Homogenization and Corrector Theory for Linear Transport in Random Media

Guillaume Bal Wenjia Jing

October 21, 2009

Abstract

We consider the theory of correctors to homogenization in stationary transport equations with rapidly oscillating, random coefficients. Let $\varepsilon \ll 1$ be the ratio of the correlation length in the random medium to the overall distance of propagation. As $\varepsilon \downarrow 0$, we show that the heterogeneous transport solution is well-approximated by a homogeneous transport solution. We then show that the rescaled corrector converges in (probability) distribution and weakly in the space and velocity variables, to a Gaussian process as an application of a central limit result. The latter result requires strong assumptions on the statistical structure of randomness and is proved only for random processes constructed by means of a Poisson point process.

1 Introduction

Partial differential equations with rapidly varying coefficients arise naturally in many important applications, such as e.g. composite material sciences, nuclear sciences, porous media equations, and Earth science as in e.g. climate modeling. It is often necessary to model such heterogeneous structures at the macroscopic level because the computational costs at the fine structure are prohibitive and because the microscopic structure is typically not well known. Macroscopic equations are derived usually for two types of heterogeneities: periodic heterogeneities and random heterogeneities. In both cases, a single parameter $\varepsilon \ll 1$, the correlation length in the heterogeneous medium, models the size of the heterogeneities related to the overall size of observation of the phenomenon. Under fairly generic ergodicity assumptions, the heterogeneous solution is shown to converge to the solution of a homogeneous (homogenized) equation; see e.g. [12, 25]. At this level, there is relatively little difference, except possibly at the mathematical level, between homogenization in periodic media and homogenization in random media.

It is often important to understand the error caused by replacing a heterogeneous solution by a homogenized approximation, for instance when such an error generates errors in the solution of an estimation (inverse) problem; see e.g. [11]. At the level of correctors, modeling the heterogeneities as random or periodic yields very different answers. Whereas correctors to homogenization are often well understood (and are typically of order $O(\varepsilon)$) in the periodic setting [12], this is not the case in the random setting, where random correctors can be arbitrarily larger than their deterministic, periodic, counterparts. In spite of its importance, the theory of random correctors to homogenization is rather poorly understood. For some of the available results in the setting of elliptic equations, we refer the reader to [4, 5, 7, 15, 23, 33].

This paper concerns the theory of correctors to the homogenization of linear transport (linear Boltzmann) equations. We consider the stationary case here although the results extend to the evolution equation as well. Homogenization theory for transport equations is well understood in fairly arbitrary ergodic random media, see e.g. [22, 26, 28]; see also e.g. [1, 13] for homogenization of transport in the periodic case. In this paper, we develop a theory for the random corrector. We first provide a bound for the corrector in energy norm. We then show that weakly in space and velocity variables, the random corrector converges in probability to a Gaussian field. This result may be seen as an application of a central limit correction as in e.g. [5, 23]. The results are shown for a specific structure of the random coefficients based on a Poisson point process. The resulting random coefficients have then short-range interactions. Whereas the results should hold for more general processes, it is clear that much more severe restrictions than mere ergodicity as in [22] must be imposed on the random structure in order to obtain a full characterization of the limiting behavior of the corrector. This is also the case for elliptic equations as may be seen in e.g. [5, 7].

The rest of the paper is structured as follows. In section 2, we present our main assumptions and the main results of the paper on the theory of the random corrector to homogenization. Section 3 recalls results on the transport equation and Poisson point processes that are useful in the derivation. The proof of the results is postponed to sections 4 for the error estimate in energy norm and 5 for the random limit of the corrector. Some technical results are postponed to the appendix.

2 Main results on the theory of random correctors

The linear transport equation finds applications in many areas of science, including neutron transport [20, 29], atmospheric science [16, 27], propagation of high frequency waves [3, 30, 31] and the propagation of photons in many medical imaging applications [2, 6]. In many settings, the coefficients in the transport equation oscillate at a very fine scale and may not be known explicitly. In such situations, it is necessary to model such coefficients as random [22, 26].

The density of particles $u_\varepsilon(x, v)$ at position x and velocity v is modeled by the following transport equation with random attenuation and scattering coefficients:

$$\begin{aligned} & v \cdot \nabla_x u_\varepsilon + a_\varepsilon(x, \omega) u_\varepsilon - \int_V k_\varepsilon(x, v', v; \omega) u_\varepsilon(x, v') dv' = 0, \quad && (x, v) \in X \times V, \\ & u_\varepsilon(x, v) = g(x, v), && (x, v) \in \Gamma_- \end{aligned} \tag{1}$$

Here X is an open, bounded, subset in $\mathbb{R}^d$ for $d = 2, 3$ spatial dimension, and V is the velocity space, which here will be $V = S^{d-1}$, the unit sphere to simplify the presentation. The sets $\Gamma_\pm$ are the sets of outgoing and incoming conditions, defined by

$$\Gamma_\pm := \{(x, v) | x \in \partial X, \ \pm\nu_x \cdot v > 0\} \tag{2}$$

where ∂X is the boundary of X, assumed to be smooth, and the normal vector to X at $x \in \partial X$ is denoted by ν_x.

The constitutive parameters in the transport equation are the total attenuation coefficient a_ε and the scattering coefficient k_ε. The above transport equation admits a unique solution in appropriate spaces [8, 19, 29] provided that these coefficients are non-negative and attenuation is larger than scattering (see below). When the coefficients are modeled as random, such constraints need to be ensured almost surely in the space of probability. We assume here that a_ε and k_ε are measurable random fields constructed on an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where a *stationary spatial Poisson point process* is constructed. We call $\{y_j^\varepsilon\}$ the points of the process with intensity $\varepsilon^{-d}\nu$. Properties of this process of importance in the paper are recalled in section 3. The intrinsic attenuation and scattering coefficients are constructed as follows:

$$\begin{aligned} a_{r\varepsilon}(x, \frac{x}{\varepsilon}, \omega) &:= a_{r0}(x) + \sum_{j\in\mathbb{N}} \psi(\frac{x - y_j^\varepsilon}{\varepsilon}), \\ k_\varepsilon(x, \frac{x}{\varepsilon}, \omega) &:= k_0(x) + \sum_{j\in\mathbb{N}} \varrho(\frac{x - y_j^\varepsilon}{\varepsilon}), \end{aligned} \tag{3}$$

where a_{r0} and k_0 are positive deterministic continuous functions and where ψ and ϱ are smooth, non-negative, compactly supported, functions in the unit ball (to simplify). The physical importance of this model is that the constitutive parameters consist of two parts: a continuous low frequency background media and random inclusions that increase attenuation and scattering.

We thus assume that scattering in (1) is isotropic, i.e., that k_ε is independent of the velocities v and v' of the particles before and after collision. Here, $a_{r\varepsilon}$ is the intrinsic attenuation, corresponding to particles that are absorbed by the medium and whose energy is transformed into heat. The total attenuation coefficient is defined as

$$a_\varepsilon(x) = a_{r\varepsilon}(x, \frac{x}{\varepsilon}) + c_d k_\varepsilon(x, \frac{x}{\varepsilon}, \omega),$$

where c_d is the volume of the unit sphere in dimension d.

Note that the above random coefficients are bounded in X $\mathbb{P}$-a.s. since the probability of infinite clustering of points in a given bounded domain is zero. However, clustering may occur so that a_ε and k_ε are not bounded uniformly in the variable ω. By construction, since $a_{r\varepsilon}$ is a positive function on X and a_ε and k_ε are positive and bounded $\mathbb{P}$-a.s., classical theories [8, 19, 29] of existence of unique solutions to (1) may be invoked $\mathbb{P}$-a.s.

Let $a := \mathbb{E}a_\varepsilon$, $k := \mathbb{E}k_\varepsilon$, and $a_r := \mathbb{E}a_{r\varepsilon}$ where $\mathbb{E}$ is the mathematical expectation associated to measure $\mathbb{P}$. Let us then define u_0 as the solution to (1) where a_ε and k_ε are replaced by their averages a and k, respectively. Then, consistently with the results shown in [22], we expect u_ε to converge to u_0. Our first result is to obtain an error estimate for the corrector $u_\varepsilon - u_0$ in the "energy" norm $L^2(\Omega, L^2(X \times V))$. More precisely, we have the following result.

Theorem 2.1. *Let dimension $d \geq 2$. Suppose that the random coefficients $a_\varepsilon, k_\varepsilon$ are constructed as in (3) and that a_{r0} is bounded from above by a positive constant. Suppose also that $g \in L^\infty(\Gamma_-)$ so that $u_0 \in L^\infty(X \times V)$. Then we have the following estimate*

$$\big(\mathbb{E}\|u_\varepsilon - u_0\|_{L^2}^2\big)^{\frac{1}{2}} \leq C\varepsilon^{\frac{1}{2}} \longrightarrow 0, \tag{4}$$

as ε goes to zero.

The above result shows that the corrector $\zeta_\varepsilon := u_\varepsilon - u_0$ may be as large as $\sqrt{\varepsilon}$. It turns out that the size of the corrector ζ_ε very much depends on the scale at which we observe it. Pointwise, ζ_ε is indeed of size $\sqrt{\varepsilon}$. However, once it is averaged over a sufficiently large domain (in space and velocities), then it may take very different values. Firstly, ζ_ε needs to be decomposed as $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$ plus $\mathbb{E}\{u_\varepsilon - u_0\}$. The latter term corresponds to deterministic correctors, which may be larger than the random corrector. The next two theorems are devoted to the limits of these correctors.

Let $\delta a_{r\varepsilon} := a_{r\varepsilon} - \mathbb{E}\{a_{r\varepsilon}\}$, and let $\delta k_\varepsilon = k_\varepsilon - k$. By construction, they are mean zero, stationary random fields. We can then define the autocorrelation function of $\delta a_{r\varepsilon}$ as

$$R_{a\varepsilon}(x) = \mathbb{E}\{\delta a_{r\varepsilon}(y)\delta a_{r\varepsilon}(y+x)\}. \tag{5}$$

By stationarity, the above right-hand side is independent of y. As we will show in Section 3, we have

$$R_{a\varepsilon}(x) = R_a(\frac{x}{\varepsilon}), \quad \text{where} \quad R_a(x) = \nu \int_{\mathbb{R}^d} \psi(x-y)\psi(y)dy = \nu\psi * \psi(x). \tag{6}$$

Similarly, we can define $R_{k\varepsilon}$ as the autocorrelation function of δk_ε, and define $R_{ak\varepsilon}$ as the cross-correlation function of the two fields, and they can be written as $R_k(\frac{x}{\varepsilon})$ and $R_{ak}(\frac{x}{\varepsilon})$ respectively where

$$R_k(x) = \nu\varrho * \varrho(x), \quad \text{and} \quad R_{ak}(x) = \nu\psi * \varrho(x). \tag{7}$$

We also denote the integration over $\mathbb{R}^d$ of the autocorrelation functions R_a and R_k by

$$\sigma_a^2 = \int_{\mathbb{R}^d} R_a(x)dx = \nu\Big(\int_{\mathbb{R}^d}\psi(x)dx\Big)^2, \ \sigma_k^2 = \int_{\mathbb{R}^d} R_k(x)dx = \nu\Big(\int_{\mathbb{R}^d}\psi(x)dx\Big)^2, \tag{8}$$

with σ_a and σ_k non-negative numbers. We then verify that the integration over $\mathbb{R}^d$ of the cross-correlation functions R_{ak} is $\sigma_a\sigma_k$. That is, the correlation of the fields is $\rho_{ak} = 1$. This is not surprising considering our construction, and (3) can be modified as in (15) below to yield $\rho_{ak} < 1$. For instance, if y_j^ε in the second line in (3) is replaced by z_j^ε, where the latter is another Poisson point process independent of y_j^ε, then we find that $\rho_{ak} = 0$. To simplify, we shall present all derivations with the model (3) knowing that all results extend to more complex models such as (15) below.

Consider a point $x \in X$, and $v \in V$ and let us denote the traveling time from x to ∂X along direction v (respectively $-v$) by $\tau_+(x,v)$ (respectively $\tau_-(x,v)$) given by

$$\tau_\pm(x,v) = \sup\{t > 0 : x \pm tv \in X\}.$$

Let x, y be two points in X, we define the amount of attenuation between x and y as

$$E(x,y) = \exp\{-\int_0^{|x-y|} a(x - s\frac{x-y}{|x-y|})ds\}.$$

We also define $E(x,y,z) = E(x,y)E(y,z)$. Recall the definition of u_0 and denote its angular integral by $\bar{u}_0$. Then we have the following limit for the deterministic corrector.

Theorem 2.2. *Let dimension $d = 2, 3$. Under the same conditions of the previous theorem, we have*

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}\{u_\varepsilon\} - u_0}{\varepsilon}(x,v) \quad = \quad U(x,v) \tag{9}$$

weakly, where $U(x,v)$ is the solution of the homogeneous (deterministic) transport equation

$$v \cdot \nabla_x U + a(x)U - k(x)\int_V U(x,v')dv' = q(x,v), \tag{10}$$

with a volume source term $q(x,v)$ given by:

$$q(x,v) = \int_{\mathbb{R}}\bigg(R_a(tv)u_0(x,v) - R_{ak}(tv)\bar{u}_0(x) - \int_V \big(R_{ak}(tw)u_0(x,w) - R_k(tw)\bar{u}_0(x)\big)dw\bigg)dt. \tag{11}$$

The above theorem presents a convergence of the corrector weakly in space. Under mild assumptions, we can show that the deterministic corrector is of order $O(\varepsilon)$ also point-wise in (x, v), and is thus independent of the scale at which it is observed. This is not the case for the random corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$. Let ε^γ be the size of the latter term. Point-wise, this term is of size $\sqrt{\varepsilon}$ with $\gamma = \frac{1}{2}$ (i.e., is a random variable of variance $O(\varepsilon)$). However, weakly in space and velocities, this term is of order $\varepsilon^{\frac{d}{2}}$ with thus $\gamma = \frac{d}{2}$ in all dimensions, hence much smaller than its point-wise value in dimension $d \geq 2$.

In this paper, we concentrate on the size of the corrector taken as a distribution in the space, velocity, and random, variables. Calculations in [8] shows the following behaviors for point-wise values of the corrector.

1. For a fixed $(x, v) \in X \times V$, the variance of the random variable $\omega \to \zeta_\varepsilon(x, v; \omega)$ is of order ε for all dimensions $d \geq 2$ so that $\gamma = \frac{1}{2}$. This property, which arises from integrating random fields along (one-dimensional) lines, is quite different from the behavior of solutions to elliptic equations considered in e.g. [5, 23].

2. For a fixed $x \in X$, let us consider the average of ζ_ε over directions and introduce the random variable $J_\varepsilon(x, \omega) := \int_V \zeta_\varepsilon(\cdot, v) dv$. The variance of $J_\varepsilon(x)$ is of order $\varepsilon^2 |\log \varepsilon|$ in dimension two (with ε^γ replaced by $\varepsilon |\log \varepsilon|^{\frac{1}{2}}$), and ε^2 in dimension $d \geq 3$ with then $\gamma = 1$. Angular averaging therefore significantly reduces the variance of the corrector.

3. Let us consider the random variable $Y_\varepsilon(\omega)$ as the average of J_ε over all positions. The variance of Y_ε is of order ε^d in dimension $d \geq 2$ with then $\gamma = \frac{d}{2}$. This is consistent with the main result described below. Weakly in space and velocities, the random corrector is of smallest size.

The scaling γ therefore depends on the scale at which the corrector is observed.

