Corrector Theory in Random Homogenization of Partial Differential Equations

Wenjia Jing

Submitted in partial fulfillment of the
Requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011

ABSTRACT

Corrector Theory in Random Homogenization of Partial Differential Equations

Wenjia Jing

We derive systematically a theory for the correctors in random homogenization of partial differential equations with highly oscillatory coefficients, which arise naturally in many areas of natural sciences and engineering. This corrector theory is of great practical importance in many applications when estimating the random fluctuations in the solution is as important as finding its homogenization limit.

This thesis consists of three parts. In the first part, we study some properties of random fields that are useful to control corrector in homogenization of PDE. These random fields mostly have parameters in multi-dimensional Euclidean spaces. In the second part, we derive a corrector theory systematically that works in general for linear partial differential equations, with random coefficients appearing in their zero-order, i.e., non-differential, terms. The derivation is a combination of the studies of random fields and applications of PDE theory. In the third part of this thesis, we derive a framework of analyzing multiscale numerical algorithms that are widely used to approximate homogenization, to test if they succeed in capturing the limiting corrector predicted by the theory.

Contents

ACKNOWLEDGMENT

First of all I would like to thank my advisor, Professor Guillaume Bal, to whom I own more than I can express. It has been a great pleasure to work with Guillaume over the past years. I thank him for giving me the problems investigated in this thesis, for his continuous and insightful guidance during my studies, and for his advices and encouragements.

I am grateful to Professor Rama Cont, Professor Julien Dubedat, Professor Lorenzo Polvani and Professor Michael Weinstein, for serving on my thesis defense committee, and for their inspiring comments and questions during the interrogation.

I wish to thank Professor David Keyes and Professor Chris Wiggins for serving on my oral exam and thesis proposal committees.

Michael and I had many interesting discussions, from which I benefit a lot. He also served on my oral exam and thesis proposal committees. I also thank Rama and Julien for carefully reading and correcting the draft of my thesis.

I am indebted to Professor Kui Ren of UT Austin for his continuous help. I enjoy our conversations. I also thank him for inviting me to Austin; I had a wonderful time there.

Most of the works contained in this thesis are joint with Guillaume. Some of the results in chapter 6 are obtained from collaboration with Professor Josselin Garnier of Université Paris VII, and with Yu Gu. I thank them for their excellent work.

I thank the exceptional curriculums and colloquiums at Columbia University and Courant Institute which offered me solid mathematical training. I thank many professors, Guillaume Bal, Panagiota Daskalopoulos, Martin Hairer, Ioannis Karatzas, Chiu-Chu Liu, D.H. Phong, Ovidiu Savin, Sylvia Serfaty, Jalal Shatah and Michael Weinstein for their wonderful courses.

The outstanding administrative staff in the Department of Applied Physics and Applied Mathematics at Columbia provided me enormous help during my PhD years. Their efficient work saved me a tremendous amount of time so that I could focus on my research. I express my great gratitude to Montserrat Fernandez-Pinkley, Dina Amin, Christina Rohm, Ria Miranda, and Marlene Arbo, for helping me on various issues and for their contributions to the department.

I thank many friends, for making progress together with me, and for making my PhD years at Columbia joyful. Last but not least, I wish to thank my family in China for always loving me, having confidence in me and being supportive.

Chapter 1

Motivations and Overview

This chapter briefly reviews the homogenization theory, introduces the principal concerns of corrector theory, and outlines the contents of this thesis.

1.1 Correctors in Random Homogenization

Partial differential equations with rapidly varying coefficients arise naturally in many important applications, such as composite material sciences, nuclear sciences, porous media equations, and geophysical science. Because the microscopic structure is typically not well known and because the computational costs at the fine structure are prohibitive, it is often necessary to model such heterogeneous structures at the macroscopic level by deriving the homogenized equation, which captures the effective properties of the heterogeneous media.

a. Periodic and random homogenization

We describe the main ideas of homogenization through the following classical example. Let $u(x)$ be the temperature distribution over a complex material which occupies some domain $X \subset \mathbb{R}^d$. Suppose that this material has conductivity tensor $A\left(\frac{x}{\varepsilon}\right)$ and internal heat source $f(x)$, and is immersed in a mixture of ice and water. Consequently, $u_\varepsilon(x)$ solves,

$$-\nabla \cdot A\left(\frac{x}{\varepsilon}\right) \nabla u_\varepsilon(x) = f(x), \quad \text{for } x \in X, \tag{1.1}$$

with Dirichlet boundary condition $u_\varepsilon = 0$ at ∂X. Very often in homogenization theory, the unscaled coefficient $A(x)$ is modeled either as a periodic function, or a stationary and ergodic random field in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Homogenization result is well established in both cases: u_ε converges weakly in $H^1_0(X)$ to u_0, which solves the homogenized equation:

$$-\nabla \cdot A^* \nabla u_0(x) = f(x), \quad \text{for } x \in X, \tag{1.2}$$

with Dirichlet boundary condition, and the effective conductivity A^* given by

$$A^*_{ij} = \mathbb{E}\left(A_{ij}(y) + A_{ik}(y)\frac{\partial \chi^j}{\partial y_k}\right),$$

Here, the vectors $\chi^1, \cdots, \chi^d$ solve the auxiliary equation:

$$-\frac{\partial}{\partial x_i} A_{ij}(x) \frac{\partial \chi^k}{\partial x_j} = \frac{\partial A_{ik}}{\partial x_i}.$$

In periodic homogenization [20, 1, 86], $\mathbb{E}$ here denotes the average over the unit cell of the periodic function; the equation above is posed on this cell with periodic boundary condition, and is called the *cell* problem. In random homogenization [74, 91, 92, 90], $\mathbb{E}$ is the mathematical expectation with respect to the probability measure $\mathbb{P}$, and the equation above is posed over the whole $\mathbb{R}^d$.

Up to reformulation, the periodic setting is a special case of the stationary ergodic setting. Then homogenization theory is essentially the ergodic theory, or the law of large numbers if one wishes. The success of homogenization theory is far beyond the above linear equation; for instance, see [55, 56] for periodic and [31, 30] for random homogenization of fully nonlinear elliptic equations, and [79, 80] for homogenization of Hamilton-Jacobi equations.

b. Corrector theory in random homogenization

The primary interest of this thesis is the corrector in random homogenization, by which we mean the difference between the random solution u_ε and the homogenized solution u_0. Homogenization theory for nonlinear PDE and the study of convergence rate remain active research fields, but are out of the scope of this thesis. We concentrate on linear equations, because even for them corrector theory is not well established. Observe, however, that the dependence of the solution to a PDE on the coefficients of the PDE is usually nonlinear, even when the PDE itself is linear. More precisely, we would like to understand the following issues, assuming the homogenization is known:

1. *What is the convergence rate, say in $L^2(\Omega, L^2(X))$? That is, is there a power γ for which we can show that $\mathbb{E}\|u_\varepsilon - u_0\|^2_{L^2(X)} \le C\varepsilon^{2\gamma}$?*

2. *What is the size of the deterministic corrector $\mathbb{E}\{u_\varepsilon\} - u_0$? What is the size of the (mean-zero) stochastic corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$? Which one is larger?* To make life easier, we may formulate these questions in the weak sense; that is, after integrating the correctors with test functions.

3. (Characterization of the limiting process). *After dividing the random corrector by its amplitude, can we characterize its limiting distribution?* That is, suppose we know the random corrector has size ε^{γ_2} for some $\gamma_2 > 0$, do we have

$$\frac{u_\varepsilon - u_0}{\varepsilon^{\gamma_2}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \text{some probability distribution in certain sense?}$$

Further for the part of the deterministic corrector that is larger than the random part, can we capture their limits as well?

Before answering the questions above, we should ask first:

0. *What assumptions should we put on the random coefficients? Is the stationarity and ergodicity enough?*

The answer is negative. Though stationarity and ergodicity are sufficient for homogenization theory, they are too mild to provide any fine information about the corrector; more information about the random coefficient is indispensable. Compared with homogenization, corrector theory requires more quantitative studies of random fields and PDE. Due to this, corrector theory is less well established. For some of the available results in this setting, we refer the reader to [8, 9, 59, 112]. In this thesis, we only consider *partial differential equations where the random coefficients appear in the zero-order terms, that is, non-differential terms.* In particular, the machinery we develop here does not work for (1.1) in two or higher dimensional spaces.

c. An example: corrector theory for the divergence equation in 1D

In one dimensional spaces, the corrector theory for (1.1) is available. In particular, it verifies the above remark that corrector theory requires finer information of the random structures. In this setting, the equation becomes

$$\begin{cases} -\dfrac{d}{dx}a(\dfrac{x}{\varepsilon},\omega)\dfrac{d}{dx}u_\varepsilon(x,\omega) = f(x),\ x \in (0,1), \\ u_\varepsilon(0,\omega) = u_\varepsilon(1,\omega) = 0. \end{cases} \tag{1.3}$$

Here, the diffusion coefficient $a(\frac{x}{\varepsilon},\omega)$ is modeled as a random process, where ω denotes the realization in an abstract probability space $(\Omega,\mathcal{F},\mathbb{P})$ in which the random process and all limits considered here are constructed. The correlation length ε is much smaller than the length of the domain, which makes the random coefficient highly oscillatory.

The homogenization theory says: When $a(x,\omega)$ is stationary, ergodic, and uniformly elliptic, i.e., $\lambda \le a(x,\omega) \le \Lambda$ for almost every x and ω. Then the solution u_ε converges to the following homogenized equation with deterministic and constant coefficient:

$$\begin{cases} -\dfrac{d}{dx}a^*\dfrac{d}{dx}u_0(x) = f(x),\ x \in (0,1), \\ u_0(0) = u_0(1) = 0. \end{cases} \tag{1.4}$$

The coefficient a^* is the harmonic mean $\left(\mathbb{E}\frac{1}{a(x,\omega)}\right)^{-1}$.

When the corrector $u_\varepsilon - u_0$ is considered, further assumptions on the random coefficient $q(x,\omega) = \frac{1}{a(x,\omega)} - \frac{1}{a^*}$, have to be specified, because they may lead to different conclusions.

Case 1. If $q(x,\omega)$ is mixing with integrable mixing coefficient and hence has short-range correlation (see Section 2.2.1 for the notions), then the corrector theory in [28] shows that

$$\frac{u_\varepsilon - u_0}{\sqrt{\varepsilon}}(x) \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma \int_0^1 (a^*)^2 \frac{\partial G_0}{\partial y}(x,t)u_0(t)dW_t,$$

where $W(t)$ is the standard one dimensional Brownian motion and $G_0(x,y)$ is the Green's function of the homogenized equation. The convergence above is in the sense of distribution in the space of continuous paths.

Case 2. If the random field $q(x,\omega)$ does not de-correlate fast enough, the normalization factor $\sqrt{\varepsilon}$ is no longer correct. In fact, if $q(x,\omega)$ is constructed as a function of Gaussian random field (see Definition 2.27) with covariance function that decays like $\kappa_g|x|^{-\alpha}$ for $\alpha \in (0,1)$, the convergence result above has to be modified [11]:

$$\frac{u_\varepsilon - u_0}{\sqrt{\varepsilon^\alpha}}(x) \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sqrt{\frac{\kappa}{H(2H-1)}} \int_0^1 (a^*)^2 \frac{\partial G_0}{\partial y}(x,t)u_0(t)dW_t^H,$$

where $W^H(t)$ is a fractional Brownian motion with Hurst index $H = 1-\frac{\alpha}{2}$, and the constant κ is related to κ_g; for the details, see Section 2.5 and Chapter 7.

1.2 Motivations for Corrector Theory

Corrector theory is of vast practical importance, as in parameter estimation, uncertainty quantification and algorithm testing, which we will address in Chapter 7. Here, we discuss in detail its application to PDE-based inverse problems.

Bayesian formulation of inverse problems. In a typical inverse problem, one has data Y obtained from some unknown input X through the forward relation:

$$Y = F(X) + E, \tag{1.5}$$

where F is the forward model that maps the input to the output, and E is the error in the data, which may accounts for modeling or measurement errors. The goal of inverse problem is to reconstruct X given $Y = y$. Since noise E is inevitable and the inverse of F is usually unbounded, the reconstruction X is very often obtained by minimizing the discrepancy, i.e., $F(x) - y$ measured in some proper norm with some type of *regularization*, among trials of x in some proper space. In the Bayesian approach to regularization, this boils down to the following. View X, Y, E as realizations in some probability space. From experience or other *a priori* information, one has beforehand a prior distribution $\pi_{\rm pr}(x)$ of X. The probability density of Y given $X = x$ is then called the *likelihood* $\pi(y|x)$. Suppose that we know the distribution of the noise E is given by $\pi_{\rm noise}(e)$ and it is independent with X, we deduce from the relation (1.5) that

$$\pi(y|x) = \pi_{\rm noise}(y - F(x)).$$

Consequently, the probability density of X given $Y = y$ is provided by the Bayes' formula

$$\pi(x|y) \propto \pi(y|x)\pi(x) = \pi_{\rm noise}(y - F(x))\pi_{\rm prior}(x).$$

Therefore, given the distributions π_{noise} and π_{prior}, we can maximize the distribution $\pi(x|Y = y)$ to get a reasonable reconstruction of X.

Application of corrector theory to PDE-based inverse problems. Many inverse problems in application are based on PDE; for instance, the Computed Tomography (CT) used in medical imaging is based on the transport equation, which describes propagation of X-ray in body tissues. In these settings, the unknown input consists of parameters of some PDE; the output is the solution to the PDE or functionals of it, and the above Bayesian formulation should be applied in some functional space setting.

There is one more issue to address. Due to the smoothing property of the forward map F which averages out high frequency modes of the input, only low frequency components of X can be stably reconstructed. In many cases, however, the high frequency parts still significantly affect the data. In such a situation, let q_0 and q_ε be the low and high frequency components of the input, which accounts for some coefficient of the governing PDE. Then the output, the solution u_ε, can be viewed as corrupted data:

$$u_\varepsilon = F(q_0) + E, \tag{1.6}$$

where $F(q_0) = u_0$ is the forward map for the PDE with low frequency coefficient q_0 only. Then E is the corrector $u_\varepsilon - u_0$, and the corrector theory for the homogenization of u_ε to u_0 provides a precise statistical model for the error term E above. Now, with a good prior model for q_0, one can apply the Bayesian formulation to approximate the smooth part of the unknown parameter. In summary, corrector theory is very useful in PDE-based inverse problems, because it provides accurate model for the effect of high frequency component on the low frequency part of the unknown parameter.

1.3 Overview

This thesis is structured as follows.

a. Random Fields and Oscillatory Integrals

Chapter 2 is a detailed study of the random fields that will be used in this thesis. As remarked before, the corrector theory, which is very often of the central limit theorem type, requires fine knowledge of the random fields beyond stationarity and ergodicity. In this chapter, we show that different decorrelation rates lead to different limiting distributions of oscillatory integrals involving the random fields. We also provide formulas for high order moments of random fields under certain conditions. In later chapters of the thesis, these formulas are useful in controlling nonlinear functionals of the random field; such functionals are almost always present even for linear PDE. Some explicit random models, such as superposition of Poisson bumps and function of Gaussian random fields, are studied in detail.

b. Corrector Theory in Random Homogenization of Equations

Here, we develop corrector theory for several partial differential (or integro-differential) equations with random coefficients, where the randomness appears in zero-order terms, i.e., not in the differential terms.

Chapter 3 reviews the solution operator of the stationary linear transport equation, as a preparation for the next chapter. We show that the norm of the solution operator as a transform in the L^2 space can be bounded independent of the structure and L^∞ norm of the constitutive parameter. This allows random perturbation of these parameters in Chapter 4. The Schwartz kernel estimate of the solution operator also makes analysis of the corrector in Chapter 4 much easier.

Chapter 4 investigates linear transport equations with random constitutive parameters. The homogenization of such equations was known to be obtained by averaging; we recover this result and capture explicit convergence rates using a random model based on Poisson point process. We then study the limiting distribution of the corrector in this random homogenization. In a weak sense, this corrector converges in distribution to some Gaussian process whose covariance structure can be explicitly characterized.

Chapter 5 is again a preparation for Chapter 6. It reviews some of the main properties of the steady-state diffusion equation with absorbing potential, the fractional Laplacian equation with absorbing potential, and introduces a pseudo-differential equation resulted from a Robin problem. These equations share the following properties: The solution operator is a transform on L^2 and its operator norm can be bounded independent of the non-negative potential; the Green's function $G(x, y)$ is of order $|x - y|^{-d+\beta}$ near the diagonal for some $\beta \in (0, d)$. These properties define a family of PDE for which the corrector theory developed in this thesis works.

Chapter 6 investigates the corrector theory in random homogenization for the family of PDE mentioned above. Under some conditions, we explicitly characterize the limiting Gaussian distributions of the random correctors. We emphasize two factors that largely determine the main features of the limiting distributions: The singularity of the Green's function near the origin, i.e., the factor β defined above, and the decorrelation rate of the random coefficient in the PDE. The fluctuation in the corrector is larger when the Green's function is more singular and when the random coefficient is longer correlated.

c. Corrector Tests for Multiscale Numerical Algorithms

Chapter 7 proposes a benchmark to test multiscale numerical algorithms that have been widely used in scientific computing to capture the homogenization; the goal is to see if these methods manage to obtain the limiting distribution suggested by the corrector theory stated in Section 1.1. Finite element method based multiscale methods are considered, and two algorithms are analyzed in detail. Our analysis suggests that though partial sampling of a PDE with random coefficient may capture the homogenization, as long as corrector is considered, it is sensitive to the decorrelation rate of the random coefficient.

1.4 Notes

Section 1.1 There are quite a few books and monographs that cover homogenization of PDE. We recommend the book by Bensoussan, Lions and Papanicolaou [20] for periodic homogenization; a more extensive book by Jikov, Kozlov and Oleinik [70] covers random homogenization also. The book by Pavliotis and Stuart [95] surveys a broad range of methods for multiscale PDE and other systems and contains an easy-to-access review of homogenization.

Section 1.2 We recommend the book by Kaipio and Somersalo [71] as a primer on the theory and computational implementations of finite dimensional inverse problems. A thorough review of the Bayesian formulation in functional spaces, with applications to PDE-based inverse problems, is given by Stuart [105]; Theoretical formulations of applying corrector theory to PDE-based inverse problems can be found in the papers by Bal and Ren [18], Nolen and Papanicolaou [88], which also include numerical experiments.

Section 1.3 Before reading the rest of this thesis, the reader should read the short list of notations located after the last chapter, though they are mostly the standard ones. We assume that the reader are familiar with elementary theory of partial differential equations, at the level of the first six chapters of Evans [57], basic real and functional analysis, at the level of Lieb and Loss [78], Hunter [68], and Reed and Simon [99]. For probability and stochastic analysis, we assume the reader is familiar with the basic theories at the level of Breiman [29] and Chung [38], and has working knowledge of stochastic integrals at the level of Kuo [75].

Chapter 2

Random Fields

This chapter studies random field models that will be used in later chapters. We introduce notations that are widely used throughout this thesis, and characterize limiting distributions of oscillatory integrals involving random fields. Specific examples of random fields, estimates and explicit formulas for high-order moments are also provided.

2.1 Random Fields

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with sample space Ω, the σ-algebra $\mathcal{F}$ of measurable sets or *events*, and a probability measure $\mathbb{P}$ for elements in $\mathcal{F}$.

A random field is nothing but a collection of random variables $\{X(t,\omega) \mid t \in T\}$ for some T consists of points in $\mathbb{R}^d$. If T is $\mathbb{N}$, the random field is just a random sequence $\{X_n(\omega)\}$; if $T = \mathbb{R}$, the random field is often written as $X_t(\omega)$ and bears the name random process. These two cases are the most discussed in standard textbooks, and the parameter t naturally plays the role of "time". In this thesis and in the context of corrector theory for random homogenization of PDE, $X(t,\omega)$ models some parameter in a PDE, and T should model the physical domain where the PDE is posed. Consequently, $t \in T$ should play the role of "position" rather than time.

As we have seen in Chapter 1, to model the highly heterogeneous properties of the background media, the random models for PDE coefficients of have the form $A\left(\frac{x}{\varepsilon}, \omega\right)$. Therefore, though the parameter x in the PDE may vary on some compact set X, the parameter in the random model itself should allow points in $\varepsilon^{-1}X$. As ε is sent to zero eventually, the parameters for the random field exhaust $\mathbb{R}^d$. For this reason, we assume $T = \mathbb{R}^d$.

Definition 2.1 (Random Field)**.** A random field on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of random variables parametrized by $\mathbb{R}^d$; that is, $\{q(x,\omega) \mid x \in \mathbb{R}^d\}$. When $d = 1$, the name random process is more standard.

Remark 2.2. Note that there are other ways to model the scaling in $A\left(\frac{x}{\varepsilon}\right)$. For instance, if we are interested in the limiting distribution of a family of random processes, we may simultaneously change the probability space $(\Omega^\varepsilon, \mathcal{F}^\varepsilon, \mathbb{P}^\varepsilon)$ as ε varies, since convergence in distribution (law) does not require the family to be defined on the same probability space. In Section 2.3, we shall see such an example based on Poisson point process. Also, $\mathbb{R}^d$ may be replaced by some symmetric space that is easy to be rescaled, say S^{d-1}. We do not go further in these directions. □

Definition 2.3 (Stationarity)**.** A random field $q(x,\omega)$ is called *stationary* if for any $n \in \mathbb{N}$, and any n-tuple $(x_1, \cdots, x_n)$, $x_i \in \mathbb{R}^d$, and any $z \in \mathbb{R}^d$, the following holds:

$$(q(x_1,\omega), \cdots, q(x_n,\omega)) \stackrel{\text{law}}{=} (q(x_1 + z,\omega), \cdots, q(x_n + z,\omega)), \tag{2.1}$$

where $\stackrel{\text{law}}{=}$ denotes equality in law.

Suppose the coefficient $q\left(\frac{x}{\varepsilon}, \omega\right)$ of a PDE is constructed from a stationary field $q(x,\omega)$. Then though for each realization the background media is spatially heterogeneous, the statistics of it is still homogeneous.

For a stationary random field $q(x,\omega)$, there exists a natural group of measure-preserving transformations $\tau_x : \Omega \to \Omega$, so that $\mathbb{P}(\tau_x^{-1}A) = \mathbb{P}(A)$ for any $A \in \mathcal{F}$.

Definition 2.4 (Ergodicity)**.** A measure-preserving transform τ_x is said to be *ergodic* if all invariant events under the map τ_x are trivial. That is to say,

$$\tau_x A = A \quad \Longrightarrow \quad \mathbb{P}(A) \in \{0,1\}. \tag{2.2}$$

A stationary random field $q(x,\omega)$ is said to be *ergodic* if the group of measure-preserving transformations $\{\tau_x | x \in \mathbb{R}^d\}$ is ergodic. Ergodicity is not easy to check. Sufficient conditions include the strong mixing property; see Definition 2.10.

Example 2.5. Consider a random field $A(x,\omega)$, $x \in \mathbb{R}^d$, consists of independent identically distributed (i.i.d.) random variables. Then $A(x,\omega)$ is stationary and ergodic.

Example 2.6 (Gaussian random field)**.** A random field $\{\dot{W}(x,\omega)\}$ with parameter space $\mathbb{R}^d$ is said to be *Gaussian* if for any $J \in \mathbb{N}$, and any $(x_1, \cdots, x_J)$ where $x_j \in \mathbb{R}^d$, the random vector $(\dot{W}(x_1), \cdots, \dot{W}(x_J))$ is an $\mathbb{R}^J$-valued Gaussian random vector. As for any random field, we can associate a mean field $\mathbb{E}\dot{W}(x)$ and the covariance function $R(x,y) = \mathbb{E}\dot{W}(x)\dot{W}(y) - \mathbb{E}\dot{W}(x)\mathbb{E}\dot{W}(y)$ to a Gaussian random field. More importantly, these two factors determine a Gaussian random field. This is the characterization enjoyed only by Gaussian random field.

Example 2.7 (Canonical representation)**.** An important setting, which in fact we should always keep in mind, is the following: Take Ω to be certain subset of Lebesgue measurable functions on $\mathbb{R}^d$ so that the PDE is well-posed. The value of $\omega \in \Omega$ at $x \in \mathbb{R}^d$ is defined almost everywhere and is denoted by $\omega(x,\omega)$. Expressed in a different way, Ω is the set of all admissible coefficients.

Let $\mathcal{F}$ be the σ-algebra generated by cylinder sets with base points that have rational coordinates in $\mathbb{R}^d$ and range sets which are product of open intervals in $\mathbb{R}$. That is to say,

$\mathcal{F}$ is generated by sets of the form:

$$\{\omega(x) \text{ admissible coefficients, } \mid \omega(x_1) \in I_1, \cdots, \omega(x_J) \in I_J\},$$

for some $J \in \mathbb{N}$, points $x_1, \cdots, x_J$ in $\mathbb{Q}^d$, and open sets I_j with rational center and rational length. In this way, $\mathcal{F}$ is countably generated.

The probability measure $\mathbb{P}$ is defined on the measurable space $(\Omega, \mathcal{F})$ so that it is invariant with respect to the translation group $\tau_x : \Omega \to \Omega$ defined by

$$\tau_x \omega\ (y) = \omega(y - x), \quad x, y \in \mathbb{R}^d.$$

We assume that the group τ_x is ergodic.

For a stationary random field $q(x, \omega)$, the corresponding mean field $\mathbb{E}\{q(x)\}$ is a constant. Here and in the following, we always denote by $\mathbb{E}$ the mathematical expectation with respect to the probability measure $\mathbb{P}$. That is to say, $\mathbb{E}\{f(x,\omega)\} = \int_\Omega f(x,\omega) d\mathbb{P}(\omega)$ for any random field $f(x,\omega)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. To simplify notations, we will make the dependency on ω implicit henceforth. We may consider only stationary random fields that are mean-zero, without loss of generality.

Definition 2.8 (Correlation function)**.** The *correlation function* of a mean-zero stationary random field $q(x)$ is defined to be:

$$R(x) := \mathbb{E}\{q(y)q(y + x)\}. \tag{2.3}$$

Note the above definition is independent of y since q is stationary. In the literature, R is also frequently called the *autocorrelation* function , or the *covariance* function especially when q is Gaussian. We remark that the $R(x)$ defined above is essentially the standard correlation between $q(y)$ and $q(y+x)$, that is,

$$\frac{\mathbb{E}\{q(y)q(y+x)\} - \mathbb{E}\{q(y)\}\mathbb{E}\{q(y+x)\}}{\sqrt{\text{Var }\{q(y)\} \;\; \text{Var }\{q(y+x)\}}},$$

up to the denominator, which is a constant for stationary random fields. Some immediate properties of $R(x)$ are worth recording.

Proposition 2.9. *Let $R(x)$ be defined as above. We have*

(1) *R is* symmetric, *i.e., $R(x) = R(-x)$.*

(2) *R is* bounded *if $q(0) \in L^2(\Omega)$.*

(3) *R is* semi-positive definite. *That is to say, for any $J \in \mathbb{N}$, J-tuple $(x_1, \cdots, x_J)$ where $x_j \in \mathbb{R}^d$, the matrix formed by $\{R(x_i - x_j)\}_{i,j=1}^J$ is a semi-positive definite matrix. In other words, for any $(\xi_1, \cdots, \xi_J)$ where $\xi_j \in \mathbb{C}$, we have*

$$\sum_{i=1}^{J} \sum_{j=1}^{J} \overline{\xi}_i R(x_i - x_j)\xi_j \geq 0. \tag{2.4}$$

(4) *As a consequence of* (3), $\int_{\mathbb{R}^d} R(x)dx \geq 0$.

Proof. The first two items are trivial. The third one is obvious once we observe that the left hand side of (2.4) is nothing but

$$\mathbb{E}\left|\sum_{i=1}^{J} \xi_i q(x_i)\right|^2.$$

The fourth item is an immediate consequence of a nontrivial result in Fourier analysis, namely the Bochner's theorem, which asserts that the Fourier transforms of semi-positive definite functions, hence that of R, are exactly positive measures. The integral in item four is nothing but the value of the Fourier transform of R evaluated at zero. □

Very often we abuse notations and do not distinguish semi-positive definite from positive definite, as long as it does not cause trouble. If $\widehat{R}(0)$ is a positive and finite number, we define

$$\sigma^2 := \int_{\mathbb{R}^d} R(y)dy. \tag{2.5}$$

We remark that there exists random field such that the above integral is zero. For instance, take a white-noise field and color it by covariance function which integrates to zero.

We would like to say a random field $q(x)$ has *short-range* correlation if its correlation function $R(x)$ is integrable over $\mathbb{R}^d$, and *long-range* correlation otherwise. In the next couple of sections, we shall investigate the following problem: What is the limiting distribution of the following oscillatory integral over some domain $X \subset \mathbb{R}^d$,

$$\int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx, \tag{2.6}$$

for some nice function f? As it turns out, the answer can be quite different for short-range and long-range random fields. However, we need more assumptions in addition to the integrability of R to give precise answers to the question. The details are provided in the next section.

Let us conclude this introduction by addressing the following perspective of random fields constructed above. Namely, we can view a random process $q(x, \omega)$ as a functional-space-valued random variable. As in Example 2.7, suppose that the coefficient $q(x, \omega)$ of some PDE belongs to some Hilbert space $\mathcal{H}$. We can then view $q(\omega)$ as an $\mathcal{H}$-valued random variable, and view $q(x, \omega)$ as the $\mathbb{R}$-valued random variable resulted from applying the linear functional x on it, by $x(q) := q(x)$. The correlation function R then maps $\mathcal{H}^* \times \mathcal{H}^*$ to $\mathbb{R}$. The advantage of this point of view is: We can consider a sequence of random fields $F_\varepsilon(x, \omega)$ as paths in certain Hilbert space, and investigate the weak convergence of the probability measures on that Hilbert space induced by these random fields.

2.2 Strong Mixing Random Fields

For the moment, let $O_\varepsilon[f]$ denote the oscillatory integral in (2.6). We observe that this random variable is mean zero. To determine its limiting distribution as ε goes to zero, we first calculate the order of its variance.

$$\mathrm{Var}\ O_\varepsilon[f] = \mathbb{E}\{O_\varepsilon^2[f]\} = \int_{X^2} R\left(\frac{x-y}{\varepsilon}\right) f(x)f(y)dxdy.$$

Suppose that $f \in L^2(X)$; we apply Proposition 2.39 to the above integral and deduce that the variance of $O_\varepsilon[f]$ is of order ε^d. Therefore, to determine the limiting distribution of $O_\varepsilon[f]$, we divide it by $\varepsilon^{d/2}$, and investigate the following quantity which is on a finer scale:

$$I_\varepsilon[f] := \frac{1}{\sqrt{\varepsilon^d}} \int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx. \tag{2.7}$$

Characterizing the limiting distribution of this integral is the main goal of this section.

2.2.1 Mixing of random fields

We have introduced the correlation function $R(x)$ of a stationary mean-zero random field $q(x)$, which quantifies the correlation of the field at two points that are x apart. To have a central limit theorem type of result, we need stronger control of the correlation of the field.

Suppose $q(x)$ is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Given a Borel set $A \subset \mathbb{R}^d$, we denote by $\mathcal{F}_A$ the sub-σ-algebra generated by $\{q(x) \mid x \in A\}$, that is, all the measurable sets regarding information of the random field restricted to A. Further, we denote by $\mathcal{L}(\mathcal{F}_A)$ the set of square integrable random variables that are measurable with respect to $\mathcal{F}_A$.

Definition 2.10 (α-mixing)**.** A stationary random field $q(x)$ is *α-mixing* with mixing coefficient $\alpha(r)$ if there exists some function $\alpha : \mathbb{R}_+ \to [0,1]$ and $\lim_{r\to\infty} \alpha(r) = 0$, such that for any Borel sets $A, B \subset \mathbb{R}^d$, we have

$$|\mathbb{P}(S \bigcap T) - \mathbb{P}(S)\mathbb{P}(T)| \leq \alpha(\text{dist}(A, B)), \quad \forall S \in \mathcal{F}_A, T \in \mathcal{F}_B. \tag{2.8}$$

Here $\text{dist}(A, B)$ is the distance between the two sets A and B.

Definition 2.11 (ρ-mixing)**.** A stationary random field $q(x)$ is *ρ-mixing* with mixing coefficient $\rho(r)$ if there exists some function $\rho : \mathbb{R}_+ \to [0, 1]$ and $\lim_{r\to\infty} \rho(r) = 0$, such that for any Borel sets $A, B \subset \mathbb{R}^d$, we have

$$|\text{Corr}(\xi, \eta)| = \left| \frac{\mathbb{E}\ \xi\eta - \mathbb{E}\xi\ \mathbb{E}\eta}{\sqrt{\text{Var}\ (\xi)\ \text{Var}\ (\eta)}} \right| \leq \rho(\text{dist}(A, B)), \quad \forall \xi \in \mathcal{L}(\mathcal{F}_A), \eta \in \mathcal{L}(\mathcal{F}_B). \tag{2.9}$$

Here $\text{dist}(A, B)$ is the distance between the two sets A and B.

The above definitions of mixing coefficients are just two examples of the various mixing coefficients used in the statistics literature. In general, mixing coefficients quantify decorrelation of the information about the random fields over separated regions in terms of the distance between these regions.

Suppose the random field $q(x)$ is i.i.d, then of course it is mixing (in fact with $\alpha(r) = \rho(r) = 0$ for $r > 0$). Therefore, mixing can be thought as a measure of weak dependency, a generalization of the concept of independency.

Recall that the most classical central limit theorem is for an i.i.d. sequence of random variables $X_1, X_2, \cdots$ with $\mathbb{E}X_1 = 0$ and $\mathbb{E}X_1^2 = \sigma^2 < \infty$. It says that the rescaled sample average $S_n/\sqrt{\sigma^2 n} = \frac{1}{\sigma\sqrt{n}}(X_1 + \cdots + X_n)$ converges to the normal distribution $\mathcal{N}(0, 1)$. Since the random coefficients of PDE that we consider in this thesis are in general not independent at different points, and is usually parametrized in dimension two or higher, we need a more general central limit theorem which accounts for weakly dependent random sequence with multi-dimensional indices. We record such a result below.

Let $X_z, z \in \mathbb{Z}^d$ be a random field. We can adapt the definitions of mixing coefficients above to the current setting, using $d(z_1, z_2) = \max_{1\le i\le d} |z_1(i) - z_2(i)|$ to measure the distance between two points $z_1, z_2 \in \mathbb{Z}^d$. Here, $z(i)$ is the ith coordinate of z. We have the following central limit theorem for weakly dependent random field.

Theorem 2.12 (Bolthausen). *Let $X_z, z \in \mathbb{Z}^d$ be a stationary random field with mean zero. Suppose X_z is α-mixing with mixing coefficient $\alpha(m)$. Further, assume that there exists some $\delta > 0$ such that*

$$\mathbb{E}|X_z|^{2+\delta} < \infty, \quad \text{and} \quad \sum_{m=1}^{\infty} m^{d-1}\alpha(m)^{\delta/(2+\delta)} < \infty.$$

Then $\sum_{z\in\mathbb{Z}^d} |\mathrm{cov}(X_0, X_z)| < \infty$ and if $\sigma^2 = \sum_{z\in\mathbb{Z}^d} \mathrm{cov}(X_0, X_z) > 0$, then the law of $S_n/\sigma|\Lambda_n|^{1/2}$ converges to the standard normal one. Here, S_n is the sum over $X_z, z \in \Lambda_n$, where $\{\Lambda_n\}$ is a sequence of subsets of $\mathbb{Z}^d$ which increases to $\mathbb{Z}^d$ and satisfies that $\lim_{n\to\infty} |\partial\Lambda_n|/|\Lambda_n| = 0$. The cardinality of the set Λ is denoted by $|\Lambda|$.

Remark 2.13. In the paper by Bolthausen [27], this theorem was proved for even weaker conditions. In fact, he defined α_{kl}-mixing coefficients where the sets A, B in (2.8) are only taken over $|A| \le k$ and $|B| \le l$. Then in his theorem, only $\alpha_{2,\infty}$ is needed.

Suppose that

$$\alpha(m) \sim O(m^{-d-\delta'}) \text{ for some } \delta' > 0, \tag{2.10}$$

then there exists some δ so that $\sum_m m^{d-1}\alpha(m)^{\delta/(2+\delta)}$ is finite. Suppose further that X_z has sufficient large moment; then the theorem can be applied.

Remark 2.14. The above theorem can be stated using the ρ-mixing coefficients as well. In fact, it is not very difficult to see that the α-mixing coefficient is actually weaker than the ρ-mixing coefficients. Given a ρ-mixing random field, one can choose $\alpha(r)$ so that

$$\alpha(r) \le 4\rho(r).$$

This is best seen by considering $\mathbf{1}_A$ and $\mathbf{1}_B$, the indicator functions of A and B, in (2.8).

2.2.2 A central limit theorem for the oscillatory integral

The central limit theorem for discrete random fields in the previous section can be employed to find the limiting distribution of the oscillatory integral (2.7).

Theorem 2.15 (Oscillatory Integral in Short-range Media)**.** *Let* $I_\varepsilon[f]$ *be as in* (2.7)*. Let* $f \in L^2(X)$*,* $q(x,\omega)$ *be stationary, mean-zero,* ρ*-mixing with mixing coefficient* $\rho(r)$ *of order* $O(r^{-d-\delta})$ *for large* r *for some positive* δ*. Suppose also that* σ *defined in* (2.5) *is positive. Assume also that the boundary of* X *is sufficiently smooth. Then,*

$$I_\varepsilon[f] = \frac{1}{\varepsilon^{\frac{d}{2}}} \int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma \int_X f(x)dW(x). \tag{2.11}$$

Here, $W(x)$ *is the standard real-valued multi-parameter (*$\mathbb{R}^d$*-parameter) Wiener process, and* dW *therefore is the standard White-noise measure. The convergence is understood as convergence of random variables in distribution.*

Note that $f \in L^2$ is required for the limiting Gaussian variable to have finite variance. This theorem was proved by Bal in [9] for the case of $f \in C(\overline{X})$ and his proof can be easily generalized to the L^2 setting. It follows quite easily from Theorem 2.12, but it is central to the corrector theory that we will develop in later chapters. So we record its proof here in detail.

Proof. 1. We prove first that it suffices to consider $f \in C(\overline{X})$. Indeed, for a general $f \in L^2(X)$, we can find a sequence $f_n \in C(\overline{X})$ such that $\|f_n - f\|_{L^2} \longrightarrow 0$. Then it follows that $I_\varepsilon[f_n] \to I_\varepsilon[f]$ in $L^2(\Omega)$ as $n \to \infty$ uniformly in ε. To see this, we calculate

$$\mathbb{E}|I_\varepsilon[f_n] - I_\varepsilon[f]|^2 = \frac{1}{\varepsilon^d} \int_{X^2} R\left(\frac{x-y}{\varepsilon}\right)(f_n - f)(x)(f_n - f)(y)dxdy. \tag{2.12}$$

Apply Proposition 2.39 to this integral; it is bounded by $\|R\|_{L^1}\|f_n - f\|_{L^2}^2$. As a result,

$I_\varepsilon[f_n]$ converges to $I_\varepsilon[f]$ in $L^2(\Omega)$.

Now suppose that the theorem holds for continuous functions. Consider an arbitrary $f \in L^2(X)$ and fix an arbitrary real number ξ. For any ϵ, there exists a continuous function $\tilde{f}$ such that $\|f - \tilde{f}\|_{L^2}^2 \le \epsilon/(3|\xi|^2\|R\|_{L^1})$. For this $\tilde{f}$, there exists an $\delta(\epsilon)$ such that

$$|\mathbb{E}\; e^{i\xi I_\varepsilon[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_0[\tilde{f}]}| < \frac{\epsilon}{3}, \quad \forall \varepsilon < \delta(\epsilon).$$

Here and below, we denote by $I_0[g]$ the right hand side of (2.11) with integrand $g \in L^2$. Consequently, we have that

$$|\mathbb{E}\; e^{i\xi I_\varepsilon[f]} - \mathbb{E}\; e^{i\xi I_0[f]}| \le |\mathbb{E}\; e^{i\xi I_\varepsilon[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_0[\tilde{f}]}| + |\mathbb{E}\; e^{i\xi I_0[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_0[f]}| + |\mathbb{E}\; e^{i\xi I_\varepsilon[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_\varepsilon[f]}|.$$

By the choice of ε, the first term is less than $\epsilon/3$. Meanwhile, the third term is bounded by

$$|\mathbb{E}\; e^{i\xi I_\varepsilon[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_\varepsilon[f]}| \le \mathbb{E}|\xi(I_\varepsilon[f] - I_\varepsilon[\tilde{f}])| \le |\xi|^2\mathbb{E}|I_\varepsilon[f] - I_\varepsilon[\tilde{f}]|^2.$$

By the choice of $\tilde{f}$, this term is bounded by $\epsilon/3$. For the middle term, we have

$$|\mathbb{E}\; e^{i\xi I_0[\tilde{f}]} - \mathbb{E}\; e^{i\xi I_0[f]}| \le |\xi|^2\mathbb{E}|I_0[f] - I_0[\tilde{f}]|^2 = |\xi|^2\|f - \tilde{f}\|_{L^2}^2.$$

The last equality is due to the Itô isometry. By our choice of $\tilde{f}$, this term is bounded by $\epsilon/3$ as well. In summary, for any $\xi \in \mathbb{R}$, we have shown that $\mathbb{E}\; e^{i\xi I_\varepsilon[f]}$ converges to $\mathbb{E}\; e^{i\xi I_0[f]}$. That is, $I_\varepsilon[f]$ converges in distribution to $I_0[f]$, completing the proof of the theorem.

2. Starting in this step, we prove the theorem for continuous f. In particular, f is uniformly bounded. Let $\{Q_j^0, j \in \mathbb{Z}^d\}$ denote the cubes of unit size that tiles up $\mathbb{R}^d$. Let $h > 0$ be a small number, and let $\{Q_j, j \in \mathbb{Z}^d\}$ be the scaled cubes $Q_j = hQ_j^0$. The total number of cubes that overlap with X are of order h^{-d}. We divide them into two categories, those that contain part of the boundary ∂X and those that are in the interior of X. Since

the boundary ∂X is smooth, the number of cubes that are in the first category is of order h^{-d+1}. Suppose f_h is a piece-wise constant function with constant value on each cube of the second category and is zero on cubes of the first category. Then, we have

$$\mathbb{E}|I_\varepsilon[f] - I_\varepsilon[f_h]|^2 \leq C\|f - f_h\|^2_{L^\infty},$$

for any ε for some constant C that is independent of h and ε. Therefore, upon reducing to another approximating sequence, we can assume f in (2.11) is in fact such a piece-wise function. That is,

$$f(x) = \sum_{j \in \mathbb{Z}^d_*} f_j \mathbf{1}_{Q_j}(x).$$

Here $\mathbb{Z}^d_*$ contains indices such that Q_j belongs to the second category.

3. In this step, we assume f has the form of the previous formula. In particular, define random variables $\{I^j_\varepsilon, j \in \mathbb{Z}^d_*\}$ by

$$I^j_\varepsilon := \frac{1}{\varepsilon^{\frac{d}{2}}} \int_{Q_j} q\left(\frac{x}{\varepsilon}\right) f(x) dx = f_j \frac{1}{\varepsilon^{\frac{d}{2}}} \int_{Q_j} q\left(\frac{x}{\varepsilon}\right) dx.$$

The task of this step is to show that these random variables are *asymptotically independent*. That is, for any $\boldsymbol{\xi} := \{\xi_j \in \mathbb{R} \mid j \in \mathbb{Z}^d_*\}$, we have

$$\mathcal{E}(\boldsymbol{\xi}) := \left| \mathbb{E}\, e^{i \sum_{j \in \mathbb{Z}^d_*} \xi_j I^j_\varepsilon} - \prod_{j \in \mathbb{Z}^d_*} \mathbb{E} e^{i \xi_j I^j_\varepsilon} \right| \longrightarrow 0, \quad \text{as } \varepsilon \to 0. \tag{2.13}$$

Let η be a number between zero and $\frac{h}{2}$. Denote by $Q_{j\eta}$ the cube which shares the center of Q_j but have sides of length η. Define

$$I^j_{\varepsilon\eta} := f_j \frac{1}{\varepsilon^{\frac{d}{2}}} \int_{Q_{j\eta}} q\left(\frac{x}{\varepsilon}\right) dx, \quad P^j_{\varepsilon\eta} = I^j_\varepsilon - I^j_{\varepsilon\eta}.$$

Let us adopt an arbitrary numbering of the set $\mathbb{Z}^d_*$. One of the cubes is then Q_1, and accordingly there are I_ε^1, $I_{\varepsilon\eta}^1$ and $P_{\varepsilon\eta}^1$. Now we write

$$\mathbb{E}\ e^{i\sum_j \xi_j I_\varepsilon^j} = \mathbb{E}\{(e^{i\xi_1 P_{\varepsilon\eta}^1}-1)e^{i\xi_1 I_{\varepsilon\eta}^1 + i\sum_{j\neq 1}\xi_j I_\varepsilon^j}\} + \mathbb{E}\{e^{i\xi_1 I_{\varepsilon\eta}^1 + i\sum_{j\neq 2}\xi_j I_\varepsilon^j}\}.$$

Using the ρ-mixing condition (2.9), we find that

$$\left|\mathbb{E}\{e^{i\xi_1 I_{\varepsilon\eta}^1 + i\sum_{j\neq 1}\xi_j I_\varepsilon^j}\} - \mathbb{E}\{e^{i\xi_1 I_{\varepsilon\eta}^1}\}\mathbb{E}\{e^{i\sum_{j\neq 1}\xi_j I_\varepsilon^j}\}\right| \leq C\rho\left(\frac{\eta}{\varepsilon}\right).$$

Consequently, we have

$$\left|\mathbb{E}\ e^{i\sum_{j\in\mathbb{Z}_*^d}\xi_j I_\varepsilon^j} - \mathbb{E}\{e^{i\xi_1 I_\varepsilon^1}\}\prod_{j\neq 1}\mathbb{E}e^{i\xi_j I_\varepsilon^j}\right| \leq C\rho\left(\frac{\eta}{\varepsilon}\right) + \left|\mathbb{E}\{(e^{i\xi_1 P_{\varepsilon\eta}^1}-1)e^{i\xi_1 I_{\varepsilon\eta}^1 + i\sum_{j\neq 1}\xi_j I_\varepsilon^j}\}\right|$$
$$+\left|\mathbb{E}\{(e^{i\xi_1 P_{\varepsilon\eta}^1}-1)e^{i\xi_1 I_{\varepsilon\eta}^1}\}\prod_{j\neq 1}\mathbb{E}e^{i\xi_j I_\varepsilon^j}\right|.$$

For the last two terms, we use the fact that the exponential function is bounded uniformly on the unit circle of the complex plane $\mathbb{C}$, and the fact that $|e^{i\theta}-1| \leq |\theta|$. They are bounded by

$$2\mathbb{E}|\xi_1 P_{\varepsilon\eta}^1| \leq 2|\xi_1|\left(\mathbb{E}|P_{\varepsilon\eta}^1|^2\right)^{\frac{1}{2}}.$$

The second moment of $P_{\varepsilon\eta}^1$ can be estimated as (2.12), and is of size η^d.

$$\left|\mathbb{E}\ e^{i\sum_{j\in\mathbb{Z}_*^d}\xi_j I_\varepsilon^j} - \mathbb{E}\{e^{i\xi_1 I_\varepsilon^1}\}\prod_{j\neq 1}\mathbb{E}e^{i\xi_j I_\varepsilon^j}\right| \leq C\rho\left(\frac{\eta}{\varepsilon}\right) + C\eta^{\frac{d}{2}}.$$

Now iterate the above argument for $j = 2, 3, \cdots, M$, where M is the cardinality of the set $\mathbb{Z}_*^d$. At the end, we have

$$\mathcal{E} \leq CM\left[\rho\left(\frac{\eta}{\varepsilon}\right) + \eta^{\frac{d}{2}}\right].$$

So, if we choose $\eta = \varepsilon^{\frac{2}{3}}$, then as ε goes to zero, $\mathcal{E} \sim \varepsilon^{\frac{d}{3}}$ which converges to zero. This shows that the random variables $\{I_\varepsilon^j \mid j \in \mathbb{Z}_*^d\}$ are asymptotically independent.

4. Due to the asymptotic independency, it suffices to investigate the limiting distribution of each I_ε^j separately. These random variables have the same form and hence can be treated once for all. In particular, it suffices to show

$$I_{\varepsilon h} := \frac{1}{\varepsilon^{\frac{d}{2}}} \int_{Q_h} q\left(\frac{x}{\varepsilon}\right) dx \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma \mathcal{N}(0, h^d), \tag{2.14}$$

where Q_h is the cube centered at the origin with sides of length h paralleling the coordinate axes. To show this convergence, we break the cube Q_h into smaller cubes with side length ε. There are totally $N = h/\varepsilon$ (which we assume is integral) such cubes. Denote the small cubes by $\{Q_{hj} \mid j \in \mathbb{Z}^d\}$ and define

$$K_j := \int_{Q_{hj}} \frac{1}{\varepsilon^d} q\left(\frac{x}{\varepsilon}\right) dx = \int_{Q_j^0} q(y) dy, \quad j \in \mathbb{Z}^d. \tag{2.15}$$

Here $\{Q_j^0 \mid j \in \mathbb{Z}^d\}$ are the image of $\{Q_{hj} \mid j \in \mathbb{Z}^d\}$ under the map $s : x \mapsto x\varepsilon^{-1}$. The random variables $\{K_j \mid j \in \mathbb{Z}^d\}$ are stationary mixing random variables. Moreover, $I_{\varepsilon h}$ can be viewed as

$$I_{\varepsilon h} = \varepsilon^{\frac{d}{2}} \sum_{s^* Q_j^0 \in Q_h} K_j = \left(\frac{h}{N}\right)^{\frac{d}{2}} \sum_{j \in \mathbb{Z}^d, |j| \leq N} K_j. \tag{2.16}$$

Here, s^* is the pullback of the map s. Therefore, the sum above accounts for Q_{hj}'s that are inside Q_h. In the second equality, an index $j = (j_1, \cdots, j_d)$ belongs to $\mathbb{Z}^d$, and $|j|$ denotes its infinity norm $\max(|j_1|, \cdots, |j_d|)$. As ε approaches zero, the sum has the central limit scaling $N^{-\frac{d}{2}}$ but is weighted by $h^{\frac{d}{2}}$. Applying the central limit theorem of Bolthausen, Theorem 2.12, we have

$$\frac{1}{N^{\frac{d}{2}}} \sum_{j \in \mathbb{Z}^d, |j| \leq N} \xrightarrow{\text{distribution}} \sigma \mathcal{N}(0, 1),$$

where $\sigma^2 = \sum_{j \in \mathbb{Z}^d} \mathbb{E}\{K_0 K_j\}$. This in turn proves (2.14). To check the mixing conditions in Theorem 2.12, we observe that the ρ-mixing coefficient of $\{K_j \mid j \in \mathbb{Z}^d\}$ is given by that of $\{q(x) \mid x \in \mathbb{R}^d\}$, and $\sum_{j\in\mathbb{Z}^d} |j|^{d-1}\rho(j) < \infty$. We also verify that

$$\sigma^2 = \sum_{j\in\mathbb{Z}^d} \mathbb{E}\{K_0 K_j\} = \int_{Q_0^0} \left(\sum_{j\in\mathbb{Z}^d} \int_{Q_j^0} R(x-y_j)dy_j \right) dx = \left(\int_{\mathbb{R}^d} R(z)dz \right) |Q_0^0|,$$

which agrees with the σ^2 defined in (2.5).

6. We apply (2.14) and find $I_\varepsilon^j \xrightarrow{\text{distribution}} \sigma\mathcal{N}(0, f_j^2 h^d)$ where I_ε^j are defined in step three. Since they are asymptotically independent. We have the following convergence of our main object $I_\varepsilon[f]$, which is nothing but the sum $\sum_j I_\varepsilon^j$:

$$I_\varepsilon[f] \xrightarrow{\text{distribution}} \sigma\mathcal{N}(0, \sum_{j\in\mathbb{Z}_*^d} f_j^2 h^d) = \sigma\mathcal{N}(0, \|f\|_{L^2}^2). \tag{2.17}$$

By Itô isometry, this proves (2.11) for piece-wise functions in step two. Recall the approximating arguments in step one and two to complete the proof of the theorem. □ [1]

2.3 Superposition of Poisson Bumps

The purpose of this section is to explicitly construct a random field that has short-range correlations. In a nutshell, our random field is a superposition of bumps whose centers follow the distribution of a Poisson point process. We start this section with a short review of this process.

2.3.1 The Poisson point process

An $\mathbb{N}$-valued random variable X is said to have a Poisson distribution with parameter λ, denoted by $X \sim \mathcal{P}(\lambda)$, if its probability density function is given by

$$\mathbb{P}(\{X = m\}) = \frac{e^{-\lambda}\lambda^m}{m!}. \tag{2.18}$$

[1] One may wonder whether the theorem still holds after ρ-mixing is replaced by α-mixing. Once I thought I had a proof of this, and used it in a paper [13]. While writing my thesis, I could not reproduce the proof.

The one dimensional Poisson process $\{N(t) \mid t \geq 0\}$ is a continuous time process so that: $N(0) = 0$; $N(t)$ has stationary and independent increments, and for any two times $0 < t_1 < t_2$, the increment $N(t_2) - N(t_1)$ has distribution $\mathcal{P}(t_2 - t_1)$. From another point of view, the jump points of $N(t)$ are a collection of points on the half line $\mathbb{R}_+$. In particular, $N(t)$ induces a random counting measure on intervals of the form (a, b) (which counts how many jump points land in this interval) by

$$N([a,b]) = N(b) - N(a) \sim \mathcal{P}(|(a,b)|),$$

where $|(a,b)|$ denotes the length of this interval. This interpretation of the Poisson process is readily generalized to higher dimensions.

Definition 2.16 (Poisson Point Process)**.** A Poisson point process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of countably many points in $Y = \{y_j \mid j \in \mathbb{N}\} \subset \mathbb{R}^d$ so that for any Borel set $A \subset \mathbb{R}^d$, the cardinality of $Y \bigcap A$, which is denoted by $N(A)$, has Poisson distribution $\mathcal{P}(|A|)$.

Remark 2.17. To put the definition in more abstract form, a Poisson point process is a random variable from $(\Omega, \mathcal{F}, \mathbb{P})$ to the measure space $(\mathfrak{N}, \mathcal{N})$. Here, $\mathfrak{N}$ is the space of locally finite counting measures, i.e., $\mathfrak{N}(A)$ is finite for any compact set $A \subset \mathbb{R}^d$, and $\mathcal{N}$ is the smallest σ-algebra which renders the map $\mathfrak{N} \ni \mathfrak{n} \mapsto \mathfrak{n}(B)$ measurable for any compact sets B. $\square$

A slightly modification of the above definition, in the same way that pure jump Lévy process generalizes the Poisson process, can be formulated by adding a parameter called "intensity" to the Poisson point process.

Definition 2.18. A Poisson point process with intensity $\nu > 0$ is defined as before except that $N(A) \sim \mathcal{P}(\nu|A|)$. We denote such a process by (Y, ν).

Proposition 2.19. *Let (Y, ν) be a Poisson point process; let $A \subset \mathbb{R}^d$ be a Borel set. Conditioned on $\{N(A) = m\}$, the m Poisson points $y_1, \cdots, y_m$ that land in the set A are independently and uniformly located in A.*

This is an important property of the Poisson point process, which shows that the process has complete randomness. We refer the reader to [41] for the proof and an extensive discussion of point process.

2.3.2 Superposition of Poisson bumps

Now we are ready to construct the random field involving Poisson bumps.

Definition 2.20 (Bump)**.** A bump function $\psi(x) : \mathbb{R}^d \to \mathbb{R}$ is a C^∞ function that is compactly supported on the unit ball B_1.

Definition 2.21 (Superposition of Poisson Bumps)**.** Let (Y, ν) be a Poisson point process on $\mathbb{R}^d$, let $\psi(x)$ be a bump function. The superposition of Poisson bumps denoted by ψ_Y is a random field given by

$$\psi_Y(x) = \sum_{j=1}^{\infty} \psi(x - y_j). \tag{2.19}$$

Here $\{y_j \mid j \in \mathbb{N}\}$ are the points in Y. We call ψ the profile of the Poisson bumps.

Remark 2.22. We remark that ψ_Y is not uniformly bounded from above. Indeed, the probability $\mathbb{P}\{N(B_1(x)) = M\}$ is positive (though small) for arbitrary large M. Consequently, with small possibility a large amount of Poisson points can accumulate near x, rendering $\psi_Y(x)$ arbitrarily large. □

The following result shows that ψ_Y is stationary and strong mixing. In particular, it is also ergodic.

Proposition 2.23. *Let $\psi_Y(x)$ be a superposition of Poisson bumps as defined in* (2.19)*. We have*

(1) *$\psi_Y(x)$ is a stationary random field.*

(2) *$\psi_Y(x)$ is ρ-mixing with mixing coefficient $\rho(r) \in C_c(\mathbb{R}_+)$.*

Proof. The stationarity of ψ_Y is due to the fact that the random counting measure $N(A)$ in the definition of (Y, ν) only depends on $|A|$, hence translation invariant.

For the second item, we observe that the σ-algebra $\mathcal{F}_A$ (cf. section 2.2.1) for a set A depends on the Poisson points in the set $A_{+1} := \{y \mid d(y, A) \le 1\}$ (here d is the distance function in Euclidean space). This is because the support of the profile function ψ is B_1. Consequently, as long as $d(A, B) > 2$, the set A_{+1} and B_{+1} will be disjoint which implies that $\mathcal{F}_A$ and $\mathcal{F}_B$ are independent due to Proposition 2.19. Therefore, the mixing coefficient $\rho(r)$ is compactly supported in $[0, 2]$. This completes the proof. □

Remark 2.24. A random field satisfies the second item is very often called m-independent, which is a much weaker dependency than ρ- or α-mixing with any decay rate. In particular, $\psi_Y(x)$ has short range correlations. □

2.3.3 Moments of superposition of Poisson bumps

We move on to derive a systematic formula for the moments of the random field ψ_Y constructed above. As a warm-up, we consider the second moment first, which already reveals the key techniques that allow us to obtain explicit moment formulas.

Since ψ_Y is stationary, its mean is a constant. Fix an $x \in \mathbb{R}^d$; then $\psi_Y(x)$ only depends on Poisson points that land in $B_1(x)$, the unit ball centered at x. The mean of $\psi_Y(x)$ can be calculated conditioning on $N(B_1(x))$, the number of Poisson points inside the ball. We have

$$\mathbb{E}\psi_Y(x) = \mathbb{E} \sum_{y_j \in B_1(x)} \psi(x - y_j) = \sum_{m=1}^{\infty} \left(\mathbb{P}\{N(B_1(x) = m\} \mathbb{E} \left[\sum_{j=1}^{m} \psi(x - y_j') | N(B_1(x)) = m \right] \right).$$

Here, we denote the m points that land in $B_1(x)$ as y'_j. Recall Proposition 2.19, we have

$$\mathbb{E}\psi_Y(x) = \sum_{m=1}^{\infty} \left(e^{-\nu\pi_d} \frac{(\nu\pi_d)^m}{m!} \left[m \int_{B_1(x)} \psi(x-z) \frac{dz}{\pi_d} \right] \right) = \nu \int_{\mathbb{R}^d} \psi(z)dz = \nu\widehat{\psi}(0).$$

Here, π_d denotes the volume of the unit ball in $\mathbb{R}^d$; $\widehat{\psi}$ is the Fourier transform of ψ, which we use only to simplify notation. We observe that the mean of $\psi_Y(x)$ is a constant.

