Self-averaging in time reversal for the parabolic wave equation

Guillaume Bal [*] George Papanicolaou [†] Leonid Ryzhik [‡]

October 29, 2002

Abstract

We analyze the self-averaging properties of time-reversed solutions of the paraxial wave equation with random coefficients, which we take to be Markovian in the direction of propagation. This allows us to construct an approximate martingale for the phase space Wigner transform of two wave fields. Using a priori L^2-bounds available in the time-reversal setting, we prove that the Wigner transform in the high frequency limit converges in probability to its deterministic limit, which is the solution of a transport equation.

1 Introduction

In time-reversal experiments a signal emitted by a localized source is recorded at an array of transducers. It is then re-emitted into the medium reversed in time, that is, the part of the signal that is recorded first is sent back last. Because of the time-reversibility of the wave equation the back-propagated signal refocuses approximately at the location of the original source since the array is limited in size. A striking experimental observation is that the presence of inhomogeneities in the medium improves the refocusing resolution. The explanation for this super-resolution is multipathing: waves in complex media that are captured by the recording array have undergone multiple scattering making it effectively larger than its physical size. Another important feature of super-resolution in time reversal is that the refocused signal does not depend on the realization of the random medium. That is, the refocused signal is deterministic. Super-resolution and self-averaging of refocused signals in complicated media has been observed both in laboratory experiments (see reviews [14, 16] and references therein) and in underwater acoustic wave propagation over long distances (tens of kilometers) [12, 19]. Time-reversal techniques have numerous applications ranging from medicine to communications and, more recently, imaging in random media [7, 10, 17].

The first mathematical analysis of time-reversal in random media was given by Clouet and Fouque [11], who analyzed refocusing and self-averaging of time-reversed pulses in a one-dimensional layered random medium. Super-resolution in spatial refocusing and its statistical stability for multi-dimensional waves in random media was analyzed in [9, 21], in a remote-sensing regime where the paraxial or parabolic wave equation can be used. The refocusing of the average signal in a full three-dimensional medium, in the regimes of random geometrical optics and radiative transfer (transport), was studied in [2, 3]. We also mention that another source of multipathing is the mixing of waves by the boundaries in an ergodic cavity. This has been studied experimentally in [15] and mathematically in [6].

[*]Department of Applied Physics and Applied Mathematics, Columbia University, New York NY, 10027; gb2030@columbia.edu

[†]Department of Mathematics, Stanford CA, 94305; papanico@math.stanford.edu

[‡]Department of Mathematics, University of Chicago, Chicago IL, 60637; ryzhik@math.uchicago.edu

The purpose of this paper is to analyze time reversal in the radiative transfer regime using the parabolic wave equation, when the waves interact fully with the random inhomogeneities. We prove mathematically that the refocused signal is self-averaging, which means that it does not depend on the realization of the random medium. The mathematical quantity that we analyze is the Wigner measure of a pair of oscillatory solutions of the random Schrödinger equation. In the present setting, the random potential depends in a Markovian way on the variable z, the main direction of propagation of the waves. This allowed us to use in [4] a martingale method to prove that the average of the Wigner distribution converges to a solution of the radiative transfer equation. In this paper we use additional regularity of the Wigner measure, available in time-reversal when there is some blurring at the recording array, to show that the whole Wigner distribution, and not only its average, converges weakly, as a Schwartz distribution and in probability, to the deterministic solution of the transport equation. The blurring at the recording array provides a priori bounds for the Wigner transform in L^2. These bounds and the Markovianity of the random potential in the direction of propagation make the time-reversal problem more tractable mathematically and allow us to prove in a fairly simple and straightforward manner self-averaging of the time-reversed signal.

We recall that the Wigner transform is a convenient tool to analyze high frequency wave propagation in deterministic [18, 20, 27] and random media [24]. Introduced by Wigner in 1932 [28], it has been used extensively in the mathematical literature recently. Convergence of the average Wigner distribution to the solution of the radiative transfer equation was first proved by H. Spohn in [25] for time-independent potentials on small time intervals. This result was extended to global in time convergence by L. Erdös and H.-T. Yau [13]. These proofs involve infinite Neumann (diagrammatic) expansions for the solution of the Schrödinger equation and are quite involved technically. The corresponding problem with time-dependent potentials is much simpler mathematically. It was treated by us in [4] in the Markovian case, and by F. Poupaud and A. Vasseur [22] in the case of finite-range time correlations. In this paper we use the fact that the Wigner family arising in time-reversal problems is more regular than the usual one because blurring is added at the recording array. This provides some additional regularity usually obtained by considering mixtures of states as, for instance, in [20, 22, 25].

The paper is organized as follows: we describe the scaling and obtain an expression for the back-propagated signal in terms of the Wigner transform in Section 2. The main result and assumptions on the random medium are formulated in Section 3. The proofs are presented in Section 4.

2 The back-propagated signal in the parabolic approximation

2.1 The back-propagated signal

The pressure field $p(z, \mathbf{x}, t)$ satisfies the scalar wave equation

$$\frac{1}{c^2(z,\mathbf{x})}\frac{\partial^2 p}{\partial t^2} - \Delta p = 0. \tag{1}$$

Here $c(z, \mathbf{x})$ is the local wave speed that we will assume to be random, and the Laplacian operator includes both direction of propagation, z, and the transverse variable $\mathbf{x} \in \mathbb{R}^d$. In the physical setting, we have $d = 2$. We consider dimensions $d \geq 1$ to stress that our analysis is independent of the dimension. If we assume that at time $t = 0$, the wave field has a “beam-like” structure in the z direction, and if back-scattering may be neglected, we can replace the wave equation by its parabolic approximation [26]. More precisely, the pressure p may be approximated as

$$p(z, \mathbf{x}, t) \approx \int_{\mathbb{R}} e^{i(-c_0\kappa t+\kappa z)}\psi(z, \mathbf{x}, \kappa)c_0\, d\kappa, \tag{2}$$

where ψ satisfies the Schrödinger equation

$$\begin{aligned}&2i\kappa\frac{\partial\psi}{\partial z}(z,\mathbf{x},\kappa)+\Delta_{\mathbf{x}}\psi(z,\mathbf{x},\kappa)+\kappa^2(n^2(z,\mathbf{x})-1)\psi(z,\mathbf{x},\kappa)=0,\\&\psi(z=0,\mathbf{x},\kappa)=\psi_0(\mathbf{x},\kappa)\end{aligned}\tag{3}$$

with $\Delta_{\mathbf{x}}$ the transverse Laplacian in the variable $\mathbf{x}$. The refraction index $n(z,\mathbf{x})=c_0/c(z,\mathbf{x})$, and c_0 in (2) is a reference speed. The rigorous passage to the parabolic approximation from the wave equation has been analyzed in [5] in the deterministic case, and [1] in a one dimensional random medium. A formal derivation of the paraxial approximation that leads to the radiative transfer regime is given below in section 2.2.

We assume that the original source is located in the plane $z=0$, and the time-reversal mirror is located in the plane $z=L$ as depicted in Figure 1.

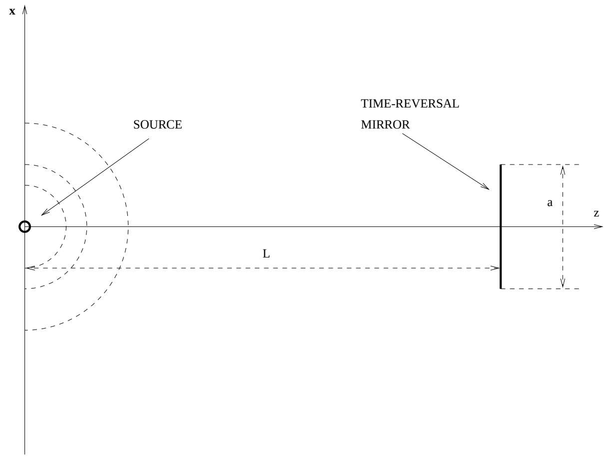

Figure 1: Geometry of the time-reversal experiment.

During the first stage of a time reversal experiment the signal propagates for a time T, or equivalently, over a distance $L=c_0T$, so that the signal arriving at the time-reversal mirror is given by

$$\psi_-(L,\mathbf{x})=\int_{\mathbb{R}^d}G(L,\mathbf{x},\kappa;\mathbf{y})\psi_0(\mathbf{y})d\mathbf{y},$$

where the Green's function solves

$$\begin{aligned}&2i\kappa\frac{\partial G}{\partial z}(z,\mathbf{x},\kappa;\mathbf{y})+\Delta_{\mathbf{x}}G(z,\mathbf{x},\kappa;\mathbf{y})+\kappa^2(n^2(z,\mathbf{x})-1)G(z,\mathbf{x},\kappa;\mathbf{y})=0,\\&G(0,\mathbf{x},\kappa;\mathbf{y})=\delta(\mathbf{x}-\mathbf{y}).\end{aligned}\tag{4}$$

Then the signal is time reversed. For three-dimensional acoustic pulses, this means that the pressure field is kept unchanged and that the sign of its time derivative is reversed. In the parabolic approximation, this is equivalent to phase conjugation, where ψ is replaced by its complex conjugate ψ^*. We assume that the recording array occupies a compact subset of the plane $z=L$, and introduce a real-valued aperture function $\chi(\mathbf{x})$. It represents the restriction of the signal onto the array, and possible amplification by the array that may vary from one receiver to another. In the absence of amplification it is given by the characteristic function of an array set $\Omega\subset\mathbb{R}^d$. We also allow for some blurring of the recorded signal, modeled by a convolution with kernel $f(\mathbf{x})$. The signal at L

after time reversal takes then the form

$$\begin{aligned}\psi_+(L,\mathbf{x},\kappa) &= \int_{\mathbb{R}^d} \chi(\mathbf{x}) f(\mathbf{x}-\mathbf{y}) \psi_-^*(L,\mathbf{y};\kappa)\chi(\mathbf{y})d\mathbf{y} \\ &= \int_{\mathbb{R}^{2d}} \chi(\mathbf{x}) G^*(L,\mathbf{y},\kappa;\mathbf{y}')\chi(\mathbf{y}) f(\mathbf{x}-\mathbf{y})\psi_0^*(\mathbf{y}',\kappa)d\mathbf{y}d\mathbf{y}'.\end{aligned}$$

The last step consists in letting the signal propagate back to the origin $z=0$ and time reversing it one more time

$$\psi^B(\mathbf{x},\kappa) = \int_{\mathbb{R}^{3d}} G^*(L,\mathbf{x},\kappa;\boldsymbol{\eta}) G(L,\mathbf{y},\kappa;\mathbf{y}')\chi(\boldsymbol{\eta})\chi(\mathbf{y}) f(\boldsymbol{\eta}-\mathbf{y})\psi_0(\mathbf{y}',\kappa)d\mathbf{y}d\mathbf{y}'d\boldsymbol{\eta}. \tag{5}$$

The last step, phase-conjugating at the source, is not performed in real physical experiments but is convenient if we need to compare the back-propagated signal to the original signal. It affects neither the degree of refocusing nor the self-averaging effect. The back-propagated signal in time is given at the plane $z=0$ by

$$\begin{aligned}\psi^B(\mathbf{x},t) &= \int_{\mathbb{R}} e^{-ic_0\kappa t}\psi^B(\mathbf{x},\kappa)d\kappa \\ &= \int_{\mathbb{R}^{3d+1}} G^*(L,\mathbf{x},\kappa;\boldsymbol{\eta}) e^{-ic_0\kappa t} G(L,\mathbf{y},\kappa;\mathbf{y}')\chi(\boldsymbol{\eta})\chi(\mathbf{y}) f(\boldsymbol{\eta}-\mathbf{y})\psi_0(\mathbf{y}',\kappa)d\mathbf{y}d\mathbf{y}'d\boldsymbol{\eta}d\kappa.\end{aligned} \tag{6}$$

We will be interested in the sequel in the refocusing and self-averaging properties of the back-propagated signal $\psi^B(\mathbf{x},\kappa)$ for each fixed frequency. Self-averaging of the time signal $\psi^B(\mathbf{x},t)$ then follows from (6) when $\psi_0(\mathbf{y},\kappa)$ has compact support in κ. An interesting conclusion of this paper is that the additional blurring by convolution with the kernel f makes the signal $\psi^B(\mathbf{x},\kappa)$ self-averaging for every frequency κ. This should be contrasted with the situation studied in [9], where no blurring was introduced and self-averaging was observed only for the full time signal.

