

## TIME SPLITTING FOR WAVE EQUATIONS IN RANDOM MEDIA

GUILLAUME BAL<sup>1</sup> AND LENYA RYZHIK<sup>2</sup>

**Abstract.** Numerical simulation of high frequency waves in highly heterogeneous media is a challenging problem. Resolving the fine structure of the wave field typically requires extremely small time steps and spatial meshes. We show that capturing macroscopic quantities of the wave field, such as the wave energy density, is achievable with much coarser discretizations. We obtain such a result using a time splitting algorithm that solves separately and successively propagation and scattering in the simplified regime of the parabolic wave equation in a random medium. The mathematical theory of the convergence and statistical properties of the algorithm is based on the analysis of the Wigner transforms in random media. Our results provide a step toward understanding time and space discretizations that are needed in order for the numerical algorithm to capture the correct macroscopic statistics of the wave energy density in a random medium.

**Mathematics Subject Classification.** 65C50, 65M12, 74J20.

Received: February 2, 2004.

## 1. INTRODUCTION

Wave propagation in highly heterogeneous media has many applications such as light in a turbulent atmosphere, microwaves in wireless communication, acoustic waves in underwater communication, and seismic waves generated by earthquakes [11, 23, 32, 33]. Numerical simulations of wave equations have been thoroughly analyzed; see, *e.g.*, [13, 15] for recent monographs. Most numerical techniques are adapted to the low-to-moderate frequency regime where the size of the calculation domain is not too large compared to the typical wavelength of the system. More recent works consider high frequency wave propagation in the semiclassical regime, which corresponds to the high frequency regime with slowly varying underlying media [8, 26, 27].

Comparatively little is known about high frequency wave propagation in highly heterogeneous media. This is the regime where the wavelength is much smaller than the overall size of the propagation domain while the scale of the variations in the medium is comparable to the wavelength. Such problems arise in many of the aforementioned applications. There, waves fully interact with the medium and the precise numerical simulation of this interaction requires computational capabilities that are usually not available. Our objective is to provide some guidelines as to how to faithfully capture macroscopic properties of the wave propagation numerically at a minimal computational cost. More precisely we wish to devise schemes to solve the wave equation that accurately compute the wave energy density and in particular have the correct statistical properties while not

---

*Keywords and phrases.* High frequency waves in random media, time splitting, multiscale analysis.

<sup>1</sup> Department of Applied Physics & Applied Mathematics, Columbia University, New York, NY 10027, USA.  
e-mail: gb2030@columbia.edu

<sup>2</sup> Department of Mathematics, University of Chicago, Chicago, IL 60637, USA. e-mail: ryzhik@math.uchicago.edu

necessarily capturing small scale structures of the wave field. These small scale details are usually of little importance in practice.

In this paper we restrict ourselves to a setting that is practically useful yet much simpler than the full wave equation so that a complete mathematical theory can be obtained. This setting is the parabolic wave approximation, which is being used and can be justified when the wave field has a beam-like structure [36]. The resulting equation has the same form as a Schrödinger equation with a time-dependent potential. Our theory can thus also be used to consider quantum wave propagation in time dependent media. We solve the parabolic equation using a time-splitting algorithm [35]. The time-splitting consists of separating propagation (in a homogeneous media) from interaction with the inhomogeneities of the medium (scattering). The advantage of the time splitting algorithm is that wave propagation is easily solved in the Fourier domain, at least as far as infinite domain or domains with periodical boundary conditions are concerned. Scattering is also easily solved as an ordinary differential equation at each spatial point. The time splitting algorithm to solve wave equations in this setting is also known as the phase screen method [37, 38]. The “time” splitting algorithm for the Schrödinger equation corresponds to a spatial discretization for the parabolic wave equation. Although our results do not extend mathematically to the full wave equation, the results we obtain for the parabolic approximation still indicate what we may expect in the spatial discretization of the former.

We model the heterogeneous medium by a random potential with the correlation length  $l$  comparable to the typical wavelength  $\lambda$  of the system. Both are much smaller than the propagation distance  $L$  so that  $\varepsilon = l/L = \lambda/L \ll 1$  is a small parameter. The relative strength of the fluctuations of the potential is of the order  $O(\sqrt{\varepsilon})$ . This ensures a full coupling of the wave field with the heterogeneities after propagation over distances of order  $O(L)$ . We consider our problem in non-dimensional coordinates so that the propagation distance  $L = O(1)$ . In the limit  $\varepsilon \rightarrow 0$  the wave energy density solves a radiative transfer equation [5, 11, 31]. This is also known as the weak coupling regime in the mathematical physics literature [16, 34]. In order to capture all the details of the wavefield (in the  $L^2$  sense for instance) the time-splitting algorithm requires that the time step  $\Theta_\varepsilon$  satisfy  $\Theta_\varepsilon \ll \varepsilon^{3/2}$ . When the Strang time splitting is used [35], convergence is ensured provided that  $\Theta_\varepsilon \ll \varepsilon^{5/4}$ . Our main result in this regime is that the time step  $\Theta_\varepsilon \ll \varepsilon$  ensures that the macroscopic wave energy density is captured correctly. More precisely, in the limit  $\varepsilon \rightarrow 0$  the time splitting algorithm converges to the correct energy density of the wave field provided that  $\Theta_\varepsilon \ll \varepsilon$ . Moreover taking the time step  $\Theta_\varepsilon = \Theta\varepsilon$  with  $\Theta > 0$  fixed, we obtain that the limiting wave energy obtained by the time splitting algorithm as  $\varepsilon \rightarrow 0$  solves a radiative transfer equation with the scattering kernel  $R_\Theta(\mathbf{k}, \mathbf{p})$  that depends on the parameter  $\Theta$ . The correct scattering kernel is recovered in the limit  $\Theta \rightarrow 0$  while a scattering kernel that corresponds to white noise fluctuations in the direction of propagation is obtained in the opposite limit of a large time step  $\Theta \rightarrow \infty$ . Therefore unless the white noise spectrum truly corresponds to the fluctuations of the medium, wave energy predicted by the time-splitting scheme solves in the limit  $\varepsilon \rightarrow 0$  a radiative transfer equation with the wrong scattering kernel.

Our theory is based on the analysis of the Wigner transform of the wave field. Wigner transforms have been successfully used recently in the microlocal analysis (analysis in the phase space) of wave fields [5, 19, 25, 30, 31]. It turns out that the Wigner transform also satisfies an evolution equation and the aforementioned time splitting algorithm for the Wigner transform consequently converges to the correct limiting radiative transfer equation provided that  $\Theta_\varepsilon \ll \varepsilon$ . We show that a modification of the time splitting algorithm allows us to replace the latter constraint by the optimal  $\Theta_\varepsilon \ll 1$ . This modification however is only possible because the Wigner transform lives in the phase space. We believe that no such modification is possible directly on the time splitting algorithm for the wave equation. Moreover, the modified time-splitting algorithm does not admit as simple a solution as the original algorithm so that its main advantage is that it allows to bypass the numerically costly advection step in the phase space.

Our objectives are similar to those in works on the commutativity of mesh size convergence and “small parameter” convergence, to a common limiting equation in homogenization problems. For instance in [22] the small parameter is the size of the cell in a periodic domain and the mesh size is that of the spatial discretization. The limiting equation is a homogeneous diffusion equation. In [20] the small parameter is the mean free path in

a transport equation, the mesh size that of the spatial discretization and the limiting equation a homogeneous diffusion equation. As in the aforementioned works [26, 27], we are interested in quantities that are quadratic in the wave field and not the wave field itself. A common feature of all these problems is that the quantities of interest may be well approximated by macroscopic homogenized equations. That makes the existence of “almost macroscopic” numerical schemes quite natural.

The paper is organized as follows. Section 2 briefly describes the passage from the wave equation to the parabolic approximation (a Schrödinger equation) and the precise asymptotic regime we are interested in. It also introduces the time-splitting scheme for the Schrödinger equation. Section 3 introduces the Wigner transform of the wave field and presents our main result of convergence of the wave energy density as the wavelength  $\varepsilon \rightarrow 0$ . To underline the main aspects of the calculation of the statistics of the time splitting scheme we present in this section a formal analysis based on the integral formulation of the Wigner transform. The rigorous justification of that analysis as well as our assumptions on the random medium are contained in Section 4. The modified time splitting algorithm in the phase space is presented in Section 5. Finally we discuss in Section 6 possible extensions of our results to more general equations and discretizations. We also briefly discuss application of our results to the numerical simulation in time reversal experiments.

## 2. PARABOLIC MODEL FOR WAVE PROPAGATION AND TIME SPLITTING

Let us recall how the parabolic approximation is obtained. Wave propagation in an inhomogeneous medium is described by the scalar wave equation for the pressure field  $p(\boldsymbol{\xi}, t)$ :

$$\frac{1}{c^2(\boldsymbol{\xi})} \frac{\partial^2 p}{\partial t^2} = \Delta_{\boldsymbol{\xi}} p. \tag{1}$$

Here  $c(\boldsymbol{\xi})$  is the local sound speed and  $\Delta_{\boldsymbol{\xi}}$  is the usual Laplacian operator in the spatial variable  $\boldsymbol{\xi} \in \mathbb{R}^n$ , where  $n = 2, 3$  in practical applications. We are interested in the numerical simulation of  $p(\boldsymbol{\xi}, t)$  when  $c(\boldsymbol{\xi})$  is highly oscillatory. The profile  $c(\boldsymbol{\xi})$  will be modeled as a given realization of a random medium.

Very little can be obtained rigorously for the full wave equation (1) in a highly heterogeneous medium. Its parabolic wave approximation is valid for beam-like structures. Let  $z$ -axis be the direction of the front propagation, and let  $\mathbf{x}$  be the orthogonal coordinates so that  $\boldsymbol{\xi} = (z, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^d$ , where  $d = 1, 2$  in practical applications. Assuming that back-scattering can be neglected, (1) can be simplified as follows. First we introduce the complex amplitude  $\psi(z, \mathbf{x}; k)$  implicitly through the relation

$$p(z, \mathbf{x}, t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ik(z-c_0t)} \psi(z, \mathbf{x}; k) c_0 dk, \tag{2}$$

where  $c_0$  is the statistical mean of the sound speed  $c(z, \mathbf{x})$  assumed to be constant here. The amplitude  $\psi(z, \mathbf{x}; k)$  at position  $(z, \mathbf{x})$  of waves with frequency  $\omega = c_0 k$  satisfies the equation

$$\frac{\partial^2 \psi}{\partial z^2} + 2ik \frac{\partial \psi}{\partial z} + \Delta_{\mathbf{x}} \psi + k^2(n^2 - 1)\psi = 0, \tag{3}$$

where the index of refraction is defined by  $n(z, \mathbf{x}) = c_0/c(z, \mathbf{x})$ . In the beam approximation one assumes that  $|\psi_{zz}| \ll k|\psi_z|$  which allows us to drop the first term in (3) and arrive at the parabolic equation

$$2ik \frac{\partial \psi}{\partial z} + \Delta_{\mathbf{x}} \psi + k^2(n^2 - 1)\psi = 0. \tag{4}$$

We are interested in the high frequency waves that are characterized by large values of  $kL \gg 1$ , where  $L$  is the overall propagation distance. We introduce the small parameter  $\varepsilon = (kL)^{-1}$  to model the typical wavelength of the system. We assume that the underlying medium has a correlation length also of order  $\varepsilon L$  and an amplitude

of order  $\sqrt{\varepsilon}$ . In this regime waves fully interact with the random medium [31]. The size of the amplitude is suitably chosen so that macroscopic effects can be observed. We rescale  $\mathbf{x}$  and  $z$  as  $L\mathbf{x}$  and  $Lz$ . In these variables and with the above assumptions, the refraction index takes the form

$$(n^2 - 1)(z, \mathbf{x}) \rightarrow -2\sqrt{\varepsilon}V\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right).$$

Here  $V$  is a random potential with statistics independent of  $\varepsilon$ . Equation (4) becomes now dimensionless:

$$\frac{\partial\psi_\varepsilon}{\partial z} = \frac{i\varepsilon}{2}\Delta_{\mathbf{x}}\psi_\varepsilon - \frac{i}{\sqrt{\varepsilon}}V\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right)\psi_\varepsilon. \tag{5}$$

We refer to [36] for additional details on the derivation and validity of the parabolic wave approximation. In the sequel we use (5) as our model of wave propagation to understand the interaction of spatial discretization (here the variable  $z$ ) with an underlying random medium with fast variations.

Observe that (5) has the form of a Schrödinger equation. In the latter the  $z$  variable is replaced by time  $t$  and we obtain in a more conventional form

$$i\varepsilon\frac{\partial\psi_\varepsilon}{\partial t} + \frac{\varepsilon^2}{2}\Delta_{\mathbf{x}}\psi_\varepsilon - \sqrt{\varepsilon}V\left(\frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right)\psi_\varepsilon = 0, \tag{6}$$

where  $\varepsilon$  is the Planck constant. Note that the potential in (6) oscillates both in space  $\mathbf{x}$  and the new “time” variable.

We can thus view the theory below as an analysis of spatial discretization (in  $z$ ) of the parabolic wave equation (5) or as a temporal discretization of the Schrödinger equation (6) with a time dependent potential. To fix notation, we treat the discretized variable as a “time” variable in the sequel.

A very classical idea to solve (6) is based on realizing that the evolution is driven by two processes, one involving dispersion in a homogeneous medium, and one characterizing interaction with the heterogeneities of the medium. One can then advance in time by treating these two processes successively on small intervals. This is the time splitting method [35].