For the second moment, we have

$$\mathbb{E}\{\psi_Y(x_1)\psi_Y(x_2))\} = \mathbb{E} \left(\sum_{j=1}^{\infty} \psi(x_1 - y_j) \ \sum_{k=1}^{\infty} \psi(x_2 - y_k) \right).$$

Since ψ is compactly supported on the unit ball, only those y_j's that are in the set $A = B_1(x_1) \bigcup B_1(x_2)$ contribute to the product, and A is a bounded set. Again, we calculate the expectation conditioning on $N(A)$. The object is now:

$$\sum_{m=1}^{\infty} e^{-\nu|A|} \frac{(\nu|A|)^m}{m!} \mathbb{E}\Big[\sum_{j=1}^{m} \psi(x_1 - y'_j)\psi(x_2 - y'_j) + \sum_{i,j=1, i\neq j}^{m} \psi(x_1 - y'_i)\psi(x_2 - y'_j)$$
$$|N(A) = m\Big] = \nu \int_{B(x_1)\bigcap B(x_2)} \psi(x_1 - z)\psi(x_2 - z)dz + (\nu\widehat{\psi}(0))^2,$$

where we have used Proposition 2.19 again.

Now if we consider the mean-zero random field $\delta\psi_Y := \psi_Y(x) - \mathbb{E}\ \psi_Y$, its correlation function can be written as

$$R(x) = \mathbb{E}\{\psi_Y(0)\psi_Y(x)\} - (\mathbb{E}\psi_Y)^2 = \nu \int_{\mathbb{R}^d} \psi(0 - z)\psi(x - z)dz. \tag{2.20}$$

We note that $R(x)$ is compactly supported in this case.

The preceding calculation reveals three key steps in deriving formulas for the moments $\mathbb{E}\{\prod_{k=1}^{M} \psi_Y(x_k)\}$. First, the moments can be calculated by conditioning on the number of Poisson points in some set. Second, we need a systematic method of tracking the distribution of these Poisson points among $\psi_Y(x_k)$. Third, we need to use Proposition 2.19. The following terminologies borrowed from combinatorics will be helpful for step two. For an positive integer n, let $\mathbb{N}_{\leq n}$ denote the set $\{1, 2, \cdots, n\}$.

Definition 2.25 (Partition of an Integer and Partition of a Set of Integers)**.** Let n be a positive integer, and $\mathbb{N}_{\leq n}$ defined as above.

(1) A *partition of* n is a set of array $(n_1, n_2, \cdots, n_k)$ satisfying that:

$$1 \leq n_1 \leq n_2 \leq \cdots \leq n_k, \text{ which satisfies } n_1 + \cdots + n_k = n.$$

The set of all such partitions is denoted by $\mathcal{P}_n$. A partition of n is called *non-single* if $n_1 \geq 2$. The set of non-single partitions of n is denoted by $\mathscr{G}_n$.

(2) A *partition of* $\mathbb{N}_{\leq n}$ is a collection of nonempty subsets $\{A_i \subset \mathbb{N}_{\leq n}\}$ satisfying

$$\bigcup A_i = \mathbb{N}_{\leq n}, \text{ and } A_i \bigcap A_j = \emptyset \text{ for } i \neq j.$$

If each A_i contains at least two points, the partition is said to be *non-single*.

The total number of all possible partitions of $\mathbb{N}_{\leq n}$ is finite and they are exhausted by first finding all partitions of n, and then for any fixed partition $(n_1, \cdots, n_k) \in \mathcal{P}_n$, finding all possible ways to divide the set $\mathbb{N}_{\leq n}$ into different subsets of cardinalities $n_i, i = 1, \cdots, k$. Observe also that for any given $\{x_1, \cdots, x_n\}$, it can be identified with $\mathbb{N}_{\leq n}$ under the obvious isomorphism. Therefore, these two steps also exhaust all possible ways to divide the set $\{x_i\}, 1 \leq i \leq n$ into disjoint subsets. For a generic term among these grouping methods, a point can be labeled as $x_i^{(\ell, n_j)}$ where $n_j, 1 \leq j \leq k$ comes from the partition of n; once $\{n_j\}$ fixed, ℓ counts the way to divide $\mathbb{N}_{\leq n}$ (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. Here and below, $C_n^{n_1, \cdots, n_k}$ is the *multinomial coefficient* $\frac{n!}{n_1! n_2! \cdots n_k!}$. We denote the *permutation coefficient* by $P_n^k = \frac{n!}{(n-k)!}$.

Now, we calculate the nth moment $\mathbb{E} \prod_{i=1}^n \psi_Y(x_i)$ by conditioning on $N(A)$ where $A =$

$\bigcup B(x_i)$. We have,

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

The product of sums can be written as

$$\prod_{i=1}^{n}\sum_{j=1}^{m}\psi(x_i-y_j)=\sum_{(n_1,\cdots,n_k)\in\mathcal{P}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\sum_{p=1}^{P_m^k}\prod_{j=1}^{k}\prod_{i=1}^{n_j}\psi(x_i^{\ell,n_j}-y_j^p). \tag{2.21}$$

Here $P_m^k, m\geq k$ 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.

$$\begin{aligned}&\sum_{m=1}^{\infty}e^{-\nu|A|}\frac{(\nu|A|)^m}{m!}\mathbb{E}\left[\sum_{(n_1,\cdots,n_k)\in\mathcal{P}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\sum_{p=1}^{P_m^k}\prod_{j=1}^{k}\prod_{i=1}^{n_j}\psi(x_i^{\ell,n_j}-y_j^p)\big|N(A)=m\right]\\ =&\sum_{(n_1,\cdots,n_k)\in\mathcal{P}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^{k}\nu\int\prod_{i=1}^{n_j}\psi(x_i^{\ell,n_j}-z)dz \\ =&\sum_{(n_1,\cdots,n_k)\in\mathcal{P}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^{k}T^{n_j}(x_1^{\ell,n_j},\cdots,x_{n_j}^{\ell,n_j}).\end{aligned} \tag{2.22}$$

Here, we defined T^{n_j} to be $T^{n_j}(x_1,\cdots,x_{n_j}):=\nu\int\prod_{i=1}^{n_j}\psi(x_i-z)dz$. In the second step in the derivation above, we used Proposition 2.19 again, which implies

$$\mathbb{E}\left[\sum_{p=1}^{P_m^k}\prod_{j=1}^{k}\prod_{i=1}^{n_j}\psi(x_i^{\ell,n_j}-y_j^p)\big|N(A)=m\right]=\sum_{m=1}^{\infty}e^{-\nu|A|}\frac{(\nu|A|)^m}{m!}P_m^k\prod_{j=1}^{k}\int_A\prod_{i=1}^{n_j}\psi(x_i^{\ell,n_j}-z)\frac{dz}{|A|}.$$

To derive higher order moments of the mean-zero random field $\delta\psi_Y$, we observe that

$$\prod_{i=1}^{n}\delta\psi_Y(x_i)=\prod_{i=1}^{n}[\psi_Y(x_i)-\nu\widehat{\psi}(0)]=\sum_{m=0}^{n}(-\nu\widehat{\psi}(0))^m\sum_{s=1}^{C_n^{n-m}}\prod_{i=1}^{n-m}\sum_{j=1}^{\infty}\psi(x_i^{s,n-m}-y_j). \tag{2.23}$$

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 2.26. *Let $\mathscr{G}_n$ be defined as before. For the mean-zero process $\delta\psi_Y$, we have*

$$\mathbb{E}\prod_{i=1}^{n}\delta\psi_Y(x_i)=\sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^{k}T^{n_j}(x_1^{\ell,n_j},\cdots,x_{n_j}^{\ell,n_j}). \tag{2.24}$$

The only difference of this formula with that of the higher order moments of ψ_Y 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 $\nu\widehat{\psi}(0)$, 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 $\mathbb{E}\prod\sum\psi(x_i-y_j)$ and the expression of $\prod\delta\psi_Y(x_i)$, we observe that the moment $\mathbb{E}\prod\delta\psi_Y(x_i)$ consists of terms of the form:

$$\pm(\nu\widehat{\psi}(0))^l\prod_{j=1}^{k}T^{n_j} \tag{2.25}$$

where $n_j \geq 2$, $k \leq n-l$ and $\sum n_j = n-l$. The terms with $l=0$ are exactly those in (2.24). We show that all the other terms with $l \geq 1$ vanish. Without loss of generality, we consider the term

$$(\nu\widehat{\psi}(0))^l T^{n_1}(x_1,\cdots,x_{n_1})T^{n_2}(x_{n_1+1},\cdots,x_{n_2})\cdots T^{n_k}(x_{n_{k-1}+1},\cdots,x_{n_k}). \tag{2.26}$$

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 $(\nu\widehat{\psi}(0))^l$.

This term appears in the expectation of the right hand side of (2.23) 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 $(\nu\widehat{\psi}(0))$, 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 (2.23) 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

$$\sum_{k=0}^{l}(-1)^k C_l^k = 0. \tag{2.27}$$

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

2.3.4 Scaling of the intensity

As mentioned earlier in this chapter, we will scale the parameter in the random field $q(x,\omega)$ properly so that it models some heterogeneity of the background media on which some PDEs are posed. For instance, a typical realization $q(x,\omega)$ oscillates about its mean value zero as x varies; assume the correlation length, the length scale on which $q(x)$ varies between its local minimum and maximum, is of order one. Now the scaled version $q\left(\frac{x}{\varepsilon},\omega\right)$ has correlation length of order ε. In other words, the scaled version is of high frequency with order ε^{-1}.

Use the scaling procedure above; we get a highly oscillating random field $\psi_Y\left(\frac{x}{\varepsilon}\right)$. In the literature, however, other ways of generating highly oscillatory random field of superposition of Poisson point process can be used, as in [12, 17]. Let $(Y_\varepsilon, \varepsilon^{-d}\nu)$ be a Poisson point process with intensity $\varepsilon^{-d}\nu$. Then one can define

$$\widetilde{\psi_{Y_\varepsilon}} := \sum_{j=1}^{\infty} \psi\left(\frac{x-y_j^\varepsilon}{\varepsilon}\right).$$

Let us show that these two definitions of scaling are equivalent in the sense of distribution.

By their definitions, it suffices to show that $\varepsilon^{-1}(Y_\varepsilon, \varepsilon^{-d}\nu)$ has the same distribution as

(Y, ν). This is easily verified by:

$$\mathbb{P}\{N(A; \varepsilon^{-1} Y_\varepsilon) = m\} = \mathbb{P}\{N(\varepsilon A; Y_\varepsilon) = m\} = \frac{e^{-\nu\varepsilon^{-d}|\varepsilon A|}(\nu\varepsilon^{-d}|\varepsilon A|)^m}{m!}, \forall A \in \mathscr{B}(\mathbb{R}^d).$$

Since $\varepsilon^{-d}|\varepsilon A| = |A|$. The probability above is precisely $\mathbb{P}\{N(A; Y) = m\}$. This completes the proof.

Finally, we observe that many functions associated with the scaled random field $q\left(\frac{x}{\varepsilon}\right)$, such as the correlation function and moments formulas, can be obtained by scaling the variables of the random field $q(x)$. In particular, the correlation function $R_\varepsilon(x)$ of $q\left(\frac{x}{\varepsilon}\right)$ is precisely $R\left(\frac{x}{\varepsilon}\right)$.

2.4 Functions of Gaussian Random Fields

Let $\{g(x) \mid x \in \mathbb{R}^d\}$ be a stationary real-valued Gaussian random field given on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Without loss of generality, we assume $g(x)$ is mean-zero and variance-one. Let $R_g(x) := \mathbb{E}\{g(0)g(x)\}$ be its covariance function that satisfies

$$R_g(x) \sim \frac{\kappa_g}{|x|^\alpha}, \qquad \text{for } |x| \text{ large.} \tag{2.28}$$

When $\alpha < d$, this is a Gaussian field with long-range correlation. Note that the covariance function itself is enough to determine such a Gaussian random field. We will refer the following process as "function of Gaussian".

Definition 2.27 (Function of Gaussian)**.** A random field $q(x, \omega)$ constructed by a function of Gaussian is defined as $\Phi \circ g(x, \omega)$, i.e., $\Phi(g(x))$, for some bounded real function $\Phi : \mathbb{R} \to \mathbb{R}$ satisfying

$$\int_{\mathbb{R}} \Phi(s) e^{-\frac{s^2}{2}} ds = 0. \tag{2.29}$$

We observe, in particular, that $q(x)$ such defined is uniformly bounded by $\|\Phi\|_{L^\infty}$, and is mean-zero. The motivation for this definition is to construct a field that is bounded (note Gaussian is not uniformly bounded) and for which some explicit calculation can be done (thanks to the underlying Gaussian field). As the first example of such explicit calculations, we show that $R(x)$ has the same asymptotic behavior as R_g in (2.28).

Lemma 2.28. *Let $q(x)$ be the random field above. Define $V_1 = \mathbb{E}\{g_0\Phi(g_0)\}$ where g_x is the underlying Gaussian random field. There exist some $T, C > 0$ such that the autocorrelation function $R(x)$ of q satisfies*

$$|R(x) - V_1^2 R_g(x)| \leq C R_g^2(x), \quad \textit{for all } |x| \geq T, \tag{2.30}$$

where R_g is the correlation function of g. Further,

$$|\mathbb{E}\{g(y)q(y+x)\} - V_1 R_g(x)| \leq C R_g^2(x), \quad \textit{for all } |x| \geq T. \tag{2.31}$$

Proof. A proof of this lemma can be found in [11]; we record it here for the reader's convenience.

$$R(x) = \frac{1}{2\pi\sqrt{1 - R_g^2(x)}} \int_{\mathbb{R}^2} \Phi(g_1)\Phi(g_2) \exp\left(-\frac{g_1^2 + g_2^2 - 2R_g(x)g_1g_2}{2(1 - R_g^2(x))}\right) dg_1 dg_2.$$

For large $|x|$, the coefficient $R_g(x)$ is small and we can expand the value of the double integral in powers of $R_g(x)$. The zeroth order term is the integration of $\Phi(g_1)\Phi(g_2)$ with respect to $\exp(-|g|^2/2)dg$ where dg is short for dg_1dg_2; this term vanishes due to (2.29). The first order term is integration of $\Phi(g_1)\Phi(g_2)g_1g_2$ with respect to the $\exp(-|g|^2/2)dg$, which gives $V_1^2 R_g(x)$.

Similarly, for the second item in the lemma, we first write

$$\mathbb{E}\{g(y)\Phi(g(y+x))\} = \frac{1}{2\pi\sqrt{1-R_g^2(x)}} \int_{\mathbb{R}^2} g_1 \Phi(g_2) \exp\left(-\frac{g_1^2+g_2^2-2R_g(x)g_1g_2}{2(1-R_g^2(x))}\right) dg_1 dg_2.$$

Then we expand the value of the double integral in powers of R_g and characterize the first two orders as before. □

It follows that $R(x)$ behaves like $\kappa|x|^{-\alpha}$, where $\kappa = V_1^2\kappa_g$, for large $|x|$. In particular, there exists some constant C so that $|R(x)| \leq C|x|^{-\alpha}$. When $\alpha < d$, R in not integrable and $q(x)$ has long-range correlations.

Similarly, $q(x)$ is uniformly bounded and strong-mixing provided that the underlying Gaussian random field is strong mixing and the function Φ is uniformly bounded.

Proposition 2.29. *Let $q(x,\omega)$ be the random field model in Definition 2.27 with some Φ satisfying the conditions there. Suppose that $|\Phi|$ is uniformly bounded by some positive number q_0. Assume also that the underlying Gaussian random field $g(x)$ is strong mixing with mixing coefficient $\alpha(r)$ satisfying the condition* (2.10)*. Then $q(x,\omega)$ is uniformly bounded and has the same strong mixing properties.*

Proof. From the definition of q and the bound on $|\Phi|$ it is obvious that $q(x,\omega)$ is uniformly bounded. Also from the definition of q, we see that the σ-algebra $\mathcal{F}_A$ generated by variables $q(x,\omega), x \in A$ is in fact generated by the underlying Gaussian random variables $g(x,\omega), x \in A$. Hence q shares the same stationarity and strong mixing coefficient $\alpha(r)$ with g. □

2.4.1 Fourth order moments formulas

As for the Poisson bumps model, we wish to develop high-order moments formulas for the model $\Phi \circ g$. In the general case, it is difficult to obtain formulas for arbitrary moments. Nevertheless, we derive an estimate of fourth order moments assuming an additional condition on the function Φ in the model. We form the estimate in terms of the fourth order cumulants.

Some terminologies are in order. Let $F = \{1, 2, 3, 4\}$ and $\mathcal{U}$ be the collections of two pairs of unordered numbers in F, i.e.,

$$\mathcal{U} = \big\{p = \{\big(p(1), p(2)\big), \big(p(3), p(4)\big)\} \mid p(i) \in F, p(1) \neq p(2), p(3) \neq p(4)\big\}. \tag{2.32}$$

As members in a set, the pairs $(p(1), p(2))$ and $(p(3), p(4))$ are required to be distinct; however, they can have one common index. There are three elements in $\mathcal{U}$ whose indices $p(i)$ are all different. They are precisely $\{(1, 2), (3, 4)\}$, $\{(1, 3), (2, 4)\}$ and $\{(1, 4), (2, 3)\}$. Let us denote by $\mathcal{U}_*$ the subset formed by these three elements, and its complement by $\mathcal{U}^*$.

Intuitively, we can visualize $\mathcal{U}$ in the following manner. Draw four points with indices 1 to 4. There are six line segments connecting them. The set $\mathcal{U}$ can be visualized as the collection of all possible ways to choose two line segments among the six. $\mathcal{U}_*$ corresponds to choices so that the two segments have disjoint ends, and $\mathcal{U}^*$ corresponds to choices such that the segments share one common end.

Definition 2.30. We say that $q(x, \omega)$ has *controlled fourth order cumulants* with control functions $\{\phi_p \in L^\infty(\mathbb{R}^d \times \mathbb{R}^d) \mid p \in \mathcal{U}^*\}$ if: There exists such control functions, and for any four point set $\{x_i\}_{i=1}^4$, $x_i \in \mathbb{R}^d$, we have the following condition on the fourth order cross-moment of $\{q(x_i, \omega)\}$:

$$\begin{aligned} &\Big|\mathbb{E}\prod_{i=1}^4 q(x_i) - \sum_{p \in \mathcal{U}_*} \mathbb{E}\{q(x_{p(1)})q(x_{p(2)})\}\mathbb{E}\{q(x_{p(3)})q(x_{p(4)})\}\Big| \\ &\leq \sum_{p \in \mathcal{U}^*} \phi_p(x_{p(1)} - x_{p(2)}, x_{p(3)} - x_{p(4)}). \end{aligned} \tag{2.33}$$

Observe that since $\mathbb{E}q(x, \omega) \equiv 0$, the left hand side is the (joint) *cumulant* of $\{q(x_i, \omega)\}$, and hence the notation for this property. In the sequel, we will denote the cumulant of

$\{q(x_i)\}_{i=1}^4$ by $\vartheta(q(x_1), \cdots, q(x_4))$.

Remark 2.31. This definition is motivated by Gaussian random fields for which all but two cumulants vanish and hence we can set ϕ_p to be zero for all p in (2.33). Although it satisfies the condition above, a Gaussian random field is not bounded and large negative values of q_ε may yield non-uniqueness of PDE. The above condition on the cumulants hence provides a "decomposition" of fourth order moments into pairs just as Gaussian random fields up to an error we wish to control.

With a further assumption on the function Φ, we show that the model in Definition 2.27 has controlled fourth order cumulants.

Proposition 2.32. *Let $q(x, \omega)$ be the random field in Definition 2.27 with some Φ satisfying the conditions there. Further assume that the Fourier transform of Φ satisfies that*

$$\int_{\mathbb{R}} |\hat{\Phi}(\xi)|\big(1 + |\xi|^3\big) < \infty; \tag{2.34}$$

Denote by κ_c the value of this integral which is a finite positive real number.

Then $q(x, \omega)$ has controlled fourth order moments with control functions $\{81\kappa_c^4 |R_g \otimes R_g|\}$.

Proof. Recall the definition of $q(x)$ and the underlying Gaussian random field $g(x)$. Fix any four points $\{x_i\}_{i=1}^4$ and let ϑ be the joint cumulant of $\{q(x_i)\}$; in the Fourier domain it can be expressed as

$$\vartheta = \int_{\mathbb{R}^4} \prod_{j=1}^4 \hat{\Phi}(\xi_j) e^{-\frac{\xi^t \xi}{2}} \Big(\prod_{i=1}^3 e^{-\frac{1}{2}\xi^t D_i \xi} - \sum_{i=1}^3 e^{-\frac{1}{2}\xi^t D_i \xi} \Big) d^4 \xi. \tag{2.35}$$

Here ξ^t denotes the transpose of ξ, and the matrices $D_i, i = 1, 2, 3$ are defined as follows:

$$D_1 = \begin{pmatrix} 0 & \rho_{12} & 0 & 0 \\ \rho_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_{34} \\ 0 & 0 & \rho_{34} & 0 \end{pmatrix}, D_2 = \begin{pmatrix} 0 & 0 & \rho_{13} & 0 \\ 0 & 0 & 0 & \rho_{24} \\ \rho_{13} & 0 & 0 & 0 \\ 0 & \rho_{24} & 0 & 0 \end{pmatrix}, D_3 = \begin{pmatrix} 0 & 0 & 0 & \rho_{14} \\ 0 & 0 & \rho_{23} & 0 \\ 0 & \rho_{23} & 0 & 0 \\ \rho_{14} & 0 & 0 & 0 \end{pmatrix},$$

where $\rho_{ij} := R_g(x_i - x_j)$ is the covariance of $g(x_i)$ and $g(x_j)$. We apply the following identity to the product and the sum inside the parenthesis in (2.35).

$$abc - a - b - c = (a-1)(b-1)(c-1) + (a-1)(b-1) + (a-1)(c-1) + (b-1)(c-1) - 2,$$

We then use (2.29) to argue that the constant two above does not contribute to (2.35). Hence we have

$$\vartheta = \int_{\mathbb{R}^4} \prod_{j=1}^{4} \hat{\Phi}(\xi_j) e^{-\frac{\xi^t \xi}{2}} \left(\prod_{i=1}^{3} [e^{-\frac{1}{2}\xi^t D_i \xi} - 1] + \sum_{i<k} [e^{-\frac{1}{2}\xi^t D_i \xi} - 1][e^{-\frac{1}{2}\xi^t D_k \xi} - 1] \right).$$

For each fixed ξ, we use the Taylor expansion for exponential function and write

$$e^{-\frac{1}{2}\xi^t D_i \xi} - 1 = -\frac{1}{2}\xi^t D_i \xi e^{-\frac{1}{2}\xi^t (c_i D_i)\xi},$$

where the real number c_i depends on ξ and D_i but is always an element in $[0, 1]$. Therefore, we have

$$\vartheta = \int_{\mathbb{R}^4} \prod_{j=1}^{4} \hat{\Phi}(\xi_j) \Bigg(-e^{-\frac{1}{2}\xi^t (I + \sum_{i=1}^3 c_i D_i)\xi} \prod_{i=1}^{3} \frac{1}{2}\xi^t D_i \xi +$$
$$+ \sum_{i<k} e^{-\frac{1}{2}\xi^t (I + c_i D_i + c_k D_k)\xi} [\frac{1}{2}\xi^t D_i \xi][\frac{1}{2}\xi^t D_k \xi] \Bigg) d^4\xi.$$

Observe that $I + D_i$, $I + D_i + D_j$ with $(i < j)$ for $i, j = 1, 2, 3$, and $I + \sum_{i=1}^3 D_i$ are non-negative definite matrices. Since $c_i \in [0,1]$, we deduce that $I + c_i D_i + c_k D_k$ for any $i < k$, and $I + \sum_{i=1}^3 c_i D_i$ are all non-negative definite. Indeed, we can rewrite them as a sum of non-negative definite matrices. For instance, without loss of generality we assume c_i is increasing in i, and then

$$I + \sum_{i=1}^{3} c_i D_i = c_1(I + \sum_{i=1}^{3} D_i) + (c_2 - c_1)(I + \sum_{i=2}^{3} D_i) + (c_3 - c_2)(I + D_3) + (1 - c_3)I.$$

Each of the matrices on the right hand side above is non-negative definite.

Therefore, we can bound the exponential terms in the integral by one, and conclude that

$$|\vartheta| \le \int_{\mathbb{R}^4} \prod_{j=1}^{4} \left|\hat{\Phi}(\xi_j)\right| \left(\prod_{i=1}^{3} \left|\frac{1}{2}\xi^t D_i \xi\right| + \sum_{i<k} \left|\frac{1}{2}\xi^t D_i \xi\right| \cdot \left|\frac{1}{2}\xi^t D_k \xi\right| \right).$$

Now the products in the parenthesis above are just polynomials in the $|\xi_j|$ variables, and for each ξ_j, the highest possible power on it is three. The coefficients in those polynomials are products of two or three ρ_{ij} functions. Since $|\rho_{ij}| \le 1$ by definition, we can bound the $\xi^t D_1 \xi$ of the first member in the parenthesis above by $|\xi_1 \xi_2| + |\xi_3 \xi_4|$. Then after evaluating the product, the coefficients in the polynomial of $|\xi_j|$ variables are products of two ρ_{ij} functions. With this in mind, it is easy to verify that

$$\begin{aligned}
&\left|\vartheta(q(x_1), \cdots, q(x_4))\right| \le \big(|\rho_{12}\rho_{13}| + |\rho_{12}\rho_{24}| + |\rho_{34}\rho_{13}| + |\rho_{34}\rho_{24}| \\
&\quad + |\rho_{12}\rho_{14}| + |\rho_{12}\rho_{23}| + |\rho_{34}\rho_{14}| + |\rho_{34}\rho_{23}| \\
&\quad + |\rho_{13}\rho_{14}| + |\rho_{13}\rho_{23}| + |\rho_{24}\rho_{14}| + |\rho_{24}\rho_{23}|\big) \int_{\mathbb{R}^4} \prod_{j=1}^{4} \hat{\Phi}(\xi_j)\big(|\xi_j|^3 + |\xi_j|^2 + |\xi_j| + 1\big) d^4\xi.
\end{aligned}$$

Thanks to (2.34), the last integral is finite and can be bounded by $3^4 \kappa_c^4$. Compare the above inequality with the cumulant condition, i.e., (2.33); we see that all pairs of indices in the products of ρ functions above lie in $\mathcal{U}^*$ where $\mathcal{U}$ is defined in (2.32). Then for each $p \in \mathcal{U}^*$, we set $\phi_p := 81\kappa_c^4 |R_g \otimes R_g|$. We see (2.33) is indeed satisfied. This completes the proof. $\square$

The above model is not the only type that has controlled fourth order cumulants. Recall the moments formula (2.24) for the Poisson bumps model $\varphi_Y(x)$ in (2.19). If we define $q(x,\omega)$ to be its mean-zero part, then the joint cumulant of $\{q(x_i,\omega)\}_{i=1}^4$ has the following expression;

$$\begin{aligned}\vartheta(q(x_1),\cdots,q(x_4)) &= \nu\int \varphi(z)\varphi(x_2-x_1+z)\varphi(x_3-x_1+z)\varphi(x_4-x_1+z)dz\\ &\le \nu\|\varphi\|_{L^\infty}\int \varphi(z)\varphi(x_2-x_1+z)\varphi(x_3-x_1+z)dz.\end{aligned} \tag{2.36}$$

We verify that the last integral above is bounded uniformly in the variables x_2-x_1 and x_3-x_1 since φ is bounded. In other words, the cumulant function ϑ satisfies (2.33), for we can set ϕ_p to be the last integral in (2.36) for $p=\{(1,2),(1,3)\}$ and $\phi_p\equiv 0$ for all other p. This verifies that $q(x,\omega)$ defined above has controlled cumulants. In fact, these ϕ_p functions are integrable in their variables since the profile function φ is compactly supported.

2.5 Random Fields with Long-range Correlations

In this section, we revisit the oscillatory integral $O_\varepsilon[f] := \int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx$, for some stationary mean-zero random field $q(x)$, an L^2 function f on some domain $X\in\mathbb{R}^d$, in the case when $q(x)$ has long range correlations. That is, when the correlation function $R(x)$ of q fails to be integrable, so that Theorem 2.15 on this oscillatory integral ceases to work.

There is no central limit type results for a general long range correlated random field. Therefore, we constrain ourselves to the case when $q(x,\omega)$ is a function of long-range Gaussian defined as follows.

Definition 2.33 (Function of Long-Range Gaussian)**.** A function of Gaussian $q(x,\omega)$ defined in Definition 2.27 is said to have long range correlation if the correlation function of the underlying Gaussian random field satisfies (2.28) with $\alpha<d$.

2.5.1 Convergence in distribution results

As in the case of strong mixing random field, we are interested in the limiting distribution of oscillatory integrals of the form

$$I_\varepsilon[f] := \frac{1}{\sqrt{\varepsilon^\alpha}} \int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx. \tag{2.37}$$

Note that the scaling factor is $\varepsilon^{\frac{\alpha}{2}}$, which is longer than $\varepsilon^{\frac{d}{2}}$ in the strong mixing case. This is indeed the correct scaling, because the variance of the integral $\int_X q_\varepsilon(x)f(x)dx$ is of order ε^α, which can be easily checked.

In the strong mixing case, the limiting distribution of the oscillating integral is captured by Theorem 2.15, in which the limit is written as a stochastic integral with respect to Brownian motion. In the long range case, we will see that the limit can be writen as a stochastic integral again, but one with respect to the fractional Brownian motion (fBm) which, unless the Brownian motions, has correlated increments. For the convenience of the reader, we briefly review some essential properties of fBm, and stochastic integral with fBm integrator.

A fBm $W^H(t)$ with Hurst index H is a mean-zero Gaussian process with $W^H(0) = 0$, stationary increments and H-self-similarity, that is, for $a > 0$,

$$\{W^H(at)\}_{t\in\mathbb{R}} \stackrel{\mathscr{D}}{=} \{a^H W^H(t)\}_{t\in\mathbb{R}}, \tag{2.38}$$

where $\stackrel{\mathscr{D}}{=}$ means the equality in the sense of finite dimensional distributions. From this similarity relation, we deduce $\mathbb{E}[(W^H(t))^2] = |t|^{2H}\mathbb{E}[(W^H(1))^2]$. In particular, if $\mathbb{E}[(W^H(1))^2] = 1$ we say the fBm is standard. It follows from the stationarity of increments that the covariance function of $W^H(t)$ is given by

$$R^H(t,s) = \mathbb{E}\{W^H(t)W^H(s)\} = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |s-t|^{2H}\right). \tag{2.39}$$

When $H = 1/2$, the increments of W^H are independent and the fBm reduces to the usual Brownian motion. For $H \neq 1/2$, the increments are stationary but not independent.

Stochastic integrals with respect to fBm can be defined on many functional spaces. Note that $H = 1 - \frac{\alpha}{2}$ is in the interval $(\frac{1}{2}, 1)$. In this case, a convenient functional space to define stochastic integral is

$$|\Gamma|^H = \left\{ f : \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x)||f(y)||x-y|^{2(H-1)} dxdy < \infty \right\}. \tag{2.40}$$

It is easy to check, for instance from the Hardy-Littlewood-Sobolev lemma [78, §4.3], that $L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \subset L^{1/H} \subset |\Gamma|^H$. Stochastic integrals against fBm do not satisfy Itô isometry; instead, we have

$$\mathbb{E}\left\{ \int_{\mathbb{R}} f(t) dW_t^H \int_{\mathbb{R}} h(s) dW_s^H \right\} = H(2H-1) \int_{\mathbb{R}^2} \frac{f(t)h(s)}{|t-s|^{2(1-H)}} dtds. \tag{2.41}$$

The right hand side is a double integral. Heuristically we can write $\mathbb{E}\{dW^H(t)dW^H(s)\} = |t-s|^{-2(1-H)}dtds$. For a nice review on stochastic integral with respect to fractional Brownian motion, we refer the reader to [96]. Now we are ready to consider the oscillatory integral (2.37).

a. One-dimensional case

In the one dimensional case, we have the following theorem.

Theorem 2.34 (Oscillatory Integral in Long-Range Media). *Let $q(x, \omega)$ be a function of long-range Gaussian with decorrelation rate $\alpha < 1$ as in Definition 2.33. Let F be a function that is both bounded and integrable on $\mathbb{R}$. Then*

$$\frac{1}{\varepsilon^{\frac{\alpha}{2}}} \int_{\mathbb{R}} q\left(\frac{x}{\varepsilon}\right) F(x) dx \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sqrt{\frac{\kappa}{H(2H-1)}} \int_{\mathbb{R}} F(x) dW^H(x). \tag{2.42}$$

Here, $W^H(x)$ is the fractional Brownian motion with Hurst index $H = 1 - \frac{\alpha}{2}$. The constant

κ is defined to be $\kappa_g \left(\mathbb{E}\{g_0\Phi(g_0)\}\right)^2$ where g_0 is the value of the underlying Gaussian process at zero.

This theorem is stated as Theorem 3.1 of [11]. A proof of it can be found there. We reiterate that a hidden condition $\mathbb{E}\{g_0\Phi(g_0)\} \neq 0$ is assumed. When this quantity is zero, the limit above is zero, indicating that the scaling $\varepsilon^{-\alpha/2}$ is not optimal. In fact, when $\alpha < 1/2$, the integral $\int_{\mathbb{R}} q_\varepsilon F(x)dx$ has variance of order $\varepsilon^{2\alpha}$. Divided by ε^α, the resulted integral has non-Gaussian limit. See the notes at the end of this chapter.

b. High-dimensional case

Fix N arbitrary test functions $\{\psi_k(x);\ 1 \le k \le N\}$ in $L^2(X)$. Consider the law of random vectors of the form $(J_1^\varepsilon(\omega), \cdots, J_N^\varepsilon(\omega))$, where

$$J_j^\varepsilon(\omega) := -\frac{1}{\varepsilon^{\alpha/2}} \int_X q_\varepsilon(y)\psi_j(y)dy. \tag{2.43}$$

We have the following result characterizing the limiting joint law of them.

Lemma 2.35. *The random vector $(J_1^\varepsilon, J_2^\varepsilon, \cdots, J_N^\varepsilon)$ converges in distribution to the centered Gaussian random vector $(J_1, J_2, \cdots, J_N)$ whose covariance matrix is given by*

$$C_{ik} = \mathbb{E}\{J_i J_k\} = \int_{X^2} \frac{\kappa\ \psi_i(y)\psi_k(z)}{|y-z|^\alpha} dydz. \tag{2.44}$$

Moreover, the random variable J_k admits the following stochastic integral representation.

$$J_k = -\int_X \psi_k(y) W^\alpha(dy). \tag{2.45}$$

Here $W^\alpha(dy)$ is as formally defined as $\dot{W}^\alpha(y)dy$ and $\dot{W}^\alpha(y)$ is a Gaussian random field with covariance function given by $\mathbb{E}\{\dot{W}^\alpha(x)\dot{W}^\alpha(y)\} = \kappa|x-y|^{-\alpha}$.

Proof. We want to show that $\forall t_1, t_2, \cdots, t_N \in \mathbb{R}$, $\sum_{i=1}^N t_i J_i^\varepsilon$ converges in distribution to

$\sum_{i=1}^N t_i J_i$. Since

$$\sum_{i=1}^N t_i J_i^\varepsilon = -\frac{1}{\varepsilon^{\alpha/2}} \int_X q_\varepsilon(y) \sum_{i=1}^N t_i \psi_i(y) dy,$$

$$\sum_{i=1}^N t_i J_i = -\int_X (\sum_{i=1}^N t_i \psi_i(y)) W^\alpha(dy),$$

and $\sum_{i=1}^N t_i \psi_i(y) \in L^2(X)$, we only need to show

$$-\frac{1}{\varepsilon^{\alpha/2}} \int_X q_\varepsilon(y) f(y) dy \xrightarrow[\varepsilon\to 0]{\text{distribution}} -\int_X f(y) W^\alpha(dy) \tag{2.46}$$

for any $f \in L^2(X)$.

We prove this convergence in two steps: First, we show it holds when $q(x) = g(x)$, i.e., q is a centered stationary Gaussian field. Second, we generalize the result to the case when $q(x) = \Phi(g(x))$.

The Gaussian case. When $q(x) = g(x)$, the random variable $-\varepsilon^{-\alpha/2} \int_X q_\varepsilon(y) f(y) dy$ is centered, Gaussian, with variance $S_\varepsilon := \varepsilon^{-\alpha} \int_{X^2} R_g(\frac{y-z}{\varepsilon}) f(y) f(z) dydz$, so it suffices to show

$$S_\varepsilon \longrightarrow \int_{X^2} \frac{\kappa_g f(y) f(z)}{|y-z|^\alpha} dydz =: \text{ Var } \left(-\int_X f(y) W^\alpha(dy)\right) \tag{2.47}$$

as $\varepsilon \to 0$. The equality above holds by the definition of our stochastic integral. Note that in this case, $q(x) = \Phi(g(x))$ with $\Phi(s) = s$; consequently, the κ in the covariance function of W^α in Theorem 6.9 is precisely κ_g, because $\mathbb{E}\{g(0)\Phi(g(0))\} = \mathbb{E}\{g(0)^2\} = 1$.

Since $R_g(x) \sim \kappa_g |x|^{-\alpha}$, for any $\delta > 0$, there exists an $M > 0$ so that $|x| > M$ implies $|R_g(x) - \kappa_g |x|^{-\alpha}| < \delta \kappa_g |x|^{-\alpha}$. According to this, we have

$$\begin{aligned} \left| S_\varepsilon - \int_{X^2} \frac{\kappa_g f(y) f(z)}{|y-z|^\alpha} dydz \right| \le \int_{|y-z|>M\varepsilon} \frac{\delta\kappa_g |f(y) f(z)|}{|y-z|^\alpha} \, dydz + \\ + \int_{|y-z|\le M\varepsilon} |f(y) f(z)| \left(\varepsilon^{-\alpha} + \frac{\kappa_g}{|y-z|^\alpha} \right) dydz := (I) + (II) + (III). \end{aligned}$$

We have used the fact $\|R\|_\infty = 1$. It is easy to see that $(I) \leq C\delta$, $(II) + (III) \leq C\varepsilon^{d-\alpha}$. First let $\varepsilon \to 0$, then let $\delta \to 0$, we prove (2.47).

The case of a function of the Gaussian field. In this case, $q(x) = \Phi(g(x))$ for more general Φ. Recall that $V_1^2 = \mathbb{E}\{g(0)\Phi(g(0))\}$ and V_1 is assumed to be positive. we claim that the difference between the random variables $\varepsilon^{-\alpha/2}\int_X q_\varepsilon(y)f(y)dy$ and $\varepsilon^{-\alpha/2}\int_X V_1 g_\varepsilon(y)f(y)dy$ converges to zero in probability. Then (2.46) follows from this, the Gaussian case, and the fact $\kappa = \kappa_g V_1^2$.

To show the convergence in probability, we estimate the second moment as follows:

$$\begin{aligned} &\mathbb{E}\left(\frac{1}{\varepsilon^{\alpha/2}}\int_X (q_\varepsilon(y) - V_1 g_\varepsilon(y))f(y)dy\right)^2 \\ = \;& \frac{1}{\varepsilon^\alpha}\int_{X^2} \mathbb{E}\{(q_\varepsilon(y) - V_1 g_\varepsilon(y))(q_\varepsilon(z) - V_1 g_\varepsilon(z))\}f(y)f(z)dydz. \end{aligned}$$

The expectation term inside the integral can be written as

$$\begin{aligned} R_\varepsilon(y-z) - V_1^2(R_g)_\varepsilon(y-z) + V_1[V_1(R_g)_\varepsilon(y-z) - \mathbb{E}\{g_\varepsilon(y)q_\varepsilon(z)\}] \\ + V_1[(R_g)_\varepsilon(y-z) - \mathbb{E}\{g_\varepsilon(z)q_\varepsilon(y)\}]. \end{aligned}$$

Recall (2.31) of Lemma 2.28 to estimate these terms. We can bound the second moment above by

$$C\varepsilon^{-\alpha}\int_{|y-z|\leq T\varepsilon} |f(y)f(z)|dydz + C\varepsilon^{-\alpha}\int_{|y-z|>T\varepsilon} \frac{\varepsilon^{2\alpha}|f(y)f(z)|}{|y-z|^{2\alpha}}dydz := (I) + (II).$$

Carry out the routine analysis we have developed for this type of integrals; it is easy to verify that $(I) \leq C\varepsilon^{d-\alpha}$ and (II) is of order ε^α if $2\alpha < d$, of order $\varepsilon^\alpha|\log\varepsilon|$ if $2\alpha = d$, and of order $\varepsilon^{d-\alpha}$ if $2\alpha > d$. In all cases, we have $(I) + (II)$ converges to zero, which completes the proof. $\square$

2.6 Convergence in Distribution in Functional Spaces

So far we have only considered the limiting distribution of oscillatory integrals of the form (2.7) or (2.37). Suppose we have a random process $V(x,\omega)$ related to $q(x,\omega)$ by

$$V(x,\omega) = \int_X G(x,y) q\left(\frac{y}{\varepsilon},\omega\right) dy, \tag{2.48}$$

for some nice integration kernel $G(x,y)$. The limit theorems so far are enough to investigate the limit of $\langle V(x,\omega), \varphi(x)\rangle$ for proper test functions φ, because we can write this pairing as $\int_X q_\varepsilon(x) f(x)$ for $f = \int_X G(y,x)\varphi(y)dy$.

This type of weak products are random variables ($\mathbb{R}$-valued) which only contain integrated information of $V(x)$. Quite often, we can obtain better limit theorems of $V(x)$ as $\mathcal{S}$-valued random variables for some proper measure space $\mathcal{S}$ equipped with natural Borel σ-algebra. This belongs to the deep theory of weak convergence of probability measures on general measure spaces, which is beyond the scope of this dissertation. We only record two special cases in this theory that we will apply in the later chapters.

2.6.1 Convergence in distribution in $C([0,1])$

Proposition 2.36. *Suppose* $\{M_\varepsilon\}_{\varepsilon\in(0,1)}$ *is a family of random processes parametrized by* $\varepsilon \in (0,1)$ *with values in the space of continuous functions* $C([0,1])$ *and* $M_\varepsilon(0) = 0$. *Then* M_ε *converges in distribution to* M_0 *as* $\varepsilon \to 0$ *if the following holds:*

(i) (Finite-dimensional distributions) *for any* $0 \le x_1 \le \cdots \le x_k \le 1$, *the joint distribution of* $(M_\varepsilon(x_1), \cdots, M_\varepsilon(x_k))$ *converges to that of* $(M_0(x_1), \cdots, M_0(x_k))$ *as* $\varepsilon \to 0$.

(ii) (Tightness) *The family* $\{M_\varepsilon\}_{\varepsilon\in(0,1)}$ *is a tight sequence of random processes in* $C([0,1])$. *A sufficient condition is the Kolmogorov criterion:* $\exists \delta, \beta, C > 0$ *such that*

$$\mathbb{E}\left\{|M_\varepsilon(s) - M_\varepsilon(t)|^\beta\right\} \le C|t-s|^{1+\delta}, \tag{2.49}$$

uniformly in ε *and* $t, s \in (0,1)$.

For a proof, see for instance [72, p.64].

Remark 2.37. The standard Kolmogorov criterion for tightness requires the existence of $t \in [0,1]$ and some exponent ν so that $\sup_\varepsilon \mathbb{E}|M_\varepsilon(t)|^\nu \le C$ for C independent of ε and ν. In our cases, since $M_\varepsilon(0) = 0$ for all ε, this condition is always satisfied.

2.6.2 Convergence in distribution in Hilbert spaces

Proposition 2.38. *Suppose* $\{M_\varepsilon\}_{\varepsilon\in(0,1)}$ *is a family of random processes parametrized by* $\varepsilon \in (0,1)$ *with values in some separable Hilbert space* $\mathcal{H}$. *Let* $\{\phi_n \mid n = 1, 2\cdots\}$ *be an orthonormal basis of* $\mathcal{H}$ *and let* P_N *be the projection to the finite dimensional space spanned by* $\phi_1, \cdots, \phi_N$. *Then* M_ε *converges in distribution to* M_0 *as* $\varepsilon \to 0$ *if the following holds:*

(i) (Finite-dimensional distributions) *for any* $k \in \mathbb{N}$ *and any* k *basis functions* $\phi_{i_1}, \cdots, \phi_{i_k}$, *the joint distribution of* $(\langle M_\varepsilon, \phi_{i_1}\rangle \cdots, \langle M_\varepsilon, \phi_{i_k}\rangle)$ *converges to that of* $(\langle M_0, \phi_{i_1}\rangle, \cdots, \langle M_0, \phi_{i_k}\rangle)$ *as* $\varepsilon \to 0$.

(ii) (Tightness) *The family* $\{M_\varepsilon\}_{\varepsilon\in(0,1)}$ *is a tight sequence of random processes in* $\mathcal{H}$. *A sufficient condition is*

$$\sup_{\varepsilon\in[0,1]} \mathbb{E}\, \|M_\varepsilon\|_{\mathcal{H}}^2 < \infty, \tag{2.50}$$

and

$$\sup_{\varepsilon\in[0,1]} \mathbb{E}\, \|M_\varepsilon - P_N M_\varepsilon\|_{\mathcal{H}}^2 \xrightarrow{N\to\infty} 0. \tag{2.51}$$

This proposition follows from the definition of tightness of general probability measure on metric spaces, and the structure of separable Hilbert spaces; see [24, 83].

2.7 Appendix: Integrals Involving Two Scales

Very often in the homogenization and corrector theory, we need to deal with integrals involving variables of two scales. A typical example is the oscillatory integral

$$I_\varepsilon = \frac{1}{\varepsilon^d}\int_X R\left(\frac{x-y}{\varepsilon}\right) f(x)g(x)dxdy. \tag{2.52}$$

For such integrals, we have the following estimates.

Proposition 2.39. *Suppose that $R(x) \in L^1(\mathbb{R}^d)$, and $f \in L^p(X)$ and $g \in L^q(X)$ with (p,q) a Hölder pair. Then the above integral is uniformly bounded as follows:*

$$|I_\varepsilon| \le \|R\|_{L^1}\|f\|_{L^p}\|g\|_{L^q}. \tag{2.53}$$

Proof. We change variables:

$$\frac{x-y}{\varepsilon} \to y, \quad x \to x.$$

Then the integral becomes, with $\tilde{f}(x)$ denote $(f\mathbf{1}_A)(x)$ where $\mathbf{1}_A$ is the indicator function of set A,

$$\int_{\mathbb{R}^{2d}} R(y)\tilde{f}(x-\varepsilon y)\tilde{g}(x)dydx \le \int_{\mathbb{R}^d}|R(y)|\int_{\mathbb{R}^d}|\tilde{f}(x-\varepsilon y)\tilde{g}(x)|dxdy.$$

Note that the ε^d is cancelled by the Jacobian from the change of variable. The inner integral is bounded uniformly in y by $\|\tilde{f}\|\|g\|$, thanks to the Hölder inequality. Since $\|\tilde{f}\|_{L^p(\mathbb{R}^d)} = \|f\|_{L^p(X)}$, we have (2.53) above. □

2.8 Notes

Sections 2.2 and 2.3 Mixing properties of random fields is by definition very technical. We recommend the monograph of Doukhan [47] for a thorough discussion. For the superposition of Poisson bumps model, we are able to derive systematic moment formulas thanks to many nice properties of the underlying Poisson point process. A good review of general spatial point processes can be found in Cox and Isham [41].

Section 2.5 We commented that there is a hidden assumption $\mathbb{E}\{g(0)\Phi(g(0))\} \neq 0$ in Theorem 2.34. Recall that the random coefficient $q(x)$ is given by $\Phi(g(x))$, and we are interested in the limiting distribution of oscillatory integrals of the form $\varepsilon^{-\frac{\alpha}{2}} \int_X q\left(\frac{x}{\varepsilon}\right) f(x)dx$. By assumption, $\mathbb{E}\{\Phi(g(0))\} = 0$, which can be written as

$$\int_{\mathbb{R}} 1 \cdot \Phi(y) d^g y = 0,$$

where $d^g y$ is the standard Gaussian measure $\frac{1}{\sqrt{2\pi}} e^{-\frac{|y|^2}{2}} dy$. The hidden assumption above can be written as

$$\int_{\mathbb{R}} y\Phi(y) d^g y \neq 0.$$

Recall the Hermite polynomials $\{H_n\}_{n=1}^{\infty}$, which form an orthonormal basis of $L^2(\mathbb{R}, d^g y)$. In particular, the first two Hermite polynomials are $H_0(y) \equiv 1$ and $H_1(y) = y$. The above requirements can be rewritten as

$$\langle H_0, \Phi \rangle_{L^2(d^g)y} = 0, \quad \langle H_1, \Phi \rangle_{L^2(d^g y)} \neq 0.$$

Following the notation of the paper by Taqqu [107], for a function Φ in $L^2(\mathbb{R}, d^g y)$, we define its Hermite rank to be the smallest integer n such that $\langle H_n, \Phi \rangle_{L^2(d^g y)}$ is nonzero. Then the above requirement can be restated as: Φ *has Hermite rank one.*

When Φ has Hermite rank two, the limit in Theorem 2.34 degenerates and hence is not optimal. In fact, we can refine the argument in the proof of Lemma 2.28 and show that the correlation function $R(x)$ of the random field $\Phi(g(x))$ is asymptotically $\kappa'|x|^{-2\alpha}$ for some κ'. So, the integral $\int_{\mathbb{R}} q_\varepsilon(x)F(x)dx$ has variance of order $\varepsilon^{2\alpha}$. That is, the right scaling in the theorem should be $\varepsilon^{-\alpha}$ rather than $\varepsilon^{-\alpha/2}$.

In the case of $\alpha < 1/2$, Taqqu [106] showed that the limit under weak convergence of normalized partial sums of $\Phi(g_i)$, where g_i, $i = 1, 2, \cdots$ is a stationary Gaussian sequence with long-range correlation and Φ has Hermite rank two, is the Rosenblatt process. Applying his result to the problem considered in Theorem 2.34, we find that the denominator on the left hand side should be ε^α, and the limit on the right hand side should be an integral with respect to the Rosenblatt increment. As a result, the limit is non-Gaussian.

Chapter 3

The Linear Transport Equations

The transport equation arises in physics and engineering as a basic model for propagation of particles or energy density of certain waves. Various properties of this equation are well explored in the literature of mathematical physics, where it bears the name "Boltzmann equation". Here, we only consider its simplest form, namely, the stationary and linear transport equation, which finds applications in many areas of science, including neutron transport [45, 85], atmospheric science [32, 81], propagation of high frequency waves [7, 100, 101] and the propagation of photons in many medical imaging applications [5, 10].

The first section reviews the physical importance of the transport equation. Well-posedness is recalled in the second section, with an emphasis on the following fact: The solution operator of the transport equation is a bounded linear map on L^p space, and its operator norm does not depend on the bound of the optical parameters. This fact is crucial for corrector analysis of transport in random media which we address in the next chapter. In the third section we present several less discussed properties of the transport equations which, again, are tailored for corrector analysis in the next chapter.

3.1 Introduction

Transport equation models propagation of particles, such as neutrons and X-rays, or energy package of certain waves, like acoustic or elastic waves, in some background medium. In the steady state case, it takes the following form.

$$-v \cdot \nabla_x u(x,v) - a_r u - \int_V k(x,v',v)u(x,v)dv' + \int_V k(x,v,v')u(x,v')dv' = 0. \tag{3.1}$$

In the context of neutron transport, the unknown function $u(x,v)$ is the density of particles which are identified by their position $x \in X \subset \mathbb{R}^d$ and velocity $v \in V \subset \mathbb{R}^d$. Particles propagate through the medium with velocity v and meanwhile get absorbed at a rate of a_r which by assumption is *isotropic*, i.e., depending only on x. Their trajectories are straight lines except at places where they collide with nuclei of the background medium and get scattered into other directions. We denote by $k(x,v',v)$ the fraction of particles at x with velocity v being scattered to direction v'.

Now it is clear that the first term in (3.1) is the rate of "loss" of particles (x,v) due to "transport", i.e., particles propagating away; the second term is loss of particles due to absorption; the third term is "loss" of particles due to scattering to other directions, and the fourth term is "gain" or "creation" of particles due to scattering from other directions. With this picture in mind, the equation (3.1) is nothing but an expression of conservation of particles. However, one may notice that since the only real loss of particles is due to absorption, there should be no "conservation" or steady state. Indeed, to maintain such a steady state of particle transport, sources $g(x,v)$ are placed on the incoming boundary Γ_- which, along with exiting boundary Γ_+, is defined as follows.

$$\Gamma_\pm := \{(x,v)|x \in \partial X,\ \pm\nu_x \cdot v > 0\} \tag{3.2}$$

where ∂X is the boundary of X and is assumed to be C^2. The normal vector of $x \in \partial X$ is denoted by ν_x. In other words, Γ_- is the ensemble of boundary points with local velocities entering the physical domain X, while Γ_+ is the ensemble of boundary points with local velocities exiting the domain.

Observe that the third term on the left hand side of equation (3.1) can be written as $-a_s u$ with

$$a_s := \int_V k(x, v', v)dv'.$$

In the literature, it is convention to define $a := a_r + a_s$ which is called the total attenuation. Meanwhile, a_r is called the *real* or *intrinsic* attenuation and a_s is the attenuation due to scattering. Using these notations, we rewrite the transport equation with boundary conditions as follows:

$$\begin{aligned} & -v \cdot \nabla_x u(x,v) - au + \int_V k(x,v,v')u(x,v')dv' = 0, \quad && (x,v) \in X \times V, \\ & u(x,v) = g(x,v), && (x,v) \in \Gamma_- \end{aligned} \tag{3.3}$$

We point out that the equation is posed on the *phase space* $X \times V$ where X is an open bounded and convex subset of $\mathbb{R}^d$ which represents the physical domain of the background media, and V represents the domain of velocity, which for simplicity is assumed to be S^{d-1}; i.e., particles propagate with unit velocity.

The attenuation coefficient a and the scattering coefficients k are usually called the *optical* or *constitution* parameters of transport equation. In the future we will make further simplifications on them. When only consider the case of isotropic scattering in the next chapter, for instance.

3.2 Well-posedness of the Linear Transport Equation

The linear transport equation (3.3) is well posed provided that the optical parameters are *admissible* which we now define.

As the opposite of a_s, the attenuation due to scattering, the *creation* due to scattering is defined as

$$a_c(x,v) := \int_V k(x,v,v')dv'. \tag{3.4}$$

Note that $a_s = a_c$ when k is symmetric in v and v', i.e., $k(x,v,v') = k(x,v',v)$.

Definition 3.1. 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. $k(x,v,\cdot)$ is integrable for $a.e.$ (x,v) and $a_s \in L^\infty$.

We say the problem is *subcritical* if in addition

$$a_r = a - a_s \geq \beta > 0,\ a - a_c \geq \beta > 0. \tag{3.5}$$

for some real number $\beta > 0$. The physical importance of this condition is that the net creation is negative.

Next we review some fundamental theories of the transport equations equipped with interior sources and absorbing boundary conditions. We start by introducing the following standard notations.

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

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

We define the following Banach spaces tailored for the equations:

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

On these spaces we define the following differential or integro-differential operators:

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

The fact that a function in $\mathcal{W}^p$ has trace on $\Gamma_\pm$ is discussed in [44].

Now we consider a transport equation similar to (3.3) though equipped with interior sources $f(x, v)$ and absorbing condition $g \equiv 0$; we can write it in the following concise form.

$$\mathbf{T} u = f. \tag{3.6}$$

Note that we did not make the choice of p explicit but it should always be read from the context. For simplicity, one can assume we are in the $\mathcal{W}^1 \bigcap \mathcal{W}^\infty$ setting.

When k is zero everywhere, the above equation reduces to $\mathbf{T}_1 u = f$; this is called the *non-scattering* case or the *free* transport. It is a first order PDE and its solution is obtained by the method of characteristics and has the following explicit expression.

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

where $\tau_-(x, v)$ is the backward travel time to the boundary; together with the forward travle time $\tau_+(x, v)$, it is defined by

$$\tau_\pm(x, v) = \inf\{t > 0 | x \pm tv \in X^c\}.$$

The function $E(x, y)$ is the amount of attenuation between x and y. More specifically,

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

In the sequel, we also use $E(x, y, z) := E(x, y)E(y, z)$ which is the amount of attenuation along a broken line connecting x, y and z. Straightforward calculation verifies the following property of free transport.

Proposition 3.2. *Assume that the problem is subcritical with parameter β. Then the solution operator $\mathbf{T}_1^{-1}$ is bounded on L^p for all $p \in [1, \infty]$. In fact,*

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

where δ is the diameter of the domain X, i.e, $\delta := \sup\{d(x, y)|x, y \in X\}$.

Proof. We show boundedness on L^p for $p = 1$ and $p = \infty$ respectively and the result follows from Riesz-Thorin interpolation. Since $a \geq a_r \geq \beta$, we have

$$|\mathbf{T}_1^{-1} f| \leq \|f\|_{L^\infty} \int_0^{\tau_-} E(x, x - tv)dt \leq \|f\|_{L^\infty} \int_0^{\delta} e^{-t\beta} dt.$$

Calculate this integral and we get the desired estimate. This proves boundedness on L^∞. For the L^1 boundedness, we use the change of variable $x - tv \to y$ and observe that

$$\int_{X \times V} |\mathbf{T}_1^{-1} f| dxdv \leq \int_{X \times V} |f(y, v)| \int_0^{\tau_+(y,v)} E(y + tv, y)dtdydv \leq \|f\|_{L^1} \int_0^{\delta} e^{-t\beta} dt.$$

This completes the proof of the L^1 setting and hence proves the proposition. □

Remark 3.3. We reiterate that the bound does not depend on $\|a\|_{L^\infty}$. □

For the equation with scattering (3.6), existence and uniqueness results have been established as a perturbation to the non-scattering case, using either semigroup theory or integral equation technique. We review the latter method and show that the solution operator remains to be a linear transform on L^p; further, its operator norm can be bounded by a constant that depends only on the geometry of the domain and subcriticality parameter β. The fact that this bound does not depend on the L^∞ norm of a, k is important when we consider transport equations in random media. In the next chapter, we will introduce some natural random field models for these parameters, and their values can be arbitrarily large for different realizations.

An application of the method of characteristics converts the transport equation into the following integral equation:

$$(I+K)u = \mathbf{T}_1^{-1} f,$$

where

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

We define also the operator $\mathcal{K}$ as follows:

$$\begin{aligned}\mathcal{K}u := A_2 \mathbf{T}_1^{-1} u &= -\int_V \int_0^{\tau_-(x,v')} E(x, x-tv') k(x, v, v') u(x-tv', v') dt dv' \\ &= -\int_X \frac{E(x,y)k(x,v,v')}{|x-y|^{d-1}} u(y, v') dy\end{aligned} \tag{3.7}$$

with $v' = (x-y)/|x-y|$ above.

The following theorem extends the conclusion of Proposition 3.2 to the solution operator of the full linear transport equation. It benefits from the subcritical condition of parameters in a fascinating manner, which allows us to show that K and $\mathcal{K}$ above are bounded linear operators on L^∞ and L^1 respectively.

Theorem 3.4. *Suppose the coefficients (a, k) are admissible and satisfy the subcriticality condition (3.5). Then, the solution operator $\mathbf{T}^{-1}$ is a bounded linear transform on $L^p(X \times V)$ for all $p \in [1, \infty]$. In fact, we have*

$$\|\mathbf{T}^{-1} f\|_{L^p} \le C \|f\|_{L^p}$$

where C is a constant depends on p, δ, β but not on $\|a\|_{L^\infty}$ or $\|k\|_{L^\infty}$. Actually, we can choose $C = (e^{\beta\delta} - 1)/\beta$.