2.2 Localized Source and Scaling

We recall the formal passage from the reduced wave equation to the parabolic equation (3) described in the Appendix to [4] and explain how the radiative transfer scaling arises in this context. We start with the reduced wave equation

$$\Delta\hat{p} + \kappa^2 n^2(z,\mathbf{x})\hat{p} = 0, \tag{7}$$

and look for solutions of (7) in the form $\hat{p}(z,\mathbf{x}) = e^{i\kappa z}\psi(z,\mathbf{x})$. Then we obtain

$$\frac{\partial^2\psi}{\partial z^2} + 2i\kappa\frac{\partial\psi}{\partial z} + \Delta_{\mathbf{x}}\psi + \kappa^2(n^2-1)\psi = 0. \tag{8}$$

The refraction index $n(z,\mathbf{x})$ is weakly fluctuating so that it has the form

$$n^2(z,\mathbf{x}) = 1 - 2\sigma V\left(\frac{z}{l_z},\frac{\mathbf{x}}{l_x}\right),$$

where V models random fluctuations that will be described in detail in section 3.3. Here, l_x and l_z are the correlation lengths of V in the transverse and longitudinal directions, respectively, and the small parameter σ measures the strength of the fluctuations. The waves propagate over a distance L_x in the $\mathbf{x}$-plane and L_z in the z-direction, and we rescale $\mathbf{x}$ and z accordingly. We also introduce a carrier wave number κ_0 and replace $\kappa = \kappa_0\kappa'$, κ' being the non-dimensional wave number (we drop the prime below). The physical parameters determined by the medium are the length scales l_x, l_z and the non-dimensional parameter $\sigma \ll 1$.

We now explain the scaling of the parameters L_x, L_z and κ_0 in the radiative transfer regime. Equation (8) in the non-dimensional variables $z' = z/L_z$, $\mathbf{x}' = \mathbf{x}/L_x$ becomes after we drop the primes

$$\frac{1}{L_z^2}\frac{\partial^2\psi}{\partial z^2} + \frac{2i\kappa\kappa_0}{L_z}\frac{\partial\psi}{\partial z} + \frac{1}{L_x^2}\Delta_{\mathbf{x}}\psi - 2\kappa^2\kappa_0^2\sigma V\left(\frac{zL_z}{l_z}, \frac{\mathbf{x}L_x}{l_x}\right)\psi = 0. \tag{9}$$

We introduce the following small parameters:

$$\delta_x = \frac{l_x}{L_x}, \quad \delta_z = \frac{l_z}{L_z}, \quad \gamma_x = \frac{1}{\kappa_0 l_x}, \quad \gamma_z = \frac{1}{\kappa_0 l_z}$$

and rewrite (9) as

$$\gamma_z\delta_z\frac{\partial^2\psi}{\partial z^2} + 2i\kappa\frac{\partial\psi}{\partial z} + \frac{\delta_x^2\gamma_x^2}{\delta_z\gamma_z}\Delta_{\mathbf{x}}\psi - \frac{2\kappa^2\sigma}{\gamma_z\delta_z}V\left(\frac{z}{\delta_z}, \frac{\mathbf{x}}{\delta_x}\right)\psi = 0. \tag{10}$$

Let us assume now that

$$\delta_x = \delta_z \ll 1, \quad \gamma_z = \gamma_x^2 \ll 1, \quad \sigma = \gamma_z\sqrt{\delta_x} \tag{11}$$

and denote $\varepsilon = \delta_x$. Then (10) becomes after multiplication by $\varepsilon/2$

$$\frac{\gamma_z\varepsilon^2}{2}\frac{\partial^2\psi}{\partial z^2} + i\kappa\varepsilon\frac{\partial\psi}{\partial z} + \frac{\varepsilon^2}{2}\Delta_{\mathbf{x}}\psi - \kappa^2\sqrt{\varepsilon}V\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right)\psi = 0. \tag{12}$$

Observe that when $\kappa = O(1)$ and $\gamma_z \ll 1$, the first term in (12) is small and may be neglected in the leading order since $|\varepsilon^2\psi_{zz}| = O(1)$. Then (12) becomes

$$i\kappa\varepsilon\frac{\partial\psi}{\partial z} + \frac{\varepsilon^2}{2}\Delta_{\mathbf{x}}\psi - \kappa^2\sqrt{\varepsilon}V\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right)\psi = 0 \tag{13}$$

which is the parabolic wave equation (3) in the radiative transfer scaling.

The second condition in (11) implies that the carrier wave number should be chosen as $\kappa_0 = l_z/l_x^2$. Then $\gamma_z = l_x^2/l_z^2 \ll 1$ implies that $l_x \ll l_z$. Therefore the correlation length in the longitudinal direction z should be much longer than in the transverse plane $\mathbf{x}$. Furthermore, we should have $\varepsilon = (\sigma/\gamma_z)^2 = \sigma^2 l_z^4/l_x^4 \ll 1$ which in turn implies that the fluctuation strength $\sigma \ll (l_x/l_z)^4$. Given that these constraints are satisfied the last condition in (11) implies that the spatial scales L_x and L_z should be chosen according to

$$L_x = l_x\frac{l_x^4}{\sigma^2 l_z^4}, \qquad L_z = l_z\frac{l_x^4}{\sigma^2 l_z^4}.$$

Note that the constraint $l_x \ll l_z$ implies that $L_x \ll L_z$, which is the usual constraint for the validity of the parabolic approximation. The connection between the relations of the physical parameters and the resulting scaling of the parabolic approximation is further discussed in [21]. The rigorous passage to the parabolic approximation in a three-dimensional layered random medium in a similar scaling is discussed in [1].

To define what we mean by the quality of the refocused signal, we assume that the initial pulse is centered around a point $\mathbf{x}_0$ with support comparable to the transverse correlation length. Then the initial data for the Schrödinger equation (13) becomes in the rescaled variables

$$\psi(z = 0, \mathbf{x}, \kappa) = \psi_0(\frac{\mathbf{x} - \mathbf{x}_0}{\varepsilon}, \kappa). \tag{14}$$

We also replace $f(\mathbf{x})$ by $\varepsilon^{-d} f(\mathbf{x}/\varepsilon)$. Finally, we are interested in the refocusing properties of $\psi^B(\mathbf{x})$ in the vicinity of $\mathbf{x}_0$ and consequently introduce the scaling $\mathbf{x} = \mathbf{x}_0 + \varepsilon\boldsymbol{\xi}$. We then recast (5) as

$$\psi_\varepsilon^B(\boldsymbol{\xi}, \kappa; \mathbf{x}_0) = \int_{\mathbb{R}^{3d}} G^*(L, \mathbf{x}_0 + \varepsilon\boldsymbol{\xi}, \kappa; \boldsymbol{\eta}) G(L, \mathbf{y}, \kappa; \mathbf{x}_0 + \varepsilon\mathbf{y}') \chi(\boldsymbol{\eta}, \mathbf{y}) \psi_0(\mathbf{y}', \kappa) d\mathbf{y}' d\mathbf{y} d\boldsymbol{\eta}, \tag{15}$$

where

$$\chi(\boldsymbol{\eta}, \mathbf{y}) = \chi(\boldsymbol{\eta})\chi(\mathbf{y}) f(\frac{\boldsymbol{\eta} - \mathbf{y}}{\varepsilon}). \tag{16}$$

We observe that $G(L, \mathbf{x}, \kappa; \mathbf{y}) = G(L, \mathbf{y}, \kappa; \mathbf{x})$, so that

$$\psi_\varepsilon^B(\boldsymbol{\xi}, \kappa; \mathbf{x}_0) = \int_{\mathbb{R}^{3d}} G^*(L, \mathbf{x}_0 + \varepsilon\boldsymbol{\xi}, \kappa; \boldsymbol{\eta}) G(L, \mathbf{x}_0 + \varepsilon\mathbf{y}', \kappa; \mathbf{y}) \chi(\boldsymbol{\eta}, \mathbf{y}) \psi_0(\mathbf{y}', \kappa) d\mathbf{y}' d\mathbf{y} d\boldsymbol{\eta}. \tag{17}$$

Let us now introduce the auxiliary function Q_ε as in [2, 3],

$$Q_\varepsilon(L, \mathbf{x}, \kappa; \mathbf{q}) = \int_{\mathbb{R}^d} G(L, \mathbf{x}, \kappa; \mathbf{y}) \chi(\mathbf{y}) e^{-i\mathbf{q}\cdot\mathbf{y}/\varepsilon} d\mathbf{y}, \tag{18}$$

which satisfies the initial value problem

$$\begin{aligned} & i\varepsilon\kappa \frac{\partial Q_\varepsilon}{\partial z}(z, \mathbf{x}; \kappa) + \frac{\varepsilon^2}{2} \Delta_{\mathbf{x}} Q_\varepsilon(z, \mathbf{x}; \kappa) - \kappa^2 \sqrt{\varepsilon} V\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right) Q_\varepsilon(z, \mathbf{x}, \kappa) = 0, \\ & Q_\varepsilon(z = 0, \mathbf{x}; \kappa) = \chi(\mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}/\varepsilon}. \end{aligned} \tag{19}$$

It physically describes the component with wave vector $\mathbf{q}$ that is sent back to the plane $z = 0$ by the array. We then define the Wigner transform as

$$W_\varepsilon(L, \mathbf{x}, \mathbf{k}, \kappa) = \int_{\mathbb{R}^d} \hat{f}(\mathbf{q}) U_\varepsilon(L, \mathbf{x}, \mathbf{k}, \kappa; \mathbf{q}) d\mathbf{q}, \tag{20}$$

where

$$U_\varepsilon(L, \mathbf{x}, \mathbf{k}, \kappa; \mathbf{q}) = \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{y}} Q_\varepsilon(L, \mathbf{x} - \frac{\varepsilon\mathbf{y}}{2}, \kappa; \mathbf{q}) Q_\varepsilon^*(L, \mathbf{x} + \frac{\varepsilon\mathbf{y}}{2}, \kappa; \mathbf{q}) \frac{d\mathbf{y}}{(2\pi)^d}. \tag{21}$$

The Wigner transform W_ε is more regular than U_ε because of additional averaging in the wave vector $\mathbf{q}$. More precisely, while the family U_ε is not uniformly bounded in L^2, the family W_ε has a uniform bound in L^2. This regularizing effect is essentially the same as the one obtained by considering mixtures of states for the Schrödinger equation as in [20, 22, 25].

We can then recast (17) as

$$\psi_\varepsilon^B(\boldsymbol{\xi}, \kappa; \mathbf{x}_0) = \int_{\mathbb{R}^{2d}} e^{i\mathbf{k}\cdot(\boldsymbol{\xi} - \mathbf{y})} W_\varepsilon(L, \mathbf{x}_0 + \varepsilon\frac{\mathbf{y} + \boldsymbol{\xi}}{2}, \mathbf{k}, \kappa) \psi_0(\mathbf{y}, \kappa) \frac{d\mathbf{y} d\mathbf{k}}{(2\pi)^d}. \tag{22}$$

The above formula shows that the asymptotic behavior of $\psi_\varepsilon^B(\boldsymbol{\xi}, \kappa; \mathbf{x}_0)$ as $\varepsilon \to 0$ is characterized by that of the Wigner transform $W_\varepsilon(L, \mathbf{x}, \mathbf{k}, \kappa)$.