This paper considers $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$ as a random field and aims at characterizing its limit as $\varepsilon \to 0$ weakly in space and velocity. The correct scaling will be $\gamma = \frac{d}{2}$. Let us consider a collection of sufficiently smooth functions $M_l, 1 \leq l \leq L$, and we seek for the limit distribution of $\langle M_l, u_\varepsilon - \mathbb{E}\{u_\varepsilon\}\rangle$ where $\langle \cdot, \cdot \rangle$ denote the integration of a pair of Hölder conjugate functions.

Let $\tilde{M}_l$ be the solution of the following adjoint transport equation:

$$\begin{aligned} & -v \cdot \nabla_x \tilde{M}_l + a\tilde{M}_l - \int_V k(x, v, v')\tilde{M}_l(x, v')dv' = M_l, \quad && (x, v) \in X \times V, \\ & \tilde{M}_l(x, v) = 0, && (x, v) \in \Gamma_+, \end{aligned} \tag{12}$$

and define $m_l := (m_{l1}, m_{l2})'$, where

$$m_{l1} = -\int_V \tilde{M}_l(x, v)u_0(x, v)dv, \text{ and } m_{2l} = -c_d m_{l1} + \int_V u_0(x, v)dv \int_V \tilde{M}_l(x, v)dv.$$

The limiting distribution of the stochastic corrector weakly in space and velocity is shown to be Gaussian. More precisely, we have the following theorem.

Theorem 2.3. *Let dimension $d = 2, 3$. Under the same condition of Theorem 2.1, we have*

$$\langle M_l, \frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\varepsilon^{\frac{d}{2}}} \rangle \xrightarrow{\mathscr{D}} I_l := \int_X m_l(y) \cdot dW(y). \tag{13}$$

The convergence here should be interpreted as convergence in distribution of random variables. The two-dimensional multivariate Wiener process $W(y) = (W_a(y), W_k(y))'$ satisfies that

$$\mathbb{E}dW(y) \otimes dW(y) = \Sigma dy := \begin{pmatrix} \sigma_a^2 & \rho_{ak}\sigma_a\sigma_k \\ \rho_{ak}\sigma_a\sigma_k & \sigma_k^2 \end{pmatrix} dy. \tag{14}$$

The notation $\otimes$ above denotes the outer product *of vectors.*

Remark 2.4. **More general attenuation and scattering models.** In the above construction using (3), $\rho_{ak} = 1$ as we mentioned so that in fact, $W_k = \frac{\sigma_k}{\sigma_a} W_a$ in distribution. The above theorem generalizes to more complex models of attenuation and scattering. For instance, consider

$$\begin{aligned} a_{r\varepsilon}(x, \frac{x}{\varepsilon}, \omega) &:= a_{r0}(x) + \sum_{l=1}^{L} \sum_{j \in \mathbb{N}} \psi_l(\frac{x - y_j^{\varepsilon,l}}{\varepsilon}), \\ k_\varepsilon(x, \frac{x}{\varepsilon}, \omega) &:= k_0(x) + \sum_{l=1}^{L} \sum_{j \in \mathbb{N}} \varrho_l(\frac{x - y_j^{\varepsilon,l}}{\varepsilon}). \end{aligned} \tag{15}$$

Here, the profile functions ψ_l and ϱ_l for $1 \le l \le L < \infty$ are smooth compactly supported non-negative functions, and the Poisson point processes $\{y_j^{\varepsilon,l}\}_{1 \le l \le L}$ are independent possibly with different intensities $\varepsilon^{-d}\nu_l$. Physically, these Poisson point processes model different types of inclusions that may absorb and/or scatter. The matrix Σ still takes the form above while σ_a, σ_k and ρ_{ak} now take the form:

$$\sigma_a^2 = \sum_{l=1}^{L} \nu_l \left(\int_{\mathbb{R}^d} \psi_l(x) dx \right)^2, \ \sigma_k^2 = \sum_{l=1}^{L} \nu_l \left(\int_{\mathbb{R}^d} \varrho_l(x) dx \right)^2,$$

$$\rho_{ak} = (\sigma_a \sigma_k)^{-1} \sum_{l=1}^{L} \nu_l \int_{\mathbb{R}^d} \psi_l(x) dx \int_{\mathbb{R}^d} \varrho_l(x) dx.$$

To simplify the presentation, we shall only consider the model (3) of random media.

Remark 2.5. We can rewrite I_l as

$$I_l(\omega) = \int_X \sigma_l(y) dW(y) := \int_X \sqrt{m_l \otimes m_l : \Sigma} \, dW(y) \tag{16}$$

where : is the Frobenius inner product of matrices, and $W(y)$ is the standard one dimensional multivariate Wiener process. The equivalence of the two formulations is easily verified by computing their variance. The formulation in (13) displays the linear dependence of the correctors in $\delta a_r, \delta k$ at the price of introducing two correlated Wiener measures as for elliptic equations [5].

Remark 2.6. Recall the adjoint transport equation of the form (12). Let $G_*(x, v, y, v')$ be the Greens function of this equation, i.e., the solution when the source term is $\delta_y(x)\delta_{v'}(v)$, and define

$$\begin{aligned} \kappa_a(x, v, y) &:= \int_V G_*(x, v, y, v') dv' u_0(x, v) \\ \kappa_k(x, v, y) &:= c_d \kappa_a + \int_V G_*(x, v, y, v') dv' \bar{u}_0(x). \end{aligned} \tag{17}$$

The convergence in the theorem can be restated as

$$\frac{u_\varepsilon - \mathbb{E}u_\varepsilon}{\varepsilon^{\frac{d}{2}}}(x, v) \Longrightarrow \int (\kappa_a(x, v; y), \kappa_k(x, v; y)) \cdot dW(y) \tag{18}$$

where $W(y)$ is as in the theorem. This convergence is weak in space and velocity and in distribution. As we remarked earlier, the convergence does not hold point-wise in (x, v).

3 Transport equation and random structures

In the first part of this section, we recall and adapt classical results on linear transport equations. We show that the solution operator is a continuous linear transform on $L^p(X \times V)$ for all $p \in [1, \infty]$, and that the operator norm can be bounded uniformly when the coefficients are not uniformly bounded. This property is crucial to us because $\|a_\varepsilon\|_{L^\infty}, \|k_\varepsilon\|_{L^\infty}$ are not uniform in ε (and ω) due to possible clustering. We also show that sufficiently high orders of scattering admits bounded Schwartz kernel, which simplifies the analysis greatly.

In the second part of this section, we derive useful features of the processes $\delta a_\varepsilon, \delta k_\varepsilon$ introduced earlier. We show that they are stationary, ρ-mixing, and that high-order statistical moments admit explicit expressions.

3.1 Transport equations and regularity results

We observe that the corrector ζ_ε satisfies

$$v \cdot \nabla_x \zeta_\varepsilon + a\zeta_\varepsilon - k \int_V \zeta_\varepsilon(x, v')dv' = -\delta a_\varepsilon u_\varepsilon + \delta k_\varepsilon \int_V u_\varepsilon(x, v')dv', \tag{19}$$

with vanishing conditions on Γ_-. The mathematical theory for such an equation is well-established; see [19, Chap. XXI] and [13, 17, 9, 32]. To simplify our presentation, we introduce the standard notation:

$$T_0 f = v \cdot \nabla_x f, \quad A_1 f = af, \quad A_2 f = -\int_V k(x, v, v')f(x, v')dv'.$$

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

Let us also set

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

and define the following differential or integro-differential operators:

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

The fact that a function in $\mathcal{W}^p$ has trace on $\Gamma_\pm$ is proved in [17, 19]. The transport equation with volume source $f(x, v)$ and vanishing boundary condition can be written in the following compact form

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

When (a, k) is replaced by $(a_\varepsilon, k_\varepsilon)$, then the corresponding operators are denoted by $\mathbf{T}_\varepsilon$ and $\mathbf{T}_{\varepsilon 1}$.

We say the coefficients (a, k) are *admissible* if the following conditions are satisfied.

1. $a, k \geq 0$, *a.e.* and $a \in L^\infty$.
2. $k(x, v, v')$ is integrable in v' for *a.e.* (x, v) and is integrable in v for *a.e.* (x, v').

We say the problem is *subcritical* if in addition

$$a_r = a - \int_V k(x, v, v')dv' \geq \beta > 0 \text{ and } a - \int_V k(x, v', v)dv' \geq \beta > 0. \tag{20}$$

for some real number $\beta > 0$. When k is isotropic, the last condition is simply $a - c_d k \geq \beta > 0$.

For the free transport equation $\mathbf{T}_1 u = f$, an application of the method of characteristics yields the following explicit solution:

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

We easily verify the following property for $\mathbf{T}_1^{-1}$.

Proposition 3.1. *Let the transport equation be subcritical with parameter β. Then the solution operator $\mathbf{T}_1^{-1}$ is a bounded linear transform on $L^p(X \times V)$ for all $p \in [1, \infty]$. Moreover,*

$$\|\mathbf{T}_1^{-1} f\|_{L^p} \leq e^{-\beta\delta} \delta \|f\|_{L^p},$$

where δ is the diameter of the domain X, i.e, $\delta := \sup_{x,y \in X} |x-y|$.

The full transport equation may be seen as a perturbation of free transport. It is proved in [19] using semi-group techniques that $\mathbf{T}^{-1}$ is a bounded linear transform on L^p for all $1 \leq p \leq \infty$. We will need a more accurate descriptions of the transport solution operator written in integral form and thus introduce the operators K and $\mathcal{K}$ as:

$$Ku := \mathbf{T}_1^{-1} A_2 u = -\int_0^{\tau_-(x,v)} E(x, x-tv) \int_V k(x-tv, v, v') u(x-tv, v') dv'.$$

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

with $v' = (x-y)/|x-y|$ above. Their importance may be understood from the relations

$$\mathbf{T}^{-1} = (I+K)^{-1} \mathbf{T}_1^{-1}, \text{ and } \mathbf{T}^{-1} = \mathbf{T}_1^{-1} (I+\mathcal{K})^{-1}.$$

The two equalities are obtained in the L^∞ and L^1 settings by standard algebraic manipulations. In [9], it is shown using this approach that the operator norm of $\mathcal{K}$ can be bounded by $1 - e^{-\beta\delta}$. Hence by Proposition 3.1, $\mathbf{T}^{-1}$ is bounded by δ in the L^1 setting while similar techniques apply in the L^∞ setting with K. Hence, we have the following result as an application of the Riesz-Thorin interpolation.

Proposition 3.2. *Let the transport equation be subcritical with parameter β. Then, the solution operator $\mathbf{T}^{-1}$ is a bounded linear transform on $L^p(X \times V)$ for all $p \in [1, \infty]$ and*

$$\|\mathbf{T}^{-1} f\|_{L^p} \leq \delta \|f\|_{L^p}.$$

Remark 3.3. The above bound does not depend on the value of $\|a\|_{L^\infty}, \|k\|_{L^\infty}$. This ensures that $\|\mathbf{T}_\varepsilon^{-1}\|_{L^p \to L^p}$ is uniformly bounded as long as the subcritical condition is satisfied with a uniform bound β even when $\|a_\varepsilon\|_{L^\infty}, \|k_\varepsilon\|_{L^\infty}$ are not uniformly controlled.

We verify that $\mathbf{T}^{-1}$ admits the following Neumann series expansion:

$$\mathbf{T}^{-1} = \mathbf{T}_1^{-1} - \mathbf{T}_1^{-1} \mathcal{K} + \mathbf{T}_1^{-1} \mathcal{K}^2 + \cdots.$$

The analysis of $\mathbf{T}^{-1}$ can be done term by term. However, to avoid dealing with an infinite series, we group the scattering contributions of sufficient large order together so as to analyze a finite sum of operators. This is done by showing that $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$ has a bounded kernel so that multiple scatterings of order at least $d+1$ can be grouped together. To show this, we observe that when acting on functions of the spatial variable only, $\mathcal{K}: L^1(X) \to L^1(X)$ has the following expression:

$$\mathcal{K}f(x) = -\int_X \frac{k(x)E(x,y)}{|x-y|^{d-1}} f(y)dy.$$

Then we have the following estimate on the kernel of $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$.

Lemma 3.4. *Let the coefficients (a,k) be subcritical, then the operator $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$ admits a Schwartz kernel that is a bounded function. That is to say,*

$$(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}f(x) = \int_X \alpha(x,y)f(y)dy \tag{21}$$

and $\|\alpha(x,y)\|_{L^\infty(X\times X)} < \infty$.

PROOF: The kernel of $\mathcal{K}$ is a function given by $\kappa(x,y) = -\frac{k(x)E(x,y)}{|x-y|^{d-1}}$, so that the kernel of $\mathcal{K}^{d+1}$ is:

$$\kappa_{d+1}(x,y) = (-1)^{d+1}\int_{X^d} \frac{k(x)lE(x,z_1)k(z_1)E(z_1,z_2)\cdots k(z_d)E(z_d,y)}{|x-z_1|^{d-1}|z_1-z_2|^{d-1}\cdots|z_d-y|^{d-1}} dz_1\cdots dz_d.$$

Thanks to the convolution lemma A.1, we see that this kernel is bounded.

For the kernel of $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$, we first write it as $\mathcal{K}^{d+1}(I+\mathcal{K})^{-1}$ and denote it by $\mathcal{L}$. Then for any $\phi \in L^1(X)$ and $\psi \in L^1(X)$, we verify the following.

$$\langle \mathcal{L}\phi,\psi\rangle_X = \int_{X^2} \kappa_{d+1}(x,y)\big((I+\mathcal{K})^{-1}\phi\big)(y)\psi(x)dxdy \le \|\kappa_{d+1}\|_{L^\infty}\|(I+\mathcal{K})^{-1}\|_{L^1\to L^1}\|\phi\|_{L^1}\|\psi\|_{L^1}.$$

The last inequality holds because the integration in x and y are separated. Therefore, we have shown that

$$\|\mathcal{L}\phi\|_{L^\infty} \le C_0\|\phi\|_{L^1}, \tag{22}$$

where C_0 is the constant appearing in the preceding inequality. The Schwartz kernel theorem [24] shows the existence of a distribution α on $X \times X$ such that $\langle \mathcal{L}\phi,\psi\rangle_X = \langle \alpha, \phi\otimes\psi\rangle_{X\times X}$ for smooth functions ϕ and ψ in $\mathcal{D}(X)$. Moreover, thanks to (22), we obtain that α is a function in the x variable and that

$$\|\sup_{\|\phi\|_{L^1(X)}=1} \langle \alpha(x,\cdot),\phi\rangle_X\|_{L^\infty_x(X)} \le C_0.$$

Thus for a.e. $x \in X$, $\sup_{\|\phi\|_{L^1(X)}=1}\langle \alpha(x,\cdot),\phi\rangle_X \le C_0$ so that for a.e. $x\in X$, the linear form $\langle\alpha(x,\cdot),\phi\rangle$ satisfies

$$\langle \alpha(x,\cdot),\phi\rangle_X = \int_X \alpha(x,y)\phi(y)dy,$$

with $\alpha(x,y) \in L^\infty_y(X)$ by the Riesz representation stating that $(L^1(X))' = L^\infty(X)$. This shows that $\mathcal{L}$ may be represented as in (21) with $\alpha \in L^\infty(X\times X)$. $\Box$

An immediate corollary is the following.

Corollary 3.5. *Under the same condition as above, we have the following decomposition.*

$$\mathbf{T}^{-1}f = \mathbf{T}_1^{-1}(f - \mathcal{K}f + \tilde{\mathcal{K}}\mathcal{K}f) \tag{23}$$

where $\tilde{\mathcal{K}}$ is a weakly singular integral operator with a kernel bounded by $C|x-y|^{-d+1}$, $d=2,3$.