Let  $\Theta_\varepsilon$  be a small time interval. We define the solution  $\psi_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon)$  for all  $n \geq 0$ , as an approximation to  $\psi_\varepsilon(n\Theta_\varepsilon)$  as follows. The initial condition  $\psi_\varepsilon^{\Theta_\varepsilon}(0) = \psi_\varepsilon(0)$  is known. We first solve

$$\begin{aligned} \frac{\partial u_\varepsilon}{\partial t} &= A_\varepsilon u_\varepsilon, & n\Theta_\varepsilon < t < (n+1)\Theta_\varepsilon, & & n \geq 0 \\ u_\varepsilon(n\Theta_\varepsilon) &= \psi_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon), \end{aligned} \tag{7}$$

followed by

$$\begin{aligned} \frac{\partial v_\varepsilon}{\partial t} &= B_\varepsilon(t)v_\varepsilon, & n\Theta_\varepsilon < t < (n+1)\Theta_\varepsilon, & & n \geq 0 \\ v_\varepsilon(n\Theta_\varepsilon) &= u_\varepsilon((n+1)\Theta_\varepsilon), \end{aligned} \tag{8}$$

and finally set

$$\psi_\varepsilon^{\Theta_\varepsilon}((n+1)\Theta_\varepsilon) = v_\varepsilon((n+1)\Theta_\varepsilon), \quad n \geq 0. \tag{9}$$

Here the operators  $A_\varepsilon$  and  $B_\varepsilon$  are defined by

$$A_\varepsilon u = \frac{i\varepsilon}{2}\Delta u, \quad B_\varepsilon(t)u = \frac{-i}{\sqrt{\varepsilon}}V\left(\frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon}\right)u. \tag{10}$$

The main advantage of such a decomposition is that both (7) and (8) are easily solved since the former involves propagation in a homogeneous medium and the latter scattering at a fixed point in space. The solutions can be

written as

$$\begin{aligned}
 u_\varepsilon((n+1)\Theta_\varepsilon) &= e^{\Theta_\varepsilon A_\varepsilon} u_\varepsilon(n\Theta_\varepsilon) \\
 v_\varepsilon((n+1)\Theta_\varepsilon) &= \exp \left\{ \int_{n\Theta_\varepsilon}^{(n+1)\Theta_\varepsilon} B_\varepsilon(\tau) d\tau \right\} v_\varepsilon(n\Theta_\varepsilon).
 \end{aligned}
 \tag{11}$$

Let us introduce the notation

$$B_\varepsilon^n = \frac{1}{\Theta_\varepsilon} \int_{n\Theta_\varepsilon}^{(n+1)\Theta_\varepsilon} B_\varepsilon(\tau) d\tau.$$

The local difference between the time splitting algorithm and the exact solution is expressed as

$$\psi_\varepsilon((n+1)\Theta_\varepsilon) = e^{\Theta_\varepsilon(A_\varepsilon + B_\varepsilon^n)} \psi_\varepsilon(n\Theta_\varepsilon), \tag{12}$$

$$\psi_\varepsilon^{\Theta_\varepsilon}((n+1)\Theta_\varepsilon) = e^{\Theta_\varepsilon B_\varepsilon^n} e^{\Theta_\varepsilon A_\varepsilon} \psi_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon). \tag{13}$$

Since  $A_\varepsilon$  and  $B_\varepsilon$  do not commute the above two evolutions are different. It is well known [35] that the time splitting algorithm introduces an error of the order

$$[A_\varepsilon, B_\varepsilon^n] \Theta_\varepsilon^2$$

at each time step, where  $[A, B] = AB - BA$  is the commutator as can be seen from Taylor expansions in (12) for small values of  $\Theta_\varepsilon$ . A straightforward estimates shows that  $[A_\varepsilon, B_\varepsilon^n] = O(\varepsilon^{-3/2})$ . Thus the final (cumulative) error at a time  $T = n\Theta_\varepsilon$  of order  $O(1)$  is of order

$$[A_\varepsilon, B_\varepsilon^n] \Theta_\varepsilon = O\left(\varepsilon^{-3/2}\right) \Theta_\varepsilon. \tag{14}$$

This implies that we need to consider time steps  $\Theta_\varepsilon \ll \varepsilon^{3/2}$  to simply get convergence of the time splitting algorithm.

The above time splitting algorithm can be improved by using the Strang algorithm [35], which consists in replacing the discrete evolution operator in (13) by

$$\psi_\varepsilon^{\Theta_\varepsilon}((n+1)\Theta_\varepsilon) = e^{\frac{\Theta_\varepsilon}{2} A_\varepsilon} e^{\Theta_\varepsilon B_\varepsilon^n} e^{\frac{\Theta_\varepsilon}{2} A_\varepsilon} \psi_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon). \tag{15}$$

Except for an initialization step, this algorithm is as costly as the time splitting algorithm we have just introduced since  $e^{\frac{\Theta_\varepsilon}{2} A_\varepsilon} e^{\frac{\Theta_\varepsilon}{2} A_\varepsilon} = e^{\Theta_\varepsilon A_\varepsilon}$ . The advantage of this algorithm is that the commutator (14) vanishes in the Taylor expansion comparing (12) and (15). The leading term of the error made at every time step in our context is of the form

$$A_\varepsilon^2 B_\varepsilon^n \Theta_\varepsilon^3 = O\left(\varepsilon^{-5/2}\right) \Theta_\varepsilon^3.$$

At the final time  $T$  the error is thus of the order  $O(\varepsilon^{-5/2}) \Theta_\varepsilon^2$  which implies that we need  $\Theta_\varepsilon \ll \varepsilon^{5/4}$  in order to get convergence of the scheme.

If the detailed structure of the field  $\psi_\varepsilon(t, \mathbf{x})$  is to be resolved, then the above constraints on the interval size  $\Theta_\varepsilon$  are optimal. However we will show that  $\Theta_\varepsilon \ll \varepsilon$  is sufficient if one is interested in getting the correct macroscopic energy density of the waves and not all the details of the wave field. Moreover we will show that  $\Theta_\varepsilon \ll 1$  is actually sufficient if one uses a time splitting algorithm for the Wigner transform, which is posed in the phase space, of the wave field, instead of (7)–(9). However, with the time-step  $O(1)$  the modification of the operator  $B_\varepsilon$  defined in phase space does not admit a simple solution and equation (8) has to be solved by a different numerical method.

### 3. PHASE SPACE ENERGY AND FORMAL HIGH FREQUENCY LIMIT

We now describe what we mean by macroscopic energy density of waves and how we can compute it numerically. The dynamics of high frequency waves propagating in slowly varying media can be approximated by WKB type expansions of the form

$$\psi_\varepsilon(t, \mathbf{x}) = A(t, \mathbf{x})e^{iS(t, \mathbf{x})/\varepsilon} + O(\varepsilon), \tag{16}$$

where the phase  $S$  and amplitude  $A$  are obtained by solving an eikonal and a transport equation, respectively [24]. In practice, the ansatz (16) is often sufficient for an accurate description of the wave field. A natural question is then to understand the size of  $\Theta_\varepsilon$  in the time splitting algorithm that ensures that  $A$  and  $S$  are well approximated by the algorithm instead of the full wave field  $\psi_\varepsilon$ .

However, the above expansion (16) is essentially useless in random media. The reason is that the wave field cannot be represented by *one* generalized plane wave of the form (16). Instead an infinite number of them is required because of multiple scattering. The theory of Wigner transforms in the phase space is a convenient way to overcome this difficulty and resolve the wave energy locally over all possible directions. The Wigner transform is defined as follows

$$W_\varepsilon(t, \mathbf{x}, \mathbf{k}) = \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{y}} \psi_\varepsilon\left(t, \mathbf{x} - \frac{\varepsilon\mathbf{y}}{2}\right) \psi_\varepsilon^*\left(t, \mathbf{x} + \frac{\varepsilon\mathbf{y}}{2}\right) \frac{d\mathbf{y}}{(2\pi)^d}. \tag{17}$$

Here  $*$  denotes complex conjugation. The Wigner transform is thus the Fourier transform of the two-point correlation function of the field  $\psi_\varepsilon$ . It roughly describes the energy density of waves at time  $t$  and position  $\mathbf{x}$  propagating with wave vector  $\mathbf{k}/\varepsilon$ .

Wigner transforms have been used intensively in recent years to understand high frequency wave propagation in highly oscillatory periodic and random media [4, 19, 25, 31]. The spatial energy density (probability density for quantum waves) is the integral of the Wigner transform over the wave vectors:

$$|\psi_\varepsilon|^2(t, \mathbf{x}) = \int_{\mathbb{R}^d} W_\varepsilon(t, \mathbf{x}, \mathbf{k}) d\mathbf{k}. \tag{18}$$

Our objective is therefore to understand how we can capture the main features of the Wigner transform (17) and then recover the physical space energy *via* (18). More precisely we want to obtain the critical size of  $\Theta_\varepsilon$  so that the statistical ensemble average  $\mathbb{E}\{W_\varepsilon\}(t, \mathbf{x}, \mathbf{k})$  has the correct limit as  $\varepsilon \rightarrow 0$ . This allows us to describe the statistics of the spatial energy density without capturing the detailed structure of  $\psi_\varepsilon$ .

The rigorous analysis of the Wigner transform (17) of wave fields  $\psi_\varepsilon$  that satisfy (6) has been performed in [5, 6] using the martingale technique under the assumption that  $V$  is a mean zero spatially homogeneous random potential with a two-point correlation function  $R(t, \mathbf{x})$ :

$$\mathbb{E}\{V(t, \mathbf{x})\} = 0, \quad \mathbb{E}\{V(t, \mathbf{x})V(t + s, \mathbf{x} + \mathbf{y})\} = R(s, \mathbf{y}). \tag{19}$$

It has been shown that in the limit  $\varepsilon \rightarrow 0$  the Wigner transform converges in probability to the solution of the radiative transport equation

$$\frac{\partial \bar{W}}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} \bar{W} = \int_{\mathbb{R}^d} \hat{R}\left(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}, \mathbf{p} - \mathbf{k}\right) (\bar{W}(t, \mathbf{x}, \mathbf{p}) - \bar{W}(t, \mathbf{x}, \mathbf{k})) \frac{d\mathbf{p}}{(2\pi)^d}. \tag{20}$$

Here  $\hat{R}$ , the Fourier transform of the two-point correlation function, is the power spectrum of the random process  $V$ . We extend this analysis to the case of discretized equations in time in Section 4. The main result is that with the time-step  $\Theta_\varepsilon = \Theta\varepsilon$  the Wigner transform  $W_\varepsilon^\xi$  of the solution of the time-splitting scheme (7)–(9)

converges as  $\varepsilon \rightarrow 0$  to the solution of a radiative transport equation of the form

$$\frac{\partial \bar{W}^\Theta}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} \bar{W}^\Theta = \int_{\mathbb{R}^d} R_\Theta(\mathbf{k}, \mathbf{p})(\bar{W}^\Theta(t, \mathbf{x}, \mathbf{p}) - \bar{W}^\Theta(t, \mathbf{x}, \mathbf{k})) \frac{d\mathbf{p}}{(2\pi)^d}. \tag{21}$$

The new scattering kernel  $R_\Theta(\mathbf{k}, \mathbf{p})$  is given by (46) below. Before proceeding with a rigorous proof, we present in the rest of this section a formal but shorter path to the answer.

**3.1. Radiative transfer equation in the continuous case**

First we formally derive the radiative transfer equation obtained in [5,6] that  $W_\varepsilon$  satisfies in the limit  $\varepsilon \rightarrow 0$ . This is helpful in the understanding of the formal analysis of the semi-classical limit of the Wigner transform of the solution of the time-splitting scheme (7)–(9) presented in Section 3.2.

We recall that the Wigner transform (17) solves the following equation [6, 31]

$$\frac{\partial W_\varepsilon}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}} \int_{\mathbb{R}^d} e^{i\mathbf{p}\cdot\mathbf{x}/\varepsilon} \tilde{V}\left(\frac{t}{\varepsilon}, \mathbf{p}\right) \left[ W\left(\mathbf{k} - \frac{\mathbf{p}}{2}\right) - W\left(\mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \frac{d\mathbf{p}}{(2\pi)^d}. \tag{22}$$

Here and in the sequel we denote the Fourier transform only in space by

$$\tilde{f}(t, \mathbf{p}) = \int_{\mathbb{R}^d} e^{-i\mathbf{p}\cdot\mathbf{x}} f(t, \mathbf{x}) d\mathbf{x}, \tag{23}$$

and the Fourier transform both in  $t$  and  $\mathbf{x}$  by

$$\hat{f}(t, \mathbf{p}) = \int_{\mathbb{R}^{d+1}} e^{-i\omega t - i\mathbf{p}\cdot\mathbf{x}} f(t, \mathbf{x}) dt d\mathbf{x}. \tag{24}$$

Equation (22) may be recast as

$$\frac{\partial W_\varepsilon}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} W_\varepsilon = \int_{\mathbb{R}^d} K_\varepsilon(t, \mathbf{x}, \mathbf{k} - \mathbf{p}) W_\varepsilon(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}, \tag{25}$$

where  $K_\varepsilon$  is given by

$$K_\varepsilon(t, \mathbf{x}, \mathbf{q}) = \frac{1}{i\pi^d \sqrt{\varepsilon}} \left[ \tilde{V}\left(\frac{t}{\varepsilon}, 2\mathbf{q}\right) e^{i2\mathbf{q}\cdot\mathbf{x}/\varepsilon} - \tilde{V}\left(\frac{t}{\varepsilon}, -2\mathbf{q}\right) e^{-i2\mathbf{q}\cdot\mathbf{x}/\varepsilon} \right]. \tag{26}$$

Inverting the free transport operator  $\partial_t + \mathbf{k} \cdot \nabla_{\mathbf{x}}$  we recast (25) as an integral equation

$$W_\varepsilon(\tau, \mathbf{x}, \mathbf{k}) = W_\varepsilon(0, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) + \int_0^\tau \int_{\mathbb{R}^d} K_\varepsilon(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k} - \mathbf{p}) W_\varepsilon(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p}) d\mathbf{p} ds. \tag{27}$$

One more iteration yields

$$\begin{aligned} W_\varepsilon(\tau, \mathbf{x}, \mathbf{k}) &= W_\varepsilon(0, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) + \int_0^\tau \int_{\mathbb{R}^d} K_\varepsilon(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k} - \mathbf{p}) W_\varepsilon(0, \mathbf{x} - s\mathbf{k} - (\tau - s)\mathbf{p}, \mathbf{p}) d\mathbf{p} ds \\ &+ \int_0^\tau \int_{\mathbb{R}^d} K_\varepsilon(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k} - \mathbf{p}) \int_0^{\tau-s} \int_{\mathbb{R}^d} K_\varepsilon(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, \mathbf{p} - \mathbf{q}) \\ &\times W_\varepsilon(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, \mathbf{q}) d\mathbf{q} du d\mathbf{p} ds. \end{aligned} \tag{28}$$

Let us now consider ensemble averaging in the above equation. We assume that the initial condition  $W_\varepsilon(t = 0)$  is deterministic and that the potential  $V$  has mean zero. This implies that the second term on the right-hand

side of (28) vanishes after ensemble averaging. To deal with the third term we assume that  $W_\varepsilon$  is statistically independent of the rest of the term. This step cannot be justified at this stage but it does provide the correct answer in the limit  $\varepsilon \rightarrow 0$ . By doing so we obtain that

$$\begin{aligned} \mathbb{E} \{W_\varepsilon\}(\tau, \mathbf{x}, \mathbf{k}) &= W_\varepsilon(0, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) \\ &+ \int_0^\tau \int_0^{\tau-s} \int_{\mathbb{R}^{2d}} \mathbb{E} \{K_\varepsilon(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k} - \mathbf{p})K_\varepsilon(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, \mathbf{p} - \mathbf{q})\} \\ &\times \mathbb{E} \{W_\varepsilon\}(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, \mathbf{q})d\mathbf{q}du d\mathbf{p}ds. \end{aligned} \tag{29}$$

