here. Additional comments on scaling are provided in Appendix A.

It is instructive to express the constraints (2.6) and (2.7) in terms of the dimensional parameters of the problem. First, both the size of the transverse scale $L_{\mathbf{x}}$ and the propagation distance L_z should be much larger than the correlation length l of the medium. Moreover, (2.6) implies that the longitudinal and transverse scales should be related by

$$\frac{L_z}{L_{\mathbf{x}}}=\left(\frac{\delta}{\sigma^2}\right)^{1/3}\gg 1$$

so that we are indeed in the beam approximation. The first inequality in (2.7) implies that

$$\frac{L_z}{L_{\mathbf{x}}}\ll\sqrt{k_0 l}=\frac{1}{\sqrt{\gamma}},$$

and with the above choice of L_z this implies that

$$\frac{\gamma^{3/2}}{\sigma^2}\ll\frac{L_{\mathbf{x}}}{l}\ll\frac{1}{\sigma^2}.$$

2.2. The high frequency limit. A convenient tool for the study of the high frequency limit, especially in random media, is the Wigner distribution. It is often used in the context of energy propagation [19, 28] but it is also useful in analyzing time reversal phenomena [2, 3, 7]. Let $\phi_\theta(\mathbf{x})$ be a family of functions oscillating on a small scale θ. The Wigner distribution is a function of the physical space coordinate $\mathbf{x}$ and wave vector $\mathbf{p}$ defined as

$$W_\theta(\mathbf{x},\mathbf{p})=\int_{\mathbb{R}^d}\frac{d\mathbf{y}}{(2\pi)^d}e^{i\mathbf{p}\cdot\mathbf{y}}\phi_\theta(\mathbf{x}-\frac{\theta\mathbf{y}}{2})\overline{\phi_\theta(\mathbf{x}+\frac{\theta\mathbf{y}}{2})}. \tag{2.8}$$

The family W_θ is bounded in the space of Schwartz distributions $\mathcal{S}'(\mathbb{R}^d\times\mathbb{R}^d)$ if the functions ϕ_θ are uniformly bounded in $L^2(\mathbb{R}^d)$. Therefore there exists a subsequence $\theta_k\to 0$ such that W_{θ_k} converges weakly as $k\to\infty$ to a limit measure $W(\mathbf{x},\mathbf{p})$. This limit $W(\mathbf{x},\mathbf{p})$ is non-negative and is customarily interpreted as the limit phase space energy density because

$$|\phi_{\theta_k}(\mathbf{x})|^2\to\int_{\mathbb{R}^d}W(\mathbf{x},\mathbf{p})d\mathbf{p}\quad\text{as }\theta\to 0 \tag{2.9}$$

in the weak sense. This allows one to think of $W(\mathbf{x},\mathbf{p})$ as a local energy density.

Let $W_\theta(z,\mathbf{x},\mathbf{p})$ be the Wigner distribution of the solution ψ of the Schrödinger equation (2.5), in the transversal space-variable $\mathbf{x}$. A straightforward calculation shows that $W_\theta(z,\mathbf{x},\mathbf{p})$ satisfies in a weak sense the linear evolution equation

$$\begin{aligned}&\frac{\partial W_\theta}{\partial z}+\frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}}W_\theta\\&=\frac{ik\delta}{2\sqrt{\varepsilon}}\int e^{i\mathbf{q}\cdot\mathbf{x}/\delta}\hat{\mu}\left(q,\frac{z}{\varepsilon}\right)\frac{W_\theta\left(\mathbf{p}-\frac{\theta\mathbf{q}}{2\delta}\right)-W_\theta\left(\mathbf{p}+\frac{\theta\mathbf{q}}{2\delta}\right)}{\theta}\frac{d\mathbf{q}}{(2\pi)^d}.\end{aligned} \tag{2.10}$$

In the limit $\theta\to 0$ the solution converges weakly in $\mathcal{S}'$, for each realization, to the (weak) solution of the random Liouville equation

$$\frac{\partial W}{\partial z}+\frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}}W+\frac{k}{2\sqrt{\varepsilon}}\nabla_{\mathbf{x}}\mu\left(\frac{\mathbf{x}}{\delta},\frac{z}{\varepsilon}\right)\cdot\nabla_{\mathbf{p}}W=0. \tag{2.11}$$

The initial condition at $z=0$ is $W(0,\mathbf{x},\mathbf{p})=W_I(\mathbf{x},\mathbf{p})$, the limit Wigner distribution of the initial wave function.

2.3. The white noise limit. In this Section we take the white noise limit $\varepsilon \to 0$ in the random Liouville equation (2.11) whose solution we now denote by W_ε. We can do this using the asymptotic theory of stochastic differential equations and flows [22, 6, 21, 26] as follows. Using the method of characteristics, the solution of the Liouville equation (2.11) may be written in the form

$$W_\varepsilon(t, \mathbf{x}, \mathbf{p}) = W_I(\mathbf{X}_\varepsilon(t; \mathbf{x}, \mathbf{p}), \mathbf{P}_\varepsilon(t; \mathbf{x}, \mathbf{p})),$$

where the processes $\mathbf{X}_\varepsilon(t; \mathbf{x}, \mathbf{p})$ and $\mathbf{P}_\varepsilon(t; \mathbf{x}, \mathbf{p})$ are solutions of the characteristic equations

$$\frac{d\mathbf{X}_\varepsilon}{dz} = -\frac{1}{k}\mathbf{P}_\varepsilon; \quad \frac{d\mathbf{P}_\varepsilon}{dz} = -\frac{k}{2\sqrt{\varepsilon}}\nabla_{\mathbf{x}}\mu\left(\frac{\mathbf{X}_\varepsilon}{\delta}, \frac{z}{\varepsilon}\right)$$

with the initial conditions $\mathbf{X}_\varepsilon(0) = \mathbf{x}$ and $\mathbf{P}_\varepsilon(0) = \mathbf{p}$. The asymptotic theory of random differential equations with rapidly oscillating coefficients implies that, under suitable conditions on μ, in the limit $\varepsilon \to 0$ the processes $\mathbf{X}_\varepsilon$, $\mathbf{P}_\varepsilon$ converge weakly (in the probabilistic sense), and uniformly on compact sets in $\mathbf{x}, \mathbf{p}$ to the limit processes $\mathbf{X}(t)$, $\mathbf{P}(t)$ that satisfy a system of stochastic differential equations

$$d\mathbf{P} = -\frac{k}{2}d\mathbf{B}(z), \quad d\mathbf{X} = -\frac{1}{k}\mathbf{P}dz, \quad \mathbf{X}(0) = \mathbf{x}, \quad \mathbf{P}(0) = \mathbf{p}.$$

The random process $\mathbf{B}(z)$ is a Brownian motion with the covariance function

$$E\left\{B_i(z_1)B_j(z_2)\right\} = -\frac{\partial^2 R_0(0)}{\partial x_i \partial x_j}ds z_1 \wedge z_2 \\ = \delta_{ij}\left(-R_0''(0)\right) z_1 \wedge z_2, \tag{2.12}$$

in the isotropic case, where

$$R_0(\mathbf{x}) = \int_{-\infty}^{\infty} R(\mathbf{x}, s)ds \tag{2.13}$$

is a function of $|\mathbf{x}|$. This implies that the average Wigner distribution $W_\varepsilon^{(1)}(z, \mathbf{x}, \mathbf{p}) = E\left\{W_\varepsilon(z, \mathbf{x}, \mathbf{p})\right\}$ converges as $\varepsilon \to 0$ uniformly on compact sets to the solution of the advection-diffusion equation in phase space

$$\frac{\partial W^{(1)}}{\partial z} + \frac{\mathbf{p}}{k} \cdot \nabla_{\mathbf{x}} W^{(1)} = \frac{k^2 D}{2}\Delta_{\mathbf{p}} W^{(1)} \tag{2.14}$$

with the initial data $W^{(1)}(0, \mathbf{x}, \mathbf{p}) = W_I(\mathbf{x}, \mathbf{p})$. Here the diffusion coefficient D is given by

$$D = -\frac{R_0''(0)}{4}. \tag{2.15}$$

The one-point moments $E\left\{[W_\varepsilon(z, \mathbf{x}, \mathbf{p})]^N\right\}$ converge as $\varepsilon \to 0$ to the functions $W^{(N)}(z, \mathbf{x}, \mathbf{p})$ that satisfy the same equation (2.14) but with the initial data $W^{(N)}(0, \mathbf{x}, \mathbf{p}) = [W_0(\mathbf{x}, \mathbf{p})]^N$. This is similar to the spot dancing phenomenon [11], where all one-point moments are governed by the same Brownian motion. In particular we have that

$$W^{(2)}(z, \mathbf{x}, \mathbf{p}) \neq \left[W^{(1)}(z, \mathbf{x}, \mathbf{p})\right]^2$$

so that the process W_ε does not converge to a deterministic one, in the strong sense pointwise.

2.4. Multi-point moment equations. As in the previous Section we may also study the white noise limit $\varepsilon \to 0$ of the higher moments of $W_\varepsilon(z,\mathbf{x},\mathbf{p})$ at different points

$$W_\varepsilon^{(N)}(z,\mathbf{x}^1,\ldots,\mathbf{x}^N,\mathbf{p}^1,\ldots,\mathbf{p}^N) = E\left\{[W_\varepsilon(z,\mathbf{x}^1,\mathbf{p}^1)]^{r_1}\cdot\ldots\cdot[W_\varepsilon(z,\mathbf{x}^N,\mathbf{p}^N)]^{r_N}\right\}.$$

Here the points $(\mathbf{x}^m,\mathbf{p}^m)$ are all distinct, $(\mathbf{x}^n,\mathbf{p}^n)\neq(\mathbf{x}^m,\mathbf{p}^m)$. We may account for moments that have different powers of W_ε at different points by taking different powers r_j of $W_\varepsilon(\mathbf{x}^j,\mathbf{p}^j)$.

