

# Large Deviation Theory for a Homogenized and “Corrected” Elliptic ODE

Guillaume Bal <sup>\*</sup>      Roger Ghanem <sup>†</sup>      Ian Langmore <sup>\*</sup>

December 4, 2010

## Abstract

We study a one-dimensional elliptic problem with highly oscillatory random diffusion coefficient. We derive a homogenized solution and a so-called Gaussian corrector. We also prove a “pointwise” large deviation principle (LDP) for the full solution and approximate this LDP with a more tractable form. Applications to uncertainty quantification are considered.

**Keywords:** Differential equations, probability theory, random coefficients, homogenization, applied mathematics, uncertainty quantification

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Asymptotic expansion of the solution <math>u_\varepsilon</math></b>	<b>4</b>
<b>3</b>	<b>Homogenization and Gaussian Corrector</b>	<b>5</b>
3.1	Homogenization . . . . .	5
3.2	Gaussian corrector . . . . .	6
<b>4</b>	<b>Large Deviations</b>	<b>7</b>
4.1	Moderate and large deviations for our problem . . . . .	7
4.2	Independent sums and basic definitions . . . . .	9
4.3	The Gärtner-Ellis theorem and contraction principle . . . . .	10
4.4	Large deviations for $u_\varepsilon$ . . . . .	10
4.4.1	Parameterized, independent, uniformly elliptic media . . . . .	11
4.4.2	Convolved media with no uniform (in $\omega$ ) lower bound . . . . .	13
<b>5</b>	<b>LDP approximations and generalizations</b>	<b>20</b>
5.1	Rate functions for approximate solutions . . . . .	20
5.2	Generalizations to incompletely characterized media . . . . .	21

---

<sup>\*</sup>Department of Applied Physics and Applied Mathematics, Columbia University; email: ianlangmore@gmail.com

<sup>†</sup>University of Southern California

<b>6</b>	<b>Model Problems</b>	<b>21</b>
6.1	Numerical results for parameterized media . . . . .	21
6.2	Numerical results for convolved media . . . . .	27
<b>7</b>	<b>Proof of Homogenization/Gaussian Corrector Results</b>	<b>29</b>
7.1	Proof of theorem 3.1 . . . . .	30
7.1.1	$L^2$ convergence . . . . .	31
7.1.2	<i>a.s.</i> convergence . . . . .	32
7.2	Proof of theorem 3.2 . . . . .	33
7.2.1	One dimensional oscillatory integral . . . . .	33
7.2.2	Asymptotic expansion of the solution $u_\varepsilon$ . . . . .	38

# 1 Introduction

Partial differential equations whose (deterministic or random) coefficients have fine-scale structure are notoriously difficult to solve. Here we consider the following elliptic problem: Given probability space  $(\Omega, \mathcal{S}, P)$ , find  $u_\varepsilon \in L^2(\Omega; H_0^1([0, 1]))$  satisfying

$$-\frac{d}{dx}A(x, \xi, \frac{x}{\varepsilon}, \omega)\frac{d}{dx}u_\varepsilon = f(x), \quad x \in (0, 1), \omega \in \Omega, \varepsilon \ll 1. \quad (1)$$

Here  $\xi : \Omega \rightarrow \mathbb{R}^n$  is a random vector. In our case we assume either

- (i)  $A_\varepsilon(x) = A(x, \xi, \frac{x}{\varepsilon}, \omega) = a(x, \xi) + b(\frac{x}{\varepsilon}, \omega)$
- (ii)  $A_\varepsilon(x) = A(\frac{x}{\varepsilon})$ , with  $A(y)^{-1} = \gamma_{\lfloor y \rfloor}$  for a stationary dependent sequence  $\gamma_k$  (which depends on  $\xi$ ).  $\lfloor \cdot \rfloor$  denotes the “floor” function.