3 The main results and assumptions on the random medium

3.1 The main results

This section presents our two main results. The first one describes the self-averaging of the back-propagated signal ψ^B in (22). Its proof is based on the second theorem, of independent interest, which shows that the Wigner transform converges in probability to a unique deterministic limit, solution of a transport equation.

Theorem 3.1 *Let the array function $\chi(\mathbf{y})$ and the radially symmetric filter $f(\mathbf{y})$ be in $L^1\cap L^\infty(\mathbb{R}^d)$, while $\psi_0 \in L^2(\mathbb{R}^d)$ for a given $\kappa \in \mathbb{R}$. Assume also that the refraction index $n(z,\mathbf{x})$ satisfies assumptions in section 3.3 below. Then for each $\boldsymbol{\xi} \in \mathbb{R}^d$ the back-propagated signal $\psi_\varepsilon^B(\boldsymbol{\xi},\mathbf{x}_0,\kappa)$ given by (22) converges in probability and weakly in $L^2_{\mathbf{x}_0}(\mathbb{R}^d)$ as $\varepsilon \to 0$ to the deterministic function*

$$\psi^B(\boldsymbol{\xi},\kappa;\mathbf{x}_0) = \int_{\mathbb{R}^{2d}} e^{i\mathbf{k}\cdot(\boldsymbol{\xi}-\mathbf{y})}\overline{W}(L,\mathbf{x}_0,\mathbf{k},\kappa)\psi_0(\mathbf{y},\kappa)\frac{d\mathbf{y}d\mathbf{k}}{(2\pi)^d}. \tag{23}$$

The function $\overline{W}$ satisfies the transport equation

$$\frac{\partial\overline{W}}{\partial z} + \frac{1}{\kappa}\mathbf{k}\cdot\nabla_{\mathbf{x}}\overline{W} = \kappa\mathcal{L}\overline{W}, \tag{24}$$

with initial data $\overline{W}_0(\mathbf{x},\mathbf{k}) = \hat{f}(\mathbf{k})|\chi(\mathbf{x})|^2$ and where the operator $\mathcal{L}$ is defined by

$$\mathcal{L}\lambda = \int_{\mathbb{R}^d}\frac{d\mathbf{p}}{(2\pi)^d}\hat{R}(\frac{|\mathbf{p}|^2-|\mathbf{k}|^2}{2},\mathbf{p}-\mathbf{k})(\lambda(\mathbf{p})-\lambda(\mathbf{k})). \tag{25}$$

Here $\hat{R}(\omega,\mathbf{p})$ is the Fourier transform of the correlation function of V, defined by (32) below.

The convergence of the Wigner transform is described by the following theorem.

Theorem 3.2 *Under the assumptions of Theorem 3.1 the Wigner distribution W_ε converges in probability and weakly in $L^2(\mathbb{R}^{2d})$ to the solution $\overline{W}$ of the transport equation (24). More precisely, for any test function $\lambda \in L^2(\mathbb{R}^{2d})$ the process $\langle W_\varepsilon(z),\lambda\rangle$ converges to $\langle\overline{W}(z),\lambda\rangle$ in probability as $\varepsilon \to 0$, uniformly on finite intervals $0 \le z \le L$.*

Here, $\langle\cdot,\cdot\rangle$ is the usual scalar product in $L^2(\mathbb{R}^{2d})$. Theorem 3.1 is then a corollary of Theorem 3.2 and Lemma 4.2 below.

Theorem 3.1 provides a mathematical explanation for the important properties of the time-reversal experiment. First, it ensures that the back-propagated signal is essentially independent of the realization of the random medium in the high-frequency limit since $\psi^B(\boldsymbol{\xi},\kappa;\mathbf{x}_0)$ is deterministic. Second, it provides a quantitative description of the back-propagated field, which may be written as

$$\psi^B(\cdot,\kappa;\mathbf{x}_0) = F(\cdot,\kappa;\mathbf{x}_0)\star\psi_0(\cdot,\kappa).$$

The kernel $F(\boldsymbol{\xi},\kappa;\mathbf{x}_0)$ is the Fourier transform in $\mathbf{k}\to\boldsymbol{\xi}$ of the solution of (24). We deduce from this formula that the back-propagated signal has much better refocusing properties in random media than in homogeneous media. As is explained in [3], the reason is that the limiting Wigner measure $\bar{W}$ is more regular in its $\mathbf{k}$ variable when the right-hand side in (24) is present. This leads to better spatial decay of F and hence ψ_B is localized tighter in a random medium than in a homogeneous medium. The improvement of the refocusing in random media has been carefully analyzed theoretically and numerically in [9], in the Fokker-Plank approximation of (24).

3.2 Outline of the proof

We summarize here the main steps of the derivation of the results stated in Theorems 3.1 and 3.2. We first obtain in Lemma 4.1 uniform bounds in L^2 for the Wigner transform W_ε independently of the realization of the random medium. Lemma 4.2 then shows that Theorem 3.1 is a consequence of Theorem 3.2.

The main mathematical restriction is to assume that the random potential $V(z,\mathbf{x})$ is Markovian in its first variable. We refer to section 3.3 for the details of its construction. Following our previous work in [4] (based on a nethod developed in [23]), we show the convergence of the expectation $\mathbb{E}\{\langle W_\varepsilon, \lambda\rangle\} \to \langle \overline{W}, \lambda\rangle$ for every test function λ. Its proof is simpler (and less general) than in [4] because of the available a priori L^2 bounds on the Wigner transform, which satisfies the following Cauchy problem

$$\begin{aligned} \frac{\partial W_\varepsilon}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}} W_\varepsilon &= \mathcal{L}_\varepsilon W_\varepsilon \\ W_\varepsilon(0,\mathbf{x},\mathbf{k}) &= W_\varepsilon^0(\mathbf{x},\mathbf{k}), \end{aligned} \tag{26}$$

with

$$\mathcal{L}_\varepsilon W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}} \int_{\mathbb{R}^d} \frac{d\tilde{V}(\frac{z}{\varepsilon},\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{x}/\varepsilon} \left[W_\varepsilon(\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}) - W_\varepsilon(\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2})\right].$$

Since κ is a fixed parameter, we set $\kappa = 1$ without loss of generality. The solution of (26) with initial data in L^2 is understood in the sense that for every smooth test function $\lambda(z,\mathbf{x},\mathbf{k})$, we have

$$\langle W_\varepsilon(z), \lambda(z)\rangle - \langle W_\varepsilon^0, \lambda(0)\rangle = \int_0^z \langle W_\varepsilon(s), \left(\frac{\partial}{\partial s} + \mathbf{k}\cdot\nabla_{\mathbf{x}} + \mathcal{L}_\varepsilon\right)\lambda(s)\rangle ds.$$

Here, we have used that $\mathcal{L}_\varepsilon$ is a self-adjoint operator for $\langle\cdot,\cdot\rangle$. Therefore for a smooth function $\lambda_0(\mathbf{x},\mathbf{k})$ we obtain $\langle W_\varepsilon(z), \lambda_0\rangle = \langle W_\varepsilon^0, \lambda_\varepsilon(0)\rangle$, where $\lambda_\varepsilon(s)$ is the solution of the backward problem

$$\frac{\partial \lambda_\varepsilon}{\partial s} + \mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda_\varepsilon + \mathcal{L}_\varepsilon\lambda_\varepsilon(s) = 0, \quad 0 \le s \le z$$

with the terminal condition $\lambda_\varepsilon(z,\mathbf{x},\mathbf{k}) = \lambda_0(\mathbf{x},\mathbf{k})$. This defines the process $W_\varepsilon(z)$ in $L^2(\mathbb{R}^{2d})$ and generates a corresponding measure P_ε on the space $C([0,L];L^2(\mathbb{R}^{2d}))$ of continuous functions in time with values in L^2. The measure P_ε is actually supported on paths inside a ball $X = \left\{W \in L^2 :\ \|W\|_{L^2} \le C\right\}$ with the constant C as in Lemma 4.1. The set X is the state space for the random process $W_\varepsilon(z)$. We have proved in [4] and will use in the sequel that the family P_ε is tight:

Lemma 3.3 *The family of measures P_ε is weakly compact.*

The proof of convergence of W_ε to its deterministic limit is obtained in two steps. Let us fix a deterministic test function $\lambda(z,\mathbf{x},\mathbf{k})$. We use the Markovian property of the random field $V(z,\mathbf{x})$ in z to construct a first functional $G_\lambda\colon C([0,L];X) \to C[0,L]$ by

$$G_\lambda[W](z) = \langle W, \lambda\rangle(z) - \int_0^z \langle W, \frac{\partial\lambda}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda + \mathcal{L}\lambda\rangle(\zeta)d\zeta \tag{27}$$

and show that it is an approximate P_ε-martingale. More precisely, we show that

$$\left|\mathbb{E}^{P_\varepsilon}\left\{G_\lambda[W](z)|\mathcal{F}_s\right\} - G_\lambda[W](s)\right| \le C_{\lambda,L}\sqrt{\varepsilon} \tag{28}$$

uniformly for all $W \in C([0,L];X)$ and $0 \le s < z \le L$. Lemma 3.3 implies that there exists a subsequence $\varepsilon_j \to 0$ so that P_{ε_j} converges weakly to a measure P supported on $C([0,L];X)$. Weak convergence of P_ε and the strong convergence (28) together imply that $G_\lambda[W](z)$ is a P-martingale so that

$$\mathbb{E}^P\left\{G_\lambda[W](z)|\mathcal{F}_s\right\} - G_\lambda[W](s) = 0. \tag{29}$$

Taking $s = 0$ above we obtain as in [4] the transport equation (24) for $\overline{W} = \mathbb{E}^P\left\{W(z)\right\}$ in its weak formulation.

The second step is to show that for every test function $\lambda(z,\mathbf{x},\mathbf{k})$ the new functional

$$G_{2,\lambda}[W](z) = \langle W,\lambda\rangle^2(z) - 2\int_0^z \langle W,\lambda\rangle(\zeta)\langle W, \frac{\partial\lambda}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda + \mathcal{L}\lambda\rangle(\zeta)d\zeta$$

is also an approximate P_ε-martingale. We then obtain that $\mathbb{E}^{P_\varepsilon}\left\{\langle W,\lambda\rangle^2\right\} \to \langle \overline{W},\lambda\rangle^2$, which implies convergence in probability. It follows that the limit measure P is unique and deterministic, and that the whole sequence P_ε converges.

That $G_{2,\lambda}[W](z)$ is an approximate P_ε-martingale uses very explicitly the uniform a priori L^2 bound on the Wigner distribution W_ε. When the a priori bound is available only in a much larger space (the space $\mathcal{A}'$ of [4]), we are not able to prove that the functional $G_{2,\lambda}$ is an approximate P_ε-martingale, and actually suspect that the result is not true. We expect to observe that some long-range correlations survive for sufficiently singular initial data, which would imply that the limiting measure P is no longer deterministic.

3.3 The random refraction index

We describe here the construction of the random potential $V(z,\mathbf{x})$. The refraction index $n(z,\mathbf{x})$ in non-dimensional variables is a random function of the form

$$n^2(z,\mathbf{x}) = 1 - 2\sigma V(z,\mathbf{x}),$$

where the non-dimensional parameter σ measures the strength of fluctuations.

The mathematical analysis of the Wigner distribution in random media can be obtained by the method of diagrammatic expansions of the solution of the Schrödinger equation as in [13, 25]. However, here we assume that the random field $V(z,\mathbf{x})$ is a Markov process in z, as in [4], and we are able to analyze the Wigner transform W_ε directly. The Markovian hypothesis is crucial to simplify the mathematical analysis because it allows us to treat the process $z \mapsto (V(z/\varepsilon,\mathbf{x}/\varepsilon), W_\varepsilon(z,\mathbf{x},\mathbf{k}))$ as jointly Markov.