PROOF: It remains to rewrite the Neumann series as

$$\begin{aligned}\mathbf{T}^{-1}f &= \mathbf{T}_1^{-1}f - \mathbf{T}_1^{-1}\mathcal{K}f + \mathbf{T}_1^{-1}\mathcal{K}^2 f + \cdots + (-1)^d \mathbf{T}_1^{-1}\mathcal{L}\mathcal{K}f \\ &= \mathbf{T}_1^{-1}(f - \mathcal{K}f + (\sum_{j=1}^{d}(-1)^{j-1}\mathcal{K}^j + (-1)^d\mathcal{L})\mathcal{K}f).\end{aligned} \tag{24}$$

The lemma shows that $\mathcal{L}$ admits a bounded kernel. We are left with a finite number of operators $\mathcal{K}^j, j = 1, 2, \cdots, d$ to consider. The Schwartz kernels of these operators are explicit and can be estimated using the convolution lemma A.1. They are all bounded by $C|x-y|^{1-d}$, and so is the kernel of $\tilde{\mathcal{K}} := \sum_{j=1}^{d}(-1)^{j-1}\mathcal{K}^j + (-1)^d\mathcal{L}$. $\square$

The theories developed for the forward transport equation apply with no modification to the adjoint transport equation, which we have used in the definition of $\tilde{M}_l$ in (12). We denote

$$\mathbf{T^*}_1 u = -T_0 u + A_1 u, \;\; \mathbf{T}^* u = \mathbf{T^*}_1 u - A_2' u, \;\; D(\mathbf{T^*}_1) = D(\mathbf{T}^*) = \{u \in \mathcal{W}^p, u|_{\Gamma_+} = 0\}. \tag{25}$$

Here A_2' is of the same form as A_2 with v and v' switched in the function k. In our case, since k is assumed to be isotropic, A_2' and A_2 are the same. We obtain similar expressions for $\mathbf{T^*}_1^{-1}$, K^* and $\mathcal{K}^*$. Under the same subcritical condition as before, we also have that $\mathbf{T^*}_1^{-1}$ and $\mathbf{T}^{*-1}$ are bounded linear transforms on $L^p(X \times V)$ for all $p \in [1, \infty]$. Also, for any pair of functions that are Hölder continuous, we find that:

$$\langle u, \mathbf{T}^{-1}w\rangle = \langle \mathbf{T}^{*-1}u, w\rangle. \tag{26}$$

3.2 Random inclusions of attenuation and scattering kernel

In this section, we exhibit some useful properties of the random model (3). In the model, the deterministic functions a_{r0} and k_0 are continuous and slowly varying, and satisfies the subcriticality condition. The random part can be viewed as additional heterogeneous inclusions of attenuation and scattering kernels. The centers of these inclusions are y_j^ε and they are distributed as a spatial Poisson point process. The profiles of the inclusions are ψ and ϱ, which are nonnegative smooth functions compactly supported on the unit ball. Similar models have been considered [23, 10].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space. A *stationary spatial Poisson point process* with *intensity* ν is a countable subset $Y_\nu := \{y_j\} \subset \mathbb{R}^d$ such that for any set A in the Borel algebra $\mathscr{B}(\mathbb{R}^d)$, the random variable $N(A)$ defined as the cardinality of $A \cap Y_\nu$, satisfies Poisson distribution with intensity $\nu|A|$, i.e.,

$$\mathbb{P}\{N(A) \le m\} = \frac{e^{-\nu|A|}(\nu|A|)^m}{m!}, \tag{27}$$

where $|A|$ is the Lebesgue measure of A. Further, for any positive integer $n \ge 2$ and mutually disjoint Borel sets $A_1, \cdots, A_n$, the random variables $N(A_i)$, $1 \le i \le n$ are independent. The map $A \mapsto N(A)$ is a counting measure on $\mathscr{B}(\mathbb{R}^d)$ and is called the *Poisson counting measure*.

Stationary property. The stationarity of such a Poisson point process is due to the fact that the distribution of $N(A)$ depends only on $|A|$ but not on the position of A. Therefore, $N(A_1 + z), \cdots N(A_n + z)$ have the same distribution with $N(A_1), \cdots, N(A_n)$ for any $z \in \mathbb{R}^d$ and $A_i \in \mathscr{B}(\mathbb{R}^d)$. By construction, the coefficients $a_{r\varepsilon} - a_{r0}$ and $k_\varepsilon - k_0$ are also stationary.

Scaling property. Let $Y_\nu = \{y_j\}$ be a Poisson point process with intensity ν, and $Y_{\varepsilon^{-d}\nu} = \{y_j^\varepsilon\}$ be the one defined before with intensity $\varepsilon^{-d}\nu$. Then $\varepsilon^{-1}Y_{\varepsilon^{-d}\nu}$ has the same distribution as Y_ν.

This is easily verified by:

$$\mathbb{P}\{\sum_j \chi_{\varepsilon A}(y_j^\varepsilon) \le m\} = \mathbb{P}\{N(\varepsilon A; Y_{\nu\varepsilon^{-d}})\} = \frac{e^{-\nu\varepsilon^{-d}|\varepsilon A|}(\nu\varepsilon^{-d}|\varepsilon A|)^m}{m!},$$

and that $\varepsilon^{-d}|\varepsilon A| = |A|$. As a result, if we define

$$a_r{}^\nu(x;\omega) = \sum_j \psi(x - y_j(\omega)),\ k^\nu(x;\omega) = \sum_j \varrho(x - y_j(\omega)), \tag{28}$$

then $a_r{}^\nu(\frac{\cdot}{\varepsilon})$ and $k^\nu(\frac{\cdot}{\varepsilon})$ have the same distribution with $a_{r\varepsilon} - a_{r0}$ and $k_\varepsilon - k_0$ respectively. Therefore, to derive moments formula for the latter, it suffices to derive them for the former and evaluate the results at position $\frac{\cdot}{\varepsilon}$.

Mixing property. Recall that a random field $q(x,\omega)$ is said to be strong *mixing* if for any two Borel sets $A, B \subset \mathbb{R}^d$, the σ-algebra $\mathcal{F}_A$ and $\mathcal{F}_B$ generated by $q|_A$ and $q|_B$ respectively decorrelate fast enough as the distance of the two sets increases. We refer the readers to [21] for more accurate definition and quantitative characterization of varieties of mixing properties. For Poisson point process, as long as A and B are disjoint, two random variables which depend only on points in A and B respectively are independent. By construction, as long as A and B are separated more than twice the support of the profile functions, the processes e.g., δa_r, restricted on A and B are independent. Hence, the random model (3) is strong mixing, actually they are m-independent; see [21].

3.2.1 Moments formulas for the random fields

Mean and autocorrelation functions. Since $a_r{}^\nu$ is stationary, its mean is a constant. It can be calculated conditioning on $N(B_1(x))$ as follows.

$$\mathbb{E}a_r{}^\nu = \mathbb{E}\sum_j \psi(x - y_j) = \sum_{m=1}^\infty \mathbb{P}\{N(B_1(x) = m\}\mathbb{E}(\sum_{j=1}^m \psi(x - y_j')|N(B_1(x)) = m).$$

Here, we denote the m points that land in $B_1(x)$ as y_j'. Recall that conditioned on $N(A) = m$, the m points y_j' are independent, identically and uniformly distributed on A, see [18, pp. 47]. We have

$$\mathbb{E}a_r{}^\nu = \sum_{m=1}^\infty e^{-\nu\alpha_d}\frac{(\nu\alpha_d)^m}{m!} m \int_{B_1(x)} \psi(x - z)\frac{dz}{\alpha_d} = \nu\int_{\mathbb{R}^d}\psi(z)dz =: \nu\hat{\psi}(0).$$

Here, α_d denotes the volume of the unit ball in $\mathbb{R}^d$. Then it follows that $\mathbb{E}a_{r\varepsilon}(x) = a_{r0}(x) + \nu\hat{\psi}(0)$. Recall the definition of $\delta a_{r\varepsilon}$, we see it has the same distribution as $\delta a_r{}^\nu(\frac{x}{\varepsilon}) := a_r{}^\nu(\frac{x}{\varepsilon}) - \nu\hat{\psi}(0)$ and is stationary and mean-zero. Similarly, $\mathbb{E}k_\varepsilon = k_0 + \nu\hat{\psi}$, and δk_ε can be identified with $\delta k^\nu(\frac{x}{\varepsilon}) := k^\nu(\frac{x}{\varepsilon}) - \nu\hat{k}(0)$, and δa_ε with the combination of the two.

Now we calculate the autocorrelation function of $\delta a_r{}^\nu$ and then evaluate the result at $\frac{\cdot}{\varepsilon}$ to get the autocorrelation function of $\delta a_{r\varepsilon}$. For simplicity, we will drop the superscript ν in the notation. We have,

$$\mathbb{E}(\delta a_r(x_1)\delta a_r(x_2)) = \mathbb{E}\prod_{i=1}^2\sum_{j\ge1}\psi(x_i - y_j) - (\nu\hat{\psi}(0))^2.$$

Since ψ is compactly supported on the unit ball, only those y_j's that are in the set $A = \cup B(x_i)$ contribute to the product, and A is a bounded set. Again, we calculate the expectation conditioning on $N(A)$. The object is now:

$$\begin{aligned}
&\sum_{m=1}^{\infty} e^{-\nu|A|}\frac{(\nu|A|)^m}{m!}\mathbb{E}(\sum_{j=1}^{m}\psi(x_1-y_j')\psi(x_2-y_j') + \sum_{i,j=1,i\neq j}^{m}\psi(x_1-y_i')\psi(x_2-y_j')|N(A)=m)\\
=&\nu\int_{B(x_1)\cap B(x_2)}\psi(x_1-z)\psi(x_2-z)dz + (\nu\hat{\psi}(0))^2
\end{aligned}$$

where we have used again the property that conditioned on $N(A) = m$, the m points are independent and uniformly distributed in A. Hence, we also have

$$R_a(x-y) = \mathbb{E}(\delta a_r(x)\delta a_r(y)) = \nu\psi * \psi(x-y).$$

We verify that this is a function of the variable $x-y$ and it is compactly supported in this variable. Now, the autocorrelation function of $\delta a_{r\varepsilon}$ is just $R_a(\frac{\cdot}{\varepsilon})$. Similarly we can derive formulas for R_k and R_{ak} and verify that they are as given in Section 2.

Higher order moments of $\delta a_{r\varepsilon}$ and δk_ε. Our proof of the main results depends on the fact that we have explicit expressions for moments (up to the eighth order for $d = 2, 3$) of the random fields. A systematic formula for higher order moments of δa_r is derived in Appendix B. We cite the results here.

We denote the set $\{1, 2, \cdots, n\}$ by $[n]$. An array $(n_1, n_2, \cdots, n_k)$ satisfying that $\sum_{i=1}^k n_i = n$ and $1 \le n_1 \le n_2 \le \cdots \le n_k$ is called a *partition of* n. If such a partition satisfies further that $n_1 \ge 2$, then we say it is *non-single*. Let $\mathcal{P}_n$ be the set of all partitions of n, and let $\mathscr{G}_n$ be the set of non-single partitions of n. Then for δk (and similarly for other coefficients), we have

$$\mathbb{E}\{\prod_{i=1}^{n}\delta k(x_i)\} = \sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^{k}T^{n_j}(x_1^{(\ell,n_j)},\cdots,x_{n_j}^{(\ell,n_j)}). \tag{29}$$

For a given $(n_1, \cdots, n_k) \in \mathscr{G}_n$, the index ℓ represents the choice of dividing $[n]$ (hence $\{x_i\}, 1 \le i \le n$) into disjoint groups of size n_i. The label ℓ, n_j on x represents the particular choice of partition. The functions $T_\varepsilon^k(\cdot)$ are defined as:

$$T^{n_j}(x_1,\cdots,x_{n_j}) := \nu\int\prod_{i=1}^{n_j}\psi(x_i-z)dz. \tag{30}$$

Then for δk_ε, we just need to evaluate the above formula at $\frac{\cdot}{\varepsilon}$.

In particular, the fourth order moments of the random model, say δk_ε, is given by

$$\begin{aligned}
\mathbb{E}\{\prod_{i=1}^{4}\delta k_\varepsilon(x_i)\} &= \nu\int\prod_{i=1}^{4}\varrho(\frac{x_i}{\varepsilon}-z)dz\\
&+ R_\varepsilon(x_1-x_2)R_\varepsilon(x_3-x_4) + R_\varepsilon(x_2-x_3)R_\varepsilon(x_1-x_4) + R_\varepsilon(x_1-x_3)R_\varepsilon(x_2-x_4).
\end{aligned} \tag{31}$$

Here and below, we will use the notation that $T_\varepsilon^{n_j} = T^{n_j}(\frac{\cdot}{\varepsilon})$ and $R_\varepsilon = R(\frac{\cdot}{\varepsilon})$.

3.2.2 L^p boundedness of the random fields

From the construction of $a_{r\varepsilon}, k_\varepsilon$, we see that they are not uniformly bounded due to the possible (though rare) clustering of y_j's in a small set. Nevertheless, when the L^p norm is considered, the random fields are bounded uniformly in ε. In fact, we have

Lemma 3.6. *The random fields such defined are in $L^n(\Omega, L^n(X))$ for $n \geq 1$:*

$$\mathbb{E}\|a_{r\varepsilon}\|_{L^n}^n + \mathbb{E}\|k_\varepsilon\|_{L^n}^n \leq C(n)$$

where $C(n)$ does not depend on ε.

PROOF: Since the result for n odd follows from the result for $n+1$, which is even, we set $n = 2m$ and have

$$\mathbb{E}\{\|\delta k_\varepsilon\|_{L^{2m}}^{2m}\} = \int_X \mathbb{E}(\delta k_\varepsilon(x))^{2m} dx.$$

We use the formula for high order moments, and since in our case all the $2m$ variables are the same, we need to evaluate the terms T^j in (29) at 0. Since we assumed that the function ϱ is C_c^∞, all the integrals are finite. Therefore, we obtain a bound independent of ε. Control of the attenuation coefficient is obtained in the same way. $\square$

In the next two sections, we prove the main results with the random field model (3). However, the same procedure of proof applies to more general random models. The main required features of the process are: (i) $a_{r\varepsilon}$ and k_ε are nonnegative, stationary, and $\mathbb{P}$-a.s bounded; (ii) The mean-zero process $a_{r\varepsilon} - \mathbb{E}\{a_{r\varepsilon}\}$ and $k_\varepsilon - \mathbb{E}\{k_\varepsilon\}$ have the same distribution as $\delta a_r(\frac{\cdot}{\varepsilon})$ and $\delta k(\frac{\cdot}{\varepsilon})$ respectively for some stationary random fields δa_r and δk; (iii) The random fields δa_r and δk have correlation functions $\{R_a, R_{ak}, R_k\}$ that are integrable in all directions and over the whole domain; (iv) The random fields δa_r and δk admit explicit expressions for their moments up to the eighth order (assuming $d = 2, 3$); see the proofs below for a more quantitative statement.

4 Proof of Theorem 2.1 and homogenization theory

In this section, we prove Theorem 2.1, which states that the solutions to the random equations converge in energy norm to the solution of the homogenized equation. We show that the corrector can be decomposed into two parts. The leading part satisfies a homogeneous transport equation with a random volume source, and the other part is much smaller. This theorem works for all dimension $d \geq 2$.