We now consider the joint process $(\mathbf{X}_\varepsilon(z;\mathbf{x}^m,\mathbf{p}^m),\mathbf{P}_\varepsilon(z;\mathbf{x}^m,\mathbf{p}^m))$, $m=1,\ldots,N$. As $\varepsilon\to 0$ it converges to the solution of the system of stochastic differential equations

$$d\mathbf{P}_i^m = -\frac{k}{2}\sum_{n=1}^N\sum_{j=1}^d \sigma_{ij}\left(\frac{\mathbf{X}^m-\mathbf{X}^n}{\delta}\right)dB_j^n(z),\quad d\mathbf{X}^m = -\frac{1}{k}\mathbf{P}^m dz, \tag{2.16}$$

with the initial conditions

$$\mathbf{X}^m(0)=\mathbf{x}^m,\quad \mathbf{P}^m(0)=\mathbf{p}^m.$$

The d-dimensional Brownian motions $\mathbf{B}^m$, $m=1,\ldots,N$ have the standard covariance tensor

$$E\left\{B_i^m(z_1)B_j^n(z_2)\right\} = \delta_{mn}\delta_{ij}z_1\wedge z_2,\quad i,j=1,\ldots,d,\ m,n=1,\ldots,N.$$

The symmetric tensor $\sigma_{ij}(\mathbf{x})$ is determined from

$$\sum_{k=1}^N \sigma_{ik}(\mathbf{x})\sigma_{jk}(\mathbf{x}) = -\left(\frac{\partial^2 R_0(\mathbf{x})}{\partial x_i\partial x_j}\right). \tag{2.17}$$

We assume that equation (2.17) has a solution that is differentiable in $\mathbf{x}$, which is compatible with the fact that the matrix on the right is, by Bochner's theorem, non-negative definite.

The moments $W_\varepsilon^{(N)}$ converge as $\varepsilon\to 0$ to the solution of the advection-diffusion equation

$$\begin{aligned}\frac{\partial W^{(N)}}{\partial t} + \sum_{m=1}^N \frac{\mathbf{p}^m}{k}\cdot\nabla_{\mathbf{x}^m}W^{(N)} &= \frac{k^2D}{2}\sum_{m=1}^N \Delta_{\mathbf{p}^m}W^{(N)}\\ &-\frac{k^2}{4}\sum_{\substack{n,m=1\\ n>m}}^N\sum_{i,j=1}^d \frac{\partial^2 R_0((\mathbf{x}^n-\mathbf{x}^m)/\delta)}{\partial x_i\partial x_j}\frac{\partial^2 W^{(N)}}{\partial p_i^n\partial p_j^m}\end{aligned} \tag{2.18}$$

with the initial data

$$W^{(N)}(0,\mathbf{x}_1,,\ldots,\mathbf{x}^N,\mathbf{p}^1,\ldots,\mathbf{p}^N) = [W_I(\mathbf{x}^1,\mathbf{p}^1)]^{r_1}\cdot\ldots\cdot[W_I(\mathbf{x}^N,\mathbf{p}^N)]^{r_N}.$$

From (2.18) we can calculate moments of functionals of W_ε of the form

$$W_{\varepsilon,\phi}(z) = \int W_\varepsilon(z,\mathbf{x},\mathbf{p})\phi(\mathbf{x},\mathbf{p})d\mathbf{x}d\mathbf{p}.$$

For example, as $\varepsilon\to 0$ we have that

$$E\left\{[W_{\varepsilon,\phi}(z)]^2\right\}\to\int W^{(2)}(z,\mathbf{x}_1,\mathbf{p}_1,\mathbf{x}_2,\mathbf{p}_2)\phi(\mathbf{x}_1,\mathbf{p}_1)\phi(\mathbf{x}_2,\mathbf{p}_2)d\mathbf{x}_1d\mathbf{p}_1d\mathbf{x}_2d\mathbf{p}_2.$$

A convenient way to deal with not only the limit of N-point moments but with the full limit process $W(z,\mathbf{x},\mathbf{p})$, at all points $\mathbf{x},\mathbf{p}$ simultaneously, is provided by the theory of stochastic flows [23]. For this we need to show that $W_\varepsilon(z,\mathbf{x},\mathbf{p})$ converges weakly (in the probabilistic sense) as $\varepsilon \to 0$ to the process $W(z,\mathbf{x},\mathbf{p})$ that satisfies the stochastic partial differential equation

$$dW_\delta = \left[-\frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}}W_\delta + \frac{k^2D}{2}\Delta_{\mathbf{p}}W_\delta\right]dz - \frac{k}{2}\nabla_{\mathbf{p}}W_\delta\cdot d\mathbf{B}(\frac{\mathbf{x}}{\delta},z). \tag{2.19}$$

Here the Gaussian random field $\mathbf{B}(\mathbf{x},z)$ has the covariance

$$E\{B_i(\mathbf{x}_1,z_1)B_j(\mathbf{x}_2,z_2)\} = -\left(\frac{\partial^2 R_0((\mathbf{x}_1-\mathbf{x}_2))}{\partial x_i\partial x_j}\right)z_1\wedge z_2.$$

We call equation (2.19) the Liouville-Ito equation. It allows us to treat all equations of the form (2.18) simultaneously and is a convenient tool for simulation and analysis. The dimensionless wave number k can be scaled out of (2.19) by writing $W(z,\mathbf{x},\mathbf{p};k)=W(z,\mathbf{x},\frac{\mathbf{p}}{k};1)$ so that we need only consider (2.19) with $k=1$. We will use this scaling in Section 3.1.

Note that unlike the single Brownian motion (2.12) that governs the evolution of one-point moments, the Brownian field that enters the SPDE (2.19) depends explicitly on the dimensionless correlation length δ in the transverse direction. Therefore the limit process also depends on δ and we denote it by W_δ.

2.5. Statistical stability in the broad beam limit. We will now consider the limit $\delta\to 0$ of the process $W_\delta(z,\mathbf{x},\mathbf{p})$ when the transverse dimension $d\geq 2$. We are particularly interested in the behavior of functionals of W_δ as $\delta\to 0$. The analysis of one-point moments in Section 2.3 showed that they do not depend on δ and are governed by a standard Brownian motion. Therefore the process W_δ does not have a pointwise deterministic limit. However, we will show that functionals of W_δ become deterministic in the limit $\delta\to 0$. We refer to this phenomenon as **statistical stabilization** and give conditions for it to happen. Stabilization plays an important role in time reversal, imaging and other applications, as discussed in the Introduction.

THEOREM 2.1. *Assume that $\phi(\mathbf{p})$ is a smooth test function of rapid decay, the transverse correlation function $R_0(\mathbf{x})$ has compact support, the initial Wigner distribution $W_I(\mathbf{x},\mathbf{p})$ is uniformly bounded and Lipschitz continuous, and the transverse dimension $d\geq 2$. Define*

$$I_{\delta,\phi}(z,\mathbf{x}) = \int W_\delta(z,\mathbf{x},\mathbf{p})\phi(\mathbf{p})d\mathbf{p}. \tag{2.20}$$

Then

$$\lim_{\delta\to 0} E\left\{I^2_{\delta,\phi}(z,\mathbf{x})\right\} = E^2\left\{I_{\delta,\phi}(z,\mathbf{x})\right\} \tag{2.21}$$

where $E\left\{I_{\delta,\phi}(z,\mathbf{x})\right\}$ is independent of δ.

The assumption of compact support for $R_0(\mathbf{x})$ is not essential but simplifies the proof. We have already noted that the Wigner distribution W_δ itself does not stabilize. However, (2.21) implies that

$$\lim_{\delta\to 0} Var\left\{I_{\delta,\phi}\right\} = \lim_{\delta\to 0} E\left\{I^2_{\delta,\phi}(z)\right\} - E^2\left\{I_{\delta,\phi}\right\} = 0. \tag{2.22}$$

Therefore, any smooth functional of the form (2.20) stabilizes in the limit $\delta\to 0$, that is,

$$I_{\delta,\phi}\approx E\{I_{\delta,\phi}\}, \tag{2.23}$$

in mean square, and the expectation of $I_{\delta,\phi}$ does not depend on δ. We prove Theorem 2.1 in Appendix B.

In the applications of the asymptotic theory to time reversal we need not only functionals $I_{\delta,\phi}$ of the form (2.20) but also of the form

$$J_\delta(z,\mathbf{x}) = \int W_\delta(z,\mathbf{x},\mathbf{p})d\mathbf{p}. \tag{2.24}$$

We need to show that such functionals are well defined with probability one and to analyze their behavior as $\delta \to 0$. This is done in the following theorem.

THEOREM 2.2. *Under the same hypotheses of Theorem 2.1 and with a non-negative initial Wigner distribution $W_I \geq 0$, the functional J_δ is bounded, continuous and non-negative with probability one. In the limit $\delta \to 0$ we have*

$$\lim_{\delta\to 0} E\left\{J_\delta^2(z,\mathbf{x})\right\} = E^2\left\{J_\delta(z,\mathbf{x})\right\} \tag{2.25}$$

where $E\{J_\delta(z,\mathbf{x})\}$ does not depend on δ.

The proof of this theorem is given in Appendix B.

What is important in both Theorems 2.1 and 2.2 is that we do integrate over the wave numbers $\mathbf{p}$ because there is no pointwise stabilization. In time reversal applications, as in section 3.1, we actually need Theorem 2.2 when the integration is only over a line segment in $\mathbf{p}$ space, and the dimension of the latter is $d \geq 2$. Its proof follows from the one of Theorem 2.2.

