Uncertainty modeling and propagation in linear kinetic equations

Guillaume Bal Wenjia Jing Olivier Pinaud

May 19, 2017

Abstract

This paper reviews recent work in two complementary aspects of uncertainty quantification of linear kinetic models. First, we review the modeling of uncertainties in linear kinetic equations as high frequency limits of models for wave propagations in heterogeneous media. Second, we analyze the propagation of stochasticity from the constitutive coefficients to the solutions of linear kinetic equations. Such uncertainty quantifications find many important applications, e.g. in physics-based modeling of errors in inverse problems measurement.

1 Introduction

Uncertainties in measurements of interest come from a very broad spectrum of reasons. One such component arises from the propagation of uncertainty in the constitutive coefficients of a differential equation to the solutions of said equation. Some recent results obtained in the context of elliptic equations were summarized in the review [4]. Here, we consider such a propagation in the context of (phase space linear) transport equations and summarize results obtained primarily by the authors.

Two different propagations, occurring at physically different scales, are presented. The first one concerns the derivation of the transport equation from models of high frequency waves propagating in heterogeneous (scattering) media. Transport equations are deterministic models for wave field correlations (or wave field energies) of waves propagating in heterogeneous media modeled as random media. Their derivation may be seen as a homogenization (law of large number) result for phase space descriptions of field-field correlations; see the review [14]. As such, however, the field-field correlations are modeled by a deterministic equation even though the underlying wave fields are inherently random. Characterizing the random fluctuations (random corrections) in the field-field correlations remains a relatively little studied subject. Section 2 present several results obtained in this direction.

Once a kinetic model has been derived, either as an approximation for the energy of wave fields as described above or by any other means, it typically involves constitutive coefficients, such as scattering and absorbing coefficients, that depend on phase space variables and are typically not perfectly known. Such uncertainties have an effect on the transport solution. Recent results obtained in this direction are summarized in section 4 after relevant material and notation on the transport equation are presented in section 3. Several results in section 4 are based on moment estimates proved in [7] for specific random models. These moment estimates are presented in detail and generalized to a large class of sufficiently mixing coefficients in section 4.4.2.

The characterization of the random fluctuations in a transport solution is a problem of independent interest, and allows us to quantify the uncertainty in various functionals of the transport solution of interest. As for the elliptic case considered in [4], we mention two additional applications. The first one pertains to the calibration of upscaling numerical codes. We refer the reader to [10, 11] for such applications in the context of elliptic equations. The second one concerns inverse problems. In typical inverse problems, the reconstruction of the high frequency of the constitutive coefficients is unaccessible from inevitably noisy measurements. Yet, such not-reconstructed components, which we may as well model as random, have an influence on the solutions and hence the available measurements. Uncertainty propagation provides quantitative, physics-based, models for such an influence and allow for more accurate reconstructions of the low frequency components of the coefficients. For an application in the reconstruction of potentials from spectral information, see [21].

2 Uncertainties in the derivation of kinetic equations

2.1 Setting of the problem

In the context of wave propagation in heterogeneous media, kinetic equations generally describe quadratic quantities in the wavefield, for instance the wave energy. They are derived in the high frequency limit, and offer therefore an approximate description of the propagation up to some errors due to the finiteness of the frequency. Our goal in this section is to quantify these errors, in particular to obtain optimal convergence estimates, and when possible to characterize the first order corrector. We first review in section 2.2 the derivation of transport models and address the corrector analysis in section 2.3.

Kinetic models can be derived for several types of waves, e.g. acoustic, electromagnetic, quantum, and elastic waves. We will focus here on acoustic waves described by the scalar wave equation; see [47] for more general models. For $p(t,x)$ the wavefield, $\rho(x)$ the medium density, and $\kappa(x)$ its compressibility, our starting point is the wave equation

$$\frac{\partial^2 p}{\partial t^2} = \kappa(x)^{-1}\nabla\cdot\rho(x)^{-1}\nabla p + f(t,x), \qquad x\in\mathbb{R}^d,\ t>0, \tag{2.1}$$

supplemented with initial conditions $p(t=0,x)$ and $\frac{\partial p}{\partial t}(t=0,x)$. Above, f is a source term and d is spatial dimension. While the large scale features of the underlying heterogeneous medium are often known, or at least can be reconstructed, the fine details might not be accessible and are therefore modeled by a random medium with a given statistics. We then assume the following form for ρ and κ: $\rho(x)=\rho_0=1$ for simplicity (generalizations are possible), and

$$\kappa(x)^{-1} = \kappa_0(x)^{-1}\left(1+\sigma_0 V\left(\frac{x}{\ell_c}\right)\right).$$

In the latter equation, κ_0 is the background compressibility modeling the large scale structure of the medium (that we recall is supposed to be known), and V accounts for random fluctuations of strength σ_0 and correlation length ℓ_c, which model the fine details. The term V is a mean zero, stationary random field with correlation function

$$\mathbb{E}\{V(x)V(y)\} = R(x-y).$$

Above, $\mathbb{E}\{\cdot\}$ denotes ensemble average over the different realizations of the random medium.

We will present the kinetic equations derived from (2.1) in section 2.2 further. The wave equation is often reduced to a simpler model of propagation in order to make the mathematical analysis more amenable. This is done in the paraxial approximation that we introduce below.

The paraxial regime. The main assumption in this regime is that the waves propagate along a privileged direction, say z, and that backscattering is negligible. We then write $x = (z, x_\perp)$ accordingly, for $x_\perp \in \mathbb{R}^{d-1}$. We suppose moreover that κ_0 is constant, and introduce $c(x) = (\rho(x)\kappa(x))^{-1/2}$ as well as $c_0 = (\rho_0\kappa_0)^{-1/2}$. We also assume that the source f is supported in the region $z < 0$, and that the initial conditions vanish, that is $p(t = 0, x) = 0$ and $\frac{\partial p}{\partial t}(t = 0, x) = 0$. For $\hat{p}_\omega(z, x_\perp)$ the Fourier transform of p with respect to the variable t (after appropriate extension to $t < 0$) and with dual variable ω, we obtain the following Helmholtz equation:

$$\partial_z^2 \hat{p}_\omega + \Delta_{x_\perp} \hat{p}_\omega + \frac{\omega^2}{c(x)^2} \hat{p}_\omega = 0, \qquad z > 0. \tag{2.2}$$

For $k_\omega = \omega/c_0$, plugging the ansatz $\hat{p}_\omega(z, x_\perp) = e^{ik_\omega z}\psi_\omega(z, x_\perp)$ in (2.2), where the function ψ_ω is assumed to vary slowly in the z variable, and neglecting as a consequence the term $\partial_z^2\psi_\omega$, one obtains the Schrödinger equation

$$ik_\omega \partial_z \psi_\omega + \Delta_{x_\perp} \psi_\omega + \sigma_0 k_\omega^2 V \psi_\omega = 0, \qquad z > 0. \tag{2.3}$$

The equation is augmented with an initial condition $\psi_\omega(z = 0, x_\perp) = \psi_\omega^0(x_\perp)$, that depends on the source term f. See [50, 22] for more details about the paraxial approximation in heterogeneous media.

In the next section, we present the kinetic models obtained from asymptotics of (2.1) and (2.3). We start with the wave equation, and continue with the Schrödinger equation.

2.2 High frequency limit

2.2.1 The wave equation

We give here a formal derivation following the lines of [47]. Comments and references about rigorous results are given at the end of the section. We begin with the scalings.

We suppose here that the source term vanishes, i.e. $f = 0$, with non zero initial conditions. The kinetic limit is done in the regime of weak coupling [26, 48], where it is assumed that the strength of the fluctuations σ_0 is weak and that the correlation length of the medium ℓ_c and the wavelength of the initial condition λ are same order. The stochastic homogenization case $\lambda \gg \ell_c$ leads to waves propagating in an effective medium, see [2], while the case $\lambda \ll \ell_c$ leads to random Liouville equations [13]. If L is a typical distance of propagation, we then set

$$\frac{\ell_c}{L} = \frac{\lambda}{L} = \sigma_0^2 = \varepsilon \ll 1.$$

The fact that $\sigma_0 = \sqrt{\varepsilon}$ ensures the random medium has a non negligible effect at the macroscopic level. We rewrite (2.1) to obtain the following high frequency wave equation:

$$\frac{\partial^2 p^\varepsilon}{\partial t^2} = c_\varepsilon^2(x)\Delta p^\varepsilon, \qquad p^\varepsilon(t = 0, x) = p_0\left(\frac{x}{\varepsilon}\right), \qquad \frac{\partial p^\varepsilon}{\partial t}(t = 0, x) = p_1\left(\frac{x}{\varepsilon}\right), \tag{2.4}$$

where $c_\varepsilon(x) = (\rho_0 \kappa_\varepsilon(x))^{-1/2}$ with

$$\kappa_\varepsilon(x)^{-1} = \kappa_0(x)^{-1} \left(1 + \sqrt{\varepsilon} V\left(\frac{x}{\varepsilon}\right)\right).$$

The asymptotic analysis of (2.4) as $\varepsilon \to 0$ is done by means of Wigner transforms. We recast first the scalar wave equation as a first-order hyperbolic system on the acoustic field $\mathbf{u}^\varepsilon = (\mathbf{v}^\varepsilon, p^\varepsilon)$, where $\mathbf{v}^\varepsilon$ is velocity, and obtain the following system:

$$\rho_0 \frac{\partial \mathbf{v}^\varepsilon}{\partial t} + \nabla p^\varepsilon = 0, \qquad \kappa_\varepsilon \frac{\partial p^\varepsilon}{\partial t} + \nabla \cdot \mathbf{v}^\varepsilon = 0.$$

The system is augmented with initial conditions $p^\varepsilon(t=0,x)$ and $\mathbf{v}^\varepsilon(t=0,x) = \nabla \varphi^\varepsilon(x)$ where the pressure potential φ^ε is obtained by solving

$$\Delta \varphi^\varepsilon = -\kappa_\varepsilon \frac{\partial p^\varepsilon}{\partial t}(t=0,\cdot).$$

Kinetic equations. Wigner transforms provide a phase space description of the propagation of the wave energy, see [36, 44] for a detailed mathematical analysis of their properties. In the high frequency limit, the wave energy satisfies transport equations whose constitutive parameters are deduced from the sound speed c_ε. The Wigner transform of the field $\mathbf{u}^\varepsilon$ is defined as the following matrix-valued function,

$$W^\varepsilon(t,x,k) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\,k\cdot y}\, \mathbf{u}^\varepsilon(t, x - \frac{\varepsilon y}{2}) \otimes \mathbf{u}^\varepsilon(t, x + \frac{\varepsilon y}{2})\, dy.$$

It is shown in [47] that in the limit $\varepsilon \to 0$, the expectation of the Wigner transform $\mathbb{E}\{W^\varepsilon\}$ converges to a measure W admitting the following decomposition (there are no vortical modes because of the form of the initial condition):

$$W(t,x,k) = a^+(t,x,k)\, \mathbf{b}^+(k) \otimes \mathbf{b}^+(k) + a^-(t,x,k)\, \mathbf{b}^-(k) \otimes \mathbf{b}^-(k),$$

where we have defined

$$\mathbf{b}^\pm(k) = \frac{1}{\sqrt{2\rho_0}} \begin{pmatrix} \hat{k} \\ \pm \rho_0 c_0^{-1} \end{pmatrix}, \qquad \hat{k} = \frac{k}{|k|}.$$

The amplitude a^+ solves the radiative transfer equation below,

$$\frac{\partial a^+}{\partial t} + c_0\, \hat{k} \cdot \nabla_x a^+ + \Sigma(k) a^+ = Q(a^+), \tag{2.5}$$

where Q and Σ^{-1} are the collision operator and the mean free time, respectively. For δ the Dirac measure, they are given by,

$$Q(a)(k) = \int_{\mathbb{R}^d} a(p)\sigma(k,p)\delta(c_0|p| - c_0|k|)dp, \qquad \Sigma(k) = \int_{\mathbb{R}^d} \sigma(k,p)\delta(c_0|p| - c_0|k|)dp.$$

The cross section $\sigma(k,p)$ appearing in these expressions is

$$\sigma(k,p) = \frac{\pi c_0^2 |k|^2}{2(2\pi)^d} \hat{R}(k-p),$$

where $\hat{R}$ is the Fourier transform of the correlation function R with the convention

$$\hat{R}(p) = \int_{\mathbb{R}^d} e^{-ip\cdot x} R(x) dx.$$

A similar equation is obtained for $a^-(t,x,k) = a^+(t,x,-k)$. It is interesting to notice that the transport equation depends on the fluctuations of the random medium only through its power spectrum $\hat{R}$. The amplitude a^+ is related to the wave energy as follows. Defining the latter by

$$\mathcal{E}^\varepsilon(t,x) = \frac{1}{2}\left(\kappa_\varepsilon(x)(p^\varepsilon(t,x))^2 + \rho_0|\mathbf{v}^\varepsilon(t,x)|^2\right),$$

we have

$$\lim_{\varepsilon\to 0} \mathbb{E}\{\mathcal{E}^\varepsilon\}(t,x) = \int_{\mathbb{R}^d} a^+(t,x,k) dk.$$

The initial condition for (2.5) is the limit of the Wigner transform of the initial condition after appropriate projection, see [47].