We can now view (19) as $\mathbf{T}\zeta_\varepsilon = A_\varepsilon u_\varepsilon$ where A_ε is an operator defined by $A_\varepsilon f = -\delta a_\varepsilon f + \delta k_\varepsilon \bar{f}$. Let $\chi_\varepsilon = \mathbf{T}^{-1} A_\varepsilon u_0$ and we verify that

$$\zeta_\varepsilon = \chi_\varepsilon + z_\varepsilon,$$

where $\mathbf{T}_\varepsilon z_\varepsilon = A_\varepsilon \chi_\varepsilon$. Hence, we introduce the following key lemmas on solutions of transport equations with interior source of the form $A_\varepsilon q$, and $A_\varepsilon \chi_\varepsilon$ and vanishing boundary conditions.

Lemma 4.1. *Assume $d \geq 3$. Let $q(x, v) \in L^\infty(X \times V)$ and define*

$$\chi_\varepsilon(x, v) = \int_0^{\tau_-(x,v)} E(x, x - tv)(-\delta a_\varepsilon(x - tv) q(x - tv, v) + \delta k_\varepsilon(x - tv)\bar{q}(x - tv))dt,$$

the solution to $\mathbf{T}_1\chi_\varepsilon = A_\varepsilon q$. Then for any integer $n \geq 1$, we have

$$\mathbb{E}\|\chi_\varepsilon\|^n_{L^n} \leq C_n\varepsilon^{\frac{n}{2}}\|q\|^n_{L^\infty}, \qquad \mathbb{E}\|\bar{\chi}_\varepsilon\|^n_{L^n} \leq C_n\varepsilon^n\|q\|^n_{L^\infty}. \tag{32}$$

Further, solving the equation $\mathbf{T}_1 u = \delta a_\varepsilon\chi_\varepsilon$ yields

$$\mathbb{E}\|\mathbf{T}_1^{-1}\delta a_\varepsilon\chi_\varepsilon\|^n_{L^n} \leq C_n\varepsilon^n\|q\|^n_{L^\infty}. \tag{33}$$

When $d = 2$, the term ε^n in the second and third estimates should be replaced by $\varepsilon^n|\log\varepsilon|^{\frac{n}{2}}$.

Proof: Since the domain $X \times V$ is bounded, we only need to consider the case n even.

1. *Control of χ_ε without averaging.* We can rewrite χ_ε as a sum of integrals of a_ε and k_ε. Using Minkowski's inequality, it is sufficient to control them separately and the proof for both terms is handled similarly. We consider

$$I_1 = \int_{X\times V}\left(\int_0^{\tau_-(x,v)} E(x, x-tv)\delta a_\varepsilon(x-tv)q(x-tv,v)\right)^n dxdv.$$

Taking expectation, we have

$$\mathbb{E}I_1 = \int_{X\times V}\int_0^{\tau_-} dt_1 \cdots \int_0^{\tau_-} dt_n \left(\prod_{i=1}^n E(x, x-t_iv)q(x-t_iv,v)\right)\mathbb{E}\prod_{i=1}^n \delta a_\varepsilon(x-t_iv)dxdv$$

where $\tau_- \equiv \tau_-(x,v)$. Recall that we have an expression for the $n-$th order moments:

$$\mathbb{E}\prod_{i=1}^n \delta a_\varepsilon(x-t_iv) = \sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^k T^{n_j}(\frac{t_2^{\ell,n_j}-t_1^{\ell,n_j}}{\varepsilon}v,\cdots,\frac{t_{n_j}^{\ell,n_j}-t_1^{\ell,n_j}}{\varepsilon}v)$$

This expression is a sum of integrable functions. Hence, for each (ℓ, n_j), we change variable $(t_i^{\ell,n_j} - t_1^{\ell,n_j})/\varepsilon \to t_i^{\ell,n_j}$, and assume that u_0 is uniformly bounded. Then we have

$$\mathbb{E}I_1 \leq C\int_{X\times V}\sum_{(n_1,\cdots,n_k)\in\mathscr{G}}\sum_{\ell=1}^{C^n_{n_1,\cdots,n_k}}\prod_{j=1}^k \varepsilon^{n_j-1}\int_0^{\tau_-} dt_1^{\ell,n_j}\int_{\mathbb{R}^{n_j-1}} T^{n_j}(t_2^{\ell,n_j}v,\cdots,t_{n_j}^{\ell,n_j}v).$$

Since T^{n_j} is integrable in all directions, we see that all the integrals above are finite and hence we find that

$$\mathbb{E}I_1 \leq C_n\|q\|^n_{L^\infty}\varepsilon^{\min_k(n-k)}.$$

From the definition of non-single partition of n, we know $k \leq \frac{n}{2}$ to make sure that $n_j \geq 2, j = 1,\cdots,k$. Hence $\min_k(n-k) = \frac{n}{2}$. This yields the first estimate. We mention that C_n depends on n through p_n' and on the size of X.

2. *Control of the average of χ_ε.* Again, we consider the a_ε term only. Recall the change of variable

$$\int_V\int_0^{\tau_-(x,v)} f(x-tv,v)dtdv = \int_X \frac{f(y,v)}{|x-y|^{d-1}}|_{v=\frac{x-y}{|x-y|}}dy. \tag{34}$$

We rewrite the term as

$$\int_V\int_0^{\tau_-} E(x,x-tv)(-\delta a_\varepsilon(x-tv))q(x-tv,v)dtdv = -\int_X\frac{E(x,y)\delta a_\varepsilon(y)}{|x-y|^{d-1}}q(y,\frac{x-y}{|x-y|})dy.$$

The term we wish to analyze is now

$$
\begin{aligned}
I_2 &= \int_X \left(\int_X \frac{E(x,y)\delta a_\varepsilon(y)}{|x-y|^{d-1}} q(y, \frac{x-y}{|x-y|}) dy \right)^n dx \\
&= \int_X dx \int_{X^n} \left(\prod_{i=1}^n \frac{E(x,y_i)}{|x-y_i|^{d-1}} q(y_i, \frac{x-y_i}{|x-y_i|}) \right) \prod_{i=1}^n \delta a_\varepsilon(y_i) d[y_1 \cdots y_n].
\end{aligned}
$$

Here and in the sequel, $d[y_1 \cdots y_n] \equiv dy_1 \cdots dy_n$. Upon taking expectation, we have

$$
\mathbb{E} I_2 \leq C \|q\|_{L^\infty}^n \int_X \int_{X^n} \prod_{i=1}^n \frac{1}{|x-y_i|^{d-1}} \mathbb{E} \prod_{i=1}^n \delta a_\varepsilon(y_i)
$$

Now we use the formula for high-order moments again and obtain

$$
\mathbb{E} I_2 \leq C \sum_{(n_1,\cdots,n_k)\in\mathscr{G}} \sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}} \int_X \prod_{j=1}^k \int_{X^{n_j}} \frac{T^{n_j}(\frac{y_2^{\ell,n_j}-y_1^{\ell,n_j}}{\varepsilon}, \cdots, \frac{y_{n_j}^{\ell,n_j}-y_1^{\ell,n_j}}{\varepsilon})}{|x-y_1^{\ell,n_j}|^{d-1} \cdots |x-y_{n_j}^{\ell,n_j}|^{d-1}}.
$$

There are many terms to estimate, which are all analyzed in the same manner. We consider a term with fixed k, n_j and ℓ. For each T^{n_j}, it is a function of $y_i^{\ell,n_j} - y_1^{\ell,n_j}$. Hence, we change variables

$$
\frac{y_i^{\ell,n_j} - y_1^{\ell,n_j}}{\varepsilon} \to y_i^{\ell,n_j},\ x - y_1^{\ell,n_j} \to y_1^{\ell,n_j},\ i = 1, \cdots, n_j, j = 1, \cdots, k.
$$

Then the integral of this term becomes

$$
\int_X dx \prod_{j=1}^k \varepsilon^{d(n_j-1)} \int_{x-X} dy_1^{\ell,n_j} \int_{(X/\varepsilon)^{n_j-1}} \frac{T^{n_j}(y_2^{\ell,n_j}, \cdots, y_{n_j}^{\ell,n_j})}{|y_1^{\ell,n_j}|^{d-1} \prod_{i=2}^{n_j} |\varepsilon y_i^{\ell,n_j} - y_1^{\ell,n_j}|^{d-1}} d[y_2^{\ell,n_j} \cdots y_{n_j}^{\ell,n_j}].
$$

We will call this term I_{n_j}. We need to control the k integrals inside the product sign, which are controlled in the same manner. We use the so-called *Voronoi diagram* of points. Let us consider one of integral with (ℓ, n_j) fixed and denote y_i^{ℓ,n_j} as y_i' to simplify the notation. Now we only need to control

$$
\int_{2X} dy_1' \int_{\mathbb{R}^{d(n_j-1)}} \frac{T^{n_j}(y_2', \cdots, y_{n_j}')}{|y_1'|^{d-1} \prod_{i=2}^{n_j} |y_1' - \varepsilon y_i'|^{d-1}}.
$$

For almost all $y_2', \cdots, y_{n_j}'$, we can consider the Voronoi diagram formed by $\varepsilon y_2', \cdots, \varepsilon y_{n_j}'$. For any fixed i, when y_1' is inside the cell of $\varepsilon y_i'$,

$$
|y_1' - \varepsilon y_l'| \geq \frac{1}{2} |\varepsilon y_i' - \varepsilon y_l'|,\ \forall l \neq i.
$$

Then if we replace $y_1' - \varepsilon y_l'$ by $\varepsilon(y_i' - y_l')/2$, the integral increases. Hence we have:

$$
I_{n_j} \leq \varepsilon^{d(n_j-1)} \sum_{i=1}^{n_j} \int_{2X} \frac{1}{|y_1'|^{d-1} |y_1' - \varepsilon y_i'|^{d-1}} dy_1' \int_{\mathbb{R}^{(n_j-1)d}} \frac{T^{n_j}(y_2', \cdots, y_{n_j}')}{\prod_{l\neq 1,i} (2^{-1}\varepsilon)^{d-1} |y_i' - y_l'|^{d-1}} d[y_2' \cdots y_{n_j}']. \tag{35}
$$

When $d = 3$, after integrating in y_1', we thus obtain the bound

$$
C\varepsilon^{(n_j-1)d-(n_j-2)(d-1)-(d-2)} \int_{\mathbb{R}^{(n_j-1)d}} \frac{T^{n_j}(y_2', \cdots, y_{n_j}')}{|y_i'|^{d-2} \prod_{l\neq 1,i} |y_i' - y_l'|^{d-2}} d[y_2' \cdots y_{n_j}'], \tag{36}
$$

Here, we used Lemma A.1 for the integral over y_1'. Hence this term is of order ε^{n_j} once we obtain that the integral is bounded. Recalling the definition of T^{n_j}, we can rewrite this integral as

$$\begin{aligned}&\int_{\mathbb{R}^{n_j d}} \psi(z)\frac{\psi(z-y_i')}{|y_i'|^{d-2}}\prod_{l\neq 1,i}^{n_j}\frac{\psi(z-y_l')}{|y_i'-y_l'|^{d-1}}dzd[y_2'\cdots y_{n_j}'] \\ &= \int_{\mathbb{R}^d} dz\psi(z)\int_{\mathbb{R}^d} dy_i'\frac{\psi(z-y_i')}{|y_i'|^{d-2}}\prod_{l\neq 1,i}\int_{\mathbb{R}^d}\frac{\psi(z-y_i'-y_l')}{|y_l'|^{d-1}}dy_l'.\end{aligned} \tag{37}$$

The integrals inside the product sign are bounded uniformly in $z-y_i'$ since

$$\int_{|y_l'|\leq 1}\frac{\psi(z-y_i'-y_l')}{|y_l'|^{d-1}}dy_l' + \int_{|y_l'|>1}\frac{\psi(z-y_i'-y_l')}{|y_l'|^{d-1}}dy_l' \leq \|\psi\|_{L^\infty}c_d + \|\psi\|_{L^1}.$$

Thus we need to estimate

$$\int_{\mathbb{R}^{2d}}\frac{\psi(z)\psi(z-y_i')}{|y_i'|^{d-1}}dy_i'dz = \int_{\mathbb{R}^d}\psi(z)\left(\psi * \frac{1}{|\cdot|^{d-1}}\right)(z)dz.$$

This integral is clearly bounded since $\psi * |\cdot|^{-d+1}$ is bounded and ψ is compactly supported.

Hence each I_{n_j} is of order ε^{n_j} and therefore I_2 is of order ε^n. In the case when $n=2$, by Lemma (A.1), the integral over y_1' above should be replaced by a logarithm function, and each I_{n_j} has a contribution of $\varepsilon^{n_j}|\log\varepsilon|$; therefore, I_2 is of order $\varepsilon^n|\log\varepsilon|^{\max k}$. Again, $k\geq\frac{n}{2}$ for all the non-single partitions. Hence, I_2 is of order $\varepsilon^n|\log\varepsilon|^{\frac{n}{2}}$.

3. *Proof of the third estimates.* The third estimate is a consequence of the first two. First we can write $\mathbf{T}_1^{-1}A_\varepsilon\chi_\varepsilon$ as

$$\mathbf{T}_1^{-1}\delta a_\varepsilon\mathbf{T}_1^{-1}\delta a_\varepsilon q - \mathbf{T}_1^{-1}\delta a_\varepsilon\mathbf{T}_1^{-1}\delta k_\varepsilon\bar{q} + \mathbf{T}_1\delta k_\varepsilon(\bar{\chi}_\varepsilon).$$

The first two terms are analyzed as in 1. While considering the L^n norm of this term, we have $2n$ terms of $\delta a_\varepsilon, \delta k_\varepsilon$, which all yield contributions of order ε^n. For the third term, we use the inequality that

$$\mathbb{E}\|\mathbf{T}_1^{-1}\delta k_\varepsilon\bar{\chi}_\varepsilon\|_{L^n}^n \leq C[\mathbb{E}\|\delta k_\varepsilon\|_{L^{2n}}^{2n}]^{\frac{1}{2}}[\mathbb{E}\|\bar{\chi}_\varepsilon\|_{L^{2n}}^{2n}]^{\frac{1}{2}}$$

and the fact that $\mathbb{E}\|k_\varepsilon\|_{L^{2n}}^{2n}$ is bounded. Application of the second estimate completes the proof. □

We can generalize these estimates to the case when $\mathbf{T}_1$ is replaced by $\mathbf{T}$ above. For $\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon q$, we have:

Corollary 4.2. *Under the same condition, replace $\mathbf{T}_1$ in previous lemma by $\mathbf{T}$. Then for any integer $n\geq 1$, we have that when $d\geq 3$,*

$$\mathbb{E}\|\mathbf{T}^{-1}A_\varepsilon q\|_{L^n}^n \leq C_n\varepsilon^{\frac{n}{2}}\|q\|_{L^\infty}^n, \qquad \mathbb{E}\|\overline{\mathbf{T}^{-1}A_\varepsilon q}\|_{L^n}^n \leq C_n\varepsilon^n\|q\|_{L^\infty}^n \tag{38}$$

and if we iterate again,

$$\mathbb{E}\|\mathbf{T}^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon q\|_{L^n}^n \leq C_n\varepsilon^n\|q\|_{L^\infty}^n. \tag{39}$$

In dimension two, ε^n in the second and third estimate is replaced by $\varepsilon^n|\log\varepsilon|^{\frac{n}{2}}$.