3. Application to time reversal in a random medium. We will now apply these results to the time reversal problem [7] described in the Introduction. A wave emitted from the plane $z = 0$ propagates through the random medium and is recorded on the time reversal mirror at L. It is then *reversed* in time and re-emitted into the medium. The back-propagated signal refocuses approximately at the source, as shown in Figure 1.1. There are two striking features of this refocusing in random media. One is that it is statistically stable, that is, it does not depend on the particular realization. The other is super-resolution, that is, the refocused spot is tighter than in the deterministic case. We discuss these two issues in this section.

3.1. The time-reversed and back-propagated field. We assume that the wave source at $z = 0$ is distributed on a scale σ_s around a point $\mathbf{x}_0$, that is,

$$\psi_\theta(z=0,\mathbf{x};k) = e^{i\mathbf{p}_0\cdot(\mathbf{x}-\mathbf{x}_0)/\theta}\psi_0(\frac{\mathbf{x}-\mathbf{x}_0}{\sigma_s};k),$$

where ψ_0 is a rapidly decaying and smooth function of $\mathbf{x}$ and k. The width of the source σ_s could be large or small compared to the Fresnel number θ, and this affects the statistical stability of the time-reversed, back-propagated field, as we explain in this Section. The Green's function, $G_\theta(z,\mathbf{x};\xi)$, solves the parabolic wave equation (2.5) with a point source at $(\mathbf{x},z) = (\xi,0)$. Using its symmetry properties and the fact that time reversal $t \to -t$ is equivalent to $\omega \to -\omega$ or $k \to -k$, the back-propagated, time-reversed field on the plane of the source has the form

$$\begin{aligned}\psi_\theta^B&(L,\mathbf{x}_0,\xi;k) = \\ &\iint G_\theta(L,\mathbf{x};\mathbf{x}_0+\theta\xi;k)\overline{G_\theta(L,\mathbf{x}_0+\eta;\eta;k)}e^{i\mathbf{p}_0\cdot\eta/\theta}\psi_0(\frac{\eta}{\sigma_s};-k)\chi_A(\mathbf{x})d\mathbf{x}d\eta.\end{aligned} \tag{3.1}$$

The complex field amplitude ψ_θ^B is evaluated at $\mathbf{x}_0+\theta\xi$, in the plane $z=0$. We scale the observation point off $\mathbf{x}_0$ by θ because we expect that the spot size of the refocused signal will be comparable to the lateral spread of the initial wave function. We denote with χ_A the aperture function of the time reversal mirror. It could be its characteristic function, occupying the region A in the plane $z = L$

$$\chi_A(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in A \\ 0, & \mathbf{x} \notin A \end{cases},$$

or a more general aperture function like a Gaussian. The time reversal mirror is located in the plane $z = L$.

After changing variables, the back-propagated field is given by

$$\begin{aligned}\psi_\theta^B(L, \mathbf{x}_0, \xi; k) &= \theta^d \int G_\theta(L, \mathbf{x}; \mathbf{x}_0 + \theta\xi; k)\overline{G_\theta(L, \mathbf{x}; \mathbf{x}_0 + \theta\eta; k)} e^{i\mathbf{p}_0\cdot\eta}\psi_0(\frac{\theta\eta}{\sigma_s}; -k)\chi_A(\mathbf{x})d\mathbf{x}d\eta \\ &= \theta^d \int G_\theta(L, \mathbf{x}_0 + \theta\xi, \mathbf{x}; k)\overline{G_\theta(L, \mathbf{x}_0 + \theta\eta, \mathbf{x}; k)} e^{i\mathbf{p}_0\cdot\eta}\psi_0(\frac{\theta\eta}{\sigma_s}; -k)\chi_A(\mathbf{x})d\mathbf{x}d\eta.\end{aligned}$$

It is now convenient to introduce the Wigner distribution

$$W_\theta(z, \mathbf{x}_0, \mathbf{p}; k) = \int \frac{\theta^d e^{i\mathbf{p}\cdot\mathbf{y}}}{(2\pi)^d} G_\theta(z, \mathbf{x}_0 - \mathbf{y}\theta/2, \mathbf{x}; k)\overline{G_\theta(z, \mathbf{x}_0 + \mathbf{y}\theta/2, \mathbf{x}; k)}\chi_A(\mathbf{x})d\mathbf{x}d\mathbf{y}, \tag{3.2}$$

and express the back-propagated field as

$$\psi_\theta^B(L, \mathbf{x}_0, \xi; k) = \int e^{i\mathbf{p}\cdot(\xi-\eta)} W_\theta(L, \mathbf{x}_0 + \frac{\theta(\xi + \eta)}{2}, \mathbf{p}; k) e^{i\mathbf{p}_0\cdot\eta}\psi_0(\frac{\theta\eta}{\sigma_s}; -k)d\mathbf{p}d\eta. \tag{3.3}$$

The Wigner distribution is scaled here differently from (2.8) because of the way we have scaled the source function.

In the high frequency limit $\theta \to 0$, $W_\theta(z, \mathbf{x}, \mathbf{p}; k)$ tends to $W(z, \mathbf{x}, \mathbf{p}; k)$, which solves the random Liouville equation (2.11). Then, in the white noise limit, it solves the Liouville-Ito equation (2.19). The mean of W solves (2.14), in the high-frequency and white noise limit, with initial data

$$W(0, \mathbf{x}, \mathbf{p}; k) = \frac{\chi_A(\mathbf{x})}{(2\pi)^d}. \tag{3.4}$$

Let

$$\beta = \frac{\sigma_s}{\theta} \tag{3.5}$$

be the ratio of the width of the source to the Fresnel number and assume that it remains fixed as $\theta \to 0$. In this limit, the time-reversed and back-propagated field is given by

$$\begin{aligned}\psi^B(L, \mathbf{x}_0, \xi; k) &= \int e^{i\mathbf{p}\cdot(\xi-\eta)} W(L, \mathbf{x}_0, \frac{\mathbf{p}}{k}) e^{i\mathbf{p}_0\cdot\eta}\psi_0(\eta/\beta; -k)d\mathbf{p}d\eta \\ &= \int e^{i\mathbf{p}\cdot\xi} W(L, \mathbf{x}_0, \frac{\mathbf{p}}{k})\beta^d\hat{\psi}_0(\beta(\mathbf{p} - \mathbf{p}_0); -k)d\mathbf{p}.\end{aligned} \tag{3.6}$$

Here we have used the scaling $W(z, \mathbf{x}, \mathbf{p}; k) = W(z, \mathbf{x}, \frac{\mathbf{p}}{k}; 1)$ in (2.19) and we have dropped the last argument $k = 1$.

3.2. Statistical stability. From the form (3.6) of the back-propagated and time-reversed field we see that when $\beta = O(1)$ (or small), which means that σ_s is comparable to the Fresnel number θ (or smaller), we can apply the results of Section 2.5 and conclude that it is statistically stable or self-averaging in the broad beam limit $\delta \to 0$. Theorems 2.1 and 2.2 are exactly what is needed for this. The fact that the initial function (3.4) may be discontinuous at the boundary of the set A is not a problem. This is because, we may approximate the function χ_A from above and below by two smooth positive functions, to which we may apply Theorems 2.1 and 2.2, and then use the maximum principle to deduce the decorrelation property when the initial data is χ_A. We have, therefore,

$$\psi^B(L, \mathbf{x}_0, \xi; k) \approx \langle \psi^B(L, \mathbf{x}_0, \xi; k) \rangle$$

in the sense of convergence in probability or in mean square, in the broad beam limit $\delta \to 0$, for each fixed frequency $\omega = kc_0$. Statistical stability of time reversal does not depend on having a broad-band signal if the source is localized in space. This is true in the regime of parameters reflected by the scaling $\theta \ll \varepsilon \ll \delta$ considered here, which is a high frequency regime encountered in optical or infrared applications like ladar. The numerical experiments in [7] and [8] are closer to the regime of ultrasound experiments [16] and in underwater sound propagation, which is different from the high frequency regime analyzed here.

For distributed sources the parameter β is large and we cannot apply Theorems 2.1 and 2.2 to (3.6). It is necessary for statistical stability in this case to have broad-band signals. For β large the time reversed and back propagated signal in the time domain has the form

$$\begin{aligned}(3.7)\qquad &\psi^B(L, \mathbf{x}_0, \xi, t) \\ &= (2\pi)^d e^{i(\mathbf{p}_0\cdot\xi - k_0 c_0 t)} \psi_0(\xi/\beta) \int W(L, \mathbf{x}_0, \frac{\mathbf{p}_0}{k_0 + k}) e^{-ikc_0 t} \hat{g}(-c_0 k) \frac{c_0 dk}{2\pi} \\ &= (2\pi)^d e^{i(\mathbf{p}_0\cdot\xi - \omega_0 t)} \psi_0(\xi/\beta) \int W(L, \mathbf{x}_0, \frac{c_0 \mathbf{p}_0}{\omega_0 + \omega}) e^{-i\omega t} \hat{g}(-\omega) \frac{d\omega}{2\pi}\end{aligned}$$

with $\hat{g}(c_0 k)$ the Fourier transform of the initial pulse relative to the central frequency $\omega_0 = c_0 k_0$. This means that we have replaced the actual wave number k by $k_0 + k$, or ω by $\omega_0 + \omega$, with the new ω, the baseband frequency, bounded by Ω, the bandwidth, $|\omega| \le \Omega < \omega_0$. The integration is over the bandwidth $[-\Omega, \Omega]$. This integral is well defined with probability one and is self-averaging in the broad beam limit $\delta \to 0$ by Theorem 2.2 and the remark following it. We will compute its average in Section 3.4.