We now have to compute the ensemble average of the product of two functions  $K_\varepsilon$ . Taking the Fourier transform of (19) in  $\mathbf{x} \rightarrow \mathbf{p}$  yields

$$\mathbb{E} \left\{ \tilde{V}(t, \mathbf{p})\tilde{V}(t + s, \mathbf{q}) \right\} = (2\pi)^d \tilde{R}(s, \mathbf{p})\delta(\mathbf{p} + \mathbf{q}). \tag{30}$$

Straightforward algebra shows that

$$\begin{aligned} \mathbb{E} \{K_\varepsilon(t, \mathbf{y}, \mathbf{k} - \mathbf{p})K_\varepsilon(t - u, \mathbf{y} - u\mathbf{p}, \mathbf{p} - \mathbf{q})\} &= \\ &\frac{1}{\pi^d \varepsilon} \tilde{R}\left(\frac{u}{\varepsilon}, 2(\mathbf{p} - \mathbf{k})\right) \left( e^{2iu(\mathbf{k}-\mathbf{p})\cdot\mathbf{p}/\varepsilon} + e^{-2iu(\mathbf{k}-\mathbf{p})\cdot\mathbf{p}/\varepsilon} \right) (\delta(\mathbf{k} + \mathbf{q} - 2\mathbf{p}) - \delta(\mathbf{k} - \mathbf{q})). \end{aligned} \tag{31}$$

Thus the ensemble average of the last term in (29) becomes

$$\begin{aligned} &\int_0^\tau \int_0^{\tau-s} \int_{\mathbb{R}^d} \tilde{R}\left(\frac{u}{\varepsilon}, 2(\mathbf{p} - \mathbf{k})\right) \left( e^{2iu(\mathbf{p}-\mathbf{k})\cdot\mathbf{p}/\varepsilon} + e^{-2iu(\mathbf{p}-\mathbf{k})\cdot\mathbf{p}/\varepsilon} \right) \\ &\times \left( \mathbb{E} \{W_\varepsilon\}(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, 2\mathbf{p} - \mathbf{k}) - \mathbb{E} \{W_\varepsilon\}(\tau - s - u, \mathbf{x} - s\mathbf{k} - u\mathbf{p}, \mathbf{k}) \right) \frac{d\mathbf{p}du ds}{\pi^d \varepsilon}. \end{aligned}$$

After the change of variables  $2\mathbf{p} - \mathbf{k} \rightarrow \mathbf{p}$  and  $u \rightarrow \varepsilon u$ , we get

$$\begin{aligned} &\int_0^\tau \int_0^{(\tau-s)/\varepsilon} \int_{\mathbb{R}^d} \frac{d\mathbf{p}du ds}{(2\pi)^d} \tilde{R}(u, \mathbf{p} - \mathbf{k}) \left( e^{iu\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}} + e^{-iu\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}} \right) \\ &\times \left( \mathbb{E} \{W_\varepsilon\} \left( \tau - s - \varepsilon u, \mathbf{x} - s\mathbf{k} - \varepsilon u\frac{\mathbf{p} + \mathbf{k}}{2}, \mathbf{p} \right) - \mathbb{E} \{W_\varepsilon\} \left( \tau - s - \varepsilon u, \mathbf{x} - s\mathbf{k} - \varepsilon u\frac{\mathbf{p} + \mathbf{k}}{2}, \mathbf{k} \right) \right). \end{aligned} \tag{32}$$

Assuming that  $W_\varepsilon$  is smooth in the limit  $\varepsilon \rightarrow 0$  we thus obtain that (32) in the limit  $\varepsilon \rightarrow 0$  yields

$$\int_0^\tau \int_{\mathbb{R}^d} \hat{R}\left(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}, \mathbf{p} - \mathbf{k}\right) (\mathbb{E} \{W\}(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p}) - \mathbb{E} \{W\}(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k})) \frac{d\mathbf{p}ds}{(2\pi)^d}, \tag{33}$$

where  $\hat{R}(\omega, \mathbf{p})$  is the power spectrum of the random potential  $V$

$$\hat{R}(\omega, \mathbf{p}) = \int_{-\infty}^\infty e^{-i\omega u} \tilde{R}(\mathbf{p}, u)du. \tag{34}$$

We replace the right side of (29) by (33) and obtain the integral equation for  $\overline{W}$ , the limit of  $\mathbb{E} \{W_\varepsilon\}$  as  $\varepsilon \rightarrow 0$ :

$$\overline{W}(\tau, \mathbf{x}, \mathbf{k}) = \overline{W}(0, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) + \int_0^\tau \int_{\mathbb{R}^d} \hat{R}\left(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}, \mathbf{p} - \mathbf{k}\right) (\overline{W}(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p}) - \overline{W}(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k})) \frac{d\mathbf{p}ds}{(2\pi)^d}. \tag{35}$$

Equation (35) is nothing but the integral form of the transport equation (20). We thus obtain that in the high frequency limit (as  $\varepsilon \rightarrow 0$ ) the average wave energy density is given by the solution to the macroscopic equation (20) that no longer involves the small parameter  $\varepsilon$ .

**3.2. Radiative transfer equation for the time splitting algorithm**

We now wish to find the largest possible size of  $\Theta_\varepsilon$  so that the Wigner transform of  $\psi_\varepsilon^{\Theta_\varepsilon}$ , solution of the time-splitting scheme (12)–(13) is an accurate approximation of (20) in the limit  $\varepsilon \rightarrow 0$ . We construct the Wigner transform of  $\psi_\varepsilon^{\Theta_\varepsilon}$  as follows. For a time  $n\Theta_\varepsilon$  we define

$$W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^-, \mathbf{x}, \mathbf{k}) = \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{y}} \psi_\varepsilon^{\Theta_\varepsilon} \left( n\Theta_\varepsilon, \mathbf{x} - \frac{\varepsilon\mathbf{y}}{2} \right) \psi_\varepsilon^{\Theta_\varepsilon*} \left( n\Theta_\varepsilon, \mathbf{x} + \frac{\varepsilon\mathbf{y}}{2} \right) \frac{d\mathbf{y}}{(2\pi)^d}. \tag{36}$$

We then define

$$W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^+, \mathbf{x}, \mathbf{k}) = W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^-, \mathbf{x} - \Theta_\varepsilon\mathbf{k}, \mathbf{k}). \tag{37}$$

This corresponds to solving the Wigner equation (25) with  $K_\varepsilon = 0$  for a time  $\Theta_\varepsilon$  and with initial data  $W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^-, \mathbf{x}, \mathbf{k})$ . We verify that

$$W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^+, \mathbf{x}, \mathbf{k}) = \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{y}} u_\varepsilon \left( (n+1)\Theta_\varepsilon, \mathbf{x} - \frac{\varepsilon\mathbf{y}}{2} \right) u_\varepsilon^* \left( (n+1)\Theta_\varepsilon, \mathbf{x} + \frac{\varepsilon\mathbf{y}}{2} \right) \frac{d\mathbf{y}}{(2\pi)^d},$$

where  $u_\varepsilon$  is the solution of (7). Finally we assume that  $W_\varepsilon^{\Theta_\varepsilon}$  satisfies the equation

$$\frac{\partial W_\varepsilon^{\Theta_\varepsilon}}{\partial t} = \int_{\mathbb{R}^d} K_\varepsilon(t, \mathbf{x}, \mathbf{k} - \mathbf{p}) W_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}, \tag{38}$$

for  $t \in (n\Theta_\varepsilon^+, (n+1)\Theta_\varepsilon^-)$ . We then verify that

$$W_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{k}) = \int_{\mathbb{R}^d} e^{i\mathbf{k}\cdot\mathbf{y}} v_\varepsilon \left( t, \mathbf{x} - \frac{\varepsilon\mathbf{y}}{2} \right) v_\varepsilon^* \left( t, \mathbf{x} + \frac{\varepsilon\mathbf{y}}{2} \right) \frac{d\mathbf{y}}{(2\pi)^d},$$

on the interval  $t \in (n\Theta_\varepsilon, (n+1)\Theta_\varepsilon)$ , where  $v_\varepsilon$  solves (8). Therefore  $W_\varepsilon^{\Theta_\varepsilon}((n+1)\Theta_\varepsilon^-)$  is indeed given by (36) where  $n$  is replaced by  $n+1$ .

This is the interpretation of the time splitting algorithm for the Wigner transform of the discretized in time wave field  $\psi_\varepsilon^{\Theta_\varepsilon}$ . We note that the above procedure (37)–(38) also corresponds to the time-splitting scheme for the Wigner equation (22) with the operator  $A = \mathbf{k} \cdot \nabla_{\mathbf{x}}$  and the operator  $B$  defined by the right side of (22). We now show formally that  $W_\varepsilon^{\Theta_\varepsilon}$  satisfies an integral equation similar to (27). We first define

$$[|t|]_\mu = |t| \bmod \mu, \quad [t]_\mu = \text{sgn}(t)[|t|]_\mu, \tag{39}$$

so that for  $t > 0$  we have  $t = [t]_\mu + \eta$ , with  $0 \leq \eta < \mu$ . We next define

$$[t, s]_\mu = [t]_\mu - [t - s]_\mu. \tag{40}$$

In what follows it is useful to notice that  $[t, s]_\mu$  is not very different from  $s$  when  $\mu$  is small. The integral formulation of (38) is

$$\begin{aligned} W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon + s, \mathbf{x}, \mathbf{k}) &= W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^+, \mathbf{x} - s\mathbf{k}, \mathbf{k}) \\ &+ \int_0^s \int_{\mathbb{R}^d} K_\varepsilon(n\Theta_\varepsilon + s - u, \mathbf{x}, \mathbf{k} - \mathbf{p}) W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon + s - u, \mathbf{x}, \mathbf{p}) d\mathbf{p} du, \end{aligned} \tag{41}$$

for  $0 < s < \Theta_\varepsilon$ . The difference with respect to (27) is that the spatial position  $\mathbf{x}$  is not modified by advection. We can then use (37) to replace  $W_\varepsilon^{\Theta_\varepsilon}(s\Theta_\varepsilon^+, \mathbf{x} - \tau\mathbf{k}, \mathbf{k})$  in (42) and then use again (41) on the interval  $((n-1)\Theta_\varepsilon, n\Theta_\varepsilon)$  and keep repeating the same process downwards until time  $t = 0$ . This yields

$$W_\varepsilon^{\Theta_\varepsilon}(\tau, \mathbf{x}, \mathbf{k}) = W_\varepsilon^{\Theta_\varepsilon}(0^+, \mathbf{x} - [\tau]_{\Theta_\varepsilon}\mathbf{k}, \mathbf{k}) + \int_0^\tau ds \int_{\mathbb{R}^d} d\mathbf{p} K_\varepsilon(\tau - s, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k}, \mathbf{k} - \mathbf{p}) W_\varepsilon^{\Theta_\varepsilon}(\tau - s, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k}, \mathbf{p}). \tag{42}$$

We can replace  $W_\varepsilon^{\Theta_\varepsilon}$  in the integral term above by its expression in (42) and then take ensemble average of all terms in the equation. With the same assumptions that led to (29) we get

$$\begin{aligned} \mathbb{E}\{W_\varepsilon^{\Theta_\varepsilon}\}(\tau, \mathbf{x}, \mathbf{k}) &= W_\varepsilon^{\Theta_\varepsilon}(0^+, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) + \int_0^\tau \int_0^{\tau-s} \int_{\mathbb{R}^{2d}} \mathbb{E}\{K_\varepsilon(\tau - s, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k}, \mathbf{k} - \mathbf{p}) \\ &\quad \times K_\varepsilon(\tau - s - u, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon}\mathbf{p}, \mathbf{p} - \mathbf{q})\} \\ &\quad \times \mathbb{E}\{W_\varepsilon^{\Theta_\varepsilon}\}(\tau - s - u, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon}\mathbf{p}, \mathbf{q}) d\mathbf{q} du d\mathbf{p} ds. \end{aligned} \tag{43}$$

Following the same calculations as in the continuous case we deduce that the second term on the right-hand side in the above equation is given by

$$\begin{aligned} &\int_0^\tau \int_0^{\tau-s} \int_{\mathbb{R}^d} \frac{d\mathbf{p} du ds}{(2\pi)^d} \frac{1}{\varepsilon} \tilde{R}\left(\frac{u}{\varepsilon}, \mathbf{p} - \mathbf{k}\right) \left( e^{i[\tau-s, u]_{\Theta_\varepsilon} \frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2\varepsilon}} + e^{-i[\tau-s, u]_{\Theta_\varepsilon} \frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2\varepsilon}} \right) \\ &\quad \times \left( \mathbb{E}\{W_\varepsilon^{\Theta_\varepsilon}\}(\tau - [\tau, s]_{\Theta_\varepsilon} - [\tau - s, u]_{\Theta_\varepsilon}, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon}\mathbf{p}, \mathbf{p}) \right. \\ &\quad \left. - \mathbb{E}\{W_\varepsilon^{\Theta_\varepsilon}\}(\tau - [\tau, s]_{\Theta_\varepsilon} - [\tau - s, u]_{\Theta_\varepsilon}, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon}\mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon}\mathbf{p}, \mathbf{k}) \right). \end{aligned} \tag{44}$$

Let us assume that  $\Theta_\varepsilon = \varepsilon\Theta$  and that  $W_\varepsilon^{\Theta_\varepsilon}$  has a smooth limit so it does not oscillate at the  $\varepsilon$  scale. Moreover we realize that as  $\varepsilon \rightarrow 0$ ,  $[\tau, s]_{\Theta_\varepsilon} \rightarrow s$  and change variables  $u/\varepsilon \rightarrow u$ . This allows us to obtain that (44) in the limit  $\varepsilon \rightarrow 0$  converges to

$$\int_0^\tau \int_{\mathbb{R}^d} d\mathbf{p} \hat{R}_\Theta(\mathbf{k}, \mathbf{p}) (W^\Theta(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k}) - W^\Theta(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p})) d\mathbf{p}, \tag{45}$$

where the power spectrum  $\hat{R}_\Theta$  is given by

$$\hat{R}_\Theta(\mathbf{k}, \mathbf{p}) = \int_0^\infty du \frac{1}{\Theta} \int_0^\Theta d\tau \tilde{R}(u, \mathbf{p} - \mathbf{k}) \left( e^{i[u+\tau]_\Theta \frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}} + e^{-i[u+\tau]_\Theta \frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}} \right). \tag{46}$$

We insert (45) into the right side of (43) and obtain in the limit  $\varepsilon \rightarrow 0$  that  $\mathbb{E}\{W_\varepsilon^{\Theta_\varepsilon}\}$  converges to  $\bar{W}^\Theta$ , solution of

$$\bar{W}^\Theta(\tau, \mathbf{x}, \mathbf{k}) = \bar{W}^\Theta(0, \mathbf{x} - \tau\mathbf{k}, \mathbf{k}) + \int_0^\tau \int_{\mathbb{R}^d} d\mathbf{p} \hat{R}_\Theta(\mathbf{k}, \mathbf{p}) (\bar{W}^\Theta(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k}) - \bar{W}^\Theta(\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p})) d\mathbf{p}. \tag{47}$$

This is the integral form of the radiative transport equation

$$\frac{\partial \bar{W}^\Theta}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} \bar{W}^\Theta = \int_{\mathbb{R}^d} \hat{R}_\Theta(\mathbf{k}, \mathbf{p}) [\bar{W}^\Theta(t, \mathbf{x}, \mathbf{p}) - \bar{W}^\Theta(t, \mathbf{x}, \mathbf{k})] \frac{d\mathbf{p}}{(2\pi)^d}. \tag{48}$$

This equation is nothing but (20) with the exact power spectrum  $\hat{R}$  replaced by its modification  $\hat{R}_\Theta$ . Therefore  $W_\varepsilon^{\Theta_\varepsilon}$  will have the correct limit  $\mathbb{E}\{W\}$  provided that the power spectrum  $\hat{R}_\Theta$  is a good approximation of  $\hat{R}$ .