Proof: First, we have

$$\mathbf{T}^{-1} = \mathbf{T}_1^{-1} - \mathbf{T}^{-1}\mathcal{K} = \mathbf{T}_1^{-1} - \mathbf{T}^{-1}A_2\mathbf{T}_1^{-1}. \tag{40}$$

Since $\mathbf{T}^{-1}A_2$ is bounded $L^n \to L^n$, we can replace $\mathbf{T}_1$ by $\mathbf{T}$ in the first estimate and in the first instance where $\mathbf{T}$ appears in third estimates. For the second estimate, we have

$$\overline{\mathbf{T}^{-1}A_\varepsilon q} = \overline{\mathbf{T}_1^{-1}A_\varepsilon q} - \overline{\mathbf{T}^{-1}\mathcal{K}A_\varepsilon q}.$$

The first term above is exactly the third item in the previous lemma. The second term above is bounded by $C\|\mathcal{K}A_\varepsilon q\|_{L^n}$ and it is exactly the second estimate in the previous lemma.

For the replacement of second $\mathbf{T}_1$ in the third estimate, we first write

$$\mathbf{T}_1^{-1}A_\varepsilon\mathbf{T}^{-1}A_\varepsilon q = \mathbf{T}_1^{-1}A_\varepsilon\mathbf{T}_1^{-1}A_\varepsilon q - \mathbf{T}_1^{-1}A_\varepsilon\mathbf{T}^{-1}\mathcal{K}A_\varepsilon q.$$

The first term is that in the lemma, and the second terms is estimated as follows:

$$\|\mathbf{T}_1^{-1}A_\varepsilon\mathbf{T}^{-1}\mathcal{K}A_\varepsilon q\|_{L^n} \leq \|\mathbf{T}_1^{-1}\|_{L^n\to L^n}(\|\delta a_\varepsilon\|_{L^{2n}} + \|\delta k_\varepsilon\|_{L^{2n}})\|\mathbf{T}^{-1}\|_{L^{2n}\to L^{2n}}\|\mathcal{K}A_\varepsilon q\|_{L^{2n}}.$$

Then we use the inequality $(a+b)^n \leq C_n(a^n + b^n)$ for $a, b \geq 0$, take the expectation and apply the Cauchy-Schwarz inequality to get the result. $\square$

Remark 4.3. All the results hold when $\mathbf{T}$ is replaced by $\mathbf{T}^*$ in the lemmas.

We are now ready to prove the first main result.

Proof of Theorem 2.1: We assume that $d \geq 3$. Only slight modifications left to the reader are needed when $d = 2$. Assume $u_0 \in L^\infty$ which is verified when $g \in L^\infty(\Gamma_-)$. Let $\chi_\varepsilon = \mathbf{T}^{-1}A_\varepsilon u_0$. We write $\zeta_\varepsilon = \chi_\varepsilon + z_\varepsilon$ and $\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \leq C\varepsilon$ by the previous lemmas, and it remains to analyze z_ε, which can be rewrite as the sum of of $z_{1\varepsilon} := -\mathbf{T}_\varepsilon^{-1}\delta a_\varepsilon\chi_\varepsilon$ and $z_{2\varepsilon} := \mathbf{T}_\varepsilon^{-1}\delta k_\varepsilon\bar{\chi}_\varepsilon$. From the previous lemma and the fact that δk_ε is in L^4, we conclude that

$$\mathbb{E}\|k_\varepsilon\bar{\chi}_\varepsilon\|_{L^2}^2 \leq [\mathbb{E}\|k_\varepsilon\|_{L^4}^4]^{\frac{1}{2}}[\mathbb{E}\|\bar{\chi}_\varepsilon\|_{L^4}^4]^{\frac{1}{2}} \leq C\varepsilon^2.$$

Then we recall that $\mathbf{T}_\varepsilon^{-1}$ is a bounded linear transform on L^2 and the bound is uniform in ε as long as we have a uniform subcriticality condition, which can be verified if $a_{r0} > \beta$. Therefore, we have

$$\mathbb{E}\|\mathbf{T}_\varepsilon^{-1}k_\varepsilon\bar{\chi}_\varepsilon\|_{L^2}^2 \leq \|\mathbf{T}_\varepsilon^{-1}\|_{L^2\to L^2}^2\mathbb{E}\|k_\varepsilon\bar{\chi}_\varepsilon\|_{L^2}^2 \leq C\varepsilon^2.$$

To control $z_{1\varepsilon}$, first we observe that

$$z_{1\varepsilon} = \mathbf{T}^{-1}(-\delta a_\varepsilon)\chi_\varepsilon + (\mathbf{T}_\varepsilon^{-1} - \mathbf{T}^{-1})(-\delta a_\varepsilon)\chi_\varepsilon = z_{11\varepsilon} + z_{12\varepsilon}.$$

For $z_{11\varepsilon}$, we use the third estimate in Corollary 4.2 and $\mathbb{E}\|z_{11\varepsilon}\|_{L^2}^2 \leq C\varepsilon^2$. For the $z_{12\varepsilon}$ term, we notice that it satisfies the equation

$$\mathbf{T}_\varepsilon z_{12\varepsilon} = A_\varepsilon z_{11\varepsilon}.$$

We then control the L^2 norm of $z_{12\varepsilon}$ by that of $A_\varepsilon z_{11\varepsilon}$. We have

$$\mathbb{E}\|z_{12\varepsilon}\|_{L^2}^2 \leq C\|\mathbf{T}_\varepsilon^{-1}\|_{L^2\to L^2}^2[\mathbb{E}\|a_\varepsilon\|_{L^4}^4 + \mathbb{E}\|\delta k_\varepsilon\|_{L^4}^4]^{\frac{1}{2}}[\mathbb{E}\|z_{11\varepsilon}\|_{L^4}^4]^{\frac{1}{2}} \leq C\varepsilon^2.$$

Hence we have shown that $\mathbb{E}\|z_\varepsilon\|_{L^2}^2 \leq C\varepsilon^2$. The proof is now complete. $\square$

5 Proof of Theorem 2.2 and 2.3: a weak CLT result

The main steps of the proof are as follows. As an application of the central limit theorem, we expect the fluctuations to be of order $\varepsilon^{\frac{d}{2}}$ with thus a variance of order $O(\varepsilon^d)$. Any contribution smaller than the latter order can thus be neglected. However, there are deterministic corrections of order larger than or equal to $\varepsilon^{\frac{d}{2}}$. We need to capture such correctors explicitly.

The deterministic and random correctors are obtained by expanding (19) as $\mathbf{T}\zeta_\varepsilon = A_\varepsilon u_0 + A_\varepsilon \zeta_\varepsilon$ in powers of A_ε. The number of terms in the expansion depends on dimension. We first consider the simpler case $d = 2$ and then address the case $d = 3$. Higher-order dimensions could be handled similarly but require tedious higher-order expansions in A_ε which are not considered here.

The derivation of the results are shown for random processes based on the Poisson point process described earlier for simplicity. As will become clear in the proof, what we need is that moments of order $2+2d$ (i.e., 6 in $d = 2$ and 8 in $d = 3$) of the random process be controlled. Any process that satisfies similar estimates would therefore lead to the same structure of the correctors as in the case of Poisson point process. Such estimates are however much more constraining than assuming statistical invariance and ergodicity, which is sufficient for homogenization [22]. For similar conclusions for elliptic equations, we refer the reader to e.g. [5, 7, 23].

5.1 The case of two dimensions

As outlined above, we have the iteration formula:

$$\zeta_\varepsilon = \mathbf{T}^{-1} A_\varepsilon u_0 + \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0 + \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \zeta_\varepsilon. \tag{41}$$

Let M by a test functions on $X \times V$, say continuous and compactly supported on X. After integration against this function on both sides of the expansion, we have

$$\langle \zeta_\varepsilon, M\rangle = \langle A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \zeta_\varepsilon, m\rangle. \tag{42}$$

Here we define $m = \mathbf{T}^{*-1} M$. We need to estimate the mean and variance of each term on the right hand side. We will show that in two dimensions, this expansion suffices. Weakly, the first term is mean-zero but is the leading-order term for the variance. The second term has a component whose mean is of order ε and converges to U as in Theorem 2.2. The other components of the second and third terms are shown to be smaller than $\varepsilon^{\frac{d}{2}}$ both in mean and in variance (weakly). The following lemmas prove these statements.

Let us call the terms in (42) as J_1, J_2 and R_1, respectively. Since u_0 is deterministic, and δa_ε and δk_ε are mean-zero, we obtain that J_1 is mean-zero. Its variance is easily seen to be of order ε^d and will be investigated later in detail. For the term J_2, we use the decomposition of $\mathbf{T}^{-1}$ and recast it as

$$J_2 = \langle A_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}_1^{-1} \mathcal{K} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}_1^{-1} \mathcal{K}\tilde{K} A_\varepsilon u_0, m\rangle,$$

and call the terms J_{21}, J_{22}, and J_{23}. Then we have the following estimates for them.

Lemma 5.1. *Assume the same condition of Theorem 2.1 hold. Let $d = 2$. Then we have*
(i) The mean of J_{21} is of order ε and more precisely,

$$\mathbb{E}\langle A_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, m\rangle = \varepsilon\langle U, M\rangle + o(\varepsilon) \tag{43}$$

where $U(x,v)$ is the solution to (10).
(ii) For the variance of J_{21}, we have

$$\mathrm{Var}\ \{J_{21}\} \le C\varepsilon^{d+2}|\log\varepsilon| \ll \varepsilon^d. \tag{44}$$

(iii) For J_{22} and J_{23}, we have

$$\mathbb{E}J_{22}^2 \le C\varepsilon^{2d}|\log\varepsilon|^2,\ \mathbb{E}J_{23}^2 \le C\varepsilon^{2d}. \tag{45}$$

Hence $\mathbb{E}|J_{2j}|$ for $j=2,3$ are much smaller than $\varepsilon^{\frac{d}{2}}$.
In dimension three, (i) is similar, and the logarithm in (ii) and (iii) can be dropped.

PROOF: (1) *The mean of J_{21}.* This term has an explicit expression.

$$\begin{aligned}J_{21} = &\int_{X\times V} m(x,v)\int_0^{\tau_-(x,v)} \big[E(x,x-tv)\big(\delta a_\varepsilon(x)\delta a_\varepsilon(x-tv)u_0 - \delta a_\varepsilon(x)\delta k_\varepsilon(x-tv)\bar u_0\big)\\ &+\int_V\int_0^{\tau_-(x,w)} E(x,x-sw)\big(-\delta k_\varepsilon(x)\delta a_\varepsilon(x-sw)u_0 + \delta k_\varepsilon(x)\delta k_\varepsilon(x-sw)\bar u_0\big)dw\big]dtdvdx.\end{aligned}$$

After taking expectation, we need to estimate

$$\begin{aligned}\mathbb{E}J_{21} = &\int_{X\times V} m(x,v)\int_0^{\tau_-(x,v)} \big[E(x,x-tv)\big(R_a(\frac{tv}{\varepsilon})u_0(x-tv,v) - R_{ak}(\frac{tv}{\varepsilon})\bar u_0(x-tv)\big)\\ &+\int_V\int_0^{\tau_-(x,w)} E(x,x-sw)\big(-R_{ak}(\frac{sw}{\varepsilon})u_0(x-sw,w) + R_k(\frac{sw}{\varepsilon})\bar u_0(x-sw)\big)dw\big]dtdvdx.\end{aligned}$$

Then we change variables $\frac{t}{\varepsilon}$ to t and $\frac{s}{\varepsilon}$ to s and obtain the following limit:

$$\begin{aligned}\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathbb{E}J_{21} = &\int_{X\times V} m(x,v)\int_{\mathbb{R}} \big[R_a(tv)u_0(x,v) - R_{ak}(tv)\bar u_0(x)\big)\\ &+\int_V\big(-R_{ak}(tw)u_0(x,w) + R_k(tw)\bar u_0(x)\big)dw\big]dtdvdx.\end{aligned}$$

The right hand side above is exactly $\langle M, U\rangle$ by definition.

(2) *The variance of J_{21}.* The moments and cross-correlations of the random coefficients δa_ε and δk_ε satisfy similar estimates. In the analysis of J_{21}, we therefore focus on the term that is quadratic in δa_ε knowing that the other three terms involving δk_ε are estimated in the same manner. We call I_1 the term quadratic in δa_ε to simplify notation, and using the change of variables (34), rewrite it as

$$\begin{aligned}I_1 &= \int_{X\times V} m(x,v)\delta a_\varepsilon(x)\int_0^{\tau_-(x,v)} E(x,x-tv)\delta a_\varepsilon(x-tv)u_0(x-tv,v)dtdxdv\\ &= \int_{X^2} m(x,\frac{x-y}{|x-y|})E(x,y)\frac{\delta a_\varepsilon(x)\delta a_\varepsilon(y)}{|x-y|^{d-1}}u_0(y,\frac{x-y}{|x-y|})dtdxdy.\end{aligned}$$

Then $\mathrm{Var}\ (I_1) = \mathbb{E}(I_1-\mathbb{E}I_1)^2$ can be written as

$$\begin{aligned}&\int_{X^4}\frac{m(x,v)m(x',v')u_0(y,v)u_0(y,v')E(x,y)E(x',y')}{|x-y|^{d-1}|x'-y'|^{d-1}}\\ &\quad(\mathbb{E}[\delta a_\varepsilon(x)\delta a_\varepsilon(y)\delta a_\varepsilon(x')\delta a_\varepsilon(y')] - \mathbb{E}[\delta a_\varepsilon(x)\delta a_\varepsilon(y)]\mathbb{E}[\delta a_\varepsilon(x')\delta a_\varepsilon(y')])d[x'y'xy].\end{aligned}$$

Now recalling the formula (31) for the fourth-order moment, we see that in the three choices of pairing the four points, the one that pairs x with y and x' with y' is the most singular term. Indeed, it is precisely $\mathbb{E}I_1^2$ and we've shown it is of order ε^2. However, this terms does not contribute to the variance, where only smaller terms appear.

Indeed, assuming that m and u_0 are uniformly bounded, we have

$$\begin{aligned}\mathrm{Var}\ (I_1) \le C \int_{X^4} \frac{1}{|x-y|^{d-1}|x'-y'|^{d-1}} \Big(& |T^4(\frac{y-x}{\varepsilon}, \frac{x'-x}{\varepsilon}, \frac{y'-x}{\varepsilon})| + \\ & |R(\frac{x'-x}{\varepsilon})R(\frac{y'-y}{\varepsilon})| + |R(\frac{x'-y}{\varepsilon})R(\frac{y'-x}{\varepsilon})| \Big).\end{aligned}$$

We estimate the three integrals. For the first integral, we change variables $(y-x)/\varepsilon \to y$, $(x'-x)/\varepsilon \to x'$, and $(y'-x)/\varepsilon \to y'$. Then the integral becomes

$$\varepsilon^{3d-2(d-1)} \int_X dx \int_{(\frac{X-x}{\varepsilon})^3} \frac{|T^4(y,x',y')|}{|y|^{d-1}|x'-y'|^{d-1}} d[yx'y'].$$

We replace the integration domain of $[y, x', y']$ to $\mathbb{R}^{3d}$. The resulting integral is finite:

$$\int_{\mathbb{R}^{3d}} \frac{|T^4(y,x',y')|}{|y|^{d-1}|x'-y'|^{d-1}} d[yx'y'] \le \left(\int \psi(z) \psi * \frac{1}{|y|^{d-1}}(z) dz \right)^2.$$

The first integral gives a contribution of order ε^{d+2} to the variance.