Rigorous results. The main question is to justify of the convergence of W^ε to W, and to define in which sense this takes place. This is a difficult matter, and up to our knowledge there are only two references in the literature. The article [45] deals with lattice waves, described by an equation of the form (2.4) with Laplacian replaced by its finite differences approximation. Under various assumptions on the random potential V, it is shown that the average $\mathbb{E}\{W^\varepsilon\}$ converges in the distribution sense to W. The thesis [25] concerns precisely (2.4). The obtained results are much stronger: it is proved that the random process W^ε converges almost surely and in the distribution sense to the deterministic W. The quantity W^ε is then referred to as statistically stable, in the sense that it is weakly random for ε small. Such a property is at the core of transport-based imaging techniques in random media, see [18, 20]. Note that the proofs presented in [45] and [25] are technically involved and based on diagrammatic expansions.

2.2.2 The Schrödinger equation

The analysis is somewhat simpler than in the previous case. There are many possible scalings for (2.3), we present here only one and point the reader to [14] for other regimes. We consider a semi-classical Schrödinger equation of the form

$$i\varepsilon\partial_z\psi_\varepsilon + \frac{\varepsilon^2}{2}\Delta_{x_\perp}\psi_\varepsilon + \sqrt{\varepsilon}V\left(\frac{z}{\varepsilon}, \frac{x_\perp}{\varepsilon}\right)\psi_\varepsilon = 0, \qquad z > 0, \tag{2.6}$$

augmented with $\psi_\varepsilon(z=0,x_\perp) = \psi_\varepsilon^0(x_\perp)$. The Wigner transform is now scalar and given by

$$W^\varepsilon(z,x_\perp,k) = \frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}} e^{i\,k\cdot y}\,\psi_\varepsilon(z,x_\perp - \frac{\varepsilon y}{2})\psi_\varepsilon^*(z,x_\perp + \frac{\varepsilon y}{2})\,dy.$$

It is then shown that $\mathbb{E}\{W^\varepsilon\}$ converges to W, solution to

$$\frac{\partial W}{\partial z} + k\cdot\nabla_{x_\perp}W + \Sigma(k)W = Q(W), \tag{2.7}$$

where Q and Σ now read

$$\begin{aligned} Q(W)(k) &= \frac{1}{(2\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} \hat{R}\left(\frac{|p|^2}{2} - \frac{|k|^2}{2}, p-k\right) W(p)dp \\ \Sigma(k) &= \frac{1}{(2\pi)^{d-1}} \int_{\mathbb{R}^{d-1}} \hat{R}\left(\frac{|p|^2}{2} - \frac{|k|^2}{2}, p-k\right) dp. \end{aligned}$$

In these definitions, the power spectrum $\hat{R}(\omega, p)$ is still the Fourier transform of $R(x)$, written as $R(x) = R(z, x_\perp)$. The initial condition for (2.7) is the limit of the Wigner transform of the initial condition ψ_ε^0.

Rigorous results. More results are available than in the case of the wave equation, see [14] for a review. It is generally assumed that the random potential V is either a Markov process in the z variable, or a gaussian field. With these hypotheses, the proof of convergence follows from martingale and perturbed test functions methods, see e.g. [16, 37, 32]. The convergence holds in probability and in distribution. When V is independent of z, the analysis is more delicate, and a proof of convergence of $\mathbb{E}\{W^\varepsilon\}$ can be found in [31].

2.3 Corrector analysis

Now that the convergence of W^ε to W is established, we turn to the main topic of this section of the chapter, that is the error analysis. We are not aware of any results about the wave equation, and concentrate therefore on the paraxial regime. The first set of results concerns an even simpler description of the wave propagation offered by the Itô-Schrödinger equation. In a second step, we present the results that are available for more general models.

2.3.1 The Itô-Schrödinger regime

In this regime, the fluctuations of the random medium in the direction of propagation are supposed to be much faster than in the transverse plane. After an appropriate rescaling, and based on central limit type arguments, the random potential can be approximated (in a statistical sense) by a Brownian field. This leads to a Schrödinger equation of the form

$$id\psi_\varepsilon + \frac{\varepsilon}{2}\Delta_{x_\perp}\psi_\varepsilon + \psi_\varepsilon \circ dB\left(z, \frac{x_\perp}{\varepsilon}\right) = 0, \qquad z > 0, \tag{2.8}$$

where $\circ$ stands for the Stratonovich product and B is a Brownian field with autocorrelation

$$\mathbb{E}\{B(z, x'_\perp)B(z', x_\perp + x'_\perp)\} = \min(z, z')R_0(x_\perp).$$

This is the most amenable regime for an error analysis since Itô calculus can be used and yields closed-form equations. See [1, 33, 38] for a rigorous derivation of the Itô-Schrödinger equation. With W^ε the Wigner transform as before and W its limit, the goal is to quantify the error $W^\varepsilon - W$ as $\varepsilon \to 0$.

Full characterization of the corrector for a particular initial condition. The following results are taken from [42]. From (2.8), direct calculations show that the Wigner transform satisfies the equation

$$dW^\varepsilon(z,x_\perp,k)+k\cdot\nabla_{x_\perp}W^\varepsilon(z,x_\perp,k)$$
$$=\frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}}e^{ip\cdot x_\perp/\varepsilon}\left(W^\varepsilon(z,x_\perp,k-\frac{p}{2})-W^\varepsilon(z,x_\perp,k+\frac{p}{2})\right)\circ d\hat{B}(z,p),$$

where $\hat{B}(z,p)$ is the Fourier transform of $B(z,x_\perp)$ in the variable $x_\perp$. Suppose that initial Wigner transform satisfies

$$W^\varepsilon(z=0,x_\perp,k)=\delta(x_\perp)\varphi(k), \tag{2.9}$$

where δ is the Dirac measure and φ is a smooth function, and define the corrector

$$Z^\varepsilon(z,x_\perp,k)=\frac{W^\varepsilon(z,x_\perp,k)-W(z,x_\perp,k)}{\sqrt{\varepsilon}}.$$

The main result of [42] characterizes the limit of Z^ε as follows, see therein for the mathematical details and technical assumptions: the process Z^ε converges weakly in law to a process Z solution to the radiative transfer equation

$$\frac{\partial Z(z,x_\perp,k)}{\partial z}+k\cdot\nabla_{x_\perp}Z(z,x_\perp,k)=\frac{1}{(2\pi)^{d-1}}\int_{\mathbb{R}^{d-1}}\hat{R}_0(k-p)\left(Z(z,x_\perp,p)-Z(z,x_\perp,k)\right)dp, \tag{2.10}$$

with initial condition

$$Z(0,x_\perp,k)=\delta(x_\perp)X(k),$$

where $X(k)$ is a distribution valued Gaussian random variable. Its (somewhat complex) expression can be found in [42], the main information that it yields is that X is linear in the random potential B. The result can then be interpreted as follows: the leading instabilities are created in a boundary layer around the initial position $z=0$, and then propagate according to the kinetic equation (2.10). These instabilities are generated by the single scattering of the ballistic part of the limit W (i.e. the part of W that propagates freely in the medium and is exponentially decreasing) by the potential B.

This behavior is actually not universal as we will see in the next section. Indeed, other types of initial conditions lead to instabilities of different amplitudes and nature.

Characterization of the covariance of the corrector for more general initial conditions. The results presented here are taken from [17]. The error $W^\varepsilon - W$ is recast as $W^\varepsilon-\mathbb{E}\{W^\varepsilon\}+\mathbb{E}\{W^\varepsilon\}-W$. It is direct to see that the second piece is simply the propagation via the radiative transfer equation of the difference between the initial Wigner transform and its limit. We therefore focus on the first contribution $W^\varepsilon-\mathbb{E}\{W^\varepsilon\}$ to the error that is the most interesting. We analyze the covariance of the corrector $W^\varepsilon-\mathbb{E}\{W^\varepsilon\}$ and not the process itself, which allows us to consider more general settings. More specifically, we introduce the scintillation function J^ε defined by

$$J^\varepsilon(z,x_\perp,k,x'_\perp,k')=\mathbb{E}\{(W^\varepsilon-\mathbb{E}\{W^\varepsilon\})(z,x_\perp,k)(W^\varepsilon-\mathbb{E}\{W^\varepsilon\})(z,x'_\perp,k')\}.$$

In the Itô-Schrödinger regime, J^ε satisfies a closed-form equation (it does not for other classes of potentials), that reads

$$\Big(\frac{\partial}{\partial z} + \mathcal{T}_2 + 2R_0(0) - Q_2 - \mathcal{K}_\varepsilon\Big) J^\varepsilon = \mathcal{K}_\varepsilon a_\varepsilon \otimes a_\varepsilon, \tag{2.11}$$

equipped with vanishing initial conditions $J^\varepsilon(0, x_\perp, k, x'_\perp, k') = 0$ when the initial condition of the Schrödinger equation is deterministic. Here, we have defined

$$\begin{aligned}
a_\varepsilon &= \mathbb{E}\{W^\varepsilon\} \\
\mathcal{T}_2 &= k \cdot \nabla_{x_\perp} + p \cdot \nabla_{x'_\perp} \\
Q_2 h &= \int_{\mathbb{R}^{2(d-1)}} \Big(\hat{R}_0(k-p)\delta(k'-p') + \hat{R}_0(k'-p')\delta(k-p)\Big) h(x_\perp, p, x'_\perp, p') \frac{dpdp'}{(2\pi)^{2(d-1)}} \\
\mathcal{K}_\varepsilon h &= \sum_{\epsilon_i, \epsilon_j = \pm 1} \epsilon_i \epsilon_j \int_{\mathbb{R}^{2(d-1)}} \hat{R}_0(u) e^{i\frac{(x_\perp - x'_\perp)\cdot u}{\varepsilon}} h\left(x_\perp, k + \epsilon_i \frac{u}{2}, x'_\perp, k' + \epsilon_j \frac{u}{2}\right) \frac{du}{(2\pi)^{d-1}}.
\end{aligned}$$

Equation (2.11) is obtained by computing the fourth moment of the wave function. Consider an initial condition of the Schrödinger equation that has the form of a coherent state,

$$\psi^0_\varepsilon(x_\perp) = \frac{1}{\varepsilon^{\frac{(d-1)\alpha}{2}}} \chi\Big(\frac{x_\perp}{\varepsilon^\alpha}\Big) e^{i\frac{x_\perp \cdot k_0}{\varepsilon}},$$

where $\alpha \in [0,1]$, χ is a smooth function with compact support, and $k_0 \in \mathbb{R}^{d-1}$ is the direction of propagation in the transverse plane. The associated Wigner transform reads

$$W^\varepsilon_0(x_\perp, k) = \frac{1}{\varepsilon^{d-1}} W_0\left(\frac{x_\perp}{\varepsilon^\alpha}, \frac{k - k_0}{\varepsilon^{1-\alpha}}\right), \tag{2.12}$$

where W_0 is the Wigner transform of the rescaled initial condition $\psi^0_{\varepsilon=1}$ with $k_0 = 0$. The parameter α measures the concentration of the initial conditions in the spatial variables which, according the uncertainty principle, quantifies as well for coherent states the concentration in the momentum variables. Note that the initial condition (2.9) of the previous section corresponds essentially to the case $\alpha = 1$.

The limit of J^ε as $\varepsilon \to 0$ for initial conditions of the form (2.12) was characterized in [17]. We summarize here the most important points and refer the reader to [17] for complete results and formulas. We focus on the physical case $d = 3$.