In addition, $V(z,\mathbf{x})$ is assumed to be stationary in $\mathbf{x}$ and z, mean zero, and is constructed in the Fourier space as follows. Let $\mathcal{V}$ be the set of measures of bounded total variation with support inside a ball $B_L = \{|\mathbf{p}| \le L\}$

$$\mathcal{V} = \left\{\hat{V} :\ \int_{\mathbb{R}^d} |d\hat{V}| \le C,\ \text{supp}\ \hat{V} \subset B_L,\ \hat{V}(\mathbf{p}) = \hat{V}^*(-\mathbf{p})\right\} \tag{30}$$

and let $\tilde{V}(z)$ be a mean-zero Markov process on $\mathcal{V}$ with generator Q. The random potential $V(z,\mathbf{x})$ is given by

$$V(z,\mathbf{x}) = \int_{\mathbb{R}^d} \frac{d\tilde{V}(z,\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{x}}$$

and is real and uniformly bounded:

$$|V(z,\mathbf{x})| \le C.$$

We assume that the process $V(z,\mathbf{x})$ is stationary in z and $\mathbf{x}$ with correlation function $R(z,\mathbf{x})$

$$\mathbb{E}\{V(s,\mathbf{y})V(z+s,\mathbf{x}+\mathbf{y})\} = R(z,\mathbf{x}) \quad \text{for all } \mathbf{x},\mathbf{y}\in\mathbb{R}^d, \text{ and } z,s\in\mathbb{R}.$$

In terms of the process $\tilde{V}(z,\mathbf{p})$ this means that given any two bounded continuous functions $\hat{\phi}(\mathbf{p})$ and $\hat{\psi}(\mathbf{p})$ we have

$$\mathbb{E}\left\{\langle\tilde{V}(s),\hat{\phi}\rangle\langle\tilde{V}(z+s),\hat{\psi}\rangle\right\} = (2\pi)^d\int_{\mathbb{R}^d} d\mathbf{p}\tilde{R}(z,\mathbf{p})\hat{\phi}(\mathbf{p})\hat{\psi}(-\mathbf{p}). \tag{31}$$

Here $\langle\cdot,\cdot\rangle$ is the usual duality product on $\mathbb{R}^d\times\mathbb{R}^d$, and the power spectrum $\tilde{R}$ is the Fourier transform of $R(z,\mathbf{x})$ in $\mathbf{x}$:

$$\tilde{R}(z,\mathbf{p})=\int_{\mathbb{R}^d}d\mathbf{x}e^{-i\mathbf{p}\cdot\mathbf{x}}R(z,\mathbf{x}).$$

We assume that $\tilde{R}(z,\mathbf{p})\in\mathcal{S}(\mathbb{R}\times\mathbb{R}^d)$ for simplicity and define $\hat{R}(\omega,\mathbf{p})$ as

$$\hat{R}(\omega,\mathbf{p})=\int_{\mathbb{R}}dze^{-i\omega z}\tilde{R}(z,\mathbf{p}),\tag{32}$$

which is the space-time Fourier transform of R.

We assume that the generator Q is a bounded operator on $L^\infty(\mathcal{V})$ with a unique invariant measure $\pi(\hat{V})$

$$Q^*\pi=0,$$

and that there exists $\alpha>0$ such that if $\langle g,\pi\rangle=0$ then

$$\|e^{rQ}g\|_{L^\infty_{\mathcal{V}}}\leq C\|g\|_{L^\infty_{\mathcal{V}}}e^{-\alpha r}.\tag{33}$$

The simplest example of a generator with gap in the spectrum and invariant measure π is a jump process on $\mathcal{V}$ where

$$Qg(\hat{V})=\int_{\mathcal{V}}g(\hat{V}_1)d\pi(\hat{V}_1)-g(\hat{V}),\qquad\int_{\mathcal{V}}d\pi(\hat{V})=1.$$

Given (33), the Fredholm alternative holds for the Poisson equation

$$Qf=g,$$

provided that g satisfies $\langle\pi,g\rangle=0$. It has a unique solution f with $\langle\pi,f\rangle=0$ and $\|f\|_{L^\infty_V}\leq C\|g\|_{L^\infty_V}$. The solution f is given explicitly by

$$f(\hat{V})=-\int_0^\infty dre^{rQ}g(\hat{V}),$$

and the integral converges absolutely because of (33).

The particular Markovian model adopted in this paper is somewhat restrictive. However, the only information about the process that enters into the main result, the transport equation for the Wigner distribution in the limit $\varepsilon\to 0$, is the two-point correlation function of V. Therefore, we expect that self-averaging of the Wigner transform and convergence to the solution of the transport equation hold for much more general classes of random perturbations of the refraction index.

4 Convergence of the Wigner transform

4.1 A priori bounds on the Wigner transform

The Wigner transform satisfies the following uniform bound.

Lemma 4.1 *There exists a constant C that is independent of ε so that*

$$\|W_\varepsilon(z)\|_{L^2(\mathbb{R}^{2d})}\leq C\|f\|^2_{L^2(\mathbb{R}^d)}\|\chi\|^4_{L^4(\mathbb{R}^d)}\tag{34}$$

for all $z\geq 0$.

Proof. It is shown in [20] that equation (26) preserves the L^2-norm of W_ε: $\|W_\varepsilon(z)\|_{L^2} = \|W_\varepsilon(z = 0)\|_{L^2}$. The initial conditions for W_ε are given by

$$W_\varepsilon(0,\mathbf{x},\mathbf{k}) = \int_{\mathbb{R}^{2d}} \frac{d\mathbf{y}d\mathbf{q}}{(2\pi)^d} e^{i\mathbf{k}\cdot\mathbf{y}+i\mathbf{q}\cdot\mathbf{y}} \hat{f}(\mathbf{q})\chi(\mathbf{x}-\frac{\varepsilon\mathbf{y}}{2})\chi(\mathbf{x}+\frac{\varepsilon\mathbf{y}}{2}) = \int_{\mathbb{R}^d} d\mathbf{y} e^{-i\mathbf{k}\cdot\mathbf{y}} f(\mathbf{y})\chi(\mathbf{x}+\frac{\varepsilon\mathbf{y}}{2})\chi(\mathbf{x}-\frac{\varepsilon\mathbf{y}}{2})$$

so that

$$\begin{aligned}\int_{\mathbb{R}^{2d}} d\mathbf{x}d\mathbf{k}|W_\varepsilon(0,\mathbf{x},\mathbf{k})|^2 =& \int_{\mathbb{R}^{4d}} d\mathbf{y}d\mathbf{y}_1 d\mathbf{x}d\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{y}_1 - i\mathbf{k}\cdot\mathbf{y}} f(\mathbf{y})\overline{f(\mathbf{y}_1)}\chi(\mathbf{x}-\frac{\varepsilon\mathbf{y}}{2})\chi(\mathbf{x}+\frac{\varepsilon\mathbf{y}}{2}) \\ \times\chi(\mathbf{x}-\frac{\varepsilon\mathbf{y}_1}{2})\chi(\mathbf{x}+\frac{\varepsilon\mathbf{y}_1}{2}) =& (2\pi)^d \int_{\mathbb{R}^{2d}} d\mathbf{x}d\mathbf{y}|f(\mathbf{y})|^2 \left|\chi(\mathbf{x}-\frac{\varepsilon\mathbf{y}}{2})\right|^2 \left|\chi(\mathbf{x}+\frac{\varepsilon\mathbf{y}}{2})\right|^2\end{aligned}$$

and (34) follows.

The above calculation also shows that the initial Wigner distribution $W_\varepsilon^0(\mathbf{x},\mathbf{k})$ converges to

$$W_0(\mathbf{x},\mathbf{k}) = \hat{f}(\mathbf{k})|\chi(\mathbf{x})|^2.$$

Notice however that one cannot expect strong convergence in L^2 of $W_\varepsilon(z)$ to $\overline{W}(z)$, the solution of the transport equation (24), because the latter has an L^2-norm that decreases as z increases while the L^2-norm of W_ε is preserved.

The next Lemma shows that one may drop the term $\varepsilon(\mathbf{y}+\boldsymbol{\xi})/2$ in the argument of W_ε in expression (22) for the back-propagated signal.

Lemma 4.2 *Let $\phi(\mathbf{x}_0) \in L^2(\mathbb{R}^d)$, then*

$$\left|\int_{\mathbb{R}^d} d\mathbf{x}_0 \psi_\varepsilon^B(\boldsymbol{\xi};\mathbf{x}_0)\phi^*(\mathbf{x}_0) - \int_{\mathbb{R}^{2d}} d\mathbf{x}_0 d\mathbf{k} e^{i\mathbf{k}\cdot\boldsymbol{\xi}} W_\varepsilon(L,\mathbf{x}_0,\mathbf{k})\phi^*(\mathbf{x}_0)\hat{\psi}_0(\mathbf{k})\right| \to 0 \quad as\ \varepsilon\to 0 \tag{35}$$

for all $\boldsymbol{\xi}$.

Proof. We have

$$\begin{aligned}&\int_{\mathbb{R}^d} \psi_\varepsilon^B(\boldsymbol{\xi};\mathbf{x}_0)\phi^*(\mathbf{x}_0)d\mathbf{x}_0 - \int_{\mathbb{R}^{2d}} e^{i\mathbf{k}\cdot\boldsymbol{\xi}} W_\varepsilon(L,\mathbf{x}_0,\mathbf{k})\phi(\mathbf{x}_0)\hat{\psi}_0(\mathbf{k})d\mathbf{x}_0 d\mathbf{k} \\ &= \int_{\mathbb{R}^{2d}} \hat{\phi}^*(\mathbf{q})\hat{W}_\varepsilon(L,\mathbf{q},\mathbf{y}-\boldsymbol{\xi})\psi_0(\mathbf{y})\left(e^{-i\varepsilon\mathbf{q}\cdot(\mathbf{y}+\boldsymbol{\xi})/2} - 1\right)\frac{d\mathbf{y}d\mathbf{q}}{(2\pi)^d}\end{aligned}$$

while

$$\int_{\mathbb{R}^{2d}} |\hat{\phi}(\mathbf{q})|^2|\psi_0(\mathbf{y})|^2 \left|1 - e^{-i\varepsilon\mathbf{q}\cdot(\mathbf{y}+\boldsymbol{\xi})/2}\right|^2 d\mathbf{y}d\mathbf{q} \to 0$$

by the Lebesgue dominated convergence theorem. Therefore Lemma 4.2 follows from Lemma 4.1.

Theorem 3.1 then follows from Lemma 4.2 and Theorem 3.2, which is proved in the following sections.