Proof. Again, we prove the cases for $p = 1$ and $p = \infty$ and use interpolation afterwards. In the L^∞ setting, we use the integral equation $(I+K)u = \mathbf{T}_1^{-1}f$. Our goal is to show that the operator norm $\|K\|_{\mathcal{L}(L^\infty)}$ is strictly less than one, so that we can write $\mathbf{T}^{-1} = (I+K)^{-1}\mathbf{T}_1^{-1}$. Indeed, recall the definition of a_c in (3.4) and the relation $a_c < a$; we have

$$\begin{aligned}|Kf(x,v)| &\le \|f\|_{L^\infty} \int_0^{\tau_-} e^{-\int_0^t a(x-sv)ds} \int_V k(x-tv,v,v')dv'dt \\ &\le \|f\|_{L^\infty} \int_0^{\tau_-} a(x-tv)e^{-\int_0^t a(x-sv)ds} dt.\end{aligned}$$

Now recognize the integrand as a total derivative; we have

$$\begin{aligned}|Kf(x,v)| &\le \|f\|_{L^\infty} \int_0^{\tau_-} -\frac{d}{dt} e^{-\int_0^t a(x-sv)ds} dt \\ &\le \|f\|_{L^\infty}(1-e^{-\beta\delta}).\end{aligned}$$

When $\beta > 0$, we verify that $\|K\|_{\mathcal{L}(L^\infty)} < 1$. Then $(I+K)^{-1}$ can be written as a Neumann series with bounded operator norm. In particular, we have

$$\|\mathbf{T}^{-1}\|_{\mathcal{L}(L^\infty)} \le \|(I+K)^{-1}\|_{\mathcal{L}(L^\infty)}\|\mathbf{T}_1^{-1}\|_{\mathcal{L}(L^\infty)} \le (1-(1-e^{-\beta\delta}))^{-1}\frac{1-e^{-\beta\delta}}{\beta} = \frac{e^{\beta\delta}-1}{\beta}.$$

For the L^1 setting, we first observe that

$$\mathbf{T} = \mathbf{T}_1 + A_2 = (I+\mathcal{K})\mathbf{T}_1.$$

Hence, $\mathbf{T}^{-1} = \mathbf{T}_1^{-1}(I+\mathcal{K})^{-1}$. It now suffices to show $\|\mathcal{K}\|_{\mathcal{L}(L^1)} < 1$ so that $(I+\mathcal{K})^{-1}$ indeed makes sense. This holds, again, thanks to the subcriticality condition in (3.5). By recognizing the definition of a_s in the following calculation, we have

$$\begin{aligned}\int_{X\times V} |\mathcal{K}f(x,v)|dxdv &\le \int_{X\times V}\int_V \int_0^{\tau_-(x,v')} k(x,v,v')E(x,x-tv')|f(x-tv',v')|dtdv'dvdx \\ &= \int_X\int_V\int_0^{\tau_-(x,v')} a_s(x,v')E(x,x-tv')|f(x-tv',v')|dtdv'dx.\end{aligned}$$

Now use the fact that $a_s < a$ and change variable $x - tv \to y$ to obtain

$$\begin{aligned}\|\mathcal{K}f\|_{L^1} &\le \int_{X\times V} |f(y,v')| \left(\int_0^{\tau_+(y,v')} a_s(y+tv',v') e^{-\int_0^t a(y+sv')ds} ds \right) dv'dy \\ &\le \int_{X\times V} |f(y,v')| \left(\int_0^{\tau_+(y,v')} -\frac{d}{dt} e^{-\int_0^t a(y+sv')ds} dt \right) dv'dy \le (1-e^{-\beta\delta})\|f\|_{L^1}.\end{aligned}$$

Since $\beta > 0$, we verify that $\|\mathcal{K}\|_{\mathcal{L}(L^1)} < 1$. Hence, as before we have

$$\|\mathbf{T}^{-1}\|_{\mathcal{L}(L^1)} \le \|\mathbf{T}_1^{-1}\|_{\mathcal{L}(L^1)} \|(I+\mathcal{K})^{-1}\|_{L^1\to L^1} \le \delta.$$

Application of Riesz-Thorin completes the proof. □

Remark 3.5. The proof shows that K is suitable in the L^∞ setting while $\mathcal{K}$ is for the L^1 setting. Nevertheless, both $(I+K)^{-1}$ and $(I+\mathcal{K})^{-1}$ are well defined in the $L^1 \bigcap L^\infty$ setting. This is seen from the algebraic relation

$$(I+K)^{-1} = I - \mathbf{T}_1^{-1}(I+\mathcal{K})^{-1}A_2, \qquad (I+\mathcal{K})^{-1} = I - A_2(I+K)^{-1}\mathbf{T}_1^{-1}, \tag{3.8}$$

and the boundedness of $\mathbf{T}_1^{-1}$ and A_2 in both settings. □

Corollary 3.6. *Suppose that (a,k) are continuous functions on X and the boundary ∂X is C^2. Let $f \in C(X)$ be a continuous source in X. Suppose further that either*

(i) *f and k are compactly supported in X or*

(ii) *the curvature of ∂X is bounded from below by γ.*

Then the transport solution $\mathbf{T}^{-1}f$ is also continuous.

Proof. 1. We first show that $\mathbf{T}_1^{-1}$ maps continuous function on X to continuous functions on $X \times V$. This is done by showing continuity in x and v separately.

Fix any $x \in X$ and $v, v' \in \mathrm{S}^{d-1}$. The functions $\mathbf{T}_1^{-1}f(x,v)$ and $\mathbf{T}_1^{-1}f(x,v')$ depend on the backward characteristics traced back from x in direction v and v' respectively. Set

$L = \mathrm{diam}(X)$. For any $\varepsilon > 0$, there exists a δ' such that

$$|a(y) - a(z)| < \varepsilon/4L^2, \quad |f(y) - f(z)| < \varepsilon/4L, \text{ if } |y - z| \le \delta',$$

since a and f are continuous, Let $\delta = \delta'/2L$ and $\tau_- = \min(\tau_-(x,v), \tau_-(x,v'))$. The two backward characteristics can be parametrized by the same "time" variable up to τ_-. With this parametrization, we have

$$\mathbf{T}_1^{-1}f(x,v) - \mathbf{T}_1^{-1}f(x,v') = \int_0^{\tau_-} e^{-\int_0^t a(x-sv)ds} f(x-tv) - e^{-\int_0^t a(x-sv')ds} f(x-tv')dt + e,$$

for some error term e. The first part involves a and f on a cone which is at most $|v - v'|L$ apart. Hence when $|v - v'| \le \delta$, the first term is bounded by $\varepsilon/2$, thanks to the continuity of f and a. The error term e involves integration near the boundary. It can be shown negligible, but the argument is considerably technical and is postponed for now.

2. Fix a direction $v \in \mathrm{S}^{d-1}$ and consider two points x and y. Again we can parametrize the characteristics for $\mathbf{T}_1^{-1}f(x,v)$ and $\mathbf{T}_1^{-1}f(y,v)$ using a "synchronized" time except an extra error term accounting for the situation near the boundary. The analysis is essentially the same as above.

3. Now let us show that the error term e in the last two steps are small. If f and k are compactly supported in X, they are necessarily small near the boundary, then e will be small.

Even in the case when f and a does not vanish on the boundary, e can be controlled provided that the geometry of X is nice. For instance when $X = B(0,R)$ and hence $\partial X = \mathrm{S}(R)$, the difference between the lengths of any two parallel straight line segments inside X, with δ being the distance between these lines, is smaller than $2\sqrt{2R\delta - \delta^2}$. (This supremum is "achieved" in the limiting case when one of the lines shrinks to a point on the boundary). Similarly, for any two lines (x,v) and (x,v') with v and v' apart with a small angle δ, the part that they don't overlap has length less than $2R\delta$. For a general convex domain X with smooth boundary, we can find a map between X and $\mathrm{S}(R)$ with controlled derivatives. We ignore this technicality here.

4. The fact that A_2 maps continuous function on $X \times V$ to continuous function on X is trivial and hence omitted.

5. Now by definition $K = \mathbf{T}_1^{-1} A_2$ maps continuous functions to continuous functions on $X \times V$, so does K^j for any j, and $\sum_{j=1}^{M} K^j$ for any finite M.

Since $(I+K)^{-1} = \lim_{M\to\infty} \sum_{j=1}^{M} K^j$ and the convergence is in the Banach space $C(X \times V)$ equipped with the uniform norm, we conclude that $(I+K)^{-1}$ maps continuous function to continuous functions as well.

6. Now recall the relation $\mathbf{T}^{-1} = (I+K)^{-1}\mathbf{T}_1^{-1}$ to complete the proof. $\square$

Definition 3.7. According to Theorem 3.4, the solution operator of transport equation (3.6) can be written as the following Neumann series:

$$\mathbf{T}^{-1} = \mathbf{T}_1^{-1} - K\mathbf{T}_1^{-1} + K^2\mathbf{T}_1^{-1} - K^3\mathbf{T}_1^{-1} + \cdots . \tag{3.9}$$

The first term, which is the same as free transport solution, is called the *ballistic* part. The second term, linear in scattering coefficient, is called the *first scattering* part. In general, the term that involves the p scattering is called the *pth scattering* part.

Theorem 3.4 allows us to control of the solutions of (3.6) by the L^p norm of the sources f. Meanwhile, we observe that the transport solution is not smoothing. This is expected somehow since the ballistic part is literally transport and no mixing happens there; In contrast, the scattering process is more or less diffusive and turns out to be smoothing. This will be one of our main focus in the next section.

3.3 Further properties of transport equations

We develop some further properties of the transport equation which, though less discussed in the literature, are interesting in nature and quite helpful in simplifying our analysis in the next chapter.

So far we have used a very general scattering kernel $k(x, v, v')$. In the rest of this section, the scattering kernel is assumed to be isotropic. That is, $k(x, v, v') = k(x)$. In this case, $a_s = a_c = \varpi_d k(x)$ where ϖ_d is the volume of the unit sphere S^{d-1}.

3.3.1 Boundedness of multiple scattering kernel

The smoothing property of scattering is probably best seen from the singular decomposition of the Schwartz kernel of the operator $\mathbf{T}^{-1}$; see the work of Choulli and Stefanov, Bal and Jollivet [15, 37] for instance. The kernel of the ballistic part can be seen as a Dirac measure type distribution focused on a line; the first scattering part is smoother though still singular, and can be seen as a Dirac measure type distribution centered on a plane; multiple scattering is even smoother and admits a kernel that is a function. For instance it is shown in the cited papers that scattering of order equal or higher than three admits a kernel that is in $L^\infty(X \times V, L^1(X \times V))$ in quite general settings.

In the case of isotropic optical parameters we aim to show that multiple scattering of order $d+1$ and up, i.e., $\sum_{j=d+1}^{\infty}(-1)^j \mathcal{K}^j$ has bounded kernel. As one can imagine, this type of result will simplify our analysis of $\mathbf{T}^{-1}$ greatly since it allows us to combine all the high scatterings together and to avoid dealing with the infinite Neumann series (3.9).

We group all the scatterings of order $d+1$ and up together as $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$ and denote it as $\mathcal{L}$. The goal is to show that the Schwartz kernel of $\mathcal{L}$ is a bounded function. There are at least two natural approaches to achieve this end.

The first idea is to use the fact that $\mathcal{K}$, as seen from (3.7), is a *weakly singular integral operator* of order $d-1$. Therefore it is attempting to conclude that $\mathcal{K}$ maps Sobolev space

H^s to H^{s+1}. Suppose this is true, then $\mathcal{K}^m$ for $m > d/2$ will map $H^0 \equiv L^2$ to $H^{\frac{d}{2}+(m-\frac{d}{2})}$ which is contained in the Hölder space $C^{0,m-\frac{d}{2}}$ by Morrey's lemma. Since $(I+\mathcal{K})^{-1}$ is bounded on H^0, we conclude that $\mathcal{L}$, which can be written as $\mathcal{K}^m(I+\mathcal{K})^{-1}$, maps H^0 to C, the space of continuous functions. By duality it also maps bounded measure to H^0. Hence choosing $m = (d+1)/2$, we conclude that $\mathcal{L}$ maps bounded measure to C. It follows immediately by duality that the Schwartz kernel of $\mathcal{L}$, denoted by $\alpha(x,y)$, is a bounded function. That is,

$$\alpha(x,y) = \langle \mathcal{L}\delta_x, \delta_y \rangle \leq M,$$

for some constant M.

The difficulty of this approach lies in the inhomogeneity of the kernel of $\mathcal{K}$, which has the following expression:

$$\kappa(x,y) := \frac{k(x)E(x,y)}{|x-y|^{d-1}}. \tag{3.10}$$

It is clear that $E(x,y)$ is not *homogeneous* in the variable $x-y$. Homogeneity of kernel is assumed in classical theories on singular integral operators. Nevertheless, when the optical parameters are assumed to be smooth it is possible to show similar smoothing properties of $\mathcal{K}$, say, $\mathcal{K}$ mapping H^s to H^{s+1} for $s=0$ provided that the optical parameters are C^2. It gets more and more difficult in this approach for larger s. We encourage the interested readers to work on the problem.

The second approach is to derive a Schwartz kernel theorem type of result. The idea is the following: We first show, by a duality argument, that for each x there exists a kernel function $g_x(y)$ for the linear transform $\mathcal{L}$ pinned at x. Then we show that $g_x(y)$ is the kernel of $\mathcal{L}$ and is uniformly bounded. To take care of measure-theoretic difficulties, we consider the case when $\mathcal{L}$ maps continuous functions to continuous functions. The following lemma indicates that this is not a very severe constraint. For now, we assume that $\mathcal{K}$ operates on functions that depend only on the spatial variable and therefore the kernel of $\mathcal{K}$ is given by

(3.10). We assume (a, k) are continuous functions up to boundary of X. Therefore we may assume X is compact in the following.

We show that $\mathcal{K}$ maps L^p functions to Hölder continuous functions for sufficient large p.

Lemma 3.8. *Let X be a convex compact subset in $\mathbb{R}^d$ with smooth boundary. Let the optical parameters (a, k) be Lipschitz continuous functions on X. Then we have that*

$$\|\mathcal{K}f\|_{C^{0,1-\frac{d}{p}}} \le C\|f\|_{L^p}, \tag{3.11}$$

provided that $d < p < \infty$. The constant C depends only on Lipschitz continuity coefficient of a and k, the index p, the dimensionality d, and diameter of X.

Proof. 1. We assumed that k is isotropic and Lipschitz continuous. Recall the expression of the kernel of $\mathcal{K}$ in (3.10). Since Hölder space $C^{0,1-\frac{d}{p}}$ forms an algebra, we may set $k \equiv 1$ without loss of generality.

2. Also from boundedness of X we see that $\mathcal{K}f$ is bounded. Hence we only need to quantize the modulus of continuity for $\mathcal{K}f$. We have

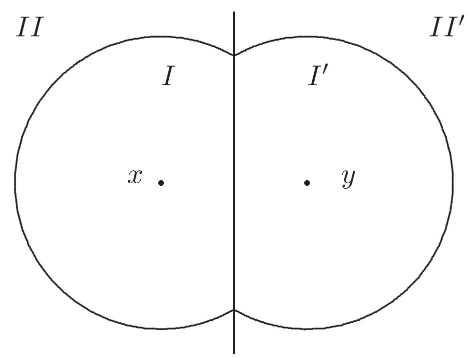
$$\mathcal{K}f(x) - \mathcal{K}f(y) = \int_X \Big(\frac{E(x,z)}{|x-z|^{d-1}} - \frac{E(y,z)}{|y-z|^{d-1}} \Big) f(z) dz.$$

Set $\rho = |x - y|$; the spheres centered at x and y respectively break the integration region into the following parts as shown in Fig. 3.1.

3. When z is in II, we use the following.

$$\begin{aligned}\frac{E(x,z)}{|x-z|^{d-1}} - \frac{E(y,z)}{|y-z|^{d-1}} = E(x,z)\Big(&\frac{1}{|x-z|^{d-1}} - \frac{1}{|y-z|^{d-1}}\Big) \\ &+ \frac{E(x,z) - E(y,z)}{|y-z|^{d-1}} =: I_1 + I_2.\end{aligned}$$

Figure 3.1: Integration region for $\mathcal{K}$.



For the first term we recall the equality that

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}),$$

and the fact that $|x-z|/2 \le |y-z| \le 2|x-z|$ on this region; the conclusion is that I_1 is bounded by $C\rho|x-z|^{-d}$.

For the second term, we recall the Taylor expansion for exponential function:

$$e^x - e^y = e^y(e^{x-y} - 1) = e^y\left(x - y + \frac{1}{2}e^{s(x-y)}(x-y)^2\right),$$

for some $s \in [0,1]$, and boundedness of E terms. It follows then

$$|I_2| \le C\Big|\int_0^{|x-z|} a(x - t\frac{x-z}{|x-z|})dt - \int_0^{|y-z|} a(y - t\frac{y-z}{|y-z|})dt\Big| =: \triangle_E(x,y).$$

We hence need to consider difference of two path integrals. Observe that we can synchronize the parametrization of the two paths by setting z the starting point. Since z is in domain

II for the moment, we have $|x-z|<|y-z|$. Then $\triangle_E(x,y)$ can be written as

$$\int_0^{|y-z|} a(z+t\hat{v}_1)-a(z+t\hat{v}_2)dt+\int_{|x-z|}^{|y-z|} a(z+t\hat{v}_1)dt.$$

Here we set $\hat{v}_1$ to be the direction vector $(x-z)/|x-z|$ and $\hat{v}_2$ the direction between z and y. Observe that the integration path of the second term is shorter than $|x-y|$. Since $\|a\|_{L^\infty}$ is finite, we see this term is bounded by $C\rho$.

For the first integral, we observe that $|z+t\hat{v}_1-(z+t\hat{v}_2)|\le\rho$ since the end points of the truncated paths are inside both spheres. Since a is Lipschitz continuous, we have for each $t\in[0,|y-z|]$,

$$\left|a(z+t\hat{v}_1)-a(z+\hat{v}_2)\right|\le \mathrm{Lip}(a)\rho.$$

In summary, we have that $I_1+I_2\le C\rho(|y-z|^{-d}+|y-z|^{-d+1})$, bounded by $C\rho|y-z|^{-d}$.

We next obtain that

$$\left|\int_{II}\frac{E(x,z)}{|x-z|^{d-1}}-\frac{E(y,z)}{|y-z|^{d-1}}f(z)dz\right|\le\int_{II}\frac{C\rho|f(z)|}{|y-z|^d}dz\le C\rho\|f\|_{L^p}\left\|\frac{1}{|y-z|^d}\right\|_{L^{p'}(II)}.$$

The last term can be estimated by

$$C\left(\int_\rho^{\mathrm{diam}(X)}\frac{1}{r^{dp'}}r^{d-1}dr\right)^{\frac{1}{p'}}\le C(1+\rho^{d(1-p')/p'})=C(1+\rho^{-\frac{d}{p}}).$$

Hence integration on region II has a contribution of size $\rho^{1-\frac{d}{p}}$. By symmetry, we have the same estimate on region II'. Note that $p>d$ is needed so that $1-d/p>0$ and $p<\infty$ is needed to integrate $|y-z|^{-dp'}$.

4. When $z\in I\bigcup I'$, we can bound the integral by

$$\int_{I\cup I'}\left(\frac{1}{|x-z|^{d-1}}+\frac{1}{|y-z|^{d-1}}\right)f(z)dz.$$

Carry out these integrals as in the last step, we find a bound of the form

$$C\|f\|_{L^p}\left(\int_0^\rho \frac{r^{d-1}}{r^{(d-1)p'}}\right)^{\frac{1}{p'}} \leq C\|f\|_{L^p}\rho^{(d-(d-1)p')/p'} = C\|f\|_{L^p}\rho^{1-\frac{d}{p}}.$$

Now combine step 3 and 4 to complete the proof. □

Theorem 3.9. *Let the coefficients (a,k) be subcritical and Lipschitz continuous on $\overline{X}$; then the operator $(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}$ admits a Schwartz kernel that is a bounded function. That is to say,*

$$(I+\mathcal{K})^{-1}\mathcal{K}^{d+1}f(x) = \int \alpha(x,y)f(y)dy. \tag{3.12}$$

Moreover, $\|\alpha(x,y)\|_{L^\infty(X\times X)} \leq C_0$ for some constant C_0 defined in (3.14) below and depends on size of X, β and $\|k\|_{L^\infty}$.

Proof. Since we assume that the optical parameters and sources are continuous up to boundary, we may treat X as $\overline{X}$ in the following.

1. First we observe that the kernel of $\mathcal{K}^{d+1}$ is given by:

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

Thanks to the convolution lemma 3.11, this function is bounded.

2. Now fix any two functions $\phi \in L^1(X)$ and $\psi \in L^1(X)$, and observe that $\mathcal{L}$ can be written as $\mathcal{K}^{d+1}(I+\mathcal{K})^{-1}$; then

$$\langle \mathcal{L}\phi, \psi\rangle := \int_X \mathcal{L}\phi(x)\psi(x)dx = \int_{X^2} \kappa_{d+1}(x,y)(I+\mathcal{K})^{-1}\phi(y)\psi(x)d[xy].$$

Since κ_{d+1} is uniformly bounded, we can pull it out. The last integral is then separated in variables x and y. Recall that $(I+\mathcal{K})^{-1}$ is a bounded operator; we have

$$|\langle \mathcal{L}\phi, \psi\rangle| \leq \|\kappa_{d+1}\|_{L^\infty}\|(I+\mathcal{K})^{-1}\|_{\mathcal{L}(L^1)}\|\phi\|_{L^1}\|\psi\|_{L^1}. \tag{3.13}$$

Denote by C_0 the constant appearing on the right hand side above. We conclude that:

$$\|\mathcal{L}\phi\|_{L^\infty} \leq C_0 \|\phi\|_{L^1}. \tag{3.14}$$

In particular, this inequality holds for all continuous functions in $C(X)$.

3. Since the optical parameters are continuous, $\mathcal{K}$ maps continuous functions to continuous functions due to Lemma 3.8. Therefore $\mathcal{L}\phi$ is continuous when ϕ is continuous. Hence, for each $x \in X$ we can define a linear functional on $C(X)$ by setting $\mathcal{L}_x(\phi) := (\mathcal{L}\phi)(x)$. The estimate (3.14) shows that this functional is continuous.

Since $C(X)$ is dense in L^1, by the bounded linear transformation theorem, $\mathcal{L}_x$ extends to a continuous functional on $L^1(X)$. Now since the dual space of L^1 is L^∞, there exists a function $g_x(y) \in L^\infty(X_y)$ such that $\mathcal{L}_x(f) = \langle g_x(\cdot), f\rangle$ for any $f \in L^1(X)$. Furthermore, $\|g_x(y)\|_{L^\infty} \leq C_0$. Indeed if otherwise the set $\{|g_x(y)| > C_0\}$ has positive measure, then we construct $\psi(y)$ as the normalized indicator function of this set with opposite sign of g and then (3.13) will be violated.

4. Now define $\alpha(x, y) := g_x(y)$. We verify that (3.12) holds. Since the bound on g_x is independent of x, this completes the proof. □

An immediate corollary is the following.

Corollary 3.10. *Under the same condition as above, we have the following decomposition.*

$$\mathbf{T}^{-1} f = \mathbf{T}_1^{-1}(f - \mathcal{K}f + \tilde{\mathcal{K}}\mathcal{K}f) \tag{3.15}$$

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

$$\mathbf{T}^{-1} f = \mathbf{T}_1^{-1} f - \mathbf{T}_1^{-1}\mathcal{K}f + \mathbf{T}_1^{-1}\mathcal{K}^2 f + \cdots + (-1)^d \mathbf{T}_1^{-1}\mathcal{L}\mathcal{K}f, \tag{3.16}$$

and define

$$\tilde{\mathcal{K}} := \sum_{j=1}^{d} (-1)^{j-1}\mathcal{K}^j + (-1)^d \mathcal{L}. \tag{3.17}$$

The lemma shows that $\mathcal{L}$ admits a bounded kernel. For all the $\mathcal{K}_j$ with $j = 1, \cdots, d$, their Schwartz kernels are explicit as and can be estimated using the convolution Lemma 3.11. They are all bounded by $C|x-y|^{1-d}$, and hence so is that of $\tilde{\mathcal{K}}$. Finally we verify that (3.15) holds. $\square$

3.3.2 The adjoint transport equation

We conclude this section by introducing the *adjoint* of $\mathbf{T}$ which we denote by $\mathbf{T}^*$.

$$\mathbf{T}^* u = -T_0 u + A_1 u + A_2' u, \; D(\mathbf{T}^*) = \{u \in \mathcal{W}^p, u|_{\Gamma_+} = 0\}. \tag{3.18}$$

Here A_2' is of the same form as A_2 with v and v' swapped in the function k. When k is assumed to be isotropic, then $A_2' = A_2$. We obtain similar expressions for $\mathbf{T}_1^{*-1}$, K^* and $\mathcal{K}^*$.

Consider the adjoint transport equation

$$\mathbf{T}^* u = f. \tag{3.19}$$

Our definition of subcriticality ensures that the well-posedness of $\mathbf{T}^*$ is exact as that of $\mathbf{T}$. The operator $\mathbf{T}^{*-1}$ shares similar estimates developed in this section. In particular, under the same subcritical condition as before, we have that $\mathbf{T}_1^{*-1}$ and $\mathbf{T}_1^{*-1}$ 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 conjugate, we find that:

$$\langle u, \mathbf{T}^{-1} w \rangle = \langle \mathbf{T}^{*-1} u, w \rangle. \tag{3.20}$$

3.4 Appendix: Convolution of Potentials

We introduce a pair of lemmas which provide estimates of convolution of *potentials* by which we mean functions on $\mathbb{R}^d$ of the type

$$f : \mathbb{R}^d \to \mathbb{R}, \quad f(x) = \frac{1}{|x|^\alpha}, \ \alpha \in (0, d). \tag{3.21}$$

Since α is positive, this function blows up at the origin and hence is *singular*; nevertheless, $\alpha < d$ implies that this function is locally integrable. Analysis of potentials is one of the main themes in *singular integral operator* theory.

a. Convolution of two potentials on a bounded domain

We start with convolution of potentials on a bounded domain. It is clear that the order of the product of two potentials is the sum of the orders; the following lemma says the order of convolution of two potentials is less than the sum.

Lemma 3.11. *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*

$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} dz \leq C \frac{1}{|x-y|^{\alpha+\beta-d}} \tag{3.22}$$

2. *If $\alpha + \beta = d$, then*

$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} dz \leq C(|\log|x-y|| + 1) \tag{3.23}$$

3. *If $\alpha + \beta < d$, then*

$$\int_X \frac{1}{|z-x|^\alpha} \cdot \frac{1}{|z-y|^\beta} dz \leq C. \tag{3.24}$$

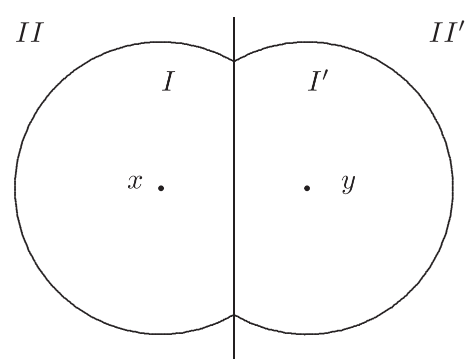
The convolution of logarithms with a weak singular potential turns out to be finite as follows:

$$\int_X |\log|z-x|| \frac{1}{|z-y|^\alpha} dz \lesssim 1. \tag{3.25}$$

The above constants depend only on the diam (X) *and dimension d but not on $|x-y|$.*

Proof. Let $\rho = |x-y|$. Let the S_x, S_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 I and I' denote the two parts respectively and II and II' denote the remaining part in X; see Fig. 3.2.

Figure 3.2: Integration region of the convolution of two potentials.



On I, $|z-y| \geq \rho$, hence

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

Similarly we have the same conclusion on I'. On $II \cup II'$, it is clear that $|z-x|/2 \leq |z-y| \leq$

$2|z-x|$, and hence we have

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

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. □

b. Convolution of two potentials on $\mathbb{R}^d$

Now we want to "extend" the result to convolution of potentials on the whole Euclidean space $\mathbb{R}^d$. This clearly will not work for f defined above since it is not integrable on the whole space. Therefore we only consider $e^{-\lambda|x|}f$. It is not hard to see that the exponential function can be replaced by any radial function which is bounded, monotone decreasing for large $|x|$ and integrable on the whole space.

The following lemma asserts that convolution of damped potentials preserves the fast decay and integrability at infinity while the order of its singularity at the origin is still one less than the sum of orders.

Lemma 3.12. *Fix two distinct points $x, y \in \mathbb{R}^d$. Let α, β be positive numbers in $(0,d)$, and*

λ another positive number. We have the following convolution results.

$$\int_{\mathbb{R}^d} \frac{e^{-\lambda|z-x|}}{|z-x|^\alpha} \frac{e^{-\lambda|z-y|}}{|z-y|^\beta} dz \leq \begin{cases} Ce^{-\lambda|x-y|}|x-y|^{d-(\alpha+\beta)}, & \text{if } \alpha+\beta > d; \\ Ce^{-\lambda|x-y|}(|\log|x-y|\mathbf{1}_{\{|y-x|<1\}}|+1), & \text{if } \alpha+\beta = d; \\ Ce^{-\lambda|x-y|} & \text{if } \alpha+\beta = d. \end{cases} \tag{3.26}$$

Here, $\mathbf{1}$ is the indicator function. Similarly, we also have that

$$\int_{\mathbb{R}^d} \frac{e^{-\lambda|z-x|}}{|z-x|^\alpha} e^{-\lambda|z-y|}|\log|z-y||dz \leq Ce^{-\lambda|x-y|}. \tag{3.27}$$

The above constants depend only on the diam (X) *and dimension d but not on $|x-y|$.*

Proof. It suffices to slightly modify the proof of the previous lemma. Still using the partition of integration domain as shown in Fig. 3.2. On I and similarly on I', we use $|z-x|+|z-y| \geq |x-y|$, and define $\rho = |x-y|$. Then we have

$$\int_I \frac{e^{-\lambda|z-x|}}{|z-x|^\alpha} \frac{e^{-\lambda|z-y|}}{|z-y|^\beta} dz \leq \frac{\pi_d e^{-\lambda|x-y|}}{\rho^\beta} \int_0^\rho \frac{r^{d-1}}{r^\alpha} dr.$$

The last integral can be calculated explicitly and yields $\rho^{d-\alpha}/(d-\alpha)$. Hence the integration over $I \cup I'$ can be bounded by

$$\frac{(2d-\alpha-\beta)\pi_d e^{-\lambda|x-y|}}{(d-\alpha)(d-\beta)|x-y|^{\alpha+\beta-d}}. \tag{3.28}$$

Now on the unbounded domain II, we observe that $|z-y| > \rho$ and $|z-y| > |z-x|$, and similar relations on II'. Therefore the integration on $II \cup II'$ is bounded from above by

$$2e^{-\lambda|x-y|} \int_{II \cup II'} \frac{e^{-\lambda|z-x|}}{|z-x|^{\alpha+\beta}} dz \leq 4\pi_d e^{-\lambda|x-y|} \int_\rho^\infty \frac{e^{-\lambda r}}{r^{\alpha+\beta-d+1}} dr.$$

Now we estimate the last integral which we call $A(\rho)$. When $\alpha+\beta < d$, the integrand is integrable over $\mathbb{R}_+$, the nonnegative real line; therefore $A(\rho)$ is bounded by some constant, actually a multiple of $\Gamma(d-\alpha-\beta)$. This together with the bound (3.28) proves the third case in (3.26).

Now we consider the case $\alpha+\beta=d$. If $\rho=|x-y|>1$, then $A(\rho)$ is bounded from above by $e^{-\lambda}/\lambda$. If $\rho\le 1$, then an integration by parts yields

$$\int_\rho^\infty = -e^{-\lambda\rho}\log\rho + \lambda\int_\rho^\infty e^{-\lambda r}\log r dr.$$

The last integral is finite over $\mathbb{R}_+$ and hence $|A(\rho)|\le Ce^{-\lambda\rho}(1+|\log\rho|)$. This together with the bound (3.28) proves the second case in (3.26).

When $\alpha+\beta>d$, let us denote $-\alpha-\beta+d-1=s$. Several iterations of integration by parts yield

$$\begin{aligned}A(\rho) = \int_\rho^\infty \frac{e^{-\lambda r^s}}{r}dr &= \frac{\lambda^\gamma}{\prod_{j=1}^\gamma(s+j)}\int_\rho^\infty \frac{e^{-\lambda}}{r^{s+\gamma}}dr\\ -e^{-\lambda\rho}\Big(\frac{\rho^{s+1}}{s+1} &+ \frac{\lambda\rho^{s+2}}{(s+1)(s+2)}+\cdots+\frac{\lambda^{\gamma-1}\rho^{s+\gamma}}{(s+1)\cdots(s+\gamma)}\Big).\end{aligned}$$

Here γ is the largest integer that is smaller than or equal to $\alpha+\beta-d$. When they are equal the right hand side above need some slight modification and the first integral involves a logarithm function. In both cases, the first integrable is finite and the second term is bounded by $Ce^{-\lambda\rho}(1+\rho^{d-\alpha-\beta})$. This together with the bound (3.28) proves the second case in (3.26).

The claim (3.27) follows from a similar and easier analysis which we omit. □

3.5 Notes

Section 3.2 For the mathematical formulation of the transport equation, we recommend the classic volumes of Dautray and Lions [43, 44], which contains the fundamental theories for very general transport equations with several interesting approaches. A nice probability representation is given by Bensoussan, Lions and Papanicolaou [21].

Section 3.3 The singular decomposition of the solution operator plays an important rule in inverse problems, and is investigated in e.g. Choulli and Stefanov [37], Bal and Jollivet [15, 16]. In Section 3.3.1, we discussed briefly the idea of proving the smoothing property of $\mathcal{K}$ using singular integral operator theory. I tried this approach first during my research, without full success. The $L^2 \to H^1$ smoothing of multiple scattering was proved by Stefanov and Uhlmann [102] using this approach. The procedure there was technical already due to the inhomogeneity of the kernel of $\mathcal{K}$. To show $L^2 \to H^{(d+1)/2}$ smoothing is presumably much more complicated. Therefore, I stopped and adopted the more practical approach taking full advantage of the explicit solution of the transport equation. This theory of singular integral operators itself, however, is mathematical beautiful with enormous practical importance. It was explored in detail in the monograph of Mikhlin and Prössdorf [84], and the classic books of Stein [103, 104]. Another regularity result for transport is by Golse, Lions, Perthame and Sentis [62].

Chapter 4

Corrector Theory in Random Homogenization of the Linear Transport Equation

This chapter concerns the corrector theory in random homogenization of the linear transport equation. 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. [48, 76]; see also e.g. [2, 21] for homogenization of transport in the periodic case, and [82] for the nonlinear case.

In this chapter, 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. [9, 59]. The results are shown for a specific structure of the random coefficients based on a Poisson point process. The resulting random coefficients then have 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 [48] 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 we have seen in Chapter 1.

4.1 Linear Random Transport Equation

In the previous chapter, we have reviewed some general properties of the linear transport equation. 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 desirable to model such coefficients as random [48, 76].

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:

$$\begin{aligned} &v \cdot \nabla_x u_\varepsilon + a_\varepsilon(x,\omega) u_\varepsilon - \int_V k_\varepsilon(x,v',v;\omega) u_\varepsilon(x,v') dv' = 0, \quad && (x,v) \in X \times V, \\ &u_\varepsilon(x,v) = g(x,v), && (x,v) \in \Gamma_- \end{aligned} \tag{4.1}$$

Here X is an open, convex, and 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 in (3.2).

The constitutive parameters in the transport equation are the total attenuation coefficient a_ε and the scattering coefficient k_ε. The small parameter $\varepsilon \ll 1$ models the scale of the heterogeneities that we want to model as random, typically because it corresponds to high frequency oscillations on which detailed information is not available. For example, in a PDE-based inverse problem, ε may model the spatial scale below which the parameters can no longer be estimated accurately. As was shown in e.g.[17], this high frequency part of the parameters still influences the reconstruction of the low frequency part. Modeling high frequency part as random then improves the statistical reconstruction of the low frequency part.

The above transport equation admits a unique solution in appropriate spaces [14, 44, 85] provided that these coefficients are non-negative and attenuation is larger than scattering (see Chapter 3). 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 have high frequency parts which are random fields constructed on an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. More precisely, we assume the intrinsic attenuation and scattering coefficients are constructed as follows:

$$\begin{aligned} a_{r\varepsilon}(x, \frac{x}{\varepsilon}, \omega) &:= a_{r0}(x) + \psi_Y\left(\frac{x}{\varepsilon}, \omega\right) = a_{r0}(x) + \sum_{j\in\mathbb{N}} \psi(\frac{x - y_j^\varepsilon}{\varepsilon}), \\ k_\varepsilon(x, \frac{x}{\varepsilon}, \omega) &:= k_0(x) + \varrho_Y\left(\frac{x}{\varepsilon}, \omega\right) = k_0(x) + \sum_{j\in\mathbb{N}} \varrho(\frac{x - y_j^\varepsilon}{\varepsilon}), \end{aligned} \tag{4.2}$$

where a_{r0} and k_0 are positive deterministic continuous functions and where ψ_Y and ϱ_Y are superposition of Poisson bumps with profile functions ψ and ϱ. We refer the reader to Chapter 2 for the definition of such random fields and their properties. 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 (4.1) 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

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

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

Note that the above random coefficients are bounded in the bounded domain 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; cf. Remark 2.23. 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 [14, 44, 85] of existence of unique solutions to (4.1) may be invoked $\mathbb{P}$-a.s. Before proceeding, let us introduce more notations which will be used throughout this chapter. 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 the probability measure $\mathbb{P}$. Set further $\delta a_{r\varepsilon} := a_{r\varepsilon} - \mathbb{E}\{a_{r\varepsilon}\}$, and $\delta k_\varepsilon = k_\varepsilon - k$. By construction, they are mean zero, stationary random fields. Then the autocorrelation function of $\delta a_{r\varepsilon}$, $R_{a\varepsilon}(x) = \mathbb{E}\{\delta a_{r\varepsilon}(y)\delta a_{r\varepsilon}(y+x)\}$, has the form

$$R_{a\varepsilon}(x) = R_a\left(\frac{x}{\varepsilon}\right), \quad \text{where} \quad R_a(x) = \nu \int_{\mathbb{R}^d} \psi(x-z)\psi(-z)dz,$$

as we have seen in (2.20). Similarly, we can define $R_{k\varepsilon}$, the autocorrelation function of δk_ε, and $R_{ak\varepsilon}$, the cross-correlation function of the two fields. Further, they can be written as $R_k\left(\frac{x}{\varepsilon}\right)$ and $R_{ak}\left(\frac{x}{\varepsilon}\right)$, respectively, where

$$R_k(x) = \nu \int_{\mathbb{R}^d} \varrho(x-z)\varrho(0-z)dz \quad \text{and} \quad R_{ak}(x) = \nu \int_{\mathbb{R}^d} \psi(x-z)\varrho(0-z)dz.$$

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

$$\begin{aligned} \sigma_a^2 &= \int_{\mathbb{R}^d} R_a(x)dx = \nu\Big(\int_{\mathbb{R}^d} \psi(x)dx\Big)^2, \\ \sigma_k^2 &= \int_{\mathbb{R}^d} R_k(x)dx = \nu\Big(\int_{\mathbb{R}^d} \psi(x)dx\Big)^2, \end{aligned} \tag{4.3}$$

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 (4.2) can be modified as in (4.23) below to yield $\rho_{ak} < 1$. For instance, if y_j^ε in the second line in (4.2) 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 (4.2) knowing that all results extend to more complex models such as (4.23) below.

We recall that higher order moments of random fields δk, the mean-zero part of (4.2), have explicit formulas:

$$\mathbb{E}\prod_{i=1}^{n}\delta k(x_i) = \sum_{(n_1,\cdots,n_k)\in\mathscr{G}_n}\sum_{\ell=1}^{C_n^{n_1,\cdots,n_k}}\prod_{j=1}^{k}T^{n_j}(x_1^{(\ell,n_j)},\cdots,x_{n_j}^{(\ell,n_j)}), \tag{4.4}$$

where the functions $T^{n_j}(\cdot)$ are defined as:

$$T^{n_j}(x_1,\cdots,x_{n_j}) := \nu\int\prod_{i=1}^{n_j}\psi(x_i - z)dz. \tag{4.5}$$

We refer to Section 2.3 for the details. Similarly, the higher moments of δa, and higher cross-moments of the two random fields all have similar formulas. Finally, the moments of the scaled random fields δk_ε, etc., are obtained by simple scaling. In particular, the fourth order moments of the random model, say δk_ε, is given by

$$\begin{aligned}\mathbb{E}\prod_{i=1}^{4}\delta k_\varepsilon(x_i) = \nu\int\prod_{i=1}^{4}\varrho(\frac{x_i}{\varepsilon} - z)dz + R_\varepsilon(x_1 - x_2)R_\varepsilon(x_3 - x_4)\\ + R_\varepsilon(x_2 - x_3)R_\varepsilon(x_1 - x_4) + R_\varepsilon(x_1 - x_3)R_\varepsilon(x_2 - x_4).\end{aligned} \tag{4.6}$$

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

We also recall the notations for transport equation in Chapter 3 that will be used intensively. 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

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

We also define $E(x,y,z) = E(x,y)E(y,z)$. We also denote by $\overline{f}$ the angular average (over v) for some function $f(x,v)$ defined on the phase space.

4.2 Homogenization Theory of Random Transport

Let us then define u_0 as the solution to (4.1) where a_ε and k_ε are replaced by their averages a and k, respectively. Then, consistently with the results shown in [48], 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 4.1 (Homogenization of Random Transport)**.** *Let dimension $d \geq 2$. Suppose that the random coefficients $a_\varepsilon, k_\varepsilon$ are constructed as in* (4.2) *and that a_{r0} is bounded from above by a positive constant β. Suppose also that $g \in L^\infty(\Gamma_-)$ so that $u_0 \in L^\infty(X \times V)$. Then we have the following estimate*

$$\left(\mathbb{E}\|u_\varepsilon - u_0\|_{L^2}^2\right)^{\frac{1}{2}} \leq C\varepsilon^{\frac{1}{2}} \longrightarrow 0, \tag{4.7}$$

as ε goes to zero. The constant C depends on the diameter of X, $\|g\|_{L^\infty}$ and β but is independent of ε.

The following estimate of the L^p norms of the random coefficients will be useful for the proofs in this chapter.

a. 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 4.2. *The random fields defined in* (4.2) *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 (4.4) at 0. Since we assumed that the function ϱ is compactly supported, all the integrals are finite. Therefore, we obtain a bound independent of ε. Control of the attenuation coefficient is obtained in the same way. □

4.2.1 Proof of the homogenization theorem

In this section, we prove Theorem 4.1, 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 $d \geq 2$.

Since the scattering kernel k_ε is isotropic by assumption, the transport equation (4.1) reads:

$$v \cdot \nabla_x \zeta_\varepsilon + a\zeta_\varepsilon - k \int_V \zeta_\varepsilon(x, v')dv' = -\delta a_\varepsilon u_\varepsilon + \delta k_\varepsilon \int_V u_\varepsilon(x, v')dv'. \tag{4.8}$$

We can view this equation as $\mathbf{T}\zeta_\varepsilon = A_\varepsilon u_\varepsilon$ where A_ε is an operator defined by $A_\varepsilon f = -\delta a_\varepsilon f + \delta k_\varepsilon \overline{f}$. 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 4.3. *Assume $d \geq 3$. Let $q(x, v) \in L^\infty(X \times V)$ and define*

$$\chi_{1\varepsilon}(x, v) = \int_0^{\tau_-(x,v)} E(x, x - tv)(-\delta a_\varepsilon(x - tv)q(x - tv, v) + \delta k_\varepsilon(x - tv)\overline{q}(x - tv))dt,$$

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

$$\mathbb{E}\|\chi_{1\varepsilon}\|_{L^n}^n \leq C_n \varepsilon^{\frac{n}{2}} \|q\|_{L^\infty}^n, \qquad \mathbb{E}\|\overline{\chi}_{1\varepsilon}\|_{L^n}^n \leq C_n \varepsilon^n \|q\|_{L^\infty}^n. \tag{4.9}$$

Further, solving the equation $\mathbf{T}_1 u = \delta a_\varepsilon \chi_{1\varepsilon}$ yields

$$\mathbb{E}\|\mathbf{T}_1^{-1} \delta a_\varepsilon \chi_{1\varepsilon}\|_{L^n}^n \leq C_n \varepsilon^n \|q\|_{L^\infty}^n. \tag{4.10}$$

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 $\chi_{1\varepsilon}$ without averaging.* We can rewrite $\chi_{1\varepsilon}$ 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

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

where $\tau_- \equiv \tau_-(x,v)$. Using the n-th order moments formula (4.4) we have:

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

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 see that $\mathbb{E}I_1$ is bounded from above by

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

where the last integral is performed on $t_i^{(\ell,n_j)}$ for $2 \le i \le n_j$. Since T^{n_j} is integrable in all directions, we see that all the integrals above are finite and hence we find that

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

From the definition of non-single partition of n, we know $k \le \frac{n}{2}$ to make sure that $n_j \ge 2, j = 1, \cdots, k$. Hence $\min_k(n-k) = \frac{n}{2}$. This yields the first estimate. We mention that the constant depends on n and on the size of X; hence we used C_n to make this dependence explicit.

2. *Control of the average of* $\chi_{1\varepsilon}$. Again, we consider the a_ε term only. Recall the change of variables

$$\int_V \int_0^{\tau_-(x,v)} f(x-tv,v)dtdv = \int_X \frac{f(y,v)}{|x-y|^{d-1}}\Big|_{v=\frac{x-y}{|x-y|}} dy. \tag{4.11}$$

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

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

We recall the notation $d[y_1\cdots y_n] \equiv dy_1\cdots dy_n$. Upon taking expectation, we have

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

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

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

There are many terms to estimate, which are all analyzed in the same manner. Let us fix $k, (n_1,\cdots,n_k)$ and ℓ. Then we need to estimate the product of k integrals involving $T^{n_1},\cdots,T^{n_k}$. Now fix j and T^{n_j} is a function of $y_i^{\ell,n_j} - y_1^{\ell,n_j}$. Hence, we change variables

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

The integral with T^{n_j} becomes

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

Denote y_i^{ℓ,n_j} as y_i' to simplify the notation. 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* (see Figure 4.1) 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 see that I_{n_j} is bounded from above by

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

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

$$C\varepsilon^{(n_j-1)d-(n_j-2)(d-1)-(d-2)} \int_{\mathbb{R}^{(n_j-1)d}} \frac{T^{n_j}(y_2',\cdots,y_{n_j}')}{|y_i'|^{d-2}\prod_{l\neq 1,i}|y_i' - y_l'|^{d-1}} d[y_2'\cdots y_{n_j}'], \tag{4.12}$$

Here, we used Lemma 3.11 for the integral over y_1'. Recall definition of T^{n_j} and the integral above is

$$\begin{aligned} &\int_{\mathbb{R}^{n_j d}} \psi(z)\frac{\psi(z-y_i')}{|y_i'|^{d-2}} \prod_{l\neq 1,i}^{n_j} \frac{\psi(z-y_l')}{|y_i' - y_l'|^{d-1}} dz d[y_2'\cdots y_{n_j}'] \\ =& \int_{\mathbb{R}^d} dz \psi(z) \int_{\mathbb{R}^d} dy_i' \frac{\psi(z-y_i')}{|y_i'|^{d-2}} \prod_{l\neq 1,i} \int_{\mathbb{R}^d} \frac{\psi(z-y_i'-y_l')}{|y_l'|^{d-1}} dy_l'. \end{aligned} \tag{4.13}$$

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

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

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 3.11, 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, $k \le \frac{n}{2}$ 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 $\mathbf{T}_1^{-1}A_\varepsilon\chi_{1\varepsilon}$ as

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

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

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

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 4.4. *Under the same condition as in the previous lemma with* $\mathbf{T}_1$ *replaced by* $\mathbf{T}$*, then for any integer* $n \geq 1$*, we have that when* $d \geq 3$*,*

$$\mathbb{E}\|\mathbf{T}^{-1}A_\varepsilon q\|_{L^n}^n \leq C_n \varepsilon^{\frac{n}{2}} \|q\|_{L^\infty}^n, \qquad \mathbb{E}\|\overline{\mathbf{T}^{-1}A_\varepsilon q}\|_{L^n}^n \leq C_n \varepsilon^n \|q\|_{L^\infty}^n. \tag{4.14}$$

Iterating once more, we have

$$\mathbb{E}\|\mathbf{T}^{-1}A_\varepsilon \mathbf{T}^{-1}A_\varepsilon q\|_{L^n}^n \leq C_n \varepsilon^n \|q\|_{L^\infty}^n. \tag{4.15}$$

In dimension two, the term ε^n *in the second and third estimates should be replaced by* $\varepsilon^n |\log \varepsilon|^{\frac{n}{2}}$*.*

Proof. First, we have

$$\mathbf{T}^{-1} = \mathbf{T}_1^{-1} - \mathbf{T}^{-1}\mathcal{K} = \mathbf{T}_1^{-1} - \mathbf{T}^{-1}A_2\mathbf{T}_1^{-1}. \tag{4.16}$$

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 second item in the previous lemma. The second term above is bounded by $C\|\mathcal{K}A_\varepsilon q\|_{L^n}$ and therefore is also controlled by 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:

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

The constant above is $\|\mathbf{T}_1^{-1}\|_{L^n\to L^n}\|\mathbf{T}^{-1}\|_{L^{2n}\to L^{2n}}$. Then observe that $(a+b)^n \le C_n(a^n+b^n)$ for $a, b \ge 0$, take the expectation and apply the Cauchy-Schwarz inequality to get the result.
□

Remark 4.5. 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 4.1. We assume that $d \ge 3$. 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 $\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \le C\varepsilon$ by the previous lemmas, and it remains to analyze z_ε, which can be rewritten as the sum of $z_{1\varepsilon} := -\mathbf{T}_\varepsilon^{-1}\delta a_\varepsilon\chi_\varepsilon$ and $z_{2\varepsilon} := \mathbf{T}_\varepsilon^{-1}\delta k_\varepsilon\overline{\chi}_\varepsilon$. From the previous lemma and the fact that δk_ε is in L^4, we conclude that

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

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

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

To control $z_{1\varepsilon}$, 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 4.4 and $\mathbb{E}\|z_{11\varepsilon}\|_{L^2}^2 \le C\varepsilon^2$. 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. □

4.3 Corrector Theory for Random Transport

The result in the previous section, Theorem 4.1 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. Point-wise, 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. This section is devoted to two theorems concerning the limits of these correctors.

4.3.1 Limits of deterministic and random correctors

In this section we investigate the weak limits of the deterministic and random parts of the corrector $u_\varepsilon - u_0$. For the deterministic corrector, we have the following theorem capturing its weak limit.

Theorem 4.6. *Let dimension $d = 2, 3$. Under the same conditions of the previous theorem,* *we have*

$$\lim_{\varepsilon\downarrow 0}\frac{\mathbb{E}\{u_\varepsilon\}-u_0}{\varepsilon}(x,v) \quad = \quad U(x,v) \tag{4.17}$$

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

$$v\cdot\nabla_x U + a(x)U - k(x)\int_V U(x,v')dv' = q(x,v), \tag{4.18}$$

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

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

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. An interesting observation is that this size depends on the scale at which observations are made. More precisely, we can consider three types of observations:

1. For a fixed $(x,v)\in X\times V$, the variance of the random variable $\omega\mapsto\zeta_\varepsilon(x,v;\omega)$ is of order ε for all dimensions $d\geq 2$ so that $\gamma=\frac{1}{2}$, cf. [14]. 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. [9, 59].

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 $Z_\varepsilon(\omega)$ as the average of J_ε over all positions.

The variance of Z_ε is of order ε^d in dimension $d \geq 2$ with then $\gamma = \frac{d}{2}$. The random corrector is therefore of smallest size when averaged over the whole phase space and is consistent with the application of the central limit theorem.

The main concern of this section is the stochastic corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$. We aim at characterizing its limit as $\varepsilon \to 0$ weakly in space and velocity. The correct scaling in this case is thus $\gamma = \frac{d}{2}$. Let us consider a collection of sufficiently smooth functions $M_l, 1 \leq l \leq L$. We seek for the limit distribution of $\langle M_l, u_\varepsilon - \mathbb{E}\{u_\varepsilon\}\rangle$, where $\langle \cdot, \cdot \rangle$ denotes the integration of a pair of Hölder conjugate functions.

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

$$\begin{aligned} & -v \cdot \nabla_x \tilde{M}_l + a\tilde{M}_l - \int_V k(x,v,v')\tilde{M}_l(x,v')dv' = M_l, \quad && (x,v) \in X \times V, \\ & \tilde{M}_l(x,v) = 0, && (x,v) \in \Gamma_+, \end{aligned} \tag{4.19}$$

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

$$m_{l1} = -\int_V \tilde{M}_l(x,v)u_0(x,v)dv, \quad \text{and} \quad m_{2l} = c_d m_{l1} + \int_V u_0(x,v)dv \int_V \tilde{M}_l(x,v)dv. \tag{4.20}$$

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 4.7. *Let dimension $d = 2, 3$. Under the same condition of Theorem 4.1, we have*

$$\langle M_l, \frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\varepsilon^{\frac{d}{2}}} \rangle \xrightarrow{\mathscr{D}} I_l := \int_X m_l(y) \cdot dW(y). \tag{4.21}$$

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*

$$\mathbb{E} dW(y) \otimes dW(y) = \Sigma dy := \begin{pmatrix} \sigma_a^2 & \rho_{ak}\sigma_a\sigma_k \\ \rho_{ak}\sigma_a\sigma_k & \sigma_k^2 \end{pmatrix} dy. \tag{4.22}$$

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

Remark 4.8. **More general attenuation and scattering models.** In the construction using (4.2), we have $\rho_{ak} = 1$ so that $W_k = \frac{\sigma_k}{\sigma_a} W_a$ in distribution. The above theorem generalizes to more complex models of attenuation and scattering. For instance, consider

$$\begin{aligned} a_{r\varepsilon}(x, \frac{x}{\varepsilon}, \omega) &:= a_{r0}(x) + \sum_{l=1}^{L} \sum_{j \in \mathbb{N}} \psi_l(\frac{x - y_j^{\varepsilon,l}}{\varepsilon}), \\ k_\varepsilon(x, \frac{x}{\varepsilon}, \omega) &:= k_0(x) + \sum_{l=1}^{L} \sum_{j \in \mathbb{N}} \varrho_l(\frac{x - y_j^{\varepsilon,l}}{\varepsilon}). \end{aligned} \tag{4.23}$$

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:

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

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

To simplify the presentation, we shall only consider the model (4.2) of random media.

Remark 4.9. We can rewrite I_l as

$$I_l(\omega) = \int_X \sigma_l(y) dW(y) := \int_X \sqrt{m_l \otimes m_l : \Sigma}\, dW(y) \tag{4.24}$$

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 variances. The formulation in (4.21) displays the linear dependence of the correctors in $\delta a_r, \delta k$ at the price of introducing two correlated Wiener processes as for elliptic equations [9].

Remark 4.10. Recall the adjoint transport equation of the form (4.19). Let $G_*(x,v,y,v')$ be the Green's function of this equation, i.e., the solution when the source term is $\delta_y(x)\delta_{v'}(v)$, and define

$$\begin{aligned} \kappa_a(x,v,y) &:= \int_V G_*(x,v,y,v')dv' u_0(x,v), \\ \kappa_k(x,v,y) &:= c_d \kappa_a + \int_V G_*(x,v,y,v')dv' \overline{u}_0(x). \end{aligned} \tag{4.25}$$

The convergence in the theorem can be restated as

$$\frac{u_\varepsilon - \mathbb{E}u_\varepsilon}{\varepsilon^{\frac{d}{2}}}(x,v) \Longrightarrow \int (\kappa_a(x,v;y), \kappa_k(x,v;y)) \cdot dW(y) \tag{4.26}$$

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).

4.3.2 Proof of the corrector theorems

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 (4.8) 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 [48]. For similar conclusions for elliptic equations, we refer the reader to e.g. [9, 11, 59].

a. The case of two dimensions

As outlined above, we have the iteration formula:

$$\zeta_\varepsilon = \mathbf{T}^{-1}A_\varepsilon u_0 + \mathbf{T}^{-1}A_\varepsilon \mathbf{T}^{-1}A_\varepsilon u_0 + \mathbf{T}^{-1}A_\varepsilon \mathbf{T}^{-1}A_\varepsilon \zeta_\varepsilon. \tag{4.27}$$

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

$$\langle \zeta_\varepsilon, M\rangle = \langle A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1}A_\varepsilon \zeta_\varepsilon, m\rangle. \tag{4.28}$$

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 4.6. 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. The following lemmas prove these statements.

Let us call the terms in (4.28) 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

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

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

Lemma 4.11. *Assume the same condition of Theorem 4.1 hold. Let $d = 2$. Then we have*
(i) *The mean of J_{21} is of order ε and more precisely,*

$$\mathbb{E}\langle A_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, m\rangle = \varepsilon\langle U, M\rangle + o(\varepsilon) \tag{4.29}$$

where $U(x, v)$ is the solution to (4.18).
(ii) *For the variance of J_{21}, we have*

$$\mathrm{Var}\ \{J_{21}\} \le C\varepsilon^{d+2}|\log\varepsilon| \ll \varepsilon^d. \tag{4.30}$$

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

$$\mathbb{E}J_{22}^2 \le C\varepsilon^{2d}|\log\varepsilon|^2,\ \mathbb{E}J_{23}^2 \le C\varepsilon^{2d}. \tag{4.31}$$

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:

$$J_{21} = \int_{X\times V} m(x,v)\Big[\int_0^{\tau_1} E\big(\delta a_\varepsilon(x)\delta a_\varepsilon(x-tv)u_0 - \delta a_\varepsilon(x)\delta k_\varepsilon(x-tv)\overline{u}_0\big)dt$$
$$+ \int_V\int_0^{\tau_2} E\big(-\delta k_\varepsilon(x)\delta a_\varepsilon(x-sw)u_0 + \delta k_\varepsilon(x)\delta k_\varepsilon(x-sw)\overline{u}_0\big)dwds\Big]dvdx.$$

Here $\tau_1 = \tau_-(x,v)$ and $\tau_2 = \tau_-(x,w)$. After taking expectation, we need to estimate

$$\int_{X\times V} m(x,v)\Big[\int_0^{\tau_1} E(x,x-tv)\big(R_a(\frac{tv}{\varepsilon})u_0(x-tv,v) - R_{ak}(\frac{tv}{\varepsilon})\overline{u}_0(x-tv)\big)dt$$
$$-\int_V\int_0^{\tau_2} E(x,x-sw)\big(R_{ak}(\frac{sw}{\varepsilon})u_0(x-sw,w) - R_k(\frac{sw}{\varepsilon})\overline{u}_0(x-sw)\big)dwds\Big]dvdx.$$

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

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathbb{E}J_{21} = \int_{X\times V} m(x,v)\int_{\mathbb{R}}\Big[R_a(tv)u_0(x,v) - R_{ak}(tv)\overline{u}_0(x)$$
$$+\int_V\big[-R_{ak}(tw)u_0(x,w) + R_k(tw)\overline{u}_0(x)\big]dw\Big]dtdvdx.$$

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 (4.11), rewrite it as

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

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

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

Now recalling the formula (4.6) 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

$$\text{Var }(I_1) \le C \int_{X^4} \frac{1}{|x-y|^{d-1}|x'-y'|^{d-1}} \Bigg(|T^4(\frac{y-x}{\varepsilon}, \frac{x'-x}{\varepsilon}, \frac{y'-x}{\varepsilon})| +$$
$$|R(\frac{x'-x}{\varepsilon})R(\frac{y'-y}{\varepsilon})| + |R(\frac{x'-y}{\varepsilon})R(\frac{y'-x}{\varepsilon})| \Bigg).$$

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. It can be written as

$$\int_{X^4} \frac{|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})|}{|x'-y+(x-x')|^{d-1}|x'-y+(y-y')|^{d-1}} d[x'y'xy].$$

We change variables $(x-x')/\varepsilon \to x$, $(y-y')/\varepsilon \to y$, and $(x'-y)/\varepsilon \to x'$. Then this term is bounded by

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

For the integral in x', we use the convolution lemma 3.11. In dimension two, the last integral is bounded by $C(|\log|x-y|| + |\log \varepsilon|)$. Hence, the above integral is bounded by

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

The first integral is clearly bounded. 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

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

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

$$\begin{aligned} &\int_{X\times V} m(x,v)\delta a_\varepsilon(x) \int_0^{\tau_1} \int_X \frac{E(x, x-tv, y) k(x-tv) \delta a_\varepsilon(y)}{|x-tv-y|^{d-1}} u_0(y, \frac{x-tv-y}{|x-tv-y|}) \\ &= \int_{X^3} m(x, \frac{x-y}{|x-y|}) \frac{E(x,z,y)\delta a_\varepsilon(x)\delta a_\varepsilon(y) k(z)}{|x-z|^{d-1}|z-y|^{d-1}} u_0(y, \frac{z-y}{|z-y|}) d[xzy]. \end{aligned}$$

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 verify that $\mathbb{E}J_{22}^2$ is bounded from above by

$$C\int_{X^6}\big(|T^4(\frac{y-x}{\varepsilon},\frac{z-x}{\varepsilon},\frac{y'-x}{\varepsilon})|+|R(\frac{x-y}{\varepsilon})R(\frac{x'-y'}{\varepsilon})|+|R(\frac{x-y'}{\varepsilon})R(\frac{x'-y}{\varepsilon})| \\ +|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})|\big)\frac{1}{|x-z|^{d-1}|z-y|^{d-1}|x'-z'|^{d-1}|z'-y'|^{d-1}}d[xyzx'y'z'].$$

We integrate over z and z' first. Using the convolution lemma, the term above is bounded by

$$C\int_{X^4}\big(|T^4(\frac{y-x}{\varepsilon},\frac{x'-x}{\varepsilon},\frac{y'-x}{\varepsilon})|+|R(\frac{x-y}{\varepsilon})R(\frac{x'-y'}{\varepsilon})|+|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})| \\ +|R(\frac{x-y'}{\varepsilon})R(\frac{x'-y}{\varepsilon})|\big)|\log|x-y|\log|x'-y'||d[xyx'y'].$$

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 3.11 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

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

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

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

In dimension three, the logarithm terms can be eliminated.