It is not difficult to check that

$$\hat{R}_\Theta(\mathbf{k}, \mathbf{p}) = \hat{R}\left(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}, \mathbf{p} - \mathbf{k}\right) + O(\Theta) \tag{49}$$

as  $\Theta \rightarrow 0$  since  $[u + \tau]_\Theta = u + O(\Theta)$ . Therefore in this limit we recover the correct power spectrum, as in the radiative transport equation (20). The other limit  $\Theta \rightarrow \infty$  is also quite interesting. One finds that

$$\hat{R}_\Theta(\mathbf{k}, \mathbf{p}) = \hat{R}(0, \mathbf{p} - \mathbf{k}) + O\left(\frac{1}{\Theta}\right). \tag{50}$$

The scaling  $\Theta_\varepsilon = \varepsilon\Theta$  thus completely characterizes the quality of the time splitting discretization as far as getting the correct macroscopic energy limit is concerned. When  $\Theta_\varepsilon \ll \varepsilon$  we obtain the correct energy limit with an error of order  $\Theta_\varepsilon\varepsilon^{-1}$ . When  $\Theta_\varepsilon \sim \varepsilon$  we get a transport equation with the wrong power spectrum. When  $\Theta_\varepsilon \gg \varepsilon$  we also obtain a transport equation in the limit but with a *white noise* power spectrum. Indeed the power spectrum  $\hat{R}(0, \mathbf{p} - \mathbf{k})$  is the Fourier transform of a two point correlation function

$$\mathbb{E}\{V(t, \mathbf{x})V(t + s, \mathbf{x} + \mathbf{y})\} = R(\mathbf{y})\delta(s),$$

which means that potentials at different times are totally uncorrelated. This result does not come totally as a surprise. In the time splitting algorithm averaging occurs at one spatial position in (8) since advection is shut off. When the interval over which this averaging takes place is long, correlations with other parts of the system are lost and one converges to a power spectrum corresponding to fluctuations with full decorrelation in time.

#### 4. THE HIGH-FREQUENCY LIMIT OF THE TIME-SPLITTING SCHEME

In this section we analyze rigorously the passage to the limit  $\varepsilon \rightarrow 0$  in the time-splitting scheme (7)–(9). The analysis is performed in the framework of the time-splitting scheme (37)–(38) for the Wigner equation that is induced by (7)–(9). In particular we justify the modified radiative equation (48) in the small  $\varepsilon$  limit.

The method of proof is based on the ideas of [5, 6]. In those works two types of convergences are considered depending on the *a priori* bounds the Wigner transform  $W(t, \mathbf{x}, \mathbf{k})$  satisfies. For the Wigner transform of a *pure state* such as defined by (17) the *a priori* bound is in  $\mathcal{A}'(\mathbb{R}^{2d})$  uniformly in time, where  $\mathcal{A}'(\mathbb{R}^{2d})$  is a distribution space bigger than the space of bounded measures [5, 25]. We can also assume that the initial data for the Wigner transform has a strong limit in the smaller space  $L^2(\mathbb{R}^{2d})$ , which implies an *a priori* bound in the same space uniformly in time for the Wigner transform. Such a bound can be obtained by considering a *mixture of states*, which corresponds to taking the initial data  $\psi_\varepsilon$  for the Schrödinger equation random and suitably averaging over this randomness first. The  $L^2$  setting also appears naturally in the mathematical theory of time reversal experiments (see [5, 6] for details and our conclusions section).

In this section we assume the latter  $L^2$  *a priori* bound for two reasons. First the proofs are simpler in this setting and second we show that not only does the average Wigner transform converge to the solution of (48), but it actually converges to its deterministic limit in probability. This means that it is a self-averaging quantity essentially independent of the realization of the random medium. This important property in time reversal applications has received some attention recently [1, 6, 29]. When the main object of interest is evolution of a pure state then the limit Wigner transform does not belong to  $L^2$ . Theorem 4.1 below still holds but only in the sense that the expectation of the Wigner transform converges to the solution of (48). The proof is similar to what we present below with modifications as in [5].

##### 4.1. Assumptions on the random potential

We begin with the assumptions on the random potential  $V(t, \mathbf{x})$ . We assume that the random field  $V(t, \mathbf{x})$  is a Markov process in  $t$ , as in [5, 6]. 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 and to apply the martingale method. We will also make a number of assumptions on the regularity of  $V$  that allow to simplify the technicalities to a great extent. Unlike the Markovianity of  $V$  those assumptions may be significantly weakened without major changes in the proof.

The random process  $V(t, \mathbf{x})$  is assumed to be stationary in  $\mathbf{x}$  and  $z$  and have mean zero. It 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  $B_L = \{|\mathbf{p}| \leq L\}$

$$\mathcal{V} = \left\{ \hat{V} : \int_{\mathbb{R}^d} |d\hat{V}| \leq C, \text{ supp } \hat{V} \subset B_L, \hat{V}(\mathbf{p}) = \hat{V}^*(-\mathbf{p}) \right\} \tag{51}$$

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

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

and is real and uniformly bounded:

$$|V(t, \mathbf{x})| \leq C.$$

Moreover, all spatial derivatives of  $V$  are also bounded by deterministic constants. We assume that the process  $V(t, \mathbf{x})$  is stationary in  $t$  and  $\mathbf{x}$  with the correlation function  $R(t, \mathbf{x})$

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

and power spectrum  $\hat{R}(\omega, \mathbf{p})$ :

$$\hat{R}(\omega, \mathbf{p}) = \int_{\mathbb{R}} \int_{\mathbb{R}^d} e^{-i\omega t - i\mathbf{p} \cdot \mathbf{z}} R(t, \mathbf{z}) d\mathbf{z} dt.$$

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

$$\mathbb{E} \left\{ \langle \tilde{V}(s), \hat{\phi} \rangle \langle \tilde{V}(t + s), \hat{\psi} \rangle \right\} = (2\pi)^d \int_{\mathbb{R}^d} d\mathbf{p} \tilde{R}(t, \mathbf{p}) \hat{\phi}(\mathbf{p}) \hat{\psi}(-\mathbf{p}). \tag{52}$$

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  $R(t, \mathbf{x})$  in  $\mathbf{x}$  only. We assume that  $R(t, \mathbf{x}) \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$  for simplicity.

We assume that the generator  $Q$  of the Markov process  $\tilde{V}(t)$  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

$$\|e^{rQ} g\|_{L^\infty} \leq C \|g\|_{L^\infty} e^{-\alpha r}. \tag{53}$$

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

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

Given (53), 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  $\|f\|_{L^\infty} \leq C \|g\|_{L^\infty}$ . The solution  $f$  is given explicitly by

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

and the integral converges absolutely because of (53). More generally, the mean-zero bounded solution of

$$\frac{\partial f}{\partial \tau} + Qf = g(\tau, \hat{V}) \tag{54}$$

with the right side  $g \in L^\infty([0, T] \times V)$  is given by

$$f(\tau, \hat{V}) = - \int_0^\infty e^{rQ} g(\tau + r, \hat{V}) dr \tag{55}$$

provided that  $\langle \pi, g \rangle(\tau) = 0$  for all  $\tau > 0$ .

**4.2. The convergence theorem**

Let us start with the Wigner equation (22) written as

$$\frac{\partial W_\varepsilon}{\partial t} + \mathbf{k} \cdot \nabla W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}} \int \frac{d\hat{V}(t/\varepsilon, \mathbf{p})}{(2\pi)^{2d}} e^{i\mathbf{p} \cdot \mathbf{x}/\varepsilon} \left[ W_\varepsilon \left( t, \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} \right) - W_\varepsilon \left( t, \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2} \right) \right]. \tag{56}$$

It is convenient to formulate the time-splitting algorithm (37)–(38) in a somewhat more general framework as follows. We replace the exact equation (56) by

$$\frac{\partial W_\varepsilon}{\partial t} + \phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \cdot \nabla W_\varepsilon = \mathcal{L}_\varepsilon W_\varepsilon \tag{57}$$

with

$$\mathcal{L}_\varepsilon f(t, \mathbf{x}, \mathbf{k}) = \frac{1}{i\sqrt{\varepsilon}} \int \frac{d\hat{V}(t/\varepsilon, \mathbf{p})}{(2\pi)^{2d}} e^{i\mathbf{p} \cdot \mathbf{x}/\varepsilon} \left[ f \left( t, \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} \right) - f \left( t, \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2} \right) \right] \psi \left( \frac{t}{\varepsilon} \right). \tag{58}$$

The functions  $\phi(\tau)$  and  $\psi(\tau)$  are periodic in  $\tau = t/\varepsilon$  with the period  $\Theta > 0$ , which corresponds to the period  $\Theta\varepsilon$  on the large time scale. A generalized time-splitting algorithm corresponds to periodically shutting down the two operators in (56) so that

$$\psi(\tau) = \begin{cases} 0, & \tau \in [0, \tau_0) \\ \frac{\Theta}{\Theta - \tau_0}, & \tau \in [\tau_0, \Theta), \end{cases} \quad \text{and} \quad \phi(\tau) = \begin{cases} \frac{1}{\tau_0}, & \tau \in [0, \tau_0) \\ 0, & \tau \in [\tau_0, \Theta). \end{cases}$$

The classical time-splitting scheme (37)–(38) that was considered in Section 3 corresponds to  $\psi(\tau) = 1$  and  $\phi(\tau) = \sum_{n=-\infty}^\infty \delta(\tau - n\Theta)$ . We allow for more general distributions  $\psi$  and  $\phi$  in order to investigate other possibilities but we impose the constraints

$$\frac{1}{\Theta} \int_0^\Theta \phi(\tau) d\tau = 1, \quad \frac{1}{\Theta} \int_0^\Theta \psi(\tau) d\tau = 1. \tag{59}$$

This is a natural restriction ensuring that both operators in the time-splitting procedure have equal weight and that time is not re-scaled. Let us define the scattering operator

$$\begin{aligned} \mathcal{L}_\Theta f = & \frac{1}{\Theta} \int_0^\Theta \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \psi(\tau + s) \psi(\tau) \\ & \times \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\tau + s) - \Phi(\tau)]} + e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\tau) - \Phi(\tau + s)]} \right] \frac{d\mathbf{p} ds d\tau}{(2\pi)^d} \end{aligned} \tag{60}$$

where  $\Phi(s)$  is an anti-derivative of  $\phi$ :  $\frac{d\Phi}{ds} = \phi$ . Since only increments of  $\Phi$  appear in our results the choice of a particular anti-derivative is irrelevant. The main result of this section is the following theorem.

**Theorem 4.1.** *Let the initial data  $W_\varepsilon^0(\mathbf{x}, \mathbf{k})$  for (57) converge to  $W_0(\mathbf{x}, \mathbf{k})$  strongly in  $L^2(\mathbb{R}^{2d})$ . Then the modified Wigner distribution  $W_\varepsilon$ , solution of (57) converges in probability and weakly in  $L^2(\mathbb{R}^d)$  to the solution  $\overline{W}$  of the modified transport equation*

$$\frac{\partial \overline{W}}{\partial t} + \mathbf{k} \cdot \nabla \overline{W} = \mathcal{L}_\Theta \overline{W} \tag{61}$$

with the initial data  $W_0(\mathbf{x}, \mathbf{k})$ . More precisely, for any test function  $\lambda \in L^2(\mathbb{R}^d)$  the random process

$$\langle W, \lambda(t) \rangle = \int_{\mathbb{R}^{2d}} W_\varepsilon(t, \mathbf{x}, \mathbf{k}) \lambda(\mathbf{x}, \mathbf{k}) d\mathbf{x} d\mathbf{k}$$

converges in probability to  $\langle \overline{W}, \lambda \rangle$  as  $\varepsilon \rightarrow 0$  uniformly on finite time intervals  $t \in [0, T]$ .

It is instructive to consider two limiting cases:  $\Theta \rightarrow 0$  and  $\Theta \rightarrow \infty$  as in the discussion at the end of Section 3. The first case corresponds to time splitting with a very small time step, when we expect a very good approximation to the true statistics. The second arises when the time step is very large and the oscillatory random potential has large variations inside each time step. Then we expect that such variations would appear as white noise. Indeed, we replace for convenience  $\psi \rightarrow \psi(\tau/\Theta)$  and  $\phi \rightarrow \phi(\tau/\Theta)$  and observe as  $\Theta \rightarrow \infty$

$$\begin{aligned} \mathcal{L}_\Theta f &= \frac{1}{\Theta} \int_0^\Theta \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \psi\left(\frac{\tau+s}{\Theta}\right) \psi\left(\frac{\tau}{\Theta}\right) \\ &\quad \times \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\frac{\tau+s}{\Theta}) - \Phi(\frac{\tau}{\Theta})]} + e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\frac{\tau}{\Theta}) - \Phi(\frac{\tau+s}{\Theta})]} \right] \frac{d\mathbf{p} ds d\tau}{(2\pi)^d} \\ &= \int_0^1 \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \psi\left(\tau + \frac{s}{\Theta}\right) \psi(\tau) \\ &\quad \times \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\tau + \frac{s}{\Theta}) - \Phi(\tau)]} + e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2} [\Phi(\tau) - \Phi(\tau + \frac{s}{\Theta})]} \right] \frac{d\mathbf{p} ds d\tau}{(2\pi)^d} \\ &\rightarrow \left( \int_0^1 |\psi(\tau)|^2 d\tau \right) \int \hat{R}(0, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \frac{d\mathbf{p}}{(2\pi)^d}. \end{aligned}$$

It is interesting to observe that in the limit  $\Theta \rightarrow \infty$  the scattering operator does not depend on the choice of the function  $\phi$  at all, so that advection may be carried out in any fashion. The only dependence on the function  $\psi$  is that it controls the magnification of the scattering operator *via* its  $L^2$ -norm. Moreover, the scattering kernel coincides with that in (50). As we have discussed previously the kernel  $\hat{R}(\omega, \mathbf{p})$  that is independent of  $\omega$  corresponds to the white noise limit in time. However, even in the white noise regime the correct amplification factor is obtained only if  $\|\psi\|_{L^2} = 1$  which together with (59) implies that  $\psi = 1$ . This means that in order to capture the correct limit in the white noise scaling one has to keep the scattering “switched on” at all times. Otherwise information is lost which leads to an incorrect limit.