Case (i) is motivated by a Karhunen-Loève expansion, which for  $Q : [0, 1] \times \Omega \rightarrow \mathbb{R}$  having continuous covariance  $\Gamma : [0, 1]^2 \rightarrow \mathbb{R}$ ,  $\Gamma(x, y) := \mathbb{E}Q(x)Q(y)$ , is given by

$$Q(x, \omega) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} \psi_j(\omega) h_j(x), \quad \text{where} \quad \mathbb{E} \{ \psi_i \psi_j \} = \int h_i(x) h_j(x) dx = \delta_{ij}.$$

See [18, 16]. This is multi-scale, but could be approximated by two scales. In case (i), the high frequency randomness  $\frac{x}{\varepsilon}$  is “decoupled from  $\xi$ ” in the sense that, after conditioning on  $\xi$ ,  $A(x, \xi, y, \omega)$  is stationary in  $y$ . So in the model  $A = a + b$  we are assuming  $b$  is a stationary random field. For simplicity we always assume  $f(x)$  is deterministic.

Case (ii) is an example of dependent media with “short range correlations.” Use of different kernels  $h_k$  allows some flexibility in modeling. Although we only consider the case where the final  $\gamma_j$  are stationary, generalizations (using e.g.  $h_j(x)$ ) would not be difficult.

In some PDE of interest (e.g. (1), or other elliptic equations [1], and also linear transport [4]) the solution admits an expansion of the form

$$u_\varepsilon(x) = u_0(x) + v_\varepsilon(x) + R_\varepsilon(x), \quad (2)$$

where  $u_0$  is a *homogenized* solution that is dominant in the limit  $\varepsilon \rightarrow 0$  [17, 20], the remainder  $R_\varepsilon$  is negligible in some sense, and  $v_\varepsilon$  is given by an oscillatory integral

$$v_\varepsilon(x) = \int G(x, y) q(y, \frac{y}{\varepsilon}) dy. \quad (3)$$

In these cases one expects (and can often prove) that  $\varepsilon^{-\alpha}(u_\varepsilon - u_0)$  converges in distribution to a Gaussian process  $\varepsilon^\alpha v(x)$  (often  $\alpha = d/2$ ). Thus, one is justified in approximating  $u_\varepsilon$  by  $u_0$  plus a Gaussian corrector:

$$u_\varepsilon \approx u_0 + \varepsilon^\alpha v(x). \quad (4)$$

From an uncertainty quantification (UQ) perspective, this represents a significant simplification. Computation of the *homogenized* solution  $u_0$  is much less expensive than  $u_\varepsilon$ . The corrector  $v(x)$  has an explicit form in terms of e.g. an Itô integral. This allows explicit calculation of the correlation function. Moreover, draws from the random process  $\varepsilon^\alpha v(x)$  can be done with minimal (compared to calculation of  $u_\varepsilon$ ) effort. Another utility of corrector results is for validation of numerical homogenization schemes. For example, it is known that the numerical homogenization techniques MsFEM and HMM give solutions  $u_\varepsilon^h$  that converge to the correct homogenization limit  $u_0$  as  $\varepsilon, h \rightarrow 0$ . The question as to whether  $\varepsilon^{-\alpha}(u_\varepsilon^h - u_0)$  converges to the correct limit is explored (for (1)) in [3].

As a downside, central limit approximations such as (4) are expected to work well only for moderate deviations, i.e. for  $|u_\varepsilon - u_0| \sim O(\varepsilon^\alpha)$ . Sometimes of interest in UQ applications are questions related to large deviations, e.g.  $P[u_\varepsilon > \ell]$  for some  $\ell \sim O(1)$ .

Our main contribution is an investigation of the large-deviation behavior of  $u_\varepsilon$ , the solution to (1). As it turns out, it is possible to derive a large deviation principle (LDP) for the solution  $u_\varepsilon(x)$  (for fixed  $x$ ). This gives asymptotic limits of e.g.  $\varepsilon \log P[u_\varepsilon(x) \geq \ell]$  for  $\ell \sim O(1)$ . The resultant *rate function* is given implicitly (see theorems 4.4, 4.5) as a solution to two (one convex and one non-convex) four-dimensional optimization problems. We also derive an approximate LDP (proposition 5.1) that corresponds to the approximation  $u_\varepsilon \approx u_0 + v_\varepsilon$ . Since  $v_\varepsilon$  is given explicitly as an oscillatory integral, the rate function is “more explicit”, being the result of a one-dimensional convex optimization problem. We verify numerically that the approximate rate function works well when  $\varepsilon \ll 1$  and  $\ell \sim O(1)$  but “not too large.” This sort of approximate LDP should be available in other situations where the solution can be approximated by a homogenized term and an oscillatory integral. Along the way we also derive a large deviation principle for some one dimensional oscillatory integrals, which appears to be new as well.