(4) *The absolute mean of* J_{23}. Using formula (4.11), we have

$$J_{23} = \int_{X^4}\frac{m(x,v)E(x,\xi)\Theta(\xi,z)E(z,y)k(z)u_0(y,v')\delta a_\varepsilon(x)\delta a_\varepsilon(y)}{|x-\xi|^{d-1}|z-y|^{d-1}}d[x\xi zy],$$

where $v=(x-\xi)|x-\xi|^{-1}$ and $v'=(z-y)|z-y|^{-1}$, and Θ is the kernel of $\tilde{\mathcal{K}}$. Assume that m and u_0 are bounded. Then $\mathbb{E}J_{23}^2$ can be bounded by

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

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

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

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

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 4.12. *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

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

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

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

Using lemma 4.2 and corollary 4.4, and the fact that $\mathbb{E}\|z_\varepsilon\|_{L^2}^2 \le C\varepsilon^2 |\log \varepsilon|$ derived in the proof of Theorem 4.1, the three terms on the right-hand side above are of size $\varepsilon|\log \varepsilon|^{\frac{1}{2}}$, order $O(1)$, and $\varepsilon^{\frac{1}{2}}$, 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 decomposi-

tion 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 4.2 and corollary 4.4 again to obtain the desired control for I_2 and similarly for I_3.

For I_1, it suffices to consider $\langle {\mathbf{T}^*}_1^{-1} A_\varepsilon m, \delta a_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0\rangle$ 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:

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

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. □

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 $d = 3$.

b. 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:

$$\begin{aligned}\langle \zeta_\varepsilon, M\rangle = \langle A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon u_0, m\rangle \\ + \langle A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \mathbf{T}^{-1} A_\varepsilon \zeta_\varepsilon, m\rangle.\end{aligned} \tag{4.32}$$

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 4.12. Now it suffices to estimate R_2. We first rewrite this term as

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

Here z_ε is defined in the proof of Theorem 4.1. Call them R_{21} and R_{22} respectively and we have the following lemma.

Lemma 4.13. *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.

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

(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 4.12 applies and it suffices to consider $\langle \delta a_\varepsilon \mathbf{T}_1^{-1} \delta a_\varepsilon \mathbf{T}_1^{-1} A_\varepsilon u_0, {\mathbf{T}^*}_1^{-1} A_\varepsilon m\rangle$. It has the expression:

$$\int_{X\times V}\int_0^{\tau_-(x,v)}\int_0^{\tau_-(x-tv,v)}\int_0^{\tau_+(x,v)} \delta a_\varepsilon(x)\big(u_0 E(x, x-tv, x-tv-t_1 v)\delta a_\varepsilon(x-tv) \\ \times u_0 \delta a_\varepsilon(x-tv-t_1v) u_0 E(x, x+sv)\delta a_\varepsilon(x+sv)\big) d[tsxv].$$

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

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

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

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

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. □

With the results above, the proof of Theorem 4.6 is immediate.

Proof of Theorem 4.6. In dimensions 2 and 3, considering the expansion (4.28) or (4.32), 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 4.11. □

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

Lemma 4.14. *Under the same conditions as in Theorem 4.1 with* $d=2,3$*, we have*

$$\mathbb{E}|\langle\varphi,\frac{\zeta_\varepsilon-\mathbb{E}\zeta_\varepsilon}{\varepsilon^{\frac{d}{2}}}-\varepsilon^{-\frac{d}{2}}\mathbf{T}^{-1}A_\varepsilon u_0\rangle|\le C\varepsilon^{\frac12}|\log\varepsilon|^{\frac12}\longrightarrow 0. \tag{4.33}$$

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:

$$\langle\varphi,\varepsilon^{-\frac{d}{2}}\mathbf{T}^{-1}A_\varepsilon u_0\rangle=-\varepsilon^{-\frac{d}{2}}(\langle\mathbf{T}^{*-1}\varphi,\delta a_{r\varepsilon}u_0\rangle+\langle\mathbf{T}^{*-1}\varphi,\delta k_\varepsilon(-\overline{u}_0+c_d u_0)\rangle).$$

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

$$I_{l\varepsilon}=\varepsilon^{-\frac{d}{2}}\int_{\mathbb{R}^d} m_l\cdot(\delta a_r(\frac{x}{\varepsilon}),\delta k(\frac{x}{\varepsilon}))dx, \tag{4.34}$$

where m is defined in (4.20).

As in Remark 4.9, proving Theorem 4.7 is equivalent to proving 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

$$\begin{aligned} \mathbb{E}I_iI_j &= \int m_{1i}m_{1j}\sigma_a^2 + (m_{1i}m_{2j} + m_{2i}m_{1j})\rho_{ak}\sigma_a\sigma_k + m_{2i}m_{2j}\sigma_k^2 dx \\ &= \int m_i \otimes m_j : \Sigma dx. \end{aligned} \tag{4.35}$$

Here, the covariance matrix Σ is defined in (4.22).

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 (CLT) 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. [27]. Generalizations to oscillatory integrals are done in [9] under similar *mixing* conditions. We refer to Theorem 2.15 for the details.

Proof of Theorem 4.7. For simplicity we assume that u_0 and hence m_l are continuous.

1. By the same argument of step one, two and three in the proof of Theorem 2.15, we can assume m_l is piece-wise constant functions m_{lh} on a system of cubes $\{Q_j\}$ of length size h, which covers the domain of interests. Further, it suffices to consider the limit of the integral in (4.34) on each cube. On a typical cube Q_j, the integral we are interested becomes,

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

where the vector m_{lhj} is some constant.

2. For a typical $I_{lh\varepsilon_j}$, as in step four in the proof of Theorem 2.15, we view it as a sum

of $N = h/\varepsilon$ random variables $\hat{q}_i^j$ divided by the CLT scaling factor $N^{\frac{d}{2}}$. More specifically,

$$\hat{q}_i^j = \int_{Q_j^i} q(y)dy,\ q(y) = m_{lh0} \cdot (\delta a_r(y,\omega), \delta k(y,\omega))'.$$

Here Q_j^i is the ith subcube of the cube Q_j chosen in the previous step. Apply the CLT for *mixing* processes parameterized by lattice points; we get $I_{lh\varepsilon j} \to \mathcal{N}(0, \sigma_j^2 h^d)$ where

$$\begin{aligned}
\sigma_j^2 &= \sum_{i\in\mathbb{Z}^d} \mathbb{E}(\hat{q}_0^j \hat{q}_i^j) = \sum_{i\in\mathbb{Z}^d} \mathbb{E} \int_{\mathcal{C}_0} m_{lhj} \cdot (\delta a_r, \delta k)(y)dy \int_{\mathcal{C}_i} m_{lhj} \cdot (\delta a_r, \delta k)(z)dz \\
&= m_{lhj} \otimes m_{lhj} \int_{\mathcal{C}_0} dy \int_{\mathbb{R}^d} dz \mathbb{E}[(\delta a_r, \delta k)(y) \otimes (\delta a_r, \delta k)(z)] \\
&= m_{lhj} \otimes m_{lhj} : \begin{pmatrix} \sigma_a^2 & \rho_{ak}\sigma_a\sigma_k \\ \rho_{ak}\sigma_a\sigma_k & \sigma_k^2 \end{pmatrix}.
\end{aligned} \tag{4.36}$$

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

This proves (4.35) for piece-wise functions, and the general case follows by approximations. This completes the proof. □

Remark 4.15. The CLT of oscillatory integral developed in [9] 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 resulting Gaussian variable to have a bounded variance, we need $m \in L^2$. □

Remark 4.16. 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 4.17. **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. □

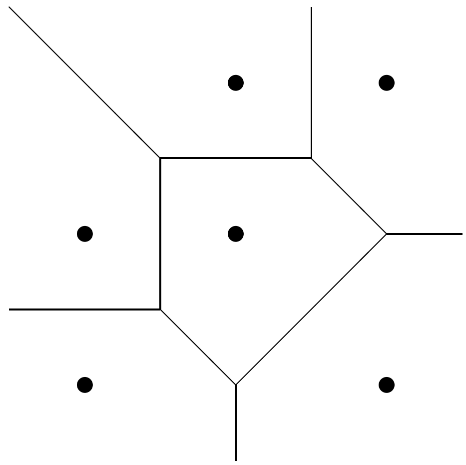
Remark 4.18. So far we have proved the main results with the random field model (4.2). However, the same procedure of proof applies to more general random models. The main required features of the process are: (i) $a_{r\varepsilon}$ and k_ε are nonnegative, stationary, and $\mathbb{P}$-a.s bounded; (ii) 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 of the main theorems for a more quantitative statement. □

4.4 Appendix: Voronoi Diagram

We have used the Voronoi diagram in deriving formulas for the high-order moments of the Poisson bumps model. An illustration of such a diagram for six points on the plane is shown in Figure 4.1.

In general, let $\{x_i\}_{i=1}^N$ be a collection of N distinct points in $\mathbb{R}^d$. The Voronoi diagram

Figure 4.1: The Voronoi diagram of six points on a plane.



for this collection of points is the unique way to divide the space $\mathbb{R}^d$ into N disjoint regions X_i, so that $\bigcup_{i=1}^N \overline{X_i} = \mathbb{R}^d$, and for any point $y \in X_i$, it satisfies that $|y - x_i| < |y - x_j|$ for any $j \neq i$. The construction of such a diagram is intuitively easy. Namely, the boundaries of these disjoint sets are the perpendicular bisectors of these points.

For a more extensive discussion of the Voronoi diagram and its applications, see [6].

Chapter 5

Linear Elliptic Equations with Potentials

This chapter investigates three linear partial differential and pseudo-differential equations. Namely, the steady-state diffusion equation with absorbing (non-negative) potential, and its modification with fractional Laplacian operator, the second one being a pseudo-differential equation. The third equation arises in the pseudo-differential operator method for Robin boundary problems. These three equations are prototypes for the general equations that will be considered in the next chapter.

The main goal of this chapter is to exhibit some of the common features that these equations share. Firstly, the well-posedness depends mildly on the potential. Secondly, the solution operator is a bounded linear transform on L^2 space and its operator norm can be bounded independent of the non-negative potential. Thirdly, the Green's functions of these equations can be bounded by Riesz potentials with certain exponents. These features define a family of PDE or pseudo-differential equations for which we will investigate the homogenization and corrector theory in the next chapter.

5.1 Stationary Diffusion Equations with Linear Potential

We record here some of the important properties of the following Dirichlet problem.

$$\begin{cases} -\Delta u(x) + q(x)u(x) = f(x), & x \in X, \\ u(x) = 0, & x \in \partial X. \end{cases} \tag{5.1}$$

One can think this equation as a stationary diffusion with absorbing potential $q(x) \geq 0$ and internal source $f(x)$. Note also that the Laplacian operator can be replaced by $-\nabla \cdot A(x)\nabla$ as long as $A(x)$ is some nice positive matrix which accounts for anisotropic diffusion.

5.1.1 Several important properties

This equation is just a very special case of the Dirichlet problem of a second order elliptic equation, for which many results are well established. Below, we record some of them which are very useful for us in the sequel.

Theorem 5.1 (Stationary Diffusion Equation)**.** *Let the potential function $q(x)$ to be non-negative, and the internal source f to be in $L^2(X)$, then there exists a unique weak solution $u \in H_0^1(X)$ of the equation. Further we have the following:*

(i) (Solution operator) *Let $L = -\Delta + q(x)$ be the differential operator; then $S := L^{-1}$ which maps $L^2(X)$ to itself is a bounded linear compact operator. In particular, $\|S\|_{\mathcal{L}(L^2)}$ can be bounded by a constant only depending on the geometry of X.*

(ii) (Spectra) *The eigenvalues of L are real. If we repeat each eigenvalue according to its (finite) multiplicity, the eigenvalues are*

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots,$$

and $\lambda_k \to \infty$ as k goes to infinity. Furthermore, there exists an orthonormal basis $\{\phi_n(x)\}_{n=1}^{\infty}$ of $L^2(X)$ such that each $\phi_n \in H_0^1(X)$ is the eigenfunction corresponding to

λ_n.

(iii) (Estimate of eigenvalues) *Let the eigenvalues be ordered as above, then we have*

$$\lambda_n \geq (2\pi)^2 \pi_d^{-\frac{2}{d}} \left(\frac{n}{|X|} \right)^{\frac{2}{d}}. \tag{5.2}$$

Here π_d is the volume of the unit ball in $\mathbb{R}^d$; $|X|$ is the volume of the domain X.

(iv) (Green's function) *Let $G(x,y)$ be the Green's function of the equation, that is, the solution corresponding to Dirac source located at point $y \in X$. Then there exists some $C \geq 1$ only depend on the geometry of the domain, such that*

$$C^{-1}|x-y|^{-n+2} \leq |G(x,y)| \leq C|x-y|^{-n+2},$$

when $n \neq 2$. When $n = 2$, the potential above should be replaced by logarithm function.

(v) (Regularity) *Suppose that $f \in C^k(\overline{X})$ and $q \in C^{k+2}(\overline{X})$, then u is also in $C^{k+2}(\overline{X})$.*

Most of the results above are now standard and appear in textbooks, with the exception of item three and four. The estimate of eigenvalues above is asymptotically an equality for large n; this is the so-called Weyl's asymptotic formula. The fourth item is essential a comparison between the Green's functions of the Dirichlet Laplacian operator $-\Delta$ and the Dirichlet Laplacian with potential $-\Delta + q$. We refer the reader to Li and Yau [77] for the eigenvalue estimate, and Chung and Zhao [42, 113] for the Green's function estimate.

The result about the solution operator can be proved as follows. It suffices to define a bilinear form on $H_0^1 \times H_0^1$ given by

$$B[u,v] = \int_X Du(x) \cdot Dv(x) + q(x)u(x)v(x)dx,$$

and verify that $B[u,v]$ satisfies the conditions of Lax-Milgram theorem. Namely, we have

$$|B[u,v]| \leq \|Du\|\|Dv\| + \|q\|_{L^\infty}\|u\|\|v\| \leq (1 + \|q\|_{L^\infty})\|u\|_{H^1}\|v\|_{H^1},$$

and

$$B[u,u] \geq \|Du\|^2 \geq (1+C_p^2)^{-1}\|u\|_{H^1}^2.$$

Here and in the sequel, we use $\|\cdot\|$ to denote the L^2 norm. The constant C_p above is the one in the Poincaré inequality $\|u\| \leq C_p\|Du\|$ for $u \in H_0^1(X)$. In particular, C_p can be chosen only depending on the geometry of X. The existence and uniqueness of the solution then follows from the Lax-Milgram theorem. The compactness of S follow from the second estimate above and the Rellich-Kondrachov compactness theorem.

Now the property of the spectra of L follows from the fact that S is a compact operator on L^2 and the fact that S is symmetric, in the sense that $(Sf,g) = (f,Sg)$ for any $f,g \in L^2(X)$. This symmetry follows from integration by parts on the equation. Apply the spectral theory for compact symmetric and positive operators to obtain the property of the spectra of S. Then translate the result to the spectra of L by observing that eigenvalues of L are the reciprocals of those of S.

The regularity result of second order elliptic equations is a very deep yet technical result. An extensive reference for it is [61].

5.1.2 Brownian motion and Feynman-Kac formula

In this section, we recall one example of the beautiful interplay between probability and partial differential equations, namely the Feynman-Kac formula for the elliptic equation considered above.

Let W_t be a standard d-dimensional Brownian motion with free staring point, that is, W_0 not specified. Since Brownian motion exits any finite balls in finite time almost surely, the following stopping time (with respect to the natural filtration of Brownian motion)

$$\tau_X := \inf\{t > 0 \ : \ W_t \in X^c\}, \quad \textit{i.e.}, \text{ the first time exiting } X, \tag{5.3}$$

is almost surely finite. Let $\mathbb{E}^x$ denote the expectation conditioned on $W_0 = x$. The solution $u(x)$ of the Dirichlet problem (5.1) has the following stochastic representation:

$$u(x) = \mathbb{E}^x \int_0^{\tau_X} \frac{1}{2} f(W_t) \exp\left\{ - \int_0^t \frac{1}{2} q(W_s) ds \right\} dt. \tag{5.4}$$

This is the celebrated Feynman-Kac formula (or Kac formula in this special case). We refer to [72] and for the details. The main purpose of this short introduction will be clear in a moment.

5.2 Fractional Laplacian Operator with Linear Potential

In this section, we replace the Laplacian operator $-\Delta$ above by its fractional exponent $(-\Delta)^{\beta/2}$ for some $\beta \in (0, 2)$, and consider the following "Dirichlet problem":

$$\begin{cases} (-\Delta)^{\beta/2} u(x) + q(x) u(x) = f(x), & x \in X, \\ u(x) = 0, & x \in X^c. \end{cases} \tag{5.5}$$

The fractional Laplacian is defined as

$$(-\Delta)^{\frac{\beta}{2}} u := \frac{2^\alpha \Gamma\left(\frac{d+\alpha}{2}\right)}{\pi^{\frac{d}{2}} \Gamma\left(-\frac{\alpha}{2}\right)} \text{p.v.} \int_{\mathbb{R}^d} \frac{u(x) - u(y)}{|x-y|^{d+\beta}} dy.$$

The principal value is taken over $\{|y-x| > \varepsilon\}$ as ε goes to zero. Consequently, the fractional Laplacian is not a local operator; this forces the boundary condition to be defined on the whole complement of X, in contrast to the boundary ∂X for the Laplacian case.

We will record some of the very important properties of this fractional Laplacian equation with potential. Many of them are copied from the probabilistic literature. There, the nice interplay between the Laplacian and the Brownian motion, or more generally the relationship between divergence form operator and diffusion process, is generalized to processes with discontinuities.

5.2.1 Stable Lévy process and fractional Laplacian

Definition 5.2 (Lévy Process). A *Lévy process* is a stochastic process V_t which maps $\Omega \times [0, \infty)$ to $\mathbb{R}^d$ satisfying:

(i) X_t has stationary and independent increments;

(ii) X_t has *càdlàg* (right continuous with left limits) paths.

Examples of Lévy processes include Brownian motions, whose paths are continuous, and the one-dimensional Poisson processes, whose paths are piecewise constant functions. For fractional Laplacian, the following family of Lévy processes are important.

Definition 5.3 (β-stable Lévy process). A *symmetric β-stable Lévy process* V_t on $\mathbb{R}^d$ is a Lévy process whose transition density $p(t, y)$, i.e., the limit of $\mathbb{P}(V_t \in dy \mid V_0 = 0)/|dy|$, has Fourier transform $\int_{\mathbb{R}^d} e^{iy\cdot\xi} p(t, y) dy = e^{-t|\xi|^\beta}$. Here, β is in $(0, 2]$.

When $\beta = 2$, this is just the Brownian motion. Note however, according to the characteristic function of the increment above, the time clock of this Brownian motion is running twice faster than the standard one. In the sequel, stable Lévy processes are referred to the case when $\beta \in (0, 2)$. Such processes are now widely used in physics, operator research, mathematical finance and risk estimation [40, 60], mainly because the discontinuity of Lévy paths can model, e.g., jumps in the price of financial assets.

Suppose X is a C^2 domain; that is, ∂X is a finite union of rotations of graphs of C^2 functions. Adjoint an extra point ∂ to X, the point "outside X" also known as the *cemetery* state, and set

$$V_t^X(\omega) = \begin{cases} V_t(\omega), & \text{if } t < \tau_X, \\ \partial, & \text{if } t \geq \tau_X. \end{cases}$$

Again, τ_X is the first time exiting X. This process is the so-called symmetric β-stable process killed upon leaving X, or simply the killed symmetric β-stable process on X. Again, Let $\mathbb{E}^x$ denote the expectation conditioned on $V_0^X = x$, the solution to the Dirichlet problem (5.5) has the following stochastic interpretation:

$$u(x) = \mathbb{E}^x \int_0^{\tau_X} f(V_t^X) \exp\left\{ -\int_0^t q(V_s^X) ds \right\} dt. \tag{5.6}$$

This is a Feynman-Kac type formula for the β-stable process. We will not investigate this stochastic representation further. The reason of including this section is to provide a helpful perspective of the somewhat complicated pseudo-differential equation (5.5). Throughout this thesis, we use the terminology pseudo-differential operator for non-local operators like $(-\Delta)^{\beta/2}$. Though it is possible, we do not rigorously justify our usage of this notion since the deep theories in that field, e.g. in [109, 110], are not used in this thesis *per se*.

5.2.2 Important properties of fractional Laplacian

In this section, we record some of the main properties of the pseudo-differential equation involving fractional Laplacian introduced above.

Theorem 5.4 (Fractional Laplacian with Potential)**.** *Let the potential function $q(x)$ to be non-negative. We have the following:*

(i) (Solution operator) *Let $L = (-\Delta)^{\frac{\beta}{2}} + q(x)$ be the differential operator with the boundary condition in* (5.5)*. Then $S := L^{-1}$ is a bounded operator from $L^p(X)$, $1 \le p \le \infty$, to $L^\infty(X)$ and to itself.*

(ii) (Spectra) *The eigenvalues of L are real. If we repeat each eigenvalue according to its (finite) multiplicity, the eigenvalues are*

$$0 < \lambda_1 \le \lambda_2 \le \lambda_3 \le \cdots,$$

and $\lambda_k \to \infty$ as k goes to infinity. Furthermore, there exists an orthonormal basis $\{\phi_n(x)\}_{n=1}^\infty$

of $L^2(X)$ such that each ϕ_n is the eigenfunction corresponding to λ_n.

(iii) (Estimate of eigenvalues) *Let the eigenvalues be ordered as above, then we have*

$$\lambda_n \geq C \left(\frac{n}{|X|} \right)^{\frac{\beta}{d}},$$

for some constant C only depend on the dimension.

(iv) (Green's function) *Let $G(x, y)$ be the Green's function of the equation, that is, the solution corresponding to Dirac source located at point $y \in X$. Then there exists some $C \geq 1$ only depend on the geometry of the domain, such that*

$$C^{-1}|x-y|^{-n+\beta} \leq |G(x,y)| \leq C|x-y|^{-n+\beta},$$

for $n \geq 2$.

(v) (Regularity) *Suppose the domain X has C^2 boundary and $f \in C_0(X)$ and $q \in C(\overline{X})$, then u is also in $C_0(X)$ as well.*

Most of the results above are systematically developped by Chen and Song [34, 35], Bogdan and Byczkowski and so on [25, 26], following the probabilistic approach of Chung and Zhao [39], which dealt with Brownian motion and corresponding results in Theorem 5.1. The eigenvalue estimate interests us in particular, and it is a combination of the result in [36] and the Li-Yau estimate mentioned before.

5.3 Pseudo-Differential Method for a Robin Boundary Problem

Pseudo-differential operators can be used to solve partial differential equations with mixed boundary conditions. As an illustration, we consider the following steady-state diffusion over the half space with Robin boundary conditions. That is,

$$\begin{cases} (-\Delta+\lambda^2)u(x)=0, & x\in\mathbb{R}^n_+, \\ \dfrac{\partial}{\partial\nu}u(x')+q_0u(x')=f(x'), & x'\in\mathbb{R}^d. \end{cases} \tag{5.7}$$

We also impose that the solution decays sufficiently fast as $|x|$ tends to infinity. Above, we identified the boundary $\partial\mathbb{R}^n_+$ with $\mathbb{R}^d$ where $d=n-1$. For simplicity we assume that the damping coefficient λ^2 is a constant with $\lambda>0$, and the impedance q_0 in the Robin boundary condition is also a positive constant. Under this condition (5.7) is well-posed, to see this we relate the equation for u above to the equation satisfied by its trace on the boundary $\mathbb{R}^d$. In the sequel and to simplify notation, we still use x, instead of x', to denote a point in $\mathbb{R}^d$.

Let us define the standard Dirichlet-to-Neumann (DtN) operator Λ as follows:

$$\Lambda g(x) := \frac{\partial}{\partial\nu}\tilde{g}(x). \tag{5.8}$$

Here, the function $g(x)$ is defined on the boundary $\mathbb{R}^d$ and $\tilde{g}$ is the solution of the volume problem (5.7) with a Dirichlet boundary condition $\tilde{g}|_{\partial\mathbb{R}^n_+}=g$. Hence, Λ maps the boundary value to the boundary flux. Either by calculating the symbol of Λ or by verifying it directly, we observe that $\Lambda=\sqrt{-\Delta+\lambda^2}$; see section 5.3.1. Note that Δ here is the Laplacian on $\mathbb{R}^d$, i.e., the surface Laplacian $\Delta_\perp$. To simplify notation, we will use Δ to denote both of the Laplacians on $\mathbb{R}^n$ and $\mathbb{R}^d$. The volume problem (5.7) is then equivalent to the following pseudo-differential equation posed on the whole space $\mathbb{R}^d$,

$$(\sqrt{-\Delta+\lambda^2}+q_0)u=f. \tag{5.9}$$

Indeed by definition, the trace of the solution to (5.7) satisfies equation (5.9), and the lift $\tilde{u}$ of solution to (5.9) solves equation (5.7). Thanks to the fact that q_0 is positive, (5.9) admits a unique weak solution in $H^{\frac{1}{2}}(\mathbb{R}^d)$ provided that $f \in H^{-\frac{1}{2}}(\mathbb{R}^d)$; see section 5.3.1 for the proof. We assume $f \in L^2(\mathbb{R}^d)$ and consequently both the pseudo-differential equation (5.9) and the diffusion equation (5.7) in the volume are well-posed.

Let $\mathcal{G}$ be the solution operator of (5.9) and let $G(x, y)$ be the corresponding Green's function, i.e., the Schwartz kernel of $\mathcal{G}$. By homogeneity, we observe that G is of the form $G(|x - y|)$. This Green's function will be investigated further in section 5.3.1. The latter function decays exponentially at infinity and behaves like $|x|^{-d+1}$ near the origin when $d \geq 2$. The exponential decay allows us to easily work in infinite domain. The singularity at the origin shows that G fails to be locally square integrable. Hence the Robin problem under investigation provides another example whose Green's function is more singular than that of the Laplace equation. In fact, we will verify that $\beta = 1$ in this case.

Equation (5.7) has an important application in biology, which we will discuss further in the next chapter. The physical domain in this application has $n = 3$ and hence $d = 2$. Our results are presented in that setting of practical interest.

5.3.1 Properties of the Green's function

In this section, we first show that the Robin problem (5.7) is equivalent to the pseudo-differential equation (5.9) by calculating the symbol of the Dirichlet-to-Neumann map Λ. Using this symbol we show that (5.9) admits a well defined solution operator $\mathcal{G}$ and derive an expression for the corresponding Green's function G.

a. Symbol of the Dirichlet-to-Neumann map

We now verify the claim that the DtN map Λ equals the pseudo-differential operator $\sqrt{-\Delta + \lambda^2}$ defined as

$$\sqrt{-\Delta + \lambda^2} f = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \sqrt{|\xi|^2 + \lambda^2} \hat{f}(\xi) d\xi, \tag{5.10}$$

where $\hat{f}$ is the Fourier transform of f defined as

$$\hat{f}(\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx. \tag{5.11}$$

We will also denote by $\mathscr{F}$ the Fourier transform operator, and by $\mathscr{F}^{-1}$ its inverse.

By definition (5.8), $\Lambda g(x)$ is the normal derivative of $\tilde{g}(x, x_n)$, the function satisfying:

$$\begin{cases} -\Delta\tilde{g}(x, x_n) + \lambda^2\tilde{g}(x, x_n) = 0, & (x, x_n) \in \mathbb{R}^n_+, \\ \tilde{g}(x, 0) = g(x), & x \in \mathbb{R}^d \equiv \partial\mathbb{R}^n_+. \end{cases} \tag{5.12}$$

Taking Fourier transform in the variable x, we obtain a second order ordinary differential equation in x_n, i.e.,

$$\begin{cases} -\partial^2_{x_n}\hat{\tilde{g}}(\xi, x_n) + (|\xi|^2 + \lambda^2)\hat{\tilde{g}} = 0, \\ \hat{\tilde{g}}(\xi, 0) = \hat{g}(\xi). \end{cases} \tag{5.13}$$

Solve this ODE with the assumption that $\hat{\tilde{g}}$ decays for large frequency to get

$$\hat{\tilde{g}}(\xi, x_n) = \hat{g}(\xi)\exp(-x_n\sqrt{|\xi|^2 + \lambda^2}).$$

Take derivative in the $-x_n$ direction, i.e. the outward normal direction and send x_n to zero to obtain Fourier transform of the function Λg. It has the form

$$\widehat{\Lambda g}(\xi) = \sqrt{|\xi|^2 + \lambda^2}\hat{g}(\xi). \tag{5.14}$$

This verifies that the symbol of Λ is $\sqrt{|\xi|^2 + \lambda^2}$. Compare this symbol with (5.10) and we see $\Lambda = \sqrt{-\Delta + \lambda^2}$. Therefore, (5.7) and (5.9) are equivalent by the argument below (5.9).

b. Solution of the pseudo-differential equation

As an immediate result, we show that (5.9) admits a solution operator $\mathcal{G} : H^{-\frac{1}{2}}(\mathbb{R}^d) \to$

$H^{\frac{1}{2}}(\mathbb{R}^d)$ given by:

$$\mathcal{G}f(x) := \mathscr{F}^{-1}\frac{\hat{f}}{\sqrt{|\xi|^2+\lambda^2}+q_0} \equiv \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} e^{i\xi\cdot x}\frac{\hat{f}}{\sqrt{|\xi|^2+\lambda^2}+q_0}d\xi. \tag{5.15}$$

In particular, the map $\mathcal{G} : f \to \mathcal{G}f$ is continuous from $L^2(\mathbb{R}^d)$ to itself, and the operator norm is bounded by a constant that only depends on λ provided that the impedance is non-negative.

We recall some definitions. The Sobolev space H^s for $s \in \mathbb{R}$ is defined as

$$H^s(\mathbb{R}^d) := \left\{ v \in \mathcal{S}' \mid \hat{v}\langle\xi\rangle^s \in L^2(\mathbb{R}^d) \right\}, \tag{5.16}$$

where $\mathcal{S}'$ is the space of tempered distributions, i.e., linear functionals of the Schwartz space $\mathcal{S}$, and $\langle\xi\rangle = (1+|\xi|^2)^{1/2}$. To simplify notation, we will denote $H^{\frac{1}{2}}$ by H, and the corresponding norm is

$$\|f\|_H := \left(\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2\langle\xi\rangle d\xi\right)^{\frac{1}{2}}. \tag{5.17}$$

To prove that (5.9) is well-posed, we first write a variational formulation of it. To do so, multiply (5.9) by a smooth test function v, and integrate. We have

$$B[u,v] = \langle f, v\rangle, \tag{5.18}$$

where $B[u,v]$ is a bilinear form defined as

$$B[u,v] := \langle \Lambda u, v\rangle + \langle q(x)u, v\rangle. \tag{5.19}$$

From its symbol we see that Λ maps $H^{1/2}$ to $H^{-1/2}$. As a result, the bilinear form $B[\cdot,\cdot]$ above is well defined on $H \times H$. We say u is a weak solution of (5.9) if (5.18) holds for arbitrary $v \in H$.

The following proposition states that the bilinear form B satisfies the conditions of the Lax-Milgram theorem and its corollary says (5.9) admits a unique solution in H. For the moment, we allow the impedance in (5.9) to be a non-negative function denoted by $q(x)$.

Proposition 5.5. *Let λ in (5.9) be a positive constant. Let $q(x)$ in (5.19) be a non-negative function and assume $\|q\|_{L^\infty}$ is finite. Set $\alpha = \|q\|_{L^\infty} + \max(1, \lambda)$, $\gamma = \min(1, \lambda)$. Then the bilinear form $B[u, v]$ in (5.19) satisfies the following:*

(i) $|B[u, v]| \leq \alpha \|u\|_H \|v\|_H$, for all $u, v \in H$, and

(ii) $\gamma \|u\|_H^2 \leq B[u, u]$, for all $u \in H$.

Proof. The following inequalities hold for all ξ.

$$\gamma \leq \sqrt{\frac{|\xi|^2 + \lambda^2}{|\xi|^2 + 1}} \leq \max(1, \lambda). \tag{5.20}$$

Using the second inequality, formula (5.14), and Cauchy-Schwarz, we get

$$|\langle \Lambda u, v \rangle| = \Big| \int_{\mathbb{R}^d} \sqrt{\lambda^2 + |\xi|^2} \hat{u} \overline{\hat{v}} d\xi \Big| \leq \max(1, \lambda) \left(\int_{\mathbb{R}^d} |\hat{u}|^2 \langle \xi \rangle d\xi \right)^{1/2} \left(\int_{\mathbb{R}^d} |\hat{v}|^2 \langle \xi \rangle d\xi \right)^{1/2}.$$

Since $\|u\|_{L^2} \leq \|u\|_H$ for all $u \in H$, we have

$$|B[u, v]| \leq \max(1, \lambda) \|u\|_H \|v\|_H + \|q\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} \leq \alpha \|u\|_H \|v\|_H,$$

which verifies (i). For the second inequality, since $q(x)$ is non-negative, we have

$$B[u, u] \geq \langle \Lambda u, u \rangle = \int_{\mathbb{R}^d} |\hat{u}|^2 \sqrt{\lambda^2 + |\xi|^2} d\xi \geq \gamma \int_{\mathbb{R}^d} |\hat{u}|^2 \langle \xi \rangle d\xi.$$

In the last inequality we applied (5.20). This verifies (ii) and completes the proof. □

Corollary 5.6. *Let $\lambda, q(x)$ and γ be the same as in the preceding proposition. Assume also that f is in $H^{-1/2}$. Then (5.9) admits a weak solution $u \in H$ satisfying (5.18). In* *particular, if $f \in L^2$, then we have that*

$$\|u\|_{L^2} \leq \gamma^{-1}\|f\|_{L^2}. \tag{5.21}$$

Proof. The first claim follows immediately from the preceding proposition and the Lax-Milgram theorem. The second one is due to the following estimate which is clear from (ii) of Proposition 5.5 and Cauchy-Schwarz inequality.

$$\gamma\|u\|_{L^2}^2 \leq \gamma\|u\|_H^2 \leq B[u,u] = \langle f, u\rangle \leq \|f\|_{L^2}\|u\|_{L^2}.$$

This completes the proof. □

Now it is a simple matter to check that $\mathcal{G}$ defined in (5.15) gives the solution operator. Therefore, the corollary above shows that the operator norm of $\mathcal{G}$ as a transformation on $L^2(\mathbb{R}^d)$ is bounded by the constant γ^{-1}.

Remark 5.7. The explicit bound γ^{-1} in estimate (5.21) is crucial for us when the random equation is considered. Suppose the potential q_0 in perturbed by a random potential $q_\varepsilon(x,\omega)$, and let $\mathcal{G}_\varepsilon$ denote the solution operator of the perturbed equation. This corollary shows that $\mathcal{G}_\varepsilon$ is well defined as long as $q_0 + q_\varepsilon$ is non-negative (which is true thanks to the uniform bound of q_ε and the operator norm $\|\mathcal{G}_\varepsilon\|_{\mathcal{L}(L^2)}$ is bounded uniformly for almost every realization.

c. Decomposition of Green's function

Let $G(x,y)$ be the Green's function associated to the solution operator $\mathcal{G}$ of (5.9). By homogeneity $G(x,y) = G(x-y)$ and $G(x)$ solves

$$\left(\sqrt{-\Delta + \lambda^2} + q_0\right) G(x) = \delta_0(x).$$

Take Fourier transform on both sides. Our choice of the definition of Fourier transform (5.11) implies that $\mathscr{F}\delta_0(x) \equiv (2\pi)^{-d/2}$. Hence, $G(x)$ is recovered by the inversion formula as follows;

$$G(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\xi \cdot x} (\sqrt{|\xi|^2 + \lambda^2} + q_0)^{-1} d\xi. \tag{5.22}$$

In dimension two, we have the following explicit characterization.

Lemma 5.8. *Let $d = 2$. Let λ, q_0 in (5.9) be positive constants and $d = 2$. The Green's function $G(x)$ defined above can be decomposed into three terms as follows:*

$$G(x) = \frac{1}{2\pi} \left(\frac{\exp(-\lambda|x|)}{|x|} - q_0 K_0(\lambda|x|) + G_r(|x|) \right). \tag{5.23}$$

Here K_0 is the modified Bessel function with index zero and the function $G_r(|x|)$ is smaller than $C_b \exp(-b|x|)$ for any positive real number $b < \lambda' \equiv \lambda/\sqrt{2}$.

Remark 5.9. In the sequel, we will call the first term on the right G_s and the second one G_b. Clearly, G_s has singularity of order $|x|^{-1}$ near the origin and has exponential decay at infinity; G_r is smooth near the origin and has exponential decay at infinity. Asymptotic analysis of Bessel functions shows that G_b has a logarithmic singularity near the origin and exponential decay at infinity, cf. [111]. In summary, we have

$$|G(x)| \leq C_\lambda \frac{\exp(-\lambda'|x|)}{|x|}, \tag{5.24}$$

where C_λ is a constant depending on λ and q_0.

Proof. We first decompose the Fourier transform of G into three parts as follows.

$$2\pi \hat{G}(\xi) = \frac{1}{\sqrt{|\xi|^2 + \lambda^2}} - \frac{q_0}{|\xi|^2 + \lambda^2} + \frac{q_0^2}{(|\xi|^2 + \lambda^2)[q_0 + \sqrt{|\xi|^2 + \lambda^2}]}. \tag{5.25}$$

Now the first two terms can be inverted explicitly. For instance, the second one is a standard example in textbooks on Fourier analysis or PDE, cf. Taylor [108, Chap. 3], Evans

[57, Chaper 4]. In our case the dimension equals two, and its inversion is the following.

$$-\frac{1}{2\pi}\int_{\mathbb{R}^2}\frac{q_0\ e^{ix\cdot\xi}}{|\xi|^2+\lambda^2} = -\frac{q_0}{2}\int_0^\infty \frac{e^{-\frac{|x|^2}{4t}-t}}{t}dt = -q_0 K_0(\lambda|x|). \tag{5.26}$$

Here K_0 is the modified Bessel function of the *second* kind with index 0. It has logarithmic singularity near the origin and decays exponentially at infinity.

In dimension two, the first term admits an explicit expression as well. Indeed, thanks to (5.14), $(\sqrt{|\xi|^2+\lambda^2})^{-1}$ can be viewed as the symbol of Λ^{-1}, i.e., the Neumann-to-Dirichlet operator which maps the Neumann boundary condition of a diffusion equation of the form (5.12) to its solution evaluated at the boundary. Therefore, G_s can be obtained by taking the trace of G_D, by which we denote the Green's function associated to (5.12) with Neumann boundary. Since $d = 2$ and $n = 3$, G_D can be calculated explicitly using the method of *images* as we show now. The fundamental solution of (5.12) posed on whole $\mathbb{R}^3$ is given by $\exp(-\lambda|x|)/4\pi|x|$, cf. Reed and Simon [98, Chap. IX.7]. By the method of images, the Green's function for the Neumann problem on the upper half space is given by

$$G_D(x,y) = \frac{1}{4\pi}\frac{\exp(-\lambda|y-x|)}{|y-x|} + \frac{1}{4\pi}\frac{\exp(-\lambda|y-\tilde{x}|)}{|y-\tilde{x}|},$$

for x in the upper space and $\tilde{x}$ denotes its image in the lower half space. Evaluating G_D for x on the boundary, we obtain that

$$G_s(x,y) = \frac{1}{2\pi}\frac{\exp(-\lambda|y-x|)}{|y-x|}.$$

Clearly, it has singularity of order $|x-y|^{-1}$ near the origin and decays exponentially at infinity.

Now we are left with the third term of (5.25). We won't give an explicit formula for its Fourier inversion. Nevertheless, we can show that its inversion decays exponentially at infinity and has no singularity near the origin. The proof is a little more involved and we wrote it as Lemma 5.10. It essentially uses the Paley-Wiener theorem. Now the proof is complete. □

Lemma 5.10. *Let λ and q_0 be positive real numbers and let $\xi \in \mathbb{R}^2$. Set $\lambda' \equiv \lambda/\sqrt{2}$. Then, for any positive real number $b < \lambda'$, there exists a finite constant C_b such that*

$$\left| \mathscr{F}^{-1} \frac{q_0^2}{(|\xi|^2 + \lambda^2)(q_0 + \sqrt{|\xi|^2 + \lambda^2})} \right| \leq C_b e^{-b|x|}. \tag{5.27}$$

Proof. 1. Let us denote by $h(\xi)$ the function whose inverse Fourier transform is considered in (5.27). Let us also define $h(z)$ to be the same function with ξ replaced by $z = (z_1, z_2) \in \mathbb{C}^2$, a complex valued function of two complex variables. Set

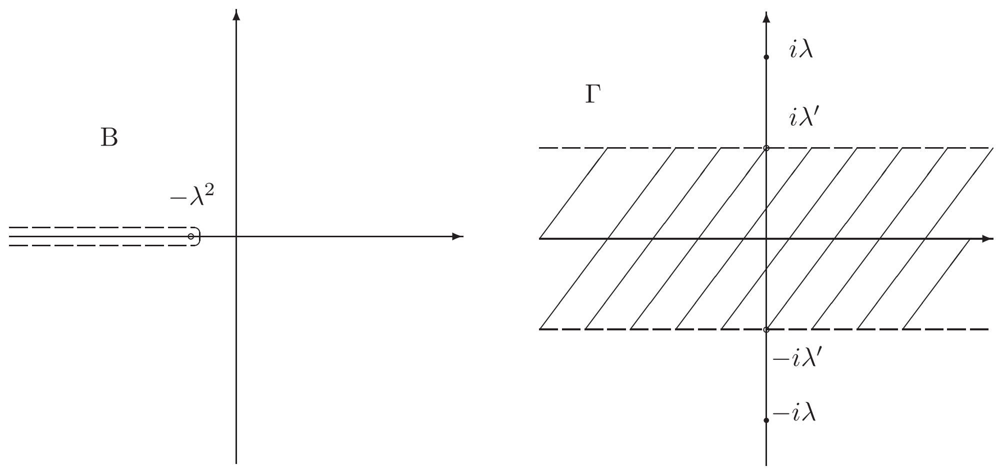
$$\Gamma := \{z \in \mathbb{C} || \mathrm{Im}(z)| \leq \lambda'\}. \tag{5.28}$$

We claim that h is holomorphic on the region Γ^2, i.e. $\Gamma \times \Gamma$.

Indeed, let $w(z_1, z_2)$ be the function $z_1^2 + z_2^2$. It is clearly entire on $\mathbb{C}^2$. Define $g(w) := \sqrt{w + \lambda^2}$ as a function of one complex variable. It is holomorphic on the branched region $\mathrm{B} := \mathbb{C} \backslash (-\infty, -\lambda^2]$ as shown in Fig. 5.1. Now when $(z_1, z_2) \in \Gamma^2$, we verify that $w \in \mathrm{B}$ and hence $g(w(z))$ is holomorphic on Γ^2. This is because composition of holomorphic functions is again holomorphic; see [58]. Since $\lambda > q_0$, we verify that $g(w(z)) + q_0$ does not vanish. Thus, $h(z)$ is holomorphic on Γ^2.

The above arguments show that for any $\eta \in \mathbb{R}^2$ so that $|\eta_j| < \lambda'$, $i = 1, 2$, the function $h(\xi + i\eta)$ is analytic. Furthermore, it is easy to check that $\|h(\xi + i\eta)\|_{L^1}$ is bounded uniformly in η. Hence we apply Theorem IX.14 of [98], which says that under such conditions, for each $0 < b < \lambda'$, there exists C_b so that $|\mathscr{F}^{-1} h| \leq C_b e^{-b|x|}$. This completes the proof. □

Figure 5.1: Holomorphic region of the function $h(z)$. The first picture shows the holomorphic region of $g(w) = \sqrt{w + \lambda^2}$; the second one shows the shadowed region Γ such that Γ^2 is the holomorphic region of $g(z_1^2 + z_2^2)$. Here $\lambda' = \lambda/\sqrt{2}$.



5.4 Notes

Sections 5.1 and 5.2 We do not use the Feynman-Kac formulas in this thesis, but they are very useful in random homogenization of PDEs with one spatial dimension. We refer the reader to the many works of Pardoux and his colleagues [46, 94, 69, 93]. I learnt the properties of equation (5.5) mainly through the probability literature where the infinitesimal generator L of the killed β-stable Lévy process is intensively studied. The comparison between the Green's functions of L, $L + q$ and $(-\Delta)^{\beta/2}$ is established by Chen and Song [34, 35], Bogdan and Byczkowski [25]. Non-probability approach is also available in Hansen [63]. The nice comparison between eigenvalues of L and $-\Delta$, which will be very useful for us later, is established by Chen and Song in [36].

Chapter 6

Corrector Theory in Random Homogenization of Linear Elliptic Equations with Potentials

In this chapter, we consider random perturbation of the elliptic partial differential or pseudo-differential equations introduced in the previous chapter. That is,

$$P(x, D)u_\varepsilon + \tilde{q}_\varepsilon\left(x, \frac{x}{\varepsilon}, \omega\right) u_\varepsilon = f(x), \tag{6.1}$$

for x in an open subset $X \subset \mathbb{R}^d$ with appropriate boundary conditions on ∂X if necessary. Here, $\tilde{q}_\varepsilon\left(x, \frac{x}{\varepsilon}, \omega\right)$ is composed of a low frequency part $q_0(x)$ and a high frequency part $q\left(\frac{x}{\varepsilon}, \omega\right)$, which is a re-scaled version of $q(x, \omega)$, a stationary mean zero random field defined on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with (possibly multi-dimensional) parameter $x \in \mathbb{R}^d$. The equations are parametrized by the realization $\omega \in \Omega$ and by the small parameter $0 < \varepsilon \ll 1$ modeling the correlation length of the random media. We denote by $\mathbb{E}$ the mathematical expectation with respect to the probability measure $\mathbb{P}$.

In Chapter 5 we have seen several examples, except that we did not consider the random structures of the potential q. Nevertheless, the results on the solution operators of these equations are still valid as long as, e.g., $\tilde{q}_\varepsilon$ is non-negative. We will impose this condition on the random coefficient, so that (6.1) is well-posed.

Under mild conditions, the homogenization of this random equation is obtained by averaging $\tilde{q}_\varepsilon$, that is, replacing $\tilde{q}_\varepsilon$ by its low frequency part q_0. Assuming that q_0 has nice properties such as uniform continuity, the results in the previous chapter apply and show that the homogenized equation has many nice properties. The main objective of this chapter is to investigate the corrector in this random homogenization, i.e., the difference between the random solution u_ε and the solution u_0 to the homogenized equation. In particular, we will capture the limiting distribution of the fluctuations in the corrector, and estimate the deterministic terms in the corrector that are larger than the fluctuation.

In the next section, we set up the main assumptions of the elliptic equations for which the corrector theory developed here works in general. We emphasize two important factors: the decorrelation rate of the random potential, and the singularity of the Green's function of the equation. These two factors together determine the size of the fluctuation, the stochasticity (the relative strength between the mean and the fluctuation), and the limiting distribution of the corrector.

6.1 Set-up of Corrector Theory in Random Homogenization of Elliptic Equations

We rewrite the random equation as follows, with low and high frequency parts of the potential separated,

$$\begin{cases} P(x,D)u_\varepsilon(x,\omega) + \left(q_0(x) + q_\varepsilon(x,\omega)\right) u_\varepsilon(x,\omega) = f(x), & x \in X, \\ u_\varepsilon(x,\omega) = 0, & x \in \partial X. \end{cases} \tag{6.2}$$

The corrector theory we develop in this chapter works for general elliptic operator $P(x, D)$ that satisfies the following conditions.

(P1) Suppose that $\tilde{q}(x)$ is a non-negative bounded function. Then the differential operator $P(x, D) + \tilde{q}$, with Dirichlet boundary condition, is invertible in $L^2(X)$. Further, the norm of the solution operator, as a transform on $L^2(X)$, can be bounded independent of the smoothness of $\tilde{q}$.

(P2) Suppose that $q_0(x)$ is a non-negative function continuous on $\overline{X}$. Then the Green's function $G(x, y)$ associated to the differential operator $P(x, D) + q_0$, with Dirichlet boundary conditions, satisfies

$$|G(x, y)| \leq \frac{C}{|x - y|^{d-\beta}}, \tag{6.3}$$

for some bounded positive constant C and some real number $\beta \in (0, d)$, which measures how singular the Green's function is near the diagonal $x = y$.

(P3) Suppose that $q_0(x)$, $f(x)$ and the boundary ∂X are sufficiently regular, then the solution u of $(P(x, D) + q_0)u = f$ with Dirichlet boundary condition is also regular, say continuous. Here the subscript D denotes Dirichlet boundary condition.

We verify that the equations considered in Chapter 5 are typical examples that satisfies the above conditions. We remark also: the theory in this chapter works also if the potential in (P2) is replaced by logarithmic function, as one can easily check following our derivation.

The main assumptions on the random process $q(x, \omega)$ are as follows.

a. Short-range random media.

As before, by short-range correlation we mean that the correlation function of the random media is an integrable function. The main assumptions, which include additional restrictions, are listed as follows.

(S1) The random field $q(x,\omega)$ is stationary, mean-zero, and uniformly bounded so that $q_0 + q(x,\omega)$ is non-negative.

(S2) The random field $q(x,\omega)$ is strong mixing, with ρ-mixing coefficient $\rho(r)$ satisfying the decay rate in (2.10), i.e., $\rho(r) \sim o(r^{-d})$ for large r.

(S3) The random field $q(x,\omega)$ has controlled fourth order cumulants with integrable control functions ϕ_p, in the sense of Definition 2.30. Further, assume these control functions are integrable in each variable.

We observe that (S2) implies that the random field is ergodic. Further, the correlation function $R(x)$ is integrable because

$$R(x) = \mathrm{Corr}\big(q(0), q(x)\big)\mathrm{Var}\big(q(0)\big) \le \rho(|x|)\|q\|^2_{L^\infty}, \tag{6.4}$$

and the last member is integrable. In particular, $\sigma^2 := \int_{\mathbb{R}^d} R(x)dx$ as defined in (2.5) is finite and we assume that $\sigma > 0$.

b. Long-range random media.

We also consider the case when $q(x,\omega)$ has long-range correlation. The main assumptions in this case are:

(A1) $q(x)$ is defined as $q(x) = \Phi(g(x))$, where $g(x)$ is a centered stationary Gaussian random field with unit variance. Furthermore, the correlation function of $g(x)$ has heavy tail of the form:

$$R_g(x) := \mathbb{E}\{g(y)g(y+x)\} \sim \kappa_g |x|^{-\alpha} \text{ as } |x| \to \infty, \tag{6.5}$$

for some positive constant κ_g and some real number $\alpha \in (0,d)$.

(A2) The function $\Phi : \mathbb{R} \to \mathbb{R}$ satisfies $|\Phi| \le \gamma \le q_0$ and

$$\int_{\mathbb{R}} \Phi(s) e^{-\frac{s^2}{2}} ds = 0. \tag{6.6}$$

Further, Φ as Hermite rank one; see Notes of Chapter 2.

(A3) The function Φ satisfies

$$\int_{\mathbb{R}} |\hat{\Phi}(\xi)|\big(1+|\xi|^3\big) < \infty,$$

where $\hat{\Phi}$ denotes the Fourier transform of Φ.

The upper bound of Φ above ensures that $|q(x)| \le \gamma$. Consequently, $q_0 + q_\varepsilon$ is non-negative, and (6.2) is well-posed almost surely with solution operator bounded uniformly with respect to q. Due to the construction above and (6.6), $q(x)$ is mean-zero and stationary, and has long-range correlation function that decays like $|x|^{-\alpha}$ as we have shown in Section 2.4.

6.2 Corrector Theory for Elliptic Equations in Short-Range Media, through a Random Robin Problem

In this section, we develop the corrector theory for elliptic equation of the form (6.2), in the case when the random part of the potential, i.e., $q(x,\omega)$, has short-range correlations.

With the "uniform" set-up of the equation in the previous section, it is possible to develop corrector theory for general differential operators satisfying the aforementioned conditions. In fact, we will do so for the case when $q(x,\omega)$ has long-range correlations. In the current case, however, we establish the theory through an explicit example—the Robin boundary problem for steady diffusion in the half space. That is, (5.7) and its equivalent formulation (5.9) obtained by applying the Dirichlet-to-Neumann map.

We add a random perturbation in the potential term of this Robin problem, and consider

$$\begin{cases} (-\Delta + \lambda^2) u_\varepsilon(x,\omega) = 0, & x = (x', x_n) \in \mathbb{R}^n_+, \\ \dfrac{\partial}{\partial \nu} u_\varepsilon + (q_0 + q(\dfrac{x'}{\varepsilon}, \omega)) u_\varepsilon = f(x'), & x = (x', 0) \in \partial\mathbb{R}^n_+. \end{cases} \tag{6.7}$$

As before, the boundary $\partial\mathbb{R}_+$ is identified with $\mathbb{R}^d$ with $d = n-1$; λ and q_0 are assumed to be positive constants. Using the DtN map as before, the above equation is equivalent with the following pseudo-differential equation about the trace of u_ε, which for simplicity is still denoted by u_ε:

$$(\sqrt{-\Delta + \lambda^2} + q_0 + q_\varepsilon(x,\omega))u_\varepsilon = f, \tag{6.8}$$

where Δ is the Laplacian on $\mathbb{R}^d$, obtained from the Laplacian on $\mathbb{R}^n$ with $\partial^2_{x_n}$ eliminated. The notation q_ε is simply $q\left(\frac{x}{\varepsilon}\right)$ as usual.

This type of boundary problems have applications in chemical physics and biology. For instance, in the context of cell communication by diffusing signals, the equation in (6.7) models the diffusion of signaling molecules in a bulk of extracellular medium which is covered at the bottom by a monolayer of cells forming a layer of epithelium. The cells on the epithelium layer can secrete and absorb signaling molecules, depending on levels of gene expression in the cells. The boundary condition in (6.7) models the action between the cells and the signaling molecules.

Now we state the main results for the random Robin problem about in the version of $d = 2$. This dimension is the physical dimension concerning the biological application.

Theorem 6.1. *Let u_ε and u solve* (6.8) *and* (5.9) *respectively and $d = 2$. Suppose λ, q_0 in those equations are positive constants and f is in $L^2(\mathbb{R}^d)$. Assume that the random field $q(x,\omega)$ satisfies* (S1) *and $R(x)$ is integrable. Then we have*

$$\mathbb{E}\|u_\varepsilon - u\|^2_{L^2(\mathbb{R}^2)} \le C\varepsilon^2|\log\varepsilon|\|f\|^2_{L^2}, \tag{6.9}$$

where the constant C only depends on the parameter λ, q_0, dimension d and $\|R\|_{L^1}$, but not on ε.

This theorem says u_ε and u are close in the energy norm $L^2(\Omega, L^2(\mathbb{R}^d))$. Let us denote the corrector by ξ_ε. We can decompose it into two parts as follows:

$$\xi_\varepsilon = (\mathbb{E}\{u_\varepsilon\} - u) + (u_\varepsilon - \mathbb{E}\{u_\varepsilon\}). \tag{6.10}$$

We call them the *deterministic corrector* and the *stochastic corrector*, respectively.

For the deterministic corrector, we can calculate its limit explicitly. Let us define

$$\tilde{R} := \int_{\mathbb{R}^2} \frac{R(y)}{2\pi|y|} dy. \tag{6.11}$$

Since R is integrable and bounded, this integral is finite. With this notation and recall that $\mathcal{G}$ denotes the solution operator of (5.9), we have the following theorem on the limit of the deterministic corrector.

Theorem 6.2. *Let u_ε and u solve* (6.8) *and* (5.9) *respectively and $d = 2$. Let $q(x, \omega)$ satisfy the same conditions as in the previous theorem. Then we have,*

$$\lim_{\varepsilon \to 0} \frac{\mathbb{E}\{u_\varepsilon\} - u}{\varepsilon} = \tilde{R}\mathcal{G}u. \tag{6.12}$$

Here the limit is taken in the weak *sense. That is, for an arbitrary test function $M \in C_c^\infty(\mathbb{R}^2)$, the real number $\varepsilon^{-1}\langle M, \mathbb{E}\{\xi_\varepsilon\}\rangle$ converges to $\langle \mathcal{G}M, \tilde{R}u\rangle$.*

Note that $\mathcal{G}$ is self-adjoint. In general, the solution operator of (6.1) is not self-adjoint, and the term $\mathcal{G}M$ above should be replaced by $\mathcal{G}^*M$ where $\mathcal{G}^*$ denotes the adjoint operator.

For the stochastic corrector, we have the following central limit theorem.

Theorem 6.3. *Let u_ε and u solve* (6.8) *and* (5.9) *respectively and $d = 2$. Let $q(x, \omega)$ satisfies* (S1)-(S3)*. Then:*

$$\frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\varepsilon} \xrightarrow{\text{distribution}} -\sigma \int_{\mathbb{R}^2} G(x-y)u(y)dW_y, \tag{6.13}$$

where σ is defined in (2.5) *and W_y is the standard multi-parameter Wiener process in $\mathbb{R}^2$. The convergence here is weakly in $\mathbb{R}^2$ and in probability distribution.*

Remark 6.4. We refer the reader to [73] for theory of multi-parameter processes. Also, from Theorem 6.2 it is clear that we can replace $\mathbb{E}\{u_\varepsilon\}$ in the theorem above by $u + \varepsilon\tilde{R}\mathcal{G}u$ since the rest is of order smaller than ε.