- The most stable case corresponds to $\alpha = 0$, with a scintillation of order $\varepsilon^{d-1} = \varepsilon^2$, while the least stable case corresponds to $\alpha = 1$, with a scintillation of order ε. This is a somewhat intuitive result as when the support of the initial condition ψ^0_ε grows, we can expect the instabilities to be averaged over a larger domain.

- When $\alpha > 1/2$, the (appropriately rescaled) limit of J^ε satisfies a kinetic equation of the form (2.11) with $\mathcal{K}_\varepsilon = 0$ and a non vanishing initial condition. As in the previous section, instabilities are created in a boundary layer around the initial z position and then propagates according to a transport equation. When $\alpha \leq 1/2$, the limit of J^ε still satisfies a kinetic equation of the form (2.11) with $\mathcal{K}_\varepsilon = 0$, with now a vanishing initial condition and a non zero right-hand side in (2.11). In the latter configuration, instabilities are created as the wave propagates and not just around the initial position.

- There is a transition in the nature of the corrector defined by a critical value $\alpha^\star(\varepsilon)$ solution to the equation $\varepsilon^{2-3\alpha^\star(\varepsilon)} = \log \varepsilon^{2\alpha^\star(\varepsilon)-1}$ ($\alpha^\star$ is close to 2/3): when $\alpha < \alpha^\star(\varepsilon)$, the source term (when $\alpha \leq 1/2$) and the initial condition (for $\alpha > 1/2$) for the transport equation satisfied by the limit of J^ε are quadratic in the power spectrum $\hat{R}_0$; when $\alpha > \alpha^\star(\varepsilon)$, the initial condition is linear in $\hat{R}_0$. This is an interesting result as it shows that the leading instabilities are generated by the fraction of the wave that was scattered at most twice by the random medium. An interpretation of this fact is that instabilities are created by the most singular components of the wave, and since higher order scattering terms are more regular, they lead to negligible contributions.

As a conclusion, the main factor that influences the size of the corrector and its structure is the regularity of the initial condition, which was measured here in terms of concentration in phase space. The instabilities satisfy transport equations driven by either a source term or initial conditions. See [34, 35] for additional references on scintillation analysis. In the next section, we present a few results available in situations other than the Itô-Schrödinger regime we just considered.

2.3.2 Other regimes

Schrödinger equations with an Ornstein-Uhlenbeck potential. It is proved in [43], that as in the Itô-Schrödinger case addressed earlier, the corrector Z^ε converges weakly in law to a process Z_1 that has a similar structure as the limit Z solution to (2.10). Namely, Z_1 verifies a radiative transfer equation with a random initial condition. The proof is more involved than [42], as now one cannot rely on Itô calculus and one has to resort to diagrammatic expansions.

Schrödinger equations with z-independent random potentials. This case is the most difficult to study as there is not direct averaging induced by the z component of the potential. The Wigner transform of a solution to (2.6) satisfies the equation

$$\frac{\partial W^\varepsilon(z,x_\perp,k)}{\partial z} + k\cdot\nabla_{x_\perp} W^\varepsilon(z,x_\perp,k) \\ = \frac{i}{(2\pi)^{d-1}\sqrt{\varepsilon}}\int_{\mathbb{R}^{d-1}} e^{ip\cdot x_\perp/\varepsilon}\left(W(z,x_\perp,k-\frac{p}{2}) - W(z,x_\perp,k+\frac{p}{2})\right)\hat{V}(p)dp,$$

where $\hat{V}$ is the Fourier transform of the potential $V(x_\perp)$. The analysis of the scintillation J^ε is considerably more difficult in this situation since it does satisfy a closed-form equation. Motivated by the results of the previous section where it was shown that only the single and double scattering of the wave contribute to the instabilities in the Itô-Schrödinger regime, the terms in J^ε linear and quadratic in the power spectrum $\hat{R}$ were characterized in [15, 19]. The results are similar to those of the Itô-Schrödinger case as the largest corrector is of order ε and the smallest of order ε^{d-1}. There is also a transition for a critical α in the structure of the instabilities.

The case of long-range correlations in the underlying random medium is addressed in [15, 19]. This situation corresponds to a non-integrable correlation function and is modeled by a power spectrum of the form

$$\hat{R}(k) = \frac{S(k)}{|k|^\gamma}, \qquad 0 < \gamma < d.$$

The central observation made in [19] is that, contrary to the single scattering case of [15] where the scintillation is approximately of order $\sqrt{\varepsilon}$ when $\gamma \sim d$, the scintillation of the double scattering is of order $\varepsilon^{d-\delta}$ and therefore close to one when $\gamma \sim d$. This indicates that correlations in the medium have a strong influence on the size of the corrector. Determining whether or not high order scattering is statistically stable in media with long-range correlations requires the analysis of the whole series of W^ε in terms of $\hat{R}$ and remains an open problem.

3 Transport equation

For the rest of the paper, we consider the following boundary value problem for the stationary linear transport equation:

$$\begin{cases} v \cdot \nabla_x u(x,v) + a(x)u(x,v) = \displaystyle\int_V k(x,v,v')u(x,v')dv', & (x,v) \in X \times V, \\ u(x,v) = g(x,v), & (x,v) \in \Gamma_-. \end{cases} \tag{3.1}$$

Besides the applications in wave propagation considered in the preceding section, such equations model the transport and scattering of particles in the spatial domain X with velocities in V in the steady state regime. Here, $X \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded convex open set with smooth boundary ∂X. The velocity space V is in general a subset of $\mathbb{R}^d$ (typically excluding an open vicinity of 0) although only the case $V = S^{d-1}$ (with velocities constrained to the unit sphere) is considered here for simplicity. The particle density is prescribed on the incoming part of the boundary which, together with the outgoing boundary, is defined as

$$\Gamma_\pm := \{(x,v) \,:\, x \in \partial X,\, v \in V, \text{ and } \pm n_x \cdot v > 0\}. \tag{3.2}$$

Here, for a point x on the spatial boundary ∂X, n_x is the outer normal vector at the point. In particular, the projection of Γ_- onto its first component is precisely ∂X, and given $x \in \partial X$, the projection of Γ_- at x onto its second component consists of velocity vectors that point into the interior of X.

3.1 Decomposition of the transport solution

We review some basic theoretical results regarding the stationary linear transport equation with vanishing data on the incoming boundary Γ_- and with a source term f in X:

$$v \cdot \nabla_x u(x,v) + a(x)u(x,v) - \int_V k(x,v,v')u(x,v)dv' = f(x,v). \tag{3.3}$$

Such a situation arises naturally when differences of two solutions with the same incoming data are concerned; see [30, Chap. XXI] and [23, 27, 12, 49] for the detailed mathematical theory. In particular, the solution of the transport equation can be decomposed into a ballistic part, a single scattering part and the remaining multiple scattering part. The mathematical foundation for this decomposition can be explained as follows. The components of the integro-differential operator involved in (3.3) can be identified as

$$T_0\phi = v \cdot \nabla_x \phi, \qquad A_1\phi = a\phi, \qquad A_2\phi = -\int_V k(x,v,v')\phi(x,v')dv'.$$

$$T_1 = T_0 + A_1, \quad T = T_1 + A_2.$$

An appropriate functional setting, for $1 \le p \le \infty$, involves the spaces

$$\mathcal{W}^p := \{\phi \in L^p(X \times V),\ T_0\phi \in L^p(X \times V)\},$$

and the following differential or integro-differential operators:

$$\mathbf{T}_1\phi = T_1\phi\ \mathbf{T}\phi = T\phi,\ D(\mathbf{T}_1) = D(\mathbf{T}) = \{\phi \in \mathcal{W}^p,\ \phi|_{\Gamma_-} = 0\}.$$

The fact that a function in $\mathcal{W}^p$ has a trace on $\Gamma_\pm$ is proved in [27, 30].

The equation (3.3) is then understood as

$$\mathbf{T}u = f.$$

For simplicity, we restrict the problem to the so-called subcritical setting where the constitutive coefficients (a_r, k) are assumed to satisfy:

(S1) $a, k \ge 0$ a.e. on X, $a, k \in L^\infty(X)$ and k is isotropic.

(S2) There exists $\beta > 0$, such that $a - \pi_d k \ge \beta$ a.e. on X.

Here and below, π_d is the volume of the unit sphere S^{d-1}. When there is no scattering, the equation reduces to $\mathbf{T}_1 u = f$, and its solution is explicitly given by

$$u(x,v) = \mathbf{T}_1^{-1} f = \int_0^{\tau_-(x,v)} E(x, x - tv) f(x - tv, v) dt,$$

where E is a function defined by $E(x,y) := \exp\{-\int_0^{|x-y|} a(x - s\frac{x-y}{|x-y|})ds\}$. Under condition (S2), it is easy to verify (see e.g. [12]) that $\mathbf{T}_1^{-1}$ is a bounded linear transform on $L^p(X \times V)$, with an operator norm bounded by $\mathrm{diam}(X)\exp(-\beta\,\mathrm{diam}(X))$, where β is the positive constant in (S2) and $\mathrm{diam}(X)$ denotes the maximal distance between points in X.

The problem (3.3) can be viewed as a "perturbation" of the non-scattering transport. By using semi-group techniques, it is proved in [30] that $\mathbf{T}$ is invertible for all $1 \le p \le \infty$. The operator norm of $\mathbf{T}^{-1}$ has an upper bound that only depends on the parameter β in (S2). Moreover, when the constitutive coefficients are random, this bound can be made independent of ω as long as the parameter $\beta > 0$ in (S2) can be made uniform for almost all realizations; Let

$$\begin{aligned}\mathcal{K}u := A_2\mathbf{T}_1^{-1}u &= -\int_V \int_0^{\tau_-(x,v')} E(x, x - tv')k(x, v, v')u(x - tv', v')dtdv' \\ &= -\int_X \frac{E(x,y)k(x,v,v')}{|x-y|^{d-1}} u(y, v')dy,\end{aligned}$$

with $v' = (x-y)/|x-y|$. Then one checks that

$$\mathbf{T}^{-1} = \mathbf{T}_1^{-1}(I + \mathcal{K})^{-1}, \tag{3.4}$$

holds in the L^1 settings. This relation can be expanded to a truncated Neumann series with controlled remainder term. More precisely, one has

$$\mathbf{T}^{-1} f = \mathbf{T}_1^{-1}(f - \mathcal{K}f + \widetilde{\mathcal{K}}\mathcal{K}f). \tag{3.5}$$

The term $\mathbf{T}_1^{-1}f$ corresponds to the non-scattering transport and is referred to as the ballistic part, the term $-\mathbf{T}_1^{-1}\mathcal{K}f$ is the first order scattering part which takes considerations of particle trajectories that are scattered once, and $\mathbf{T}_1^{-1}\widetilde{\mathcal{K}}\mathcal{K}f$ corresponds to the multiple scattering part. Moreover, $\widetilde{\mathcal{K}}$ is a weakly singular integral operator on $L^1(X)$ with a kernel bounded by $C|x-y|^{-d+1}$ and, hence, the multiple scattering part is smoothing. The proof of this decomposition of $\mathbf{T}^{-1}$ can be found, for example, in [12, 7].

Remark 3.1. A parallel theory can be developed for the adjoint transport equation. Denote

$$\mathbf{T}_1^* u = -T_0 u + A_1 u, \;\; \mathbf{T}^* u = \mathbf{T}_1^* u - A_2' u,$$

and $D(\mathbf{T}^*{}_1) = D(\mathbf{T}^*) = \{u \in \mathcal{W}^p, u|_{\Gamma_+} = 0\}$, and A_2' is of the same form as A_2 but with the variables v and v' of k swapped. When k is assumed to be isotropic, then $A_2' = A_2$. In particular, $\mathbf{T}^{*-1}$ is a bounded linear transform on $L^p(X \times V)$ for all $p \in [1, \infty]$. The bound on the operator norm of $\mathbf{T}^{*-1}$ and the expansion formula still hold, provided that $\mathcal{K}$ is replaced by its formal adjoint. For any Hölder pair (p, q), $u \in L^p$ and $w \in L^q$, it holds that $\langle u, \mathbf{T}^{-1}w\rangle = \langle \mathbf{T}^{*-1}u, w\rangle$.

4 Uncertainty propagation in transport equations

In this section, we review the propagation of uncertainty from the constitutive coefficients to the solutions of kinetic equations. We start by modeling the uncertainty in the coefficients decomposed as a smoothly varying deterministic part plus a highly oscillating (in space) random part. We assume that the random parts are sufficiently mixing, or roughly speaking, very short-range correlated. We then review some homogenization and corrector theory that show how the uncertainty of the coefficients propagates to the random fluctuations in the kinetic solutions.