In the opposite limit  $\Theta \rightarrow 0$  the function  $\Phi(s)$  may be decomposed as

$$\Phi(s) = \int_0^s \phi\left(\frac{\xi}{\Theta}\right) d\xi = s + \Theta \int_0^{s/\Theta - [s/\Theta]} (\phi(\xi) - 1) d\xi = s + \Theta \tilde{\Phi}\left(\frac{s}{\Theta}\right),$$

where  $[s]$  is the integer part of  $s$ . The function  $\tilde{\Phi}$  is periodic in  $\Theta$  and bounded so that  $\tilde{\Phi}(s) \rightarrow s$  as  $\Theta \rightarrow 0$ . The same change of variables as in the case  $\Theta \rightarrow \infty$  shows that

$$\begin{aligned} \mathcal{L}_\Theta f \rightarrow \int_0^1 d\tau \int_0^1 d\zeta \int_0^\infty ds \int \frac{d\mathbf{p}}{(2\pi)^d} \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \psi(\tau + \zeta) \psi(\tau) \\ \times \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}s} + e^{-i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}s} \right] = \int \frac{d\mathbf{p}}{(2\pi)^d} \hat{R}\left(\frac{\mathbf{p}^2 - \mathbf{k}^2}{2}, \mathbf{p} - \mathbf{k}\right) [f(\mathbf{p}) - f(\mathbf{k})]. \end{aligned}$$

We see that the correct power spectrum is recovered in this limit. Therefore the time-splitting scheme (57) has the correct behavior for an arbitrary choice of the controls  $\phi$  and  $\psi$  that satisfy (59) provided that the time step  $\Theta_\varepsilon = \Theta\varepsilon$  with  $\Theta \ll 1$ , or equivalently  $\Theta_\varepsilon \ll \varepsilon$ .

Another important special case arises when  $\psi = 1$  and  $\phi(\tau) = \Theta \sum_{j=-\infty}^\infty \delta(\tau - j\Theta)$ . This corresponds to the time-splitting scheme (37)–(38) when scattering is accounted for at all times while advection is accounted for at times  $t = j\varepsilon\Theta$  by the correction  $W(\varepsilon j\Theta^+, \mathbf{x}, \mathbf{k}) = W(\varepsilon j\Theta^-, \mathbf{x} - \varepsilon\Theta\mathbf{k}, \mathbf{k})$ . This is exactly the algorithm analyzed formally in Section 3. Then we have  $\Phi(s) = \Theta[s/\Theta] := [s]_\Theta$  and obtain the following expression for  $\mathcal{L}_\Theta$ :

$$\mathcal{L}_\Theta f = \frac{1}{\Theta} \int_0^\Theta \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \times \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}([\tau+s]_\Theta - [\tau]_\Theta)} + e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}([\tau]_\Theta - [\tau+s]_\Theta)} \right] \frac{d\mathbf{p}ds}{(2\pi)^d}.$$

However, we have  $[\tau]_\Theta = 0$  when  $0 \leq \tau < \Theta$  and we obtain

$$\mathcal{L}_\Theta f = \frac{1}{\Theta} \int_0^\Theta \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [f(\mathbf{p}) - f(\mathbf{k})] \left[ e^{i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}[\tau+s]_\Theta} + e^{-i\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}[\tau+s]_\Theta} \right] \frac{d\mathbf{p}ds}{(2\pi)^d}. \tag{62}$$

The operator  $\mathcal{L}_\Theta$  in (62) coincides with that obtained in (46) by the formal calculation and hence the limit equation (61) is nothing but (48) in this case.

### 4.3. Outline of the proof

The proof of Theorem 4.1 follows the idea of the proof of the main result in [6]. Therefore we outline the main steps and concentrate only on the necessary modifications in the proof. First one may show that the family of measures  $P_\varepsilon$  generated by the process  $W_\varepsilon(t)$  on  $C([0, T]; L^2(\mathbb{R}^{2d}))$  is tight.

**Lemma 4.2.** *The family of measures  $P_\varepsilon$  is weakly compact.*

The proof of this lemma is very similar to that in [6] and is omitted. It is straightforward to verify that the  $L^2$ -norm of  $W_\varepsilon$  is preserved by evolution and hence  $W_\varepsilon$  takes values in a ball  $X = \{W \in L^2 : \|W\|_{L^2} \leq C\}$ .

**Lemma 4.3.** *The  $L^2$ -norm of the modified Wigner distribution is preserved:*

$$\|W_\varepsilon(t)\|_{L^2(\mathbb{R}^{2d})} = \|W_\varepsilon(0)\|_{L^2(\mathbb{R}^{2d})}. \tag{63}$$

Let  $\lambda(t, \mathbf{x}, \mathbf{k})$  be a fixed deterministic function. In order to identify the limit of  $W_\varepsilon$  we construct a functional  $G_\lambda : C([0, T]; X) \rightarrow C[0, T]$  defined by

$$G_\lambda[W](t) = \langle W, \lambda \rangle(t) - \int_0^t \left\langle W, \frac{\partial \lambda}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda + \mathcal{L}_\Theta \lambda \right\rangle(s) ds \tag{64}$$

and show that it is an approximate martingale. More precisely, we show that the following lemma holds.

**Lemma 4.4.** *There exists a constant  $C > 0$  so that*

$$|\mathbb{E}^{P_\varepsilon} \{G_\lambda[W](t)|\mathcal{F}_s\} - G_\lambda[W](s)| \leq C_{\lambda,T}\sqrt{\varepsilon} \tag{65}$$

uniformly for all  $W \in C([0, T]; X)$  and  $0 \leq s < t \leq T$ .

The proof of Lemma 4.4 is based on the construction of an exact martingale  $G_\lambda^\varepsilon[W]$  that is uniformly close to  $G_\lambda[W]$  within  $O(\sqrt{\varepsilon})$ . Lemma 4.2 implies that there exists a subsequence  $\varepsilon_j \rightarrow 0$  so that  $P_{\varepsilon_j}$  converges weakly to a measure  $P$  supported on  $C([0, T]; X)$ . Weak convergence of  $P_\varepsilon$  and the strong convergence (65) together imply that  $G_\lambda[W](t)$  is a  $P$ -martingale so that

$$\mathbb{E}^P \{G_\lambda[W](t)|\mathcal{F}_s\} - G_\lambda[W](s) = 0. \tag{66}$$

Taking  $s = 0$  above we obtain as in [5] the transport equation (61) for  $\overline{W} = e^P \{W(t)\}$  in its weak formulation. Construction of the martingale  $G_\lambda^\varepsilon$  and the proof of Lemma 4.4 are presented in detail in Section 4.4.

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

$$G_{2,\lambda}[W](t) = \langle W, \lambda \rangle^2(t) - 2 \int_0^t \langle W, \lambda \rangle(s) \left\langle W, \frac{\partial \lambda}{\partial s} + \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda + \mathcal{L}_\Theta \lambda \right\rangle(s) ds$$

is also an approximate  $P_\varepsilon$ -martingale. We then obtain that  $\mathbb{E}^{P_\varepsilon} \{ \langle W, \lambda \rangle^2 \} \rightarrow \langle \overline{W}, \lambda \rangle^2$ , which implies convergence in probability. It follows that the limit measure  $P$  is unique and deterministic, and that the whole sequence  $P_\varepsilon$  converges. All the modifications required in this step compared to [6] are very similar to those in the proof of Lemma 4.4 in the construction of the martingale  $G_\lambda^\varepsilon$  and we do not present the details in this part of the proof.

#### 4.4. The approximate martingale

To obtain the approximate martingale property (65) and prove Lemma 4.4, 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  $C([0, T]; \mathcal{V} \times X)$  generated by  $V(t/\varepsilon)$  and the Cauchy problem (57). The only functions we need to consider are actually of the form  $F(W, \hat{V}) = \langle W, \lambda(\hat{V}) \rangle$  with  $\lambda \in L^\infty(\mathcal{V}; C^1([0, T]; \mathcal{S}(\mathbb{R}^{2d})))$ . Given a function  $F(W, \hat{V})$  let us define the conditional expectation

$$\mathbb{E}_{W, \hat{V}, t}^{\tilde{P}_\varepsilon} \{ F(W, \hat{V}) \} (\tau) = \mathbb{E}^{\tilde{P}_\varepsilon} \{ F(W(\tau), \tilde{V}(\tau)) | W(t) = W, \tilde{V}(t) = \hat{V} \}, \quad \tau \geq t.$$

The weak form of the infinitesimal generator of the Markov process generated by  $\tilde{P}_\varepsilon$  is given by

$$\frac{d}{dh} \mathbb{E}_{W, \hat{V}, t}^{\tilde{P}_\varepsilon} \{ \langle W, \lambda(\hat{V}) \rangle \} (t+h) \Big|_{h=0} = \frac{1}{\varepsilon} \langle W, Q\lambda \rangle + \left\langle W, \left( \frac{\partial}{\partial t} + \phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \cdot \nabla_{\mathbf{x}} + \frac{1}{\sqrt{\varepsilon}} \mathcal{K} \left[ \hat{V}, \frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon} \right] \right) \lambda \right\rangle, \tag{67}$$

hence

$$G_\lambda^\varepsilon = \langle W, \lambda(\hat{V}) \rangle(t) - \int_0^t \left\langle W(s), \left( \frac{1}{\varepsilon} Q + \frac{\partial}{\partial s} + \phi \left( \frac{s}{\varepsilon} \right) \mathbf{k} \cdot \nabla_{\mathbf{x}} + \frac{1}{\sqrt{\varepsilon}} \mathcal{K} \left[ \hat{V}, \frac{s}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon} \right] \right) \lambda(s) \right\rangle ds \tag{68}$$

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

$$\mathcal{K}[\hat{V}, \tau, \mathbf{z}] f(\mathbf{x}, \tau, \mathbf{z}, \mathbf{k}, \hat{V}) = \frac{1}{i} \int_{\mathbb{R}^d} \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot \mathbf{z}} \left[ f \left( \tau, \mathbf{x}, \mathbf{z}, \mathbf{k} - \frac{\mathbf{p}}{2} \right) - f \left( \tau, \mathbf{x}, \mathbf{z}, \mathbf{k} + \frac{\mathbf{p}}{2} \right) \right] \psi(\tau). \tag{69}$$

The generator (67) comes from equation (57) written in the form

$$\frac{\partial W_\varepsilon}{\partial t} + \phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \cdot \nabla_{\mathbf{x}} W_\varepsilon = \frac{1}{\sqrt{\varepsilon}} \mathcal{K} \left[ \hat{V} \left( \frac{t}{\varepsilon} \right), \frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon} \right] W_\varepsilon. \tag{70}$$

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

$$\lambda_\varepsilon(t, \mathbf{x}, \mathbf{k}, \hat{V}) = \lambda(t, \mathbf{x}, \mathbf{k}) + \sqrt{\varepsilon} \lambda_1^\varepsilon(t, \mathbf{x}, \mathbf{k}, \hat{V}) + \varepsilon \lambda_2^\varepsilon(t, \mathbf{x}, \mathbf{k}, \hat{V}) \tag{71}$$

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

$$\|G_{\lambda_\varepsilon}^\varepsilon(t) - G_\lambda(t)\|_{L^\infty(\mathcal{V})} \leq C_\lambda \sqrt{\varepsilon} \tag{72}$$

for all  $t \in [0, T]$ . Here  $G_{\lambda_\varepsilon}^\varepsilon$  is defined by (68) with  $\lambda$  replaced by  $\lambda_\varepsilon$ , and  $G_\lambda$  is defined by (64). The approximate martingale property (65) follows from this.

The functions  $\lambda_1^\varepsilon$  and  $\lambda_2^\varepsilon$  are defined as follows. Let  $\lambda_1(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k}, \hat{V})$  be the mean-zero solution of

$$\frac{\partial \lambda_1}{\partial \tau} + \phi(\tau) \mathbf{k} \cdot \nabla_{\mathbf{z}} \lambda_1 + Q \lambda_1 = -\frac{1}{i} \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot \mathbf{z}} \left[ \lambda\left(\mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda\left(\mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(\tau), \tag{73}$$

where  $\hat{V}(\mathbf{p})$  is fixed and independent of  $\tau$ . This equation has a form similar to (54) and its solution is given by an appropriate modification of (55):

$$\lambda_1(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k}, \hat{V}) = \frac{1}{i} \int_0^\infty e^{sQ} \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i(\mathbf{p} \cdot \mathbf{z} + (\mathbf{k} \cdot \mathbf{p})(\Phi(\tau+s) - \Phi(\tau)))} \left[ \lambda\left(t, \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda\left(t, \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(\tau + s) ds. \tag{74}$$

The equation for  $\lambda_2$  is

$$(\partial_\tau + \phi(\tau) \mathbf{k} \cdot \nabla_{\mathbf{z}} + Q) \lambda_2 = \mathcal{L}_\Theta \lambda - \mathcal{K} \lambda_1 + [1 - \phi(\tau)] \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda. \tag{75}$$

The first term on the right may be recast as

$$\left( \mathcal{L}_\Theta \lambda - \mathcal{L}_\Theta^\tau \lambda \right) + \left( \mathcal{L}_\Theta^\tau \lambda - \mathcal{K} \lambda_1 \right).$$

We decompose  $\lambda_2$  as  $\lambda_{21} + \lambda_{22} + \lambda_{23}$ , corresponding to the arising three source terms, respectively. The operator  $\mathcal{L}_\Theta$  is defined by (60) while  $\mathcal{L}_\Theta^\tau$  is defined by

$$\mathcal{L}_\Theta^\tau \lambda = \mathbb{E} \{ \mathcal{K} \lambda_1 \}.$$

Before proceeding with the solution for  $\lambda_2$  we find the explicit form of  $\mathcal{L}_\Theta^\tau$  and verify that

$$\mathcal{L}_\Theta \lambda = \frac{1}{\Theta} \int_0^\Theta \mathcal{L}_\Theta^\tau \lambda d\tau, \tag{76}$$

where  $\Theta$  is the period of  $\phi$  and  $\psi$ . Let us first compute  $\mathcal{L}_\Theta^\tau \lambda$ :

$$\begin{aligned} \mathcal{L}_\Theta^\tau \lambda(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k}) &= -\frac{1}{i} \mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot \mathbf{z}} \left[ \lambda_1\left(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda_1\left(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(\tau) \right\} \\ &= I + I^*. \end{aligned}$$