4.2 Convergence of the expectation

To obtain the approximate martingale property (28), one has to consider the conditional expectation of functionals $F(W,\hat{V})$ with respect to the probability measure $\tilde{P}_\varepsilon$ on the space $C([0,L];\mathcal{V}\times X)$ generated by $V(z/\varepsilon)$ and the Cauchy problem (26). The only functions we need to consider are actually of the form $F(W,\hat V) = \langle W, \lambda(\hat V)\rangle$ with $\lambda \in L^\infty(\mathcal{V}; C^1([0,L]; \mathcal{S}(\mathbb{R}^{2d})))$. Given a function $F(W,\hat V)$ let us define the conditional expectation

$$\mathbb{E}^{\tilde P_\varepsilon}_{W,\hat V,z}\left\{F(W,\hat V)\right\}(\tau) = \mathbb{E}^{\tilde P_\varepsilon}\left\{F(W(\tau),\tilde V(\tau))|\ W(z)=W, \tilde V(z)=\hat V\right\},\quad \tau \geq z.$$

The weak form of the infinitesimal generator of the Markov process generated by $\tilde P_\varepsilon$ is given by

$$\frac{d}{dh}\mathbb{E}^{\tilde P_\varepsilon}_{W,\hat V,z}\left\{\langle W,\lambda(\hat V)\rangle\right\}(z+h)\Big|_{h=0} = \frac{1}{\varepsilon}\langle W, Q\lambda\rangle + \left\langle W, \left(\frac{\partial}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}} + \frac{1}{\sqrt{\varepsilon}}\mathcal{K}[\hat V, \frac{\mathbf{x}}{\varepsilon}]\right)\lambda\right\rangle, \tag{36}$$

hence

$$G^\varepsilon_\lambda = \langle W,\lambda(\hat V)\rangle(z) - \int_0^z \left\langle W, \left(\frac{1}{\varepsilon}Q + \frac{\partial}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}} + \frac{1}{\sqrt{\varepsilon}}\mathcal{K}[\hat V, \frac{\mathbf{x}}{\varepsilon}]\right)\lambda\right\rangle(s)ds \tag{37}$$

is a $\tilde P_\varepsilon$-martingale. The operator $\mathcal{K}$ is defined by

$$\mathcal{K}[\hat V,\boldsymbol{\eta}]\psi(\mathbf{x},\boldsymbol{\eta},\mathbf{k},\hat V) = \frac{1}{i}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{i\mathbf{p}\cdot\boldsymbol{\eta}}\left[\psi(\mathbf{x},\boldsymbol{\eta},\mathbf{k}-\frac{\mathbf{p}}{2}) - \psi(\mathbf{x},\boldsymbol{\eta},\mathbf{k}+\frac{\mathbf{p}}{2})\right]. \tag{38}$$

The generator (36) comes from equation (26) written in the form

$$\frac{\partial W_\varepsilon}{\partial z} + \mathbf{k}\cdot\nabla_{\mathbf{x}}W_\varepsilon = \frac{1}{\sqrt{\varepsilon}}\mathcal{K}[\tilde V(z/\varepsilon), \mathbf{x}/\varepsilon]W_\varepsilon. \tag{39}$$

Given a test function $\lambda(z,\mathbf{x},\mathbf{k}) \in C^1([0,L];\mathcal{S})$ we construct a function

$$\lambda_\varepsilon(z,\mathbf{x},\mathbf{k},\hat V) = \lambda(z,\mathbf{x},\mathbf{k}) + \sqrt{\varepsilon}\lambda_1^\varepsilon(z,\mathbf{x},\mathbf{k},\hat V) + \varepsilon\lambda_2^\varepsilon(z,\mathbf{x},\mathbf{k},\hat V) \tag{40}$$

with $\lambda^\varepsilon_{1,2}(t)$ bounded in $L^\infty(\mathcal{V};L^2(\mathbb{R}^{2d}))$ uniformly in $z\in[0,L]$. The functions $\lambda^\varepsilon_{1,2}$ will be chosen so that

$$\|G^\varepsilon_{\lambda_\varepsilon}(z) - G_\lambda(z)\|_{L^\infty(\mathcal{V})} \leq C_\lambda\sqrt{\varepsilon} \tag{41}$$

for all $z\in[0,L]$. Here $G^\varepsilon_{\lambda_\varepsilon}$ is defined by (37) with λ replaced by λ_ε, and G_λ is defined by (27). The approximate martingale property (28) follows from this.

The functions λ_1^ε and λ_2^ε are as follows. Let $\lambda_1(z,\mathbf{x},\boldsymbol{\eta},\mathbf{k},\hat V)$ be the mean-zero solution of the Poisson equation

$$\mathbf{k}\cdot\nabla_{\boldsymbol{\eta}}\lambda_1 + Q\lambda_1 = -\mathcal{K}\lambda. \tag{42}$$

It is given explicitly by

$$\lambda_1(z,\mathbf{x},\boldsymbol{\eta},\mathbf{k},\hat V) = \frac{1}{i}\int_0^\infty dr e^{rQ}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{ir(\mathbf{k}\cdot\mathbf{p})+i(\boldsymbol{\eta}\cdot\mathbf{p})}\left[\lambda(z,\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}) - \lambda(z,\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2})\right]. \tag{43}$$

Then we let $\lambda_1^\varepsilon(z,\mathbf{x},\mathbf{k},\hat V) = \lambda_1(z,\mathbf{x},\mathbf{x}/\varepsilon,\mathbf{k},\hat V)$. Furthermore, the second order corrector is given by $\lambda_2^\varepsilon(z,\mathbf{x},\mathbf{k},\hat V) = \lambda_2(z,\mathbf{x},\mathbf{x}/\varepsilon,\mathbf{k},\hat V)$ where $\lambda_2(z,\mathbf{x},\boldsymbol{\eta},\mathbf{k},\hat V)$ is the mean-zero solution of

$$\mathbf{k}\cdot\nabla_{\boldsymbol{\eta}}\lambda_2 + Q\lambda_2 = \mathcal{L}\lambda - \mathcal{K}\lambda_1, \tag{44}$$

which exists because $\mathbb{E}\{\mathcal{K}\lambda_1\} = \mathcal{L}\lambda$, and is given by

$$\lambda_2(z,\mathbf{x},\boldsymbol{\eta},\mathbf{k},\hat V) = -\int_0^\infty dr e^{rQ}\left[\mathcal{L}\lambda(z,\mathbf{x},\mathbf{k}) - [\mathcal{K}\lambda_1](z,\mathbf{x},\boldsymbol{\eta}+r\mathbf{k},\mathbf{k},\hat V)\right].$$

Using (42) and (44) we have

$$\frac{d}{dh}\mathbb{E}^{\tilde{P}_\varepsilon}_{W,\hat{V},z}\left\{\langle W,\lambda_\varepsilon\rangle\right\}(z+h)\Big|_{h=0} = \left\langle W, \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}+\frac{1}{\sqrt{\varepsilon}}\mathcal{K}[\hat{V},\frac{\mathbf{x}}{\varepsilon}]+\frac{1}{\varepsilon}Q\right)\left(\lambda+\sqrt{\varepsilon}\lambda_1^\varepsilon+\varepsilon\lambda_2^\varepsilon\right)\right\rangle$$
$$= \left\langle W, \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\right)\lambda+\mathcal{L}\lambda\right\rangle + \left\langle W, \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\right)\left(\sqrt{\varepsilon}\lambda_1^\varepsilon+\varepsilon\lambda_2^\varepsilon\right)+\sqrt{\varepsilon}\mathcal{K}[\hat{V},\frac{\mathbf{x}}{\varepsilon}]\lambda_2^\varepsilon\right\rangle$$
$$= \left\langle W, \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\right)\lambda+\mathcal{L}\lambda\right\rangle + \sqrt{\varepsilon}\langle W,\zeta_\varepsilon^\lambda\rangle$$

with

$$\zeta_\varepsilon^\lambda = \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\right)\lambda_1^\varepsilon + \sqrt{\varepsilon}\left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\right)\lambda_2^\varepsilon + \mathcal{K}[\hat{V},\frac{\mathbf{x}}{\varepsilon}]\lambda_2^\varepsilon.$$

The terms $\mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda_{1,2}^\varepsilon$ above are understood as differentiation with respect to the slow variable $\mathbf{x}$ only, and not with respect to $\boldsymbol{\eta}=\mathbf{x}/\varepsilon$. It follows that $G^\varepsilon_{\lambda_\varepsilon}$ is given by

$$G^\varepsilon_{\lambda_\varepsilon}(z) = \langle W(z),\lambda_\varepsilon\rangle - \int_0^z ds\left\langle W, \left(\frac{\partial}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}+\mathcal{L}\right)\lambda\right\rangle(s) - \sqrt{\varepsilon}\int_0^z ds\langle W,\zeta_\varepsilon^\lambda\rangle(s) \tag{45}$$

and is a martingale with respect to the measure $\tilde{P}_\varepsilon$ defined on $D([0,L];X\times\mathcal{V})$, the space of right-continuous paths with left-side limits [8]. The estimate (28) follows from the following two lemmas.

Lemma 4.3 *Let $\lambda\in C^1([0,L];\mathcal{S}(\mathbb{R}^{2d}))$. Then there exists a constant $C_\lambda>0$ independent of $z\in[0,L]$ so that the correctors $\lambda_1^\varepsilon(z)$ and $\lambda_2^\varepsilon(z)$ satisfy the uniform bounds*

$$\|\lambda_1^\varepsilon(z)\|_{L^\infty(\mathcal{V};L^2)}+\|\lambda_2^\varepsilon(z)\|_{L^\infty(\mathcal{V};L^2)}\le C_\lambda \tag{46}$$

and

$$\left\|\frac{\partial\lambda_1^\varepsilon(z)}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda_1^\varepsilon(z)\right\|_{L^\infty(\mathcal{V};L^2)}+\left\|\frac{\partial\lambda_2^\varepsilon(z)}{\partial z}+\mathbf{k}\cdot\nabla_{\mathbf{x}}\lambda_2^\varepsilon(z)\right\|_{L^\infty(\mathcal{V};L^2)}\le C_\lambda. \tag{47}$$

Lemma 4.4 *There exists a constant C_λ such that*

$$\|\mathcal{K}[\hat{V},\mathbf{x}/\varepsilon]\|_{L^2\to L^2}\le C$$

for any $\hat{V}\in\mathcal{V}$ and all $\varepsilon\in(0,1]$.

Indeed, (46) implies that $|\langle W,\lambda\rangle-\langle W,\lambda_\varepsilon\rangle|\le C\sqrt{\varepsilon}$ for all $W\in X$ and $\hat{V}\in\mathcal{V}$, while (47) and Lemma 4.4 imply that for all $z\in[0,L]$

$$\|\zeta_\varepsilon^\lambda(z)\|_{L^2}\le C \tag{48}$$

for all $\hat{V}\in\mathcal{V}$ so that (28) follows.

Proof of Lemma 4.4. Lemma 4.4 follows immediately from the definition of $\mathcal{K}$, the bound (30) and the Cauchy-Schwartz inequality.

We now prove Lemma 4.3. We will omit the z-dependence of the test function λ to simplify the notation.

Proof of Lemma 4.3. We only prove (46). Since $\lambda\in\mathcal{S}(\mathbb{R}^{2d})$, there exists a constant C_λ so that

$$|\lambda(\mathbf{x},\mathbf{k})|\le\frac{C_\lambda}{(1+|\mathbf{x}|^{5d})(1+|\mathbf{k}|^{5d})}.$$

Then we obtain using (30) and (33)

$$
\begin{aligned}
&|\lambda_1^\varepsilon(z,\mathbf{x},\mathbf{k},\hat V)| = C\left|\int_0^\infty dr e^{rQ}\int_{\mathbb{R}^d} d\hat V(\mathbf{p}) e^{ir(\mathbf{k}\cdot\mathbf{p})+i(\mathbf{x}\cdot\mathbf{p})/\varepsilon}\left[\lambda(z,\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2})-\lambda(z,\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2})\right]\right| \\
&\le C\int_0^\infty dr e^{-\alpha r}\sup_{\hat V}\int_{\mathbb{R}^d}|d\hat V(\mathbf{p})|\left[|\lambda(z,\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2})|+|\lambda(z,\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2})|\right] \\
&\le \frac{C}{(1+|\mathbf{x}|^{5d})(1+(|\mathbf{k}|-L)^{5d}\chi_{|\mathbf{k}|\ge 5L}(\mathbf{k}))}
\end{aligned}
$$

and the L^2-bound on λ_1 follows.