6.2.1 Homogenization and convergence rate

In this section, we prove the first two main theorems. The proof works for dimensions larger than three as well, and in that case the $\varepsilon^2|\log\varepsilon|$ in (6.9) should be replaced by ε^2. Let us denote by $\xi_\varepsilon = u_\varepsilon - u$ the corrector. Now subtract (5.9) from (6.8) to get

$$(\sqrt{-\Delta + \lambda^2} + q_0 + q_\varepsilon)\xi_\varepsilon = -q_\varepsilon u. \tag{6.14}$$

Recall that $\mathcal{G}$ is the solution operator $(\sqrt{-\Delta+\lambda^2} + q_0)^{-1}$, and $\mathcal{G}_\varepsilon$ is the solution operator with random impedance. Therefore, the above equation says $\xi_\varepsilon = -\mathcal{G}_\varepsilon q_\varepsilon u$. Unfortunately, $\mathcal{G}_\varepsilon$ is not as explicit as $\mathcal{G}$. Nevertheless, we will show shortly that $-\mathcal{G}q_\varepsilon u$ is the leading term of $-\mathcal{G}_\varepsilon q_\varepsilon u$ and hence it suffices to estimate the former. Let us assign it the following notation;

$$\chi_\varepsilon := -\mathcal{G}q_\varepsilon u. \tag{6.15}$$

We have the following estimate.

Lemma 6.5. *Let u solve* (5.9) *and χ_ε be defined as above and $d = 2$. Assume that the coefficients λ, q_0, and the random field $q(x,\omega)$ satisfy the same conditions as in Theorem 6.1. Then we have*

$$\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \le C\varepsilon^2|\log\varepsilon|\ \|u\|_{L^2}^2, \tag{6.16}$$

where the constant C depends on λ, q_0 and $\|R\|_{L^1}$ but not on u or ε.

Proof. 1. We first express $\|\chi_\varepsilon\|_{L^2}^2$ as a triple integral of the form

$$\int_{\mathbb{R}^{3d}} G(x-y)q_\varepsilon(y)u(y)G(x-z)q_\varepsilon(z)u(z)d[yzx].$$

Here and in the sequel, the short-hand notation $d[x_1\cdots x_n]$ is the same as $dx_1\cdots dx_n$. Take expectation and use the definition of $R(x)$ to obtain

$$\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 = \int_{\mathbb{R}^{3d}} G(x-y)G(x-z)R(\frac{y-z}{\varepsilon})u(y)u(z)d[yzx].$$

2. We integrate in x first. Use the estimate (5.24) to replace the Green's functions by potentials of the form $e^{-\lambda'|x-y|}/|x-y|$; then apply Lemma 3.12 to bound the integration in x of these potentials. We obtain

$$\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \le C\int_{\mathbb{R}^{2d}} e^{-\lambda'|y-z|}\big(\big|\log|y-z|\big|+1\big)\Big|R(\frac{y-z}{\varepsilon})u(y)u(z)\Big|d[yz]. \tag{6.17}$$

Now change variable $(y-z)/\varepsilon \to y$. This change of variable yields a Jacobian ε^d and the integral on the right hand side becomes

$$\varepsilon^d\int_{\mathbb{R}^{2d}} e^{-\varepsilon\lambda'|y|}\big(\,|\log|y|+\log\varepsilon|+1\big)\Big|R(y)u(z+\varepsilon y)u(z)\Big|d[yz].$$

3. Now, bound the exponential term by 1, and integrate in z. Use Cauchy-Schwarz to get

$$\int_{\mathbb{R}^d}\big|u(z+\varepsilon y)u(z)\big|dz \le \|u\|_{L^2}\|u(\cdot+\varepsilon y)\|_{L^2} = \|u\|_{L^2}^2. \tag{6.18}$$

Therefore, we have

$$\mathbb{E}\|\chi_\varepsilon\|_{L^2}^2 \le C\varepsilon^d\|u\|_{L^2}^2\int_{\mathbb{R}^d}\big(\,|\log|y||+1+|\log\varepsilon|\big)\big|R(y)\big|dy.$$

Recall that $R(y)$ behaves like $|y|^{-d-\delta}$ for some positive δ; see (2.10) and (6.4). Hence the function $(|\log |y|| + 1)|R|$ is integrable. Since $d = 2$ the integral above is

$$C\varepsilon^2 |\log \varepsilon| \cdot \|u\|_{L^2}^2 \|R\|_{L^1} + O(\varepsilon^2).$$

This completes the proof. We also see that the constant C only depends on λ, q_0 and $\|R\|_{L^1}$.
□

Theorem 6.1 now follows if we can control $\|\xi_\varepsilon - \chi_\varepsilon\|_{L^2}$. From (6.15) we see

$$(\sqrt{-\Delta + \lambda^2} + q_0 + q_\varepsilon)\chi_\varepsilon = -q_\varepsilon u + q_\varepsilon \chi_\varepsilon.$$

Subtract this equation from (6.14); we get an equation for $\xi_\varepsilon - \chi_\varepsilon$. Apply $\mathcal{G}_\varepsilon$ on this equation to get

$$\xi_\varepsilon = \chi_\varepsilon - \mathcal{G}_\varepsilon q_\varepsilon \chi_\varepsilon. \tag{6.19}$$

The following proof relies on this expression and the fact that the operator $\mathcal{G}_\varepsilon$ is bounded uniformly in ε and ω as we have emphasized in Remark 5.7.

Proof of of Theorem 6.1. From the expression (6.19) we have,

$$\|u_\varepsilon - u\|_{L^2} \le \|\chi_\varepsilon\|_{L^2} + \sup_{\omega \in \Omega} \|\mathcal{G}_\varepsilon\|_{\mathcal{L}} \|q\|_{L^\infty(\Omega \times \mathbb{R}^d)} \|\chi_\varepsilon\|_{L^2}.$$

Due to the uniform bound of q_ε and Corollary 5.6, we have $\|q\|_{L^\infty} \le q_0$ and $\|\mathcal{G}_\varepsilon\|_{L^\infty(\Omega, \mathcal{L}(L^2))} \le \min\{1, \lambda\}^{-1}$. We will denote the products of the two constants by C. Then we have

$$\|u_\varepsilon - u\|_{L^2} \le (1 + C)\|\chi_\varepsilon\|_{L^2}.$$

Square both sides and take expectation; then apply Lemma 6.5 to get

$$\mathbb{E}\{\|u_\varepsilon - u\|_{L^2}^2\} \le C\mathbb{E}\{\|\chi_\varepsilon\|_{L^2}^2\} \le C\varepsilon^2 |\log \varepsilon| \cdot \|u\|_{L^2}^2.$$

Now use Corollary 5.6 to replace the L^2 norm of u by that of f. Again, all constants involved do not depend on ε. This completes the proof. □

To prove Theorem 6.2 and 6.3, i.e., to characterize the limits of the deterministic and stochastic correctors, we first express ξ_ε as a sum of three terms with increasing order in q_ε. To this end, move the term $q_\varepsilon \xi_\varepsilon$ in (6.14) to the right hand side, and then apply $\mathcal{G}$ on it. We get

$$\xi_\varepsilon = -\mathcal{G} q_\varepsilon u - \mathcal{G} q_\varepsilon \xi_\varepsilon.$$

Iterate this formula one more time to get

$$\xi_\varepsilon = -\mathcal{G} q_\varepsilon u + \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u + \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon. \tag{6.20}$$

Note that the limits in both theorems are taken weakly in space, so we consider an arbitrary test function M, e.g. in C_c^∞, and integrate the above formula with M. We get

$$\langle \xi_\varepsilon, M\rangle = -\langle \mathcal{G} q_\varepsilon u, M\rangle + \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u, M\rangle + \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle. \tag{6.21}$$

Defining $m := \mathcal{G} M$, the last term can be written as $\langle q_\varepsilon \xi_\varepsilon, \mathcal{G} q_\varepsilon m\rangle$ since $\mathcal{G}$ is self-adjoint. Using this notation we now prove the second main theorem.

Proof of Theorem 6.2. Take expectation on the weak formulation (6.21). The first term vanishes since q_ε is mean zero. To estimate the third term, we observe that

$$\big|\langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle\big| = \big|\langle q_\varepsilon \xi_\varepsilon, \mathcal{G} q_\varepsilon m\rangle\big| \le \|q_\varepsilon\|_{L^\infty} \|\xi_\varepsilon\|_{L^2} \|\mathcal{G} q_\varepsilon m\|_{L^2}.$$

By assumption (S1), $\|q_\varepsilon\|_{L^\infty}$ is bounded by q_0. After taking expectations on both sides and using Cauchy-Schwarz on the right hand side, we obtain

$$\mathbb{E}\big|\langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle\big| \le C\big(\mathbb{E}\{\|\xi_\varepsilon\|^2\}\ \mathbb{E}\{\|\mathcal{G} q_\varepsilon m\|^2\}\big)^{1/2} \le C\varepsilon^2 |\log \varepsilon| \cdot \|u\|_{L^2} \|m\|_{L^2}, \tag{6.22}$$

where the last inequality follows from Theorem 6.1 and Lemma 6.5. In the limit, this term is much smaller than ε.

Now we calculate the expectation of the second term in (6.21), which can be written as:

$$\mathbb{E}\langle q_\varepsilon u, \mathcal{G} q_\varepsilon m\rangle = \int_{\mathbb{R}^{2d}} G(x-y) R(\frac{x-y}{\varepsilon}) u(x) m(y) d[xy]. \tag{6.23}$$

As in the proof of Lemma 6.5, we change variable $(x-y)/\varepsilon$ to x. The integral above now becomes

$$\varepsilon^d \int_{\mathbb{R}^{2d}} G(\varepsilon x) R(x) u(y+\varepsilon x) m(y) d[xy] \leq \|u\|_{L^2} \|m\|_{L^2} \int_{\mathbb{R}^d} \varepsilon^d G(\varepsilon|x|)|R(x)| dx. \tag{6.24}$$

The last equality is obtained by integrating in y and applying the same technique as in (6.18). Recalling Lemma 5.8 and $d=2$, G can be decomposed into three terms. We have

$$\varepsilon^2 G(\varepsilon|x|) = \frac{\varepsilon^2}{2\pi}\left(\frac{\exp(-\lambda\varepsilon|x|)}{\varepsilon|x|} - q_0 K_0(\lambda\varepsilon|x|) + G_r(\varepsilon|x|)\right).$$

Since K_0 only has logarithmic singularity at the origin and G_r is uniformly bounded as we have seen in Lemma 5.8, the last two terms above are of order $\varepsilon^2|\log\varepsilon|$ and ε^2 respectively. Their contributions to (6.24) are negligible.

Hence the leading term in (6.24) is

$$\varepsilon \int_{\mathbb{R}^2} \frac{e^{-\varepsilon\lambda|x|}}{2\pi|x|} R(x) u(y) m(y+\varepsilon x) dy dx. \tag{6.25}$$

Taking the limit and recalling the definition of $\tilde{R}$ in (6.11), we see that this term is

$$\varepsilon\tilde{R}\langle u, m\rangle + o(\varepsilon) = \varepsilon\tilde{R}\langle \mathcal{G}u, M\rangle + o(\varepsilon).$$

This completes the proof. □

6.2.2 Convergence in distribution of random correctors

Our proof of the third theorem also relies on the formula (6.21). The plan is as follows. First, we show that the leading term in the stochastic corrector $\xi_\varepsilon - \mathbb{E}\{\xi_\varepsilon\}$ is the first term in (6.21); this is done by showing that the variances of the other terms are small. Then we verify that the first term has a limiting distribution that can be written as the right hand side of (6.13); this step is rather standard and follows from a generalized central limit theorem, i.e., Theorem 2.15. For the moment, let us assume the following lemma and prove Theorem 6.3.

Lemma 6.6. *Let u solve (5.9) with $d = 2$ and M be a test function in $C_c^\infty(\mathbb{R}^d)$. Assume that the random field $q(x, \omega)$ satisfies the same conditions as in Theorem 6.3. Then we have the following estimate:*

$$\mathrm{Var}\ \ \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u, M\rangle \le C\varepsilon^{d+1}, \tag{6.26}$$

where C depends on λ, q_0, $\|u\|_{L^2}$, $\|G\|_{L^1}$, $\|M\|_{L^1}$, $\|M\|_{L^\infty}$, dimension d, $\|\phi_p\|_{L^1}$ and $\|\phi_p\|_{L^\infty}$ in (2.33), but not on ε.

Proof of Theorem 6.3. 1. We rewrite formula (6.21) as

$$\langle u_\varepsilon - u + \mathcal{G} q_\varepsilon u, M\rangle = \langle \mathcal{G}_\varepsilon q_\varepsilon \mathcal{G}_\varepsilon q_\varepsilon u, M\rangle + \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle.$$

Take expectation on both sides and note that $\mathbb{E}(\mathcal{G} q_\varepsilon u) = 0$; then we have

$$\langle \mathbb{E}\{u_\varepsilon\} - u, M\rangle = \mathbb{E}\langle \mathcal{G}_\varepsilon q_\varepsilon \mathcal{G}_\varepsilon q_\varepsilon u, M\rangle + \mathbb{E}\langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle.$$

Subtract this equation from the preceding one and divide both sides by ε; take expectation on the absolute value of both sides, and use basic inequalities to get

$$\mathbb{E}\Big|\langle \frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\varepsilon} + \frac{\mathcal{G} q_\varepsilon u}{\varepsilon}, M\rangle\Big| \le \frac{1}{\varepsilon}\big(\ \mathrm{Var}\ \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u, M\rangle\big)^{\frac{1}{2}} + \frac{2}{\varepsilon}\mathbb{E}\{|\langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon \xi_\varepsilon, M\rangle|\}.$$

The last term is of order $\varepsilon|\log\varepsilon|$ thanks to the estimate (6.22), and the next-to-last is of order $\sqrt{\varepsilon}$ due to (6.26). Therefore the right hand side above vanishes in the limit. This shows convergence of $\varepsilon^{-1}\langle u_\varepsilon - \mathbb{E}\{u_\varepsilon\}, M\rangle$ to $-\varepsilon^{-1}\langle \mathcal{G}q_\varepsilon u, M\rangle$ in $L^1(\Omega)$ which in turn implies convergence in distribution. Hence, we only need to characterize the asymptotic distribution of the latter term.

2. The random variable $\varepsilon^{-1}\langle \mathcal{G}q_\varepsilon u, M\rangle$, which is the same as $\varepsilon^{-1}\langle q_\varepsilon u, m\rangle$ where $m = \mathcal{G}M$, is of the form of an oscillatory integral. Let $v(y)$ denote $u(y)m(y)$; it is an L^2 function. We want

$$\int_{\mathbb{R}^2} \frac{1}{\varepsilon} q(\frac{y}{\varepsilon}) v(y) dy \xrightarrow{\text{distribution}} \sigma \int_{\mathbb{R}^2} v(y) dW_y, \tag{6.27}$$

where W_y is the standard two-parameter Wiener process as in Theorem 6.3. If the integration region is a bounded set in $\mathbb{R}^d$, this is precisely the convergence result in Theorem 2.15. Here, since we assumed that M is compactly supported, v decays fast and is in $L^2(\mathbb{R}^d)$, and we obtain (6.27) by applying the theorem on the ball with radius B and sending B to infinity. This completes the proof of the theorem. $\square$

It remains to prove the preceding lemma.

Proof of Lemma 6.6. We express random variable $\langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u, M\rangle$, which equals $\langle q_\varepsilon u, \mathcal{G}q_\varepsilon m\rangle$ where $m = \mathcal{G}M$, as the following integral.

$$I := \int_{\mathbb{R}^{2d}} u(x)m(y)G(x-y)q_\varepsilon(x)q_\varepsilon(y)d[xy].$$

Take the variance of this variable. Denote by ϑ the joint cumulant. We have the following expression for $\mathrm{Var}\{I\}$, i.e., $\mathbb{E}\{I^2\} - (\mathbb{E}\{I\})^2$;

$$\begin{aligned}\mathrm{Var}\{I\} = \int_{\mathbb{R}^{4d}} & u(x)m(y)u(x')m(y')G(x-y)G(x'-y')\Big[\vartheta\{q_\varepsilon(x), q_\varepsilon(y), q_\varepsilon(x'), q_\varepsilon(y')\} \\ & + R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon}) + R(\frac{x-y'}{\varepsilon})R(\frac{y-x'}{\varepsilon})\Big] d[xyx'y'].\end{aligned}$$

Then we identify x, y, x', y' with x_1, x_2, x_3, x_4. Let $\mathcal{U}$ and $\mathcal{U}^*$ be the sets defined in (2.32) and the paragraph below it. Recall that the joint cumulant $\vartheta\{q_\varepsilon(x_i)\}_{i=1}^4$ satisfies (2.33) with $\phi_p \in L^1 \cap L^\infty(\mathbb{R}^d \times \mathbb{R}^d)$; we have the following bound for $\mathrm{Var}\{I\}$:

$$\begin{aligned}\int_{\mathbb{R}^{4d}} &|u(x)m(y)u(x')m(y')|G(x-y)G(x'-y')\Big(\sum_{p\in\mathcal{U}^*}\phi_p(\frac{x_{p(1)}-x_{p(2)}}{\varepsilon},\frac{x_{p(3)}-x_{p(4)}}{\varepsilon})\\ &+\Big|R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})\Big|+\Big|R(\frac{x-y'}{\varepsilon})R(\frac{y-x'}{\varepsilon})\Big|\Big)d[xyx'y'].\end{aligned} \tag{6.28}$$

Let us denote the contributions of the last two terms in the parenthesis above by J_2 and J_3 respectively, and denote the contribution of the other term by J_1. We observe that the variables in the $R\otimes R$ functions are independent with the variables in the Green's functions, while this is not the case for the variables in the ϕ_p functions.

We first estimate J_2. It has the following expression;

$$J_2 := \int_{\mathbb{R}^{4d}} \big|u(x)m(y)u(x')m(y')G(x-y)G(x'-y')R(\frac{x-x'}{\varepsilon})R(\frac{y-y'}{\varepsilon})\big|d[xyx'y'].$$

Perform a change of variables as follows:

$$x \to x,\ \frac{x-x'}{\varepsilon} \to x',\ \frac{y-y'}{\varepsilon} \to y',\ x-y \to y.$$

This change of variables yields a Jacobian ε^{2d} and the integral above becomes

$$\varepsilon^{2d}\int_{\mathbb{R}^{4d}} \big|u(x)m(x-y)u(x-\varepsilon x')m(y-\varepsilon y')G(y)G(y-\varepsilon(x'-y'))R(x')R(y')\big|d[xyx'y']. \tag{6.29}$$

Now we observe that the function $m = \mathcal{G}M$ is uniformly bounded as follows;

$$\|m\|_{L^\infty} \le C(\|M\|_{L^\infty} + \|M\|_{L^1}). \tag{6.30}$$

Indeed, we use the estimate (5.24) for the Green's function and have

$$\begin{aligned} m(x) &= \int_{\mathbb{R}^d} G(x-y)M(y)dy \le C\int_{\mathbb{R}^d} \frac{M(y)}{|x-y|^{d-1}}dy \\ &\le C\left(\|M\|_{L^\infty} \int_{B_1(x)} \frac{1}{|x-y|^{d-1}}dy + \int_{B_1^c(x)} M(y)dy \right). \end{aligned}$$

Here we denote by $B_1(x)$ the unit ball centered at x, and by $B_1^c(x)$ its complement. The integral inside $B_1(x)$ is bound by $\pi^{\lfloor \frac{d}{2} \rfloor}$, and the integral on $B_1^c(x)$ is bounded by $\|M\|_{L^1}$. Hence we obtain (6.30). Use this bound to control the m functions in (6.29). Integrate in x and use (6.18) to control the u functions. Integrate in y for the two Green's function and view the integration as a convolution. Use (5.24) to bound them by potentials of the form $e^{-\lambda'|x|}/|x|$, and use the second inequality in (3.26) of Lemma 3.12 to get

$$\int_{\mathbb{R}^d} G(y)G(y-\varepsilon(x'-y'))dy \le Ce^{-\lambda'\varepsilon|x'-y'|}\Big(|\log(\varepsilon|x'-y'|)| \cdot \mathbf{1}_{\{\varepsilon|x'-y'|\le 1\}} + 1 \Big),$$

where $\mathbf{1}$ is the indicator function of a set. Therefore, after controlling u, m, and G, we get

$$\begin{aligned} J_2 \le C\varepsilon^{2d}\|u\|_{L^2}^2\|m\|_{L^\infty}^2 \int_{\mathbb{R}^{2d}} &\Big(\big|\log(\varepsilon|x'-y'|)\big|\mathbf{1}_{\{\varepsilon|x'-y'|\le 1\}} + 1 \Big) \\ &\times |R(x')|\cdot|R(y')|d[x'y']. \end{aligned} \tag{6.31}$$

The constant one in the parenthesis hence has a contribution of order ε^{2d} since $\|R\|_{L^1}$ is finite. For the logarithmic term, we observe that

$$\sup_{0<r\le 1} r^{d-1}|\log r| \le \frac{e^{-1}}{d-1}, \text{ for } d \ge 2. \tag{6.32}$$

Therefore, we have

$$\big|\log(\varepsilon|x'-y'|)\big|\mathbf{1}_{\{\varepsilon|x'-y'|\le 1\}} \le \frac{e^{-1}}{(d-1)\varepsilon^{d-1}|x'-y'|^{d-1}}\mathbf{1}_{\{\varepsilon|x'-y'|\le 1\}}.$$

The contribution of the logarithm term in (6.31) is bounded by

$$C\varepsilon^{d+1}\|u\|_{L^2}^2\|m\|_{L^\infty}^2\int_{\mathbb{R}^{2d}}\frac{|R(x')|\cdot|R(y')|}{|x'-y'|^{d-1}}d[x'y'].$$

Now apply the Hardy-Littlewood-Sobolev inequality, e.g. [78, §4.3], to get

$$\left|\int_{\mathbb{R}^{2d}}\frac{|R(x')|\cdot|R(y')|}{|x'-y'|^{d-1}}\right|\le C(\frac{2d}{d+1},d-1)\|R\|^2_{L^{\frac{2d}{d+1}}}. \tag{6.33}$$

Since $R\in L^1\cap L^\infty$, it is certainly in $L^{\frac{2d}{d+1}}$. We have proved that

$$J_2\le C\varepsilon^{d+1}\|u\|_{L^2}^2\|m\|_{L^\infty}^2\|R\|_{L^\infty}^{\frac{3}{2}}\|R\|_{L^1}^{\frac{1}{2}}+O(\varepsilon^{2d}), \tag{6.34}$$

where $d=2$. Similarly, J_3 can be shown to be of size smaller than ε^{d+1} as well in dimension two.

Now we consider J_1. There are $C_6^2-3=12$ terms that appear in the sum over $p\in\mathcal{U}^*$ in (6.28), and they can be divided into two groups. In the first group, the function ϕ_p shares a variable with one of the Green's functions; in the second group, the variable of one of the Green's functions is a linear combination of the two variables of the ϕ_p function.

We first consider a typical term from the first group and still call it J_1; it has the following expression:

$$J_1:=\int_{\mathbb{R}^{4d}}\left|G(x-y)G(x'-y')\phi_p(\frac{x-y}{\varepsilon},\frac{x-x'}{\varepsilon})u(x)m(y)u(x')m(y')\right|d[xyx'y'],$$

Note that the $x-y$ variable is shared by the first Green's function and ϕ_p. We perform the following change of variables:

$$x\to x,\ \frac{x-x'}{\varepsilon}\to x',\ \frac{x-y}{\varepsilon}\to y,x'-y'\to y'.$$

The Jacobian is again ε^{2d}, and then the integral becomes

$$\varepsilon^{2d}\int_{\mathbb{R}^{4d}}\big|u(x)m(x-\varepsilon y)u(x-\varepsilon x')m(x'-y')G(y')G(\varepsilon y)\big|\phi_p(y,x')d[xyx'y'].$$

Use (6.30) to control the m functions; integrate in x and use (6.18) to control the u functions; integrate in y' to control the first Green's function. We obtain the following bound for J_2.

$$J_2 \le C\varepsilon^{2d}\|u\|_{L^2}^2\|m\|_{L^\infty}^2\|G\|_{L^1}\int_{\mathbb{R}^{2d}}\frac{1}{(\varepsilon|y|)^{d-1}}\phi_p(y,x')d[yx'], \tag{6.35}$$

where we have used (5.24) for the Green's function. The scaling ε^{-d+1} resulting from the Green's function combined with the Jacobian ε^{2d} indicates that J_2 is of size ε^{d+1} once we control the following integral:

$$\int_{\mathbb{R}^{2d}}\frac{\phi_p(y,x')}{|y|^{d-1}}d[yx'].$$

Indeed, this integral is finite since $|y|^{d-1}$ is integrable near the origin and ϕ_p is integrable at infinity. To summarize we have

$$J_2 \le C\varepsilon^{d+1}\|u\|_{L^2}^2\|m\|_{L^\infty}^2\|G\|_{L^1}\left\|\frac{\phi_p(y,x')}{|y|^{d-1}}\right\|_{L^1}. \tag{6.36}$$

For a typical term from the second group in the sum over $p\in\mathcal{U}^*$ in (6.28), we can apply the same procedure exactly and in (6.35) we will have $|x'-y|^{d-1}$ on the denominator in the integral, and we can control the integral as in (6.33). Therefore, the contributions of such terms are also of size ε^{d+1} with $d=2$. This completes the proof. □

6.2.3 General setting with singular Green's function

So far, we only considered dimension $d=2$ which is physical for the Robin problem above. However, the derivation we developed works for elliptic pseudo-differential equations of the form (6.8) in general dimensions. We consider the following pseudo-differential equation with random coefficient:

$$P(x,D)u_\varepsilon(x,\omega) + (q_0(x) + q_\varepsilon(x,\omega))u_\varepsilon = f(x), \tag{6.37}$$

posed on a subset X of $\mathbb{R}^d$ with appropriate boundary condition. As before, $q_\varepsilon(x,\omega) = q(x/\varepsilon,\omega)$ and $q(x,\omega)$ is a stationary, mean zero, finite variance, strong mixing random field defined on $(\Omega,\mathcal{F},\mathbb{P})$, with parameters $x \in \mathbb{R}^d$. By assumption the solution operators,

$$\mathcal{G} := \big(P(x,D) + q_0(x)\big)^{-1}, \quad \mathcal{G}_\varepsilon := \big(P(x,D) + q_0(x) + q_\varepsilon\big)^{-1},$$

are well defined almost everywhere in Ω. Further, as transformations on $L^2(X)$, $\mathcal{G}$ and $\mathcal{G}_\varepsilon$ are bounded for all realizations, and the upper bound of the operator norm is independent of realizations.

Using the same techniques developed in previous sections, we can show that u_ε converges to the solution of a homogenized equation denoted by u in the $L^2(X\times\Omega)$ norm. We can then show that the random corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$ converges weakly and in probability distribution to a Gaussian process with variance of size ε^d. The large components, with size no less than $\varepsilon^{d/2}$, of the deterministic corrector $\mathbb{E}\{u_\varepsilon\} - u$ can also be captured. As in the main body of this paper, we need additional assumptions on some higher-order moments of the random field $q(x,\omega)$ to obtain the last two results.

To be precise, suppose the Green's function $G(x,y)$ has the following decomposition with decreasing singularities,

$$G(x,y) \sim \sum_{j=1}^{N} \frac{c_j(x,y)}{|x-y|^{\gamma_j}} + G_r(x,y). \tag{6.38}$$

Here, N is a finite integer and

$$d > \gamma_1 > \gamma_2 > \cdots > \gamma_N \geq \frac{d}{2}.$$

Let us denote the terms in the sum above as G_j. The functions $\{c_j(x, y)\}$ are uniformly bounded and decay fast enough so that $\{G_j\}$ are integrable if the domain X is unbounded. Further, $G_r(x, y)$ is a term that is both integrable and square integrable (with respect to one of the variables and uniformly in the other variable).

Then, the homogenized equation for (6.37) will be of the same form with q_ε averaged (or removed). In fact, we have the following as an analogy of Theorem 6.7.

$$\mathbb{E}\|u_\varepsilon - u\|_{L^2}^2 \le \begin{cases} C\varepsilon^{2(d-\gamma_1)}\|u\|_{L^2}^2, & \text{if } 2\gamma_1 > d, \\ C\varepsilon^d |\log\varepsilon| \|u\|_{L^2}^2, & \text{if } 2\gamma_1 = d. \end{cases} \tag{6.39}$$

These estimates show that u_ε converges to the homogenized solution u in energy norm. At this stage, we do not need the mixing property or control of higher order moments of $q(x, \omega)$.

Under certain conditions on some moments of the random field, we know that the fluctuations in the corrector are approximately weakly Gaussian and of size $\varepsilon^{d/2}$. To further approximate u_ε, we would like to capture all the terms in the corrector whose means are larger. To do this, we expand u_ε as iterations of $\mathcal{G}$ on random potentials as follows.

$$u_\varepsilon(x) - u = -\mathcal{G}q_\varepsilon\mathcal{G}f + \mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon\mathcal{G}f - \mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon\mathcal{G}f + \cdots + (-\mathcal{G}q_\varepsilon)^k\xi_\varepsilon. \tag{6.40}$$

The order k at which we terminate the iteration is chosen so that $\mathbb{E}\{\|(\mathcal{G}q_\varepsilon)^{k-2}\mathcal{G}M\|_{L^2}^2\} \le \varepsilon^\gamma$ with $\gamma > 2\gamma_1 - d$ for some test function M. Then weakly, the remainder term $(-\mathcal{G}q_\varepsilon)^k\xi_\varepsilon$ is of order less than $\varepsilon^{d/2}$. Hence, the finite terms in (6.40) before the remainder include all the components in the corrector whose means are weakly larger than the random fluctuations. Then it is a tedious routine as shown in the paper to calculate the large deterministic means of these terms and to check that their variances are less than ε^d. As a result, the limiting law of $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$ is given by the limiting law of $\frac{1}{\varepsilon^{d/2}}\mathcal{G}q_\varepsilon u$, which is Gaussian and admits a convenient stochastic integral representation.

As an example, we summarize and compare results for the diffusion equation (5.7) as the dimension n and hence d change.

When $n = 2$ and hence $d = 1$, the Green's function G has logarithmic singularity only and hence $G_j \equiv 0$ in (6.38). As a result, G is square integrable and the problem reduces to a case that is investigated in [9]. In particular, the deterministic corrector $\mathbb{E}\{u_\varepsilon - u\}$ is of size ε and does not show up in Theorem 6.9; in other words, the deterministic corrector is dominated by the random fluctuations, which are of size $\sqrt{\varepsilon}$.

When $n \geq 4$ and hence $d > 2$, then the leading term of the Green's function is given by a modified Bessel potential and has singularity of order $\gamma_1 = d - 1$ at the origin, and $2\gamma_1 > d$. Then the leading term in the deterministic corrector will be of order $\varepsilon^{d-\gamma_1}$, which is larger that $\varepsilon^{d/2}$, In other words, the deterministic corrector dominates the fluctuations., which is of size $\varepsilon^{d/2}$

The physical dimension $n = 3$ considered in the main section turns out to be the critical case when the deterministic corrector is in fact of the same size as the fluctuations, and they are of size ε.

6.3 Corrector Theory for Elliptic Equations in Long-Range Media

The moral of the previous section is: The singularity of the Green's function determine the size of the fluctuation in the corrector, the relative strength of the deterministic and random parts of the corrector, and the limiting distribution of the corrector. We showed this under the assumption that the random potential has short-range correlations, i.e., satisfying (S1)-(S3). In this section, we investigate the importance of another factor: the decorrelation rate of the random potential.

We develop the corrector theory for the general elliptic equation (6.2) which satisfies

(P1)-(P3), and we assume that the random potential $q(x, \omega)$ there satisfies (A1)-(A3). In particular, $\alpha \in (0, d)$ in (6.5) tunes the decorrelation rate.

The first main theorem concerns the homogenization of (6.2). It shows, in particular, how the competition between the de-correlation rate α and the Green's function singularity β affects the convergence rate of homogenization.

Theorem 6.7. *Let u_ε be the solution to* (6.2) *and u_0 be the solution to the same equation with q_ε replaced by its zero average. Assume that $q(x)$ is constructed as in* (A1) *and* (A2) *and that $f \in L^2(X)$. Then, assuming $2\beta < d$, we have*

$$\mathbb{E}\,\|u_\varepsilon - u_0\|^2 \le \|f\|^2 \times \begin{cases} C\varepsilon^\alpha, & \alpha < 2\beta, \\ C\varepsilon^{2\beta}|\log\varepsilon|, & \alpha = 2\beta, \\ C\varepsilon^{2\beta}, & \alpha > 2\beta. \end{cases} \tag{6.41}$$

The constants α and β are defined in (6.5) *and* (6.3) *respectively. When $2\beta \ge d$, the result on the first line above holds. The constant C depends on α, β, γ and the uniform bound on the solution operator of* (6.2).

This theorem states u_ε and u_0 are close in the energy norm $L^2(\Omega, L^2(X))$. The corrector, defined as the difference between these two solutions, can be decomposed as in (6.10). Again, we call the first first part the *deterministic corrector*, and the second mean-zero part the *stochastic corrector*. For the deterministic corrector, we have the following estimates on its size, which depend on α and β.

Theorem 6.8. *Let u_ε, u_0, $q(x)$ and f be as in the previous Theorem. Then for an arbitrary* *function* $\varphi \in L^2(X)$*, we have,*

$$|\langle \mathbb{E}\{u_\varepsilon\} - u_0, \varphi\rangle| \le \|f\|\|\varphi\| \times \begin{cases} C\varepsilon^\alpha, & \alpha < \beta, \\ C\varepsilon^\beta|\log\varepsilon|, & \alpha = \beta, \\ C\varepsilon^\beta, & \alpha > \beta. \end{cases} \tag{6.42}$$

The constant C *depends on the same factors as in the previous theorem.*

The magnitude of the stochastic corrector is always of order $\varepsilon^{\frac{\alpha}{2}}$, as we shall see later in the paper. We deduce from the above theorem that the deterministic corrector can therefore be larger than the stochastic corrector when $\alpha > 2\beta$. To describe the stochastic corrector more precisely, we characterize its limiting distribution. We need to impose the following additional assumptions:

This condition allows one to derive a (non-asymptotic) estimate, (2.33) in the appendix, for the fourth-order moments of $q(x)$, which is a technicality one encounters often in corrector theory. With this assumption, we have:

Theorem 6.9. *Let* u_ε *and* u_0 *solve* (6.2) *and the homogenized equation, respectively. Assume* $f \in L^2(X)$ *and* $q(x)$ *is constructed by* (A1-A2) *with* Φ *satisfying* (A3). *Further, assume* $\alpha < 4\beta$*. Then:*

$$\frac{u_\varepsilon - \mathbb{E}\{u_\varepsilon\}}{\varepsilon^{\alpha/2}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} -\int_X G(x,y)u_0(y)W^\alpha(dy), \tag{6.43}$$

where $W^\alpha(dy)$ *is formally defined to be* $\dot{W}^\alpha(y)dy$ *and* $\dot{W}^\alpha(y)$ *is a Gaussian random field with covariance function given by* $\mathbb{E}\{\dot{W}^\alpha(x)\dot{W}^\alpha(y)\} = \kappa|x-y|^{-\alpha}$*. Here,* $\kappa = \kappa_g\left(\mathbb{E}\{g_0\Phi(g_0)\}\right)^2$ *where* κ_g, Φ *and* g_0 *are defined in* (6.5)*. The convergence is understood in probability distribution and weakly in space; see the following remark.*

Remark 6.10. We refer the reader to [73] for the theory on multi-parameter random processes. What we mean by convergence in probability distribution weakly in space is as follows. We fix an arbitrary natural number N and a set of test functions $\{\varphi_i;\ 1 \le i \le N\}$ in $\mathcal{C}(\overline{X})$. Define $I_i^\varepsilon := \langle \varphi_i, \varepsilon^{-\alpha/2}(u_\varepsilon - \mathbb{E}\{u_\varepsilon\})\rangle$, for $i = 1, \cdots, N$. What (6.43) means is that the N-dimensional random vector $(I_1^\varepsilon, \cdots, I_N^\varepsilon)$ converges in distribution to a centered N-dimensional Gaussian vector $(I_1, \cdots, I_N)$, whose covariance matrix Σ_{ij} is given by

$$\Sigma_{ij} := \int_{X^2} \frac{\kappa}{|y-z|^\alpha} (u_0 \mathcal{G}\varphi_i)(y)(u_0\mathcal{G}\varphi_j)(z) dy dz. \tag{6.44}$$

By the definition of the stochastic integral above, we see I_i is precisely the inner product of φ_i with the right hand side of (6.43).

We deduce from Theorem 6.8 that when $\alpha < 2\beta$ we can replace $\mathbb{E}\{u_\varepsilon\}$ in (6.43) by u_0, since the deterministic corrector is asymptotically smaller. This is no longer the case for $\alpha \ge 2\beta$. The condition $\alpha < 4\beta$ in Theorem 6.9 is due to technical reasons which we explain later. The conclusion of the theorem holds in general if we can prove an estimate on high-order (more than four-order) moments of q, which is not considered in this paper. □

6.3.1 Homogenization and convergence rate

The following lemma is very useful in the sequel.

Lemma 6.11. *Let $\mathcal{G}$ be the Green's operator and $q(x)$ be the random field above. Let f be an arbitrary function in $L^2(X)$. Assume $2\beta < d$. Then, we have:*

$$\mathbb{E}\, \|\mathcal{G}q_\varepsilon f\|^2 \le \|f\|^2 \times \begin{cases} C\varepsilon^\alpha, & \alpha < 2\beta, \\ C\varepsilon^{2\beta}|\log\varepsilon|, & \alpha = 2\beta, \\ C\varepsilon^{2\beta}, & \alpha > 2\beta. \end{cases} \tag{6.45}$$

The constant C depends only on α, β, X, $\|q\|_\infty$ and the bound for $\|\mathcal{G}_\varepsilon\|_{\mathcal{L}}$. If $2\beta \ge d$, then only the first case is necessary.

Proof. The L^2 norm of $\mathcal{G}q_\varepsilon f$ has the following expression:

$$\|\mathcal{G}q_\varepsilon f\|^2 = \int_X \left(\int_X G(x,y)q_\varepsilon(y)f(y)dy \right)^2 dx.$$

After writing the integrand as a double integrals and taking expectation, we have

$$\mathbb{E}\|\mathcal{G}q_\varepsilon f\|^2 = \int_{X^3} G(x,y)G(x,z)R_\varepsilon(y-z)f(y)f(z)dydzdx. \tag{6.46}$$

Use (6.3) to bound the Green's functions. Integrate over x and apply Lemma 3.11. We get

$$\mathbb{E}\|\mathcal{G}q_\varepsilon f\|^2 \le C \int_{X^2} \frac{1}{|y-z|^{d-2\beta}} |R_\varepsilon(y-z)f(y)f(z)|dydz. \tag{6.47}$$

Change variable $(y, y-z) \mapsto (y,z)$. The above integral becomes

$$\int_X \int_{y-X} \frac{1}{|z|^{d-2\beta}} |R_\varepsilon(z)f(y)f(y-z)|dydz.$$

We can further bound the integral from above by enlarging the domain $y-X$ to some finite ball $B(2\rho)$ where $\rho = \sup_{x\in X}|x|$, because the translated region $y-X$ is included in this ball for every y. After this replacement, integrate over y first, and we have:

$$\mathbb{E}\|\mathcal{G}q_\varepsilon f\|^2 \le C\|f\|^2 \int_{B(2\rho)} \frac{|R_\varepsilon(z)|}{|z|^{d-2\beta}} dz. \tag{6.48}$$

Decompose the integration region into two parts:

$$\begin{cases} D_1 := \{|x\varepsilon^{-1}| \le T\} \cap B(2\rho), & \text{on which we have } |R_\varepsilon| \le \gamma^2; \\ D_2 := \{|x\varepsilon^{-1}| > T\} \cap B(2\rho), & \text{on which we have} |R_\varepsilon| \le C\varepsilon^\alpha |x|^{-\alpha}. \end{cases}$$

The integration on D_1 can be carried out explicitly. The restriction $|x| \le T\varepsilon$ yields that this term is of order $\varepsilon^{2\beta}$. The integration over D_2 is

$$C\int_{\varepsilon T}^{2\rho} \frac{\varepsilon^\alpha |z|^{d-1}}{|z|^{d-2\beta+\alpha}} d|z|.$$

When $2\beta = \alpha$, the integral equals $C\varepsilon^\alpha(\log(2\rho) - \log(T\varepsilon))$, and is of order $\varepsilon^\alpha|\log\varepsilon|$. When $2\beta \neq \alpha$, the integral equals $C\varepsilon^\alpha((2\rho)^{2\beta-\alpha} - (T\varepsilon)^{2\beta-\alpha})$. This estimate proves the other two cases of the lemma.

The same analysis can be done for $2\beta \geq d$. In this case, the singular term $|y-z|^{-(d-2\beta)}$ in (6.47) should be replaced by either $|\log|y-z||$ or C, which is much smoother. Consequently, $\mathbb{E}\|\mathcal{G}q_\varepsilon f\|^2$ is of order ε^α. □

Proof of Theorem 6.7. The homogenized solution satisfies $(P(x,D) + q_0)u_0 = f$. Define $\chi_\varepsilon = -\mathcal{G}q_\varepsilon u_0$, that is the solution of $(P(x,D)+q_0)\chi_\varepsilon = -q_\varepsilon u_0$. Compare these two equations with the one for u_ε, i.e. (6.2). We get

$$(P(x,D) + q_0 + q_\varepsilon)(\xi_\varepsilon - \chi_\varepsilon) = -q_\varepsilon\chi_\varepsilon,$$

where ξ_ε denotes $u_\varepsilon - u_0$. Since this equation is well-posed a.e. in Ω, we have $\xi_\varepsilon = \chi_\varepsilon - \mathcal{G}_\varepsilon q_\varepsilon \chi_\varepsilon$, which implies

$$\|\xi_\varepsilon\| \leq \|\chi_\varepsilon\| + \|\mathcal{G}_\varepsilon\|_{\mathcal{L}(L^2)}\|q\|_\infty\|\chi_\varepsilon\|. \tag{6.49}$$

Recall that the operator norm $\|\mathcal{G}_\varepsilon\|_{\mathcal{L}(L^2)}$ can be bounded uniformly in Ω; so the right hand side above is further bounded by $C\|\chi_\varepsilon\|$. Since χ_ε is of the form of $\mathcal{G}q_\varepsilon f$, we take expectation and apply the previous lemma to complete the proof. □

We decompose the corrector into the deterministic corrector $\mathbb{E}\{u_\varepsilon\} - u_0$ and the stochastic corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$. We consider their sizes and limits only in the weak sense, that is after pairing with test functions. We have the following formula for u_ε,

$$u_\varepsilon - u_0 = -\mathcal{G}q_\varepsilon u_0 + \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0 + \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon (u_\varepsilon - u_0). \tag{6.50}$$

Pairing this with an arbitrary test function $\varphi \in \mathcal{C}(\overline{X})$, we have

$$\langle u_\varepsilon - u_0, \varphi\rangle = -\langle \mathcal{G}q_\varepsilon u_0, \varphi\rangle + \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0, \varphi\rangle + \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon (u_\varepsilon - u_0), \varphi\rangle. \tag{6.51}$$

Now the deterministic corrector $\langle \mathbb{E}\{u_\varepsilon\} - u_0, \varphi\rangle$ is precisely the expectation of the expression above. In the following, we estimate the size of this corrector using the analysis developed in the proof of Lemma 6.11.

Proof of Theorem 6.8. Take expectation in (6.51). Since the first term on the right is mean zero, we have

$$\langle \mathbb{E}\{u_\varepsilon\} - u_0, \varphi\rangle = \mathbb{E}\langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0, \varphi\rangle + \mathbb{E}\langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon (u_\varepsilon - u_0), \varphi\rangle. \tag{6.52}$$

Let m denote the L^2 function $\mathcal{G}\varphi$. Rewrite the first term on the right as $\mathbb{E}\langle q_\varepsilon u_0, \mathcal{G}q_\varepsilon m\rangle$, which can be written as

$$\int_X G(x,y) R_\varepsilon(x-y) u_0(x) m(y) dx dy.$$

After controlling the Green's function by $C|x-y|^{-d+\beta}$, we have an object similar to (6.47). Following the same procedure, we can show that $|\mathbb{E}\langle q_\varepsilon u_0, \mathcal{G}q_\varepsilon m\rangle|$ can be bounded as in (6.42). To complete the proof, we only need to control the remainder term in (6.52), which can be written as $\mathbb{E}\langle q_\varepsilon (u_\varepsilon - u_0), \mathcal{G}q_\varepsilon m\rangle$. We have:

$$\mathbb{E}|\langle q_\varepsilon (u_\varepsilon - u_0), \mathcal{G}q_\varepsilon m\rangle| \le \|q_\varepsilon\|_\infty \left(\mathbb{E}\|u_\varepsilon - u_0\|^2\right)^{1/2} \left(\mathbb{E}\|\mathcal{G}q_\varepsilon m\|^2\right)^{1/2}. \tag{6.53}$$

According to Theorem 6.7 and Lemma 6.11, this term can be bounded by the right hand side of (6.45). Therefore, the remainder is smaller than the quadratic term which gives the desired estimate. □

For any fixed test function φ, the random corrector $\langle u_\varepsilon - \mathbb{E}\{u_\varepsilon\}, \varphi\rangle$ is precisely the mean-zero part of the right hand side of (6.51). We are interested in its limiting distribution. The size of its variance is given by that of $-\langle \mathcal{G} q_\varepsilon u_0, \varphi\rangle$. We calculate

$$\text{Var}\ \left(-\langle \mathcal{G} q_\varepsilon u_0, \varphi\rangle\right) = \ \text{Var}\ \left(-\langle q_\varepsilon u_0, m\rangle\right) = \int_{X^2} R_\varepsilon(x-y) u_0 m(x) u_0 m(y) dx dy.$$

Estimating this integral by decomposing the domain as in the proof of Lemma 6.11, we verify that this object is of size ε^α independent of β. Therefore, a more accurate characterization of the stochastic corrector is to find the limiting distribution of $\varepsilon^{-\alpha/2}\langle u_\varepsilon - \mathbb{E}\{u_\varepsilon\}, \varphi\rangle$. This is the task of our next step.

6.3.2 Convergence of correctors

In this section, we consider the limiting distribution of the stochastic corrector. In the analyses we are going to develop, the following estimate proves very useful. Recall that R is uniformly bounded, and there exists some T so that $|R| \le C|x|^{-\alpha}$ when $|x| > T$.

Lemma 6.12. *Recall that $R(x)$ denotes the correlation function of the random field $q(x)$ constructed in* (A1) *and* (A2)*, and that $R_\varepsilon(x)$ denotes $R(\varepsilon^{-1}x)$. Let $p \ge 1$; we have*

$$\|R_\varepsilon\|_{p,B(\rho)} \le \begin{cases} C\varepsilon^\alpha, & \alpha p < d, \\ C\varepsilon^\alpha |\log \varepsilon|^{\frac{1}{p}}, & \alpha p = d, \\ C\varepsilon^{\frac{d}{p}}, & \alpha p > d. \end{cases} \tag{6.54}$$

Here, $B(\rho)$ is the open ball centered at zero with radius ρ. The constant C depends on ρ, dimension d, and the constant in the asymptotic behavior of $R(x)$.

Proof. We break the expression for $\|R_\varepsilon\|_p^p$ into two parts as follows:

$$\int_{B(\varepsilon T)} |R_\varepsilon(x)|^p dx + \int_{B(\rho)\setminus B(\varepsilon T)} |R_\varepsilon(x)|^p dx.$$

For the first term, we bound R_ε by its uniform norm and verify this term is of order ε^d. For the second term, which we call I_2, we use the asymptotic behavior of R and have

$$I_2 \le C \int_{B(\rho)\setminus B(\varepsilon T)} \varepsilon^{\alpha p} |x|^{-\alpha p} dx \le C\varepsilon^{\alpha p} \int_{T\varepsilon}^{\rho} r^{d-1-\alpha p} dr.$$

We carry out this integral and find that it is of order $\varepsilon^{\alpha p}|\log \varepsilon|$ if $\alpha p = d$ and of order $\varepsilon^{\alpha p \wedge d}$ otherwise.

Now combine the two parts; compare the orders case by case to get the bound for $\|R_\varepsilon\|_p^p$. Then take pth roots to complete the proof. □

Lemma 6.13. *Assume $q(x)$ constructed in* (A1-A2) *satisfies* (A3)*. Let φ be an arbitrary test function in $\mathcal{C}(\overline{X})$. Then we have the following estimate of the variance of the second term in* (6.51)*:*

$$\mathrm{Var} \quad \langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0, \varphi\rangle \ll C\|u_0\|^2 \|\varphi\|_\infty^2 \varepsilon^\alpha. \tag{6.55}$$

Again, the constant C only depend on the factors as stated in Theorem 6.7.

Proof. We observe first that $m := \mathcal{G}\varphi$ is bounded since φ is uniformly bounded; a useful fact in the sequel. To simplify notation, we denote by I the variance of $\langle \mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0, \varphi\rangle$. It has the expression:

$$I = \int_{X^4} u_0(x) m(y) u_0(\xi) m(\eta) G(x,y) G(\xi,\eta) \times \\ \times \Big[\mathbb{E}\{q_\varepsilon(x)q_\varepsilon(y)q_\varepsilon(\xi)q_\varepsilon(\eta)\} - \mathbb{E}\{q_\varepsilon(x)q_\varepsilon(y)\}\mathbb{E}\{q_\varepsilon(\xi)q_\varepsilon(\eta)\}\Big] dx dy d\xi d\eta.$$

Apply Proposition 2.32 to estimate the variance of the product of q_ε above and use the bound for the Green's functions. We have

$$I \leq C \int_{X^4} |u_0(x)m(y)u_0(\xi)m(\eta)| \frac{1}{|x-y|^{d-\beta}} \frac{1}{|\xi-\eta|^{d-\beta}} \times$$
$$\times \sum_{p \neq \{(1,2),(3,4)\}} |R_\varepsilon(x_{p(1)} - x_{p(2)}) R_\varepsilon(x_{p(3)} - x_{p(4)})| dx dy d\xi d\eta.$$

Here, $p = \{(p_1, p_2), (p_3, p_4)\}$ denotes the possibilities of choosing two different pairs of indices from $\{1, 2, 3, 4\}$ in such a way that each pair contains different indices though the two pairs may share the same index. There are $C_6^2 = 15$ different choices for p; however, $p = \{(1,2),(3,4)\}$ is excluded from the sum above. Identifying (x_1, x_2, x_3, x_4) with (x, y, ξ, η), we see that there are 14 terms in the sum, and each of them is a product of two R_ε functions whose arguments are the difference vectors of points in $\{x, y, \xi, \eta\}$; more importantly, at most one of the R_ε functions shares the same argument as one of the Green's functions.

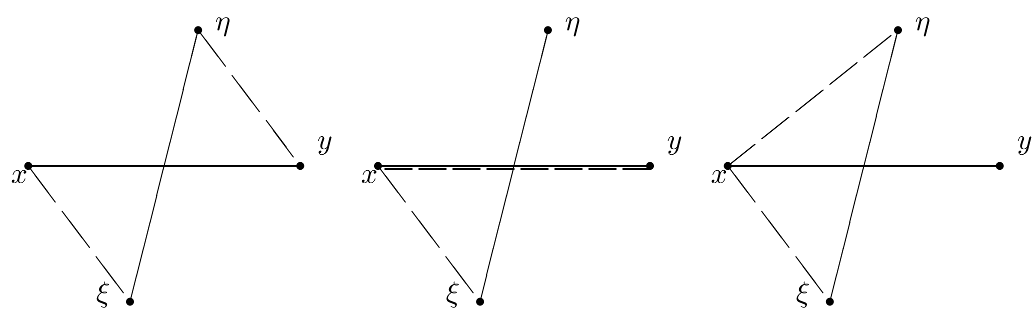


Figure 6.1: Difference vectors of four points. The *solid* lines represent arguments of the Green's functions, while the *dashed* lines represent those of the correlation functions.

We can divide the fourteen choices of p into three categories as shown in Figure 6.1. In the first category as illustrated by the first picture, the two vectors in the correlation functions are linear independent with both of the vectors in the Green's functions; in the second category, one of the Green's function shares the same argument with one of the correlation function; finally in the third category, the vector in one of the Green's function is a linear combination of the two vectors of the correlation functions.

For the first category, we consider a typical term of the form:

$$J_1 = \int_{X^4} |u_0(x)m(y)u_0(\xi)m(\eta)| \frac{1}{|x-y|^{d-\beta}} \frac{1}{|\xi-\eta|^{d-\beta}} |R_\varepsilon(x-\xi)R_\varepsilon(y-\eta)|. \tag{6.56}$$

Change variable as follows:

$$(x, x-y, x-\xi, y-\eta) \mapsto (x, y, \xi, \eta).$$

Bound m by its uniform norm. In terms of the new variables, we have

$$J_1 \le \|m\|_\infty^2 \int_X dx \int_{x-X} dy \int_{x-X} d\xi \int_{x-y-X} d\eta \; \frac{|u_0(x)u_0(x-\xi)R_\varepsilon(\xi)R_\varepsilon(\eta)|}{|y|^{d-\beta}|y-(\xi-\eta)|^{d-\beta}}.$$

We can replace the integration region of y and ξ by $B(2\rho)$, and replace that of η by $B(3\rho)$, where ρ as before denotes the maximum distance of a point in X and the origin. After doing this, we integrate over x first to get rid of the u_0 function; then integrate over y and apply Lemma 3.11 to get

$$J_1 \le \|m\|_\infty^2 \|u_0\|^2 \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|R_\varepsilon \mathbf{1}_{B(2\rho)}(\xi)| \; |R_\varepsilon \mathbf{1}_{B(3\rho)}(\eta)|}{|\xi-\eta|^{d-2\beta}} d\xi d\eta. \tag{6.57}$$

Here, $\mathbf{1}_A$ is the indicator function of a subset $A \subset \mathbb{R}^d$. We considered the case $2\beta < d$; the other cases are easier. To estimate the integral above, we apply the Hardy-Littlewood-Sobolev inequality [78, Theorem 4.3]. With $p = 2d/(d+2\beta) > 1$, we have

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|R_\varepsilon \mathbf{1}_{B(2\rho)}(\xi)| \; |R_\varepsilon \mathbf{1}_{B(3\rho)}(\eta)|}{|\xi-\eta|^{d-2\beta}} d\xi d\eta \le C(d,\beta,p) \|R_\varepsilon\|_{p,B(2\rho)} \|R_\varepsilon\|_{p,B(3\rho)}. \tag{6.58}$$

Now apply Lemma 6.54: If $\alpha p \le d$, we see J_1 is of order $\varepsilon^{2\alpha}$ or $\varepsilon^{2\alpha}|\log \varepsilon|^{2/p}$ which is much smaller than ε^α; if otherwise, J_1 is of order $\varepsilon^{2d/p} \ll \varepsilon^\alpha$ because by our choice of p we have $2d/p - \alpha = d + 2\beta - \alpha > 2\beta > 0$.

In the second category, we consider a typical term of the form:

$$J_2 = \int_{X^4} |u_0(x)m(y)u_0(\xi)m(\eta)| \frac{1}{|x-y|^{d-\beta}} \frac{1}{|\xi-\eta|^{d-\beta}} |R_\varepsilon(x-y)R_\varepsilon(x-\xi)|. \tag{6.59}$$

This time we use the following change of variables,

$$(x, x-y, x-\xi, \xi-\eta) \mapsto (x, y, \xi, \eta).$$

With this change and bounding m, we have

$$J_2 \leq \|m\|_\infty^2 \int_X dx \int_{x-X} dy \int_{x-X} d\xi \int_{x-\xi-X} d\eta \; \frac{|u_0(x)u_0(x-\xi)R_\varepsilon(\xi)R_\varepsilon(y)|}{|y|^{d-\beta}|\eta|^{d-\beta}}.$$

Enlarge the integration region of y, ξ, η as before, and then integrate over x and η. We have

$$J_2 \leq \|m\|_\infty^2 \|u_0\|^2 \int_{B^2(2\rho)} \frac{1}{|y|^{d-\beta}} |R_\varepsilon(y)||R_\varepsilon(\xi)| d\xi dy. \tag{6.60}$$

The integration over ξ yields a term of size ε^α; meanwhile, the integration over y can be estimated as in the integral in (6.48), and is of size given in (6.42). Therefore, $J_2 \ll \varepsilon^\alpha$.

For the third category, we consider a typical term of the form:

$$J_3 = \int_{X^4} |u_0(x)m(y)u_0(\xi)m(\eta)| \frac{1}{|x-y|^{d-\beta}} \frac{1}{|\xi-\eta|^{d-\beta}} |R_\varepsilon(x-\xi)R_\varepsilon(x-\eta)|. \tag{6.61}$$

Change variables according to

$$(x, x-y, x-\xi, x-\eta) \mapsto (x, y, \xi, \eta).$$

After the routine of enlarging integration domains, bounding m, and integrating the non-singular terms, we have

$$J_3 \leq \|m\|_\infty^2 \|u_0\|^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|R_\varepsilon \mathbf{1}_{B(2\rho)}(\xi)||R_\varepsilon \mathbf{1}_{B(2\rho)}(\eta)|}{|\xi - \eta|^{d-\beta}} d\xi d\eta. \tag{6.62}$$

This term can be estimated exactly as what we have done for (6.57). In particular, it is much smaller than ε^α. This completes the proof. □

To prove Theorem 6.9, we essentially consider the law of random vectors of the form $(J_1^\varepsilon(\omega), \cdots, J_N^\varepsilon(\omega))$, where

$$J_j^\varepsilon(\omega) := -\frac{1}{\varepsilon^{\alpha/2}} \int_X q_\varepsilon(y) \psi_j(y) dy, \tag{6.63}$$

for some collection of $L^2(X)$ functions $\{\psi_k(x);\ 1 \leq k \leq N\}$. We apply Lemma 2.35 which characterise their limiting joint law.

According to the interpretation in Remark 6.10, the lemma above implies that $\mathcal{G} q_\varepsilon u_0$ converges to the limit in (6.43). The other terms in the stochastic corrector $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$ are controlled by Lemmas 6.11 and 6.13. These are sufficient to prove Theorem 6.9 as follows.

Proof of Theorem 6.9. Recall the expression (6.50) for the corrector. We see its random part, i.e. $u_\varepsilon - \mathbb{E}\{u_\varepsilon\}$, can be decomposed as

$$-\mathcal{G} q_\varepsilon u_0 + (\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0 - \mathbb{E}\{\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0\}) + (\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon (u_\varepsilon - u_0) - \mathbb{E}\{\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon (u_\varepsilon - u_0)\}). \tag{6.64}$$

By (6.55), for any test function $\varphi \in \mathcal{C}(\overline{X})$, we have

$$\left\langle \frac{\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0 - \mathbb{E}\{\mathcal{G} q_\varepsilon \mathcal{G} q_\varepsilon u_0\}}{\varepsilon^{\alpha/2}}, \varphi \right\rangle \xrightarrow[\varepsilon \to 0]{\text{probability}} 0. \tag{6.65}$$

Recall estimate (6.53) and apply (6.41) and (6.45). We find that when $\alpha < 4\beta$, the size of

$\mathbb{E}|\langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon(u_\varepsilon - u_0), \varphi\rangle|$ is much smaller than $\varepsilon^{\alpha/2}$, which implies

$$\left\langle \frac{\mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon(u_\varepsilon - u_0)}{\varepsilon^{\alpha/2}}, \varphi \right\rangle \xrightarrow[\varepsilon\to 0]{\text{probability}} 0. \tag{6.66}$$

The leading term in the random corrector is therefore $\langle -\mathcal{G}q_\varepsilon u_0, \varphi\rangle$.