We have

$$\begin{aligned}
 I &= -\frac{1}{i} \mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \lambda_1 \left( \mathbf{k} + \frac{\mathbf{p}}{2} \right) \psi(\tau) \right\} \\
 &= \mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \int_0^\infty e^{sQ} \int \frac{d\hat{V}(\mathbf{q})}{(2\pi)^d} e^{i((\mathbf{q}\cdot\mathbf{z}) + ((\mathbf{k}+\mathbf{p}/2)\cdot\mathbf{q})[\Phi(\tau+s) - \Phi(\tau)])} \right\} \\
 &\quad \times \left[ \lambda \left( \mathbf{k} + \frac{\mathbf{p}}{2} - \frac{\mathbf{q}}{2} \right) - \lambda \left( \mathbf{k} + \frac{\mathbf{p}}{2} + \frac{\mathbf{q}}{2} \right) \right] \psi(\tau+s)\psi(\tau) ds \\
 &= \int_0^\infty \int \tilde{R}(s, \mathbf{p}) e^{-i(\mathbf{k}+\mathbf{p}/2)\cdot\mathbf{p}[\Phi(\tau+s) - \Phi(\tau)]} [\lambda(\mathbf{k} + \mathbf{p}) - \lambda(\mathbf{k})] \psi(\tau+s)\psi(\tau) \frac{d\mathbf{p}ds}{(2\pi)^d} \\
 &= \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) e^{i(\mathbf{k}^2 - \mathbf{p}^2)[\Phi(\tau+s) - \Phi(\tau)]/2} [\lambda(\mathbf{p}) - \lambda(\mathbf{k})] \psi(\tau+s)\psi(\tau) \frac{d\mathbf{p}ds}{(2\pi)^d}.
 \end{aligned}$$

Then we obtain

$$\begin{aligned}
 \mathcal{L}_\Theta^\tau \lambda(t, \tau, \mathbf{x}, \mathbf{k}) &= \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) [\lambda(t, \mathbf{x}, \mathbf{p}) - \lambda(t, \mathbf{x}, \mathbf{k})] \\
 &\quad \times \left[ e^{i(\mathbf{k}^2 - \mathbf{p}^2)[\Phi(\tau+s) - \Phi(\tau)]/2} + e^{i(\mathbf{k}^2 - \mathbf{p}^2)[\Phi(\tau) - \Phi(\tau+s)]/2} \right] \psi(\tau+s)\psi(\tau) \frac{d\mathbf{p}ds}{(2\pi)^d} \tag{77}
 \end{aligned}$$

and (76) follows. We observe that the operators  $\mathcal{L}_\Theta$  and  $\mathcal{L}_\Theta^\tau$  are independent of  $\hat{V}$  and  $\mathbf{z}$  and therefore the function  $\lambda_{21} = \lambda_{21}(t, \tau, \mathbf{x}, \mathbf{k})$  is also independent of these variables. It is given explicitly by

$$\lambda_{21}(t, \tau, \mathbf{x}, \mathbf{k}) = \int_0^\tau [\mathcal{L}_\Theta(t, \mathbf{x}, \mathbf{k}) - \mathcal{L}_\Theta^\tau(t, \zeta, \mathbf{x}, \mathbf{k})] \lambda(t, \mathbf{x}, \mathbf{k}) d\zeta \tag{78}$$

and is periodic in the fast variable  $\tau$ . Similarly the function  $\lambda_{23}$  is also independent of  $\hat{V}$  and  $\mathbf{z}$  and is given by

$$\lambda_{23}(t, \tau, \mathbf{x}, \mathbf{k}) = [\tau - \Phi(\tau)] \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda(t, \mathbf{x}, \mathbf{k}). \tag{79}$$

The function  $\lambda_{22}$  satisfies

$$\frac{\partial \lambda_{22}}{\partial \tau} + \phi(\tau) \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda_{22} + Q \lambda_{22} = \mathcal{L}_\Theta^\tau \lambda - \mathcal{K} \lambda_1.$$

It is given explicitly by

$$\lambda_{22}(t, \tau, \mathbf{x}, \mathbf{z}, \mathbf{k}, \hat{V}) = - \int_0^\infty e^{sQ} [\mathcal{L}_\Theta^\tau \lambda(t, \tau + s, \mathbf{x}, \mathbf{k}) - \mathcal{K} \lambda_1(t, \tau + s, \mathbf{x}, \mathbf{z} + (\Phi(\tau + s) - \Phi(\tau)) \mathbf{k}, \mathbf{k}, \hat{V})]. \tag{80}$$

Using (73) and (75) we have

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

with

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

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  $\mathbf{z} = \mathbf{x}/\varepsilon$ . It follows that  $G_{\lambda_\varepsilon}^\varepsilon$  is given by

$$G_{\lambda_\varepsilon}^\varepsilon(t) = \langle W(t), \lambda_\varepsilon \rangle - \int_0^t \left\langle W, \left( \frac{\partial}{\partial t} + \mathbf{k} \cdot \nabla_{\mathbf{x}} + \mathcal{L}_\Theta \right) \lambda \right\rangle(s) ds - \sqrt{\varepsilon} \int_0^t \langle W, \zeta_\varepsilon^\lambda \rangle(s) ds \tag{81}$$

and is a martingale with respect to the measure  $\tilde{P}_\varepsilon$  defined on  $C([0, T]; X \times \mathcal{V})$ . Lemma 4.4 and the estimate (65) follow from the following two lemmas.

**Lemma 4.5.** *Let  $\lambda \in C^1([0, T]; \mathcal{S}(\mathbb{R}^{2d}))$ . Then there exists a constant  $C_\lambda > 0$  independent of  $t \in [0, T]$  so that the correctors  $\lambda_1^\varepsilon(t)$  and  $\lambda_2^\varepsilon(t)$  satisfy the uniform bounds*

$$\|\lambda_1^\varepsilon(t)\|_{L^\infty(\mathcal{V}; L^2)} + \|\lambda_2^\varepsilon(t)\|_{L^\infty(\mathcal{V}; L^2)} \leq C_\lambda \tag{82}$$

and

$$\left\| \frac{\partial \lambda_1^\varepsilon(t)}{\partial t} + \phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda_1^\varepsilon(t) \right\|_{L^\infty(\mathcal{V}; L^2)} + \left\| \frac{\partial \lambda_2^\varepsilon(t)}{\partial t} + \phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \cdot \nabla_{\mathbf{x}} \lambda_2^\varepsilon(t) \right\|_{L^\infty(\mathcal{V}; L^2)} \leq C_\lambda. \tag{83}$$

**Lemma 4.6.** *There exists a constant  $C_\lambda$  such that*

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

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

*Proof of lemma 4.4.* Observe that (82) implies that  $|\langle W, \lambda \rangle - \langle W, \lambda_\varepsilon \rangle| \leq C\sqrt{\varepsilon}$  for all  $W \in X$  and  $\hat{V} \in \mathcal{V}$ , while (83) and Lemma 4.6 imply that for all  $t \in [0, T]$

$$\|\zeta_\varepsilon^\lambda(t)\|_{L^2} \leq C \tag{84}$$

for all  $\hat{V} \in \mathcal{V}$  so that (65) follows from the fact that (81) is a martingale. □

*Proof of Lemma 4.6.* Lemma 4.6 follows immediately from the definition of  $\mathcal{K}$ , the uniform bound on the total mass of  $\hat{V}$  in (51) and the Cauchy-Schwartz inequality. We have  $\mathcal{K}f = I + I^*$ , with

$$I(\mathbf{x}, \mathbf{k}) = \frac{1}{i} \int_{\mathbb{R}^d} \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot \mathbf{x}/\varepsilon} f \left( \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} \right) \psi \left( \frac{t}{\varepsilon} \right)$$

and

$$\begin{aligned} \int |I(\mathbf{x}, \mathbf{k})|^2 &= \int_{\mathbb{R}^{4d}} e^{i(\mathbf{p}-\mathbf{q}) \cdot \mathbf{x}/\varepsilon} f \left( \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} \right) \bar{f} \left( \mathbf{x}, \mathbf{k} - \frac{\mathbf{q}}{2} \right) \left| \psi \left( \frac{t}{\varepsilon} \right) \right|^2 \frac{d\hat{V}(\mathbf{p}) d\hat{V}(\mathbf{q})}{(2\pi)^{2d}} \\ &\leq C \int \frac{|d\hat{V}(\mathbf{p})| |d\hat{V}(\mathbf{q})|}{(2\pi)^{2d}} \|f\|_{L^2}^2 \leq C \|f\|_{L^2}^2. \end{aligned}$$

We now prove Lemma 4.5. We will omit the  $t$ -dependence of the test function  $\lambda$  to simplify notation. □

*Proof of Lemma 4.5.* We only prove (82). Since  $\lambda \in \mathcal{S}(\mathbb{R}^{2d})$ , there exists a constant  $C_\lambda$  so that

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

Then we obtain using (51) and (53)

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

and the  $L^2$ -bound on  $\lambda_1$  follows. Here  $L$  is the size of the support of  $\hat{V}$  as in (51).

We show next that  $\lambda_2^\varepsilon$  is uniformly bounded. Recall that  $\lambda_2^\varepsilon = \lambda_{21}^\varepsilon + \lambda_{22}^\varepsilon + \lambda_{23}^\varepsilon$  with the three terms given by (78), (80) and (79), respectively. The  $L^2$ -bound on  $\lambda_{23}$  follows immediately from the fact that  $|\tau - \Phi(\tau)| \leq C$  because of (59). We also note that

$$\lambda_{21}(t, \tau, \mathbf{x}, \mathbf{k}) = \int_{n\Theta}^\tau [\mathcal{L}_\Theta(t, \mathbf{x}, \mathbf{k}) - \mathcal{L}_\Theta^\tau(t, \zeta, \mathbf{x}, \mathbf{k})] \lambda(t, \mathbf{x}, \mathbf{k}) d\zeta \tag{85}$$

where  $n\Theta \leq \tau < (n + 1)\Theta$ . Moreover, we note that for all  $\tau > 0$  we have

$$|(\mathcal{L}_\Theta^\tau \lambda)(t, \tau, \mathbf{x}, \mathbf{k})| \leq C \int Z(\mathbf{p} - \mathbf{k}) [|\lambda(t, \mathbf{x}, \mathbf{p})| + |\lambda(t, \mathbf{x}, \mathbf{k})|] d\mathbf{p}, \tag{86}$$

where  $Z(\mathbf{k}) = \int_0^\infty |\tilde{R}(s, \mathbf{k})|$  and hence  $\|(\mathcal{L}_\Theta^\tau \lambda)(t, \tau, \mathbf{x}, \mathbf{k})\|_{L^2(\mathbb{R}^{2d})} \leq C_\lambda$  for all  $t$  and  $\tau$ . This in turn implies that  $\|(\mathcal{L}_\Theta \lambda)(t, \mathbf{x}, \mathbf{k})\|_{L^2} \leq C_\lambda$  so that (85) implies that  $\lambda_{21}^\varepsilon$  is uniformly bounded in  $L^2$ . We have for  $\lambda_{22}$ :

$$\begin{aligned} \lambda_{22}^\varepsilon(t, \mathbf{x}, \mathbf{k}, \hat{V}) &= - \int_0^\infty ds e^{sQ} \left\{ \mathcal{L}_\Theta^\tau \lambda\left(t, \frac{t}{\varepsilon} + s, \mathbf{x}, \mathbf{k}\right) - \frac{1}{i} \int_{\mathbb{R}^d} \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot (\mathbf{x}/\varepsilon + (\Phi(t/\varepsilon + s) - \Phi(t/\varepsilon))\mathbf{k})} \right. \\ &\quad \times \left[ \lambda_1\left(t, \frac{t}{\varepsilon} + s, \mathbf{x}, \frac{\mathbf{x}}{\varepsilon} + \left(\Phi\left(\frac{t}{\varepsilon} + s\right) - \Phi\left(\frac{t}{\varepsilon}\right)\right) \mathbf{k}, \mathbf{k} - \frac{\mathbf{p}}{2}, \hat{V}\right) \right. \\ &\quad \left. \left. - \lambda_1\left(t, \frac{t}{\varepsilon} + s, \mathbf{x}, \frac{\mathbf{x}}{\varepsilon} + \left(\Phi\left(\frac{t}{\varepsilon} + s\right) - \Phi\left(\frac{t}{\varepsilon}\right)\right) \mathbf{k}, \mathbf{k} + \frac{\mathbf{p}}{2}, \hat{V}\right) \right] \psi\left(\frac{t}{\varepsilon} + s\right) \right\}. \end{aligned}$$

The second term above may be written as  $I + I^*$  with

$$\begin{aligned} I &= \frac{1}{i} \int_{\mathbb{R}^d} \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot (\mathbf{x}/\varepsilon + (\Phi(t/\varepsilon + s) - \Phi(t/\varepsilon))\mathbf{k})} \lambda_1\left(t, \frac{t}{\varepsilon} + s, \mathbf{x}, \frac{\mathbf{x}}{\varepsilon} + \left(\Phi\left(\frac{t}{\varepsilon} + s\right) - \Phi\left(\frac{t}{\varepsilon}\right)\right) \mathbf{k}, \mathbf{k} - \frac{\mathbf{p}}{2}, \hat{V}\right) \\ &\quad \times \psi\left(\frac{t}{\varepsilon} + s\right). \end{aligned}$$

This may be re-written as

$$\begin{aligned}
 I = & - \int_{\mathbb{R}^d} \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot (\mathbf{x}/\varepsilon + (\Phi(t/\varepsilon+s) - \Phi(t/\varepsilon))\mathbf{k})} \int_0^\infty e^{rQ} \int \exp \left\{ i\mathbf{q} \cdot \left( \frac{\mathbf{x}}{\varepsilon} + \left( \Phi \left( \frac{t}{\varepsilon} + s \right) - \Phi \left( \frac{t}{\varepsilon} \right) \mathbf{k} \right) \right. \right. \\
 & \left. \left. + \left( \Phi \left( \frac{t}{\varepsilon} + s + r \right) - \Phi \left( \frac{t}{\varepsilon} + s \right) \right) \mathbf{q} \cdot \left( \mathbf{k} - \frac{\mathbf{p}}{2} \right) \right\} \left[ \lambda \left( \mathbf{k} - \frac{\mathbf{p}}{2} - \frac{\mathbf{q}}{2} \right) - \lambda \left( \mathbf{k} - \frac{\mathbf{p}}{2} + \frac{\mathbf{q}}{2} \right) \right] \\
 & \times \psi \left( \frac{t}{\varepsilon} + s \right) \psi \left( \frac{t}{\varepsilon} + s + r \right) \frac{d\hat{V}(\mathbf{q})}{(2\pi)^d} dr.
 \end{aligned}$$

Therefore we obtain

$$\begin{aligned}
 |\lambda_2^\varepsilon(t, \mathbf{x}, \mathbf{k}, \hat{V})| \leq & C \int_0^\infty ds e^{-\alpha s} \left[ \left| \mathcal{L}\Theta^\tau \lambda \left( t, \frac{t}{\varepsilon} + s, \mathbf{x}, \mathbf{k} \right) \right| + \sup_{\hat{V}} \int_{\mathbb{R}^d} |d\hat{V}(\mathbf{p})| \int_0^\infty dr e^{-\alpha r} \sup_{\hat{V}_1} \int_{\mathbb{R}^d} |d\hat{V}_1(\mathbf{q})| \right. \\
 & \left. \times \left( \left| \lambda \left( \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} - \frac{\mathbf{q}}{2} \right) \right| + \left| \lambda \left( \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2} + \frac{\mathbf{q}}{2} \right) \right| + \left| \lambda \left( \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2} - \frac{\mathbf{q}}{2} \right) \right| + \left| \lambda \left( \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2} + \frac{\mathbf{q}}{2} \right) \right| \right) \right] \\
 \leq & C \left[ \sup_{\tau \geq 0} |\mathcal{L}_\tau \lambda(\mathbf{x}, \mathbf{k})| + \frac{1}{(1 + |\mathbf{x}|^{5d})(1 + (|\mathbf{k}| - L)^{5d} \chi_{|\mathbf{k}| \geq 5L}(\mathbf{k}))} \right]
 \end{aligned}$$

and the  $L^2$ -bound on  $\lambda_2^\varepsilon$  in (82) follows because of (86). The proof of (83) is very similar and is omitted.  $\square$

This finishes the proof of Lemma 4.4. As explained in Section 4.3 the tightness of measures  $P_\varepsilon$  given by Lemma 4.2 implies then that the expectation  $\mathbb{E}\{W_\varepsilon(t, \mathbf{x}, \mathbf{k})\}$  converges weakly in  $L^2(\mathbb{R}^{2d})$  to the solution  $\overline{W}(t, \mathbf{x}, \mathbf{k})$  of the transport equation for each  $t \in [0, T]$ . The proof that actually  $W_\varepsilon$  converges to  $\overline{W}(t, \mathbf{x}, \mathbf{k})$  in probability is very similar to that in [6] with the modifications essentially identical to those we made in the proof of convergence of the expectation. Hence we omit the details.