3.3. The effective aperture of the array. From the explicit expression for the Green's function of (2.14), with $k = 1$,

$$\begin{aligned}U(z, \mathbf{x}, \mathbf{p}; \mathbf{x}^0, \mathbf{p}^0) = \int \frac{d\mathbf{w} d\mathbf{r}}{(2\pi)^{2d}} \exp\left(i\mathbf{w}\cdot(\mathbf{x} - \mathbf{x}^0) + i\mathbf{r}\cdot(\mathbf{p} - \mathbf{p}^0) - iz\mathbf{w}\cdot\mathbf{p}^0\right) \\ \times \exp\left(-\frac{Dz}{2}\left[r^2 + z\mathbf{r}\cdot\mathbf{w} + \frac{\mathbf{w}^2 z^2}{3}\right]\right),\end{aligned}$$

and with the time reversal mirror a distance L from the source and $\mathbf{x}_0 = 0$, it follows from (3.6) that

$$\begin{aligned}(3.8)\quad &\langle \psi^B(z, \xi; k) \rangle = \\ &\int \frac{d\mathbf{p} d\mathbf{y} d\mathbf{w}}{(2\pi)^{2d}} e^{i\mathbf{p}\cdot\xi} \beta^d \psi_0(\beta(\mathbf{p} - \mathbf{p}_0); -k) \chi_A(\mathbf{y}) \exp\left[-i\mathbf{w}\cdot\mathbf{y} - iz\mathbf{w}\cdot\frac{\mathbf{p}}{k} - \frac{Dz^3 w^2}{6}\right]\end{aligned}$$

The high-frequency, white-noise limit of the *self-averaging* time-reversed and back-propagated field is therefore given by a convolution

$$\langle \psi^B(L,\xi;k)\rangle = \psi_0^\beta(\cdot,-k) * \mathcal{W}(\cdot)(\xi) \tag{3.9}$$

with

$$\mathcal{W}(\eta) = \mathcal{W}(\eta;L,k) = \frac{k^d}{(2\pi L)^d}\hat{\chi}_A(\eta k/L)\; e^{-\eta^2/(2\sigma_M^2)}, \tag{3.10}$$

the **point spread function**, and

$$\psi_0^\beta(\eta,-k) = e^{i\mathbf{p}_0\cdot\eta}\psi_0(\eta/\beta)\hat{g}(-kc_0) \tag{3.11}$$

with $\psi_0(\eta/\beta)$ the spatial source distribution function and $\hat{g}$ the Fourier transform of the pulse shape function $g(t)$. This notation is consistent with (3.7), with the time factor $e^{-ik_0c_0t}$ omitted, along with the horizontal phase e^{ikz} which cancels in time reversal. We have also introduced the refocused **spot size** with multipathing

$$\sigma_M^2 = \frac{3}{DLk^2} = \frac{L^2}{k^2a_e^2} \tag{3.12}$$

and the **effective aperture** $a_e = a_e(L)$,

$$a_e = \sqrt{\frac{DL^3}{3}}, \tag{3.13}$$

which we now interpret.

If the time reversal mirror is the whole plane $z = L$, then $\chi_A \equiv 1$ and

$$\langle \psi^B(L,\xi;k)\rangle = \psi_0^\beta(\xi,-k).$$

In this case the back-propagated field is the source field reversed in time, both in the random and in the deterministic case. The point spread function $\mathcal{W}$ determines the resolution of the refocused signal for a time reversal mirror of finite aperture. Multipathing in a random medium gives rise to the Gaussian factor (3.12) whose variance is σ_M^2. We can give an interpretation of this variance, or spot size, as follows. For a square time reversal mirror of size a, the Fourier transform of χ_A is the sinc function so that

$$\mathcal{W}(\eta_1,\eta_2;L,k) = \left(\frac{1}{\pi L}\right)^2 \sin(\frac{\eta_1 ka}{2L})\sin(\frac{\eta_2 ka}{2L})e^{-(\eta_1^2+\eta_2^2)/(2\sigma_M^2)}$$

For a deterministic medium ($D = 0$) the Rayleigh resolution is the distance η_F to the first zero of the sine, the first Fresnel zone in either direction,

$$\eta_F = \frac{2\pi L}{ka} = \frac{\lambda L}{a}.$$

In general, if χ_A is supported by a region of size a we may define the Fresnel resolution, or the Fresnel **spot size**, by

$$\sigma_F = \frac{L}{ka}.$$

For **weak multipathing** we have $\sigma_M \gg \sigma_F$ and

$$\mathcal{W}(\eta; L, k) \sim \left(\frac{k}{2\pi L}\right)^d \hat{\chi}_A(\eta k/L) \ ,$$

which is the diffractive point spread function whose integral over $\eta \in R^d$ is one. If, however, we have **strong multipathing**, $\sigma_M \ll \sigma_F$, then we may approximate $\hat{\chi}_A(\eta k/L)$ by $\hat{\chi}_A(0) = a^d$ in (3.10), and the point spread function becomes

$$\mathcal{W}(\eta; L, k) \sim \left(\frac{ka}{2\pi L}\right)^d e^{-|\eta|^2/(2\sigma_M^2)}.$$

By writing the variance (spot size) σ_M^2 in the form (3.12) we can interpret a_e as an effective aperture of the time reversal mirror. We can rewrite the point spread function in terms of a normalized Gaussian as

$$\mathcal{W}(\eta; L, k) \sim \left(\frac{\sigma_M}{\sqrt{2\pi}\sigma_F}\right)^d \frac{e^{-|\eta|^2/(2\sigma_M^2)}}{(2\pi\sigma_M^2)^{d/2}}$$

with the factor in front of the normalized Gaussian also equal to

$$\left(\frac{a}{\sqrt{2\pi}a_e}\right)^d.$$

This means that when there is strong multipathing the integral of the point spread function over R^d is not equal to one but to this ratio, which can be much smaller than one if $a_e \gg a$. Multipathing produces a tighter point spread function but there is also loss of energy, as of course we should expect.

A more direct interpretation for the effective aperture can be given if the time reversal mirror has a Gaussian aperture function

$$\chi_A(\eta) = e^{-|\eta|^2/(2a^2)}.$$

The point spread function $\mathcal{W}$ has now the form

$$\mathcal{W}(\eta; L, k) = \left(\frac{ka}{\sqrt{2\pi}L}\right)^d e^{-|\eta|^2/(2\sigma_g^2)},$$

with

$$\sigma_g = \frac{L}{ka_g}$$

and the effective aperture a_g given by

$$a_g = \sqrt{a^2 + \frac{DL^3}{3}} = \sqrt{a^2 + a_e^2}.$$

Clearly, $a_g \approx a_e$ when there is strong multipathing and $a_e \gg a$. Written with a normalized Gaussian the point spread function for a Gaussian aperture has the form

$$\mathcal{W}(\eta) = \left(\frac{a}{a_g}\right)^d \frac{e^{-|\eta|^2/(2\sigma_g^2)}}{(2\pi\sigma_g^2)^{d/2}}.$$

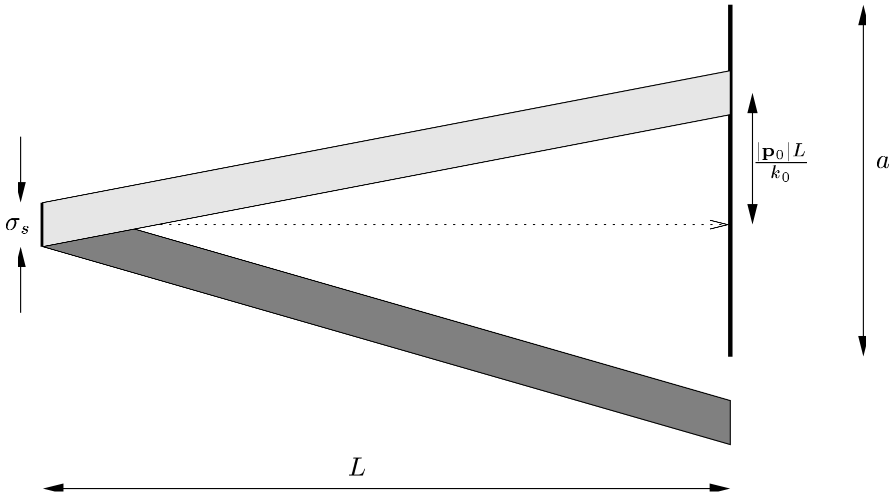

FIG. 3.1. *A directed field propagates from a distributed source of size σ_s toward the time reversal mirror of size a. The time-reversed, back-propagated field depends on the location of the mirror relative to the direction of the propagating beam.*

3.4. Broad-band time reversal for distributed sources. For a distributed source, its support σ_s is large compared to the Fresnel number θ so the ratio $\beta = \sigma_s/\theta$ is large. In this case we can compute the average of (3.7) the same way as we did in (3.8) and we find that

$$\begin{aligned}(3.14)\quad & \langle \Psi^B(L,\mathbf{x}_0,\xi,t)\rangle \\ &= (2\pi)^d e^{i(\mathbf{p}_0\cdot\xi - k_0c_0t)}\psi_0(\xi/\beta)\int \langle W(L,\mathbf{x}_0,\frac{\mathbf{p}_0}{k_0+k})\rangle e^{-ikc_0t}\hat{g}(-c_0k)\frac{c_0dk}{2\pi} \\ &= e^{i(\mathbf{p}_0\cdot\xi-k_0c_0t)}\psi_0(\xi/\beta)\int \frac{d\mathbf{y}d\mathbf{w}c_0dk}{(2\pi)^{d+1}}\chi_A(\mathbf{y})e^{i(\frac{L\mathbf{w}\mathbf{p}_0}{k_0+k}-\mathbf{w}\cdot\mathbf{y}-kc_0t)}e^{-\frac{DL^3w^2}{6}}\hat{g}(-c_0k).\end{aligned}$$

The $\mathbf{y}$ integral on the right gives the Fourier transform of the aperture function $\chi_A(\mathbf{y})$ so with $\omega_0 = c_0k_0$ and a change of variable from k to $\omega = c_0k$ we have

$$\begin{aligned}(3.15)\quad & \langle \Psi^B(L,\mathbf{x}_0,\xi,t)\rangle \\ &= e^{i(\mathbf{p}_0\cdot\xi-\omega_0t)}\psi_0(\xi/\beta)\int \frac{d\omega}{2\pi}e^{-i\omega t}\hat{g}(-\omega)\ \chi_A * \left(\frac{e^{-x^2/(2a_e^2)}}{(2\pi a_e^2)^{d/2}}\right)(\frac{Lc_0\mathbf{p}_0}{\omega_0+\omega})\end{aligned}$$

Here the star denotes convolution with respect to the spatial variables $\mathbf{x}$, and a_e is the effective aperture defined by (3.13).