Consider an arbitrary set of test functions $\{\varphi_i, 1 \le i \le N\}$. By the same argument above we can verify that the vectors $(Q_1^\varepsilon, \cdots, Q_N^\varepsilon)$, where

$$Q_i^\varepsilon := \varepsilon^{-\alpha/2}\langle \varphi_i, \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0 + \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon(u_\varepsilon - u_0)\rangle,$$

converge in probability to zero vectors. On the other hand, by Lemma 2.35 and the fact that $u_0(y)\mathcal{G}\varphi(y) \in L^2(X)$, we verify that $(I_\varepsilon^i, \cdots, I_\varepsilon^N)$ converges in distribution to $(I_1, \cdots, I_N)$, where

$$I_\varepsilon^i := \varepsilon^{-\alpha/2}\langle \varphi_i, -\mathcal{G}q_\varepsilon u_0\rangle,$$

and $(I_1, \cdots, I_N)$ is the centered Gaussian with covariance matrix given by (6.44). Combining this convergence result with (6.65) and (6.66), we see that $(I_1^\varepsilon, \cdots, I_N^\varepsilon)$, where $I_i^\varepsilon := \varepsilon^{-\alpha/2}\langle u_\varepsilon - \mathbb{E}\{u_\varepsilon\}, \varphi_i\rangle$ as defined in Remark 6.10, converges in distribution to $(I_1, \cdots, I_N)$. This completes the proof. $\square$

6.3.3 General setting with long range correlations

We considered the deterministic stochastic correctors for equation (6.2), where the coefficient in the potential term is constructed as a function of a long-range correlated Gaussian random field. We found that the stochastic corrector had magnitude $\varepsilon^{\alpha/2}$ and its limiting distribution can be characterized by a Gaussian random process in some weak sense. The deterministic corrector, however, may be larger than the stochastic corrector. We find that the threshold for this to happen is $\alpha = \beta$.

In our analysis, we assumed that the Green's function $G(x, y)$ had a singularity of the type $|x-y|^{-(d-\beta)}$ near the diagonal $x = y$. Other types of singularities, such as $G(x,y) \sim \log|x-y|$, can be analyzed using similar techniques. For the logarithmic singularity, which occurs for the steady diffusion problem when $d = 2$ and the Robin boundary equation when $d = 1$, our results still hold. The deterministic corrector is then of order ε^{α} while the stochastic corrector has an amplitude of order $\varepsilon^{\alpha/2}$.

To prove the convergence in distribution of the stochastic corrector, we have assumed $\alpha < 4\beta$. This is a technical reason related to the fact that only in this case is the estimate (7.49) enough to control the remainder term in (6.51). Generalizations to $\alpha > 4\beta$ require that we estimate sufficiently high-order moments of $q(x)$. Once we have a good estimate on the sixth-order moments for instance, we can perform one more iteration in (6.51) to get

$$\begin{aligned}\langle u_\varepsilon - u_0, \varphi\rangle = -\langle \mathcal{G}q_\varepsilon u_0, \varphi\rangle + \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0, \varphi\rangle - \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0, \varphi\rangle \\ - \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon (u_\varepsilon - u_0), \varphi\rangle.\end{aligned}$$

Supposing that the sixth-order moment estimate is sufficiently accurate to control the variance of the third item on the right, and that the estimate on four-order moments is sufficient to control the remaining terms, then the same results as stated in Theorem 6.9 hold for a larger range of values of α. We do not carry out the details of such derivations here.

6.4 Convergence of the Random Correctors in Functional Spaces

6.4.1 Convergence in the space $C([0,1])$ in one dimensional space

In this section, we restrict the dimension to be one. With further assumptions that the Green's function is Lipschitz continuous and the solution to (6.2) has continuous path, we derive a stronger convergence result of $u_\varepsilon - u_0$, in probability distribution in the space of continuous paths. The proof largely resembles and depends on [11].

For simplicity, we assume that the solution to (6.2) has continuous path. This is the case for the steady diffusion problem, where solutions belong to $H_0^1(X) \subset \mathcal{C}(X)$. We also assume that the Green's function $G(x, y)$ is Lipschitz in x with Lipschitz constant uniform in y. Again, this is the case for the steady diffusion problem. However, it is not the case for the Robin boundary equation, where even in $1D$, the Green's function has a logarithmic singularity. With these assumptions, we characterize the limiting distribution of $(u_\varepsilon - u_0)/\varepsilon^{\alpha/2}$ in the space of continuous paths $\mathcal{C}(X)$, as in [28, 11]. We have the following theorem.

Theorem 6.14. *Let X be the unit interval $[0, 1]$ in $\mathbb{R}$. Assume that the Green's function $G(x, y)$ is Lipschitz continuous in x with Lipschitz constant* $\mathrm{Lip}(G)$ *uniform in y. Let u_ε be the solution to* (6.2) *and u_0 be the homogenized solution.*

(i) (Short-Range random media). *Assume that $q(x)$ is constructed as in DEF. Then*

$$\frac{u_\varepsilon - u_0}{\varepsilon^{1/2}}(x) \xrightarrow[\varepsilon \to 0]{\text{distribution}} -\sigma \int_0^1 G(x, y) u_0(y) dW(y), \tag{6.67}$$

where $W(x)$ is the standard one dimensional Brownian motion.

(ii) (Long-Range random media). *Assume $q(x)$ is constructed as in* (A1)-(A3)*. Then*

$$\frac{u_\varepsilon - u_0}{\varepsilon^{\alpha/2}}(x) \xrightarrow[\varepsilon \to 0]{\text{distribution}} -\sqrt{\frac{\kappa}{H(2H-1)}} \int_0^1 G(x, y) u_0(y) dW_H(y), \tag{6.68}$$

where W_H is the standard fractional Brownian motion with Hurst index $H = 1 - \frac{\alpha}{2}$.

Remark 6.15. We refer the reader to [96] for a review on the definitions of fractional Brownian motions and of the stochastic integral with respect to them. In particular, the random process on the right hand side of (6.68) is a mean-zero Gaussian process which, if designated as $I_H(x)$, has the following covariance function:

$$\mathrm{Cov}[I_H](x, y) = \kappa \int_0^1 \int_0^1 \frac{G(x, t) u_0(t) G(y, s) u_0(s)}{|t - s|^{2(1-H)}} dt ds. \tag{6.69}$$

Recall the decomposition in (6.50) and write

$$\frac{u_\varepsilon - u_0}{\varepsilon^{\alpha/2}}(x) = -\varepsilon^{-\alpha/2}\mathcal{G}q_\varepsilon u_0(x) + \varepsilon^{-\alpha/2}\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon u_0(x) + \varepsilon^{-\alpha/2}\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon(u_\varepsilon - u_0)(x). \quad (6.70)$$

We call the first time on the right hand side $I_\varepsilon(x)$, the second term $Q_\varepsilon(x)$, and the third one $r_\varepsilon(x)$. We verify also that the sum of the last two terms is $\varepsilon^{-\alpha/2}\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon u_\varepsilon(x)$, which we call $Q^\varepsilon(x)$.

Our plan is as follows: First, we show that $I_\varepsilon(x)$ has the limiting distribution in $\mathcal{C}(X)$ as desired in (6.68). Second, we show that $Q^\varepsilon(x)$ converges in distribution $\mathcal{C}(X)$ to the zero function. Since the zero process is deterministic, the convergence in fact holds in probability [24, p.27]; the conclusion of Theorem 6.14 follows immediately.

To show convergence of $I_\varepsilon(x)$ and $Q^\varepsilon(x)$, we apply the standard result on weak convergence of probability measures in $C([0,1])$, Proposition 2.36.

Proof of Theorem 6.14. We carry out the aforementioned two-step plan. Let us denote by $I(x)$ the Gaussian process on the right hand side of (6.68).

Convergence of $I_\varepsilon(x)$ to $I(x)$. We first show convergence of finite dimensional distributions. Fix an arbitrary natural number N, an N-tuple $(x_1, \cdots, x_N)$, we need to show that the joint law of $(I_\varepsilon(x_1), \cdots, I_\varepsilon(x_N))$ converges to that of $(I(x_1), \cdots, I(x_N))$. It suffices to show that for arbitrary N-tuple $(\xi_1, \cdots \xi_N) \in \mathbb{R}^N$, we have

$$\sum_{i=1}^N \xi_i I_\varepsilon(x_i) \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sum_{i=1}^N \xi_i I(x_i),$$

as convergence in distribution of random variables. Recalling the exact form of I_ε and I, our goal is to show, with $\sigma_H := \sqrt{\kappa/(H(2H-1))}$, that

$$\frac{1}{\varepsilon^{\alpha/2}}\int_X \sum_{i=1}^N \xi_i G(x_i, y) q_\varepsilon(y)dy \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma_H \int_X \sum_{i=1}^N \xi_i G(x_i, y) u_0(y) dW_H(y). \quad (6.71)$$

Set $F_x(y) = \sum_{i=1}^N \xi_i G(x_i, y) u_0(y)$. We verify that $F_x \in L^1 \cap L^\infty(\mathbb{R})$ and apply the following convergence result:

$$\frac{1}{\varepsilon^{\alpha/2}} \int_X F(y) q_\varepsilon(y) dy \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma_H \int_X F(y) dW_H(y), \quad \text{for } F \in L^1 \cap L^\infty, \tag{6.72}$$

which is Theorem 3.1 of [11]. This proves the convergence of finite dimensional distributions.

To show tightness of $I_\varepsilon(x)$, we calculate $\mathbb{E}|I_\varepsilon(x) - I_\varepsilon(y)|^2$ which we denote by J_1. Calculation shows:

$$\begin{aligned} J_1 &= \frac{1}{\varepsilon^\alpha} \mathbb{E} \left(\int_X [G(x,z) - G(y,z)] q_\varepsilon(z) u_0(z) dz \right)^2 \\ &= \frac{1}{\varepsilon^\alpha} \int_{X^2} [G(x,z) - G(y,z)][G(x,\xi) - G(y,\xi)] R_\varepsilon(z - \xi) u_0(z) u_0(\xi) dz d\xi. \end{aligned}$$

Use the assumption on the Lipschitz continuity of G to obtain

$$J_1 \leq (\text{Lip}G)^2 |x-y|^2 \frac{1}{\varepsilon^\alpha} \int_X |R_\varepsilon(z-\xi) u_0(z) u_0(\xi)| dz d\xi \leq C|x-y|^2. \tag{6.73}$$

We used the fact that the integral above has size ε^α, which can be easily proved as before. This shows tightness and complete the first step.

Convergence of $Q^\varepsilon(x)$ to zero function. For convergence of the finite distributions, we show that $\sum_{i=1}^N \xi_i Q^\varepsilon(x_i)$ converges to zero in $L^2(\Omega, \mathbb{P})$, which is stronger. Since we can group $\sum_{i=1}^N \xi_i G(x_i, y)$ together as in (6.71), it suffices to show $\sup_{x \in X} \mathbb{E}|Q^\varepsilon(x)| \to 0$.

We prove this by showing $\sup_{x\in X} \mathbb{E}|Q_\varepsilon(x)|^2 \to 0$ and $\sup_{x \in X} \mathbb{E}|r_\varepsilon(x)| \to 0$. The first term, i.e., $\mathbb{E}|Q_\varepsilon(x)|^2$, has the following expression,

$$\varepsilon^{-\alpha} \int_{X^4} G(x,y)G(y,z)G(x,\xi)G(\xi,\eta) u_0(z) u_0(\eta) \mathbb{E}\{q_\varepsilon(y) q_\varepsilon(z) q_\varepsilon(\xi) q_\varepsilon(\eta)\} d\xi d\eta dz dy. \tag{6.74}$$

Bound the Green's functions and u_0 by their uniform norms. Then apply Proposition 2.32

to get

$$\mathbb{E}|Q_\varepsilon(x)|^2 \le C\varepsilon^{-\alpha}\|G\|_\infty^4\|u_0\|_\infty^2 \int_{X^4}\sum_p |R_\varepsilon(x_{p(1)}-x_{p(2)})R_\varepsilon(x_{p(3)}-x_{p(4)})|. \tag{6.75}$$

This time p runs over all 15 possible ways to choose two pairs from $\{1,2,3,4\}$. Since R_ε is bounded by $C\varepsilon^\alpha|x|^{-\alpha}$, we verify each item in the sum has a contribution of size $\varepsilon^{2\alpha}$ and so does the sum. Consequently, $\mathbb{E}|Q_\varepsilon(x)|^2 \le C\varepsilon^\alpha$ and converges to zero uniformly in x.

For $r_\varepsilon(x)$, we use Cauchy-Schwarz to get

$$|r_\varepsilon(x)| \le \varepsilon^{-\frac{\alpha}{2}}\left(\int_X |q_\varepsilon(z)(u_\varepsilon - u_0)(z)|^2 dz\right)^{\frac12}\left(\int_X\left(\int_X G(x,y)q_\varepsilon(y)G(y,z)dy\right)^2 dz\right)^{\frac12}.$$

Bound q_ε in the first integral by its uniform norm. Take expectation afterwards. We verify that $\mathbb{E}|r_\varepsilon(x)|$ is bounded by

$$C\varepsilon^{-\frac{\alpha}{2}}\left(\mathbb{E}\|u_\varepsilon - u_0\|^2\right)^{\frac12}\left(\mathbb{E}\int_{X^3} G(x,y)G(y,z)G(x,\xi)G(\xi,z)q_\varepsilon(y)q_\varepsilon(\xi)dyd\xi dz\right)^{\frac12}.$$

The integral above can be estimated as before and is of size ε^α. Expectation of $\|u_\varepsilon - u_0\|^2$ is also of size α as shown before. As a result, $\mathbb{E}|r_\varepsilon(x)| \le C\varepsilon^\alpha$ and converges to zero uniformly with respect to x.

It suffices now to prove tightness of $Q^\varepsilon(x)$. To this end, we calculate $\mathbb{E}|Q^\varepsilon(x) - Q^\varepsilon(y)|^2$ which we denote by J_2.

$$J_2 = \mathbb{E}\left(\varepsilon^{-\frac{\alpha}{2}}\int_{X^2}[G(x,z)-G(y,z)]q_\varepsilon(z)G(z,\xi)q_\varepsilon(\xi)u_\varepsilon(\xi)d\xi dz\right)^2.$$

Use Cauchy-Schwarz and the uniform bound on q_ε; we get

$$J_2 \le \varepsilon^{-\alpha}\mathbb{E}\left\{(\|q\|_\infty\|u_\varepsilon\|)^2\int_X\left(\int_X[G(x,z)-G(y,z)]q_\varepsilon(z)G(z,\xi)dz\right)^2 d\xi\right\}.$$

The term $\|u_\varepsilon\|$ can be bounded uniformly with respect to ω because the operator norm of $\mathcal{G}_\varepsilon$ is. Therefore, we have

$$J_2 \leq C\mathbb{E}\int_{X^3}[G(x,z)-G(y,z)][G(x,\eta)-G(y,\eta)]q_\varepsilon(z)q_\varepsilon(\eta)G(z,\xi)G(\eta,\xi)dzd\eta d\xi.$$

Use the Lipschitz continuity and the uniform bound of G to get

$$J_2 \leq C\varepsilon^{-\alpha}\int_{X^3}(\mathrm{Lip}G)^2|x-y|^2R_\varepsilon(z-\eta)\|G\|_\infty^2 dzd\eta d\xi \leq C|x-y|^2. \tag{6.76}$$

The second inequality holds because the integral is of size ε^α as we have seen many times. This completes the proof of Q^ε converging to zero functions. Recall the argument above Proposition 2.36 to complete the proof of the theorem. □

Remark 6.16. We assume that the random field $q(x)$ satisfies (A3) to take advantage of Proposition 2.32. However, this assumption is not necessary for Theorem 6.14 to hold. Indeed, with (A1) and (A2), we can derive the asymptotic behavior of the fourth order moment $\mathbb{E}\{q(x_1)q(x_2)q(x_3)q(x_4)\}$ when the four points are mutually far away from each other. We can use this fact to estimate (6.74) instead. The argument involves routine decomposition of integration domains, which is tedious so we omit it here.

6.4.2 Convergence in distribution in Hilbert spaces

In higher dimensional spaces, for the prototypes where $P(x,D)$ is the Laplacian or fractional Laplacian, we can show that the limit in Theorem 6.9 actually holds in distribution in appropriate Hilbert spaces. More precisely, we consider the pseudo-differential equation:

$$\left[(-\Delta)_{\mathrm{D}}^{\frac{\beta}{2}} + q_0 + q_\varepsilon(x)\right]u_\varepsilon(x) = f(x). \tag{6.77}$$

Here the exponent $\beta \in (0,2]$. The subscription D denotes "Dirichlet boundary" on X. When $\beta = 2$, the boundary condition is in the usual sense, but when β is less than two and hence the equation is pseudo-differential, the boundary condition is $u_\varepsilon = 0$ on X^c, the whole complement of X. This is necessary because the fractional Laplacian is non-local.

It turns out that the above equation admits a set of pairs $(\lambda_n^\beta, \phi_n^\beta)$, $1 \le n \le \infty$, where λ_n^β is an eigenvalue and ϕ_n^β is the corresponding eigenfunction. That is,

$$(-\Delta)_{\mathrm{D}}^{\frac{\beta}{2}} \phi_n^\beta = \lambda_n^\beta \phi_n^\beta. \tag{6.78}$$

Without loss of generality we can assume that $\{\phi_n^\beta\}$ is orthonormal in $L^2(X)$. We can then define a system of Hilbert spaces as follows, with $\mathscr{D}'$ denoting the space of Schwartz distributions,

$$\mathcal{H}_\beta^s := \left\{ f \in \mathscr{D}' \ : \ \sum_{n=1}^\infty \Big(\langle f, \phi_n^\beta \rangle (\lambda_n^\beta)^s \Big)^2 < \infty \right\}, \quad s \in \mathbb{R}. \tag{6.79}$$

The inner product and norm on $\mathcal{H}_\beta^s$ is implied in the definition. We observe from the definition that $\mathcal{H}_\beta^{-s}$ is the dual space of $\mathcal{H}_\beta^s$. Moreover, when s is an integer, $\mathcal{H}_\beta^s$ consists of distributions f such that $((-\Delta)_{\mathrm{D}}^{\beta/2})^s f$ is in $L^2(X)$.

We can view the corrector $u_\varepsilon - u_0$ as $\mathcal{H}_\beta^s$-valued random variables for certain s. With the natural metric on $\mathcal{H}_\beta^s$, we can consider the weak convergence of the probability measures on $\mathcal{H}_\beta^s$ (equipped with its Borel σ-algebra) induced by the random variables $\{u_\varepsilon - \mathbb{E}u_\varepsilon\}_{\varepsilon \in (0,1)}$, as ε goes to zero, and in the sense of [24]. That is, the laws of these random variables converges to the law of the limiting process.

Theorem 6.17. *Let u_ε be the solution of the pseudo-differential equation* (6.77) *with Laplacian exponent $\beta \in (0,2]$, and let u_0 be the homogenized solution. Suppose that q_0 and f are smooth enough so that u_0 is continuous on $\overline{X}$. Suppose also the random coefficient $q(x,\omega)$ satisfies the conditions in Theorem 6.9; in particular, assume the decorrelation rate α is* *less than* 4β. *Set* $\mu = [d/2\beta]$, *the integer part of* $d/2\beta$. *Then we have that* (6.43) *holds in distribution in the space* $\mathcal{H}_\beta^{-\mu}$.

This theorem states that the limit in Theorem 6.9 holds in a stronger sense. Namely, viewed as $\mathcal{H}_\beta^{-\mu}$-valued processes, $\{u_\varepsilon - \mathbb{E}u_\varepsilon\}_{\varepsilon\in(0,1)}$ converges in distribution to the right hand side of (6.43). In some cases, $\mathcal{H}_\beta^{-\mu}$ can be chosen as $L^2(X)$.

Let $\mathcal{H}$ denotes a separable Hilbert space with an orthonormal basis $\{\phi_n\}_{n=1}^\infty$. To prove convergence in law of $\mathcal{H}$-valued process $\{Y_\varepsilon\}_{\varepsilon\in(0,1)}$ to a $\mathcal{H}$-valued random variable Y_0, we need to show that any finite dimensional distribution of Y_ε converges to that of Y_0 and that the family of laws of $\{Y_\varepsilon\}_{\varepsilon\in(0,1)}$ is tight. The first condition boils down to

$$(\langle Y_\varepsilon, \phi_{i_1}\rangle, \cdots, \langle Y_\varepsilon, \phi_{i_k}\rangle) \xrightarrow[\varepsilon\to 0]{\text{distribution}} (\langle Y_0, \phi_{i_1}\rangle, \cdots, \langle Y_0, \phi_{i_k}\rangle), \tag{6.80}$$

as $\mathbb{R}$-valued random variables, for any $k \in \mathbb{N}$, and any k-tuple $(i_1, \cdots, i_k)$. The technicality lies in the tightness of the family $\{Y_\varepsilon\}_{\varepsilon\in(0,1)}$. A sufficient condition is Proposition 2.38 which we apply in the following proof.

Proof of Theorem 6.17. *The Laplacian case.* We first consider the case $P(x, D) = -\Delta$, and hence $\beta = 2$. For simplicity, let us denote the eigenvalues and corresponding eigenfunctions of $(-\Delta)_{\mathrm{D}}$ by $(\nu_n, \phi_n)_{n=1}^\infty$; let us also simplify the notation $\mathcal{H}_2^s$ by $\mathcal{H}^s$.

We denote by $\{Y_\varepsilon(x)\}$ the $\mathcal{H}^{-\mu}$-valued sequence $\varepsilon^{-\alpha/2}(u_\varepsilon - \mathbb{E}u_\varepsilon)$ and by $I(x)$ the process in (6.43). According to the remark preceding this proof, Theorem 6.9 proves convergence of finite-dimensional distributions of Y_ε to those of I. It remain to show that $\{Y_\varepsilon\}$ is a tight sequence in $\mathcal{H}^{-\mu}$. To this end, we apply the proposition above. We first decompose Y_ε into three parts: $Y_{1\varepsilon} := -\varepsilon^{-\alpha/2}\mathcal{G}q_\varepsilon u_0$ and

$$Y_{2\varepsilon} := \frac{\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon u_0 - \mathbb{E}\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon u_0}{\varepsilon^{\frac{\alpha}{2}}}, \quad Y_{3\varepsilon} := \frac{\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon(u_\varepsilon - u_0) - \mathbb{E}\mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon(u_\varepsilon - u_0)}{\varepsilon^{\frac{\alpha}{2}}}.$$

Both criteria in the proposition concerns $\mathcal{H}^{-\mu}$ norms, so we express those of $Y_{i\varepsilon}$ explicitly, using the orthonormal basis given by $\{\nu_n^\mu \phi_n\}_{n=1}^\infty$. We have

$$\|Y_{1\varepsilon}\|^2_{\mathcal{H}^{-\mu}} = \sum_{n=1}^\infty \langle Y_{1\varepsilon}, \nu^\mu \phi_n\rangle^2_{\mathcal{H}^{-\mu}} = \sum_{n=1}^\infty \frac{1}{\nu^{2\mu}} \langle Y_{1\varepsilon}, \phi_n\rangle^2. \tag{6.81}$$

Recall the definition of χ_ε; we have that $Y_{1\varepsilon} = \varepsilon^{-\alpha/2}\chi_\varepsilon$. Since χ_ε satisfies

$$-\Delta\chi_\varepsilon + q_0\chi_\varepsilon = -q_\varepsilon u_0,$$

we have

$$\langle Y_{1\varepsilon}, \phi_n\rangle = \left\langle \frac{(-\Delta)_{\mathrm{D}}^{-1}(-q_\varepsilon u_0 - q_0\chi_\varepsilon)}{\varepsilon^{\frac{\alpha}{2}}}, \phi_n \right\rangle = \frac{1}{\nu_n} \left\langle \frac{-q_\varepsilon u_0 - q_0\chi_\varepsilon}{\varepsilon^{\frac{\alpha}{2}}}, \phi_n \right\rangle.$$

Now write

$$\left\langle \frac{-q_\varepsilon u_0 - q_0\chi_\varepsilon}{\varepsilon^{\frac{\alpha}{2}}}, \phi_n \right\rangle = -\frac{1}{\varepsilon^{\frac{\alpha}{2}}} \int_X q\left(\frac{x}{\varepsilon}\right)[u_0(x)\phi_n(x) - u_0(x)m(x)]dx,$$

with $m(x) = \mathcal{G}(q_0\phi_n)(x)$. It follows then that the mean square of this item can be bounded by $\|u_0\|_{L^\infty}, \|q_0\|_{L^\infty}$, with uniform bound in ε and n. That is,

$$\mathbb{E}\langle Y_{1\varepsilon}, \phi_n\rangle^2 \le C/\nu_n^2,$$

with some constant C uniform in ε and n. This shows that

$$\sup_{\varepsilon\in(0,1)} \|Y_{1\varepsilon}\|^2_{\mathcal{H}^{-\mu}} \le \sum_{n=1}^\infty \frac{C}{\nu_n^{2(\mu+1)}} \le C.$$

Here we used the fact that $\nu_n \le Cn^{2/d}$ for some C only depend on the volume of the domain X; see the Li-Yau estimate [77] for $\{\nu_n\}_{n\in\mathbb{N}}$. The series above converges because asymptotically the elements in the series are $1/n^{4(\mu+1)/d}$ and μ is chosen so that $4(\mu+1)/d >$

1. This proves (2.50) for $Y_{1\varepsilon}$. Since $Y_{1\varepsilon} - P_N Y_{1\varepsilon}$ precisely consists of the coordinates with indices larger than N, the second criterion follows from the same lines above.

Now for $Y_{2\varepsilon}$ and $Y_{3\varepsilon}$, we repeat the above proof for $Y_{1\varepsilon}$. The only modification is:

$$\mathbb{E}\langle Y_{2\varepsilon}, \phi_n\rangle^2 = \varepsilon^{-\alpha} \text{ Var } \langle \mathcal{G}q_\varepsilon \mathcal{G}q_\varepsilon u_0, \varphi\rangle = \nu_n^{-2}\varepsilon^{-\alpha} \text{ Var } \langle q_\varepsilon \mathcal{G}q_\varepsilon u_0, \phi_n - m\rangle,$$

again with $m = \mathcal{G}q_0\phi_n$. The last equality can be shown by introducing $\chi_{2\varepsilon} = \mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon u_0$ and following the trick we did with χ_ε above. Now in Lemma 6.13, let u_0 play the role of φ of the lemma, and bound the L^2 norm of $\phi_n - m$ by some uniform constant. This implies $\sup_{\varepsilon\in(0,1)} \mathbb{E}\langle Y_{2\varepsilon}, \phi_n\rangle^2 \le C/\nu_n^2$. Then the criteria (2.50)-(2.51) follows for $Y_{2\varepsilon}$.

For $Y_{3\varepsilon}$, we can introduce $\chi_{3\varepsilon} = \mathcal{G}q_\varepsilon\mathcal{G}q_\varepsilon(u_\varepsilon - u_0)$ and argue as above, and use estimate (6.53), again with the roles of u_0 and $\phi_n - m$ exchanged. Since $\alpha < 4\beta$, this estimate is enough to prove the criteria for $Y_{3\varepsilon}$.

Combining the above arguments, we finally proved that $\{Y_\varepsilon\}_{\varepsilon\in(0,1)}$ is tight in $\mathcal{H}^{-\mu}$. Therefore, we proved the theorem for the case of $P(x,D)$ being the Laplacian.

The fractional Laplacian case. We use the fact that λ_n^β, the eigenvalue of $(-\Delta)_{\mathrm{D}}^{\beta/2}$, is comparable to a fractional power of ν_n, the eigenvalue of $(-\Delta)_{\mathrm{D}}$:

$$C^{-1}\nu_n^{\frac{\beta}{2}} \le \lambda_n^\beta \le C\nu_n^{\frac{\beta}{2}}, \tag{6.82}$$

for some constant C [36]. Combining the above with the Li-Yau estimate, we see that $\lambda_n^\beta \sim n^{\beta/d}$. Then the same procedure above works. This completes the proof. □

6.5 Notes

Section 6.2 In the chemical physics literature, the authors of [22, 23] have investigated a similar diffusion process of particles through a heterogeneous surface which reflects particles except on some periodically or randomly located patches that absorb particles. Hence, in their setting, the boundary condition in (6.7) is $-\partial_\nu u_\varepsilon = \kappa_{\text{disc}} u_\varepsilon$ on the patches, and is $-\partial_\nu u_\varepsilon = 0$ otherwise. Here ∂_ν denotes the partial derivative in the outer normal direction ν and κ_{disc} is the absorption rate on the patches.

This boundary condition is similar with ours except for the geometric configuration of the discs. Analyzing the data obtained from Brownian dynamics simulations, they find that as long as the diffusion away from the boundary is concerned, the heterogeneous boundary conditions above can be replaced by an effective homogeneous boundary which partially absorbs particles in a uniform rate over the entire surface, i.e., by $-\partial_\nu u_\varepsilon = \kappa u_\varepsilon$ where κ is the uniform absorption rate.

The authors of [22, 23] also proposed an expression of κ from data analysis. However, this homogenization procedure is intuitive and empirical. Other boundary conditions have also been investigated in e.g., [87]. The authors of that paper considered a reaction-diffusion equation and the boundary condition is $u_\varepsilon = v$ on small-scale patches and is $-\partial_\nu u_\varepsilon = \varepsilon^{-1} g$, where v and g are known functions. Their homogenization results are obtained by formally studying a boundary layer and matching the boundary layer solution with the solution in the interior of the domain. Again, a rigorous mathematical proof in this case is too complicated.

Section 6.4 The Hilbert spaces in which we proved convergence of the corrector is somewhat strange. In particular, for the Laplacian equation, it implies convergence in distribution in $L^2(X)$ of the corrector only for dimension $d \geq 3$. This dimension restriction on controlling the L^2 norm of the random corrector is somewhat optimal, as already observed by Bal [9] and by Bardos, Garnier and Papanicolaou [19].

The Li-Yau estimate of the eigenvalues of the Dirichlet Laplacian is the key in our proof of the convergence in distribution in Hilbert spaces. The constant in the estimate, equation (5.2), is very precise. We do not really need this precise constant; all that matters is the asymptotic relation $\nu_n \sim O(n^{-\frac{2}{d}})$. Therefore, the Weyl's law on the counting of eigenvalues of the Dirichlet Laplacian, as discussed in Hörmander [64, Theorem 17.5.3], is enough. Nevertheless, the approach of Li and Yau to the precise constant is surprisingly elementary but useful, and we recommend their paper [77].

Chapter 7

Corrector Theory for Multiscale Numerical Algorithms

Despite of the fact that PDEs with highly oscillating random coefficients can be well approximated by homogenization, finding the homogenized coefficients may be a daunting computational task and the assumptions necessary to the applicability of homogenization theory may not be met. Several numerical methodologies have been developed to find accurate approximations of the solution without solving all the details of the micro-structure [4, 50, 53, 54, 3]. Examples include the multi-scale finite element method (MsFEM) and the finite element heterogeneous multi-scale method (HMM). Such schemes are shown to perform well in the homogenization regime, in the sense that they approximate the solution to the homogenized equation without explicitly calculating any macroscopic, effective medium, coefficient. Homogenization theory thus serves as a benchmark that ensures that the multi-scale scheme performs well in controlled environments, with the hope that it will still perform well in non-controlled environments, for instance when ergodicity and stationarity assumptions are not valid.

In many applications, as seen in Chapter 1, estimating the random fluctuations (finding the random corrector) in the solution is as important as finding its homogenized limit. In this chapter we aim to present another benchmark for such multi-scale numerical schemes that addresses the limiting stochasticity of the solutions. We calculate the limiting (probability) distribution of the random corrector given by the multi-scale algorithm when the correlation length of the medium tends to 0 at a fixed value of the discretization size h. We then compare this distribution to the distribution of the corrector of the continuous equation. When these distributions are close, in the sense that the $h-$dependent distribution converges to the continuous distribution as $h \to 0$, we deduce that the multi-scale algorithm asymptotically correctly captures the randomness in the solution and passes the random corrector test.

7.1 Set-up of the Corrector Test

a. The test equation

The above proposal requires a controlled environment in which the theory of correctors is available. There are very few equations for which this is the case [9, 11, 59, 12, 13]. Here, we initiate such an analysis in the simple case of the one-dimensional second-order elliptic equation (1.3) in Chapter 1. That is,

$$-\frac{d}{dx} a\left(\frac{x}{\varepsilon}, \omega\right) \frac{d}{dx} u_\varepsilon(x, \omega) = f(x), \quad x \in (0,1), \tag{7.1}$$

with zero Dirichlet boundary condition. Under the mild conditions that $a(x, \omega)$ is stationary, ergodic and uniformly elliptic in the sense of $\lambda \le a(x) \le \Lambda$ for any x for some positive real numbers $\lambda < \Lambda$, u_ε converges weakly in $H_0^1(0,1)$ to the solution of the homogenized equation

$$-\frac{d}{dx} a^* \frac{d}{dx} u_0(x) = f(x), \quad x \in (0,1), \tag{7.2}$$

with zero Dirichlet boundary condition. The effective coefficient a^* is the harmonic mean of $a(x,\omega)$ which, together with the deviation between the two, is defined by

$$\frac{1}{a^*} := \mathbb{E}\left\{\frac{1}{a(0,\omega)}\right\}, \quad q(x,\omega) = \frac{1}{a(x,\omega)} - \frac{1}{a^*}. \tag{7.3}$$

b. Two different random fields

As we have seen in Chapter 1, the limiting distribution of the corrector $u_\varepsilon - u_0$ depends very much on the decorrelation rate of $q(x,\omega)$. We consider the following two sets of assumptions on it.

I. The first set will be referred as the case of short-range correlations:

(S1) The random process $q(x)$ is stationary ergodic and mean-zero; the coefficients a^* and $a(x,\omega)$ are uniformly elliptic.

(S2) The correlation function $R(x)$ of the random process $q(x)$ is integrable in $\mathbb{R}$. In particular, the constant

$$\sigma^2 := \int_{\mathbb{R}} R(x)dx, \tag{7.4}$$

is finite; see the remark after 2.5.

(S3) For some small $\delta > 0$, the mixing coefficient $\rho(r)$ of $q(x)$ is of order $O(r^{-d-\delta})$ for large r.

The above assumptions are quite general. In particular, (S3) implies ergodicity of $q(x)$, and (S1) plus ergodicity is the standard assumption for homogenization theory; (S3) is the standard assumption to obtain central limit theorem of oscillatory integrals with integrand $q_\varepsilon(x)$; see Theorem 2.15.

II. The second set of assumptions will be referred to as function of Gaussian random field, which is an example of random field that has long-range correlations.

(L1) The process $q(x)$ is constructed as a function of Gaussian random field as in Definition 2.27. That is,

$$q(x,\omega) = \Phi(g(x,\omega)), \tag{7.5}$$

where g_x is a stationary Gaussian random process with mean zero and variance one. Further, assume that the correlation function R_g of g_x has the following asymptotic behavior:

$$R_g(\tau) \sim \kappa_g \tau^{-\alpha}, \tag{7.6}$$

where $\kappa_g > 0$ is a constant and $\alpha \in (0,1)$.

(L2) The function $\Phi(x)$ satisfies $|\Phi| \le q_0$ for some constant q_0, so that the process $a(x,\varepsilon)$, constructed by the relation (7.3) for some positive constant a^*, satisfies uniform ellipticity with constants (λ, Λ).

The process $q(x)$ is stationary and mean-zero. More importantly, its correlation function $R(x)$ has a similar asymptotic behavior to that in (7.6) with κ_g replaced by κ, where

$$\kappa := \frac{\kappa_g}{\sqrt{2\pi}} \int_{\mathbb{R}} s\Phi(s) e^{-\frac{s^2}{2}} ds. \tag{7.7}$$

Therefore, $R(x)$ is no longer integrable and $q(x)$ has long range correlation; cf. Lemma 2.28. We note that when $\alpha > 1$, the process constructed above has short range correlation and provides an example satisfying (S2).

c. The corrector test

The corrector theory in the random homogenization of (7.1) can be summarized by,

$$\frac{u_\varepsilon - u_0}{\varepsilon^{\frac{\alpha\wedge 1}{2}}}(x) \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}_{\alpha\wedge 1}(x; W^{\alpha\wedge 1}).$$

Here and below, we use $c \wedge d$ to denote $\min\{c,d\}$. In this concise formula, $\alpha < 1$ means

Figure 7.1: A commutative diagram for the corrector test

$$
\begin{array}{ccc}
\dfrac{u_\varepsilon^h - u_0^h}{\varepsilon^{\frac{\alpha\wedge 1}{2}}}(x,\omega) & \xrightarrow[(i)]{h\to 0} & \dfrac{u_\varepsilon - u_0}{\varepsilon^{\frac{\alpha\wedge 1}{2}}}(x,\omega) \\
\varepsilon\to 0 \Big\downarrow (ii) & & (iii) \Big\downarrow \varepsilon\to 0 \\
\mathcal{U}_{\alpha\wedge 1}^h(x; W^{\alpha\wedge 1}) & \xrightarrow[(iv)]{h\to 0} & \mathcal{U}_{\alpha\wedge 1}(x; W^{\alpha\wedge 1}).
\end{array}
$$

assumptions (L1) and (L2) are used for $q(x,\omega)$, while $\alpha > 1$ means the other set of assumptions are used. Above, $W^1 = W$ is the standard Brownian motion whereas $W^\alpha = W^H$ is the fractional Brownian motion with Hurst index $H = 1 - \frac{\alpha}{2}$. The limit $\mathcal{U}_{\alpha\wedge 1}$ denotes the corresponding limit in these two situations; see Chapter 1.

Now let us apply some multiscale numerical algorithm to solve (7.1). Let h denote the discretization size and $u_\varepsilon^h(x)$ the solution of the scheme. Let $u_0^h(x)$ denote the standard finite element solution of the homogenized equation (7.2). We characterize the limiting distribution of $u_\varepsilon^h - u_0^h$ as a random process after proper rescaling by $\varepsilon^{\frac{\alpha\wedge 1}{2}}$. We say that a numerical procedure is consistent with the corrector theory and that is passes the corrector test when the diagram in Fig. 7.1 commutes.

The four convergence paths should be understood in the sense of distribution in $\mathcal{C}([0,1])$. The convergence path (iii) is the corrector theories stated in the previous concise result. In (i), h is sent to zero while ε is fixed. To check (i) is a numerical analysis question without multi-scale issues since the ε-scale details are resolved. Convergence in (i) can be obtained path-wise and not only in distribution (path (iv) may also be considered path-wise). The main new mathematical difficulties therefore lie in analyzing the paths (ii) and (iv).

7.2 Some Multiscale Numerical Algorithms

We briefly introduce the common ideas of finite element based multiscale methods. Assume $f \in L^2 \subset H^{-1}$. The weak solution to (7.1) is the unique function $u_\varepsilon \in H_0^1(0,1)$ such that

$$A_\varepsilon(u,v) = f(v), \quad \forall v \in H_0^1(0,1). \tag{7.8}$$

The associated bilinear and linear forms are defined as

$$A_\varepsilon(u,v) := \int_0^1 a_\varepsilon(x)\frac{du}{dx}\cdot\frac{dv}{dx}dx, \quad F(v) := \int_0^1 fv\, dx. \tag{7.9}$$

Existence and uniqueness of u_ε are guaranteed by the uniform ellipticity of $a_\varepsilon(x)$.

7.2.1 General finite element based multi-scale schemes

Almost all finite element based multi-scale schemes for (7.1) have the same main premise: in the weak formulation (7.8), we approximate H_0^1 by a finite dimensional space and if necessary, also approximate the bilinear form.

To describe the choices of the finite spaces, we choose a uniform partition of the unit interval into N sub-intervals with size $h = 1/N$. Let x_k denote the kth grid point, with $x_0 = 0$ and $x_N = 1$, and I_k the interval (x_{k-1}, x_k). To simplify notation in the general setting, we still denote by V_ε^h the finite space and by $\{\phi_\varepsilon^j\}_{j=1}^{N-1}$ the basis functions.

Example 7.1. The standard finite element (FEM) basis functions are piece-wise linear "hat functions", each of them peaking at one nodal point and vanishing at all other nodal points. Denote these hat functions by $\{\phi_0^j\}$ and denote the subspace of H_0^1 they span by V_0^h. The standard FEM approximates H_0^1 by V_0^h.

In a general scheme, the bilinear form in (7.8) might be modified. Nevertheless, to simplify notations, we still denote it as A_ε. The solution obtained from the scheme is then

$u_\varepsilon^h \in V_\varepsilon^h$ such that

$$A_\varepsilon(u_\varepsilon^h, v) = F(v), \quad \forall v \in V_\varepsilon^h. \tag{7.10}$$

Since V_ε^h is finite dimensional, the above condition amounts to a linear system, which is obtained by putting $u_\varepsilon^h = U_j^\varepsilon \phi_\varepsilon^j$, and by requiring the above equation to hold for all basis functions. The linear system is:

$$A_\varepsilon^h U^\varepsilon = F^\varepsilon. \tag{7.11}$$

Here, the vector U^ε is a vector in $\mathbb{R}^{N-1}$, and it has entries U_j^ε. The load vector F^ε is also in $\mathbb{R}^{N-1}$ and has entries $F(\phi_\varepsilon^j)$. The stiffness matrix A_ε^h is an $N-1$ by $N-1$ matrix, and its entries are $A_\varepsilon(\phi_\varepsilon^i, \phi_\varepsilon^j)$. Our main assumptions on the basis functions and the stiffness matrix are the following.

(N1) For any $1 \le j \le N-1$, the basis function ϕ_ε^j is supported on $I_j \cup I_{j+1}$, and it takes the value δ_j^i at nodal points $\{x_i\}$. Here δ_j^i is the Kronecker symbol.

(N2) The matrix A_ε^h is symmetric and tri-diagonal. In addition, we assume that there exists a vector $b_\varepsilon \in \mathbb{R}^N$ with entries $\{b_\varepsilon^k\}_{k=1}^N$, so that $A_{\varepsilon ii+1}^h = -b_\varepsilon^{i+1}$ for any $i = 1, \cdots, N-2$ and

$$A_{\varepsilon ii}^h = -(A_{\varepsilon ii-1}^h + A_{\varepsilon ii+1}^h), \quad i = 1, \cdots, N-1. \tag{7.12}$$

In other words, the ith diagonal entry of A_ε^h is the negative sum of its neighbors in each row. Here, $A_{\varepsilon 01}^h$ and $A_{\varepsilon N-1N}^h$ are not matrix elements and are set to be b_ε^1 and b_ε^N, respectively.

(N3) On each interval I_j for $j = 1, \cdots, N$, the only two basis functions that are nonzero are ϕ_ε^j and ϕ_ε^{j-1}, and they sum to one, i.e., $\phi_\varepsilon^j + \phi_\varepsilon^{j-1} = 1$. Equivalently, we have

$$\phi_\varepsilon^j(x) = \mathbf{1}_{I_j}\tilde{\phi}_\varepsilon^j(x) + \mathbf{1}_{I_{j+1}}(x)[1 - \tilde{\phi}^{j+1}(x)], \tag{7.13}$$

for some functions $\{\tilde{\phi}_\varepsilon^k(x)\}_{k=1}^N$, each of them defined only on I_j with boundary value 0 at the left end point and 1 at the right.

As we shall see for MsFEM, (N3) implies (N2) when the bilinear form is symmetric. The special tri-diagonal structure of A_ε^h in (N2) simplifies the calculation of its action on a vector. Let U be any vector in $\mathbb{R}^{N-1}$, we have

$$(A_\varepsilon^h U)_i = -D^+(b_\varepsilon^i D^- U_i), \qquad i = 1, \cdots, N-1. \tag{7.14}$$

Here, the symbol D^- denotes the backward difference operator, which is defined, together with the forward difference operator D^+, as

$$(D^- U)_k = U_k - U_{k-1}, \quad (D^+ U)_k = U_{k+1} - U_k. \tag{7.15}$$

The equality (7.14) is easy to check, and to make sense of the case when i equals 1 or N, we need to extend the definition of U by setting U_0 and U_N to zero. This formula has been used, for example, in [67]. It is a very useful tool in the forthcoming computations.

7.2.2 The multiscale finite element method

The idea of the multiscale finite element method, also known as MsFEM and developed in e.g. [66, 67, 65, 52, 33, 51], is to replace the hat basis functions in FEM by multi-scale basis functions $\{\phi_\varepsilon^j\}$. They are constructed as follows.

$$\begin{cases} \mathcal{L}_\varepsilon \phi_\varepsilon^j(x) = 0, & x \in I_1 \cup I_2 \cup \cdots \cup I_{N-1}, \\ \phi_\varepsilon^j = \phi_0^j, & x \in \{x_k\}_{k=0}^N. \end{cases} \tag{7.16}$$

Here $\mathcal{L}_\varepsilon$ is the differential operator in (7.1). Clearly, ϕ_ε^j has the same support as ϕ_0^j and thus satisfies (N1). Note that the $\{\phi_\varepsilon^j\}$ are constructed locally on independent intervals, and are suitable for parallel computing.

For any $k = 1, \cdots, N$, we observe that the only non-zero basis functions are ϕ_ε^k and ϕ_ε^{k-1}. Further, they sum up to one at the boundary points x_{k-1} and x_k. Since equation (7.16) is of linear divergence form, we conclude that $\phi_\varepsilon^k(x) + \phi_\varepsilon^{k-1}(x) \equiv 1$ on the interval. This shows that MsFEM satisfies (N3). In fact, the functions $\{\tilde{\phi}_\varepsilon^k\}_{k=1}^N$ for MsFEM are constructed by solving (7.16) on I_k with boundary values zero at x_{k-1} and one at x_k. Once they are constructed, $\{\phi_\varepsilon^j\}$ is given by (7.13). We can solve ϕ_ε^k analytically and obtain that

$$\tilde{\phi}_\varepsilon^j = b_\varepsilon^j \int_{x_{j-1}}^x a_\varepsilon^{-1}(t)dt, \qquad b_\varepsilon^j = \left(\int_{I_j} a_\varepsilon^{-1}(t)dt \right)^{-1}. \tag{7.17}$$

Consequently, (N1) and (N3) indicate that MsFEM also satisfies (N2). To calculate the entries of the stiffness matrix A_ε^h, we fix any $i = 2, \cdots, N-2$, and compute

$$(A_\varepsilon^h)_{i-1i} = -\int_{I_i} a_\varepsilon \left(\frac{d\tilde{\phi}_\varepsilon^i}{dx} \right)^2 dx = -\left(a_\varepsilon \frac{d\tilde{\phi}_\varepsilon^i}{dx} \right)^2 \int_{I_i} a_\varepsilon^{-1}(s)ds = -b_\varepsilon^i.$$

The last equality can be verified from the fact that $\tilde{\phi}_\varepsilon^i$ solves (7.16) and integration by parts. For $i = 1$ and N, we verify that (7.12) holds for b_ε^0 and b_ε^N given by (7.17).

Remark 7.2. (**Super-convergence in one dimension**) It is well know that when dimension $d = 1$, the standard finite element method is super-convergent, in the sense that it yields exact values at nodal points. In our case, $u_0^h(x_k) = u_0(x_k)$, where u_0 solves (7.2) and u_0^h is the FEM approximation. We observe that this property is preserved by MsFEM. Indeed, let Pu_ε be the projection of u_ε in V_ε^h, i.e., $Pu_\varepsilon = u_\varepsilon(x_j)\phi_\varepsilon^j(x)$. Then, using integrations by parts, (7.16), and the fact that $Pu_\varepsilon - u_\varepsilon$ vanishes at nodal points, we have

$$A_\varepsilon(Pu_\varepsilon, v) = A_\varepsilon(u_\varepsilon, v) = F(v), \quad \forall v \in V_\varepsilon^h.$$

Since the second equality is also satisfied by u_ε^h, it follows that $A_\varepsilon(Pu_\varepsilon - u_\varepsilon^h, v) = 0$ for any

v in V_ε^h. In particular, by choosing $v = Pu_\varepsilon - u_\varepsilon^h$ and by coersivity of $A_\varepsilon(\cdot,\cdot)$, we conclude that $Pu_\varepsilon = u_\varepsilon^h$. The super-convergence result follows.

Several useful results follow from this super-convergent property. First, $u_\varepsilon^h(x)$ of MsFEM coincides with the true solution $u_\varepsilon(x)$ at nodal points. Note that u_ε can be explicitly solved analytically and that $|u_\varepsilon(x) - u_\varepsilon(y)| \le C|x-y|$ for some universal C. We then have

$$|D^- U_k^\varepsilon| = |u_\varepsilon^h(x_k) - u_\varepsilon^h(x_{k-1})| \le Ch. \tag{7.18}$$

This improves the condition (7.39) in Proposition 7.11 and hence improves several subsequent estimates. Second, a fact which we have used extensively before, we have $|D^- G_0^h| \le Ch$ and for any fixed x_k, $G_0^h(x; x_k)$ defined in (7.51) equals the continuous Green's function $G_0(x, x_k)$ for (7.2). This is because the functions agree at the nodal points due to super-convergence and they are both piece-wise linear in x. □

7.2.3 The heterogeneous multiscale method

The goal of the FEM-based heterogeneous multiscale method, abbreviated as HMM and developed in [49, 50], is to approximate the large-scale properties of the solution to (7.1) without computing the homogenized coefficient first. Suppose we already know this effective coefficient, i.e., a^* in our case. Then the large-scale solution u_0 can be solved by minimizing the functional

$$I[u] := \frac{1}{2}A_0(u,u) - F(u) = \frac{1}{2}\int_0^1 a^* \left(\frac{du}{dx}\right)^2 \, dx - \int_0^1 fu \, dx.$$

In numerical methods, the first integral above can be computed by the following mid-point quadrature rule:

$$A_0(u,u) \approx \sum_{j=1}^N \left(a^*(x^j) \frac{du}{dx}(x^j) \right)^2 h.$$

Here $x^j = (x_{j-1} + x_j)/2$ is the mid-point of I_j. In HMM, a^* is unknown, and the idea is to approximate $(u'a^*u')(x^j)$ by averaging in a local patch around the point x^j. For instance, we can take

$$\left(a^*(x^j)\frac{du}{dx}(x^j)\right)^2 \approx \frac{1}{\delta}\int_{I_j^\delta}\left(a_\varepsilon(s)\frac{d(\mathscr{L}u)}{dx}(s)\right)^2 ds.$$

Here, I_j^δ denotes the interval $x^j + \frac{\delta}{2}(-1,1)$, that is, the small interval centered in I_j with length δ. The operator $\mathscr{L}$ maps a function w in V_0^h, i.e., the space spanned by hat functions, to the solution of the following equation:

$$\begin{cases} \mathcal{L}_\varepsilon(\mathscr{L}w) = 0, & x \in I_1^\delta \cup \cdots \cup I_{N-1}^\delta, \\ \mathscr{L}w = w, & x \in \{\partial I_j^\delta\}_{j=1}^{N-1}. \end{cases} \tag{7.19}$$

The idea here is the same as MsFEM, namely to encode small-scale structures of the random media into the construction of the bilinear form. The key difference that distinguishes HMM and MsFEM is that the above equations are solved for HMM on patches I_k^δ that are smaller than I_k. We check that $\mathscr{L}$ is a linear operator; therefore, the following approximation of $A_0(\cdot,\cdot)$ is indeed bilinear:

$$A_\varepsilon^\delta(w,v) := \sum_{j=1}^{N}\frac{h}{\delta}\int_{I_j^\delta} a_\varepsilon\,\frac{d(\mathscr{L}w)}{dx}\frac{d(\mathscr{L}v)}{dx}\,dx. \tag{7.20}$$

With this approximation of the bilinear form, HMM consists finding

$$u_\varepsilon^{h,\delta} := \underset{w\in V_0^h}{\operatorname{argmin}}\ \frac{1}{2}A_\varepsilon^\delta(w,w) - F(w).$$

This variational problem is equivalent to solving $u_\varepsilon^{h,\delta} = U_j^{\varepsilon,\delta}\phi_0^j(x)$, where $U^{\varepsilon,\delta}$ is determined by the linear system

$$A_\varepsilon^{h,\delta}U^{\varepsilon,\delta} = F^0. \tag{7.21}$$

Therefore, the above HMM can be viewed as a finite element method. The finite dimensional space here is V_0^h. Therefore HMM clearly satisfies (N1) and (N3). To check (N2), we calculate the associated stiffness matrix $A_\varepsilon^{h,\delta}$. It has entries $A_\varepsilon^\delta(\phi_0^i, \phi_0^j)$. From the defining equation (7.19), we see that $\mathscr{L}\phi_0^i$ is non-zero only on $I_i^\delta \cup I_{i+1}^\delta$, which implies that $A_\varepsilon^{h,\delta}$ is again tri-diagonal. Further, we verify that $\mathscr{L}\phi_0^i + \mathscr{L}\phi_0^{i-1} = 1$ on the interval I_i^δ, which can be obtained from integrations by parts and which implies that $A_\varepsilon^{h,\delta}$ satisfies (7.12). Therefore, HMM satisfies (N2).

In fact, we can calculate the b_ε vectors. Let us consider the $(i-1,i)$th entry of $A_\varepsilon^{h,\delta}$, where i can be $2, \cdots, N-1$. Since $(\mathscr{L}\phi_0^{i-1})' = -(\mathscr{L}\phi_0^i)'$ on I_i^δ, we have

$$\left(A^{h,\delta}\right)_{\varepsilon i-1 i} = -\frac{h}{\delta}\int_{I_i^\delta} a_\varepsilon(s)\left(\frac{d(\mathscr{L}\phi^i)}{dx}\right)^2 ds.$$

Now from (7.19), we verify that $a_\varepsilon(\mathscr{L}\phi_0^i)'$ on I_i^δ is a constant given by

$$c_i^\delta = \Big(\int_{I_i^\delta} a_\varepsilon^{-1}(s)ds\Big)^{-1} \mathscr{L}\phi_0^i\Big|_{x^i-\frac{\delta}{2}}^{x^i+\frac{\delta}{2}} = \Big(\int_{I_i^\delta} a_\varepsilon^{-1}(s)ds\Big)^{-1}\frac{\delta}{h}.$$

Therefore, we have

$$\left(A^{h,\delta}\right)_{\varepsilon i-1 i} = -(c_i^\delta)^2\frac{h}{\delta}\int_{I_i^\delta} a_\varepsilon^{-1}ds = -\frac{\delta}{h}\left(\int_{I_i^\delta} a_\varepsilon^{-1}(s)ds\right)^{-1} =: -b_{\varepsilon,\delta}^i. \tag{7.22}$$

We extend the definition of $b_{\varepsilon,\delta}^i$ to the cases of $i=1$ and $i=N$, and check that the $(1,1)$th and $(N-1,N-1)$th entries of the stiffness matrix also satisfy (7.12). In particular, the action of $A_\varepsilon^{h,\delta}$ on a vector satisfies the conservative form as in (7.14). In the sequel, to simplify notation, we drop the δ in the notations $A_\varepsilon^{h,\delta}$, $U^{\varepsilon,\delta}$ and $b_{\varepsilon,\delta}^i$.

The well-posedness of the optimization problem above, or equivalently of the linear system (7.21) is obtained by Lax-Milgram. We show that the bilinear form $A_\varepsilon^\delta(\cdot,\cdot)$ is continuous and coercive. Consider two arbitrary functions $w = W_i\phi_0^i$ and $v = V_j\phi_0^j$ in V_0^h.

Then:

$$A_\varepsilon^h(w,v) = W_i A_{\varepsilon ij}^h V_j = -\sum_i W_i D^+(b_\varepsilon^i D^- V_i) = \sum_i D^- W_i b_\varepsilon^i D^- V_i.$$

Estimating the entries of vector b_ε by its infinity norm and using Cauchy-Schwarz, we obtain

$$|A_\varepsilon^h(w,v)| \le \left(\sup_{1\le i\le N} b_\varepsilon^i \right) \|D^- W\|_{\ell^2} \|D^- V\|_{\ell^2} \le \Lambda |w|_{H^1} |v|_{H^1}.$$

In the last inequality above, we used the fact that $\lambda h^{-1} \le b_\varepsilon^i \le \Lambda h^{-1}$, which can be seen from its definition in (7.22) and the uniform ellipticity of a_ε, and that $\|D^- W\|_{\ell^2} = |w|_{H^1}\sqrt{h}$ for $w \in V_0^h$. This proves continuity. Taking $w = v$, we have

$$A_\varepsilon^h(w,w) \ge \left(\inf_{1\le i\le N} b_\varepsilon^i \right) \|D^- W\|_{\ell^2} \|D^- W\|_{\ell^2} \ge \lambda |w|_{H^1}^2.$$

This proves coercivity. Therefore, by the Lax-Milgram theorem for the finite element space [97, p.137], there exists a unique $u_\varepsilon^{h,\delta} \in V_0^h$ that solves the optimization problem. Further, we have

$$|u_\varepsilon^{h,\delta}|_{H^1} \le \frac{1}{\lambda} \sup_{w\in V_0^h} \frac{F(w)}{|w|_{H^1}} \le \frac{1}{\pi\lambda}\|f\|_2. \tag{7.23}$$

An immediate consequence is that $|D^- U^\varepsilon| \le C\sqrt{h}$ by the argument in Remark 7.10.

7.3 Main Results on the Corrector Test

In this section, we state our main convergence theorems in the setting of MsFEM and HMM although they hold for more general schemes. The sufficient conditions on these schemes will be revealed when we prove the main theorems in the forthcoming sections.

To simplify notation, we drop the dependency in ω when this does not cause confusion. We define $a_\varepsilon(x) = a(x/\varepsilon)$. For a function g in $L^p(D)$, we denote its norm by $\|g\|_{p,D}$. When D is the unit interval we drop the symbol D. The natural space for (7.1) and (7.2) is the

Hilbert space H_0^1. By the Poincaré inequality, the semi-norm of H_0^1 defined by

$$|u|_{H^1,D} = \|\frac{du}{dx}\|_{2,D}, \tag{7.24}$$

is equivalent with the standard norm. We use the notation C to denote constants that may vary from line to line. When C depends only on the elliptic constants (λ, Λ), we refer to it as a *universal* constant. Finally, the Einstein summation convention is used throughout this paper: two repeated indices, such as in $c_i d^i$, are summed over their (natural) domain of definition.

The first theorem analyze MsFEM in the setting of short-range correlations.

Theorem 7.3. *Let u_ε and u_0 be solutions to* (7.1) *and* (7.2)*, respectively. Let u_ε^h be the solution to* (7.1) *obtained by MsFEM and let u_0^h be the standard finite element approximation of u_0. Then we have:*

(i) *Suppose that $a(x)$ satisfies* (S1) *and f is continuous on* $[0,1]$*. Then*

$$|u_\varepsilon^h - u_\varepsilon|_{H^1} \le \frac{h}{\lambda\pi}\|f\|_2, \quad \|u_\varepsilon^h - u_\varepsilon\|_2 \le \frac{h^2}{\lambda\pi^2}\|f\|_2. \tag{7.25}$$

Assume further that $q(x)$ satisfies (S2)*. Then*

$$\sup_{x\in[0,1]} \left| \mathbb{E}\left(u_\varepsilon^h(x) - u_0^h(x)\right)^2 \right| \le C\frac{\varepsilon}{h}\|R\|_{1,\mathbb{R}}(1 + \|f\|_2), \tag{7.26}$$

where C is a universal constant and R is the correlation function of q.

(ii) *Now assume that $q(x)$ satisfies* (S3)*. Then,*

$$\frac{u_\varepsilon^h(x) - u_0^h(x)}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma \int_0^1 L^h(x,t)dW_t =: \mathcal{U}^h(x;W). \tag{7.27}$$

The constant σ is defined in (7.4) *and W is the standard Wiener process. The function $L^h(x,t)$ is explicitly given in* (7.50)*. The convergence is in distribution in the space $\mathcal{C}$.*

(iii) *Now let h goes to zero, we have*

$$\mathcal{U}^h(x;W) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}(x;W) := \sigma \int_0^1 L(x,t)dW_t. \tag{7.28}$$

The Gaussian process $\mathcal{U}(x;W)$ was characterized in [28]. *The kernel $L(x,t)$ is defined as*

$$L(x,t) = \mathbf{1}_{[0,x]}(t)\left(\int_0^1 F(s)ds - F(t)\right) + x\left(F(t) - \int_0^1 F(s)ds\right)\mathbf{1}_{[0,1]}(t). \tag{7.29}$$

Here and below, $\mathbf{1}$ *is the indicator function and* $F(t) = \int_0^t f(s)ds$.