4.1 Uncertainty Modeling for the Constitutive Coefficients

The constitutive coefficients consist of two parts: a deterministic part that we assume to be known and an uncertain part, which we model as random fields. In many situations and in particular with the applications to inverse problems [21] in mind, we assume the uncertainty part is noise-like, in the sense that it models fluctuations around some average occurring on a very small scale $0 < \varepsilon \ll 1$. Hence, we set:

$$a_{r\varepsilon}\left(x, \frac{x}{\varepsilon}, \omega\right) = a_{r0}(x) + \mu\left(\frac{x}{\varepsilon}, \omega\right), \qquad k_\varepsilon\left(x, \frac{x}{\varepsilon}, \omega\right) = k_0(x) + \nu\left(\frac{x}{\varepsilon}, \omega\right), \tag{4.1}$$

for the intrinsic absorption coefficient and the isotropic scattering cross section, respectively. The deterministic part of the coefficients are $a_{r0}(x)$ and $k_0(x)$, and the uncertainty parts are modeled as random fields scaled from $\mu(y, \omega)$ and $\nu(y, \omega)$, which are mean-zero random fields defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The apparent total attenuation coefficient is then

$$a_\varepsilon\left(x, \frac{x}{\varepsilon}, \omega\right) = a_{r\varepsilon} + \int_{S^{d-1}} k_\varepsilon(x) d\theta = a_0(x) + \mu\left(\frac{x}{\varepsilon}, \omega\right) + \pi_d \nu\left(\frac{x}{\varepsilon}, \omega\right), \tag{4.2}$$

where $a_0(x) = a_{r0}(x) + \pi_d k_0(x)$. We henceforth refer to (3.1), with $a(x)$ replaced by $a_\varepsilon(x, x/\varepsilon, \omega)$ and with $k(x)$ replaced by $k_\varepsilon(x, x/\varepsilon, \omega)$, as the random linear transport equation. For notational convenience, the dependence of a_ε and k_ε on x/ε and ω is usually suppressed.

4.1.1 Stationarity, ergodicity, and mixing properties

At the macroscopic length scale (much larger than ε), the random fluctuations of the constitutive coefficients have a homogenized effect. This means that the solution of the random transport equation can be well approximated by that of an effective equation with homogenized coefficients. The (stochastic) error of this approximation should then be quantified. To rigorously establish such results, we assume that the random fields $\mu(x,\omega)$ and $\nu(x,\omega)$ are stationary, ergodic, and sufficiently mixing as defined below.

Stationarity and ergodicity. A random process $\mu(x,\omega)$ on $(\Omega,\mathcal{F},\mathbb{P})$ is called $\mathbb{R}^d$-stationary if for any positive integer k, for any k-tuple $(x_1,x_2,\cdots,x_k)\in(\mathbb{R}^d)^k$, and for any $z\in\mathbb{R}^d$,

$$(\mu(x_1,\cdot),\mu(x_2,\cdot),\cdots,\mu(x_k,\cdot)) \text{ and } (\mu(x_1+z,\cdot),\mu(x_2+z,\cdot),\cdots,\mu(x_k+z,\cdot),$$

have the same probability distribution. This means the statistics of the random field is homogeneous with respect to spatial translations.

An equivalent formulation can be done on the canonical probability space of μ, which is still denoted by $(\Omega,\mathcal{F},\mathbb{P})$. Then $\mathbb{R}^d$-stationarity means: there exist a random variable $\hat{\mu}:\Omega\to\mathbb{R}$ and a group action $\{\tau_z\}_{z\in\mathbb{R}^d}$ that is $\mathbb{P}$-preserving, and $\mu(x,\omega)=\hat{\mu}(\tau_x\omega)$. The set $\{\tau_z\}_{z\in\mathbb{R}^d}$ being a $\mathbb{P}$-preserving group action means: for every $z\in\mathbb{R}^d$, $\tau_z:\Omega\to\Omega$ preserves the measure $\mathbb{P}$, i.e. $\mathbb{P}(A)=\mathbb{P}(\tau^{-1}A)$ for every $A\in\mathcal{F}$, and $\tau_{z+y}=\tau_z\tau_y$. We say μ is ergodic if the underlying group action $\{\tau_z\}_{z\in\mathbb{R}^d}$ is ergodic. That is, if $A\in\mathcal{F}$ and $\tau_z A=A$ for all $z\in\mathbb{R}^d$, then $\mathbb{P}(A)\in\{0,1\}$.

Stationary and ergodicity are the essential properties of random fields that yield qualitative homogenization result which, in some sense, only captures the mean effect of the uncertainty. To quantify the convergence and to further study the uncertainties in the solutions, however, stronger assumption on the decorrelation structure of the random field, such as mixing properties, is often needed.

Mixing properties. We quantify the decorrelation structure of μ by the so-called "maximal correlation coefficient" ϱ, which is a decreasing function $\varrho:[0,\infty)\to[0,1]$, $\rho(r)\to 0$ as $r\to\infty$, and for each $r>0$, $\varrho(r)$ is the smallest value such that the bound

$$\mathbb{E}\left(V_1(\mu)V_2(\mu)\right)\le\varrho(r)\sqrt{\mathbb{E}\left(V_1^2(\mu)\right)\,\mathbb{E}\left(V_2^2(\mu)\right)} \tag{4.3}$$

holds for any two compact sets $K_1,K_2\in\mathcal{C}$ satisfying $d(K_1,K_2)\ge r$ and for any two random variables of the form $V_i(\mu)$, $i=1,2$, such that $V_i(q)$ is $\mathcal{F}_{K_i}$-measurable and $\mathbb{E}V_i(q)=0$.

Here, $\mathcal{C}$ denotes the set of compact sets in $\mathbb{R}^d$. Given $K\subset\mathcal{C}$, $\mathcal{F}_K$ is the σ-algebra generated by the random variables $\{\mu(x):x\in K\}$. For K_1,K_2 in $\mathcal{C}$, the distance $d(K_1,K_2)$ is defined to be

$$d(K_1,K_2)=\min_{x\in K_1,y\in K_2}|x-y|.$$

Remark 4.1. It is an important fact that ϱ-mixing fields are ergodic. For a stationary random field μ, the autocorrelation function is defined by

$$R(x):=\mathbb{E}\mu(x+y,\cdot)\mu(y,\cdot).$$

Note that by stationarity, the y variable in this definition does not play any role. $R_\mu(x)$ can be bounded by ϱ. Indeed, for any $x\in\mathbb{R}^d$,

$$|R_\mu(x)|=|\mathbb{E}(\mu(x)\mu(0))|\le\varrho(|x|)\|\mu\|^2_{L^2(\Omega)}. \tag{4.4}$$

In view of this relation, the decay rate of mixing coefficient ϱ yields decay rate of second order moments of μ. Such relations for higher order moments will be derived in section 4.4.2.

Main assumptions. We impose the following assumptions on the intrinsic absorption and scattering random fields $a_{r\varepsilon}$ and k_ε.

(A) Let $d = 2, 3$. We assume that the random fields $\mu(x, \omega)$ and $\nu(x, \omega)$ in (4.1) are $\mathbb{R}^d$-stationary and they admit a maximal correlation function ϱ satisfying $\varrho^{\frac{1}{8}} \in L^1(\mathbb{R}_+, r^{d-1}dr)$, that is

$$\int_0^\infty \varrho^{\frac{1}{8}}(r) r^{d-1} dr < \infty.$$

(B) Let the deterministic part (a, k) satisfies condition (S). Let μ and ν be uniformly bounded so that $a_r + \mu \geq \beta > 0$ a.e. in Ω.

Hypothesis (B), by the discussion in Section 3, guarantees that a.e. in Ω, the random transport equations are well posed. Let $(\mathbf{T}_\varepsilon)^{-1}$ be the inverse transport operator, it has bounded operator norm on $L^p(X \times V)$ essentially uniformly in Ω.

Hypothesis (A), in view of (4.4), implies that R_μ is integrable. Random fields with correlation functions satisfying such decay properties are referred to as short-range correlated. Since μ is stationary, the autocorrelation function $R_\mu(x, y)$ is a non-negative definite function in the sense that, for any $N \in \mathbb{N}$ and for any $x_i \in \mathbb{R}^n$, $y_j \in \mathbb{R}^n$, $i, j = 1, 2, \cdots, N$, the matrix $(M_{ij}) \in \mathbb{R}^{N\times N}$ given by $M_{ij} := R_\mu(x_i, y_j)$ is non-negative definite. Let

$$\sigma_\mu^2 = \int_{\mathbb{R}^n} R_\mu(x, 0)dx, \quad \sigma_\nu^2 = \int_{\mathbb{R}^n} R_\nu(x, 0)dx.$$

Then by (A) and (4.4), σ_μ^2 and σ_ν^2 are finite real numbers and, thanks to Bôchner's theorem, they are nonnegative. Throughout the paper, we assume those numbers are positive.

When the decay rate of ϱ is much weaker, so that (A) is violated, the random variations of the coefficients are in a different setting, and the quantitative results for the random fluctuations in the transport solutions will be changed; see section 4.4.1.

Remark 4.2 (Poisson bumps model)**.** Assumptions on the mixing coefficient ϱ of random media have been used in [3, 9, 39]; we refer to these papers for explicit examples of random fields satisfying the assumptions.

A widely used mixing random field model is the so-called Poisson bumps model; see e.g. [7]. The model is constructed as follows. We start with a *spatial Poisson point process* with intensity $\rho > 0$, which is a countable random subset $Y_\rho(\omega) := \{y_j(\omega) : j \in \mathbb{N}\} \subseteq \mathbb{R}^d$ defined on an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying, for any bounded Borel set $A \subseteq \mathbb{R}^d$, that the random variable $N(A)$, which is the cardinality of $A \cap Y_\rho$, follows the Poisson distribution with intensity $\rho|A|$, i.e.,

$$\mathbb{P}\{N(A) = m\} = \frac{e^{-\rho|A|}(\rho|A|)^m}{m!}. \tag{4.5}$$

See [28] for details. For any disjoint Borel sets $A_1, \cdots, A_n$, $n \geq 2$, the random variables $N(A_i)$, $1 \leq i \leq n$ are independent. The Poisson bumps model for the constitutive coefficients are then defined by

$$a_r(x; \omega) = \sum_j \psi(x - y_j(\omega)), \qquad k(x; \omega) = \sum_j \phi(x - y_j(\omega)). \tag{4.6}$$

Here, ψ and ϕ are smooth functions satisfying

$$0 \le \psi, \phi \le 1, \quad \phi(0) = \psi(0) = 1, \quad \text{and } \psi, \phi \text{ have compact supports.}$$

Using the properties of Poisson point process [28], one can show that a_r and k so defined are stationary and have finite range correlations and, hence, are mixing and ergodic. The mean values $\mathbb{E}(a_r(x,\cdot))$ and $\mathbb{E}k(x,\cdot)$ are constant; in fact, they are given by $\rho\|\psi\|_{L^1}$ and $\rho\|\phi\|_{L^1}$, respectively. Let

$$\mu(x,\omega) = a_r(x,\omega) - \mathbb{E}(a_r(x,\cdot)) \quad \text{and} \quad \nu(x,\omega) = k(x,\omega) - \mathbb{E}k(x,\cdot).$$

They are the mean-zero random parts of a_r and k.

Finally, by scaling the spatial variables, the formulas in (4.1) provide random field models for the constitutive coefficients which satisfies (A) and (B).

4.2 Homogenization result and the convergence rate

It is well known that the random transport equation homogenizes if the random fluctuations of the constitutive coefficients are periodic or stationary ergodic random fields, see e.g. [23, 29], and the effective coefficients are then given by the statistical average. The assumption (A) gives further quantitative information of the random fluctuations, such as the convergence rates. Throughout this section, u_ε denotes the solution to the random transport equation, and u_0 denotes the solution for the transport equation with the effective (mean) coefficients.