The other two integrals are handled in a similar way. Noting the symmetry between x' and y', we consider only the second integral. We change variables $(x-x')/\varepsilon \to x$, $(y-y')/\varepsilon \to y$, and $(x'-y)/\varepsilon \to x'$. Then, we have

$$\begin{aligned}&\int_{X^4} \frac{|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})|}{|x-y|^{d-1}|x'-y'|^{d-1}} d[x'y'xy] = \int_{X^4} \frac{|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})|}{|x'-y+(x-x')|^{d-1}|x'-y+(y-y')|^{d-1}} d[x'y'xy] \\ &\le \int_X dy' \int_{\mathbb{R}^{2d}} |R(x)R(y)| d[xy] \varepsilon^{2d} \int_{2X} \frac{1}{|x'+\varepsilon x|^{d-1}|x'+\varepsilon y|^{d-1}} dx'.\end{aligned}$$

For the integral in x', we use the convolution lemma A.1. In dimension two, the integral is bounded by $C(|\log|x-y|| + |\log \varepsilon|)$. Hence, the above integral is bounded by

$$C|X| \left(\varepsilon^{2d} |\log \varepsilon| \int_{\mathbb{R}^{2d}} |R(x)R(y)| dxdy + \varepsilon^{2d} \int_{\mathbb{R}^{2d}} |R(x)R(y) \log|x-y|| dxdy \right).$$

Again, the integrals are bounded. The first one is trivial. The second one is again a convolution of a compactly supported function with a locally integrable function. This yields a contribution of order $\varepsilon^4|\log \varepsilon|$ in dimension two and ε^{d+2} in dimension three.

Observe that $2d = d+2$ in dimension two. We conclude that

$$\mathrm{Var}\ \{I_1\} \le C|X| \|u_0 m\|_{L^\infty}^2 \varepsilon^{d+2} |\log \varepsilon|.$$

(3) *The absolute mean of* J_{22}. We recast J_{22} as

$$\begin{aligned}&\int_{X\times V} m(x,v)\delta a_\varepsilon(x) \int_0^{\tau_-(x,v)} E(x, x-tv) \int_X \frac{E(x-tv,y)\delta a_\varepsilon(y)}{|x-tv-y|^{d-1}} u_0(y, \frac{x-tv-y}{|x-tv-y|}) d[ytx] \\ &= \int_{X^3} m(x, \frac{x-y}{|x-y|}) \frac{\delta a_\varepsilon(x) \delta a_\varepsilon(y)}{|x-z|^{d-1}|z-y|^{d-1}} u_0(y, \frac{z-y}{|z-y|}) d[xzy].\end{aligned}$$

Using the decomposition of the fourth order moments, the problem reduces to estimating similar integrals as was done before. Since there is another integration in z, this term is more regular than the ballistic part and the mean square of this term is negligible compared to the random fluctuations. We have

$$\begin{aligned}\mathbb{E}J_{22}^2 \leq& C\int_{X^6}\big(|T^4(\frac{y-x}{\varepsilon},\frac{z-x}{\varepsilon},\frac{y'-x}{\varepsilon})| + |R(\frac{x-y}{\varepsilon})R(\frac{x'-y'}{\varepsilon})| + |R(\frac{x-y'}{\varepsilon})R(\frac{x'-y}{\varepsilon})| \\ &+ |R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})|\big)\frac{1}{|x-z|^{d-1}|z-y|^{d-1}|x'-z'|^{d-1}|z'-y'|^{d-1}}d[xyzx'y'z'].\end{aligned}$$

We integrate over z and z' first. Using the convolution lemma, we obtain

$$\begin{aligned}\mathbb{E}J_{22}^2 \leq& C\int_{X^4}\big(|T^4(\frac{y-x}{\varepsilon},\frac{x'-x}{\varepsilon},\frac{y'-x}{\varepsilon})| + |R(\frac{x-y}{\varepsilon})R(\frac{x'-y'}{\varepsilon})| + |R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})| \\ &+ |R(\frac{x-y'}{\varepsilon})R(\frac{x'-y}{\varepsilon})|\big)|\log|x-y|\log|x'-y'||d[xyx'y'].\end{aligned}$$

The most singular term arises when the correlation function and the logarithmic functions have the same singularity. These most singular terms are treated as follows. For the integral

$$\int_{X^4}|R(\frac{x-y}{\varepsilon})R(\frac{x'-y'}{\varepsilon})||\log|x-y|\log|x'-y'||d[xyx'y'],$$

we change variables $(x-y)/\varepsilon \to y$ and $(x'-y')/\varepsilon \to y'$ and the integral is bounded by

$$\varepsilon^{2d}|\log\varepsilon|^2\int_{X^2}d[xx']\left(\int_{\mathbb{R}^d}|R(y)|dy\right)^2.$$

The integral is finite for the same reasons as before.

The other contributions in the variance of J_{22} are negligible compared to this contribution. For the third integral, which is identical with the fourth integral, we need to control

$$\int_{X^4}|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})||\log|x-y|\log|x'-y'||d[xyx'y'].$$

We first change variables $(x-x')/\varepsilon \to x'$, $(y-y')/\varepsilon \to y'$, and $x-y \to y$, and then use the convolution lemma A.1 in the integral in y. Observe that the integral of the product of log functions on bounded domains is uniformly bounded. Hence we find that this term is of order ε^{2d}.

For the first integral involving T^4, after changing variables, we need to consider

$$C|X|\left(\varepsilon^{3d}|\log\varepsilon|^2\int|T^4(y,x',y')d[yx'y'] + \varepsilon^{3d}|T^4(y,x',y')\log|y|\log|x'-y'||d[yx'y']\right)$$

and the integrals converge as before. Hence, the contribution to the variance is of order $\varepsilon^{3d}|\log\varepsilon|^2$. To summarize, we have obtained that

$$\mathrm{Var}\ (J_{22}) \lesssim \varepsilon^{2d}, \qquad \mathbb{E}J_{22}^2 \lesssim \varepsilon^{2d}|\log\varepsilon|^2.$$

In dimension three, the logarithm terms can be eliminated.

(4) *The absolute mean of J_{23}*. This term has the following expression:

$$\int_{X\times V} m(x,v)\delta a_\varepsilon(x)\int_0^{\tau_-(x,v)} E(x,x-tv)\int_X \Theta(x-tv,z)\int_X \frac{E(z,y)\delta a_\varepsilon(y)}{|z-y|^{d-1}} u_0(y,v)d[yztxv]$$
$$= \int_{X^4} \frac{m(x,v)E(x,\xi)\Theta(\xi,z)E(z,y)u_0(y,v')}{|x-\xi|^{d-1}|z-y|^{d-1}} d[x\xi zy],$$

where $v=(x-\xi)|x-\xi|^{-1}$ and $v'=(z-y)|z-y|^{-1}$. Assume that m and u_0 are bounded. Then $\mathbb{E}J_{23}^2$ can be bounded by

$$C\int_{X^8} \frac{\mathbb{E}[\delta a_\varepsilon(x)\delta a_\varepsilon(y)\delta a_\varepsilon(x')\delta a_\varepsilon(y')]}{|x-\xi|^{d-1}|\xi-z|^{d-1}|z-y|^{d-1}|x'-\xi'|^{d-1}|\xi'-z'|^{d-1}|z'-y'|^{d-1}} d[x\xi zyx'\xi'z'y'].$$

The analysis of this term is exactly as in (ii). We integrate over ξ,ξ' first and then z,z'. Then all potentials disappear in two dimensions and integrable logarithm terms emerge in three dimensions and hence we find that

$$\text{Var } (J_{23}) \lesssim \varepsilon^{2d}, \qquad \mathbb{E}J_{23}^2 \lesssim \varepsilon^{2d}.$$

This completes the proof when $d=2$. In three dimensions, the only change needed is to discard the logarithm terms in part (2) above. $\square$

Next we consider the remainder term R_1. Recall that $\zeta_\varepsilon = \chi_\varepsilon + z_\varepsilon$. We see that R_1 can be written as

$$R_1 = \langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon \mathbf{T}^{-1}A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon z_\varepsilon, m\rangle$$

We will call them R_{11} and R_{12} respectively. We have the following estimates.

Lemma 5.2. *Assume the same conditions as in the previous lemma. Then we have:*
(i) The absolute mean of R_{12} is smaller than $\varepsilon^{\frac{d}{2}}$. More precisely, we have

$$\mathbb{E}|\langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon z_\varepsilon, m\rangle| \le C\varepsilon^{\frac{3}{2}}|\log\varepsilon|^{\frac12} \ll \varepsilon^{\frac{d}{2}}$$

in dimension $d=2$.
(ii) The absolute mean of the term R_{11} is also smaller than $\varepsilon^{\frac{d}{2}}$. More precisely, we have

$$\mathbb{E}|\langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon \mathbf{T}^{-1}A_\varepsilon u_0, m\rangle| \le C\varepsilon^2|\log\varepsilon| \ll \varepsilon^{\frac{d}{2}},$$

in dimension $d=2$. When $d=3$, the size is ε^2.

PROOF: (1) *The term R_{12}*. Use the duality relation we can write this term as $\langle z_\varepsilon, A_\varepsilon \mathbf{T}^{*-1}A_\varepsilon m\rangle$. Then we have

$$\mathbb{E}|R_{12}| \le C\{\mathbb{E}\|z_\varepsilon\|_{L^2}^2\}^{\frac12}\{\mathbb{E}(\|\delta a_\varepsilon\|_{L^4}^4 + \|\delta k_\varepsilon\|_{L^4}^4)\}^{\frac14}\{\mathbb{E}\|\mathbf{T}^{*-1}A_\varepsilon m\|_{L^4}^4\}^{\frac14}.$$

Using lemma 3.6 and corollary 4.2, and the fact that $\mathbb{E}\|z_\varepsilon\|_{L^2}^2 \le C\varepsilon^2|\log\varepsilon|$ derived in the proof of Theorem 2.1, the three terms on the right-hand side above are of size $\varepsilon|\log\varepsilon|^{\frac12}$, order $O(1)$, and $\varepsilon^{\frac12}$, respectively.

(2) *The term R_{11}*. Write this term as $\langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon u_0, \mathbf{T}^{*-1}A_\varepsilon m\rangle$, and use the decomposition of $\mathbf{T}$ and $\mathbf{T}^*$. Then we have

$$R_{11} = \langle A_\varepsilon \mathbf{T}_1^{-1}A_\varepsilon u_0, \mathbf{T^*}_1^{-1}A_\varepsilon m\rangle - \langle A_\varepsilon \mathbf{T}_1^{-1}A_\varepsilon u_0, \mathbf{T}^{*-1}\mathcal{K}^*A_\varepsilon m\rangle - \langle A_\varepsilon \mathbf{T}^{-1}\mathcal{K}A_\varepsilon u_0, \mathbf{T}^{*-1}A_\varepsilon m\rangle.$$

We will call them I_1, I_2 and I_3 respectively. Then I_2 and I_3 are of the same form and can be controlled as follows:

$$\mathbb{E}|I_2| \leq C\{\mathbb{E}\|\mathbf{T}^{-1}A_\varepsilon \mathbf{T}_1^{-1}A_\varepsilon u_0\|_{L^2}^2\}^{\frac{1}{2}}\{\mathbb{E}(\|\delta a_\varepsilon\|_{L^4}^4 + \|\delta k_\varepsilon\|_{L^4}^4)\}^{\frac{1}{4}}\{\mathbb{E}\|\mathcal{K}^* A_\varepsilon m\|_{L^4}^4\}^{\frac{1}{4}}.$$

We then use lemma 3.6 and corollary 4.2 again to obtain the desired control for I_2 and similarly for I_3.

For I_1, we only need to consider the term $\langle \mathbf{T^*}_1^{-1} A_\varepsilon m, \delta a_\varepsilon \mathbf{T}_1^{-1}A_\varepsilon u_0\rangle$ because the other component is as I_2 and is controlled in the same manner. We still call this term I_1 and it has the expression:

$$\int_{X\times V}\left(\int_0^{\tau_+} E(x,x+tv)\delta a_\varepsilon(x+tv)m(x+tv,v)dt\right)$$
$$\delta a_\varepsilon(x)\left(\int_0^{\tau_-} E(x,x-sv)\delta a_\varepsilon(x-sv)u_0(x-sv,v)ds\right)dxdv$$

where $\tau_\pm$ are short for $\tau_\pm(x,v)$. Assume that m and u_0 are uniformly bounded. The mean square of I_1 is bounded by

$$C\int_{X^2\times V^2}\int_0^{\tau_+}\int_0^{\tau_-}\int_0^{\tau_+}\int_0^{\tau_-}\mathbb{E}[\delta a_\varepsilon(x+tv)\delta a_\varepsilon(x)\delta a_\varepsilon(x-sv)$$
$$\delta a_\varepsilon(x'+t'v')\delta a_\varepsilon(x')\delta a_\varepsilon(x'-s'v')]d[s't'stx'v'xv].$$

We use the high-order moment formula again, and then need to control several integrals involving T^{n_j}'s. The analysis is exactly the same as the previous lemma although there are more terms.

Let us divide the six-point set into two categories: the first one consists of $x, x+tv, x-sv$ and the second one consists of $x', x'+t'v', x'-s'v'$. The non-single partitions of a six-point set include group of (2,2,2), (2,4) and (3,3). Among these groupings, there is one term where only points from the same category are grouped together; it is the following:

$$C\int_{X^2\times V^2}\int_0^{\tau_+}\int_0^{\tau_-(x,v)} T^3(\frac{tv}{\varepsilon}, -\frac{sv}{\varepsilon})dtds\int_0^{\tau_+}\int_0^{\tau_-(x',v')} T^3(\frac{t'v'}{\varepsilon}, -\frac{s'v'}{\varepsilon})dt'ds'd[xx'vv'].$$

Change variable and recall that T^3 is integrable along all directions. We see this term is of order ε^4.

For all other partitions except some terms in the pattern (2,2,2) which we will discuss later, there is at least one point from the first category and one from the second category that are grouped together; without loss of generality we can assume x and x' are grouped together. In the (3,3) grouping pattern, there is another point from the same category of either x or x' that is grouped with them. This yields a term of the form $T^3(\frac{x-x'}{\varepsilon}, \frac{tv}{\varepsilon})$ and after routine change of variables, the integration of x' yields a term of size ε^d and the integration of t yields another multiplication by a term of order ε so that the whole integral is no larger than order ε^{d+1}. Similarly, if x and x' are grouped together in a (2,4) pattern, the same analysis holds and we still have enough variables to integrate and the term is no larger than ε^{d+1}.

For the pattern (2,2,2), the terms of the form

$$C\int_{X^2\times V^2}\int_0^{\tau_-}\int_0^{\tau_-'}\int_0^{\tau_+}\int_0^{\tau_+'} R(\frac{x-x'}{\varepsilon})R(\frac{x-x'-tv+t'v'}{\varepsilon})$$
$$\times R(\frac{x-x'+sv-s'v'}{\varepsilon}))d[xyvwtst's'],$$

needs separate consideration. For this term, we can use change of variables in $tv - t'v'$ to an integration over a two-dimension region and integration over $\frac{1}{\sin\theta}$ for some angular variable. In two dimension, this is of order $\varepsilon^{d+2}|\log\varepsilon|$, and in dimension three this is of order ε^{d+2}. The lemma is proved. $\square$

Now we are ready to prove the last two main theorems in the case of $d = 2$. However, we will postpone it after briefly discussing the case of $d = 3$.