Remark 7.4. Equivalently, the theorem says the diagram in Figure 7.1 commutes when q has short range correlation and that MsFEM passes the corrector test in this setting.

To prove (iv) of the diagram, we recast $L(x,t)$ as

$$L(x,t) = a^* \frac{\partial}{\partial y} G_0(x,t) \cdot a^* \frac{\partial}{\partial x} u_0(t). \tag{7.30}$$

Here G_0 is the Green's function of (7.2). It has the following expression:

$$G_0(x,y) = \begin{cases} {a^*}^{-1} x(1-y), & x \le y, \\ {a^*}^{-1}(1-x)y, & x > y. \end{cases} \tag{7.31}$$

In particular,G_0 is Lipschitz continuous in each variable while the other is kept fixed. □

The next theorem accounts for MsFEM in media with long-range correlations. We recall that the random process $q(x)$ below is constructed as a function of a Gaussian process with long-range correlation.

Theorem 7.5. *Let u_ε, u_0, u_ε^h and u_0^h be defined as in the previous theorem. Let $q(x,\omega)$ and $a(x,\omega)$ be constructed as in* (L1)-(L2). *Then we have*

(i)

$$\sup_{x\in[0,1]} \left| \mathbb{E} \left(u_\varepsilon^h(x) - u_0^h(x) \right)^2 \right| \le C \left(\frac{\varepsilon}{h} \right)^\alpha, \tag{7.32}$$

for some constant C depending on (λ, Λ), κ, α, and f.

(ii) *As ε goes to zero while h is fixed, we have*

$$\frac{u_\varepsilon^h(x) - u_0^h(x)}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}_H^h(x; W^H) := \sigma_H \int_0^1 L^h(x,t) dW_t^H. \tag{7.33}$$

Here $H = 1 - \frac{\alpha}{2}$, and W_t^H is the standard fractional Brownian motion with Hurst index H. The constant σ_H is defined as $\sqrt{\kappa/H(2H-1)}$. The function $L^h(x,t)$ is defined as in the previous theorem.

(iii) *As h goes to zero, we have*

$$\mathcal{U}_H^h(x; W^H) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x; W^H) := \sigma_H \int_0^1 L(x,t) dW_t^H. \tag{7.34}$$

Remark 7.6. As before, this theorem says the diagram in Figure 7.1 commutes in the current case. In particular, $\alpha < 1$, and the scaling is $\varepsilon^{\frac{\alpha}{2}}$. Thus MsFEM passes the corrector test for both short-range and long-range correlations.

The stochastic integrals in (7.33) and (7.34) have fractional Brownian motions as integrators. We give a short review of such integrals below. A good reference is [96]. □

The next theorem addresses the convergence properties of HMM.

Theorem 7.7. *Let u_ε and u_0 be the solutions to* (7.1) *and* (7.2), *respectively. Let $u_\varepsilon^{h,\delta}$ be the HMM solution and u_0^h the standard finite element approximation of u_0.*

(i) *Suppose that the random processes $a(x)$ and $q(x)$ satisfy* (S1)-(S3). *Then*

$$\frac{u_\varepsilon^{h,\delta}(x) - u_0^h(x)}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}^{h,\delta}(x; W) \xrightarrow[h\to 0]{\text{distribution}} \sqrt{\frac{h}{\delta}}\, \mathcal{U}(x; W). \tag{7.35}$$

Here, $\mathcal{U}^{h,\delta}(x; W)$ is as in (7.27) *with L^h replaced by $L^{h,\delta}(x,t)$, which is defined in* (7.93) *below. The process* $\mathcal{U}(x;W)$ *is as in* (7.28).

(ii) *Suppose instead that the random processes* $a(x)$ *and* $q(x)$ *satisfy* (L1)-(L2). *Then*

$$\frac{u_\varepsilon^{h,\delta}(x) - u_0^h(x)}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}_H^{h,\delta}(x;W^H) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x;W^H). \tag{7.36}$$

Here, $\mathcal{U}_H^{h,\delta}(x;W^H)$ *is as in* (7.33) *with* L^h *replaced by* $L^{h,\delta}$, *and* $\mathcal{U}_H(x;W^H)$ *is as in* (7.34).

Remark 7.8. HMM is computationally less expensive than MsFEM when δ is much smaller than h. However, the theorem implies that this advantage comes at a price: when the random process $q(x)$ has short-range correlation, HMM amplifies the variance of the corrector. We will discuss methods to eliminate this effect in section 7.4.7. In the case of long-range correlations, however, HMM does pass the corrector test.

Intuitively, averaging occurs at the small scale $\delta \ll h$ for short-range correlations. Since HMM performs calculations on a small fraction of each interval h, each integral needs to be rescaled by h/δ to capture the correct mean, which over-amplifies the size of the fluctuations. In the case of long-range correlations, the self-similar structure of the limiting process shows that the convergence to the Gaussian process occurs simultaneously at all scales (larger than ε) and hence at the macroscopic scale. HMM may then be seen as a collocation method (with grid size h), which does capture the main features of the random integrals.

The amplification of the random fluctuations might be rescaled to provide the correct answer. The main difficulty is that the rescaling factor depends on the structure of the random medium and thus requires prior information or additional estimations about the medium. For general random media with no clear scale separation or no stationarity assumptions, the definition of such a rescaling coefficient might be difficult. In section 7.4.7, we present a hybrid method between HMM and MsFEM that is less expensive than the MsFEM presented above while still passing the corrector test. □

7.4 Proof of the Corrector Test Results

The starting point to prove the main theorems is to derive a formula for the corrector $u_\varepsilon^h - u_0^h$ for multi-scale schemes. This can be achieved for a large class of multi-scale schemes, namely those satisfying (N1)-(N3) in the previous section.

7.4.1 Expression for the corrector and convergence as $\varepsilon \to 0$

Let u_ε^h be the solution obtained from a multiscale numerical scheme satisfying (N1)-(N3). Now we derive an expression of the corrector, i.e., the difference between u_ε^h and u_0^h, the standard FEM solution to (7.2).

The function $u_0^h(x)$ is obtained from a weak formulation similar to (7.8) with a_ε replaced by a^*, and H_0^1 replaced by V_0^h, the space spanned by hat functions $\{\phi_0^j\}$. Clearly, these basis functions satisfy (N1) and (N3). Let A_0^h denote the associated stiffness matrix; then one can verify that it satisfies (N2). In fact, the vector b is given by $b_0^k = a^*/h$. Now $u_0^h(x)$ is simply $U_j^0 \phi_0^j$, where U^0 solves

$$A_0^h U^0 = F^0. \tag{7.37}$$

Subtracting this equation from (7.11), we obtain:

$$A_0^h (U^\varepsilon - U^0) = (F^\varepsilon - F^0) - (A_\varepsilon^h - A_0^h) U^\varepsilon.$$

Let G_0^h denote the inverse of the matrix A_0^h. We have

$$U^\varepsilon - U^0 = G_0^h (F^\varepsilon - F^0) - G_0^h (A_\varepsilon^h - A_0^h) U^\varepsilon.$$

Since both A_ε^h and A_0^h satisfy (N2), the difference $A_\varepsilon^h - A_0^h$ acts on vectors in the same manner as in (7.14). Since both $\{\phi_\varepsilon^j\}$ and $\{\phi_0^j\}$ satisfy (N3), we verify that

$$(F^\varepsilon - F^0)_j = -D^+(\tilde{F}_j^\varepsilon - \tilde{F}_j^0), \qquad \tilde{F}_j^\varepsilon := \int_{I_j} f(t)\tilde{\phi}_\varepsilon^j(t)dt.$$

Using these difference forms, we have

$$\begin{aligned} U_j^\varepsilon - U_j^0 &= -\sum_{m=1}^{N-1} (G_0^h)_{jm} \left(D^+(\tilde{F}^\varepsilon - \tilde{F}^0)_m - D^+\left((b_\varepsilon^m - b_0^m) D^- U_m^\varepsilon \right) \right) \\ &= \sum_{k=1}^{N} D^- G_{0jk}^h \left((\tilde{F}^\varepsilon - \tilde{F}^0)_k - (b_\varepsilon^k - b_0^k) D^- U_k^\varepsilon \right). \end{aligned} \tag{7.38}$$

The second equality is obtained from summation by parts. Note that we have extended the definitions of U^ε and U^0 so that they equal zero when the index is 0 or N. Similarly, $(G_0^h)_{j0}$ and $(G_0^h)_{jN}$ are zero as well.

The vector $U^\varepsilon - U^0$ is the corrector evaluated at the nodal points. We have the following control of its ℓ^2 norm under some assumptions on the statistics of $\{b_\varepsilon^k\}$ and $\{\phi_\varepsilon^j\}$.

Proposition 7.9. *Let U^ε and U^0 be as above. Let the basis functions $\{\phi_\varepsilon^j\}$ and the stiffness matrix A_ε^h satisfy* (N1)-(N3)*. Suppose also that*

$$\sup_{1\le k\le N} |D^- U_k^\varepsilon| \le C\|f\|_2 h^{\frac{1}{2}}, \tag{7.39}$$

for some universal constant C.

(i) *Suppose further that for any $k = 1, \cdots, N$, and any $x \in I_k$, we have*

$$\mathbb{E}\left(\tilde{\phi}_\varepsilon^k(x) - \tilde{\phi}_0^k(x)\right)^2 \le C\frac{\varepsilon}{h}\|R\|_{1,\mathbb{R}}, \tag{7.40}$$

and

$$\mathbb{E}\left(b_\varepsilon^k - b_0^k\right)^2 \le C\frac{\varepsilon}{h^3}\|R\|_{1,\mathbb{R}}, \tag{7.41}$$

for some universal constant C. Then we have

$$\mathbb{E}\left\|U^\varepsilon - U^0\right\|_{\ell^2}^2 \leq C\frac{\varepsilon}{h^2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2), \tag{7.42}$$

for some universal C.

(ii) *Suppose instead that the right hand side of* (7.40) *is $C\left(\frac{\varepsilon}{h}\right)^\alpha$, and the right hand side of* (7.41) *is $C\frac{1}{h^2}\left(\frac{\varepsilon}{h}\right)^\alpha$. Then the estimate in* (7.42) *should be changed to $C\frac{1}{h}\left(\frac{\varepsilon}{h}\right)^\alpha$.*

Remark 7.10. The assumption (7.39) essentially says that u_ε^h should have a Hölder regularity. Suppose the weak formulation associated to the multiscale scheme admits a unique solution u_ε^h such that $\|u_\varepsilon^h\|_{H^1} \leq C\|f\|_2$. Then by Morrey's inequality [57, p.266], $u_\varepsilon^h \in C^{0,\frac{1}{2}}$ in one dimension. Consequently, (7.39) holds.

For MsFEM, we have a better estimate: $|D^- U_k^\varepsilon| \leq Ch$ due to a super-convergence result; see (7.18). Therefore, the estimate in (7.42) can be improved to be $C\frac{\varepsilon}{h}$ in case (i) and $C\left(\frac{\varepsilon}{h}\right)^\alpha$ in case (ii). □

Item (i) of this proposition is useful when the random medium $a(x)$, or equivalently $q(x)$, has short range correlation, while item (ii) is useful in the case of long range correlations. The constant C in the second item depends on (λ, Λ), f and R_g, but not on h.

We remark also that throughout our analysis, the basis functions are assumed to be exact; that is to say, we do not account for the error in constructing $\{\phi_\varepsilon^j\}$.

Proof of Proposition 7.9. To prove (i), we use a super-convergence result, which we prove in Remark 7.2, to get $|D^- G_{0jk}^h| \leq Ch$. Using this estimate together with (7.39) and (7.38), we have

$$\mathbb{E}\left|U_j^\varepsilon - U_j^0\right|^2 \leq Ch^2\sum_{k=1}^N \mathbb{E}\left|\tilde{F}_k^\varepsilon - \tilde{F}_k^0\right|^2 + Ch^3\sum_{k=1}^N \mathbb{E}\left|b_\varepsilon^k - b_0^k\right|^2.$$

For the second term, we use (7.41) and obtain

$$\sum_{k=1}^N \mathbb{E}\left|b_\varepsilon^k - b_0^k\right|^2 \leq \sum_{k=1}^N C\frac{\varepsilon}{h^3}\|R\|_{1,\mathbb{R}} = C\frac{\varepsilon}{h^4}\|R\|_{1,\mathbb{R}}. \tag{7.43}$$

For the other term, an application of Cauchy-Schwarz to the definition of $\tilde{F}^\varepsilon$ yields

$$\left|\tilde{F}_k^\varepsilon - \tilde{F}_k^0\right|^2 \le \|f\|_{2,I_k}^2 \|\tilde{\phi}_\varepsilon^k - \tilde{\phi}_0^k\|_{2,I_k}^2.$$

Using (7.40), we have

$$\mathbb{E}\|\tilde{\phi}_\varepsilon^k - \tilde{\phi}_0^k\|_{2,I_k}^2 = \int_{I_k} \mathbb{E}\left(\tilde{\phi}_\varepsilon^k - \tilde{\phi}_0^k\right)^2 (x)dx \le C\frac{\varepsilon}{h}\cdot h\|R\|_{1,\mathbb{R}} = C\varepsilon\|R\|_{1,\mathbb{R}}. \tag{7.44}$$

Therefore, we have

$$\mathbb{E}\left|U_j^\varepsilon - U_j^0\right|^2 \le \left(Ch^2\sum_{k=1}^N \|f\|_{2,I_k}\varepsilon + Ch^4\frac{\varepsilon}{h^4}\right)\|R\|_{1,\mathbb{R}} \le C\varepsilon\|R\|_{1,\mathbb{R}}(1+\|f\|_2).$$

Note that this estimate is uniform in j. Sum over j to complete the proof of (i).

Proof of item (ii) follows in exactly the same way, using the corresponding estimates. □

Now, the corrector in this general multi-scale numerical scheme is:

$$\begin{aligned} u_\varepsilon^h(x) - u_0^h(x) &= U_j^\varepsilon\phi_\varepsilon^j(x) - U_j^0\phi_0^j(x) \\ &= (U^\varepsilon - U^0)_j\phi_0^j(x) + U_j^0(\phi_\varepsilon^j - \phi_0^j)(x) + (U^\varepsilon - U^0)_j(\phi_\varepsilon^j - \phi_0^j)(x). \end{aligned} \tag{7.45}$$

We call the three terms on the right hand side $K_i(x)$, $i = 1,2,3$. Now $K_1(x)$ is the piecewise interpolation of the corrector evaluated at the nodal points; $K_2(x)$ is the corrector due to different choices of basis functions; and $K_3(x)$ is much smaller due to the previous proposition and (7.40). Our analysis shows that $K_1(x)$ and $K_2(x)$ contribute to the limit when $\varepsilon \to 0$ while h is fixed, but only a part of $K_1(x)$ contributes to the limit when $h \to 0$.

Due to self-averaging effect which are made precise in Lemma 7.13, integrals of $q_\varepsilon(x)$ are small. Therefore, our goal is to decompose the above expression into two terms: a leading term which is an oscillatory integral against q_ε, and a remainder term which contains multiple oscillatory integrals.

Proposition 7.11. *Assume that u_ε^h is the solution to* (7.1) *obtained from a multi-scale scheme, which satisfies* (N1)-(N3) *and has basis functions $\{\phi_\varepsilon^j\}$, and that u_0^h is the solution of* (7.2) *obtained by the standard FEM with hat basis functions $\{\phi_0^j\}$. Let b_ε and b_0 denote the vectors in* (N2) *of these methods. Suppose that* (7.39) *holds and that for any $k = 1, \cdots, N$, we have*

$$\begin{aligned} \tilde{\phi}_\varepsilon^k(t) - \tilde{\phi}_0^k(t) &= [1+\tilde{r}_{1k}]\frac{a^*}{h}\left(\int_{x_{k-1}}^t q_\varepsilon(s)ds - \frac{t-x_{k-1}}{h}\int_{x_{k-1}}^{x_k} q_\varepsilon(s)ds\right), \\ b_\varepsilon^k - b_0^k &= [1+\tilde{r}_{2k}]\left(-\frac{a^{*2}}{h^2}\int_{x_{k-1}}^{x_k} q_\varepsilon(t)dt\right), \end{aligned} \tag{7.46}$$

for some random variables $\tilde{r}_{1k}$ and $\tilde{r}_{2k}$.

(i) *Assume that $q(x)$ has short range correlation, i.e., satisfies* (S2), *and that*

$$\sup_{1\le k\le N} \max\{\mathbb{E}|\tilde{r}_{1k}|^2, \mathbb{E}|\tilde{r}_{2k}|^2\} \le C\frac{\varepsilon}{h}\|R\|_{1,\mathbb{R}}, \tag{7.47}$$

for some universal constant C. Then, the corrector can be written as

$$u_\varepsilon^h(x) - u_0^h(x) = \int_0^1 L^h(x,t)q_\varepsilon(t)dt + r_\varepsilon^h(x). \tag{7.48}$$

Furthermore, the remainder $r_\varepsilon^h(x)$ satisfies

$$\sup_{x\in[0,1]} \mathbb{E}|r_\varepsilon^h(x)| \le C\frac{\varepsilon}{h^2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2), \tag{7.49}$$

for some universal constant C. The function $L^h(x,t)$ is the sum of L_1^h and L_2^h defined by:

$$L_1^h(x,t) = \sum_{k=1}^{N} \mathbf{1}_{I_k}(t) \frac{a^* D^- G_0^h(x,x_k)}{h} \left(\frac{a^* D^- U_k^0}{h} + \int_t^{x_k} f(s)ds - \int_{x_{k-1}}^{x_k} f(s)\tilde{\phi}_0^k(s)ds \right),$$
$$L_2^h(x,t) = \frac{a^*}{h} D^- U_{0j(x)}^h \left(\mathbf{1}_{[x_{j(x)-1},x]}(t) - \frac{x - x_{j(x)-1}}{h} \mathbf{1}_{[x_{j(x)-1},x_{j(x)}]}(t) \right). \tag{7.50}$$

Given x, the index $j(x)$ is the unique one so that $x_{j(x)-1} < x \le x_{j(x)}$. The function $G_0^h(x,x_k)$ is defined as

$$G_0^h(x,x_k) = \sum_{j=1}^{N-1} G_{0jk}^h \phi_0^j(x). \tag{7.51}$$

G_0^h is the interpolation in V_0^h using the discrete Green's function of standard FEM.

(ii) *Assume that $q(x)$ has long range correlation, i.e.,* (L1)-(L2) *are satisfied, and that the estimate in* (7.47) *is $C\left(\frac{\varepsilon}{h}\right)^\alpha$. Then the same decomposition holds, the expression of $L^h(x,t)$ remains the same, but the estimate in* (7.49) *should be replaced by $C\left(\frac{\varepsilon}{h}\right)^\alpha$.*

Remark 7.12. Due to the super-convergent result in Remark 7.2, the function $G_0^h(x,x_k)$ above is exactly the Green's function evaluated at (x,x_k). This can be seen from the facts that they agree at nodal points and both are piece-wise linear and continuous.

Proof. We only present the proof of item (i). Item (ii) follows in exactly the same way. We point out that the assumption (7.46) and the estimates (7.47) imply (7.44) and (7.43) thanks to Lemma 7.13.

The idea is to extract the terms in the expression (7.45) that are linear in q_ε. For $K_1(x)$, we use (7.38) and write

$$\begin{aligned} K_1(x) &\approx \sum_{j=1}^{N-1} \sum_{k=1}^{N} D^- G_{0jk}^h (\tilde{F}_k^\varepsilon - \tilde{F}_k^0 - (b_\varepsilon^k - b_0^k) D^- U_k^0) \phi_0^j(x) \\ &= \sum_{k=1}^{N} D^- G_0^h(x,x_k)(\tilde{F}_k^\varepsilon - \tilde{F}_k^0 - (b_\varepsilon^k - b_0^k) D^- U_k^0). \end{aligned}$$

Note that the expression above is an approximation because we have changed D^-U^ε on the right hand side of (7.38) to D^-U^0. The error is

$$r_{11}^h(x) = -\sum_{k=1}^N D^-G_0^h(x,x_k)(b_\varepsilon^k - b_0^k)D^-(U^\varepsilon - U^0)_k. \tag{7.52}$$

Estimating $|D^-G_0^h|$ by Ch and using Cauchy-Schwarz on the sum over k and (7.43) and Lemma (7.42), we verify that $\mathbb{E}|r_{11}^h(x)| \le C\varepsilon h^{-2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2)$.

Using the expressions of $\tilde\phi_\varepsilon$ and b_ε, and the estimates of the higher order terms in them, (7.46), we can further approximate $K_1(x)$ by

$$\begin{aligned} K_1(x) \approx \sum_{k=1}^N D^-G_0^h(x,x_k)\Bigg(&\int_{x_{k-1}}^{x_k} f(t)\frac{a^*}{h}\left[\int_{x_{k-1}}^t q_\varepsilon(s)ds - \frac{t-x_{k-1}}{h}\int_{x_{k-1}}^{x_k} q_\varepsilon(s)ds\right]dt \\ &+ \frac{a^{*2}}{h^2}D^-U_k^0\int_{x_{k-1}}^{x_k} q_\varepsilon(t)dt\Bigg). \end{aligned}$$

The error in this approximation is:

$$\begin{aligned} r_{12}^h(x) = \sum_{k=1}^N D^-G_0^h(x,x_k)\Bigg(&\tilde r_{1k}\int_{x_{k-1}}^{x_k} f(t)\frac{a^*}{h}\left[\int_{x_{k-1}}^t q_\varepsilon(s)ds - \frac{t-x_{k-1}}{h}\int_{x_{k-1}}^{x_k} q_\varepsilon(s)ds\right]dt \\ &+ \tilde r_{2k}\frac{a^{*2}}{h^2}D^-U_k^0\int_{x_{k-1}}^{x_k} q_\varepsilon(t)dt\Bigg). \end{aligned} \tag{7.53}$$

Using Lemma 7.13, (7.47) and Cauchy-Schwarz, we have

$$\mathbb{E}\left|\tilde r_{1k}\int_{I_k} q_\varepsilon(s)ds\right| \le C\varepsilon.$$

Using this estimate, we verify that the mean of the absolute value of the first term in r_{12}^h is bounded by $C\varepsilon\|f\|_2\|R\|_{1,\mathbb{R}}$. A similar estimate with $|D^-U_k^0| \le Ch$ (in Remark 7.2) shows that the second term in r_{12}^h has absolute mean bounded by $C\varepsilon h^{-1}\|R\|_{1,\mathbb{R}}$. Therefore, we have $\mathbb{E}|r_{12}^h(x)| \le C\varepsilon h^{-1}(1+\|f\|_2)\|R\|_{1,\mathbb{R}}$. We remark also that in the case of long range correlations, we should apply Lemma 7.14 instead.

Moving on to $K_2(x)$, we observe that for fixed x, $K_2(x)$ reduces to a sum over at most two terms, due to the fact that ϕ_ε^j and ϕ_0^j have local support only. Let $j(x)$ be the index so that $x \in (x_{j(x)-1}, x_{j(x)}]$. We have

$$\begin{aligned} K_2(x) &= \sum_{j=1}^{N} D^- U_j^0 (\tilde{\phi}_\varepsilon^j - \tilde{\phi}_0^j)(x) = D^- U_{j(x)}^0 (\tilde{\phi}_\varepsilon^{j(x)} - \tilde{\phi}_0^{j(x)})(x) \\ &\approx D^- U_{j(x)}^0 \frac{a^*}{h} \left(\int_{x_{j(x)-1}}^{x} q_\varepsilon(t)dt - \frac{x - x_{j(x)-1}}{h} \int_{x_{j(x)-1}}^{x_{j(x)}} q_\varepsilon(t)dt \right). \end{aligned}$$

In the second step above, we used the decomposition of $\tilde{\phi}_\varepsilon$ again. The error we make in this step is

$$r_2^h(x) = \tilde{r}_{1j(x)} \frac{a^* D^- U_{j(x)}^0}{h} \left(\int_{x_{j(x)-1}}^{x} q_\varepsilon(t)dt - \frac{x - x_{j(x)-1}}{h} \int_{x_{j(x)-1}}^{x_{j(x)}} q_\varepsilon(t)dt \right). \tag{7.54}$$

We verify again that $\mathbb{E}|r_2^h(x)| \le C\varepsilon \|R\|_{1,\mathbb{R}}$.

Now for $K_3(x)$, we use Cauchy-Schwarz and have

$$\mathbb{E}|K_3(x)| \le \mathbb{E} \left(\sum_{j=1}^{N-1} |U_{\varepsilon j}^h - U_{0j}^h|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{N-1} (\phi_\varepsilon^j - \phi_0^j)^2(x) \right)^{\frac{1}{2}} \le C\frac{\varepsilon}{h} \|R\|_{1,\mathbb{R}} (1 + \|f\|_2). \tag{7.55}$$

The last inequality is due to (7.42) and (7.44).

In the approximations of $K_1(x)$ and $K_2(x)$, we change the order of summation and integration. We find that $K_1(x)$ is then $\int_0^1 L_1^h(x,t) q_\varepsilon(t)dt$ plus the error term $r_{11}^h + r_{12}^h$, and $K_2(x)$ is $\int_0^1 L_2^h(x,t) q_\varepsilon(t)dt$ plus the error term r_2^h. Therefore we proved (7.48) with $r_\varepsilon^h(x) = r_{11}^h + r_{12}^h + r_2^h + K_3(x)$. The estimates above for these error terms are uniform in x, verifying (7.49). $\square$

7.4.2 Weak convergence of the corrector of a multiscale scheme

In this section, we prove the weak convergence of the corrector $u_\varepsilon^h - u_0^h$ in the general setting. We first record two key estimates on oscillatory integrals, which we have used already. The first one accounts for short range media.

Lemma 7.13. *Let $q(x,\omega)$ be a mean-zero stationary random process with integrable correlation function $R(x)$. Let $[a,b]$ and $[c,d]$ be two intervals on $\mathbb{R}$ and assume $b-a \le d-c$. Then*

$$\left| \mathbb{E} \int_a^b \int_c^d q_\varepsilon(t) q_\varepsilon(s) dtds \right| \le \varepsilon (b-a) \|R\|_{1,\mathbb{R}}. \tag{7.56}$$

Proof. Let T denotes the expectation of the double integral. It has the following expression:

$$T = \int_a^b \int_c^d R(\frac{t-s}{\varepsilon}) dtds = \int_{\mathbb{R}} \int_{\mathbb{R}} R(\frac{t-s}{\varepsilon}) \mathbf{1}_{[a,b]}(t) \mathbf{1}_{[c,d]}(s) dtds.$$

We change variables by setting $t \to t$ and $(t-s)/\varepsilon \to s$. The Jacobian of this change of variables is ε. Then we have

$$|T| \le \varepsilon \int_{\mathbb{R}} \int_{\mathbb{R}} |R(s)| \mathbf{1}_{[a,b]}(t) dtds = \varepsilon (b-a) \|R\|_{1,\mathbb{R}}.$$

This completes the proof. □

The second one accounts for a special family of long range media. The proof is adapted from [11].

Lemma 7.14. *Let $q(x,\omega)$ be defined as in* (L1)-(L2). *Let F be a function in the space $L^\infty(\mathbb{R})$. Let (a,b) and (c,d) be two open intervals and assume $b-a \le d-c$. Then we have*

$$\left| \mathbb{E} \int_a^b \int_c^d q(\frac{t}{\varepsilon}) q(\frac{s}{\varepsilon}) F(t) F(s) dtds \right| \le C \varepsilon^\alpha (b-a)(d-c)^{1-\alpha}. \tag{7.57}$$

The constant C above depends only on κ, α and $\|F\|_{\infty,\mathbb{R}}$.

Proof. By the definition of the correlation function R, we have

$$\mathbb{E}\Big\{\frac{1}{\varepsilon^\alpha}\int_a^b\int_c^d q(\frac{t}{\varepsilon})q(\frac{s}{\varepsilon})F(t)F(s)dtds\Big\} = \int_{\mathbb{R}^2}\varepsilon^{-\alpha}R(\frac{t-s}{\varepsilon})F(t)\mathbf{1}_{[a,b]}(t)F(s)\mathbf{1}_{[c,d]}(s)dtds.$$

As shown in [11], $R(\tau)$ is asymptotically $\kappa\tau^{-\alpha}$ with κ defined in (7.7). We expect to replace R by $\kappa\tau^{-\alpha}$ in the limit. Therefore, let us consider the difference

$$\int_{\mathbb{R}^2}\left|\varepsilon^{-\alpha}R(\frac{t-s}{\varepsilon}) - \frac{\kappa}{|t-s|^\alpha}\right||F(t)|\mathbf{1}_{[a,b]}(t)|F(s)|\mathbf{1}_{[c,d]}(s)dtds.$$

By the asymptotic relation $R \sim \kappa\tau^{-\alpha}$, we have for any $\delta > 0$, the existence of T_δ such that $|R(\tau) - \kappa\tau^{-\alpha}| \le \delta\tau^{-\alpha}$. Accordingly, we decompose the domain of integration into three subdomains:

$$D_1 = \{(t,s)\in\mathbb{R}^2, |t-s| \le T_\delta\varepsilon\},$$
$$D_2 = \{(t,s)\in\mathbb{R}^2, T_\delta\varepsilon < |t-s| \le 1\},$$
$$D_3 = \{(t,s)\in\mathbb{R}^2, 1 < |t-s|\}.$$

On the first domain, we have

$$\begin{aligned}&\int_{D_1}\left|\varepsilon^{-\alpha}R(\frac{t-s}{\varepsilon}) - \frac{\kappa}{|t-s|^\alpha}\right||F(t)|\mathbf{1}_{[a,b]}(t)|F(s)|\mathbf{1}_{[c,d]}(s)dtds\\ &\quad\le \int_{D_1}\left|\varepsilon^{-\alpha}R(\frac{t-s}{\varepsilon})\right||F_1(t)||F_2(s)|dtds + \int_{D_1}\left|\frac{\kappa}{|t-s|^\alpha}\right||F_1(t)||F_2(s)|dtds.\end{aligned}$$

Here and below, we use the short hand notation $F_1(t) = F(t)\mathbf{1}_{[a,b]}(t)$ and $F_2(s) = F(s)\mathbf{1}_{[c,d]}(s)$. The above integrals are then bounded by

$$\begin{aligned}&\varepsilon^{-\alpha}\|R\|_{\infty,\mathbb{R}}\int_a^b|F(t)|\int_{t-T_\delta\varepsilon}^{t+T_\delta\varepsilon}|F_2(s)|dsdt + \int_a^b|F(t)|\int_{-T_\delta\varepsilon}^{T_\delta\varepsilon}\kappa|s|^{-\alpha}|F_2(t-s)|dsdt\\ \le&\|F\|^2_{\mathbb{R},\infty}\Big(2T_\delta\|R\|_{\mathbb{R},\infty} + \frac{2\kappa T_\delta^{1-\alpha}}{1-\alpha}\Big)\varepsilon^{1-\alpha}.\end{aligned}$$

On domain D_2, we have

$$\int_{D_2} \left| \varepsilon^{-\alpha} R(\frac{t-s}{\varepsilon}) - \frac{\kappa}{|t-s|^\alpha} \right| |F_1(t)||F_2(s)| dtds \le \delta \int_{D_2} |t-s|^{-\alpha} |F_1(t)||F_2(s)| dtds$$
$$\le 2\delta \int_a^b |F(t)| \int_{T_\delta \varepsilon}^1 |s|^{-\alpha} |F_2(t-s)| dsdt \le \frac{2\delta \|F\|^2_{\infty,\mathbb{R}}}{1-\alpha} (1 + T_\delta^{1-\alpha} \varepsilon^{1-\alpha}).$$

On domain D_3, we can bound $|t-s|^{-\alpha}$ by one, and we have

$$\int_{D_3} \left| \varepsilon^{-\alpha} R(\frac{t-s}{\varepsilon}) - \frac{\kappa}{|t-s|^\alpha} \right| |F_1(t)||F_2(s)| dtds \le \delta \int_{D_3} |t-s|^{-\alpha} |F_1(t)||F_2(s)| dtds$$
$$\le 2\delta \int_a^b |F(t)| \int_{T_\delta \varepsilon}^1 |F(t-s)| dsdt \le 2\delta \|F\|^2_{\infty,\mathbb{R}} (1 + T_\delta \varepsilon).$$

Therefore, for some constant C that does not depend on ε or δ, we have

$$\limsup_{\varepsilon \to 0} \varepsilon^{-\alpha} \left| \mathbb{E} \int_a^b \int_c^d q(\frac{t}{\varepsilon}) q(\frac{s}{\varepsilon}) F_1(t) F_2(s) dtds - \int_{\mathbb{R}^2} \frac{\kappa}{|t-s|^\alpha} F_1(t) F_2(s) dtds \right| \le C \|F\|^2_{\infty,\mathbb{R}} \delta.$$

Sending δ to zero, we see that

$$\lim_{\varepsilon \to 0} \varepsilon^{-\alpha} \mathbb{E} \int_a^b \int_c^d q(\frac{t}{\varepsilon}) q(\frac{s}{\varepsilon}) F_1(t) F_2(s) dtds = \int_{\mathbb{R}^2} \frac{\kappa}{|t-s|^\alpha} F_1(t) F_2(s) dtds. \tag{7.58}$$

Finally, from the Hardy-Littlewood-Sobolev inequality [78, §4.3], we have

$$\left| \int_{\mathbb{R}^2} \frac{F_1(t) F_2(s)}{|t-s|^\alpha} dtds \right| \le C \|F_1\|_{1,\mathbb{R}} \|F_2\|_{(1-\alpha)^{-1},\mathbb{R}} \le C \|F\|^2_\infty (b-a)(d-c)^{1-\alpha}. \tag{7.59}$$

This completes the proof. □

Now we are ready to characterize the limit of the corrector in the multiscale scheme when ε is sent to zero. As we have seen before, the scaling depends on the correlation range of the random media.

Proposition 7.15. *Let u^h_ε be the solution to* (7.1) *given by a multi-scale scheme that* *satisfies* (N1)-(N3). *Suppose* (7.39) *holds. Let* u_0^h *be the standard FEM solution to* (7.2).

(i) *Suppose that* $q(x)$ *satisfies* (S1)-(S3) *and that the conditions of item one in Proposition 7.11 hold. Then,*

$$\frac{u_\varepsilon^h - u_0^h}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma \int_0^1 L^h(x,t)dW_t. \tag{7.60}$$

(ii) *Suppose that* $q(x)$ *satisfies* (L1)-(L2) *and that the conditions of item two in Proposition 7.11 hold. Then,*

$$\frac{u_\varepsilon^h - u_0^h}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma_H \int_0^1 L^h(x,t)dW_t^H. \tag{7.61}$$

The real number σ *is defined in* (7.4) *and* σ_H *is defined in Theorem 7.5.*

These results are exactly what we need to prove the weak convergence in step (ii) of the diagram in Figure 7.1. Note our assumptions allow for general schemes other than MsFEM. A standard method to attain such weak convergence results is to use Proposition 2.36.

We will prove item one of Proposition 7.15 in detail; proof of item two follows in the same way, so we only point out the necessary modifications. Recall the decomposition in (7.48). Let I_ε denote the first member on the right hand side of this equation, i.e., the oscillatory integral. Let $\mathcal{I}^h$ denote the right hand side of (7.60). The strategy in the case of short range media is to show that $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$ converges in distribution in $\mathcal{C}$ to the target process $\mathcal{I}^h$, while $\{\varepsilon^{-\frac{1}{2}} r_\varepsilon^h\}$ converges in distribution in $\mathcal{C}$ to the zero function. Since the zero process is deterministic, the convergence in fact holds in probability; see [24, p.27]. Then (7.60) follows.

Proof. **Convergence of** $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$. We first check that the finite dimensional distributions of $I_\varepsilon(x)$ converge to those of $\mathcal{I}^h(x)$. Using characteristic functions, this amounts to showing

$$\mathbb{E}\exp\left(i\cdot\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)\textstyle\sum_{j=1}^n \xi_j L^h(x^j,t)dt\right) \xrightarrow{\varepsilon\to 0} \mathbb{E}\exp\left(i\sigma\int_0^1 \textstyle\sum_{j=1}^n \xi_j L^h(x^j,t)dW_t\right),$$

for any positive integer n, and any n-tuple $(x^1,\cdots,x^n)$ and n-tuple $(\xi_1,\cdots,\xi_n)$. We set

$m(t) = \sum_{j=1}^{n} \xi_j L^h(x^j, t)$. The convergence above holds if we can show

$$\frac{1}{\sqrt{\varepsilon}} \int_0^1 q_\varepsilon(t) m(t) dt \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma \int_0^1 m(t) dW_t, \tag{7.62}$$

for any $m(t)$ that is square integrable on $[0, 1]$. Indeed, this convergence holds, due to Theorem 2.15 since by assumption $q(x, \omega)$ is a stationary mean-zero process that admits an integrable ρ-mixing coefficient $\rho(r) \in L^1(\mathbb{R})$. Therefore, we proved the convergence of the finite distributions of $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$.

Next, we establish tightness of $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon(x)\}$ by verifying (2.49). Consider the fourth moments and recall $L^h = L_1^h + L_2^h$ in (7.50); we have

$$\begin{aligned} \mathbb{E}(I_\varepsilon(x) - I_\varepsilon(y))^4 \leq 8 \Big\{ \mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}} \int_0^1 q_\varepsilon(t)(L_1^h(x,t) - L_1^h(y,t)) dt \Big)^4 \\ + \mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}} \int_0^1 q_\varepsilon(t)(L_2^h(x,t) - L_2^h(y,t)) dt \Big)^4 \Big\}. \end{aligned} \tag{7.63}$$

We estimate the two terms on the right separately. For the first term we observe that $L_1^h(x, t)$ is Lipschitz continuous in x. This is due to the fact that $G_0^h(x, x_k)$ is Lipschitz in x with a universal Lipschitz coefficient. Since the other terms in the expression of $L_1^h(x, t)$ in (7.50) are bounded by C, we have

$$|L_1^h(x,t) - L_2^h(y,t)| \leq \frac{C}{h} |x - y|.$$

We use this fact and apply Lemma 7.21 to deduce

$$\mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}} \int_0^1 q_\varepsilon(t)(L_1^h(x,t) - L_1^h(y,t)) dt \Big)^4 \leq \frac{C}{h^4} |x - y|^4. \tag{7.64}$$

The constant C above depends on λ, Λ, and $\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}$.

To estimate the second term in (7.63), consider two distinct points $y < x$. Let j and k be the indices such that $x \in (x_{j-1}, x_j]$ and $y \in (x_{k-1}, x_k]$. Then one of the following holds:

$j-k \geq 2$, $j-k=0$ or $j-k=1$. In the first case, since $|D^-U^0| \leq Ch$ for some C depending on λ, Λ and $\|f\|_2$, we have the following crude bound.

$$|L_2^h(x,t) - L_2^h(y,t)| \leq C \leq \frac{C}{h}|x-y|.$$

The same analysis leading to (7.64) applies, and the second term in (7.63) is bounded by $C|x-y|^4/h^4$ in this case.

When $|j-k|=0$, x and y are in the same interval (x_j, x_{j+1}). We can write

$$\int_0^1 q_\varepsilon(t)(L_2^h(x,t) - L_2^h(y,t))dt = \frac{a^* D^- U_{0j}^h}{h} \left(\int_y^x q_\varepsilon(t)dt - \frac{x-y}{h} \int_{I_j} q_\varepsilon(t)dt \right). \tag{7.65}$$

Since x and y are in the same interval, the function $(x-y)/h$ is bounded by one. Now Lemma 7.21 applies and we see that the fourth moments of the members in (7.65) are bounded by

$$C\left[\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_x^y q_\varepsilon(t)dt\right)^4 + \left(\frac{x-y}{h}\right)^4 \mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_{I_k} q_\varepsilon(t)dt\right)^4\right] \leq C|x-y|^2.$$

When $j-k=1$, we have

$$\begin{aligned}\int_0^1 q_\varepsilon(t)L_2^h(y,t)dt &= \frac{a^* D^- U_{0j-1}^h}{h}\left(\int_{x_{j-2}}^y q_\varepsilon(t)dt - \frac{y-x_{j-2}}{h}\int_{x_{j-2}}^{x_{j-1}} q_\varepsilon(t)dt\right)\\ &= \frac{a^* D^- U_{0j-1}^h}{h}\left(-\int_y^{x_{j-1}} q_\varepsilon(t)dt - \frac{y-x_{j-1}}{h}\int_{x_{j-2}}^{x_{j-1}} q_\varepsilon(t)dt\right).\end{aligned}$$

Let x_{j-1} play the role of x in (7.65) and notice that $L_2^h(x_{j-1},t)=0$. We get

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(y,t)dt\right)^4 \leq C\frac{|y-x_{j-1}|^2}{h^2}.$$

Similarly, in the interval where x lands, let x_{j-1} play the role of y in (7.65). We have

$$\mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(x,t)dt\Big)^4 \le C\frac{|x-x_{j-1}|^2}{h^2}.$$

We combine these estimates and see that in this case, the second term in (7.63) is bounded by

$$\begin{aligned}&8\Big(\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(y,t)dt\Big)^4 + \mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(x,t)dt\Big)^4\right)\\ \le& C\frac{|x_{j-1}-y|^2+|x-x_{j-1}|^2}{h^2} \le C\frac{|x-y|^2}{h^2}.\end{aligned}$$

In the last inequality, we used the fact that $a^2+b^2 \le (a+b)^2$ for two non-negative numbers a and b.

Combine these three cases to conclude that for any $x, y \in [0,1]$, the second term in (7.63) is bounded by $C|x-y|^2/h^2$. This, together with (7.64), shows

$$\mathbb{E}\Big(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)(L^h(x,t)-L^h(y,t))dt\Big)^4 \le C\frac{|x-y|^2}{h^4}. \tag{7.66}$$

In other words, $\{\varepsilon^{-\frac{1}{2}}I_\varepsilon(x)\}$ satisfies (2.49) with $\beta = 4$ and $\delta = 1$, and is therefore a tight sequence. Consequently, it converges to $\mathcal{I}^h$ in distribution in $\mathcal{C}$.

Convergence of $\{\varepsilon^{-\frac{1}{2}}r_\varepsilon^h\}$. For the convergence of finite dimensional distributions, we need to show

$$\mathbb{E}\exp\left(i\cdot\frac{1}{\sqrt{\varepsilon}}\textstyle\sum_{j=1}^n \xi^j r_\varepsilon^h(x^j)\right) \to 1,$$

for any fixed n, $\{x^j\}_{j=1}^n$ and $\{\xi_j\}_{j=1}^n$. Since $|e^{i\theta}-1| \le |\theta|$ for any real number θ, the left hand side of the equation above can be bounded by

$$\frac{1}{\sqrt{\varepsilon}}\mathbb{E}\Big|\textstyle\sum_j \xi_j r_\varepsilon^h(x_j)\Big| \le \textstyle\sum_j |\xi_j|\frac{1}{\sqrt{\varepsilon}}\sup_{1\le j\le n}\mathbb{E}|r_\varepsilon^h(x_j)|.$$

The last sum above converges to zero thanks to (7.49), completing the proof of convergence of finite dimensional distributions.

For tightness, we recall that $r_\varepsilon^h(x)$ consists of r_{11}^h in (7.52), r_{12}^h in (7.53), $K_3(x)$ in (7.48) and $r_2^h(x)$ in (7.54). In the first three functions, x appears in Lipschitz continuous terms, e.g., in $D^- G_0^h(x;x_k)$ or $\phi_\varepsilon^j(x) - \phi_0^j(x)$. Meanwhile, the terms that are x-independent have mean square of order $O(\varepsilon)$ or less. Therefore, we can choose $\beta = 2$ and $\delta = 1$ in (2.49). For instance, we consider $r_{12}^h(x)$ in (7.53) and bound the terms that are not $\tilde{r}_{1k}$ or $\tilde{r}_{2k}$ in the parenthesis by some constant C. Using Lipschitz continuity of $D^- G_0^h$, we have

$$\mathbb{E}\left(\frac{r_{12}^h(x) - r_{12}^h(y)}{\sqrt{\varepsilon}}\right)^2 \le C\frac{1}{\varepsilon}|x-y|^2 \sup_k \mathbb{E}\{|\tilde{r}_{1k}|^2 + |\tilde{r}_{2k}|^2\} \le C\frac{|x-y|^2}{h},$$

thanks to the estimate (7.47). Similarly, we can control r_{11}^h and K_3.

For r_2^h in (7.54), we observe that it has the form of the main part of $K_2(x)$, which corresponds to $L_2^h(x,t)$ and the second term in (7.63), except the extra integral of q_ε. Therefore, the tightness argument for the second term in (7.63) can be repeated. The extra q_ε term is favorable: we can choose $\beta = 2$ and $\delta = 1$ in (2.49).

To summarize, $\{\varepsilon^{-\frac{1}{2}} r_\varepsilon^h/\sqrt{\varepsilon}\}$ can be shown to be tight by choosing $\beta = 2$ and $\delta = 1$ in (2.49). Therefore, it converges to the zero function in distribution in $\mathcal{C}$. We have thus established the convergence in (7.60).

The case of long range media. In this case, the scaling is $\varepsilon^{-\frac{\alpha}{2}}$. The proof is almost the same as above and we only point out the key modifications.

Let us denote the right hand side of (7.61) by $\mathcal{I}_H^h$. To show the convergence of the finite dimensional distributions of $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon\}$, instead of using (7.62), we need the following analogue for random media with long range correlations:

$$\frac{1}{\varepsilon^{\frac{\alpha}{2}}}\int_0^1 q_\varepsilon(t) m(t) dt \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma_H \int_0^1 m(t) dW_t^H, \tag{7.67}$$

where σ_H is defined below (7.33). The above holds only thanks to Theorem 2.34. Hence we conclude that the finite dimensional distributions of $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon(x)\}$ converge to those of $\mathcal{I}_H^h$.

For the tightness of $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon(x)\}$, we can follow the same procedures that lead to (7.64) and (7.65). We only need to consider second order moments when applying the Kolmogorov criterion thanks to Lemma 7.57, which says

$$\mathbb{E}\Big(\frac{1}{\varepsilon^{\frac{\alpha}{2}}}\int_x^y q_\varepsilon(t)dt\Big)^2 \le C|x-y|^{2-\alpha}. \tag{7.68}$$

In the short range case, since α equals one we only have $|x-y|$ on the right. To get an extra exponent δ, we had to consider fourth moments. In the long range case, α is less than one, so we gain a $\delta = 1-\alpha$ from the above estimate. With this in mind, we can simplify the proof we did for (7.60) to prove that $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon\}$ converges to $\mathcal{I}_H^h$. Similarly, $\{\varepsilon^{-\frac{\alpha}{2}} r_\varepsilon^h\}$ converges to the zero function in distribution, and hence in probability, in the space $\mathcal{C}$. The conclusion is that (7.61) holds. This completes the proof of Proposition 7.15. □

Remark 7.16. From the proofs of the propositions in this section, the results often hold if the conditions in item (i) or (ii) of Proposition 7.11 are violated in an ε-independent manner. For instance, if the second equation in (7.46) is modified to

$$b_\varepsilon^k - b_0^k = c(h)[1+\tilde{r}_{2k}]\left(-\frac{a^{*2}}{h^2}\int_{D_k} q_\varepsilon(t)dt\right), \tag{7.69}$$

for some function $c(h)$ and for region $D_k \subset I_k$, then this modification will be carried to $L^h(x,t)$ and following estimates, but the weak convergences in Proposition 7.15 still hold.

7.4.3 Weak convergence as h goes to 0

In the previous section, we established weak convergence of the corrector $u_\varepsilon^h - u_0^h$ of a general multi-scale scheme when the correlation length ε of the random medium goes to zero while the discretization h is fixed. In this section, we send h to zero, and characterize the limiting process. We aim to prove the following statement.

Proposition 7.17. *Let* $L^h(x,t)$ *be defined as in* (7.50). *As* h *goes to zero, the Gaussian processes on the right hand sides of* (7.60) *and* (7.61) *have the following limits in distribution in* $\mathcal{C}$*:*

$$\sigma \int_0^1 L^h(x,t) dW_t \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}(x;W), \tag{7.70}$$

where $\mathcal{U}$ *is the Gaussian process in* (7.28). *Similarly,*

$$\sigma_H \int_0^1 L^h(x,t) dW_t^H \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x;W^H), \tag{7.71}$$

where $\mathcal{U}_H$ *is the Gaussian process in* (7.34).

We consider the case of short range random media first. Recall that $\mathcal{I}^h(x)$ denotes the left hand side of (7.70). It can be split further into three terms as follows. Let us first split $L_1^h(x,t)$ into two pieces:

$$\begin{aligned} L_{11}^h(x,t) &= \sum_{k=1}^N \mathbf{1}_{I_k}(t) \frac{a^* D^- G_0^h(x,x_k)}{h} \cdot \frac{a^* D^- U_k^0}{h}, \\ L_{12}^h(x,t) &= \sum_{k=1}^N \mathbf{1}_{I_k}(t) a^* D^- G_0^h(x,x_k) \left(\frac{1}{h} \int_t^{x_k} f(s) ds - \frac{1}{h} \int_{x_{k-1}}^{x_k} f(s) \tilde{\phi}_0^k(s) ds \right). \end{aligned} \tag{7.72}$$

Then define $\mathcal{I}_i^h(x)$ by

$$\mathcal{I}_i^h(x;W) = \sigma \int_0^1 L_{1i}^h(x,t) dW_t, \ \ i=1,2. \quad \mathcal{I}_3^h(x;W) = \sigma \int_0^1 L_2^h(x,t) dW_t. \tag{7.73}$$

As it turns out, $\mathcal{I}_1^h(x;W)$ converges to the desired limit, while $\mathcal{I}_2^h(x;W)$ and $\mathcal{I}_3^h(x;W)$ converge to zero in probability.

Proof of (7.70). **Convergence of** $\{\mathcal{I}_1^h(x)\}$. By Proposition 2.36, we show the convergence of finite distributions of $\{\mathcal{I}_1^h(x)\}$ and tightness. Since all processes involved are Gaussian, for finite dimensional distribution it suffices to consider the covariance function $R_1(x,y) :=$

$\mathbb{E}\{\mathcal{I}_1^h(x)\mathcal{I}_1^h(y)\}$. By the Itô isometry of Wiener integrals, we have

$$R_1^h(x,y) = \sigma^2 \int_0^1 L_{11}^h(x,t) L_{11}^h(y,t) dt.$$

For any fixed x, $L_{11}^h(x,t)$, as a function of t, is a piecewise constant approximation of $L(x,t)$. This is obvious from the expression of $L(x,t)$ in (7.30). Therefore, $L_{11}^h(x,t)$ converges to $L(x,t)$ in (7.29) pointwise in t. Meanwhile, L_{11}^h is uniformly bounded as well. The dominant convergence theorem yields that for any x and y,

$$\lim_{h\to 0} R_1^h(x,y) = \sigma^2 \int_0^1 L(x,t)L(y,t)dt = \mathbb{E}(\mathcal{U}(x;W)\mathcal{U}(y;W)). \tag{7.74}$$

This proves convergence of finite dimensional distributions.

The heart of the matter is to show that $\{\mathcal{U}_1^h(x;W)\}$ is a tight sequence. To this end, we consider its fourth moment

$$\mathbb{E}\Big(\mathcal{I}_1^h(x) - \mathcal{I}_1^h(y)\Big)^4 = \int_{[0,1]^4} \prod_{i=1}^4 (L_{11}^h(x,t_i) - L_{11}^h(y,t_i)) \mathbb{E} \prod_{i=1}^4 dW_{t_i}. \tag{7.75}$$

Since increments in a Brownian motion are independent Gaussian random variables, we have

$$\mathbb{E}\prod_{i=1}^4 dW_{t_i} = [\delta(t_1-t_2)\delta(t_3-t_4) + \delta(t_1-t_3)\delta(t_2-t_4) + \delta(t_1-t_4)\delta(t_2-t_3)] \prod_{i=1}^4 dt_i. \tag{7.76}$$

Using this decomposition, and the fact that the L_{11}^h is piecewise constant, we rewrite the fourth moment above as

$$3\left(\int_0^1 (L_{11}^h(x,t) - L_{11}^h(y,t))^2 dt\right)^2 = 3\left[\sum_{k=1}^N \left(\frac{a^* D^-(G_0^h(x,x_k) - G_0^h(y,x_k))}{h} \frac{a^* D^- U_k^0}{h}\right)^2 h\right]^2.$$

Hence, we need to control $\|L_{11}^h(x,\cdot) - L_{11}^h(y,\cdot)\|_2$. Since G_0^h is the Green's function associated to (7.2), as commented in Remark 7.12, it admits expression (7.31). Fix $y < x$, and let j_1 and j_2 be the indices so that $y \in (x_{j_1-1}, x_{j-1}]$ and $x \in (x_{j_2-1}, x_{j_2}]$. Then we can split the above sum into three parts. In the first part, k runs from one to $j_1 - 1$. In that case, both x_k and x_{k-1} are less than y. Formula (7.31) says: $a^*(G_0^h(x, x_k) - G_0^h(y, x_k)) = x_k(y - x)$. Consequently,

$$\frac{a^* D^-(G_0^h(x, x_k) - G_0^h(y, x_k))}{h} = (y - x). \tag{7.77}$$

Since $|D^- U_k^0/h|$ is bounded, we have

$$\sum_{k=1}^{j_1-1} \left(\frac{a^* D^-(G_0^h(x, x_k) - G_0^h(y, x_k))}{h} \right)^2 \left(\frac{a^* D^- U_k^0}{h} \right)^2 h \leq C|x - y|^2 \sum_{k=1}^{j_1-1} h \leq C|x - y|^2. \tag{7.78}$$

Another part is k running from $j_2 + 1$ to N. In that case, both x_k and x_{k-1} are larger than x. The above analysis yields the same bound for this partial sum.

The remaining part is when k runs from j_1 to j_2. In this case, for some k, x_k may end up in (y, x), and we have to use different branches of (7.31) when evaluating $G_0^h(x, x_k)$ and $G_0^h(y, x_k)$. Consequently, the cancellation of h in (7.77) will not happen, and we need to modify our analysis. We observe that, due to the Lipschitz continuity of G_0^h and boundedness of $|D^- U^0/h|$, we always have

$$\sum_{k=j_1}^{j_2} \left(\frac{a^* D^-(G_0^h(x, x_k) - G_0^h(y, x_k))}{h} \right)^2 \left(\frac{a^* D^- U_k^0}{h} \right)^2 \cdot h \leq C \frac{|x - y|^2}{h^2} \sum_{k=j_1}^{j_2} h. \tag{7.79}$$

If $j_2 - j_1 \leq 1$, the last sum above is then bounded by $2C|x - y|^2/h$. In this case, it is clear that $|x - y| \leq 2h$; as a result, the sum above is bounded by $C|x - y|$.

If $j_2 - j_1 \geq 2$, the above estimate will not help much if $j_2 - j_1$ is very large. Nevertheless, since $|D^- G_0^h/h|$ is bounded by some universal constant C. We have

$$\sum_{k=j_1}^{j_2} \left(\frac{a^* D^-(G_0^h(x, x_k) - G_0^h(y, x_k))}{h} \right)^2 \left(\frac{a^* D^- U_k^0}{h} \right)^2 \cdot h \leq C \sum_{k=j_1}^{j_2} h. \tag{7.80}$$

Meanwhile, we observe that in this case

$$\begin{aligned}3|x-y| &\geq 3(x_{j_2-1}-x_{j_1}) = 3(j_2-j_1-1)h = (j_2-j_1+1)h + 2(j_2-j_1-2)h\\ &\geq (j_2-j_1+1)h.\end{aligned}$$

Consequently, the sum in (7.80) is again bounded by $C|x-y|$. Combining these estimates, we have

$$\|L^h_{11}(x,\cdot) - L^h_{11}(y,\cdot)\|_2^2 \leq C|x-y|. \tag{7.81}$$

It follows from the equation below (7.76) that $\{\mathcal{I}^h_1(x)\}$ is a tight sequence and hence converges to $\mathcal{U}(x,W)$.

Convergence of $\mathcal{I}^h_{12}$ to zero function. For the finite dimensional distributions, we consider the covariance function $R^h_2(x,y) = \mathbb{E}\{\mathcal{I}^h_2(x)\mathcal{I}^h_2(y)\}$. By Itô isometry,

$$\sigma^2 \int_0^1 L^h_{12}(x,t)L^h_{12}(y,t)dt. \tag{7.82}$$

Now from the expression of $L^h_{12}(x,t)$, (7.72), we see that $L^h_{12}(x,t)$ converges to zero point-wise in t for any fixed x. Indeed, in the above expression, $|D^-G^h_0/h|$ is uniformly bounded while the integrals of $f(s)$ and of $f(s)\tilde{\phi}^k_0(s)$ go to zero due to shrinking integration regions. Meanwhile, L^h_{12} is also uniformly bounded. The dominated convergence theorem shows $R_2(x,y) \to 0$ for any x and y, proving the convergence of finite dimensional distributions. The tightness of $\{\mathcal{I}^h_2(x)\}$ is exactly the same as $\{\mathcal{I}^h_1(x)\}$; that is to say, the properties of $D^-G^h_0$ can still be applied. We conclude that $\{\mathcal{I}^h_2(x)\}$ converges to zero.

Convergence of $\mathcal{I}^h_3(x)$ to zero. For the finite dimensional distributions, we observe that $L^h_2(x,t)$ is uniformly bounded and for any fixed x, it converges to zero point-wise in t, due to shrinking of the non-zero interval $I_{j(x)}$. The covariance function of $\mathcal{I}^h_3(x)$, therefore, converges to zero, proving convergence of finite dimensional distributions.

For tightness, we consider the fourth moment of $\mathcal{I}_3^h(x) - \mathcal{I}_3^h(y)$. By (7.76), it equals

$$\mathbb{E}(\mathcal{I}_3^h(x;W) - \mathcal{I}_3^h(y;W))^4 = 3\left(\int_0^1 (L_2^h(x,t) - L_2^h(y,t))^2 dt\right)^2. \tag{7.83}$$

Recalling the expression of $L_2^h(x,t)$ in (7.50), it is non-zero only on an interval of size h and is uniformly bounded. Let $j(x)$ be the interval where $L_2^h(x)$ is non-zero, and similarly define $j(y)$. Assume $y < x$ without loss of generality. Consider three cases: $j(x) = j(y)$, $j(y) = j(x) - 1$, and $j(x) - j(y) \geq 2$. In the first case, x and y fall in the same interval $[x_{j-1}, x_j]$ for some index j. Then we have

$$\int_0^1 (L_2^h(x,t) - L_2^h(y,t))^2 dt \leq C\int_0^1 \left(\mathbf{1}_{[x,y]}(t) - \frac{x-y}{h}\mathbf{1}_{I_j}(t)\right)^2 dt.$$

This integral can be calculated explicitly; it equals:

$$\begin{aligned}&\int_0^1 \mathbf{1}_{[x,y]}(t) - 2\frac{x-y}{h}\mathbf{1}_{[x,y]} + \frac{(x-y)^2}{h^2}\mathbf{1}_{I_j}(t)dt\\ =&(x-y) - 2\frac{x-y}{h}(x-y) + \frac{(x-y)^2}{h^2}h = (x-y)[1 - \frac{x-y}{h}].\end{aligned}$$

Since $|1 - (x-y)/h| \leq 1$ and $|D^- U_k^0/h| \leq C$, the above quantity is bounded by $C|x-y|$.