When multipathing is weak we can ignore the Gaussian factor in the convolution and we have

$$\begin{aligned}(3.16)\quad & \langle \Psi^B(L,\mathbf{x}_0,\xi,t)\rangle \\ &= e^{i(\mathbf{p}_0\cdot\xi-\omega_0t)}\psi_0(\xi/\beta)\int \frac{d\omega}{2\pi}e^{-i\omega t}\hat{g}(-\omega)\ \chi_A\left(\frac{Lc_0\mathbf{p}_0}{\omega_0+\omega}\right)\end{aligned}$$

In the opposite case, when there is strong multipathing and the effective aperture is much larger than the physical one, $a_e \gg a$, we have

$$(3.17)\quad \langle \Psi^B(L,\mathbf{x}_0,\xi,t)\rangle$$

$$= e^{i(\mathbf{p}_0\cdot\xi - \omega_0 t)}\psi_0(\xi/\beta)\left(\frac{a}{\sqrt{2\pi}a_e}\right)^d \int \frac{d\omega}{2\pi} e^{-i\omega t}\hat{g}(-\omega)e^{-\frac{1}{2}(\frac{Lc_0\mathbf{p}_0}{a_e(\omega_0+\omega)})^2}$$

To interpret these results we note first that a distributed source function of the form (3.11) can be considered as a phased array emitting an inhomogeneous plane wave, a beam, in the direction $(k, \mathbf{p}_0)$, within the paraxial or parabolic approximation. The ratio $|\mathbf{p}_0|/k$ is the tangent of the angle the direction vector makes with the z axis, and $L|\mathbf{p}_0|/k$ is the transverse distance of the beam center to the center of the phased array (see Figure 3.1). If for each ω the beam displacement vector $Lc_0\mathbf{p}_0/(\omega_0+\omega)$ is inside the set A occupied by the time reversal array, then we recover at the source the full pulse in (3.16), time-reversed,

$$\langle\Psi^B(L, \mathbf{x}_0, \xi, t)\rangle = e^{i(\mathbf{p}_0\cdot\xi-\omega_0 t)}\psi_0(\xi/\beta)g(-t).$$

If, however, for some frequencies the transverse displacement vector is outside the time reversal array, these frequencies will be nulled in the integration and a distorted time pulse will be received at the source. Depending on the position of the time reversal mirror relative to the beam, high or low frequencies may be nulled.

In a strongly multipathing medium the situation is quite different because the expression (3.17), or more generally (3.15), holds now. Even if the beam from the phased array does not intercept the time reversal mirror at all, we will still get a time reversed signal at the source but with a much diminished amplitude. If the beam falls entirely within the time reversal mirror then the time reversed pulse will be a distorted form of $g(-t)$, with its amplitude reduced by the factor $(a/a_e)^d$. An interesting and important application of the time reversal of a beam in a random medium is the possibility of **estimating** the effective aperture a_e by pointing the beam in different directions toward the time reversal mirror, measuring the time reversed signal that back propagates to the source, that is, to the phased array, and inferring a_e by fitting the measurements to (3.15).

4. Summary and conclusions. We have analyzed and explained two important phenomena associated with time reversal in a random medium:

- Super-resolution of the back-propagated signal due to multipathing
- Self-averaging that gives a statistically stable refocusing

Our analysis is based on a specific asymptotic limit (see Section 2.1) where the longitudinal distance of propagation is much larger than the size of the time-reversal mirror, which in turn is much larger than the correlation length of the medium, fluctuations in the index of refraction are weak, and the wave length is short compared to the correlation length. This asymptotic regime is more relevant to optical or infrared time reversal than it is to sonar or ultrasound. We have related the self-averaging properties of the back-propagated signal to those of functionals of the Wigner distribution. Self-averaging of these functionals implies the statistical stability of the time-reversed and back-propagated signal in the frequency domain, provided that the source function is not too broad compared to the Fresnel number (2.4). Time reversal refocusing of waves emitted from a distributed source is self-averaging only in the time domain.

We apply our theoretical results about stochastic Wigner distributions to time reversal and discuss in detail super-resolution and statistical stability in section 3.

Appendix A. The white noise limit and the parabolic approximation.

We collect here some comments on the scaling analysis of section 2.1 and refer to [1, 27, 33] for additional comments and results on scaling and asymptotics in the high-frequency and white-noise regime.

The dimensionless parameters $\delta, \varepsilon, \gamma$ introduced by (2.2) in Section 2.1, along with the Fresnel number θ defined by (2.4), lead to the scaled parabolic wave equation (2.5). If we do not make the parabolic approximation and keep the ψ_{zz} term we have the scaled Helmholtz equation, with the phase e^{ikz} removed,

$$\frac{\varepsilon^2\theta^2}{\delta^2}\psi_{zz} + 2ik\theta\psi_z + \theta^2\Delta_{\mathbf{x}}\psi + \frac{k^2\delta}{\varepsilon^{1/2}}\mu(\frac{\mathbf{x}}{\delta}, \frac{z}{\varepsilon})\psi = 0. \tag{A.1}$$

Here, as in (2.5), we relate the strength of the fluctuations σ to ε and δ by (2.6). Is the parabolic approximation valid in the ordering (2.7), $\theta \ll \varepsilon \ll \delta \ll 1$, that we have analyzed? The answer is yes but not before both θ **and** ε limits have been taken, in which case the scaled Wigner distribution (2.8) converges to the Liouville-Ito process that is defined by the stochastic partial differential equation (2.19).

It is in the white noise limit $\varepsilon \to 0$, with Fresnel number θ and δ fixed, that the parabolic approximation is valid for (A.1), as was pointed out in [1]. This is easily seen if the random fluctuations μ are differentiable in z. The parabolic approximation is clearly not valid in the high frequency limit $\theta \to 0$, before the white noise limit $\varepsilon \to 0$ is also taken. In the white noise limit, the wave function $\psi(z, \mathbf{x})$ satisfies an Ito-Schrödinger equation

$$2ik\theta d_z\psi + \theta^2\Delta_{\mathbf{x}}\psi dz + \frac{ik^3\delta^2}{4\theta}R_0(0)\psi dz + k^2\delta\psi d_z B(\frac{\mathbf{x}}{\delta}, z) = 0. \tag{A.2}$$

Here R_0 is the integrated covariance of the fluctuations μ given by (2.15) and (2.13), and the Brownian field $B(\mathbf{x}, z)$ has covariance

$$\langle B(\mathbf{x}, z_1)B(\mathbf{y}, z_2)\rangle = R_0(\mathbf{x}-\mathbf{y})z_1 \wedge z_2.$$

This Ito-Schrödinger equation is the result of the central limit theorem applied to (A.1). Let

$$B^\varepsilon(\mathbf{x}, z) = \frac{1}{\sqrt{\varepsilon}}\int_0^z \mu(\mathbf{x}, \frac{s}{\varepsilon})ds.$$

Then, as $\varepsilon \to 0$ this process converges weakly, under suitable hypotheses, to the Brownian field $B(\mathbf{x}, z)$ with the above covariance. The extra term in (A.2) is the Stratonovich correction.

The white noise limit for stochastic partial differential equations is analyzed in [10] and a rigrous theory of the Ito-Schrödinger equation is given in [11]. The ergodic theory of the Ito-Schroedinger equation is explored in [17]. Wave propagation in the parabolic approximation with white-noise fluctuations is considered in detail in [18, 30].

The scaled Wigner distribution for the process ψ, defined by (2.8), satisfies the stochastic transport equation

$$\begin{aligned} &d_z W_\theta(z, \mathbf{x}, \mathbf{p}) + \frac{\mathbf{p}}{k}\cdot\nabla_x W_\theta(z, \mathbf{x}, \mathbf{p})dz \\ &= \frac{k^2\delta^2}{4\theta^2}\int\frac{d\mathbf{q}}{(2\pi)^d}\hat{R}_0(\mathbf{q})\left(W_\theta(z, \mathbf{x}, \mathbf{p}+\frac{\theta\mathbf{q}}{\delta}) - W_\theta(z, \mathbf{x}, \mathbf{p})\right)dz \\ &+\frac{ik\delta}{2\theta}\int\frac{d\mathbf{q}}{(2\pi)^d}e^{i\mathbf{q}\cdot\mathbf{x}/\delta}\left(W_\theta(z, \mathbf{x}, \mathbf{p}-\frac{\theta\mathbf{q}}{2\delta}) - W_\theta(z, \mathbf{x}, \mathbf{p}+\frac{\theta\mathbf{q}}{2\delta})\right)d_z\hat{B}(\mathbf{q}, z), \end{aligned} \tag{A.3}$$

which is derived from (A.2) equation using the Ito calculus. The Wigner process W_θ converges in the limit $\theta \to 0$ to the Liouville-Ito process defined by the stochastic partial differential equation (2.19).