We show next that λ_2^ε is uniformly bounded. We have

$$
\begin{aligned}
\lambda_2^\varepsilon(\mathbf{x},\mathbf{k},\hat V) = -\int_0^\infty dr e^{rQ}\Bigg[&\mathcal{L}\lambda(\mathbf{x},\mathbf{k})-\frac{1}{i}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{i\mathbf{p}\cdot(\mathbf{x}/\varepsilon+r\mathbf{k})} \\
&\times\left[\lambda_1(\mathbf{x},\frac{\mathbf{x}}{\varepsilon}+r\mathbf{k},\mathbf{k}-\frac{\mathbf{p}}{2},\hat V)-\lambda_1(\mathbf{x},\frac{\mathbf{x}}{\varepsilon}+r\mathbf{k},\mathbf{k}+\frac{\mathbf{p}}{2},\hat V)\right]\Bigg].
\end{aligned}
$$

The second term above may be written as

$$
\begin{aligned}
&\frac{1}{i}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{i\mathbf{p}\cdot(\mathbf{x}/\varepsilon+r\mathbf{k})}\left[\lambda_1(\mathbf{x},\frac{\mathbf{x}}{\varepsilon}+r\mathbf{k},\mathbf{k}-\frac{\mathbf{p}}{2},\hat V)-\lambda_1(\mathbf{x},\frac{\mathbf{x}}{\varepsilon}+r\mathbf{k},\mathbf{k}+\frac{\mathbf{p}}{2},\hat V)\right] \\
&=-\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{i\mathbf{p}\cdot(\mathbf{x}/\varepsilon+r\mathbf{k})}\int_0^\infty ds e^{sQ}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{q})}{(2\pi)^d}e^{is(\mathbf{k}-\mathbf{p}/2)\cdot\mathbf{q}+i(\mathbf{x}/\varepsilon+r\mathbf{k})\cdot\mathbf{q}} \\
&\times\left[\lambda(\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}-\frac{\mathbf{q}}{2})-\lambda(\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}+\frac{\mathbf{q}}{2})\right] \\
&+\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{p})}{(2\pi)^d}e^{i\mathbf{p}\cdot(\mathbf{x}/\varepsilon+r\mathbf{k})}\int_0^\infty ds e^{sQ}\int_{\mathbb{R}^d}\frac{d\hat V(\mathbf{q})}{(2\pi)^d}e^{is(\mathbf{k}+\mathbf{p}/2)\cdot\mathbf{q}+i(\mathbf{x}/\varepsilon+r\mathbf{k})\cdot\mathbf{q}} \\
&\times\left[\lambda(\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2}-\frac{\mathbf{q}}{2})-\lambda(\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2}+\frac{\mathbf{q}}{2})\right].
\end{aligned}
$$

Therefore we obtain

$$
\begin{aligned}
&|\lambda_2^\varepsilon(\mathbf{x},\mathbf{k},\hat V)|\le C\int_0^\infty dr e^{-\alpha r}\Bigg[|\mathcal{L}\lambda(\mathbf{x},\mathbf{k})|+\sup_{\hat V}\int_{\mathbb{R}^d}|d\hat V(\mathbf{p})|\int_0^\infty ds e^{-\alpha s}\sup_{\hat V_1}\int_{\mathbb{R}^d}|d\hat V_1(\mathbf{q})| \\
&\times\Big(|\lambda(\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}-\frac{\mathbf{q}}{2})|+|\lambda(\mathbf{x},\mathbf{k}-\frac{\mathbf{p}}{2}+\frac{\mathbf{q}}{2})|+|\lambda(\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2}-\frac{\mathbf{q}}{2})|+\lambda(\mathbf{x},\mathbf{k}+\frac{\mathbf{p}}{2}+\frac{\mathbf{q}}{2})\Big)\Bigg] \\
&\le C\left[|\mathcal{L}\lambda(\mathbf{x},\mathbf{k})|+\frac{1}{(1+|\mathbf{x}|^{5d})(1+(|\mathbf{k}|-L)^{5d}\chi_{|\mathbf{k}|\ge 5L}(\mathbf{k}))}\right]
\end{aligned}
$$

and the L^2-bound on λ_2^ε in (46) follows because the operator $\mathcal{L}: L^2\to L^2$ is bounded. The proof of (47) is very similar and is omitted.

Lemma 4.3 and Lemma 4.4 together with (45) imply the bound (41). The tightness of measures P_ε given by Lemma 3.3 implies then that the expectation $\mathbb{E}\{W_\varepsilon(z,\mathbf{x},\mathbf{k})\}$ converges weakly in $L^2(\mathbb{R}^{2d})$ to the solution $\overline{W}(z,\mathbf{x},\mathbf{k})$ of the transport equation for each $z\in[0,L]$.

4.3 Convergence in probability

We now prove that for any test function λ the second moment $\mathbb{E}\{\langle W_\varepsilon,\lambda\rangle^2\}$ converges to $\langle\overline{W},\lambda\rangle^2$. This will imply the convergence in probability claimed in Theorem 3.2. The proof is similar to that for $\mathbb{E}\{\langle W_\varepsilon, \lambda\rangle\}$ and is based on constructing an appropriate approximate martingale for the functional $\langle W\otimes W, \mu\rangle$, where $\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2)$ is a test function, and $W\otimes W(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2) = W(z,\mathbf{x}_1,\mathbf{k}_1)W(z,\mathbf{x}_2,\mathbf{k}_2)$. We need to consider the action of the infinitesimal generator on functions of W and $\hat{V}$ of the form

$$F(W,\hat{V}) = \langle W(\mathbf{x}_1,\mathbf{k}_1)W(\mathbf{x}_2,\mathbf{k}_2), \mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2,\hat{V})\rangle = \langle W\otimes W, \mu(\hat{V})\rangle$$

where μ is a given function. The infinitesimal generator acts on such functions as

$$\frac{d}{dh}\mathbb{E}^{\tilde{P}_\varepsilon}_{W,\hat{V},z}\left\{\langle W\otimes W,\mu(\hat{V})\rangle\right\}(z+h)\Big|_{h=0} = \frac{1}{\varepsilon}\langle W\otimes W, Q\lambda\rangle + \langle W\otimes W, \mathcal{H}_2^\varepsilon\mu\rangle, \tag{49}$$

where

$$\mathcal{H}_2^\varepsilon\mu = \sum_{j=1}^2 \frac{1}{\sqrt{\varepsilon}}\mathcal{K}_j\left[\hat{V},\frac{\mathbf{x}^j}{\varepsilon}\right]\mu + \mathbf{k}^j\cdot\nabla_{\mathbf{x}^j}\mu, \tag{50}$$

with

$$\mathcal{K}_1[\hat{V},\boldsymbol{\eta}_1]\mu = \frac{1}{i}\int_{\mathbb{R}^d} d\hat{V}(\mathbf{p})e^{i(\mathbf{p}\cdot\boldsymbol{\eta}_1)}\left[\mu(\mathbf{k}_1-\frac{\mathbf{p}}{2},\mathbf{k}_2) - \mu(\mathbf{k}_1+\frac{\mathbf{p}}{2},\mathbf{k}_2)\right]$$

and

$$\mathcal{K}_2[\hat{V},\boldsymbol{\eta}_2]\mu = \frac{1}{i}\int_{\mathbb{R}^d} d\hat{V}(\mathbf{p})e^{i(\mathbf{p}\cdot\boldsymbol{\eta}_2)}\left[\mu(\mathbf{k}_1,\mathbf{k}_2-\frac{\mathbf{p}}{2}) - \mu(\mathbf{k}_1,\mathbf{k}_2+\frac{\mathbf{p}}{2})\right].$$

Therefore the functional

$$\begin{aligned} G_\mu^{2,\varepsilon} &= \langle W\otimes W, \mu(\hat{V})\rangle(z) \\ &- \int_0^z \left\langle W\otimes W, \left(\frac{1}{\varepsilon}Q + \frac{\partial}{\partial z} + \mathbf{k}_1\cdot\nabla_{\mathbf{x}_1} + \mathbf{k}_2\cdot\nabla_{\mathbf{x}_2} + \frac{1}{\sqrt{\varepsilon}}(\mathcal{K}_1[\hat{V},\frac{\mathbf{x}_1}{\varepsilon}] + \mathcal{K}_2[\hat{V},\frac{\mathbf{x}_2}{\varepsilon}])\right)\mu\right\rangle(s)ds \end{aligned} \tag{51}$$

is a $\tilde{P}^\varepsilon$ martingale. We let $\mu(z,\mathbf{X},\mathbf{K})\in\mathcal{S}(\mathbb{R}^{2d}\times\mathbb{R}^{2d})$ be a test function independent of $\hat{V}$, where $\mathbf{X} = (\mathbf{x}_1,\mathbf{x}_2)$, and $\mathbf{K} = (\mathbf{k}_1,\mathbf{k}_2)$. We define an approximation

$$\mu_\varepsilon(z,\mathbf{X},\mathbf{K}) = \mu(z,\mathbf{X},\mathbf{K}) + \sqrt{\varepsilon}\mu_1(z,\mathbf{X},\mathbf{X}/\varepsilon,\mathbf{K}) + \varepsilon\mu_2(\mathbf{X},\mathbf{X}/\varepsilon,\mathbf{K}).$$

We will use the notation $\mu_1^\varepsilon(z,\mathbf{X},\mathbf{K}) = \mu_1(z,\mathbf{X},\mathbf{X}/\varepsilon,\mathbf{K})$ and $\mu_2^\varepsilon(z,\mathbf{X},\mathbf{K}) = \mu_2(z,\mathbf{X},\mathbf{X}/\varepsilon,\mathbf{K})$. The functions μ_1 and μ_2 are to be determined. We now use (49) to get

$$\begin{aligned} D_\varepsilon &:= \frac{d}{dh}\Big|_{h=0}\mathbb{E}_{W,\hat{V},z}(\langle W\otimes W,\mu_\varepsilon(\hat{V})\rangle)(z+h) = \frac{1}{\varepsilon}\left\langle W\otimes W, \left(Q + \sum_{j=1}^2\mathbf{k}^j\cdot\nabla_{\boldsymbol{\eta}^j}\right)\mu\right\rangle \\ &+ \frac{1}{\sqrt{\varepsilon}}\left\langle W\otimes W, \left(Q+\sum_{j=1}^2\mathbf{k}^j\cdot\nabla_{\boldsymbol{\eta}^j}\right)\mu_1 + \sum_{j=1}^2\mathcal{K}_j\left[\hat{V},\boldsymbol{\eta}^j\right]\mu\right\rangle \\ &+ \left\langle W\otimes W, \left(Q+\sum_{j=1}^2\mathbf{k}^j\cdot\nabla_{\boldsymbol{\eta}^j}\right)\mu_2 + \sum_{j=1}^2\mathcal{K}_j\left[\hat{V},\boldsymbol{\eta}^j\right]\mu_1 + \frac{\partial\mu}{\partial z} + \sum_{j=1}^2\mathbf{k}^j\cdot\nabla_{\mathbf{x}^j}\mu\right\rangle \\ &+ \sqrt{\varepsilon}\left\langle W\otimes W, \sum_{j=1}^2\mathcal{K}_j\left[\hat{V},\boldsymbol{\eta}^j\right]\mu_2 + \left(\frac{\partial}{\partial z} + \sum_{j=1}^2\mathbf{k}^j\cdot\nabla_{\mathbf{x}^j}\right)(\mu_1+\sqrt{\varepsilon}\mu_2)\right\rangle. \end{aligned} \tag{52}$$

The above expression is evaluated at $\boldsymbol{\eta}_j = \mathbf{x}_j/\varepsilon$. The term of order ε^{-1} in D_ε vanishes since μ is independent of V and the fast variable $\boldsymbol{\eta}$. We cancel the term of order $\varepsilon^{-1/2}$ in the same way as before by defining μ_1 as the unique mean-zero (in the variables $\hat{V}$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$) solution of

$$\big(Q + \sum_{j=1}^{2} \mathbf{k}^j \cdot \nabla_{\boldsymbol{\eta}^j}\big)\mu_1 + \sum_{j=1}^{2} \mathcal{K}_j\left[\hat{V}, \boldsymbol{\eta}^j\right]\mu = 0. \tag{53}$$

It is given explicitly by

$$\begin{aligned}\mu_1(\mathbf{X}, \boldsymbol{\eta}, \mathbf{K}, \hat{V}) = \frac{1}{i}\int_0^\infty dr e^{rQ} \int_{\mathbb{R}^d} d\hat{V}(\mathbf{p}) e^{ir(\mathbf{k}_1\cdot\mathbf{p}) + i(\boldsymbol{\eta}_1\cdot\mathbf{p})} \left[\mu(\mathbf{k}_1 - \frac{\mathbf{p}}{2}, \mathbf{k}_2) - \mu(\mathbf{k}_1 + \frac{\mathbf{p}}{2}, \mathbf{k}_2)\right] \\ + \frac{1}{i}\int_0^\infty dr e^{rQ} \int_{\mathbb{R}^d} d\hat{V}(\mathbf{p}) e^{ir(\mathbf{k}_2\cdot\mathbf{p}) + i(\boldsymbol{\eta}_2\cdot\mathbf{p})} \left[\mu(\mathbf{k}_1, \mathbf{k}_2 - \frac{\mathbf{p}}{2}) - \mu(\mathbf{k}_1, \mathbf{k}_2 + \frac{\mathbf{p}}{2})\right].\end{aligned}$$