A secondary contribution is a generalization of homogenization and corrector results. Homogenization results typically start with a uniformly (over all realizations) elliptic diffusion coefficient of the form  $A(\frac{x}{\varepsilon}, \omega)$ . In this case the homogenized tensor is constant. Here we generalize these results slightly by allowing for non-constant low-frequency randomness (in the case (i)) and relaxing the uniform ellipticity requirement to (8). We do this by conditioning on the coarse-scale and bounding higher moments of  $A_\varepsilon^{-1}$ . These assumptions are more in line with those encountered in UQ. We also prove almost sure convergence of the homogenized tensor. This is motivated by the fact that in practice the homogenized tensor can be obtained by picking one high-frequency media realization and averaging over a domain of size  $\rho$  [8]. Thus, it is nice to know that with probability one this realization converges as  $\rho \rightarrow \infty$ .

Our large deviation result allows us to determine (roughly) to what degree the Gaussian corrector captures the tail behavior of the solution. This is useful in applications where one may consider replacing the full solution with the homogenized solution plus a Gaussian corrector. Also, although we don’t answer this here, questions such

as “does HMM capture the large-deviation behavior of the solution” could potentially be answered in a manner similar to the question “does HMM capture the moderate deviation behavior” as discussed above. Also of interest (and also not pursued further here) is the relation between large deviations and importance sampling of rare events. As it turns out, a large deviations result can give “asymptotically efficient” importance functions [10, 14].

In section 2 an asymptotic expansion of  $u_\varepsilon$  is presented. In section 3, theorems 3.1 and 3.2 give results on the homogenized convergence  $u_\varepsilon \rightarrow u_0$ , and the corrector characterizing moderate deviations of  $u_\varepsilon - u_0$ . In section 4 large deviations are considered. After introducing the subject we derive a large deviation principle for two types of media, each with piecewise constant high-frequency parts (theorems 4.4, 4.5). We next present our approximate LDP in 5.1 and then numerical results in section 6. Proofs of the homogenization and corrector theorems, which are generalizations of known results, are relegated to sections 7.1 and 7.2.

## 2 Asymptotic expansion of the solution $u_\varepsilon$

The boundary value problem (1) may be integrated and the solution is

$$u_\varepsilon(x) = - \int_0^x \frac{F(s)}{A_\varepsilon(s)} ds + \frac{\langle F/A_\varepsilon \rangle}{\langle 1/A_\varepsilon \rangle} \int_0^x \frac{1}{A_\varepsilon(s)} ds. \quad (5)$$

Here we define *1-average*  $\langle \cdot \rangle$  of a function  $\phi : [0, 1] \rightarrow \mathbb{R}$ , and  $F : [0, 1] \rightarrow \mathbb{R}$  by

$$\langle \phi \rangle := \int_0^1 \phi(y) dy, \quad F(s) := \int_0^s f(t) dt.$$

Define the *homogenized tensor*  $A_0$  by

$$A_0(x) := (\mathbb{E}_\xi \{ A(x, 0)^{-1} \})^{-1},$$

where for  $X : \Omega \rightarrow \mathbb{R}$ , we denote conditional expectation by

$$\mathbb{E}_\xi \{ X \} := \mathbb{E} \{ X | \xi \} = \int_\Omega X(\omega) dP_\xi(\omega).$$

Above the measure  $P_\xi = P[\cdot | \xi]$  is defined implicitly.