Theorem 4.3. *Assume (A) and (B). Assume further that the boundary data $g \in L^\infty(\Gamma_-)$ so $u_0 \in L^\infty(X \times V)$. Then there exists some constant $C > 0$ depending only on the diameter of X, $\|g\|_{L^\infty}$ and β such that, as $\varepsilon \to 0$,*

$$\left(\mathbb{E}\|u_\varepsilon - u_0\|^2_{L^2(X\times V)}\right)^{\frac{1}{2}} \le C\varepsilon^{\frac{1}{2}}. \tag{4.7}$$

This theorem implies that the convergence rate of the homogenization is of order $\sqrt{\varepsilon}$ in the energy norm $L^2(\Omega, L^2(X \times V))$. The proof of this theorem can be found in [7]. In particular, the main contribution comes from the energy of the ballistic parts which amounts to weighted average of the random fluctuations in the coefficients over lines. Since the central limit scaling of the average of random fields is $\varepsilon^{\mathfrak{d}/2}$, where $\mathfrak{d}$ is the dimension over which the average is taken, the scaling $\sqrt{\varepsilon}$ above is reasonable.

The leading terms in the homogenization error $u_\varepsilon - u$, i.e. the so-called correctors, contain two parts that play dominating roles at different scales. The mean-zero random part χ_ε is defined by

$$v \cdot \nabla_x \chi_\varepsilon + a(x)\chi_\varepsilon - k(x)\int_V \chi_\varepsilon(x,v')dv' = -\mu_\varepsilon(x)u_0(x,v) + \nu_\varepsilon(x)\int_V u_0(x,v')dv'. \tag{4.8}$$

We simply refer to it as $\mathbf{T}^{-1}A_\varepsilon u_0$, where the definition of A_ε can be read from above. χ_ε is the response of the transport equation to a source term given by $A_\varepsilon u_0$. The other part is deterministic and solves

$$v \cdot \nabla_x U + a(x)U - k(x)\int_V U(x,v')dv' = q(x,v), \tag{4.9}$$

where the source term $q(x,v)$ is given by:

$$\int_{\mathbb{R}}\Big(R_\mu(tv)u_0(x,v)-R_{\mu\nu}(tv)\overline{u}_0(x)-\int_V\big(R_{\mu\nu}(tw)u_0(x,w)-R_\nu(tw)\overline{u}_0(x)\big)dw\Big)dt.$$

Here, R_μ, R_ν and $R_{\mu\nu}$ correspond to, respectively, the auto-correlation functions of μ, ν, and the cross-correlation function of the two. Note also U is essentially the average of $\mathbf{T}^{-1}A_\varepsilon\chi_\varepsilon$. We have the following result.

Theorem 4.4. *Under the assumptions of Theorem 4.3, for any test function φ, we have*

$$|\langle\mathbb{E}(u_\varepsilon-u_0)-\varepsilon U,\varphi\rangle|\lesssim\varepsilon^2. \tag{4.10}$$

This theorem concerns the mean value of the homogenization error which lives on the large scale. Clearly, the term χ_ε is mean zero and does not contribute to $\mathbb{E}u_\varepsilon$, so the deterministic corrector U is indeed the dominating term. In the subsection below, we check that the variance of the homogenization error, integrated with a test function over space and velocity, is of order ε^d, and χ_ε is responsible for the main contribution.

4.3 Central limit theory for the random fluctuations

In this section, we study the random fluctuation $u_\varepsilon-\mathbb{E}u_\varepsilon$ and characterize their probability distribution at various observation scales.

Observations of the solutions. We consider three different scales of observation of the transport solutions:

- the pointwise data $u_\varepsilon(x,v)$ for fixed (x,v),
- the angularly averaged data $J_\varepsilon(x):=\int_V u_\varepsilon(x,v)\varphi(v)dv$ for fixed $x\in\overline{X}$ and some averaging kernel φ over V,
- the fully averaged data $\int_{X\times V}u_\varepsilon(x,v)\varphi(x,v)$ for some averaging kernel φ over $X\times V$.

When the application to inverse problems is considered, the angularly averaged or the fully averaged data are not uncommon. Below, we focus on the setting of fully averaged data since it is simpler to present. We first briefly comment on other settings of interest.

An interesting fact is that the size of the random fluctuations in u_ε depend on the observation scales. In [8], it is shown that, for pointwise data, the variance of $u_\varepsilon(x,v)$ is of order ε for all dimensions $d\geq 2$. This property arises from integrating random fields along (one-dimensional) lines. For the angularly averaged data $J_\varepsilon(x,\omega)$ defined above, its variance is of order $\varepsilon^2|\log\varepsilon|$ in dimension two and ε^2 in dimension $d\geq 3$. Angular averaging introduces additional mixing along the angular direction and therefore significantly reduces the variance of the corrector. Finally, as we show in detail below, the variance of the fully averaged data $\langle u_\varepsilon,\varphi\rangle$ is of order ε^d in dimension $d\geq 2$. The random corrector is therefore of smallest size when averaged over the whole phase space. This is consistent with the central limit theorem.

Fluctuation theory for the fully averaged data. We take a sufficiently smooth functions φ and consider the fully averaged data $\langle\varphi,u_\varepsilon-\mathbb{E}u_\varepsilon\rangle$. To describe the limiting probability distribution

of this random variable, we introduce additional notation. Let $\psi := \mathbf{T}^{*-1}\varphi$; in other words, ψ is is the solution of the following adjoint transport equation:

$$\begin{cases} -v\cdot\nabla_x\psi + a\psi - \displaystyle\int_V k(x,v,v')\psi(x,v')dv' = \varphi, & (x,v)\in X\times V,\\ \psi(x,v) = 0, & (x,v)\in\Gamma_+. \end{cases} \tag{4.11}$$

We denote by $\rho_{\mu,\nu}$ the correlation factor between the random fields μ and ν; it is defined by $\sigma_\mu^{-1}\sigma_\nu^{-1}\int_{R^n}\mathbb{E}\mu(x,\cdot)\nu(0,\cdot)dx$. Let Σ denote the non-negative matrix

$$\Sigma := \begin{pmatrix} \sigma_\mu^2 & \rho_{\mu,\nu}\sigma_\mu\sigma_\nu \\ \rho_{\mu,\nu}\sigma_\mu\sigma_\nu & \sigma_\nu^2 \end{pmatrix}.$$

We have the following result.

Theorem 4.5. *Let the dimension $d = 2, 3$. Under the same condition of Theorem 4.3, we have*

$$\left\langle \varphi, \frac{u_\varepsilon - \mathbb{E}u_\varepsilon}{\varepsilon^{d/2}} \right\rangle \xrightarrow[\varepsilon\to 0]{\text{distribution}} \int_X \left[\int_V \begin{pmatrix} -\psi(x,v)u_0(x,v) \\ \psi(x,v)[-\pi_d u_0(x,v) + \langle u_0\rangle_V(x)] \end{pmatrix} dv\right]\cdot\sqrt{\Sigma}dW(y). \tag{4.12}$$

Here, $W(y) = (W_a(y), W_k(y))'$ is a two dimensional $2d$-parameter Wiener process and $\langle u_0\rangle_V(x)$ denotes the angular average $\int_V u_0(x,v')dv'$.

This theorem says that the random fluctuations of the fully averaged data are of order $\varepsilon^{\frac{d}{2}}$, satisfy a functional central limit theory with limiting Gaussian distribution. The right-hand side of (4.12) is a Gaussian distribution $\mathcal{N}(0,\sigma^2)$ with the variance σ^2 given by

$$\int_X \left[\int_V \begin{pmatrix} -\psi(x,v)u_0(x,v) \\ \psi(x,v)[-\pi_d u_0(x,v) + \langle u_0\rangle_V(x)] \end{pmatrix} dv\right]^T \Sigma \left[\int_V \begin{pmatrix} -\psi(x,v)u_0(x,v) \\ \psi(x,v)[-\pi_d u_0(x,v) + \langle u_0\rangle_V(x)] \end{pmatrix} dv\right] dx.$$

We refer the reader to [41] for more details on multi-parameter Wiener process. An $\mathbb{R}^2$ valued Wiener process is needed in the limit since two random fields are involved in the random transport equation. For the Poisson bumps model of Remark 4.2, the correlation factor $\rho_{\mu,\nu}$ above is 1, but more general situations still based on Poisson point process may be considered; see [7].

Finally, it is worth mentioning that for $d = 2, 3$, the term $\mathbb{E}u_\varepsilon$ in the statement of the theorem above can be replaced by εU of Theorem 4.4. In other words, εU is the only term in the mean error $\mathbb{E}(u_\varepsilon - u_0)$ that is larger than the random fluctuations. The competition between the deterministic and the random parts of the homogenization error is, hence, clearly characterized in dimensions $d \le 3$.

Outline of the proof. We now outline the proofs of the main theorems. It clearly appears from these proofs that the key ingredients about the random coefficients that allow one to quantify the homogenization are moments estimates.

The required moments estimates and their decay rates were obtained in [7] for the Poisson bumps model and strongly depended on the statistical properties of Poisson point processes. At the end of this section, we extend such moments estimates to general random fields whose mixing coefficient ϱ, defined in (4.3), decays sufficiently fast. The results of [7] thus generalize to the versions stated in this review.

To get the convergence rate in Theorem 4.3, we observe that χ_ε defined in (4.8) satisfies

$$u_\varepsilon - u = \chi_\varepsilon + \mathbf{T}_\varepsilon^{-1} A_\varepsilon \chi_\varepsilon.$$

Since $\mathbf{T}_\varepsilon^{-1}$ and A_ε are uniformly bounded (in ε and $\omega \in \Omega$) linear transformations on $L^2(X \times V)$, due to assumption (B), it suffices to show that $\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \lesssim \varepsilon$. Using the representation formula (3.5) of $\mathbf{T}^{-1}$, we have rather explicit formula for χ_ε and quantifying its mean square norm is straightforward, given the second order moments of the random fields.

For Theorems 4.4 and 4.5, the starting point is the following expansion formula:

$$\begin{aligned} u_\varepsilon - u_0 = \mathbf{T}^{-1} A_\varepsilon u_0 + \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0 + \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0 \\ + \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon (u_\varepsilon - u_0). \end{aligned} \tag{4.13}$$

This formula is obtained by iterating the relation $u_\varepsilon - u_0 = \mathbf{T}^{-1} A_\varepsilon u_\varepsilon$ three times, and the iteration process could be continued further. We focus on the case $d = 2$ below. For the case $d = 3$, another iteration is needed for the argument below to hold.

Consider the inner product of (4.13). The first term on the right-hand side is mean-zero and has no contribution. The second term is $\langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, \psi \rangle$ where $\psi = \mathbf{T}^{*-1}\varphi$. Using second order moments again, the main contribution to its statistical mean comes from $\mathbb{E}\langle A_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, \psi \rangle$, and it converges to $\langle U, \varphi \rangle$. Using the third and higher order moments of mixing random fields obeying (A), all the remainder terms are of smaller orders than ε^2. Theorem 4.4 is then proved. Note that the remainder terms in the mean is even smaller than the central limit scaling $\varepsilon^{d/2}$.

For the fluctuation theory, we consider the variances of the terms in (4.13) after taking inner products with the test function and subtracting the mean values. The first term has variance of order ε^d, agreeing with the central limit setting. To prove Theorem 4.5, it suffices to establish

$$\text{Var } (\langle \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, \varphi \rangle) \ll O(\varepsilon^d), \tag{4.14}$$

$$\text{Var } (\langle \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, \varphi \rangle) \ll O(\varepsilon^d), \tag{4.15}$$

$$\mathbb{E}|\langle \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon (u_\varepsilon - u_0), \varphi \rangle| \ll O(\varepsilon^{\frac{d}{2}}), \tag{4.16}$$

$$\varepsilon^{-d/2} \langle \mathbf{T}^{-1} A_\varepsilon u_0, \varphi \rangle \text{ converges in distribution to the right hand side of (4.12).} \tag{4.17}$$

Let us start with the estimate (4.16) for the remainder. This is an $L^1(\Omega)$ estimate, which is sufficient to show that this term, divided by $\varepsilon^{d/2}$, converges to zero in distribution. We apply Hölder inequality and get

$$\begin{aligned} \mathbb{E}|\langle \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon (u_\varepsilon - u_0), \varphi \rangle| &\le C \mathbb{E}\{\|u_\varepsilon - u_0\|_{L^2} \|\mathbf{T}^{*-1} A_\varepsilon \mathbf{T}^{*-1} A_\varepsilon \psi\|_{L^2}\} \\ &\le C \left(\mathbb{E}\|u_\varepsilon - u_0\|_{L^2}^2\right)^{\frac{1}{2}} \left(\mathbb{E}\|\mathbf{T}^{*-1} A_\varepsilon \mathbf{T}^{*-1} A_\varepsilon \psi\|_{L^2}^2\right)^{\frac{1}{2}}. \end{aligned}$$

This reduce (4.16) to $\mathbb{E}\|\mathbf{T}^{*-1} A_\varepsilon \mathbf{T}^{*-1} A_\varepsilon \psi\|_{L^2}^2 \ll O(\varepsilon^{d-1})$. Using fourth-order moments, this term is of order ε^2, and hence the desired result follows for $d = 2$. This estimate also shows that, for $d = 3$, another iteration is needed in (4.13).