When μ has the form $\mu = \lambda \otimes \lambda$, then μ_1 has the form $\mu_1 = \lambda_1 \otimes \lambda + \lambda \otimes \lambda_1$ with the corrector λ_1 given by (43). Let us also define μ_2 as the mean zero with respect to π_V solution of

$$\big(Q + \sum_{j=1}^{2} \mathbf{k}^j \cdot \nabla_{\boldsymbol{\eta}^j}\big)\mu_2 + \sum_{j=1}^{2} \mathcal{K}_j\left[\hat{V}, \boldsymbol{\eta}^j\right]\mu_1 = \overline{\sum_{j=1}^{2} \mathcal{K}_j\left[\hat{V}, \boldsymbol{\eta}^j\right]\mu_1}, \tag{54}$$

where $\overline{f} = \int d\pi_V f$. The function μ_2 is given by

$$\begin{aligned}\mu_2(\mathbf{X}, \boldsymbol{\eta}, \mathbf{K}, \hat{V}) = -\int_0^\infty dr e^{rQ} \big[\overline{\mathcal{K}_1[\hat{V}, \boldsymbol{\eta}_1 + r\mathbf{k}_1]\mu_1(\mathbf{X}, \boldsymbol{\eta} + r\mathbf{K}, \mathbf{K})} \\ - [\mathcal{K}_1[\hat{V}, \boldsymbol{\eta}_1 + r\mathbf{k}_1]\mu_1](\mathbf{X}, \boldsymbol{\eta} + r\mathbf{K}, \mathbf{K}, \hat{V})\big] \\ - \int_0^\infty dr e^{rQ} \big[\overline{\mathcal{K}_2[\hat{V}, \mathbf{k}_2 + r\boldsymbol{\eta}_2]\mu_1(\mathbf{X}, \boldsymbol{\eta} + r\mathbf{K}, \mathbf{K})} \\ - [\mathcal{K}_2[\hat{V}, \boldsymbol{\eta}_2 + r\mathbf{k}_2]\mu_1](\mathbf{X}, \boldsymbol{\eta} + r\mathbf{K}, \mathbf{K}, \hat{V})\big].\end{aligned} \tag{55}$$

Unlike the first corrector μ_1, the second corrector μ_2 may not be written as an explicit sum of tensor products even if μ has the form $\mu = \lambda \otimes \lambda$ because μ_1 depends on $\hat{V}$.

The $\tilde{P}^\varepsilon$-martingale $G^{2,\varepsilon}_{\mu_\varepsilon}$ is given by

$$\begin{aligned}G^{2,\varepsilon}_{\mu} = \langle W \otimes W, \mu(\hat{V})\rangle(z) - \int_0^z \left\langle W \otimes W, \Big(\frac{\partial}{\partial z} + \mathbf{k}_1 \cdot \nabla_{\mathbf{x}_1} + \mathbf{k}_2 \cdot \nabla_{\mathbf{x}_2} + \mathcal{L}_2^\varepsilon\Big)\mu\right\rangle(s)ds \\ -\sqrt{\varepsilon}\int_0^z \langle W \otimes W, \zeta^\mu_\varepsilon\rangle(s)ds,\end{aligned} \tag{56}$$

where ζ^μ_ε is given by

$$\zeta^\varepsilon_\mu = \sum_{j=1}^{2} \mathcal{K}_j\left[\hat{V}, \frac{\mathbf{x}_j}{\varepsilon}\right]\mu_2^\varepsilon + \left(\frac{\partial}{\partial z} + \sum_{j=1}^{2} \mathbf{k}^j \cdot \nabla_{\mathbf{x}^j}\right)(\mu_1^\varepsilon + \sqrt{\varepsilon}\mu_2^\varepsilon)$$

and the operator $\mathcal{L}_2^\varepsilon$ is defined by

$$\mathcal{L}_2^\varepsilon\mu = -\frac{1}{(2\pi)^d}\int_0^\infty dr\int_{\mathbb{R}^d} d\mathbf{p}\tilde{R}(r,\mathbf{p})\Big[e^{ir(\mathbf{k}_1+\frac{\mathbf{p}}{2})\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1+\mathbf{p},\mathbf{x}_2,\mathbf{k}_2))$$
$$-e^{ir(\mathbf{k}_1-\frac{\mathbf{p}}{2})\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1-\mathbf{p},\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))\Big]$$
$$+\Big[e^{i\mathbf{p}\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}+ir\mathbf{k}_2\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2})-\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2}))$$
$$-e^{i\mathbf{p}\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}+ir\mathbf{k}_2\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2})-\mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2}))\Big]$$
$$+\Big[e^{ir\mathbf{k}_1\cdot\mathbf{p}+i\frac{\mathbf{x}_1-\mathbf{x}_2}{\varepsilon}\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2})-\mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2}))$$
$$-e^{ir\mathbf{k}_1\cdot\mathbf{p}+i\frac{\mathbf{x}_1-\mathbf{x}_2}{\varepsilon}\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2})-\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2}))\Big]$$
$$+\Big[e^{ir(\mathbf{k}_2+\frac{\mathbf{p}}{2})\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2+\mathbf{p}))$$
$$-e^{ir(\mathbf{k}_2-\frac{\mathbf{p}}{2})\cdot\mathbf{p}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2-\mathbf{p})-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))\Big]. \tag{57}$$

We have used in the calculation of $\mathcal{L}_2^\varepsilon$ that for a sufficiently regular function f, we have

$$\mathbb{E}\left[\int_{\mathbb{R}^d}\frac{d\hat{V}(\mathbf{q})}{(2\pi)^d}\int_0^\infty dr\, e^{rQ}\int_{\mathbb{R}^d} d\hat{V}(\mathbf{p})f(r,\mathbf{p},\mathbf{q})\right] = \int_0^\infty dr\int_{\mathbb{R}^d}\tilde{R}(r,\mathbf{p})f(r,\mathbf{p},-\mathbf{p})d\mathbf{p}.$$

The bound on ζ_ε^μ is similar to that on $\zeta_\varepsilon^\lambda$ obtained previously as the correctors μ_j^ε satisfy the same kind of estimates as the correctors λ_j:

Lemma 4.5 *There exists a constant $C_\mu > 0$ so that the functions $\mu_{1,2}^\varepsilon$ obey the uniform bounds*

$$\|\mu_1^\varepsilon(z)\|_{L^2(\mathbb{R}^{2d})} + \|\mu_2^\varepsilon\|_{L^2(\mathbb{R}^{2d})} \le C_\mu \tag{58}$$

and

$$\left\|\frac{\partial\mu_1^\varepsilon(z)}{\partial z} + \sum_{j=1}^2 \mathbf{k}_j\cdot\nabla_{\mathbf{x}_j}\mu_1^\varepsilon(z)\right\|_{L^2(\mathbb{R}^{2d})} + \left\|\frac{\partial\mu_2^\varepsilon(z)}{\partial z} + \sum_{j=1}^2 \mathbf{k}_j\cdot\nabla_{\mathbf{x}_j}\mu_2^\varepsilon(z)\right\|_{L^2(\mathbb{R}^{2d})} \le C_\mu \tag{59}$$

for all $z \in [0, L]$ and $V \in \mathcal{V}$.

The proof of this lemma is very similar to that of Lemma 4.3 and is therefore omitted.

Unlike the first moment case, the averaged operator $\mathcal{L}_2^\varepsilon$ still depends on ε. We therefore do not have strong convergence of the $\tilde{P}^\varepsilon$-martingale $G_{\mu_\varepsilon}^{2,\varepsilon}$ to its limit yet. However, the a priori bound on W_ε in L^2 allows us to characterize the limit of $G_{\mu_\varepsilon}^{2,\varepsilon}$ and show strong convergence. This is shown as follows. The first and last terms in (57) that are independent of ε give the contribution:

$$\mathcal{L}_2\mu = \int_0^\infty dr\int_{\mathbb{R}^d}\frac{d\mathbf{p}}{(2\pi)^d}\Big[\tilde{R}(r,\mathbf{p}-\mathbf{k}_1)e^{ir\frac{p^2-k_1^2}{2}}(\mu(z,\mathbf{x}_1,\mathbf{p},\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))$$
$$+\tilde{R}(r,\mathbf{k}_1-\mathbf{p})e^{ir\frac{k_1^2-p^2}{2}}(\mu(z,\mathbf{x}_1,\mathbf{p}_1,\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))$$
$$+\tilde{R}(z,\mathbf{p}-\mathbf{k}_2)e^{ir\frac{p^2-k_2^2}{2}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{p})-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))$$
$$+\tilde{R}(z,\mathbf{k}_2-\mathbf{p})e^{ir\frac{k_2^2-p^2}{2}}(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{p})-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))\Big]$$
$$=\int_{\mathbb{R}^d}\frac{d\mathbf{p}}{(2\pi)^d}\hat{R}(\frac{p^2-k_1^2}{2},\mathbf{p}-\mathbf{k}_1)(\mu(z,\mathbf{x}_1,\mathbf{p},\mathbf{x}_2,\mathbf{k}_2)-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2))$$
$$+\hat{R}(\frac{p^2-k_2^2}{2},\mathbf{p}-\mathbf{k}_2)(\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{p})-\mu(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2)).$$

The two remaining terms give a contribution that tends to 0 as $\varepsilon \to 0$ for sufficiently smooth test functions. They are given by

$$
\begin{aligned}
&(\mathcal{L}_2^\varepsilon - \mathcal{L}_2)\mu = \frac{1}{(2\pi)^d}\int_0^\infty dr \int_{\mathbb{R}^d} d\mathbf{p}\tilde{R}(r,\mathbf{p})\times \\
&\Big(e^{i\mathbf{p}\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}+ir\mathbf{k}_2\cdot\mathbf{p}} + e^{ir\mathbf{k}_1\cdot\mathbf{p}+i\frac{\mathbf{x}_1-\mathbf{x}_2}{\varepsilon}\cdot\mathbf{p}}\Big)\Big(\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2}) - \mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2})\Big) \\
&+\Big(e^{i\mathbf{p}\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}+ir\mathbf{k}_2\cdot\mathbf{p}} + e^{ir\mathbf{k}_1\cdot\mathbf{p}+i\frac{\mathbf{x}_1-\mathbf{x}_2}{\varepsilon}\cdot\mathbf{p}}\Big)\Big(\mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2}) - \mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2})\Big).
\end{aligned}
$$

We have

$$
\tilde{R}(z,\mathbf{p}) = \tilde{R}(-z,-\mathbf{p}) \geq 0
$$

by Bochner's theorem. Since $(\mathcal{L}_2^\varepsilon - \mathcal{L}_2)$ and λ are real quantities, we can take the real part of the above term and, after the change of variables $r \to -r$ and $\mathbf{p} \to -\mathbf{p}$, obtain

$$
\begin{aligned}
(\mathcal{L}_2^\varepsilon - \mathcal{L}_2)\mu &= \frac{1}{(2\pi)^d}\int_{-\infty}^\infty dr \int_{\mathbb{R}^d} d\mathbf{p}\tilde{R}(r,\mathbf{p})\cos(\mathbf{p}\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon})(e^{ir\mathbf{k}_2\cdot\mathbf{p}} + e^{ir\mathbf{k}_1\cdot\mathbf{p}}) \\
&\times\Big(\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2}) + \mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2}) \\
&- \mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2-\tfrac{\mathbf{p}}{2}) - \mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2})\Big) \\
&= \frac{2}{(2\pi)^d}\int_{\mathbb{R}^d} d\mathbf{p}(\hat{R}(-\mathbf{k}_1\cdot\mathbf{p},\mathbf{p}) + \hat{R}(-\mathbf{k}_2\cdot\mathbf{p},\mathbf{p}))\cos(\mathbf{p}\cdot\tfrac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}) \\
&\times\Big(\mu(z,\mathbf{x}_1,\mathbf{k}_1+\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2}) - \mu(z,\mathbf{x}_1,\mathbf{k}_1-\tfrac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\tfrac{\mathbf{p}}{2})\Big) \\
&= g_1 + g_2 + g_3 + g_4 + c.c.
\end{aligned}
$$