## 5. AN EFFICIENT TIME SPLITTING ALGORITHM FOR THE WIGNER TRANSFORM

### 5.1. A time splitting scheme for the Wigner transform

The results of the preceding sections show that  $\Theta_\varepsilon = \varepsilon/N$  with  $N \gg 1$  is necessary to obtain the correct statistics of the wave energy density by solving the time splitting algorithm (7)–(9). However, instead of trying to calculate the wave field  $\psi_\varepsilon$  directly one may try to compute numerically its Wigner transform  $W_\varepsilon$  which provides the energy distribution in the phase space. The main disadvantage of this approach is that the latter calculation requires to solve a problem in the phase space, with twice as many variables as in the physical domain where  $\psi_\varepsilon$  is defined. Yet we shall see that the interval  $\Theta_\varepsilon$  can be chosen substantially larger if one modifies the treatment of the scattering term in an appropriate manner in a time splitting scheme for the Wigner transform.

Recall that the Wigner transform satisfies the evolution equation (25). Let us consider the time splitting algorithm for (25) that differs from (37)–(38) as follows. The advective operator  $\mathbf{k} \cdot \nabla_{\mathbf{x}}$  in (25) is treated in the same way as in (38):

$$W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^+, \mathbf{x}, \mathbf{k}) = W_\varepsilon^{\Theta_\varepsilon}(n\Theta_\varepsilon^-, \mathbf{x} - \Theta_\varepsilon \mathbf{k}, \mathbf{k}). \tag{87}$$

However, the scattering is accounted for differently from (38). The new scattering equation is

$$\frac{\partial W_\varepsilon^{\Theta_\varepsilon}}{\partial t} = \int_{\mathbb{R}^d} K_\varepsilon(\mathbf{x} - (t - n\Theta_\varepsilon)\mathbf{k}, \mathbf{k} - \mathbf{p}, t) W_\varepsilon^{\Theta_\varepsilon}(\mathbf{x}, \mathbf{p}, t) d\mathbf{p}, \tag{88}$$

on  $t \in (n\Theta_\varepsilon, (n+1)\Theta_\varepsilon)$  with  $K_\varepsilon$  given by (26). Note that (88) is different from (38): the kernel  $K_\varepsilon$  is taken at the point  $\mathbf{x} - (t - n\Theta_\varepsilon)\mathbf{k}$  instead of  $\mathbf{x}$ . That means, roughly speaking, that we account for the advection of the rapidly varying part of the right side in (88) that depends on the oscillatory potential  $V$  but we do not

advection  $W_\varepsilon$  inside the integral. This modification allows us to obtain the right dynamics with much larger  $\Theta_\varepsilon$  than in previous sections because indeed  $W_\varepsilon$  has a slowly varying limit as  $\varepsilon \rightarrow 0$ . The formal analysis of this scheme proceeds as in Section 3 and we omit the tedious details. The main difference is that (44) is now replaced by

$$\int_0^\tau \int_0^{\tau-s} \int_{\mathbb{R}^d} dp du ds \frac{1}{\varepsilon} \tilde{R}\left(\frac{u}{\varepsilon}, \mathbf{p} - \mathbf{k}\right) \left( e^{i\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}u/\varepsilon} + e^{-i\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}u/\varepsilon} \right) \times \left( \mathbb{E} \{ W_\varepsilon^{\Theta_\varepsilon} \} (\tau - [\tau, s]_{\Theta_\varepsilon} - [\tau - s, u]_{\Theta_\varepsilon}, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon} \mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon} \mathbf{p}, \mathbf{p}) - \mathbb{E} \{ W_\varepsilon^{\Theta_\varepsilon} \} (\tau - [\tau, s]_{\Theta_\varepsilon} - [\tau - s, u]_{\Theta_\varepsilon}, \mathbf{x} - [\tau, s]_{\Theta_\varepsilon} \mathbf{k} - [\tau - s, u]_{\Theta_\varepsilon} \mathbf{p}, \mathbf{k}) \right). \tag{89}$$

Still assuming that  $\mathbb{E} \{ W_\varepsilon^{\Theta_\varepsilon} \}$  is sufficiently smooth in the limit  $\varepsilon \rightarrow 0$  we obtain that this term is nothing but

$$\int_0^\tau \int_{\mathbb{R}^d} ds d\mathbf{p} \hat{R}\left(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}, \mathbf{p} - \mathbf{k}\right) (\mathbb{E} \{ W \} (\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{p}) - \mathbb{E} \{ W \} (\tau - s, \mathbf{x} - s\mathbf{k}, \mathbf{k})), \tag{90}$$

up to a term of order  $O(\Theta_\varepsilon + \varepsilon)$ . This error would be of the order of  $O(\Theta_\varepsilon/\varepsilon + \varepsilon)$  in the case of the time splitting algorithm for  $\psi_\varepsilon$ . This implies the following striking result: the time splitting algorithm (87)–(88) in the limit  $\varepsilon \rightarrow 0$  provides the exact solution of (20) with an error  $\Theta_\varepsilon$ . We can thus choose  $\Theta_\varepsilon$  independent of  $\varepsilon$ . All we need is the optimal constraint  $\Theta_\varepsilon \ll 1$  for convergence.

We see that the original time splitting (7)–(9), which is equivalent to the time splitting (37)–(38) for the Wigner transform can be substantially (and optimally) improved by the algorithm (87)–(88) that still decouples advection from the scattering part. The modification is based on considering the potential  $V$  at the correct location  $(\mathbf{x} - (t - n\Theta_\varepsilon)\mathbf{k})$  instead of  $\mathbf{x}$  during the scattering treatment. This correction is inherently posed in the phase space: the correct location is obtained wavenumber by wavenumber. This cannot be done in the physical domain and we do not believe that there are modifications of (7)–(9) that would yield the correct statistics with  $\Theta_\varepsilon \ll 1$ .

However, the possibility to take a large time step comes at a price. Not only do we have to work in the phase space that has a higher dimension than the physical space but also the modified scheme (88) may be harder to solve than (38). Indeed the latter can be recast as

$$\frac{\partial \hat{W}_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{y})}{\partial t} = \frac{i}{\sqrt{\varepsilon}} \left( V\left(\frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon} - \frac{\mathbf{y}}{2}\right) - V\left(\frac{t}{\varepsilon}, \frac{\mathbf{x}}{\varepsilon} + \frac{\mathbf{y}}{2}\right) \right) \hat{W}_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{y}), \tag{91}$$

where  $\hat{W}_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{y})$  is the Fourier transform  $\mathbf{k} \rightarrow \mathbf{y}$  of  $W_\varepsilon^{\Theta_\varepsilon}(t, \mathbf{x}, \mathbf{k})$ . In this domain (91) is trivial to solve and the solution of (38) is obtained by inverse Fourier transform. Since the modified algorithm (88) breaks the convolution  $K_\varepsilon * W_\varepsilon^{\Theta_\varepsilon}$  in the wavenumber variable, it can no longer be solved by Fourier transforms. The main advantage of the modification (88) is that it allows us to bypass the numerically costly advection that has to be performed much less often than in the time-splitting scheme (37)–(38).

### 5.2. Convergence of the time-splitting scheme for the Wigner equation

We analyze now convergence of the time-splitting algorithm (87)–(88) for the Wigner transform in the small  $\varepsilon$  limit for the Wigner transform with the time-step  $\Theta$  that is independent of  $\varepsilon$ . As in the analysis in Section 4 it is convenient to introduce a somewhat more general set-up that includes (87)–(88) as a particular example. We modify (56) as follows:

$$\frac{\partial W_\varepsilon}{\partial t} + \phi(t)\mathbf{k} \cdot \nabla_{\mathbf{x}} W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}} \int \frac{d\hat{V}(t/\varepsilon, \mathbf{p})}{(2\pi)^d} e^{i\mathbf{p} \cdot (\mathbf{x} - \eta(t)\mathbf{k})/\varepsilon} \left[ W_\varepsilon\left(t, \mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - W_\varepsilon\left(t, \mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(t). \tag{92}$$

This is the analog of the modified Wigner equation (57)–(58). Once again, choosing the functions  $\phi$  and  $\psi$  equal to zero on alternating time intervals in (92) leads to a genuine time-splitting scheme. In general the periodic functions  $\phi$  and  $\psi$  are as in Section 4 with the period  $\Theta$  and average equal to one. However, there is an important difference between the general time-splitting in (92) and that in (57)–(58): now the functions  $\phi$  and  $\psi$  vary on the macroscopic time-scale and not on the microscopic one. That corresponds to taking a time-step  $\Theta$  independent of  $\varepsilon$  in the time-splitting scheme. However, in order to allow for such a large time step one has to modify the oscillatory phase in the operator on the right side of (92) by means of the function  $\eta(t)$  that is also varying on the macroscopic scale. We choose  $\eta$  to be

$$\eta(t) = \int_0^t [\phi(s) - 1] ds. \tag{93}$$

This allows us to compensate for the large time-step by appropriately adjusting the potential  $V$  accounting indirectly for advection during the long time-step. This modification need not be made in the argument of the Wigner transform since the latter has a macroscopic limit. The main result is as follows.

**Theorem 5.1.** *Let the initial data  $W_\varepsilon^0(\mathbf{x}, \mathbf{k})$  for (57) converge to  $W_0(\mathbf{x}, \mathbf{k})$  strongly in  $L^2(\mathbb{R}^{2d})$ . Then the modified Wigner distribution  $W_\varepsilon$ , solution of (92) converges in probability and weakly in  $L^2(\mathbb{R}^d)$  to the solution  $\overline{W}$  of the modified transport equation*

$$\frac{\partial \overline{W}}{\partial t} + \phi(t)\mathbf{k} \cdot \nabla \overline{W} = \psi^2(t)\mathcal{L}_0 \overline{W}, \tag{94}$$

with the initial data  $W_0(\mathbf{x}, \mathbf{k})$ , and with

$$\mathcal{L}_0 f = \int \hat{R} \left( \frac{\mathbf{k}^2 - \mathbf{p}^2}{2}, \mathbf{k} - \mathbf{p} \right) (f(\mathbf{x}, \mathbf{p}) - f(\mathbf{x}, \mathbf{k})) \frac{d\mathbf{p}}{(2\pi)^d}. \tag{95}$$

More precisely, for any test function  $\lambda \in L^2(\mathbb{R}^d)$  the random process

$$\langle W, \lambda(t) \rangle = \int_{\mathbb{R}^{2d}} W_\varepsilon(t, \mathbf{x}, \mathbf{k}) \lambda(\mathbf{x}, \mathbf{k}) d\mathbf{x} d\mathbf{k}$$

converges in probability to  $\langle \overline{W}, \lambda \rangle$  as  $\varepsilon \rightarrow 0$  uniformly on finite time intervals  $t \in [0, T]$ .

Equation (94) is nothing but a time-splitting approximation with time-step  $\Theta$  of the correct limit radiative transport equation (20) that has the form

$$\frac{\partial \overline{W}}{\partial t} + \mathbf{k} \cdot \nabla \overline{W} = \mathcal{L}_0 \overline{W}. \tag{96}$$

In particular we obtain the correct scattering kernel for all  $\Theta$ .