For the second and the third terms in the expansion (4.13), by similar estimates as above, we see that their $L^1(\Omega)$ norms are not small enough. Hence we need the variance estimates, which are enough for the convergence in distribution. For (4.14), we appeal to the fourth-order variance

estimate in Proposition 4.8, and for (4.15), we appeal to the sixth-order variance estimate in Proposition 4.11. From a technical point of view, the variances are of much smaller order because, in the corresponding variance estimates (4.20) and (4.27) for the products of the random fields, the terms which are responsible for the largeness of the $L^1(\Omega)$ controls are eliminated.

Finally, the limiting distribution of $\varepsilon^{-d/2}\langle u_\varepsilon - u_0, \varphi\rangle$ is given by that of $\frac{1}{\sqrt{\varepsilon^d}}\langle \mathbf{T}^{-1}A_\varepsilon u_0, \varphi\rangle$, which has the expression

$$\frac{1}{\sqrt{\varepsilon^d}}\int_X \mu\left(\frac{x}{\varepsilon},\omega\right)\langle\psi u_0\rangle_V(x) + \nu\left(\frac{x}{\varepsilon},\omega\right)\left[-\pi_d\langle\psi\rangle_V(x)\langle u_0\rangle_V(x) + \langle\psi u_0\rangle_V(x)\right]dx.$$

Again, $\langle\psi u_0\rangle_V$ denotes the angular average of the pointwise product ψu_0. This is an oscillatory integral of random fields with short range correlations, and the central limit theorem of such integrals was proved in [24, 3]. The above integral hence has the desired limit as in (4.12). This proves Theorem 4.5.

In three dimensions, $d = 3$, the analysis is more involved. A further iteration should be added to (4.13). The resulted remainder term, and the first three terms are controlled as above. An additional term appears, which, after taking inner product with the test function, becomes $\langle \mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon u_0, \varphi\rangle$. We consider its $L^2(\Omega)$ norm

$$\mathbb{E}\left|\langle \mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon u_0, \varphi\rangle\right|^2 = \mathbb{E}\left|\langle A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon u_0, \psi\rangle\right|^2.$$

By appealing to the eighth-order moment estimates, one can check that this term is of order $o(\varepsilon^3)$ and hence does not contribute to the limit.

4.4 Further remarks

4.4.1 Long range correlated random media

When the random fluctuations in the constitutive coefficients have long range correlations and the assumptions (A) is violated, it is still possible to quantify the convergence rate and to analyze the random part of the homogenization error, if sufficient quantitative information about the random fields is given.

For instance, in [8], random fields μ and ν of the form $\Phi(g(x,\omega))$ are studied, where $g(x,\omega)$ is some underlying Gaussian random field with heavy tail $R_g(x)$, and $R_g(r) \sim r^{-\alpha}$ for some $0 < \alpha < 1$ asymptotically at infinity. Here, Φ is a bounded real function with sufficient regularity and of Hermite rank one; see [10, 46]. It is then shown that the pointwise data have random fluctuations of order $\sqrt{\varepsilon^\alpha}$, which is much larger than $\sqrt{\varepsilon}$ for the short range setting. We refer the reader to [6, 5, 40] for more discussions on random fields with long range correlations.

4.4.2 Moments estimates for mixing random fields

In [7], Theorems 4.3, 4.4 and 4.5 were proved only for the Poisson bumps model. We now show that such results hold for more general random fields satisfying (A) as the moment and variance estimates enjoyed by the Poisson points model also hold for generalized mixing random fields.

Below, we first recall the crucial moments formulae (hence moments estimates) of the Poisson bumps model, and then show that those estimates hold for sufficiently mixing random fields. Even

though much more general results can be produced, only the first several (up to eighth-order) moments estimates are provided.

Moments formulas for the Poisson bumps model. Let n be a positive integer and let I_n denote the index set $\{1, 2, \cdots, n\}$. Given a set $\{x_1, x_2, \cdots, x_n\} \subset \mathbb{R}^d$ and $J \subseteq I_n$, x_J denotes the subset $\{x_j : j \in J\}$. For a random field $\mu(x, \omega)$, we are interested in getting estimates for

$$\Phi_\mu^{(n)}(x_{I_n}) := \mathbb{E}[\mu(x_1)\mu(x_2)\cdots\mu(x_n)],$$

which is the nth order moments of the random field μ evaluated at x_{I_n}. We note that Φ is viewed as a function of set-valued arguments, since the order of the elements in the set plays no role. In the sequel, the dependence on μ is omitted when the random field under study is clear.

For the set I_n, we say $(n_1, n_2, \cdots, n_k)$ is a k-partition of I_n if $1 \le n_1 \le n_2 \le \cdots \le n_k$ and $\sum_{i=1}^k n_i = n$. A partition is called *non-single* if $n_1 \ge 2$. We denote by $\mathscr{G}_n$ the set of all non-single partitions of I_n. Given $(n_1, n_2, \cdots, n_k)$, there are finitely many possible ways to divide I_n (hence x_{I_n}) into k disjoint subsets $J_1, \cdots, J_k$ of cardinalities $n_1, n_2, \cdots, n_k$, respectively. We denote this finite number by $C_n^{n_1,n_2,\cdots,n_k}$, and we order those possibilities following the dictionary order of the array formed by $(\max J_1, \cdots, \max J_k)$. The ℓ-th choice is hence denoted by $(J_1^\ell, \cdots, J_k^\ell)$.

Proposition 4.6 ([7]). *Let $\mu(x, \omega)$ be the mean-zero part of the Poisson bumps potential. Fix a positive integer n. For any integer $k \le n$ and any subset $J \subseteq I_n$ with cardinality k, define*

$$T^{(k)}(x_J) := \nu \int \prod_{j \in J} \psi(x_j - z)dz. \tag{4.18}$$

Then we have the following formula for $\Phi^{(n)}(x_{I_n})$

$$\Phi^{(n)}(x_{I_n}) = \sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n} \sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}} \prod_{j=1}^k T^{(n_j)}(x_{J_j^\ell}). \tag{4.19}$$

In [7], we also need to have estimates on variances of products of μ evaluated at several points. More precisely, if p is a positive integer such that $p \le n/2$, we are interested in

$$\Psi^{(p,n-p)}(x_{I_p}, x_{I_n\setminus I_p}) := \mathbb{E}\left[\left(\prod_{j\in I_p}\mu(x_j) - \Phi^p(x_{I_p})\right)\left(\prod_{k\in I_n\setminus I_p}\mu(x_k) - \Phi^{n-p}(x_{I_n\setminus I_p})\right)\right],$$

which is the covariance of the random variables $\prod_{j\in I_p}\mu(x_j)$ and $\prod_{k\in I_n\setminus I_p}\mu(x_k)$. It is clear that formulae for those covariance functions can be read from the formulae for moments of μ.

Moments estimates for mixing random fields. In the rest of this section, we fix a random field $\mu(x, \omega)$ satisfying the assumption (A). All of the moment and variance functions Φ and Ψ are understood as those of μ. The first theorem deals with the third order moment.

Proposition 4.7. *Let $C = \|\mu\|_{L^2(\Omega)}\|\mu^2\|_{L^2(\Omega)}$, and let $\eta(r) = \sqrt{\rho(r/2)}$. Then*

$$\left|\Phi^{(3)}(\{x_1, x_2, x_3\})\right| \le C\eta(|x_2 - x_1|)\eta(|x_3 - x_1|).$$

From the proof below, it is clear that after a permutation of (x_1, x_2, x_3) on the right hand side, the resulted estimate still holds. Hence, the right hand side can be thought as depending only on the set $\{x_1, x_2, x_3\}$.

Before proceeding to the proof of this result, we introduce some more notation. Recall that $I_n := \{1, 2, \cdots, n\}$. We consider dividing the set I_n into two subsets of positive cardinality p and $n - p$. For each $0 < p \leq n/2$, let

$$\mathcal{I}_n^p = \{J \,:\, J \subseteq I_n, \, \mathrm{card}(J) = p\}.$$

Given a $J \in \mathcal{I}_n^p$, let J^c denote the complement of J in I_n. Then (J, J^c) corresponds to a division of I_n, and (x_J, x_{J^c}) corresponds to a division of the set x_{I_n} into two subsets of cardinality p and $n - p$. Define

$$L_p = \max_{J \in \mathcal{I}_n^p} \mathrm{dist}\,(x_J, x_{J^c})\,.$$

Then L_p is the maximum separation distance among divisions of x_{I_n} into two subsets of cardinality p and $n - p$. Finally, let

$$L = \max_{0 < p \leq n/2} L_p.$$

Then L is the maximum separation distance between possible divisions of x_{I_n} into two subsets of positive cardinalities.

Proof of Proposition 4.7. In view of the discussion above, since $n = 3$, the largest integer smaller or equal to $n/2$ is 1. So, $L = L_1$. Let $J \in \mathcal{I}_3^1$ be $\arg\max L_1$. We consider two scenarios.

If $x_1 \in x_J$, without loss of generality, we assume $L = |x_1 - x_2|$. It is then clear that $|x_3 - x_1| \leq 2L$ because if otherwise, $|x_3 - x_2| \geq |x_3 - x_1| - |x_2 - x_1| > L$ and hence $\{x_3\} \cup \{x_1, x_2\}$ would be a division yielding larger L_1. Applying the mixing condition, we get

$$\begin{aligned}
\Phi^{(3)}(x_1, x_2, x_3) &= \mathbb{E}\{\mu(x_1)[\mu(x_2)\mu(x_3) - \mathbb{E}(\mu(x_2)\mu(x_3))]\} \\
&\leq \varrho(L)\|\mu\|^3_{L^\infty(\Omega)} \\
&\leq \varrho^{1/2}(|x_1 - x_2|)\varrho^{1/2}(|x_1 - x_3|/2)\|\mu\|^3_{L^\infty(\Omega)} \\
&\leq \eta(|x_1 - x_2|)\eta(|x_1 - x_3|).
\end{aligned}$$

If $x_1 \in x_{J^c}$, without loss of generality, assume the division is given by $\{x_2\} \cup \{x_1, x_3\}$. We consider two further sub-cases. If the separating distance is $L = |x_2 - x_1|$, then $|x_1 - x_3| \leq L$. If, instead, the separating distance is $L = |x_2 - x_3|$, then $|x_1 - x_3| \leq L$ and $L \leq |x_1 - x_2| \leq 2L$. In both cases, we have

$$\begin{aligned}
\Phi^{(3)}(x_1, x_2, x_3) &= \mathbb{E}\{V(x_2)[V(x_1)V(x_3) - \mathbb{E}(V(x_1)V(x_3))]\} \\
&\leq \varrho(L)\|\mu\|^3_{L^\infty(\Omega)} \leq \eta(|x_1 - x_2|)\eta(|x_1 - x_3|).
\end{aligned}$$

The conclusion of the theorem, hence, is established. □

The fourth-order moment was considered in [39] and in [3]; see also [10]. For the convenience of the reader, we recall the following estimate of [39].

Proposition 4.8. *Let* $C = 4(\|\mu\|_{L^2(\Omega)}\|\mu^3\|_{L^2(\Omega)} + \|\mu^2\|^2_{L^2(\Omega)})$ *and let* $\eta(r) = \sqrt{\rho(r/3)}$. *Then*

$$\left|\Phi^{(2,2)}(x_1,x_2,x_3,x_4) - R(x_1-x_2)R(x_3-x_4)\right| \leq C\eta(|x_1-x_3|)\eta(x_2-x_4)+C\eta(|x_1-x_4|)\eta(|x_2-x_3|). \tag{4.20}$$

Note that the estimate (4.4) implies $|R(x-y)| \leq \eta(|x-y|)$. The variance estimate above also yields the following moment estimate:

$$\left|\Phi^{(4)}(x_{I_4})\right| \leq \frac{C}{2}\sum_{J\in\mathcal{I}_4^2}\eta(x_J)\eta(x_{J^c}). \tag{4.21}$$

For more general results, the following fact will be helpful.