5.2 Extension to dimension three

The analysis for J_2 still holds in dimension three. However, the estimate on R_1 is not sufficient and we need to push the iteration to have one additional term:

$$\langle \zeta_\varepsilon, M\rangle = \langle A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \zeta_\varepsilon, m\rangle. \quad (46)$$

Let us call the third above term J_3 and the fourth R_2. Then J_3 is precisely the first component of R_1 in dimension two and has been estimated in Lemma 5.2. Now it suffices to estimate R_2. We first rewrite this term as

$$R_2 = \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon z_\varepsilon, m\rangle.$$

Call them R_{21} and R_{22} respectively and we have the following lemma.

Lemma 5.3. *Under the same condition of previous lemmas, let $d = 3$. We have*
(i) For the absolute mean of R_{22}, we have $\mathbb{E}|R_{22}| \le C\varepsilon^2 \ll \varepsilon^{\frac{d}{2}}$.
(ii) For the term R_{21}, we have $\mathbb{E}|R_{21}| \le C\varepsilon^2 \ll \varepsilon^{\frac{d}{2}}$.

Proof: (1) *The term R_{22}.* We can write this term as $\langle A_\varepsilon z_\varepsilon, \mathbf{T}^{*-1} A_\varepsilon \mathbf{T}^{*-1} A_\varepsilon m\rangle$. Then it is controlled as follows.

$$\mathbb{E}|R_{22}| \le C\{\mathbb{E}\|\mathbf{T}^{*-1} A_\varepsilon \mathbf{T}^{*-1} A_\varepsilon m\|_{L^4}^4\}^{\frac14}\{\mathbb{E}\|\delta k_\varepsilon\|_{L^4}^4 + \mathbb{E}\|\delta a_\varepsilon\|_{L^4}^4\}^{\frac14}\{\mathbb{E}\|z_\varepsilon\|_{L^2}^2\}^{\frac12} \le C\varepsilon^2.$$

(2) *The term R_{21}.* We can write this term as $\langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, \mathbf{T}^{*-1}\mathcal{K}^* A_\varepsilon m\rangle$. Using the decomposition of T and $\mathbf{T}^*$ we can break this term into four components. The same analysis as in Lemma 5.2 applies and we only need to consider the term $\langle \delta a_\varepsilon \mathbf{T}_1^{-1} \delta a_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, \mathbf{T^*}_1^{-1} A_\varepsilon m\rangle$. It has the expression:

$$\begin{aligned}\int_{X\times V}\int_0^{\tau_-(x,v)}\int_0^{\tau_-(x-tv,v)}\int_0^{\tau_+(x,v)} & \delta a_\varepsilon(x)\big(u_0 E(x, x-tv)\delta a_\varepsilon(x-tv)u_0 \\ & \times E(x-tv, x-tv-t_1v)\delta a_\varepsilon(x-tv-t_1v)u_0 E(x, x+sv)\delta a_\varepsilon(x+sv)\big)d[tsxv].\end{aligned}$$

Then $\mathbb{E}R_{21}^2$ is bounded by

$$\begin{aligned}C\int_{X^2\times V^2}\int_0^{\tau_{-1}}\int_0^{\tau_{-2}}\int_0^{\tau_+}\int_0^{\tau_{-1}'}\int_0^{\tau_{-2}'}\int_0^{\tau_+'} & \mathbb{E}\{\delta a_\varepsilon(x)\delta a_\varepsilon(x-tv)\delta a_\varepsilon(x-tv-t_1v)\delta a_\varepsilon(x+sv) \\ & \delta a_\varepsilon(y)\delta a_\varepsilon(y-t'w)\delta a_\varepsilon(y-t'w-t_1'w)\delta a_\varepsilon(y+s'w)\big)d[xyvwtst_1t's't_1'].\end{aligned}$$

Then we use the eighth order moments formula.

For non-single partitions of eight points, the patterns are (2,2,2,2), (2,2,4), (2,3,3), (2,6), (3,5) and (4,4). Again, we divide the points into two categories, the first one including $x, x-tv, x-tv-t_1v, x+sv$, and the second including $x', x'-t'v', x'-t'v'-t_1'v', x'+s'v'$. Now the partitions when only points from the same category are grouped together yields the following term:

$$C\left(\int_{X\times V}\int_0^{\tau_-}\int_0^{\tau_+}\mathbb{E}\{\delta a_\varepsilon(x)\delta a_\varepsilon(x-tv)\delta a_\varepsilon(x-tv-t_1v)\delta a_\varepsilon(x+sv)d[tt_1sxv]\right)^2.$$

We have seen that this term is of order $(\varepsilon^2)^2$. For all other partitions, x and x' are grouped together, and except for some terms in the pattern (2,2,2,2) which we will discuss later, there is another independent t variable that can be integrated over. Therefore, these terms are of order no larger than ε^{d+1}.

In the pattern (2,2,2,2), the terms of the form

$$\begin{aligned}C\int_{X^2\times V^2}\int_0^{\tau_{-1}}\int_0^{\tau_{-2}}\int_0^{\tau_+}\int_0^{\tau_{-1}'}\int_0^{\tau_{-2}'}\int_0^{\tau_+'} R(\frac{x-x'}{\varepsilon})R(\frac{x-x'-tv+t'v'}{\varepsilon})\\ \times R(\frac{x-x'-tv-t_1v+t'v'+t_1'v'}{\varepsilon})R(\frac{x-x'+sv-s'v'}{\varepsilon}))d[xyvwtst_1t's't_1'],\end{aligned}$$

need separate consideration. As in the previous lemma, we can change variable in $tv-t'v'$. These terms are of order ε^{d+2}. Hence the lemma is proved. □

5.3 Limit of the deterministic corrector

With the results above, the proof of Theorem 2.2 is immediate.

Proof of Theorem 2.2: In dimension 2 and 3, consider the expansion (42) or (46), we see the only term whose contribution to $\mathbb{E}\{\zeta_\varepsilon\}$ is larger than $\varepsilon^{\frac{d}{2}}$ is $\langle A_\varepsilon \mathbf{T}_1^{-1}A_\varepsilon u_0, m\rangle$ and its limit is already derived in Lemma 5.1. □

5.4 Limit distribution of the random corrector

The following result follows immediately from the lemmas proved earlier in this section.

Lemma 5.4. *Under the same conditions as in Theorem 2.1, let $d=2,3$. We have*

$$\mathbb{E}|\langle\varphi, \frac{\zeta_\varepsilon-\mathbb{E}\zeta_\varepsilon}{\varepsilon^{\frac{d}{2}}}-\varepsilon^{-\frac{d}{2}}\mathbf{T}^{-1}A_\varepsilon u_0\rangle| \lesssim \varepsilon^{\frac12}|\log\varepsilon|^{\frac12}\longrightarrow 0. \tag{47}$$

This lemma states that $(\zeta_\varepsilon-\mathbb{E}\zeta_\varepsilon)\varepsilon^{-\frac{d}{2}}$ converges to $\varepsilon^{-\frac{d}{2}}\mathbf{T}^{-1}A_\varepsilon u_0$ weakly and in mean (root mean square), which implies convergence weakly and in distribution. Therefore, we seek the limiting distribution of:

$$\langle\varphi,\varepsilon^{-\frac{d}{2}}\mathbf{T}^{-1}A_\varepsilon u_0\rangle = -\varepsilon^{-\frac{d}{2}}(\langle\mathbf{T}^{*-1}\varphi,\delta a_{r\varepsilon}u_0\rangle+\langle\mathbf{T}^{*-1}\varphi,\delta k_\varepsilon(-\bar u_0+c_d u_0)\rangle).$$

When φ is taken to be $M_l, 1\le l\le L$ as in the section on main results, the resulting random variables are

$$I_{l\varepsilon}=\varepsilon^{-\frac{d}{2}}\int_{\mathbb{R}^d} m_l\cdot(\delta a_r(\frac{x}{\varepsilon}),\delta k(\frac{x}{\varepsilon}))dx, \tag{48}$$

where m is defined in Section 2.

As in Remark 2.5, proving Theorem 2.3 is equivalent to prove that $\{I_{l\varepsilon}\}$ converge in distribution to mean zero Gaussian random variables $\{I_l(\omega)\}$, whose covariance matrix of the random variables I_l is given by

$$\mathbb{E}I_iI_j = \int m_{1i}m_{1j}\sigma_a^2 + (m_{1i}m_{2j} + m_{2i}m_{1j})\rho_{ak}\sigma_a\sigma_k + m_{2i}m_{2j}\sigma_k^2 dx = \int m_i \otimes m_j : \Sigma dx. \quad (49)$$

Here, the covariance matrix Σ is defined in (14).

Note that $I_{l\varepsilon}$ is an oscillatory integral. Convergence of such integrals to a Gaussian random variable can be seen as a generalization of the Central Limit Theorem which is classically stated for independent sequences of random variables. Generalizations to processes on lattice points which are not independent but "independent in the limit", usually shown through *mixing* properties of the process, are done in the probability literature; see e.g. [14]. Generalizations to oscillatory integrals are done in [5] under similar *mixing* conditions. We will follow the procedure of the latter paper, to which we refer the reader for omitted details.

Proof of Theorem 2.3: For simplicity we assume that u_0 and hence m_l are continuous.

1. Observe that if we approximate m_l by piece-wise constant functions m_{lh}, which we define in the next step, the corresponding integrals $I_{lh\varepsilon}$ approximate $I_{l\varepsilon}$ in mean square and uniformly in ε, that is,

$$\mathbb{E}(I_{lh\varepsilon} - I_{l\varepsilon})^2 \leq C\|m_{lh} - m_l\|_{L^\infty}^2. \quad (50)$$

This is due to the fact that R_a, R_k, R_{ak} are integrable. Hence we can consider the limit distribution of $I_{lh\varepsilon}$.

2. Now we explicit our choice of m_{lh}. We use a uniform mesh of size h to divide the domain into small cubes of size h. On each cube that is contained in X (and the cubes that intersect with the boundary will be omitted), we replace m_l in the integral $I_{l\varepsilon}$ by its value at the grid point, say the center point. That is,

$$I_{lh\varepsilon j} = \int_{\mathcal{C}_j^h} m_{lhj} \cdot \frac{1}{\varepsilon^{\frac{d}{2}}}(\delta a_r(\frac{y}{\varepsilon}, \omega), \delta k(\frac{y}{\varepsilon}, \omega))' dy$$

where $\mathcal{C}_j^h$ is the j-th cube and h is its size (size of its sides). The vector m_{lhj} is the vector function m_l evaluated at the center of $\mathcal{C}_j^h$. Neglecting small boundary contributions, $I_{lh\varepsilon} = \sum_{j=1}^J I_{lh\varepsilon j}$ where J is the total number of cubes and is of order h^{-d}.

3. Then we can show that the variables $I_{lh\varepsilon j}$ are independent in the limit $\varepsilon \downarrow 0$, so the limiting distribution of $I_{lh\varepsilon}$ is the sum of the limiting distribution of $I_{lh\varepsilon j}$.

4. Without loss of generality, let us consider the limit distribution of $I_{lh\varepsilon 0}$ which corresponds to the cube around 0 (which we assume is inside X). We divide the cube further into $N = \frac{h}{\varepsilon}$ smaller cubes uniformly. We find

$$I_{lh\varepsilon 0} = \left(\frac{h}{N}\right)^{\frac{d}{2}} \int_{\mathcal{C}^N} q(y,\omega)dy = \left(\frac{h}{N}\right)^{\frac{d}{2}} \sum_{i\in\mathbb{Z}^d\cap\mathcal{C}^N} \int_{\mathcal{C}^i} q(y,\omega)dy = h^{\frac{d}{2}} \frac{1}{N^{\frac{d}{2}}} \sum_{i\in\mathbb{Z}^d\cap\mathcal{C}^N} \hat{q}_i. \quad (51)$$

Here, $\mathcal{C}^N$ is the cube of size N and $\mathcal{C}_i$'s are the unit cubes in $\mathcal{C}^N$. Hence i runs over a part of $\mathbb{Z}^d$ that belongs to $\mathcal{C}^N$. For simplicity, we used the notation

$$\hat{q}_i = \int_{\mathcal{C}_i} q(y)dy, \; q(y) = m_{lh0} \cdot (\delta a_r(y,\omega), \delta k(y,\omega))'.$$

We also use the fact that $\delta a_{r\varepsilon}, \delta k_\varepsilon$ are stationary hence we can assume the cubes are centered at lattice points on $\mathbb{Z}^d$.

5. Then we recall CLT for *mixing* processes parameterized by lattice points, which states that

$$\frac{1}{\sigma N^{\frac{d}{2}}} \sum_{i \in \mathbb{Z} \cap \mathcal{C}^N} \hat{q}_i \longrightarrow \mathcal{N}(0,1), \tag{52}$$

in distribution as $N \to \infty$, where $\mathcal{N}(0,1)$ is the standard normal distribution, and $\sigma^2 = \sum_{i \in \mathbb{Z}^d} \mathbb{E}(\hat{q}_0 \hat{q}_i)$. Therefore, we have $I_{lh\varepsilon j} \to \mathcal{N}(0, \sigma_j^2 h^d)$ where

$$\begin{aligned}
\sigma_j^2 &= \sum_{i \in \mathbb{Z}^d} \mathbb{E}(\hat{q}_0^j \hat{q}_i^j) = \sum_{i \in \mathbb{Z}^d} \mathbb{E} \int_{\mathcal{C}_0} m_{lhj} \cdot (\delta a_r, \delta k)(y) dy \int_{\mathcal{C}_i} m_{lhj} \cdot (\delta a_r, \delta k)(z) dz \\
&= m_{lhj} \otimes m_{lhj} \int_{\mathcal{C}_0} dy \int_{\mathbb{R}^d} dz \mathbb{E}[(\delta a_r, \delta k)(y) \otimes (\delta a_r, \delta k)(z)] \\
&= m_{lhj} \otimes m_{lhj} : \begin{pmatrix} \sigma_a^2 & \rho_{ak}\sigma_a\sigma_k \\ \rho_{ak}\sigma_a\sigma_k & \sigma_k^2 \end{pmatrix}.
\end{aligned} \tag{53}$$

Then we see $I_{lh\varepsilon} \to \sum_j \mathcal{N}(0, \sigma_j^2) = I_{lh}$, which is a Gaussian random variable with variance $\int m_{lh} \otimes m_{lh} : \Sigma dy$.

6. Then we pass to the limit $h \to 0$ to get the result. $\square$

Remark 5.5. The CLT of oscillatory integral developed in [5] assumes that the function m_l is continuous. Generalization to the case when m_l is in L^2 is straightforward since continuous functions are dense in L^2. We cannot generalize this further because for the resulted Gaussian variable to have a bounded variance, we need $m \in L^2$.

Remark 5.6. From the estimates on the mean and variance of the terms on the right hand side of the expansion, we see that $\mathbb{E}\{\zeta_\varepsilon\}$ in the theorem can be replaced by the mean of $\mathbf{T}^{-1} A_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0$ because other terms have contributions to the mean of size smaller than the random fluctuations. Furthermore, when R_a and the other correlations decay fast so that rR is integrable in each direction, which is the case in our model, then $\mathbb{E}\{\zeta_\varepsilon\}$ can be replaced by $\varepsilon U(x,v)$.

Remark 5.7. **Anisotropic scattering kernel**. For simplicity, we assumed that scattering k_ε was isotropic. All the results presented here generalize to the case $k_\varepsilon(x, v', v) = k_\varepsilon(x) f(v, v')$, where $f(v, v')$ is a known, bounded, function and k_ε is defined as before. All the required L^∞ estimates used in the derivation are clearly satisfied in this setting.