Appendix B. Decorrelation of the Wigner process.

B.1. Proof of Theorem 2.1. We give here the proof of Theorems 2.1 and 2.2. We consider Theorem 2.1 first. It will follow from the Lebesgue dominated convergence theorem if we show that for $\mathbf{p}_1 \neq \mathbf{p}_2$:

$$E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)W_\delta(z,\mathbf{x},\mathbf{p}_2)\right\} - E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)\right\}E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_2)\right\} \to 0 \tag{B.1}$$

as $\delta \to 0$ because the function W_δ is uniformly bounded and $E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)\right\}$ does not depend on δ. Furthermore, the correlation function at the same spatial point but for two different values of the wave vector, $U_\delta^{(2)}(z,\mathbf{x},\mathbf{p}_1,\mathbf{p}_2) = E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)W_\delta(z,\mathbf{x},\mathbf{p}_2)\right\}$ is the solution of (2.18) with $N=2$ and the initial data

$$W_\delta^{(2)}(0,\mathbf{x}_1,\mathbf{p}_1,\mathbf{x}_2,\mathbf{p}_2) = W_I(\mathbf{x}_1,\mathbf{p}_1)W_I(\mathbf{x}_2,\mathbf{p}_2),$$

evaluated at $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$. Therefore $U_\delta^{(2)}$ may be represented as

$$U_\delta^{(2)}(z,\mathbf{x}_1,\mathbf{p}_1,\mathbf{x}_2,\mathbf{p}_2) = E\left\{W_I(\mathbf{X}_\delta^1(z),\mathbf{P}_\delta^1(z))W_I(\mathbf{X}_\delta^2(z),\mathbf{P}_\delta^2(z))\right\}.$$

The processes $\mathbf{X}_\delta^{1,2}$ and $\mathbf{P}_\delta^{1,2}$ satisfy the system of SDE's (2.16) which may be more explicitly written as

$$\begin{aligned} d\mathbf{P}_\delta^1 &= -\left[\sigma(0)d\mathbf{B}^1(z) + \frac{1}{2}\sigma\left(\frac{\mathbf{X}_\delta^1 - \mathbf{X}_\delta^2}{\delta}\right)d\mathbf{B}^2(z)\right] \\ d\mathbf{P}_\delta^2 &= -\left[\sigma(0)d\mathbf{B}^2(z) + \frac{1}{2}\sigma\left(\frac{\mathbf{X}_\delta^2 - \mathbf{X}_\delta^1}{\delta}\right)d\mathbf{B}^1(z)\right] \\ d\mathbf{X}_\delta^1 &= -\mathbf{P}_\delta^1 dz, \quad d\mathbf{X}_\delta^2 = -\mathbf{P}_\delta^2 dz \end{aligned} \tag{B.2}$$

with the initial conditions $\mathbf{X}_\delta^{1,2}(0) = \mathbf{x}$, $\mathbf{P}_\delta^m(0) = \mathbf{p}_m$, $m=1,2$. Here $\sigma^2(0) = D$, the diffusion coefficient (2.15) and the coupling matrix $\sigma(\mathbf{x})$ is given by (2.17). Recall that $W_\delta(z,\mathbf{x},\mathbf{p},k) = W_\delta(z,\mathbf{x},\mathbf{p}/k;1)$ and we need only consider the case $k=1$.

It is convenient to introduce the processes $\mathbf{X}^{1,2}$ and $\mathbf{P}^{1,2}$ that are solutions of (2.16) with no coupling:

$$\begin{aligned} &d\mathbf{P}^m = -\sigma(0)d\mathbf{B}^m(z), \quad d\mathbf{X}^m = -\mathbf{P}^m dz, \\ &\mathbf{X}^{1,2}(0) = \mathbf{x}, \quad \mathbf{P}^m(0) = \mathbf{p}_m, \quad m=1,2 \end{aligned} \tag{B.3}$$

and define the deviations of the solutions of the coupled system of SDE's (B.2) from those of (B.3): $\mathbf{Z}_\delta^m = \mathbf{X}_\delta^m - \mathbf{X}^m$, $\mathbf{S}_\delta^m = \mathbf{P}_\delta^m - \mathbf{P}^m$. Then we have

$$\begin{aligned} &d\mathbf{S}_\delta^1 = -\frac{1}{2}\sigma\left(\frac{\mathbf{X}_\delta^1 - \mathbf{X}_\delta^2}{\delta}\right)d\mathbf{B}^2(z), \quad d\mathbf{S}_\delta^2 = -\frac{1}{2}\sigma\left(\frac{\mathbf{X}_\delta^2 - \mathbf{X}_\delta^1}{\delta}\right)d\mathbf{B}^1(z) \\ &d\mathbf{Z}_\delta^1 = -\mathbf{S}_\delta^1 dz, \quad d\mathbf{Z}_\delta^2 = -\mathbf{S}_\delta^2 dz \end{aligned} \tag{B.4}$$

with the initial data $\mathbf{S}_\delta^m(0) = \mathbf{Z}^m(0) = 0$. Define

$$\begin{aligned} &\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}_\delta^1,\mathbf{Z}_\delta^2,\mathbf{S}_\delta^1,\mathbf{S}_\delta^2) \\ &= W_I(\mathbf{X}^1+\mathbf{Z}_\delta^1,\mathbf{P}^1+\mathbf{S}_\delta^1)W_I(\mathbf{X}^2+\mathbf{Z}_\delta^2,\mathbf{P}^2+\mathbf{S}_\delta^2) - W_I(\mathbf{X}^1,\mathbf{P}^1)W_I(\mathbf{X}^2,\mathbf{P}^2) \end{aligned} \tag{B.5}$$

then we have with the above notation

$$\begin{aligned} &E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)W_\delta(z,\mathbf{x},\mathbf{p}_2)\right\} - E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_1)\right\}E\left\{W_\delta(z,\mathbf{x},\mathbf{p}_2)\right\} \\ &= E\left\{\mathcal{V}(\mathbf{X}^1(z),\mathbf{X}^2(z),\mathbf{P}^1(z),\mathbf{P}^2(z),\mathbf{Z}_\delta^1(z),\mathbf{Z}_\delta^2(z),\mathbf{S}_\delta^1(z),\mathbf{S}_\delta^2(z))\right\} \\ &\le CE\left\{|\mathbf{Z}_\delta^1(z)| + |\mathbf{Z}_\delta^2(z)| + |\mathbf{S}_\delta^1(z)| + |\mathbf{S}_\delta^2(z)|\right\} \end{aligned} \tag{B.6}$$

since W_I is a Lipschitz function.

Let us assume for simplicity that the correlation function $R(\mathbf{x})$ has compact support inside the set $|\mathbf{x}| \leq M$. Then the coupling term in (B.2) is non-zero only when $|\mathbf{X}_\delta^1 - \mathbf{X}_\delta^2| \leq M\delta$. We introduce the processes $\mathbf{Q}_\delta = \mathbf{P}_\delta^1 - \mathbf{P}_\delta^2$ and $\mathbf{Y}_\delta = \mathbf{X}_\delta^1 - \mathbf{X}_\delta^2$ that govern (B.4). They satisfy the SDE's

$$
\begin{aligned}
&d\mathbf{Q}_\delta = -\left[\sigma(0) - \frac{1}{2}\sigma\left(\frac{\mathbf{Y}_\delta}{\delta}\right)\right] d\tilde{\mathbf{B}}, \quad d\mathbf{Y}_\delta = -\mathbf{Q}_\delta dz, \\
&\mathbf{Q}_\delta(0) = \mathbf{p}_1 - \mathbf{p}_2, \;\; \mathbf{Y}_\delta(0) = 0
\end{aligned} \tag{B.7}
$$

with $\tilde{\mathbf{B}} = \mathbf{B}^1 - \mathbf{B}^2$ being a Brownian motion.

In order to prove the theorem we show that the coupling term $\sigma(\cdot)$ in (B.2) introduces only lower order correction terms, that is, $\mathbf{S}_\delta^m$ and $\mathbf{Z}_\delta^m$ are small. We show first that after a small 'time', τ, the points $\mathbf{X}_\delta^m$ are driven apart since $\mathbf{Q}_\delta(0) = \mathbf{P}_\delta^1(0) - \mathbf{P}_\delta^2(0) \neq 0$. Then we show that after the points have separated the probability that they come close, so that the coupling term $\sigma(\cdot)$ becomes non-zero, is small. This "non-recurrence" condition requires that the spatial dimension $d \geq 2$. It follows that to leading order the points $\mathbf{X}_\delta^m$ are uncorrelated when $d \geq 2$ and that the coupling term introduces only lower order corrections. A similar argument for $d = 1$ would require an estimate on the time that points that are originally separated in the spatial variable spend near each other, where the coupling term in (B.2) is not zero.