Proposition 4.9. *Let* $n \in \mathbb{N}$ *and* $n \geq 2$, *and let* I_n *and* L *be defined as above, then*

$$\operatorname{dist}(x_j,x_k) \leq 2(n-1)L, \qquad \forall j,k \in I_n. \tag{4.22}$$

Proof. Let L be achieved by a division (x_J, x_{J^c}). Let $p = \operatorname{card}(J)$ and, without loss of generality, $L = \operatorname{dist}(x_1,x_2)$ for some $x_1 \in J$, $x_2 \in J^c$. We show that

$$\operatorname{dist}(x_j,\{x_1,x_2\}) \leq (n-1)L, \qquad \forall j \in I_n. \tag{4.23}$$

If this fails, say $\operatorname{dist}(x_3,x_2) > (n-1)L$, then since $L \geq L_1$, there must be a point, say x_4, such that $x_4 \in \overline{B}_L(x_3)$. Indeed, if otherwise, then $L_1 \geq \operatorname{dist}(x_3, x_{I_n\setminus\{3\}}) > L$, which is impossible. Similarly, since $L \geq L_2$, there must be another point, say x_5, such that $x_5 \in \overline{B}_L(x_3)\cup\overline{B}_L(x_4)$. Note that $x_5 \in \overline{B}_{2L}(x_3)$. By repeating this argument, we hence find p points

$$\{x_3,x_4,\cdots,x_{3+p-1}\} \subseteq \overline{B}_{(p-1)L}(x_3). \tag{4.24}$$

Let $q = n-p$. Applying the same argument above with $\{x_3\}$ replaced by $\{x_1,x_2\}$, we find that the remaining set $x_{I_n\setminus\{3,\cdots,3+p-1\}}$, which contains $\{x_1,x_2\}$, must satisfy

$$\{x_{3+p},\cdots,x_n\} \subseteq \overline{B}_{(q-2)L}(x_1)\cup\overline{B}_{(q-2)L}(x_1) \subseteq \overline{B}_{(q-1)L}(x_2). \tag{4.25}$$

Let $K = \{3,4,\cdots,3+p-1\}$ and $K^c = \{1,2,3+p,\cdots,n\}$. Then (4.24), (4.25) and the assumption $\operatorname{dist}(x_2,x_3) > (n-1)L$ imply that $\operatorname{dist}(x_K,x_{K^c}) > L$, which is impossible. Therefore, (4.23) holds and, by an application of triangle inequality, (4.22) is established. □

We move to the fifth-order moments, and obtain the following estimate.

Proposition 4.10. *Let* $C = \|\mu\|_{L^2(\Omega)}\|\mu^4\|_{L^2(\Omega)} + \|\mu^2\|_{L^2(\Omega)}\|\mu^3\|_{L^2(\Omega)} + \|\mu\|^3_{L^2(\Omega)}\|\mu^2\|_{L^2(\Omega)}$. *Let* η *be defined as in Proposition 4.7 and define* $\psi^{(3)}(\{x,y,z\})$ *as* $\varrho^{\frac{1}{4}}(|x-z|)\varrho^{\frac{1}{4}}(|y-z|)$. *Then*

$$\left|\Phi^{(5)}(x_{I_5})\right| \leq C\sum_{J\in\mathcal{I}_5^2}\eta(\tfrac{1}{4}|x_{j_1}-x_{j_2}|)\psi^{(3)}(\tfrac{1}{8}\{x_{J^c}\}). \tag{4.26}$$

Proof. Recall the definition of L, L_1 and L_2 after the statement of Proposition 4.7. We only need to consider two cases.

Case 1: $L = L_1 \geq L_2$. Without loss of generality, let $L = \mathrm{dist}(x_1, x_2) = \mathrm{dist}(\{x_1\}, \{x_2, \cdots, x_5\})$. Then we have, due to the mixing property of the random field μ,

$$\left|\Phi^{(5)}(x_{I_5})\right| = \left|\mathbb{E}\left(\mu(x_1)\left[\prod_{j=2}^5 \mu(x_j) - \Phi^{(4)}(\{x_2, \cdots, x_5\})\right]\right)\right| \leq \varrho(L)\|\mu\|_{L^2(\Omega)}\|\mu^4\|_{L^2(\Omega)}.$$

In view of (4.22), the right-hand side above is bounded by $C\varrho(|x_1 - x_2|)\psi^{(3)}(\{x_3, x_4, x_5\})$. Since ϱ and hence η are decreasing functions, this bound is smaller than some of the terms of the right-hand side of (4.26).

Case 2: $L = L_2 > L_1$. Let L be maximized by the division given by $x_J = \{y_1, y_2\}$ and $x_{J^c} = \{y_3, y_4, y_5\}$. Then, by the mixing property,

$$\left|\Phi^{(5)}(x_{I_5}) - R(y_1 - y_2)\Phi^{(3)}(\{y_3, y_4, y_5\})\right| \leq \varrho(L)\|\mu^2\|_{L^2(\Omega)}\|\mu^3\|_{L^2(\Omega)}.$$

We can find some $J' = (j_1', j_2') \in \mathcal{I}_5^2$ such that $J' \neq J$. Then in view of (4.22), the right-hand side above can be bounded by $C\varrho^{1/2}(\frac{1}{8}|x_{j_1'} - x_{j_2'}|)\psi^{(3)}(\frac{1}{8}x_{(J')^c})$.

Moreover, in view of (4.22), we have

$$\left|R(y_1 - y_2)\Phi^{(3)}(\{y_3, y_4, y_5\})\right| \leq C\varrho(|y_1 - y_2|)\psi^{(3)}(\{y_3, y_4, y_5\}).$$

The bounds we have for the preceding two quantities correspond to two different terms in the right-hand side of (4.26). The desired result is hence established. □

Next, we study the sixth-order moments of μ. We first derive a variance type estimate, from which the moment estimate follows easily. Let $\mathcal{I}_6^{2,2,2}$ denote the set

$$\{(J_1, J_2, J_3) \,:\, \mathrm{card}(J_1) = \mathrm{card}(J_2) = \mathrm{card}(J_3) = 2, \cup_{i=1}^3 J_i = I_6\},$$

which is the collection of partitions of I_6 into three disjoint subsets of cardinality 2.

Proposition 4.11. *Let $C = \|\mu\|_{L^2}\|\mu^5\|_{L^2} + \|\mu^2\|_{L^2}(\|\mu^4\|_{L^2} + \|\mu\|_{L^3}^3\|\mu\|_{L^2}) + \|\mu^3\|_{L^2}^2$ with $\|\cdot\|_{L^p} \equiv \|\cdot\|_{L^p(\Omega)}$. Let η be defined as in Proposition 4.7 and let $\psi^{(3)}$ be defined as in Proposition 4.10. Define $\xi^{(2)} = \eta^{\frac{2}{3}}$. Then*

$$\begin{aligned}\frac{1}{C}\left|\Psi^{(3,3)}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\})\right| \leq & \sum_{(J_1,J_2,J_3)\in\mathcal{I}_6^{2,2,2}} \xi^{(2)}(\tfrac{1}{20}x_{J_1})\xi^{(2)}(\tfrac{1}{20}x_{J_2})\xi^{(2)}(\tfrac{1}{20}x_{J_3}) \\ & + \frac{1}{2}\sum_{K\in\mathcal{I}_6^3\setminus\{\{1,2,3\},\{4,5,6\}\}} \psi^{(3)}(\frac{1}{20}x_K)\psi^3(\frac{1}{20}x_{K^c}).\end{aligned} \tag{4.27}$$

We note that, as in the variance estimate (4.20), the partition $\{1, 2, 3\} \cup \{4, 5, 6\}$ does not appear on the right-hand side.

Proof. Recall the definition of L, L_1, L_2 and L_3 after the statement of Proposition 4.7. We study several cases.

Case 1: $L = L_1 \geq \max(L_2, L_3)$. Without loss of generality, let $\arg\max L = \{1\}$. Then we have

$$\left|\Phi^{(6)}(x_{I_6})\right| = \left|\mathbb{E}(\mu(x_1)[\textstyle\prod_{j\in\{1\}^c}\mu(x_j) - \Phi^{(5)}(x_{\{1\}^c})])\right| \leq \varrho(L)\|\mu\|_{L^2(\Omega)}\|\mu^5\|_{L^2}.$$

Meanwhile, $\operatorname{dist}(x_1, x_j) \geq L$ for $j = 2, 3$. It follows that $\operatorname{dist}(\{x_1\}, \{x_2, x_3\}) \geq L$ and

$$\left|\Phi^{(3)}(\{x_1, x_2, x_3\})\Phi^{(3)}(\{x_4, x_5, x_6\})\right| \leq \|\mu\|^3_{L^3(\Omega)}\|\mu\|_{L^2(\Omega)}\|\mu^2\|_{L^2(\Omega)}\varrho(L). \tag{4.28}$$

Finally, in view of (4.22), we may choose any $J \in \mathcal{I}_6^3$ such that $J \neq \{1, 2, 3\}$, and we verify that

$$\varrho(L) \leq \left(\varrho^{\frac{1}{4}}(\tfrac{1}{10}|x_{j_1} - x_{j_2}|)\varrho^{\frac{1}{4}}(\tfrac{1}{10}|x_{j_1} - x_{j_3}|)\right)\left(\varrho^{\frac{1}{4}}(\tfrac{1}{10}|x_{j_4} - x_{j_5}|)\varrho^{\frac{1}{4}}(\tfrac{1}{10}|x_{j_4} - x_{j_6}|)\right),$$

where $J = \{j_1, j_2, j_3\}$ and $J^c = \{j_4, j_5, j_6\}$. It follows that in this case,

$$\left|\Psi^{(3,3)}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\})\right| \leq C\psi^{(3)}(\tfrac{1}{10}x_J)\psi^{(3)}(\tfrac{1}{10}x_{J^c}),$$

which is a term on the right hand side of (4.27).

Case 2: $L = L_2 \geq L_3$ and $L_2 > L_1$. Renaming the points, we assume L is obtained by the division $x_J = \{y_1, y_2\}$, $x_{J^c} = \{y_3, \cdots, y_6\}$ and $\operatorname{dist}(y_2, y_3) = L$. Then we note that

$$\left|\Phi^{(6)}(x_{I_6}) - \Phi^{(2)}(\{y_1, y_2\})\Phi^{(4)}(\{y_3, \cdots, y_6\})\right| \leq \varrho(L)\|\mu^2\|_{L^2(\Omega)}\|\mu^4\|_{L^2(\Omega)}.$$

Moreover, if $x_J \subseteq \{x_1, x_2, x_3\}$ or $x_J \subseteq \{x_4, x_5, x_6\}$, then (4.28) holds. If otherwise, then we may assume that $y_2 = x_1$ and $y_3 = x_4$. Since $\{x_2, x_3\}$ must contain at least one point from x_{J^c}, and because $\operatorname{dist}(\{x_1\}, x_{J^c}) = L$, we conclude that $\max_{j=2,3}\operatorname{dist}(x_1, x_j) \geq L$. It follows that

$$\max_{j\in I_3}\operatorname{dist}(\{x_j\}, x_{I_3}\setminus\{x_j\}) \geq L/2. \tag{4.29}$$

Then we have

$$\left|\Phi^{(3)}(\{x_1, x_2, x_3\})\Phi^{(3)}(\{x_4, x_5, x_6\})\right| \leq \|\mu\|^3_{L^3(\Omega)}\|\mu\|_{L^2(\Omega)}\|\mu^2\|_{L^2(\Omega)}\varrho(L/2). \tag{4.30}$$

We then repeat the argument in Case 1 to control the $\varrho(L)$ and $\varrho(L/2)$ terms. Finally, we get

$$\left|\Psi^{(3,3)}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\})\right| \leq \left|R(y_1 - y_2)\Phi^{(4)}(\{y_3, y_4, y_5, y_6\})\right| + C\psi^{(3)}(\tfrac{1}{20}x_K)\psi^{(3)}(\tfrac{1}{20}x_{K^c})$$

where both K and K^c are different from $\{1, 2, 3\}$. For the $|R\Phi^{(4)}|$ term above, we combine the estimates (4.4) and (4.21) to get

$$\left|R(y_1 - y_2)\Phi^{(4)}(\{y_3, y_4, y_5, y_6\})\right| \leq \frac{C}{2}\varrho^{1/3}(|y_1 - y_2|)\sum_J \eta(x_J)\eta(x_{\{3,4,5,6\}\setminus J})$$

where J runs in the set $\{J \subset \{3, 4, 5, 6\} \,|\, \operatorname{card}(J) = 2\}$. Since by definition $\eta \leq \xi^{(2)}$, the error bound above is dominated by some term on the right-hand side of (4.27).