Generalization to scattering kernels of the form

$$k_\varepsilon(x, v', v) = \sum_{j=1}^{J} k_{\varepsilon j}(x) Y_j(v, v')$$

where J is finite and Y_j's are the spherical harmonics and the terms $k_{\varepsilon j}$ are defined as $k_\varepsilon(x)$ above is also possible. In this case, we need to deal with a finite system of integral equations and the analysis is therefore slightly more cumbersome.

Acknowledgment: We would like to thank Alexandre Jollivet for discussions on the theory of transport. The work was supported in part by NSF Grants DMS-0554097 and DMS-0804696.

A Estimates of convolution of potentials

Lemma A.1. *Let X be an open and bounded subset in $\mathbb{R}^d$, and $x \neq y$ two points in X. Let α, β be positive numbers in $(0, d)$. We have the following convolution results.*

1. *If $\alpha + \beta > d$, then*
$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} \leq C \frac{1}{|x-y|^{\alpha+\beta-d}} \tag{54}$$

2. *If $\alpha + \beta = d$, then*
$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} \leq C(|\log|x-y|| + 1) \tag{55}$$

3. *If $\alpha + \beta < d$, then*
$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} \leq C. \tag{56}$$

The convolution of logarithms with a weak singular potential turns out to be finite as follows:

$$\int_X |\log|z-x|| \frac{1}{|z-y|^\alpha} \lesssim 1. \tag{57}$$

The above constants do not depend on the distance between x and y.

Proof: Let $\rho = |x-y|$. Let the C_x, C_y be spheres with radius ρ centered at x and y respectively, and B_x, B_y the balls enclosed. The common section of the two balls divide their union into two symmetric parts, one containing x and the other containing y. Let D_1, D_2 denote the two parts respectively and D_3 the remaining part in X. On D_1, $|z-x| \geq \rho/2$, hence

$$\int_{D_1} \frac{1}{|z-x|^\alpha} \frac{1}{|z-y|^\beta} dz \lesssim \frac{1}{\rho^\beta} \int_{B_x} \frac{1}{|z-x|^\alpha} dz \lesssim \frac{1}{\rho^{\alpha+\beta-d}}.$$

Similarly we have the same conclusion on D_2. On D_3, it is clear that $|z-x|/2 \leq |z-y| \leq 2|z-x|$, and hence we have

$$\int_{D_3} \frac{1}{|z-x|^\alpha} \frac{1}{|z-y|^\beta} dz \lesssim \int_{D_3} \frac{1}{|z-x|^{\alpha+\beta}} dz.$$

In the case of $\alpha + \beta > d$, the last integral is bounded by $\rho^{-\alpha-\beta+d}$; in the case of $\alpha + \beta < d$, the integral is bounded since the domain is bounded; in the case of $\alpha + \beta = d$, the last integral is bounded by $|\log \rho|$ plus some constant depending on the diameter of X. However, we are interested in x close to y and hence the logarithm term dominates. This completes the first part of the lemma.

Using the same procedure and the fact that

$$\int_0^R \frac{\log r}{r^{\delta-1}} dr \leq C_{R,\delta}$$

for all bounded R and $\delta > 0$, the second part is similarly proved. $\square$

B Higher order moments of random fields

As remarked in section 3, it suffices to derive the statistics of processes of the form

$$b(x,\omega) = \sum_{j=1}^{\infty} \phi(x - y_j(\omega)), \tag{58}$$

where $Y = \{y_j\}$ is a Poisson point process with intensity ν. We denote by δb the mean-zero process $b - \mathbb{E}b$. We develop a systematic formula for the n-th order moment of b and δb. The moments and cross-moments of the random model (3) then follows.

The moments of b are expectations of product of sums. We recall that $[n]$ denotes the set $\{1, 2, \cdots, n\}$, that $\mathcal{P}_n$ is the set of all partitions of n, i.e., the set of arrays $(n_1, n_2, \cdots, n_k)$ satisfying that $\sum_{i=1}^{k} n_i = n$ and $1 \le n_1 \le n_2 \le \cdots \le n_k$, and that $\mathscr{G}_n$ is the set of non-single partitions of n, i.e., $n_1 \ge 2$. Define a *partition of* $[n]$ to be a collection of nonempty mutually disjoint subsets $\{A_i\}$ such that $\cup A_i = [n]$. The total number of all possible partitions of $[n]$ is finite and they are exhausted by first finding all partitions of n, and then for any fixed $(n_1, \cdots, n_k) \in \mathcal{P}_n$, finding all possible ways to divide the set $[n]$ into different groups of size $n_i, i = 1, \cdots, k$. Observe also that for any given $\{x_1, \cdots, x_n\}$, it can be identified with $[n]$ under the obvious isomorphism. Hence, these two steps also exhaust all possible ways to divide the set $\{x_i\}, 1 \le i \le n$ into different groups. For a generic term among these grouping methods, a point can be labeled as $x_i^{(\ell, n_j)}$ where $n_j, 1 \le j \le k$ comes from the partition of n; once $\{n_j\}$ fixed, ℓ counts the way to divide $[n]$ (hence $\{x_i\}$) into groups with size n_j, and it runs from 1 to $C_n^{n_1, \cdots, n_k}$; i is the *natural* order inside the group.

Now, we calculate $\mathbb{E} \prod_{i=1}^{n} b(x_i)$ conditioning on $N(A)$ where $A = \cup B(x_i)$. We have,

$$\mathbb{E} \prod_{i=1}^{n} b(x_i) = \sum_{m=1}^{\infty} e^{-\nu |A|} \frac{(\nu |A|)^m}{m!} \mathbb{E}\big[\prod_{i=1}^{n} \sum_{j=1}^{m} \phi(x_i - y_j) | N(A) = m \big].$$

The product of sums can be written as

$$\prod_{i=1}^{n} \sum_{j=1}^{m} \phi(x_i - y_j) = \sum_{(n_1, \cdots, n_k) \in \mathcal{P}_n} \sum_{p=1}^{P_m^k} \sum_{\ell=1}^{C_n^{n_1, \cdots, n_k}} \prod_{j=1}^{k} \prod_{i=1}^{n_j} \phi(x_i^{\ell, n_j} - y_j^p). \tag{59}$$

Here $P_m^k, m \ge k$ counts the number of way to choose k different numbers from $[m]$. It corresponds to choosing k different points from the m Poisson points in the set A and assign them to the k groups, and y_j^p represents the choice. The expectation of the product of sums are calculated as follows.

$$\begin{aligned}
&\sum_{m=1}^{\infty} e^{-\nu |A|} \frac{(\nu |A|)^m}{m!} \mathbb{E}\Big[\sum_{(n_1, \cdots, n_k) \in \mathcal{P}_n} \sum_{p=1}^{p_m^k} \sum_{\ell=1}^{C_n^{n_1, \cdots, n_k}} \prod_{j=1}^{k} \prod_{i=1}^{n_j} \phi(x_i^{\ell, n_j} - y_j^p) \big| N(A) = m \Big] \\
&= \sum_{(n_1, \cdots, n_k) \in \mathcal{P}_n} \sum_{\ell=1}^{C_n^{n_1, \cdots, n_k}} \prod_{j=1}^{k} \nu \int \prod_{i=1}^{n_j} \phi(x_i^{\ell, n_j} - z) dz \qquad (60) \\
&= \sum_{(n_1, \cdots, n_k) \in \mathcal{P}_n} \sum_{\ell=1}^{C_n^{n_1, \cdots, n_k}} \prod_{j=1}^{k} T^{n_j}(x_1^{\ell, n_j}, \cdots, x_{n_j}^{\ell, n_j}).
\end{aligned}$$

We verify that this function only depends on the relative distance between the points x_i. This is due to the stationarity of the Poisson point process.

To derive higher order moments of δb, we observe that

$$\prod_{i=1}^{n} \delta b(x_i) = \prod_{i=1}^{n} [b(x_i) - \nu\hat{\phi}(0)] = \sum_{m=0}^{n} (-\nu\hat{\phi}(0))^m \sum_{s=1}^{C_n^{n-m}} \prod_{i=1}^{n-m} \sum_{j=1}^{\infty} \phi(x_i^{s,n-m} - y_j). \tag{61}$$

Here s numbers the ways to choose $n-m$ points from the x_i's and the chosen points are labeled by $s, n-m$ with (relative natural) order i. Then we have the following formula.

Lemma B.1. *Let $\mathscr{G}_n$ be defined as before. For the mean-zero process δb, we have*

$$\mathbb{E}\{\prod_{i=1}^{n} \delta b(x_i)\} = \sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n} \sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}} \prod_{j=1}^{k} T^{n_j}(x_1^{\ell,n_j}, \cdots, x_{n_j}^{\ell,n_j}). \tag{62}$$

The only difference of this formula with that of the higher order moments of b is the change from $\mathcal{P}_n$ to $\mathscr{G}_n$. This is due to the fact that all the T^1 terms, i.e., terms with $\nu\hat{\phi}(0)$, cancel out and we are left with the terms T^{n_j} with $n_j \geq 2$. The proof below follows this observation.

Proof: Combining the formula for $\mathbb{E}\prod\sum\phi(x_i - y_j)$ and the expression of $\prod b(x_i)$, we observe that the moment $\mathbb{E}\prod b(x_i)$ consists of terms of the form:

$$\pm(\nu\hat{\phi}(0))^l \prod_{j=1}^{k} T^{n_j} \tag{63}$$

where $n_j \geq 2$, $k \leq n-l$ and $\sum n_j = n-l$. The terms with $l=0$ are exactly those in (62). We show that all the other terms with $l \geq 1$ vanish. Without loss of generality, we consider the term

$$(\nu\hat{\phi}(0))^l T^{n_1}(x_1, \cdots, x_{n_1}) T^{n_2}(x_{n_1+1}, \cdots, x_{n_2}) \cdots T^{n_k}(x_{n_{k-1}+1}, \cdots, x_{n_k}). \tag{64}$$

This term corresponds to the partition that groups the points with indices between $n_{l-1}+1$ and n_l together for $1 \leq l \leq k$ (with $n_0 = 0$). The last l points contribute the term $(\nu\hat{\phi}(0))^l$.

This term appears in the expectation of the right hand side of (61) with $m = 0, 1, \cdots, l$. It is counted once in the expectation of the term with $m=0$. It is counted C_l^1 times in the expectation of terms with $m=1$. The reason is as follows. For the $m=1$ term, first we choose a point which contributes $(\nu\hat{\phi}(0))$, then we partition the set with $n-1$ points. There are C_l^1 ways to choose this point, and view the other $l-1$ points as coming from the partition of the $n-1$ points. By the same token, this term is counted C_l^2 times in (61) with $m=2$, and so on. It is counted C_l^l times with $m=l$. Note also that for different values of m, the signs of the term alternate. Now recall the combinatoric equality

$$\sum_{k=0}^{l} (-1)^k C_l^k = 0. \tag{65}$$

Hence the term we are considering vanishes. In general, all terms with $l \neq 0$ vanish. This completes the proof. □

References

[1] G. Allaire and G. Bal, *Homogenization of the criticality spectral equation in neutron transport*, M2AN Math. Model. Numer. Anal. **33** (1999), 721–746.

[2] S. R. Arridge, *Optical tomography in medical imaging*, Inverse Problems **15** (1999), R41–R93.

[3] G. Bal, *Kinetics of scalar wave fields in random media*, Wave Motion **43** (2005), 132–157. MR 2186924

[4] ______, *Homogenization in random media and effective medium theory for high frequency waves*, Disc. Cont. Dyn. Syst. B **8** (2007), no.2, 473–492.

[5] ______, *Central limits and homogenization in random media*, Multiscale Model. Simul. **7** (2008), no. 2, 677–702.

[6] ______, *Inverse transport theory and applications*, Inverse Problems **25** (2009), 053001.

[7] G. Bal, J. Garnier, S. Motsch, and V. Perrier, *Random integrals and correctors in homogenization*, Asymptot. Anal. **59** (2008), no. 1-2, 1–26.

[8] G. Bal and W. Jing, *Radiation transfer through random media*, in preparation.

[9] G. Bal and A. Jollivet, *Stability estimates in stationary inverse transport*, Inverse Probl. Imaging **2** (2008), no. 4, 427–454.

[10] G. Bal and K. Ren, *Transport-based imaging in random media*, SIAM J. Appl. Math. **68** (2008), no. 6, 1738–1762.

[11] ______, *Physics-based models for measurement correlations. application to an inverse sturm-liouville problem*, Inverse Problems **25** (2009), 055006.

[12] A. Bensoussan, J.-L. Lions, and G. C. Papanicolaou, *Asymptotic analysis for periodic structures*, Studies in Mathematics and its Applications, 5, North-Holland Publishing Co., Amsterdam-New York, 1978.

[13] ______, *Boundary layers and homogenization of transport processes*, Publ. Res. Inst. Math. Sci. **15** (1979), 53–157.

[14] E. Bolthausen, *On the central limit theorem for stationary mixing random fields*, Ann. Probab. **10** (1982), no. 4, 1047–1050.

[15] A. Bourgeat and A. Piatnitski, *Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second order operator*, Asymptot. Anal. **21** (1999), no. 3-4, 303–315.

[16] S. Chandrasekhar, *Radiative Transfer*, Dover Publications, New York, 1960. MR 0111583

[17] M. Choulli and P. Stefanov, *An inverse boundary value problem for the stationary transport equation*, Osaka J. Math. **36** (1999), no. 1, 87–104.

[18] D. R. Cox and V. Isham, *Point processes*, Chapman & Hall, London, 1980, Monographs on Applied Probability and Statistics. MR MR598033 (82j:60091)

[19] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol.6*, Springer Verlag, Berlin, 1993.

[20] B. Davison and J. B. Sykes, *Neutron transport theory*, Oxford University Press, 1957.

[21] P. Doukhan, *Mixing*, Lecture Notes in Statistics, vol. 85, Springer-Verlag, New York, 1994.

[22] L. Dumas and F. Golse, *Homogenization of transport equations*, SIAM J. Appl. Math. **60** (2000), 1447–1470.

[23] R. Figari, E. Orlandi, and G. Papanicolaou, *Mean field and Gaussian approximation for partial differential equations with random coefficients*, SIAM J. Appl. Math. **42** (1982), 1069–1077.

[24] L. Hörmander, *The analysis of linear partial differential operators I*, Springer-Verlag, New York, 1990.

[25] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of differential operators and integral functionals*, Springer-Verlag, New York, 1994.

[26] C. D. Levermore, G. C. Pomraning, D. L. Sanzo, and J. Wong, *Linear transport theory in a random medium*, J. Math. Phys. **27** (1986), 2526–2536.

[27] A. Marshak and A. B. Davis, *3d radiative transfer in cloudy atmospheres*, Springer, New-York, 2005.

[28] A. Mellet and A. Vasseur, *Homogenization of a nonlinear transport equation*, Asymptot. Anal. **51** (2007), no. 2, 157–166.

[29] M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory*, World Scientific, Singapore, 1997.

[30] L. Ryzhik, G. C. Papanicolaou, and J. B. Keller, *Transport equations for elastic and other waves in random media*, Wave Motion **24** (1996), 327–370. MR 1427483

[31] 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.

[32] P. Stefanov and G. Uhlmann, *An inverse source problem in optical modecular imaging*, Anal. PDE **1** (2008), no. 1, 115–126.

[33] V. V. Yurinskii, *Averaging of symmetric diffusion in a random medium*, Siberian Math. J. **4** (1986), 603–613, English translation of: Sibirsk. Mat. Zh. 27 (1986), no. 4, 167–180 (Russian).