The proof of Theorem 5.1 is very similar to that of Theorem 4.1. We only explain the necessary modifications. One no longer needs to introduce separately the fast spatial and temporal variables  $\mathbf{z} = \mathbf{x}/\varepsilon$  and  $\tau = t/\varepsilon$  in the construction of the correctors. The new fast variable is  $\mathbf{z} = (\mathbf{x} - \eta(t)\mathbf{k})/\varepsilon$  so that one formally has

$$\nabla_{\mathbf{x}} \rightarrow \nabla_{\mathbf{x}} + \frac{1}{\varepsilon}\nabla_{\mathbf{z}}, \quad \frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} - \frac{1}{\varepsilon}\dot{\eta}(t)\mathbf{k} \cdot \nabla_{\mathbf{z}}.$$

With our choice (93) of  $\eta(t)$  equation for the corrector  $\lambda_1$  takes a particularly simple form

$$\mathbf{k} \cdot \nabla_{\mathbf{z}}\lambda_1 + Q\lambda_1 = -\frac{1}{i} \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \left[ \lambda\left(\mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda\left(\mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(t), \tag{97}$$

since the function  $\phi(t) - \dot{\eta}(t)$  that would multiply the  $\mathbf{k} \cdot \nabla_{\mathbf{z}}$  term on the left side is equal identically to one. The function  $\lambda_1$  is given explicitly by

$$\lambda_1(\mathbf{x}, \mathbf{k}, \mathbf{z}, \hat{V}) = \frac{1}{i} \int_0^\infty e^{sQ} \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i(\mathbf{p}\cdot\mathbf{z} + s(\mathbf{k}\cdot\mathbf{p}))} \left[ \lambda\left(\mathbf{x}, \mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda\left(\mathbf{x}, \mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(t) ds.$$

Then as in the proof of Theorem 4.1 the right side of the limit equation is given by:

$$\mathcal{L}\lambda = -\frac{1}{i}\mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \left[ \lambda_1\left(\mathbf{k} - \frac{\mathbf{p}}{2}\right) - \lambda_1\left(\mathbf{k} + \frac{\mathbf{p}}{2}\right) \right] \psi(t) \right\} = I + I^*.$$

We have

$$\begin{aligned} I &= -\frac{1}{i}\mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \lambda_1\left(\mathbf{k} + \frac{\mathbf{p}}{2}\right) \psi(t) \right\} \\ &= \mathbb{E} \left\{ \int \frac{d\hat{V}(\mathbf{p})}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{z}} \int_0^\infty e^{sQ} \int \frac{d\hat{V}(\mathbf{q})}{(2\pi)^d} e^{i(\mathbf{q}\cdot\mathbf{z} + s(\mathbf{k} + \frac{\mathbf{p}}{2})\cdot\mathbf{q})} \left[ \lambda\left(\mathbf{k} + \frac{\mathbf{p}}{2} - \frac{\mathbf{q}}{2}\right) - \lambda\left(\mathbf{k} + \frac{\mathbf{p}}{2} + \frac{\mathbf{q}}{2}\right) \right] \psi^2(t) ds \right\} \\ &= \int_0^\infty \int \tilde{R}(s, \mathbf{p}) e^{-is(\mathbf{k} + \frac{\mathbf{p}}{2})\cdot\mathbf{p}} [\lambda(\mathbf{k} + \mathbf{p}) - \lambda(\mathbf{k})] \psi^2(t) \frac{d\mathbf{p}ds}{(2\pi)^d} \\ &= \int_0^\infty \int \tilde{R}(s, \mathbf{p} - \mathbf{k}) e^{is\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}} [\lambda(\mathbf{p}) - \lambda(\mathbf{k})] \psi^2(t) \frac{d\mathbf{p}ds}{(2\pi)^d}. \end{aligned}$$

The operator  $\mathcal{L}$  is given by

$$\mathcal{L}\lambda = \int_{\mathbb{R}^d} \hat{R}\left(\frac{\mathbf{k}^2 - \mathbf{p}^2}{2}, \mathbf{p} - \mathbf{k}\right) [\lambda(\mathbf{p}) - \lambda(\mathbf{k})] \psi^2(t) \frac{d\mathbf{p}}{(2\pi)^d}. \tag{98}$$

The rest of the proof of Theorem 5.1 is very similar to that of Theorem 4.1 and we omit the details.

### 6. DISCUSSION AND CONCLUSIONS

This paper considers the time splitting of the parabolic wave equation that models high frequency waves propagating in highly heterogeneous media. The role of the time splitting algorithm is to treat separately and successively propagation and scattering. Our main results are the following. The time step  $\Theta_\varepsilon$  needs to be chosen much smaller than the wavelength  $\varepsilon$  ( $\Theta_\varepsilon \ll \varepsilon^{3/2}$  for the time splitting algorithm (7)–(9) and  $\Theta_\varepsilon \ll \varepsilon^{5/4}$  for the Strang algorithm) to obtain convergence of the discrete solution to the exact solution as  $\varepsilon \rightarrow 0$  in the strong  $L^2$  sense. This however can be improved when only the macroscopic energy density of the waves is of

interest. When the time step  $\Theta_\varepsilon = \Theta\varepsilon$ , we find that the discrete energy density converges as  $\varepsilon \rightarrow 0$  to the solution of a radiative transfer equation with the scattering kernel that depends on  $\Theta$ . The correct scattering kernel is recovered only in the limit  $\Theta \rightarrow 0$  while the scattering kernel that corresponds to the (generally incorrect) white noise statistics of the random medium is obtained when  $\Theta \rightarrow \infty$ . This may have important consequences in practice as qualitative behavior of solutions of the radiative transport equations depends on the form of the power scattering kernel. For instance, the power spectrum  $\hat{R}$  has the form  $\tau_c \hat{R}(\tau_c \frac{\mathbf{p}^2 - \mathbf{k}^2}{2}, \mathbf{p} - \mathbf{k})$ , where  $\tau_c$  is temporal correlation length. It converges to  $\delta(\frac{|\mathbf{p}|^2 - |\mathbf{k}|^2}{2}) \hat{R}(\mathbf{p} - \mathbf{k})$  as  $\tau_c \rightarrow \infty$ , that is, in the limit of a quenched medium. Then the Wigner transform solves a diffusion equation [14, 31] in the long time limit. However, there is no diffusion limit in the case of white noise scattering as energy can freely cascade to high wavenumbers  $|\mathbf{k}|$ . Thus replacing the correct power spectrum numerically by a white noise approximation may lead to quite incorrect descriptions of the wave energy, even qualitatively. We note that in the specific case where the underlying medium has white noise statistics, one can actually show that the correct energy density is obtained for  $\Theta_\varepsilon \ll 1$ . This is easily obtained in the formal calculations and requires straightforward modifications in the rigorous theory.

The scaling  $\Theta_\varepsilon = \Theta\varepsilon$  marks then the transition between correct (small  $\Theta$ ) and incorrect (large  $\Theta$ ) limits for the wave energy in the generic case. Furthermore, we show that a modified time splitting algorithm allows us to obtain the correct statistics for  $\Theta_\varepsilon \ll 1$ . This algorithm, given in (87)–(88), however requires to introduce a modification in the phase space and applies to the Wigner transform of the wave field and not the wave solution itself. In this framework we observe that we can take first the limit  $\varepsilon \rightarrow 0$  and then  $\Theta_\varepsilon \rightarrow 0$  or vice-versa and obtain the same correct limiting equation for the wave energy; this is to be compared to the results [20, 22] that are similar in spirit.

We would like to stress that our results are very well adapted to the calculation of time reversed waves propagating in random media. Time reversal has received a lot of attention recently both in the physical [17, 18, 21] and mathematical [2, 3, 6, 7, 9, 10, 12, 28] literatures. The striking phenomenon of such time reversed waves is that they satisfy much better refocusing properties than the same waves propagating in a homogeneous medium. We refer to [6] for a framework that is convenient to our discussion. There we find that the refocused signal  $\psi^B(\boldsymbol{\xi}; \mathbf{x}_0)$  at a distance  $\boldsymbol{\xi}$  away from the center  $\mathbf{x}_0$  of the source term converges in the high frequency limit  $\varepsilon \rightarrow 0$  to

$$\psi^B(\boldsymbol{\xi}; \mathbf{x}_0) = \int_{\mathbb{R}^{2d}} e^{i\mathbf{k} \cdot (\boldsymbol{\xi} - \mathbf{y})} \overline{W}(L, \mathbf{x}_0, \mathbf{k}) \psi_0(\mathbf{y}) \frac{d\mathbf{y} d\mathbf{k}}{(2\pi)^d}, \tag{99}$$

where  $\psi_0$  is the source term and  $\overline{W}$  is a Wigner transform that models the time reversal experiment on a domain  $0 < z < L$  and solves a radiative transfer equation of the form (20). Moreover, convergence is stable in the sense that  $\psi^B$  is essentially independent of the realization of the random medium; see [6] for the details. The results we present in this paper show that  $\psi_{\Theta_\varepsilon}^B$ , obtained by solving the *parabolic wave* equation during the whole time reversal experiment with a time splitting method, converges to the correct limit  $\psi^B$  as  $\varepsilon \rightarrow 0$  provided that  $\Theta_\varepsilon \ll \varepsilon$ . Notice that since  $\psi^B$  is linear in  $\overline{W}$ , the whole detailed structure of the wave field  $\psi^B$  is captured and not only its energy density. Similarly when  $\Theta_\varepsilon = \Theta\varepsilon$  with  $\Theta = O(1)$ , one gradually passes from a correct characterization of the time reversed signal as  $\Theta \rightarrow 0$  to the generally incorrect time reversed signal one would obtain in a medium with white noise fluctuations as  $\Theta \rightarrow \infty$ .

Let us finally discuss possible generalizations of our results to other discretizations and equations. Although no rigorous proofs are available yet in these frameworks, we believe that our results extend to the time splitting algorithm for the full wave equation (1) and to first-order hyperbolic systems that model more general wave propagation phenomena [19, 31]. Time splitting algorithms based on separating propagation in homogeneous domain from scattering require us to use time steps  $\Theta_\varepsilon \ll \varepsilon$  to get the correct wave energy unless a modified algorithm in the phase space is used directly on the equation for the Wigner transform. A similar behavior should be expected in the discretization of the *space variables* since after all the “time splitting” algorithm corresponds to a discretization of the  $z$ -variable for the parabolic wave equation. In order to correctly capture the interplay between propagation and scattering, we believe that the typical mesh size  $\Delta x$  must satisfy  $\Delta x \ll \varepsilon$  to capture the correct statistics of the wave energy and be much smaller to capture the detailed structure of

the wave field. One should however be able to chose  $\Delta x \ll 1$  when a modified algorithm is used to directly calculate the Wigner transform in the phase space.

*Acknowledgements.* This work was supported by ONR grant N00014-02-1-0089, NSF Grants DMS-0072008 (GB) and DMS-0203537 (LR), and two Alfred P. Sloan Fellowships. This work was motivated by numerous discussions at the Stanford MGSS summer school. The authors also acknowledge the hospitality of the BIRS conference center at Banff/Lake Louise, Canada, where part of this work was completed.

## REFERENCES

- [1] G. Bal, On the self-averaging of wave energy in random media. *SIAM Multiscale Model. Simul.* **2** (2004) 398–420.
- [2] G. Bal and L. Ryzhik, Time reversal for classical waves in random media. *C. R. Acad. Sci. Paris I* **333** (2001) 1041–1046.
- [3] G. Bal and L. Ryzhik, Time reversal and refocusing in random media. *SIAM J. Appl. Math.* **63** (2003) 1475–1498.
- [4] G. Bal, A. Fannjiang, G. Papanicolaou and L. Ryzhik, Radiative transport in a periodic structure. *J. Statist. Phys.* **95** (1999) 479–494.
- [5] G. Bal, G. Papanicolaou and L. Ryzhik, Radiative transport limit for the random Schrödinger equations. *Nonlinearity* **15** (2002) 513–529.
- [6] G. Bal, G. Papanicolaou and L. Ryzhik, Self-averaging in time reversal for the parabolic wave equation. *Stochastics Dynamics* **4** (2002) 507–531.
- [7] G. Bal, T. Komorowski and L. Ryzhik, Self-averaging of the Wigner transform in random media. *Comm. Math. Phys.* **242** (2003) 81–135.
- [8] W. Bao, S. Jin and P.A. Markowich, On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime. *J. Comp. Phys.* **175** (2002) 487–524.
- [9] C. Bardos and M. Fink, Mathematical foundations of the time reversal mirror. *Asymptot. Anal.* **29** (2002) 157–182.
- [10] P. Blomgren, G. Papanicolaou and H. Zhao, Super-resolution in time-reversal acoustics. *J. Acoust. Soc. Am.* **111** (2002) 230–248.
- [11] S. Chandrasekhar, *Radiative Transfer*. Dover Publications, New York (1960).
- [12] 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.
- [13] G.C. Cohen, *Higher-order numerical methods for transient wave equations*. Scientific Computation, Springer-Verlag, Berlin (2002).
- [14] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Vol. 6, Springer-Verlag, Berlin (1993).
- [15] D.R. Durran, *Numerical Methods for Wave equations in Geophysical Fluid Dynamics*. Springer, New York (1999).
- [16] L. Erdős and H.T. Yau, Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation. *Comm. Pure Appl. Math.* **53** (2000) 667–735.
- [17] M. Fink, Time reversed acoustics. *Physics Today* **50** (1997) 34–40.
- [18] M. Fink, Chaos and time-reversed acoustics. *Physica Scripta* **90** (2001) 268–277.
- [19] P. Gérard, P.A. Markowich, N.J. Mauser and F. Poupaud, Homogenization limits and Wigner transforms. *Comm. Pure Appl. Math.* **50** (1997) 323–380.
- [20] F. Golse, S. Jin and C.D. Levermore, The convergence of numerical transfer schemes in diffusive regimes. I. Discrete-ordinate method. *SIAM J. Numer. Anal.* **36** (1999) 1333–1369.
- [21] 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. *J. Acoust. Soc. Am.* **105** (1999) 1597–1604.
- [22] T.Y. Hou, X. Wu and Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.* **227** (1999) 913–943.
- [23] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York, Academic (1978).
- [24] J.B. Keller and R. Lewis, Asymptotic methods for partial differential equations: The reduced wave equation and Maxwell’s equations, in *Surveys in applied mathematics*, J.B. Keller, D. McLaughlin and G. Papanicolaou Eds., Plenum Press, New York (1995).
- [25] P.-L. Lions and T. Paul, Sur les mesures de Wigner. *Rev. Mat. Iberoamericana* **9** (1993) 553–618.
- [26] P. Markowich, P. Pietra and C. Pohl, Numerical approximation of quadratic observables of Schrödinger-type equations in the semi-classical limit. *Numer. Math.* **81** (1999) 595–630.
- [27] P. Markowich, P. Pietra, C. Pohl and H.P. Stimming, A Wigner-measure analysis of the Dufort-Frankel scheme for the Schrödinger equation. *SIAM J. Numer. Anal.* **40** (2002) 1281–1310.

- [28] G. Papanicolaou, L. Ryzhik and K. Solna, The parabolic approximation and time reversal. *Matem. Contemp.* **23** (2002) 139–159.
- [29] G. Papanicolaou, L. Ryzhik and K. Solna, Statistical stability in time reversal. *SIAM J. App. Math.* **64** (2004) 1133–1155.
- [30] F. Poupaud and A. Vasseur, Classical and quantum transport in random media. *J. Math. Pures Appl.* **82** (2003) 711–748.
- [31] L. Ryzhik, G. Papanicolaou and J.B. Keller, Transport equations for elastic and other waves in random media. *Wave Motion* **24** (1996) 327–370.
- [32] H. Sato and M.C. Fehler, *Seismic wave propagation and scattering in the heterogeneous earth. AIP series in modern acoustics and signal processing*, AIP Press, Springer, New York (1998).
- [33] P. Sheng, *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*. Academic Press, New York (1995).
- [34] H. Spohn, Derivation of the transport equation for electrons moving through random impurities. *J. Stat. Phys.* **17** (1977) 385–412.
- [35] G. Strang, On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.* **5** (1968) 507–517.
- [36] F. Tappert, *The parabolic approximation method, Lect. notes Phys., Vol. 70, Wave propagation and underwater acoustics*. Springer-Verlag (1977).
- [37] B.J. Uscinski, *The elements of wave propagation in random media*. McGraw-Hill, New York (1977).
- [38] B.J. Uscinski, Analytical solution of the fourth-moment equation and interpretation as a set of phase screens. *J. Opt. Soc. Am.* **2** (1985) 2077–2091.