We need the following two Lemmas. The first one shows that particles that start at the same point $\mathbf{x}$ with different initial directions $\mathbf{p}_1$ and $\mathbf{p}_2$, get separated with a large probability:

LEMMA B.1. *Let $\mathbf{Y}_\delta$, $\mathbf{Q}_\delta$ solve (B.7) with $\mathbf{Y}_\delta(0) = 0$, $\mathbf{Q}_\delta(0) = \mathbf{q} \neq 0$. Then for any $\varepsilon > 0$ there exists $\tau_0(\varepsilon) > 0$ that depends only on $\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2$ but not on δ so that we have $P\left(|\mathbf{Y}_\delta(\tau)| \geq \dfrac{|\mathbf{q}|\tau}{2}\right) \geq 1 - \varepsilon$ for all $\tau \leq \tau_0(\varepsilon)$.*

The second lemma shows that after the particles are separated, the probability that they come close to each other is small:

LEMMA B.2. *Given any fixed $r > 0$ and $z > 0$, if $\mathbf{Y}_\delta$, $\mathbf{Q}_\delta$ solve (B.7) with $|\mathbf{Y}_\delta(0)| \geq r$, $\mathbf{Q}_\delta(0) = \mathbf{q} \neq 0$, then $P\left(\inf_{0 \leq s \leq z} |\mathbf{Y}_\delta(s)| \leq M\delta\right) \to 0$ as $\delta \to 0$.*

We prove Theorem 2.1 before proving Lemmas B.1 and B.2:

Proof. Let z and $\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2$ be fixed and defined as above. Given $\varepsilon > 0$, then for any $\tau < \tau_0(\varepsilon)$ (with τ_0 as defined in Lemma B.1), Lemma B.2 and the Markov property of the Brownian motion imply that

$$
P\left(\mathbf{S}_\delta^m(z) = \mathbf{S}_\delta^m(\tau) \middle| |\mathbf{Y}_\delta(\tau)| \geq \frac{\tau|\mathbf{q}|}{2}\right) \geq 1 - \varepsilon
$$

and

$$
P\left(\mathbf{Z}_\delta^m(z) = \mathbf{Z}_\delta^m(\tau) + (z-\tau)\mathbf{S}_\delta^m(\tau) \middle| |\mathbf{Y}_\delta(\tau)| \geq \frac{\tau|\mathbf{q}|}{2}\right) \geq 1 - \varepsilon
$$

for $\delta < \delta_0(\tau, \varepsilon)$. Furthermore,

$$
\begin{aligned}
&E\left\{|\mathbf{Z}_\delta^1(\tau)| + |\mathbf{Z}_\delta^2(\tau)| + |\mathbf{S}_\delta^1(\tau)| + |\mathbf{S}_\delta^2(\tau)| \middle| |\mathbf{Y}_\delta(\tau)| \geq \frac{\tau|\mathbf{q}|}{2}\right\} \\
&\leq E\left\{|\mathbf{Z}_\delta^1(\tau)| + |\mathbf{Z}_\delta^2(\tau)| + |\mathbf{S}_\delta^1(\tau)| + |\mathbf{S}_\delta^2(\tau)|\right\}/(1-\varepsilon) \leq C\tau
\end{aligned} \tag{B.8}
$$

because the function σ is uniformly bounded. Therefore we have

$$\begin{aligned}
&E\left\{\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}^1_\delta,\mathbf{Z}^2_\delta,\mathbf{S}^1_\delta,\mathbf{S}^2_\delta)\right\}\\
&= E\left\{\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}^1_\delta,\mathbf{Z}^2_\delta,\mathbf{S}^1_\delta,\mathbf{S}^2_\delta)\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right\}P\left(|\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right)\\
&+E\left\{\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}^1_\delta,\mathbf{Z}^2_\delta,\mathbf{S}^1_\delta,\mathbf{S}^2_\delta)\Big||\mathbf{Y}_\delta(\tau)|\leq\frac{\tau|\mathbf{q}|}{2}\right\}P\left(|\mathbf{Y}_\delta(\tau)|\leq\frac{\tau|\mathbf{q}|}{2}\right)=I+II.
\end{aligned}\tag{B.9}$$

The second term above is small because the probability for $\mathbf{Y}_\delta(\tau)$ to be very small is bounded by Lemma B.1. More precisely, given $\varepsilon>0$ and $\tau<\tau_0(\varepsilon)$, Lemma B.1 implies that

$$II\leq C\varepsilon.\tag{B.10}$$

The first term in (B.9) corresponds to the more likely scenario that $\mathbf{Y}_\delta$ at time τ has left the ball of radius $\tau|\mathbf{q}|/2$. We estimate it as follows. The probability that $\mathbf{Y}_\delta$ re-enters the ball of radius $M\delta$ is small according to Lemma B.2. Moreover, if $\mathbf{Y}_\delta$ stays outside this ball, the difference variables $\mathbf{Z}^m$ and $\mathbf{S}^m$ are bounded in terms of their values at time τ. The latter are small if τ is small. More precisely, using (B.8) we choose τ so small that

$$E\left\{|\mathbf{Z}^1_\delta(\tau)|+|\mathbf{Z}^2_\delta(\tau)|+|\mathbf{S}^1_\delta(\tau)|+|\mathbf{S}^2_\delta(\tau)|\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right\}\leq\varepsilon.$$

Then we obtain

$$\begin{aligned}
I&\leq E\left\{\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}^1_\delta,\mathbf{Z}^2_\delta,\mathbf{S}^1_\delta,\mathbf{S}^2_\delta)\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right\}\\
&\leq E\left\{\mathcal{V}(\mathbf{X}^1,\mathbf{X}^2,\mathbf{P}^1,\mathbf{P}^2,\mathbf{Z}^1_\delta,\mathbf{Z}^2_\delta,\mathbf{S}^1_\delta,\mathbf{S}^2_\delta)\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\text{ and }\inf_{\tau\leq s\leq z}|\mathbf{Y}_\delta(s)|\leq M\delta\right\}\\
&\times P\left(\inf_{\tau\leq s\leq z}|\mathbf{Y}_\delta(s)|\leq M\delta\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right)\\
&+E\left\{|\mathbf{Z}^1_\delta(z)|+|\mathbf{Z}^2_\delta(z)|+|\mathbf{S}^1_\delta(z)|+|\mathbf{S}^2_\delta(z)|\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\text{ and }\inf_{\tau\leq s\leq z}|\mathbf{Y}_\delta(s)|\geq M\delta\right\}\\
&\times P\left(\inf_{\tau\leq s\leq z}|\mathbf{Y}_\delta(s)|\geq M\delta\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right)=I_1+I_2.
\end{aligned}$$

The term I_1 goes to zero as $\delta\to0$ by Lemma B.2. However, if the conditions in I_2 hold, then

$$\mathbf{S}^m_\delta(z)=\mathbf{S}^m_\delta(\tau),\quad\mathbf{Z}^m_\delta(z)=\mathbf{Z}^m_\delta(\tau)-\frac{1}{k}(z-\tau)\mathbf{S}^m_\delta(\tau).$$

Therefore the term I_2 may be bounded with the help of (B.8) by

$$\begin{aligned}
I_2&\leq E\left\{|\mathbf{Z}^1_\delta(z)|+|\mathbf{Z}^2_\delta(z)|+|\mathbf{S}^1_\delta(z)|+|\mathbf{S}^2_\delta(z)|\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\text{ and }\inf_{\tau\leq s\leq z}|\mathbf{Y}_\delta(s)|\geq M\delta\right\}\\
&\leq CE\left\{|\mathbf{Z}^1_\delta(\tau)|+|\mathbf{Z}^2_\delta(\tau)|+|\mathbf{S}^1_\delta(\tau)|+|\mathbf{S}^2_\delta(\tau)|\Big||\mathbf{Y}_\delta(\tau)|\geq\frac{\tau|\mathbf{q}|}{2}\right\}\leq C\tau.
\end{aligned}$$

Putting together (B.9), (B.10) and the above bounds on I_1 and I_2, we obtain

$$E\left\{|\mathbf{Z}^1_\delta(z)|+|\mathbf{Z}^2_\delta(z)|+|\mathbf{S}^1_\delta(z)|+|\mathbf{S}^2_\delta(z)|\right\}\leq C\varepsilon$$

for $\delta<\bar{\delta}$ and Theorem 2.1 follows from (B.6). □

B.2. Proof of Lemmas B.1 and B.2. We first prove Lemma B.1.

Proof. We write

$$Q_\delta(z) = \mathbf{q} - \int_0^z \left(\sigma(0) - \frac{1}{2}\sigma(\mathbf{Y}(s)/\delta)\right)\ d\tilde{\mathbf{B}}(s) \ \equiv \ \mathbf{q} + \tilde{\mathbf{Q}}_\delta(z)$$

so that

$$\mathbf{Y}_\delta(t) = -\mathbf{q}t - \int_0^t \tilde{\mathbf{Q}}_\delta(s)ds.$$

Then we have

$$P\left(\sup_{0\le s\le\tau} |\tilde{\mathbf{Q}}_\delta(s)| \ > \ r\right) \le C\tau/r^2 \tag{B.11}$$

and hence

$$P\left(|\mathbf{Y}_\delta(\tau) + \tau\mathbf{q}| > r\tau\right) \le P\left(\sup_{0\le s\le\tau} |\tilde{\mathbf{Q}}_\delta(s)| > r\right) \le C\tau/r^2.$$

We let $r = |\mathbf{q}|/2$ in the above formula and obtain

$$P\left(|\mathbf{Y}_\delta(\tau)| < \frac{\tau|\mathbf{q}|}{2}\right) \le \frac{C}{|\mathbf{q}|^2}\tau,$$

and the conclusion of Lemma B.1 follows. □

Finally, we prove Lemma B.2.