In the second case, with j the unique index so that $y \leq x_j < x$ and using the triangle inequality, we have

$$\|L_2^h(x,t) - L_2^h(y,t)\|_2^2 \leq 2\left(\|L_2^h(x,t) - L_2^h(x_j,t)\|_2^2 + \|L_2^h(x_j,t) - L_2^h(y,t)\|_2^2\right).$$

For the first term of the right hand side above, let x_j play the role of y in the previous calculation. This term is bounded by $C(x - x_j)$. Similarly, for the second term, let x_j play the role of x, and we bound this term by $C(x_j - y)$. Consequently, we can still bound $\|L_2^h(x,\cdot) - L_2^h(y,t)\|_2^2$ by $C|x-y|$.

In the third case, we have $h \leq |x-y|$. Meanwhile, since L_2^h is uniformly bounded and is nonzero only on intervals of size h. We have

$$\|L_2^h(x,t) - L_2^h(y,t)\|_2^2 \le Ch \le C|x-y|.$$

Combining these three cases, the conclusion is:

$$\mathbb{E}(\mathcal{I}_3^h(x;W) - \mathcal{I}_3^h(y;W))^4 \le C|x-y|^2. \tag{7.84}$$

This proves tightness and completes proof of the first item of Proposition 7.17. □

Remark 7.18. In the proof above, we used the fact that $G_0^h(x)$ defined in (7.51) is in fact the real Green's function defined in (7.31). However, the analysis follows as long as $|D_k^- G_0^h(x,x_k)/h|$ is piecewise Lipschitz in x with constant independent of h, and the total number of pieces does not depend on h.

The fact that $\mathcal{I}_2^h(x)$ and $\mathcal{I}_3^h(x)$ do not contribute to the limit is quite remarkable. It says the following. As long as the limiting distribution of the corrector $u_\varepsilon^h - u_0^h$ is considered, the role of the multi-scale basis functions is mainly to construct the stiffness matrix, which is reflected by $\mathcal{I}_1^h(x)$; its roles in constructing the load vector F^ε and in assembling the global function, which are reflected in $\mathcal{I}_2^h(x)$ and $\mathcal{I}_3^h(x)$ respectively, are asymptotically not important. □

Now, we prove the second part of Proposition 7.17. The reader should read preliminary material on fractional Brownian motion in section 2.5 of Chapter 2.

Proof of (7.71). Recall that $\mathcal{I}_H^h(x)$ denotes the left hand side of (7.71). Using the same splitting of L_1^h in (7.72), we can split $\mathcal{I}_H^h$ into three pieces $\mathcal{I}_{Hi}^h(x)$, $i = 1,2,3$, as in (7.73). The only necessary modification is to replace σ by σ_H and to replace the Brownian motion W_t by the fBm W_t^H. We show that $\mathcal{I}_{H1}^h(x)$ converges to $\mathcal{U}_H$ while $\mathcal{I}_{H2}^h$ and $\mathcal{I}_{H3}^h$ converge to the zero function.

Convergence of finite dimensional distributions. For $\mathcal{I}_{H1}^h$, we consider the covariance matrix $R^h_{H1}(x,y)$ defined by $\mathbb{E}\{\mathcal{I}^h_{H1}(x)\mathcal{I}^h_{H1}(y)\}$. Using the isometry (2.41), we have

$$R^h_{H1}(x,y) = \kappa \int_0^1 \int_0^1 \frac{L^h_{11}(x,t)L^h_{11}(y,s)}{|t-s|^\alpha}\, dtds. \tag{7.85}$$

As before, the integrand in the above integral converges to $L(x,t)L(y,s)/|t-s|^\alpha$ for almost every (t,s). Meanwhile, since L^h_{11} is uniformly bounded, the integrand above is bounded by $C|t-s|^{-\alpha}$ which is integrable. The dominated convergence theorem then implies that R^h_{H1} converges to the covariance function of $\mathcal{U}_H(x;W^H)$. The convergence of finite distributions of $\mathcal{I}^h_{H2}$ and $\mathcal{I}^h_{H3}$ are similarly proved.

Tightness. Due to the long range correlations, we only need to consider the second moments in 2.49. For $\{\mathcal{I}^h_{H1}\}$, we consider

$$\mathbb{E}(\mathcal{I}^h_{H1}(x) - \mathcal{I}^h_{H1}(y))^2 = \kappa \int_{\mathbb{R}^2} \frac{(L^h_{11}(x,t) - L^h_{11}(y,t))(L^h_{11}(x,s) - L^h_{11}(y,s))}{|t-s|^\alpha} dtds,$$

using again the isometry (2.41). Now we claim that

$$\|L^h_{11}(x,t) - L^h_{11}(y,t)\|_{L^p_t} \le C|x-y|^{\frac{1}{p}}, \tag{7.86}$$

for any $p \ge 1$. Indeed, for $p=2$, this is shown in (7.81); the analysis there actually shows also that the above holds for $p=1$. For $p=\infty$, this follows from the uniform bound on L^h_{11}. For other p, this follows from interpolation; see [78, p.75].

Now, we apply the Hardy-Littlewood-Sobolev lemma [78, §4.3] to the expression of the second moment above. We obtain the bound

$$C(\alpha)\kappa\|L^h_{11}(x,\cdot) - L^h_{11}(y,\cdot)\|_{L^1}\|L^h_{11}(x,\cdot) - L^h_{11}(y,\cdot)\|_{L^{\frac{1}{1-\alpha}}} \le C|x-y|^{2-\alpha}.$$

Therefore, the Kolmogorov criterion (2.49) holds with $\beta = 2$ and $\delta = 1-\alpha$, proving tightness of $\{\mathcal{I}^h_{H1}\}$. Tightness of $\{\mathcal{I}^h_{H2}\}$ follows in the same way because L^h_{12} has the same structure as L_{11}^h as remarked before. Tightness of $\{\mathcal{I}_{H3}^h\}$ follows from the same argument above and the control on $\|L_2^h(x,\cdot) - L_2^h(y,\cdot)\|_2^2$ in the equation above (7.84). This complete the proof of (7.71). □

7.4.4 Applications to MsFEM in random media

In this section, we prove Theorems 7.3 and 7.5 as an application of the general results obtained in the preceding two sections by verifying that the multiscale finite element method (MsFEM) satisfies the conditions of Proposition 7.15.

Since MsFEM is a scheme that satisfies (N1)-(N3), in order to apply (7.42) and (7.50) in previous propositions, we only need to check that (7.46) and (7.47) hold.

Lemma 7.19. *Let* $\tilde{\phi}_\varepsilon^k$ *and* b_ε^k *be the functions in* (N1)-(N3) *for MsFEM defined in* (7.17). *Let* $\tilde{\phi}_0^k$ *and* b_0^k *be the corresponding functions for FEM.*

(i) *Suppose* $a(x,\omega)$ *and* $q(x,\omega)$ *satisfy* (S1)-(S3). *Then* (7.46) *and* (7.47) *hold and the conclusion of item one in Proposition 7.11 follows.*

(ii) *Suppose* $a(x,\omega)$ *and* $q(x,\omega)$ *satisfy* (L1)-(L2). *Then the conditions and hence the conclusions of the second item of Proposition 7.11 hold.*

Proof. From the explicit formulas (7.17), we have

$$b_\varepsilon^k - b_0^k = \left(\int_{I_k} \frac{1}{a_\varepsilon} dt\right)^{-1} - \left(\int_{I_k} \frac{1}{a^*} dt\right)^{-1} = -b_\varepsilon^k \frac{a^*}{h} \int_{I_k} q_\varepsilon(t) dt.$$

Comparing with the second equation in (7.46), we find that it is satisfied with

$$\tilde{r}_{2k} := b_\varepsilon^k \int_{I_k} q_\varepsilon(t) dt. \tag{7.87}$$

Similarly, we have

$$\tilde{\phi}^k_\varepsilon(x) - \tilde{\phi}^k_0(x) = b^k_\varepsilon \left(\int_{x_{k-1}}^{x} q_\varepsilon(s)ds - \frac{x - x_{k-1}}{h} \int_{x_{k-1}}^{x_k} q_\varepsilon(s)ds \right). \tag{7.88}$$

This shows again that (7.46) holds with $\tilde{r}_{1k}$ having the same expression as $\tilde{r}_{2k}$ defined above. In (7.87), since $0 \le b^k_\varepsilon \le \Lambda h^{-1}$, we can apply Lemma 7.13 in the case of short range media or apply Lemma 7.14 in the case of long range media to conclude that $\mathbb{E}|\tilde{r}_{2k}|^2 \le Ch^{-1}\varepsilon$ in the first setting, while $\mathbb{E}|\tilde{r}_{2k}|^2 \le C(\varepsilon h^{-1})^\alpha$ in the second setting. This completes the proof.□

Note that the estimates (7.40) and (7.41) follow directly from this lemma. Therefore, we can apply Proposition 7.9 directly. Now we prove Theorem 7.3. Estimates (7.25) and (7.26) do not follow from Propositions 7.15 and 7.17 directly and need additional considerations.

Proof of Theorem 7.3. **Finite element analysis**. We have seen that u^h_ε super-converges to u_ε; see Remark 7.2. From (7.1) and (7.16), we observe that the following equation holds on I_j for $j = 1, \cdots, N$:

$$\begin{cases} \mathcal{L}_\varepsilon(u_\varepsilon - u^h_\varepsilon) = f, & \text{in } I_j, \\ u^h_\varepsilon - u_\varepsilon = 0, & \text{on } \partial I_j. \end{cases} \tag{7.89}$$

Using the ellipticity of the diffusion coefficient and integrations by parts, we obtain

$$\begin{aligned} \lambda|u^h_\varepsilon - u_\varepsilon|^2_{H^1,I_j} &\le \int_{I_j} a_\varepsilon \frac{d}{dx}(u^h_\varepsilon - u_\varepsilon) \cdot \frac{d}{dx}(u^h_\varepsilon - u_\varepsilon)\ dx = \int_{I_j} (u^h_\varepsilon - u_\varepsilon)\mathcal{L}_\varepsilon(u^h_\varepsilon - u_\varepsilon)\ dx \\ &= \int_{I_j} f(x)(u^h_\varepsilon - u_\varepsilon)(x)\ dx \le \|f\|_{2,I_j}\|u^h_\varepsilon - u_\varepsilon\|_{2,I_j}. \end{aligned}$$

Now recall that the Poincaré-Friedrichs inequality says that

$$\|u^h_\varepsilon - u_\varepsilon\|_{2,I_j} \le \frac{h}{\pi}|u^h_\varepsilon - u_\varepsilon|_{H^1,I_j}. \tag{7.90}$$

Combining the inequalities above, we obtain

$$|u_\varepsilon^h - u_\varepsilon|_{H^1, I_j} \le \frac{h}{\lambda \pi} \|f\|_{2, I_j}.$$

Taking the sum over j, we obtain the first inequality in (7.25). To get the second inequality, we first apply the Poincare-Friedrichs inequality to the equation above to get

$$\|u_\varepsilon^h - u_\varepsilon\|_{2, I_j} \le \frac{h^2}{\lambda \pi^2} \|f\|_{2, I_j}, \tag{7.91}$$

and then sum over j. This completes the proof of (7.25) in item one of the theorem.

Energy norm of the corrector. By energy norm, we mean the $L^2(\Omega, L^2(D))$ norm. Recall the decomposition of the corrector into $K_i(x)$ in (7.45). For $K_1(x)$, we apply Cauchy-Schwarz to get the following bound for $|K_1|^2$

$$\textstyle\sum_i (U^\varepsilon - U^0)_i^2 \sum_j (\phi_0^j(x))^2 \le \sum_i (U_i^\varepsilon - U_i^0)^2 \left(\sum_j \phi_0^j(x) \right)^2 = \left\| U^\varepsilon - U^0 \right\|_{\ell^2}^2 .$$

In the above derivation, we used the fact that $\phi_0^j(x)$ is non-negative, and $\sum_j \phi_0^j(x) \equiv 1$. Now we apply (7.42) to control this term. The function $K_2(x)$, as in the proof of Proposition 7.11, can be written as $D^- U^0_{j(x)} (\tilde{\phi}_\varepsilon^{j(x)} - \tilde{\phi}_0^{j(x)})$. Then from (7.40), we have $\mathbb{E}|K_2(x)|^2 \le C\varepsilon \|R\|_{1,\mathbb{R}}$. For K_3, we have controlled $\mathbb{E}|K_3(x)|$ in (7.55). To control $\mathbb{E}|K_3(x)|^2$, we observe that $|K_3(x)| \le C\|f\|_2$. Note that all three estimates concluded in the three steps are uniform in x. Combining them, we complete the proof of (7.26).

Convergence in distribution as ε to zero. To prove item two of the theorem, we apply (7.60) of Proposition 7.15. We need to verify (7.39) in addition to (7.46) and (7.47), which we already verified in the previous lemma. But this is implied by (7.18), and hence we obtain (7.27).

Convergence in distribution as h to zero. To prove (7.28), we apply the first result in Proposition 7.17. This completes the proof of the theorem. □

Proof of Theorem 7.5. In this case, the random processes $q(x)$ and $a(x)$ are constructed by (L1)-(L2). To prove the estimate in the energy norm, we follow the same steps as in the proof above, but use item two of Proposition 7.9 to control the term $\|U^\varepsilon - U^0\|_{\ell^2}^2$ in $K_1(x)$ and use Lemma 7.14 to control the terms in $K_2(x)$ and $K_3(x)$.

To obtain the results in (7.33) and (7.34), we verify the conditions in item two of Propositions 7.15 and 7.17, applying the second case in Lemma 7.19 and following the steps in the previous proof. This completes the proof of the theorem. □

7.4.5 Applications to HMM in random media

To prove Theorem 7.7, we apply Proposition 7.11 to write the corrector $u_\varepsilon^{h,\delta} - u_0^h$ as an oscillatory integral plus a lower order term. To apply Propositions 7.15 and 7.17 and obtain the weak convergences, we need to consider the difference $b_\varepsilon^k - b_0^k$ since $\tilde{\phi}_\varepsilon^j = \tilde{\phi}_0^j$ in HMM. From the expression of b_ε^k in (7.22), we have

$$b_\varepsilon^k - b_0^k = -b_\varepsilon^k \frac{a^*}{\delta} \int_{I_k^\delta} q_\varepsilon(t)dt = -(1 + \tilde{r}_{2k}) \frac{h}{\delta} \frac{a^{*2}}{h^2} \int_{I_k^\delta} q_\varepsilon(t)dt,$$

where $\tilde{r}_{2k}$ is a random variable defined by

$$\tilde{r}_{2k} = -\frac{h}{\delta} b_\varepsilon^k \int_{I_k^\delta} q_\varepsilon(t)dt.$$

We verify that in the case of short range media, i.e., when $q(x)$ satisfies (S1)-(S3), we have

$$\mathbb{E}|\tilde{r}_{2k}|^2 \le C\frac{\varepsilon}{\delta}\|R\|_{1,\mathbb{R}}, \quad \mathbb{E}\left(b_\varepsilon^k - b_0^k\right)^2 \le C\frac{\varepsilon}{h^2\delta}\|R\|_{1,\mathbb{R}}, \tag{7.92}$$

for some universal constant C. Comparing this with (7.41) and (7.47), we observe that the estimates have been multiplied by a factor $\frac{h}{\delta}$ in the HMM case. Similarly, it can be checked that in the case of long range media, i.e., when $q(x)$ satisfies (L1)-(L2), these estimates will be multiplied by a factor of $\left(\frac{\delta}{h}\right)^{\alpha}$. With these formulas at hand, we prove the third main theorem of the paper.

Proof of Theorem 7.7. **Short range media and amplification effect**. In this case, the difference of $b_\varepsilon^k - b_0^k$ and an estimate of it was captured in (7.92) and the equation above it. We cannot apply Propositions 7.11 and 7.15 directly. However, as mentioned in Remark 7.16, similar conclusions still hold. The same procedure as in the proof of Proposition 7.11 shows that the $L^h(x,t)$ function for HMM is:

$$L^{h,\delta}(x,t) = \frac{h}{\delta}\sum_{k=1}^{N} \mathbf{1}_{I_k^\delta}(t)\frac{a^* D^- G_0^h(x,x_k)}{h}\frac{a^* D^- U_k^0}{h}. \tag{7.93}$$

The first weak convergence in (7.35) holds with this definition of $L^{h,\delta}$ as an application of a modified version of Proposition 7.15. Indeed, the proof there works with $L^{h,\delta}$ playing the role of L_{11}^h. The tightness is still obtained from the function $D^- G_0^h$, and the factor $\frac{h}{\delta}$ does not play any role at this stage.

When h goes to zero, we can follow the proof of Proposition 7.17 to verify the second convergence in (7.35). Indeed, tightness can be proved in exactly the same way. All that needs to be modified is the limit of the covariance function of $\mathcal{U}^{h,\delta}(x;W)$, which is defined to be $\sigma\int_0^1 L^{h,\delta}(x,t)dW_t$. This covariance function, by the Itô isometry, is as follows:

$$\begin{aligned} R^{h,\delta}(x,y) :&= \sigma^2\int_0^1 L^{h,\delta}(x,t)L^{h,\delta}(y,t)dt \\ &= \sigma^2\frac{h^2}{\delta^2}\sum_{k=1}^{N}\delta\frac{a^* D^- G_0^h(x;x_k)}{h}\frac{a^* D^- G_0^h(y;x_k)}{h}\left(\frac{a^* D^- U_k^0}{h}\right)^2. \end{aligned} \tag{7.94}$$

Recall the expression of L_{11}^h in (7.72). We verify that the above quantity can be written as

$$\sigma^2\frac{h}{\delta}\int_0^1 L_{11}^h(x,t)L_{11}^h(y,t)dt.$$

Now the convergence in (7.74) implies that $R^{h,\delta}$ converges to the covariance function of

$\sqrt{\frac{h}{\delta}}\mathcal{U}(x;W)$. This completes the proof of (7.35).

Long range media. The expression for $b_\varepsilon^k - b_0^k$ are given above. Therefore, we can apply Propositions 7.11 and 7.15 (with modifications), to show that as ε goes to zero while h is fixed, the HMM corrector indeed converges to $\mathcal{U}_H^{h,\delta}(x;W^H)$ defined in (7.36). When h is sent to zero, we can follow the proof of Proposition 7.17 and show that $\mathcal{U}_H^{h,\delta}$ converges in distribution to some Gaussian process. To find its expression, we calculate the covariance function of $\mathcal{U}_H^{h,\delta}$. Thanks to the isometry (2.41), it is given by

$$R_H^{h,\delta}(x,y) := \kappa \int_0^1 \int_0^1 \frac{L^{h,\delta}(x,t)L^{h,\delta}(y,s)}{|t-s|^\alpha} dtds. \tag{7.95}$$

Using the expression of $L^{h,\delta}$, and the following short-hand notations:

$$J_k(x) := \frac{a^* D^- G_0^h(x;x_k)}{h} \frac{a^* D^- U_k^0}{h},$$

the covariance function can be written as

$$\kappa \frac{h^2}{\delta^2}\Big(\sum_{k=1}^N \sum_{m=1}^k [J_k(x)J_m(y) + J_m(x)J_k(y)] \int_{I_k^\delta}\int_{I_m^\delta} \frac{dtds}{|t-s|^\alpha} + \sum_{k=1}^N J_k(x)J_k(y) \int_{I_k^\delta}\int_{I_k^\delta} \frac{dtds}{|t-s|^\alpha}\Big).$$

The integral of $|t-s|^{-\alpha}$ can be evaluated explicitly:

$$\frac{\kappa}{(1-\alpha)(2-\alpha)} \sum_{k=1}^N \sum_{m=1}^{k-1} [J_k(x)J_m(y) + J_m(x)J_k(y)] \frac{h^2}{\delta^2}\Big([(k-m)h+\delta]^{2-\alpha} + \\ -2[(k-m)h]^{2-\alpha} + [(k-m)h-\delta]^{2-\alpha}\Big) + \frac{\kappa}{(1-\alpha)(2-\alpha)} \sum_{k=1}^N 2J_k(x)J_k(y)\frac{h^2}{\delta^2}\delta^{2-\alpha}.$$

When $m < k$, the quantity between parentheses together with the δ^2 on the denominator forms a centered difference approximation of the second order derivative of the function $r^{2-\alpha}$, evaluated at $(k-m)h$, i.e., at $t-s$. This derivative is precisely $(1-\alpha)(2-\alpha)|t-s|^{-\alpha}$. Meanwhile, the h^2 on the nominator can be viewed as the size of the measure $dtds$ on each block $I_k \times I_m$. Furthermore, $J_k(x)$ is precisely $L_{11}^h(x,t)$ evaluated on I_k. The conclusion is: those terms in the above equation with $m < k$ form an approximation of

$$\kappa \int_0^1 \int_0^1 \frac{L_{11}^h(x,t) L_{11}^h(y,s)}{|t-s|^\alpha} \, dtds.$$

The second sum corresponds to the diagonal terms $k = m$. Since $|J_k|$ is bounded, this sum is of order $O(h\delta^{-\alpha})$ and does not contribute in the limit as $h \to 0$, as long as $\delta \gg h^{\frac{1}{\alpha}}$. $R_H^{h,\delta}$ converges to the covariance function of $\mathcal{U}_H(x; W^H)$, finishing the proof of (7.36). □

7.4.6 Numerical experiment

Here, we provide two numerical experiment, which verifies the theory developed above. We apply the MsFEM and HMM described in the previous sections to the random ODE (7.1), and plot the correctors (divided by proper power of ε) in Figure 7.2 and Figure 7.3.

In these experiments, the random field $q(x,\omega)$ in (7.1), that is, the deviation of $1/a(x,\omega)$ from $1/a^*$, is constructed as in (L1). The function Φ there is chosen to be one half of the sign function $\frac{1}{2}$sign, and the underlying Gaussian random process $g(x)$ is chosen as follows.

1. The short range case. We choose $g(x)$ to be the following Ornstein-Uhlenbeck process. That is, $g(x)$ solves

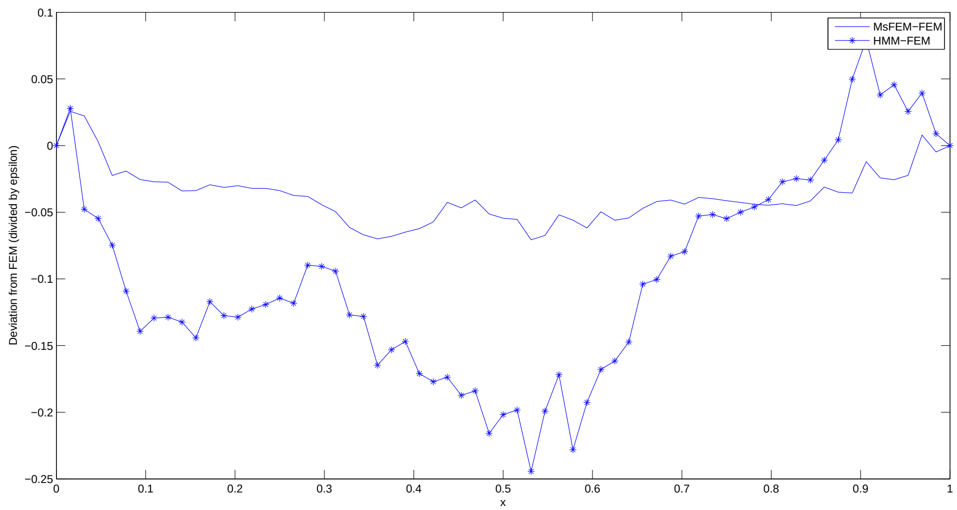
$$dg(x) = -g(x)dx + \sqrt{2} dW(x), \tag{7.96}$$

where $W(x)$ is the standard Brownian motion. $g(0)$ has the standard Gaussian distribution $\mathcal{N}(0,1)$. This process is mean-zero, stationary and has short-range correlation; see e.g. [89].

2. The long range case. We essentially use the increment of a fractional Brownian motion with Hurst index H. To generate a vector of such increments of length N (which can be thought as the total number of steps), we first generate an i.i.d. Gaussian vector of this length, i.e, Brownian motion increments. Then we color this vector by the following

Figure 7.2: Corrector in the MsFEM and HMM solutions of Equation (7.1), I.

The random field $q(x,\omega)$ are constructed by $\frac{1}{2}\text{sign}(g)$; g is the Ornstein-Uhlenbeck process in the first picture, and is the colored one in the second picture.
Upper: $\varepsilon = 2^{-14}, \delta = 2^{-9}, h = 2^{-6}$. Lower: $\varepsilon = 2^{-12}, \delta = 2^{-8}, h = 2^{-6}$.



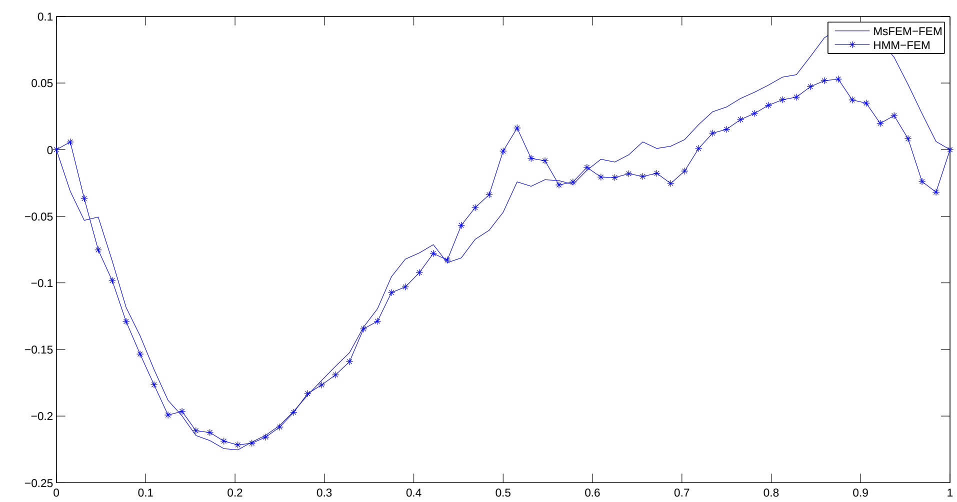
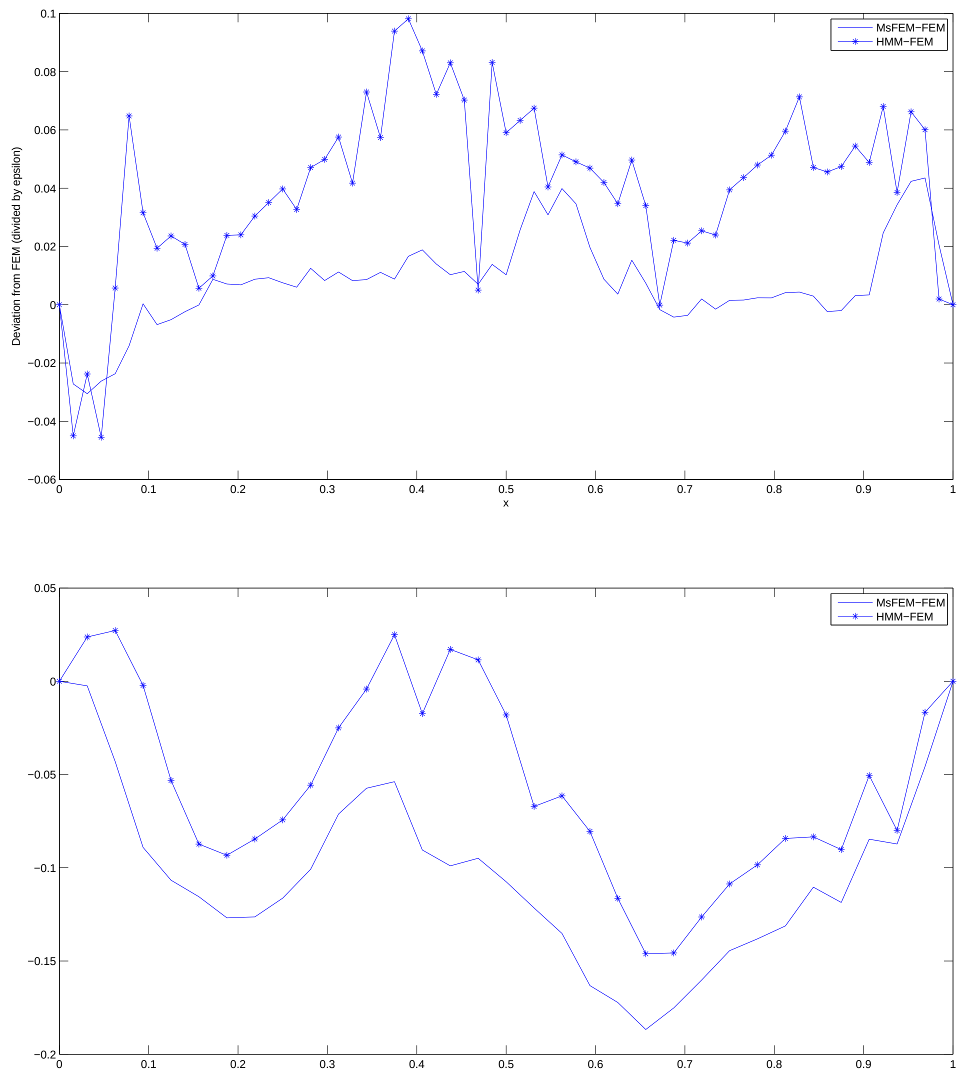


Figure 7.3: Corrector in the MsFEM and HMM solutions of Equation (7.1), II.

The construction of the random field $q(x, \omega)$ is as in Figure 7.2.
Upper: $\varepsilon = 2^{-14}, \delta = 2^{-9}, h = 2^{-6}$. Lower: $\varepsilon = 2^{-12}, \delta = 2^{-8}, h = 2^{-5}$.



covariance function:

$$C(t,s) = \frac{1}{2}\left(|t|^{2H} - 2|t-s|^{2H} + |s|^{2H}\right). \tag{7.97}$$

More precisely, we construct the covariance matrix C with $C_{ij} = C(i,j)$, and we compute the square root of C. Use this square root matrix to color the i.i.d. Gaussian vector.

7.4.7 A hybrid scheme that passes the corrector test

We now present a method that eliminates the amplification effect of HMM exhibited in item one of Theorem 7.7 when the random media has short range correlations. Such an effect arises because the short-range averaging effects occurring on the interval of size h are not properly captured by averaging occurring on an interval of size $\delta < h$.

The main idea is to subdivide the element I_k uniformly into M smaller patches and perform M independent calculations on each of these patches. This is a hybrid method that captures the idea of performing calculations on small intervals of size $\delta \ll h$ to reduce cost as in HMM while preserving the averaging property of MsFEM by solving the elliptic equation on the whole domain.

Let $\delta = h/M$ be the size of the small patch I_k^ℓ for $1 \le \ell \le M$. Define $b_{\varepsilon\ell}^k$ by

$$b_{\varepsilon\ell}^k = \left(\frac{\delta}{h}\right)^2 \left(\int_{I_k^\ell} a_\varepsilon^{-1}(s)ds\right)^{-1}. \tag{7.98}$$

This definition is motivated by (7.22). Given a function w in the space V_0^h, we define its local projection into the space of oscillatory functions in the small patches I_k^ℓ by:

$$\begin{cases} \mathcal{L}_\varepsilon(w_\ell^k) = 0, & x \in \cup_{k=1}^N \cup_{\ell=1}^M I_k^\ell, \\ w_\ell^k = w, & x \in \cup_{k=1}^N \cup_{\ell=1}^M \partial I_k^\ell, \end{cases} \tag{7.99}$$

where w_ℓ^k denotes this local projection. Recall that $\tilde{\phi}_0^k$ is the left piece of the hat basis function. Integrations by parts show that $b_{\varepsilon\ell}^k = A_\varepsilon((\tilde{\phi}_0)_\ell^k, (\tilde{\phi}_0)_\ell^k)$, where A_ε is the bilinear form defined in (7.11). HMM choose one small patch $I_k^{\ell*}$ and uses $A_\varepsilon((\tilde\phi_0)^k_{\ell*}, (\tilde\phi_0)^k_{\ell*}) = b^k_{\varepsilon\ell_*}$ to approximate the value $A_\varepsilon(\tilde\phi^k_0, \tilde\phi^k_0)$. Of course, the scaling h/δ is needed. This scaling factor turns out to amplify the variance as h goes to zero when the random medium has short range correlation.

We modify the method of HMM by constructing b_ε as follows:

$$b^k_\varepsilon := \textstyle\sum_{\ell=1}^M b^k_{\varepsilon\ell}.$$

In other words, we sum the pieces $A_\varepsilon(\tilde\phi^k_0, \tilde\phi^k_0)$ to form the entries of the stiffness matrix. With this definition, we verify that

$$\begin{aligned} b^k_\varepsilon - b^k_0 &= \sum_{\ell=1}^M \left(\frac{\delta}{h}\right)^2 \Big[\Big(\int_{I_k^\ell} a_\varepsilon^{-1} ds\Big)^{-1} - \Big(\int_{I_k^\ell} a^{*-1} ds\Big)^{-1}\Big] \\ &= \sum_{\ell=1}^M \left(\frac{a^*}{h}\right)^2 \left[-\int_{I_k^\ell} q_\varepsilon(s)ds + \Big(\int_{I_k^\ell} q_\varepsilon(s)ds\Big)^2 \Big(\int_{I_k^\ell} a_\varepsilon^{-1} ds\Big)^{-1}\right] \end{aligned}$$

Rewriting the sum of the first terms in the parenthesis, we obtain

$$b^k_\varepsilon - b^k_0 = -\left(\frac{a^*}{h}\right)^2 \int_{I_k} q_\varepsilon(s)ds + r^k_\varepsilon,$$

where r^k_ε accounts for the sum over the second terms in the parenthesis. Clearly, $\mathbb{E}|r^k_\varepsilon| \le C\varepsilon(h\delta)^{-1}$. This decomposition of $b_\varepsilon - b_0$ and the estimate of r^k_ε shows that we can apply Proposition 7.11 to obtain the decomposition of the corrector. The $L^h(x,t)$ function in this case will be $L^h_{11}(x,t)$ in (7.72). Then it follows from Propositions 7.15 and 7.17 that the corrector in this method converges to the right limit.

In this modified method, all the local informations on I_k are used to construct b^k_ε as in MsFEM. The main advantage is that the computation on $\{I_k^\ell\}_{\ell=1}^M$ can be done in a parallel manner. The calculation in MsFEM performed on a whole domain of size h is replaced by h/δ independent calculations. Accounting for the coupling between the h/δ subdomains is necessary in MsFEM. It is no longer necessary in the modified method, which significantly reduces is complexity.

7.5 Appendix: Moment Bound for Stochastic Process

In this section we provide a bound for the fourth order moment of $q(x,\omega)$ in terms of the L^1 norm of the ρ-mixing coefficient.

Let $\mathcal{P}$ be the set of all ways of choosing pairs of points in $\{1,2,3,4\}$, i.e.,

$$\mathcal{P} := \Big\{ p = \big\{\{p(1),p(2)\},\{p(3),p(4)\}\big\} \;\Big|\; p(i) \in \{1,2,3,4\} \Big\}. \tag{7.100}$$

There are $C_6^2 = 15$ elements in $\mathcal{P}$.

Lemma 7.20. *Let $q(x,\omega)$ be a stationary mean-zero stochastic process. Assume $\mathbb{E}|q(0)|^4$ is finite and $q(x,\omega)$ is ρ-mixing with mixing coefficient $\rho(r)$ that is decreasing in r. Then we have*

$$|\mathbb{E}\{\prod_{i=1}^4 q(x_i)\}| \le \mathbb{E}|q(0)|^4 \sum_{p\in\mathcal{P}} \rho^{\frac12}(|x_{p(1)} - x_{p(2)}|)\rho^{\frac12}(|x_{p(3)} - x_{p(4)}|). \tag{7.101}$$

Proof. Given four points $\{q(x_i)\}$, $i = 1,\cdots,4$, we can draw six line segments joining them. Among these line segment there is one that has the shortest length. Rearranging the indices if necessary, we assume it is the one joining x_1 and x_2. Then set $A = \{x_1, x_2\}$ and $B = \{x_3, x_4\}$. Rearranging the indices among each set if necessary, we assume also that $d(A,B)$ is obtained by $|x_1 - x_3|$. Then by the definition of ρ-mixing, we have

$$|\mathbb{E}\{\prod_{i=1}^4 q(x_i)\} - R(x_1-x_2)R(x_3-x_4)| \le \mathrm{Var}\{q(x_1)q(x_2)\}^{\frac12}\mathrm{Var}\{q(x_3)q(x_4)\}^{\frac12}\rho(|x_1-x_3|). \tag{7.102}$$

We can bound $\mathrm{Var}\{q(x_1)q(x_2)\}$, and similarly the variance of $\mathrm{Var}\{q(x_3)q(x_4)\}$, from above by $(\mathbb{E}|q(x_1)|^4\mathbb{E}|q(x_2)|^4)^{1/2}$. Therefore, the above term is bounded by $\mathbb{E}|q(0)|^4\rho(|x_1 - x_3|)$.

Since ρ is decreasing and $|x_1 - x_3| \geq |x_1 - x_2|$, we also have

$$|\mathbb{E}\{\prod_{i=1}^4 q(x_i)\} - R(x_1 - x_2)R(x_3 - x_4)| \leq \mathbb{E}|q(0)|^4\rho(|x_1 - x_2|). \tag{7.103}$$

Now observe that $\min\{a, b\} \leq (ab)^{\frac{1}{2}}$ for any two non-negative real numbers a and b. Applying this observation to the bounds of the two inequalities above, and using the triangle inequality, we obtain

$$|\mathbb{E}\{\prod_{i=1}^4 q(x_i)\}| \leq |R(x_1 - x_2)| \cdot |R(x_3 - x_4)| + \mathbb{E}|q(0)|^4\rho^{\frac{1}{2}}(|x_1 - x_2|)\rho^{\frac{1}{2}}(|x_1 - x_3|). \tag{7.104}$$

Using the definition of mixing again, we obtain

$$|R(x_1-x_2)| = |\mathbb{E}q(x_1)q(x_2)| \leq \mathrm{Var}^{\frac{1}{2}}(q(x_1))\mathrm{Var}^{\frac{1}{2}}(q(x_2))\rho(|x_1-x_2|) \leq (\mathbb{E}|q(0)|^4)^{\frac{1}{2}}\rho^{\frac{1}{2}}(|x_1-x_2|).$$

In the last step, we used the fact that $\rho \leq \rho^{\frac{1}{2}}$ since ρ can always be chosen no larger than 1. We can bound $R(x_1 - x_3)$ in the same way. Therefore, we obtain

$$|\mathbb{E}\{\prod_{i=1}^4 q(x_i)\}| \leq \mathbb{E}|q(0)|^4\Big[\rho^{\frac{1}{2}}(|x_1 - x_2|)\rho^{\frac{1}{2}}(|x_3 - x_4|) + \rho^{\frac{1}{2}}(|x_1 - x_2|)\rho^{\frac{1}{2}}(|x_1 - x_3|)\Big]. \tag{7.105}$$

This completes the proof. □ We now derive a bound for the fourth order moment of oscillatory integrals of q_ε.

Lemma 7.21. *Let $q(x, \omega)$ satisfy the conditions in the previous lemma. Assume in addition that the mixing coefficient satisfies that $\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}$ is finite. Let (x, y) be an interval in $\mathbb{R}$. Then for any bounded function $m(t)$, we have*

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_x^y q(\frac{t}{\varepsilon})m(t)dt\right)^4 \leq 60\mathbb{E}|q(0)|^4 \cdot \|\rho^{\frac{1}{2}}\|^2_{1,\mathbb{R}_+}\|m\|^4_\infty \cdot |x - y|^2. \tag{7.106}$$

Proof. The left hand side of the desired inequality is

$$I = \frac{1}{\varepsilon^2}\int_x^y\int_x^y\int_x^y\int_x^y \mathbb{E}\prod_{i=1}^4 q(\frac{t_i}{\varepsilon})\prod_{i=1}^4 m(t_i)d[t_1t_2t_3t_4]. \tag{7.107}$$

Here and below, $d[t_1\cdot t_4]$ is a short-hand notation for $dt_1\cdots d_4$. Apply the preceding lemma, we have

$$I \le \frac{\mathbb{E}|q(0)|^4\|m\|_\infty^4}{\varepsilon^2}\sum_{p\in\mathcal{P}}\int_x^y\int_x^y\int_x^y\int_x^y \rho^{\frac{1}{2}}(\frac{t_{p(1)}-t_{p(2)}}{\varepsilon})\rho^{\frac{1}{2}}(\frac{t_{p(3)}-t_{p(4)}}{\varepsilon})d[t_{p(1)}\cdots t_{p(4)}].$$

Note that we did not write absolute sign for the argument in the ρ functions. We assume ρ is extended to be defined on the whole $\mathbb{R}$ by letting $\rho(x)=\rho(|x|)$. There are 15 terms in the sum above that are estimated in the same manners. Let us look at one of them, with $p(1)=p(3)=1$, $p(2)=2$, and $p(4)=3$. We perform the following change of variables:

$$\frac{t_1-t_2}{\varepsilon}\to t_2,\quad \frac{t_1-t_3}{\varepsilon}\to t_3,\quad t_1\to t_1,\quad t_4\to t_4.$$

The Jacobian resulting from this change of variable cancels ε^2 on the denominator. The integral becomes

$$\int_x^y dt_1\int_x^y dt_4\int_{\frac{t_1-y}{\varepsilon}}^{\frac{t_1-x}{\varepsilon}}\rho^{\frac{1}{2}}(t_2)dt_2\int_{\frac{t_1-y}{\varepsilon}}^{\frac{t_1-x}{\varepsilon}}\rho^{\frac{1}{2}}(t_3)dt_3. \tag{7.108}$$

This integral is finite and is bounded from above by

$$|x-y|^2\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}}^2 = 4|x-y|^2\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}^2. \tag{7.109}$$

The other terms in the sum have the same bound. Hence we have,

$$I \le \mathbb{E}|q(0)|^4\times 15\times 4|x-y|^2\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}^2\|m\|_\infty^4. \tag{7.110}$$

This verifies (7.106) and completes the proof. □

Notations and Background

1. Euclidean Spaces

1. $\mathbb{N}$ denotes the set of natural numbers $0, 1, 2, \cdots$; $\mathbb{Z}$ and $\mathbb{Q}$ denote the set of all integers and all rational numbers respectively; $\mathbb{R}$ and $\mathbb{C}$ denote the set of all real and complex numbers.

2. $\mathbb{R}^d = d$-dimensional real Euclidean space, $\mathbb{R} = \mathbb{R}^1$.

3. A typical point in $\mathbb{R}^d$ is denoted as $x = (x_1, \cdots, x_d)$ where x_i is the i-th coordinate of x.

4. When it does not cause confusion, $(x_1, \cdots, x_n)$ denotes also an n-tuple in $(\mathbb{R}^d)^n$. Here, each $x_j \in \mathbb{R}^d$, for $j = 1, \cdots, n$.

5. $-x = (-x_1, \cdots, -x_d)$ is the symmetric point of x with respect to the origin.

6. $\mathbb{R}^d_+ = \{x = (x_1, \cdots, x_d) \in \mathbb{R}^d \mid x_d > 0\}$ = open upper half-space. $\mathbb{R}_+ = \mathbb{R}^1_+$.

7. X and other capital Latin letters very often denote open sets in $\mathbb{R}^d$. $\overline{X}$ denotes the closure of X; X^c denotes the complement of X.

8. ∂X = boundary of X, that is, $X \bigcap \overline{X}$.

9. For a set $V \subset X$, we write $X \backslash V$ to denote the relative complement of V in X, that is, $X \bigcap V^c$.

10. For a point $x \in \mathbb{R}^d$, $X + x$ is the set obtained by translating the points in X uniformly according to x, that is, $\{y + x \in \mathbb{R}^d \mid y \in X\}$.

11. $B_r = \{y \in \mathbb{R}^d \mid |y| < r\}$ = open ball in $\mathbb{R}^d$ centered at 0, the origin, with radius $r > 0$. $B_r(x) = B_r + x$.

12. π_d = volume of the unit ball B_1 in $\mathbb{R}^d$. $\varpi_d = d\pi_d$ = volume of the unit sphere ∂B_1 in $\mathbb{R}^d$.

13. $|x| = \sqrt{x_1^2 + \cdots + x_d^2}$ is the standard Euclidean norm of x. For a set X, $|X|$ = volume of X with respect to the standard Lebesgue measure.

14. For two points x, y in $\mathbb{R}^d$, $d(x, y)$ denotes the usual distance between them.

2. Notations of Functions

1. We adopt the notation of functions of Evans [57]. In particular, for a real valued function, Du or ∇u is its gradient, and $D^\alpha u$ for a multiindex α is the corresponding higher order derivative. The *Laplacian* of u is $\Delta u = \nabla \cdot \nabla u$.

2. We denote by $\mathcal{S}$ the space of smooth functions that decay faster than any polynomials at the infinity. $\mathscr{D}'$ denots its dual space.

3. A real valued function $f(x) : \mathbb{R}^d \to \mathbb{R}$ is said to be homogeneous with degree k if $f(\lambda x) = \lambda^k f(x)$ for any x and $\lambda > 0$.

4. The Fourier transform of a $\mathbb{C}$-valued function f in $\mathcal{S}$ is denoted by $\widehat{f}$ and is defined by
$$\widehat{f}(\xi) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) dx.$$
This definition works for functions in $L^1(\mathbb{R}^d)$. It also works on $L^2(\mathbb{R}^n)$ since Fourier transform is an isometry on this space. In general, one can define Fourier transform of distributions in $\mathscr{D}'$ by duality.

3. Analysis of Functions

1. **Theorem A** (Lax-Milgram Lemma). *Assume that $B[u,v]$ is a bilinear form on $H \times H$ for some Hilbert space H, satisfying*

(i) $$|B[u,v]| \leq \alpha \|u\| \|v\|, \quad \forall u, v \in H,$$

(ii) $$\beta \|u\|^2 \leq B[u,u], \quad \forall u \in H.$$

Let f be a bounded linear functional on H. Then there exists a unique $u \in H$ such that

$$B[u,v] = f(v), \qquad \forall v \in H.$$

2. **Theorem B** (Hardy-Littlewood-Sobolev Lemma) *Let $p, r > 1$ and $0 < \lambda < d$ with $1/p + \lambda/d + 1/r = 2$. Let $f \in L^p(\mathbb{R}^d)$ and $h \in L^r(\mathbb{R}^d)$. Then there exists a sharp constant $C(d, \lambda, p)$, independent of f and h, such that*

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{f(x)\ h(y)}{|x-y|^\lambda}\ dxdy \right| \leq C(d,\lambda,p) \|f\|_{L^p} \|h\|_{L^r}.$$

The sharp constant satisfies

$$C(d,\lambda,p) \leq \frac{d}{d-\lambda} \left(\frac{\pi_d}{d}\right)^{\frac{\lambda}{d}} \frac{1}{pr} \left(\left(\frac{\lambda/d}{1-1/p}\right)^{\frac{\lambda}{d}} + \left(\frac{\lambda/d}{1-1/r}\right)^{\frac{\lambda}{d}} \right).$$

Bibliography

[1] G. Allaire, *Homogenization and two-scale convergence*, SIAM J. Math. Anal., 23 (1992), pp. 1482–1518.

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

[3] G. Allaire and R. Brizzi, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul., 4 (2005), pp. 790–812 (electronic).

[4] T. Arbogast, *Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems*, SIAM J. Numer. Anal., 42 (2004), pp. 576–598.

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

[6] F. Aurenhammer and R. Klein, *Voronoi diagrams*, in Handbook of computational geometry, North-Holland, Amsterdam, 2000, pp. 201–290.

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

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

[9] ——, *Central limits and homogenization in random media*, Multiscale Model. Simul., 7 (2008), pp. 677–702.

[10] ——, *Inverse transport theory and applications*, Inverse Problems, 25 (2009), pp. 053001, 48.

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

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

[13] ——, *Corrector theory for elliptic equations in random media with singular greens function. application to random boundaries*, Commun. Math. Sci., 9 (2011), pp. 383–411.

[14] ——, *Fluctuation theory for radiative transfer in random media*, Journal of Quantitative Spectroscopy & Radiative Transfer, 112 (2011), pp. 660–670.

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

[16] ——, *Time-dependent angularly averaged inverse transport*, Inverse Problems, 25 (2009), pp. 075010, 32.

[17] G. Bal and K. Ren, *Transport-based imaging in random media*, SIAM J. Appl. Math., 68 (2008), pp. 1738–1762.

[18] ——, *Physics-based models for measurement correlations: application to an inverse Sturm-Liouville problem*, Inverse Problems, 25 (2009), pp. 055006, 13.

[19] C. Bardos, J. Garnier, and G. Papanicolaou, *Identification of Green's functions singularities by cross correlation of noisy signals*, Inverse Problems, 24 (2008), pp. 015011, 26.

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

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

[22] A. M. Berezhkovskii, Y. A. Makhnovskii, M. I. Monine, V. Y. Zitserman, and S. Y. Shvartsman, *Boundary homogenization for trapping by patchy surfaces*, J. Chem. Phys., 121 (2004).

[23] A. M. Berezhkovskii, M. I. Monine, C. B. Muratov, and S. Y. Shvartsman, *Homogenization of boundary conditions for surfaces with regular arrays of traps*, J. Chem. Phys., 124 (2006).

[24] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[25] K. Bogdan and T. Byczkowski, *Potential theory for the α-stable Schrödinger operator on bounded Lipschitz domains*, Studia Math., 133 (1999), pp. 53–92.

[26] K. Bogdan, T. Byczkowski, T. Kulczycki, M. Ryznar, R. Song, and Z. Vondraček, *Potential analysis of stable processes and its extensions*, vol. 1980 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2009. Edited by Piotr Graczyk and Andrzej Stos.

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

[28] 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), pp. 303–315.

[29] L. Breiman, *Probability*, vol. 7 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.

[30] L. A. Caffarelli and P. E. Souganidis, *Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media*, Invent. Math., 180 (2010), pp. 301–360.

[31] L. A. Caffarelli, P. E. Souganidis, and L. Wang, *Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media*, Comm. Pure Appl. Math., 58 (2005), pp. 319–361.

[32] S. Chandrasekhar, *Radiative transfer*, Dover Publications Inc., New York, 1960.

[33] Z. Chen and T. Y. Savchuk, *Analysis of the multiscale finite element method for nonlinear and random homogenization problems*, SIAM J. Numer. Anal., 46 (2007/08), pp. 260–279.

[34] Z.-Q. Chen and R. Song, *Intrinsic ultracontractivity and conditional gauge for symmetric stable processes*, J. Funct. Anal., 150 (1997), pp. 204–239.

[35] ——, *Estimates on Green functions and Poisson kernels for symmetric stable processes*, Math. Ann., 312 (1998), pp. 465–501.

[36] ——, *Two-sided eigenvalue estimates for subordinate processes in domains*, J. Funct. Anal., 226 (2005), pp. 90–113.

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

[38] K. L. Chung, *A course in probability theory*, Academic Press Inc., San Diego, CA, third ed., 2001.

[39] K. L. Chung and Z. X. Zhao, *From Brownian motion to Schrödinger's equation*, vol. 312 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1995.

[40] R. Cont and P. Tankov, *Financial modelling with jump processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.

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

[42] M. Cranston, E. Fabes, and Z. Zhao, *Potential theory for the Schrödinger equation*, Bull. Amer. Math. Soc. (N.S.), 15 (1986), pp. 213–216.

[43] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 5*, Springer-Verlag, Berlin, 1992.

[44] ——, *Mathematical analysis and numerical methods for science and technology. Vol. 6*, Springer-Verlag, Berlin, 1993.

[45] B. Davison and J. B. Sykes, *Neutron transport theory*, Oxford, at the Clarendon Press, 1957.

[46] M. A. Diop, B. Iftimie, E. Pardoux, and A. L. Piatnitski, *Singular homogenization with stationary in time and periodic in space coefficients*, J. Funct. Anal., 231 (2006), pp. 1–46.

[47] P. Doukhan, *Mixing*, vol. 85 of Lecture Notes in Statistics, Springer-Verlag, New York, 1994. Properties and examples.

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

[49] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: a review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.

[50] W. E, P. Ming, and P. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156 (electronic).

[51] Y. Efendiev, *The Multiscale Finite Element Method (MsFEM) and Its Applications*, PhD thesis, 1999.

[52] Y. Efendiev and T. Y. Hou, *Multiscale finite element methods*, vol. 4 of Surveys and Tutorials in the Applied Mathematical Sciences, Springer, New York, 2009. Theory and applications.

[53] Y. R. Efendiev, T. Y. Hou, and X.-H. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.

[54] B. Engquist and P. E. Souganidis, *Asymptotic and numerical homogenization*, Acta Numer., 17 (2008), pp. 147–190.

[55] L. C. Evans, *The perturbed test function method for viscosity solutions of nonlinear PDE*, Proc. Roy. Soc. Edinburgh Sect. A, 111 (1989), pp. 359–375.

[56] ——, *Periodic homogenisation of certain fully nonlinear partial differential equations*, Proc. Roy. Soc. Edinburgh Sect. A, 120 (1992), pp. 245–265.

[57] ——, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1998.

[58] M. J. Field, *Several complex variables and complex manifolds. I*, vol. 65 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1982.

[59] 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), pp. 1069–1077.

[60] P. Garbaczewski and J. R. Klauder, *Schrödinger problem, Lévy processes, and noise in relativistic quantum mechanics*, Phys. Rev. E (3), 51 (1995), pp. 4114–4131.

[61] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[62] F. Golse, P.-L. Lions, B. Perthame, and R. Sentis, *Regularity of the moments of the solution of a transport equation*, J. Funct. Anal., 76 (1988), pp. 110–125.

[63] W. Hansen, *Global comparison of perturbed Green functions*, Math. Ann., 334 (2006), pp. 643–678.

[64] L. Hörmander, *The analysis of linear partial differential operators. III*, Classics in Mathematics, Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[65] T. Y. Hou, *Numerical approximations to multiscale solutions in partial differential equations*, in Frontiers in numerical analysis (Durham, 2002), Universitext, Springer, Berlin, 2003, pp. 241–301.

[66] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.

[67] T. Y. Hou, X.-H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.

[68] J. K. Hunter and B. Nachtergaele, *Applied analysis*, World Scientific Publishing Co. Inc., River Edge, NJ, 2001.

[69] B. Iftimie, É. Pardoux, and A. Piatnitski, *Homogenization of a singular random one-dimensional PDE*, Ann. Inst. Henri Poincaré Probab. Stat., 44 (2008), pp. 519–543.

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

[71] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, vol. 160 of Applied Mathematical Sciences, Springer-Verlag, New York, 2005.

[72] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991.

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

[74] S. M. Kozlov, *The averaging of random operators*, Mat. Sb. (N.S.), 109(151) (1979), pp. 188–202, 327.

[75] H.-H. Kuo, *Introduction to stochastic integration*, Universitext, Springer, New York, 2006.

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

[77] P. Li and S. T. Yau, *On the Schrödinger equation and the eigenvalue problem*, Comm. Math. Phys., 88 (1983), pp. 309–318.

[78] E. H. Lieb and M. Loss, *Analysis*, vol. 14 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2001.

[79] P.-L. Lions and P. E. Souganidis, *Homogenization of degenerate second-order PDE in periodic and almost periodic environments and applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 22 (2005), pp. 667–677.

[80] ——, *Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media*, Comm. Partial Differential Equations, 30 (2005), pp. 335–375.

[81] A. Marshak and A. B. Davis, *3d Radiative transfer in cloudy atmoshperes*, Springer, New-York, 2005.

[82] A. Mellet and A. Vasseur, *Homogenization of a nonlinear transport equation*, Asymptot. Anal., 51 (2007), pp. 157–166.

[83] M. Métivier, *Convergence faible et principe d'invariance pour des martingales à valeurs dans des espaces de Sobolev*, Ann. Inst. H. Poincaré Probab. Statist., 20 (1984), pp. 329–348.

[84] S. G. Mikhlin and S. Prössdorf, *Singular integral operators*, Springer-Verlag, Berlin, 1986. Translated from the German by Albrecht Böttcher and Reinhard Lehmann.

[85] M. Mokhtar-Kharroubi, *Mathematical topics in neutron transport theory*, vol. 46 of Series on Advances in Mathematics for Applied Sciences, World Scientific Publishing Co. Inc., River Edge, NJ, 1997. New aspects, With a chapter by M. Choulli and P. Stefanov.

[86] F. Murat and L. Tartar, *H-convergence*, in Topics in the mathematical modelling of composite materials, vol. 31 of Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, Boston, MA, 1997, pp. 21–43.

[87] C. B. Muratov and S. Y. Shvartsman, *Boundary homogenization for periodic arrays of absorbers*, Multiscale Model. Simul., 7 (2008), pp. 44–61.

[88] J. Nolen and G. Papanicolaou, *Fine scale uncertainty in parameter estimation for elliptic equations*, Inverse Problems, 25 (2009), pp. 115021, 22.

[89] B. Øksendal, *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, sixth ed., 2003. An introduction with applications.

[90] G. C. Papanicolaou, *Diffusion in random media*, in Surveys in applied mathematics, Vol. 1, vol. 1 of Surveys Appl. Math., Plenum, New York, 1995, pp. 205–253.

[91] G. C. Papanicolaou and S. R. S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, in Random fields, Vol. I, II (Esztergom, 1979), vol. 27 of Colloq. Math. Soc. János Bolyai, North-Holland, Amsterdam, 1981, pp. 835–873.

[92] G. C. Papanicolaou and S. R. S. Varadhan, *Diffusions with random coefficients*, in Statistics and probability: essays in honor of C. R. Rao, North-Holland, Amsterdam, 1982, pp. 547–552.

[93] E. Pardoux and A. Piatnitski, *Homogenization of a singular random one dimensional PDE*, in Multi scale problems and asymptotic analysis, vol. 24 of GAKUTO Internat. Ser. Math. Sci. Appl., Gakkōtosho, Tokyo, 2006, pp. 291–303.

[94] E. Pardoux and A. L. Piatnitski, *Homogenization of a nonlinear random parabolic partial differential equation*, Stochastic Process. Appl., 104 (2003), pp. 1–27.

[95] G. A. Pavliotis and A. M. Stuart, *Multiscale methods*, vol. 53 of Texts in Applied Mathematics, Springer, New York, 2008. Averaging and homogenization.

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

[97] A. Quarteroni and A. Valli, *Numerical approximation of partial differential equations*, vol. 23 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1994.

[98] M. Reed and B. Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975.

[99] ——, *Methods of modern mathematical physics. I*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second ed., 1980. Functional analysis.

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

[101] H. Sato and M. C. Fehler, *Seismic wave propagation and scattering in the heterogeneous earth*, AIP Series in Modern Acoustics and Signal Processing, American Institute of Physics, New York, 1998.

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

[103] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.

[104] ——, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, vol. 43 of Princeton Mathematical Series, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[105] A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.

[106] M. S. Taqqu, *Weak convergence to fractional Brownian motion and to the Rosenblatt process*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 31 (1974/75), pp. 287–302.

[107] ——, *Convergence of integrated processes of arbitrary Hermite rank*, Z. Wahrsch. Verw. Gebiete, 50 (1979), pp. 53–83.

[108] M. E. Taylor, *Partial differential equations. I*, vol. 115 of Applied Mathematical Sciences, Springer-Verlag, New York, 1996. Basic theory.

[109] ——, *Partial differential equations. II*, vol. 116 of Applied Mathematical Sciences, Springer-Verlag, New York, 1996. Qualitative studies of linear equations.

[110] F. Trèves, *Introduction to pseudodifferential and Fourier integral operators. Vol. 1*, Plenum Press, New York, 1980. Pseudodifferential operators, The University Series in Mathematics.

[111] Z. X. Wang and D. R. Guo, *Special functions*, World Scientific Publishing Co. Inc., Teaneck, NJ, 1989. Translated from the Chinese by Guo and X. J. Xia.

[112] V. V. Yurinskiĭ, *Averaging of symmetric diffusion in a random medium*, Sibirsk. Mat. Zh., 27 (1986), pp. 167–180, 215.

[113] Z. X. Zhao, *Green function for Schrödinger operator and conditioned Feynman-Kac gauge*, J. Math. Anal. Appl., 116 (1986), pp. 309–334.

Index