We have (since μ is real-valued)

$$
\begin{aligned}
I = &\int_{\mathbb{R}^{4d}} d\mathbf{x}_1 d\mathbf{k}_1 d\mathbf{x}_2 d\mathbf{k}_2 |g_1(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2)|^2 = C\int_{\mathbb{R}^{6d}} d\mathbf{x}_1 d\mathbf{k}_1 d\mathbf{x}_2 d\mathbf{k}_2 d\mathbf{p} d\mathbf{q}\hat{R}(-\mathbf{k}_1\cdot\mathbf{p},\mathbf{p})\hat{R}(-\mathbf{k}_1\cdot\mathbf{q},\mathbf{q}) \\
&\times e^{i(\mathbf{p}-\mathbf{q})\cdot\frac{\mathbf{x}_2-\mathbf{x}_1}{\varepsilon}}\mu(z,\mathbf{x}_1,\mathbf{k}_1-\frac{\mathbf{p}}{2},\mathbf{x}_2,\mathbf{k}_2+\frac{\mathbf{p}}{2})\mu(z,\mathbf{x}_1,\mathbf{k}_1-\frac{\mathbf{q}}{2},\mathbf{x}_2,\mathbf{k}_2+\frac{\mathbf{q}}{2}).
\end{aligned}
$$

Using density arguments we may assume that μ has the form

$$
\mu(\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2) = \mu_1(\mathbf{x}_1-\mathbf{x}_2)\mu_2(\mathbf{x}_1+\mathbf{x}_2)\mu_3(\mathbf{k}_1)\mu_4(\mathbf{k}_2).
$$

Then we have

$$
\begin{aligned}
I &= C\int_{\mathbb{R}^{6d}} d\mathbf{x}_1 d\mathbf{k}_1 d\mathbf{x}_2 d\mathbf{k}_2 d\mathbf{p} d\mathbf{q}\hat{R}(-\mathbf{k}_1\cdot\mathbf{p},\mathbf{p})\hat{R}(-\mathbf{k}_1\cdot\mathbf{q},\mathbf{q}) \\
&\times e^{-i(\mathbf{p}-\mathbf{q})\cdot\frac{\mathbf{x}_1}{\varepsilon}}\mu_1^2(\mathbf{x}_1)\mu_2^2(\mathbf{x}_2)\mu_3(\mathbf{k}_1-\frac{\mathbf{p}}{2})\mu_4(\mathbf{k}_2+\frac{\mathbf{p}}{2})\mu_3(\mathbf{k}_1-\frac{\mathbf{q}}{2})\mu_4(\mathbf{k}_2+\frac{\mathbf{q}}{2}) \\
&= C\|\mu_2\|_{L^2}^2\int_{\mathbb{R}^{4d}} d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{p} d\mathbf{q}\hat{R}(-\mathbf{k}_1\cdot\mathbf{p},\mathbf{p})\hat{R}(-\mathbf{k}_1\cdot\mathbf{q},\mathbf{q})\hat{\nu}(\frac{\mathbf{p}-\mathbf{q}}{\varepsilon}) \\
&\times\mu_3(\mathbf{k}_1-\frac{\mathbf{p}}{2})\mu_4(\mathbf{k}_2+\frac{\mathbf{p}}{2})\mu_3(\mathbf{k}_1-\frac{\mathbf{q}}{2})\mu_4(\mathbf{k}_2+\frac{\mathbf{q}}{2})
\end{aligned}
$$

where $\nu(\mathbf{x}) = \mu_1^2(\mathbf{x})$. We introduce $G(\mathbf{p}) = \sup_\omega \hat{R}(\omega,\mathbf{p})$ and use the Cauchy-Schwartz inequality in $\mathbf{k}_1$ and $\mathbf{k}_2$:

$$
|I| \leq C\|\mu_2\|_{L^2}^2\|\mu_3\|_{L^2}^2\|\mu_4\|_{L^2}^2\int_{\mathbb{R}^{2d}} d\mathbf{p} d\mathbf{q} G(\mathbf{p})G(\mathbf{q})\left|\hat{\nu}(\frac{\mathbf{p}-\mathbf{q}}{\varepsilon})\right|.
$$

We use again the Cauchy-Schwartz inequality, now in $\mathbf{p}$, to get

$$
\begin{aligned}
|I| &\leq C\|\mu_2\|_{L^2}^2\|\mu_3\|_{L^2}^2\|\mu_4\|_{L^2}^2\|G\|_{L^2}\int_{\mathbb{R}^d} d\mathbf{q}G(\mathbf{q})\left(\int_{\mathbb{R}^d} d\mathbf{p}\left|\hat{\nu}(\frac{\mathbf{p}}{\varepsilon})\right|^2\right)^{1/2} \\
&\leq C\varepsilon^{d/2}\|\mu_2\|_{L^2}^2\|\mu_3\|_{L^2}^2\|\mu_4\|_{L^2}^2\|G\|_{L^2}\|G\|_{L^1}\|\nu\|_{L^2}.
\end{aligned}
$$

This proves that $\|(\mathcal{L}_2^\varepsilon - \mathcal{L}_2)\mu\|_{L^2} \to 0$ as $\varepsilon \to 0$. Notice that oscillatory integrals of the form

$$
\int_{\mathbb{R}^d} e^{i\frac{\mathbf{p}\cdot\mathbf{x}}{\varepsilon}}\mu(\mathbf{p})d\mathbf{p}
$$

are not small in the bigger space $\mathcal{A}'$, which is natural in the context of Wigner transforms and was used in [4]. In this bigger space, we cannot control $(\mathcal{L}_2^\varepsilon - \mathcal{L}_2)\mu$ and actually suspect that the limit measure P may no longer be deterministic.

We therefore deduce that

$$
G_\mu^2 = \langle W\otimes W, \mu(\hat{V})\rangle(z) - \int_0^z \left\langle W\otimes W, \Big(\frac{\partial}{\partial z} + \mathbf{k}_1\cdot\nabla_{\mathbf{x}_1} + \mathbf{k}_2\cdot\nabla_{\mathbf{x}_2} + \mathcal{L}_2\Big)\mu\right\rangle(s)ds
$$

is an approximate $\tilde{P}_\varepsilon$ martingale. The limit of the second moment

$$
W_2(z,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2) = \mathbb{E}^P\left\{W(z,\mathbf{x}_1,\mathbf{k}_1)W(z,\mathbf{x}_2,\mathbf{k}_2)\right\}
$$

thus satisfies (weakly) the transport equation

$$
\frac{\partial W_2}{\partial t} + (\mathbf{k}_1\cdot\nabla_{\mathbf{x}_1} + \mathbf{k}_2\cdot\nabla_{\mathbf{x}_2})W_2 = \mathcal{L}_2 W_2,
$$

with initial data $W_2(0,\mathbf{x}_1,\mathbf{k}_1,\mathbf{x}_2,\mathbf{k}_2) = W_0(\mathbf{x}_1,\mathbf{k}_1)W_0(\mathbf{x}_2,\mathbf{k}_2)$. Moreover, the operator $\mathcal{L}_2$ acting on a tensor product $\lambda\otimes\lambda$ has the form

$$
\mathcal{L}_2[\lambda\otimes\lambda] = \mathcal{L}\lambda\otimes\lambda + \lambda\otimes\mathcal{L}\lambda.
$$

This implies that

$$
\mathbb{E}^P\left\{W(z,\mathbf{x}_1,\mathbf{k}_1)W(z,\mathbf{x}_2,\mathbf{k}_2)\right\} = \mathbb{E}^P\left\{W(z,\mathbf{x}_1,\mathbf{k}_1)\right\}\mathbb{E}^P\left\{W(z,\mathbf{x}_2,\mathbf{k}_2)\right\}
$$

by uniqueness of the solution to the above transport equation with initial conditions given by $W_0(\mathbf{x}_1,\mathbf{k}_1)W_0(\mathbf{x}_2,\mathbf{k}_2)$. This proves that the limiting measure P is deterministic and unique (because characterized by the transport equation) and that the sequence $W_\varepsilon(z,\mathbf{x},\mathbf{k})$ converges in probability to $W(z,\mathbf{x},\mathbf{k})$.

Acknowledgment

This work was partially supported by ONR grants N00014-02-1-0088 and N00014-02-1-0089. G.B. was supported in part by NSF grants DMS-0072008 and DMS-0233549, G.P. by grants AFOSR F49620-01-1-0465 and NSF DMS-9971972, and L.R. in part by NSF grants DMS-9971742 and DMS-0203537 and the Alfred P. Sloan Foundation.

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, Radiative transport limit for the random Schrödinger equation, Nonlinearity, **15**, 2002, 513-529.

[5] A. Bamberger, B. Engquist, L. Halpern and P. Joly, Parabolic wave equation approximations in heterogeneous media, SIAM Journal on Appl. Math., **48**, 1988, 99-128.

[6] C. Bardos and M. Fink, Mathematical foundations of the time-reversal mirror, Asympt. Anal. **29** (2002) 157-182.

[7] J. Berryman, L. Borcea, G. Papanicolaou and C. Tsogka, Statistically stable ultrasonic imaging in random media, Jour. Acoust. Soc. America **112** (2002) 1509-1522.

[8] P. Billingsley, Convergence of probability measures, New York, Wiley, 1968.

[9] P. Blomgren, G. Papanicolaou, and H. Zhao, Super-Resolution in Time-Reversal Acoustics, J. Acoust. Soc. Am., **111**, 2001, 230-248.

[10] L. Borcea, G. Papanicolaou, C. Tsogka and J. Berryman, Imaging and time reversal in random media, Inverse Problems **18** (2002) 1247-1279.

[11] J. F. Clouet and J. P. Fouque, A time-reversal method for an acoustical pulse propagating in randomly layered media, Wave Motion, **25**, 1997, 361–368.

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

[13] L. Erdös and H. T. Yau, Linear Boltzmann equation as the weak coupling limit of a random Schrödinger Equation, Comm. Pure Appl. Math., **53**, 2000, pp. 667–735.

[14] M. Fink, Time reversed acoustics, Physics Today, **50**, 1997, 34–40.

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

[16] 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.

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

[18] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud, Homogenization limits and Wigner transforms, Comm. Pure Appl. Math., **50**, 1997, 323–380.

[19] 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.

[20] P.-L. Lions and T. Paul, Sur les mesures de Wigner, Rev. Mat. Iberoamericana, **9**, 1993, 553–618.

[21] G. Papanicolaou, L. Ryzhik, and K. Solna, Statistical stability in time reversal, Preprint, 2002.

[22] F. Poupaud and A. Vasseur, Classical and Quantum Transport in Random Media, Preprint, 2001.

[23] G. Papanicolaou and S. Weinryb, A functional limit theorem for waves reflected by a random medium, Appl. Math. Optimiz., **30**, 1994, 307-334.

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

[25] H. Spohn, Derivation of the transport equation for electrons moving through random impurities, Jour. Stat. Phys., **17**, 1977, 385–412.

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

[27] L. Tartar, H-measures, a new approach for studying homogenization, oscillations and concentration effects in partial differential equations, Proc. Roy. Soc. Edinburgh, **115A**, 1990, 193-230.

[28] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev., **40**, 1932, 749-759.