Case 3: $L = L_3 > \max(L_1, L_2)$. If $\text{dist}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_5\}) = L$, then we have

$$|\Psi^{(3,3)}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\})| \le \varrho(L)\|\mu^3\|^2_{L^2(\Omega)}.$$

If otherwise, $\text{dist}(\{x_1, x_2, x_3\}, \{x_4, x_5, x_5\}) < L$ and there exists $x_J = \{y_1, y_2, y_3\}$, with $x_J \neq \{1, 2, 3\}$ but $x_J \cap \{x_1, x_2, x_3\} \neq \emptyset$, such that $\text{dist}(x_J, x_{J^c}) = L$. Then we have

$$\left|\Phi^{(6)}(x_{I_6}) - \Phi^{(3)}(x_J)\Phi^{(3)}(x_{J^c})\right| \le \varrho(L)\|\mu^3\|^2_{L^2(\Omega)}.$$

Without loss of generality, assume $x_1 \in J$, $x_4 \in J^c$ and $\text{dist}(x_1, x_4) = L$. Then $\{x_2, x_3\} \cap J^c$ is non-empty, and the element in this intersection has distance larger than L from x_1. Hence, (4.29) holds, and the rest of the analysis can be carried out as in Case 2.

In all three cases, we can bound the left-hand side of (4.27) by some terms on the right-hand side, and, hence, the desired result is established. □

As a corollary, we have the following estimate for the full sixth-order moments:

Corollary 4.12. *Let C be defined as in Proposition 4.11. Then*

$$\frac{1}{C}\left|\Phi^{(6)}(x_{I_6})\right| \le \sum_{(J_1,J_2,J_3)\in\mathcal{I}_6^{2,2,2}} \xi^{(2)}(\tfrac{1}{10}x_{J_1})\xi^{(2)}(\tfrac{1}{10}x_{J_2})\xi^{(2)}(\tfrac{1}{10}x_{J_3}) + \tfrac{1}{2}\textstyle\sum_{K\in\mathcal{I}_6^3}\psi^{(3)}(\tfrac{1}{20}x_K)\psi^3(\tfrac{1}{20}x_{K^c}).$$

Finally, we have the following result for the eighth-order moments. Define

$$\begin{aligned}\mathcal{I}_8^{2,2,2,2} &:= \{(J_1, J_2, J_3, J_4) : \text{card}(J_1) = \text{card}(J_2) = \text{card}(J_3) = \text{card}(J_4) = 2, \cup_{i=1}^4 J_i = I_8\},\\ \mathcal{I}_8^{2,3,3} &:= \{(J_1, J_2, J_3) : \text{card}(J_1) = 2, \text{card}(J_2) = \text{card}(J_3) = 3, \cup_{i=1}^3 J_i = I_8\}.\end{aligned}$$

$\mathcal{I}_8^{2,2,2,2}$ is the collection of partitions of I_8 into four mutually disjoint subsets, each of which has cardinality two. $\mathcal{I}_8^{2,3,3}$ is the collection of partitions of I_8 into four mutually disjoint subsets of cardinalities two, three and three, respectively.

Proposition 4.13. *Define $\zeta^{(2)} = \sqrt{\eta}$ and $\phi^{(3)} = \sqrt{\psi^{(3)}}$. Then there exists some constant $C > 0$ so that*

$$\begin{aligned}\frac{1}{C}\left|\Phi^{(8)}(x_{I_8})\right| \le &\sum_{(J_1,J_2,J_3,J_4)\in\mathcal{I}_8^{2,2,2,2}} \zeta^{(2)}(\tfrac{1}{14}x_{J_1})\zeta^{(2)}(\tfrac{1}{14}x_{J_2})\zeta^{(2)}(\tfrac{1}{14}x_{J_3})\zeta^{(2)}(\tfrac{1}{14}x_{J_4})\\ &+ \sum_{J\in\mathcal{I}_8^{2,3,3}} \eta(\tfrac{1}{14}x_{J_1})\phi^{(3)}(\tfrac{1}{28}x_{J_2})\phi^{(3)}(\tfrac{1}{28}x_{J_3}).\end{aligned}$$

This result can be proved using the same methods as in the proofs of Proposition 4.10 and Proposition 4.11. We do not reproduce the details.

The function $\phi^{(3)}$ defined above has the expression $\phi^{(3)}(\{x_1, x_2, x_3\}) = \varrho^{\frac{1}{8}}(x_2 - x_1)\varrho^{\frac{1}{8}}(x_3 - x_1)$. For the fluctuation theory that will be reviewed in Section 4, we need $\phi^{(3)}$ to be integrability for each of its variables, and hence $\varrho^{\frac{1}{8}}$ should decay sufficiently fast. This explains the integrability condition of the maximal correlation function ϱ that is required in assumption (A).

Acknowledgments

GB's contribution was partially funded by NSF Grant DMS-1408867 and ONR Grant N00014-17-1-2096. OP's work is supported by NSF CAREER Grant DMS-1452349.

References

[1] F. Bailly, J. F. Clouet, and J.-P. Fouque. Parabolic and gaussian white noise approximation for wave propagation in random media. *SIAM J. Appl. Math*, 56(5):1445–1470, 1996.

[2] G. Bal. Homogenization in random media and effective medium theory for high frequency waves. *Discrete Contin. Dyn. Syst. Ser. B*, 8(2):473–492 (electronic), 2007.

[3] G. Bal. Central limits and homogenization in random media. *Multiscale Model. Simul.*, 7(2):677–702, 2008.

[4] G. Bal. *Propagation of Stochasticity in Heterogeneous Media and Applications to Uncertainty Quantification*, pages 1–24. Springer International Publishing, Cham, 2016.

[5] G. Bal, J. Garnier, Y. Gu, and W. Jing. Corrector theory for elliptic equations with long-range correlated random potential. *Asymptot. Anal.*, 77(3-4):123–145, 2012.

[6] G. Bal, J. Garnier, S. Motsch, and V. Perrier. Random integrals and correctors in homogenization. *Asymptot. Anal.*, 59(1-2):1–26, 2008.

[7] G. Bal and W. Jing. Homogenization and corrector theory for linear transport in random media. *Discrete Contin. Dyn. Syst.*, 28(4):1311–1343, 2010.

[8] G. Bal and W. Jing. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 112(4):660–670, 2011.

[9] G. Bal and W. Jing. Corrector theory for elliptic equations in random media with singular Green's function. Application to random boundaries. *Commun. Math. Sci.*, 19(2):383–411, 2011.

[10] G. Bal and W. Jing. Corrector theory for MSFEM and HMM in random media. *Multiscale Model. Simul.*, 9(4):1549–1587, 2011.

[11] G. Bal and W. Jing. Corrector analysis of a heterogeneous multi-scale scheme for elliptic equations with random potential. *ESAIM Math. Model. Numer. Anal.*, 48(2):387–409, 2014.

[12] G. Bal and A. Jollivet. Stability estimates in stationary inverse transport. *Inverse Probl. Imaging*, 2(4):427–454, 2008.

[13] G. Bal, T. Komorowski, and L. Ryzhik. Self-averaging of Wigner transforms in random media. *Comm. Math. Phys.*, 242(1-2):81–135, 2003.

[14] G. Bal, T. Komorowski, and L. Ryzhik. Kinetic limits for waves in random media. *Kinetic Related Models*, 3(4):529 – 644, 2010.

[15] G. Bal, I. Langmore, and O. Pinaud. Single scattering estimates for the scintillation function of waves in random media. *J. Math. Phys.*, 51(2):022903, 18, 2010.

[16] G. Bal, G. Papanicolaou, and L. Ryzhik. Self-averaging in time reversal for the parabolic wave equation. *Stochastics and Dynamics*, 4:507–531, 2002.

[17] G. Bal and O. Pinaud. Dynamics of Wave Scintillation in Random Media. *CPDE*, 35(7):1176–1235, 2010.

[18] G. Bal and O. Pinaud. Imaging using transport models for wave-wave correlations. *M3AS*, 21(5):1071–1093, 2011.

[19] G. Bal and O. Pinaud. Analysis of the double scattering scintillation of waves in random media. *CPDE*, 38(6):945–984, 2013.

[20] G. Bal and K. Ren. Transport-based imaging in random media. *SIAM Applied Math.*, 68(6):1738–1762, 2008.

[21] G. Bal and K. Ren. Physics-based models for measurement correlations: application to an inverse Sturm-Liouville problem. *Inverse Problems*, 25(5):055006, 13, 2009.

[22] A. Bamberger, E. Engquist, L. Halpern, and P. Joly. Parabolic wave equation approximations in heterogeneous media. *SIAM J. Appl. Math.*, 48:99–128, 1988.

[23] A. Bensoussan, J.-L. Lions, and G. C. Papanicolaou. Boundary layers and homogenization of transport processes. *Publ. Res. Inst. Math. Sci.*, 15(1):53–157, 1979.

[24] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Ann. Probab.*, 10(4):1047–1050, 1982.

[25] M. Butz. Kinetic limit for Wave Propagation in a Continuous, Weakly Random Medium. *Ph.D. thesis, TU Munich*, 2015.

[26] S. Chandrasekhar. *Radiative Transfer*. Dover Publications, New York, 1960.

[27] M. Choulli and P. Stefanov. An inverse boundary value problem for the stationary transport equation. *Osaka J. Math.*, 36(1):87–104, 1999.

[28] D. R. Cox and V. Isham. *Point processes*. Chapman & Hall, London, 1980. Monographs on Applied Probability and Statistics.

[29] A.-L. Dalibard. Homogenization of linear transport equations in a stationary ergodic setting. *Comm. Partial Differential Equations*, 33(4-6):881–921, 2008.

[30] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6*. Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.

[31] L. Erdös and H. T. Yau. Linear Boltzmann equation as the weak coupling limit of a random Schrödinger Equation. *Comm. Pure Appl. Math.*, 53(6):667–735, 2000.

[32] A. C. Fannjiang. Self-averaging scaling limits for random parabolic waves. *Arch. Ration. Mech. Anal.*, 175(3):343–387, 2005.

[33] J. Garnier and K. Sølna. Coupled paraxial wave equations in random media in the white-noise regime. *Ann. Appl. Probab.*, 19(1):318–346, 2009.

[34] J. Garnier and K. Sølna. Scintillation in the white-noise paraxial regime. *CPDE*, 39:626–650, 2014.

[35] J. Garnier and K. Sølna. Fourth-moment analysis for wave propagation in the white-noise paraxial regime. *ARMA*, 220:37–81, 2016.

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

[37] C. Gomez. Radiative transport limit for the random Schrödinger equation with long-range correlations. *J. Math. Pures. Appl.*, 98:295–327, 2012.

[38] C. Gomez and O. Pinaud. Fractional white-noise limit and paraxial approximation for waves in random media. *submitted*.

[39] M. Hairer, E. Pardoux, and A. Piatnitski. Random homogenisation of a highly oscillatory singular potential. *Stoch. Partial Differ. Equ. Anal. Comput.*, 1(4):571–605, 2013.

[40] W. Jing. Limiting distribution of elliptic homogenization error with periodic diffusion and random potential. *Anal. PDE*, 9(1):193–228, 2016.

[41] D. Khoshnevisan. *Multiparameter processes*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2002. An introduction to random fields.

[42] T. Komorowski, S. Peszat, and L. Ryzhik. Limit of fluctuations of solutions of Wigner equation. *Comm. Math. Phys.*, 292(2):479–510, 2009.

[43] T. Komorowski and L. Ryzhik. Fluctuations of solutions to wigner equation with an ornstein-uhlenbeck potential. *Discr. Cont. Dyn. Sys. B*, 17:871–914, 2012.

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

[45] J. Lukkarinen and H. Spohn. Kinetic limit for wave propagation in a random medium. *Arch. Ration. Mech. Anal.*, 183:93–162, 2007.

[46] V. Pipiras and M. S. Taqqu. Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields*, 118(2):251–291, 2000.

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

[48] P. Sheng. *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*. Academic Press, New York, 1995.

[49] P. Stefanov and G. Uhlmann. An inverse source problem in optical molecular imaging. *Anal. PDE*, 1(1):115–126, 2008.

[50] F. Tappert. *The parabolic approximation method, Lecture notes in physics, vol. 70. Wave Propagation and Underwater Acoustics*, Ed. J.B. Keller and J.S. Papadakis, Springer-Verlag, pp. 224-287, 1977.