Proof. Let τ_δ be the first time $\mathbf{Y}_\delta(z)$ enters the ball of radius $M\delta$:

$$\tau_\delta = \inf\left\{z : \ |\mathbf{Y}_\delta(z)| \le M\delta\right\},$$

with $\mathbf{Y}_\delta(0) = \mathbf{Y}^0 \ne 0$. For $0 < \alpha < 1$ let $\Delta z = \delta^{1-\alpha}$, $n = \lceil z/\Delta z\rceil$, $J_i = (i\Delta z, (i+1)\Delta z)$ and $p < 1$. Note that until the time τ_δ the process $(\mathbf{Y}_\delta, \mathbf{Q}_\delta)$ coincides with the process $(\mathbf{Y}, \mathbf{Q})$ governed by (B.7) without the coupling term $\sigma(\mathbf{Y}_\delta/\delta)$. We find

$$P(\tau_\delta < z) \le \sum_{i=0}^{n-1}\left\{P\left(|\mathbf{Y}(i\Delta z)| < M\delta^p\right) + P\left(\inf_{s\in J_i}|\mathbf{Y}(s)| < M\delta \,\Big|\, |\mathbf{Y}(i\Delta z)| \ge M\delta^p\right)\right\}.$$

The process $\mathbf{Y}(s)$ is Gaussian with mean $\mathbf{Y}^0$ and variance $\mathcal{O}(s^2)$. Therefore, there is a $\bar{\delta} > 0$ such that for $\delta < \bar{\delta}$

$$P(|\mathbf{Y}(i\Delta z)| < M\delta^p) \le C\delta^{dp}.$$

If we assume

$$p < 1 - \alpha \tag{B.12}$$

then also

$$\begin{aligned}
P(\tau_\delta < z) &\le nC\left(\delta^{dp} + P\left(\sup_{0<s<\Delta z}|\mathbf{Y}(s) - \mathbf{Y}^0| \ge M[\delta^p - \delta]\right)\right)\\
&\le C(\delta^{dp+\alpha-1} + \delta^{\alpha-1}P\left(\sup_{0<s<\Delta z}|\mathbf{B}(s)| \ge M[\delta^p - \delta]/\Delta z\right))\\
&\le C(\delta^{dp+\alpha-1} + \delta^{\alpha-1}\frac{E\left\{\mathbf{B}(\Delta z)^{2r}\right\}\Delta z^{2r}}{(\delta^p - \delta)^{2r}})\\
&\le C\left[\delta^{dp+\alpha-1} + \delta^{\alpha-1-rp+3r(1-\alpha)/2}\right].
\end{aligned}$$

Note that with $p < 1-\alpha$ and r large enough, there is a $q > 0$ so that

$$P(\tau_\delta < z) \le C\delta^q$$

if $d \ge 2$ and Lemma B.2 follows. □

B.3. Proof of Theorem 2.2. We need to show first that

$$J_\delta(z,\mathbf{x}) = \int W_\delta(z,\mathbf{x},\mathbf{p})d\mathbf{p} \tag{B.13}$$

is finite with probability one. The stochastic flow $(\mathbf{X}_\delta(t,\mathbf{x},\mathbf{p}), \mathbf{P}_\delta(t,\mathbf{x},\mathbf{p})$ is continuous in $(t,\mathbf{x},\mathbf{p})$ with probability one, so $W_\delta(z,\mathbf{x},\mathbf{p}) = W_I(\mathbf{X}_\delta(t,\mathbf{x},\mathbf{p}), \mathbf{P}_\delta(t,\mathbf{x},\mathbf{p}))$ is bounded and continuous. It is, moreover, non-negative if $W_I \ge 0$. We know that

$$\int E\{W_\delta(z,\mathbf{x},\mathbf{p})\}d\mathbf{p}$$

is finite and independent of δ, and the order of integration and expectation can be interchanged by Tonelli's theorem. This theorem implies in addition that $J_\delta(z,\mathbf{x})$ is finite with probability one.

We can now consider

$$E\{J_\delta^2(z,\mathbf{x})\} = \int E\{W_\delta(z,\mathbf{x},\mathbf{p}_1)W_\delta(z,\mathbf{x},\mathbf{p}_2)\}d\mathbf{p}_1 d\mathbf{p}_2.$$

The integrand is bounded by an integrable function uniformly in δ because

$$E\{W_\delta(z,\mathbf{x},\mathbf{p}_1)W_\delta(z,\mathbf{x},\mathbf{p}_2)\} \le E^{1/2}\{W_\delta^2(z,\mathbf{x},\mathbf{p}_1)\}E^{1/2}\{W_\delta(z,\mathbf{x},\mathbf{p}_2)\},$$

the right side does not depend on δ, and is integrable. Therefore by the Lebesgue dominated convergence theorem and the results of the previous Section we have that

$$\lim_{\delta\to 0} E\{J_\delta^2(z,\mathbf{x})\} = E^2\{J_\delta(z,\mathbf{x})\}$$

and the right side does not depend on δ. This completes the proof of Theorem 2.2.

REFERENCES

[1] F. Bailly, J.F. Clouet and J.P. Fouque, Parabolic and white noise approximation for waves in random media, SIAM Journal on Applied Mathematics **56**, 1996, 1445-1470.

[2] G.Bal and L. Ryzhik, Time reversal for classical waves in random media, Comptes rendus de l'Académie des sciences - Série I - Mathématique, **333**, 2001, 1041-1046.

[3] G.Bal and L. Ryzhik, Time reversal for waves in random media, Preprint, 2002.

[4] G.Bal, G. Papanicolaou and L. Ryzhik, Self-averaging in time reversal for the parabolic wave equation Preprint, 2002.

[5] G.Bal, G. Papanicolaou and L. Ryzhik, Radiative transport limit for the random Schrödinger equation, Nonlinearity, **15**, 2002, 513-529.

[6] G. Blankenship and G. C. Papanicolaou, Stability and Control of Stochastic Systems with Wide-Band Noise Disturbances, SIAM J. Appl. Math., **34**, 1978, 437-476.

[7] P. Blomgren, G. Papanicolaou, and H. Zhao, Super-Resolution in Time-Reversal Acoustics, J. Acoust. Soc. Am., **111**, 2002, 230-248.

[8] L. Borcea, C. Tsogka, G. Papanicolaou and J. Berryman, Imaging and time reversal in random media, to appear in Inverse Problems, 2002.

[9] J. Berryman, L. Borcea, G. Papanicolaou and C. Tsogka, Statistically stable ultrasonic imaging in random media, Preprint, 2002.

[10] R. Bouc and E. Pardoux, Asymptotic analysis of PDEs with wide-band noise disturbances and expansion of the moments, Stochastic Analysis and Applications, **2**, 1984, 369-422.

[11] D. Dawson and G. Papanicolaou, A random wave process, Appl. Math. Optim., **12**, 1984, 97–114.

[12] D. Dowling and D. Jackson, Phase conjugation in underwater acoustics, Jour. Acoust. Soc. Am., **89**, 1990, 171-181

[13] D. Dowling and D. Jackson, Narrow-band performance of phase-conjugate arrays in dynamic random media, Jour. Acoust. Soc. Am.,**91**, 1992, 3257-3277.

[14] M. Fink and J. de Rosny, Time-reversed acoustics in random media and in chaotic cavities, Nonlinearity, **15**, 2002, R1-R18.

[15] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.L. Thomas and F. Wu, Time-reversed acoustics, Rep. Progr. Phys., **63**, 2000, 1933-1995.

[16] M. Fink and C. Prada, Acoustic time-reversal mirrors, Inverse Problems, **17**, 2001, R1–R38.

[17] J.P. Fouque, G.C. Papanicolaou and Y. Samuelides, Forward and Markov Approximation: The Strong Intensity Fluctuations Regime Revisited, Waves in Random Media, **8**, 1998, 303-314.

[18] K. Furutsu, Random Media and Boundaries: Unified Theory, Two-Scale Method, and Applications, Springer Verlag, 1993.

[19] P.Gérard, P.Markovich, N.Mauser and F.Poupaud, Homogenization limits and Wigner transforms, Comm.Pure Appl. Math., **50**, 1997, 323-380.

[20] W. Hodgkiss, H. Song, W. Kuperman, T. Akal, C. Ferla and D. Jackson, A long-range and variable focus phase-conjugation experiment in a shallow water, Jour. Acoust. Soc. Am., **105**, 1999, 1597-1604.

[21] H. Kesten and G. Papanicolaou, A Limit Theorem for Turbulent Diffusion, Comm. Math. Phys., **65**, 1979, 97-128.

[22] G. Papanicolaou and W. Kohler, Asymptotic analysis of deterministic and stochastic equations with rapidly varying components, Comm. Math. Phys. **45**, 217–232, 1975.

[23] H. Kunita, *Stochastic flows and stochastic differential equations.* Cambridge Studies in Advanced Mathematics, 24. Cambridge University Press, Cambridge, 1997.

[24] W. Kuperman, W. Hodgkiss, H. Song, T. Akal, C. Ferla and D. Jackson, Phase-conjugation in the ocean, Jour. Acoust. Soc. Am., **102**, 1997, 1-16.

[25] W. Kuperman, W. Hodgkiss, H. Song, T. Akal, C. Ferla and D. Jackson, Phase conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror, J. Acoust. Soc. Am., **103**, 1998, 25-40.

[26] H. Kushner, Approximation and weak convergence methods for random processes, with applications to stochastic systems theory, MIT Press Series in Signal Processing, Optimization, and Control, MIT Press, 1984.

[27] B. Nair and B. White, High-frequency wave propagation in random media- a unified approach, SIAM J. Appl. Math., **51**, 1991, 374-411.

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

[29] F. Tappert, The parabolic approximation method, Lecture notes in physics, vol. 70, *Wave propagation and underwater acoustics*, Springer-Verlag, 1977.

[30] V. I. Tatarskii, A. Ishimaru and V. U. Zavorotny, editors, Wave Propagation in Random Media (Scintillation), SPIE and IOP, 1993.

[31] Thomas J. L. and M. Fink, *Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: Application to transskull therapy*, IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 43, 1122-1129, (1996).

[32] C. Tsogka and G. Papanicolaou, Time reversal through a solid-liquid interface and super-resolution, Preprint, 2001.

[33] B. White, The stochastic caustic, SIAM Jour. Appl. Math., **44**, 1984, 127-149.