Now re-write the integrals appearing in (5) as

$$\int_0^x \frac{1}{A_\varepsilon(s)} ds = \int_0^x \frac{1}{A_0(s)} ds + X_x^\varepsilon, \quad \int_0^x \frac{F(s)}{A_\varepsilon(s)} ds = \int_0^x \frac{F(s)}{A_0(s)} ds + Y_x^\varepsilon.$$

Defining the *homogenized solution*  $u_0$  by the equation (5) with  $A_0$  rather than  $A_\varepsilon$  (or equivalently the weak solution to (1) with  $A_0$  rather than  $A_\varepsilon$ ), we have the following expansion.

$$\begin{aligned} u_\varepsilon(x) &= u_0(x) + v_\varepsilon(x) + R_\varepsilon(x), \\ v_\varepsilon(x) &:= -Y_x^\varepsilon + \left( Y_1^\varepsilon - X_1^\varepsilon \frac{\langle F/A_0 \rangle}{\langle 1/A_0 \rangle} \right) \frac{1}{\langle 1/A_0 \rangle} \int_0^x \frac{1}{A_0(s)} ds + X_x^\varepsilon \frac{\langle F/A_0 \rangle}{\langle 1/A_0 \rangle}, \\ R_\varepsilon(x) &:= (X_1^\varepsilon)^2 \frac{\langle F/A_\varepsilon \rangle}{\langle 1/A_0 \rangle^2 \langle 1/A_\varepsilon \rangle} \int_0^x \frac{1}{A_\varepsilon(s)} ds \\ &\quad - \frac{X_1^\varepsilon}{\langle 1/A_0 \rangle^2} \left[ Y_1^\varepsilon \int_0^x \frac{1}{A_\varepsilon(s)} ds + \langle F/A_\varepsilon \rangle X_x^\varepsilon \right] + \frac{Y_1^\varepsilon X_x^\varepsilon}{\langle 1/A_0 \rangle}. \end{aligned} \quad (6)$$

The deterministic homogenized solution  $u_0$  is dominant in the limit  $\varepsilon \rightarrow 0$ . As we will show, the remainder  $R_\varepsilon$  is  $O(\varepsilon)$  in  $L^1(\Omega \times [0, 1])$ . The term  $v_\varepsilon$  can be re-written

$$v_\varepsilon(x) = \int_0^1 q_\varepsilon(s) G(x, s) \, ds, \quad q_\varepsilon(s) := \frac{1}{A_\varepsilon(s)} - \frac{1}{A_0(s)},$$

$$G(x, s) := \begin{cases} \left( F(s) - \frac{\langle F/A_0 \rangle}{\langle 1/A_0 \rangle} \right) \left( \frac{1}{\langle 1/A_0 \rangle} \int_0^x \frac{1}{A_0(t)} \, dt - 1 \right), & 0 \leq s \leq x \\ \left( F(s) - \frac{\langle F/A_0 \rangle}{\langle 1/A_0 \rangle} \right) \frac{1}{\langle 1/A_0 \rangle} \int_0^x \frac{1}{A_0(t)} \, dt, & x < s \leq 1. \end{cases} \quad (7)$$

Note that  $G(x, s)$  is piecewise continuous in  $s$  and so long as  $A_0^{-1}$  is bounded,  $G(x, s)$  is uniformly (in  $s$ ) Lipschitz in  $x$ . We also show that  $v_\varepsilon$  is  $O(\sqrt{\varepsilon})$  (in  $L^2(\Omega \times [0, 1])$ ) and has a limit that can be characterized well. Note that these are slight generalizations of previous results. In particular this limit was studied by [8], and in the case of media with long-range correlations in [2]. This paper adds the feature that the random media is allowed to be non-stationary in one variable, and has no uniform (in  $\omega$ ) ellipticity lower bound.

### 3 Homogenization and Gaussian Corrector

Here we obtain homogenization and Gaussian corrector results that are a generalization of known results (e.g. [17, 20, 7, 1]) to the case of media that has no uniform (in  $\omega$ ) upper or lower bounds. Proofs are deferred till section 7.

First, we assume  $A_\varepsilon(x)$  satisfies

$$0 < \nu_1(\omega, \varepsilon) \leq A_\varepsilon(x) \leq \nu_2(\omega, \varepsilon) < \infty. \quad (8)$$

We abuse notation by writing  $A_\varepsilon(x) := A(x, \frac{x}{\varepsilon}) = A(x, \frac{x}{\varepsilon}, \omega)$ . The form  $A(x, \frac{x}{\varepsilon}, \omega)$  emphasizes that  $A$  is a random field defined on a probability space  $(\Omega, \mathcal{S}, P)$ . The form  $A(x, \xi, \frac{x}{\varepsilon}, \theta)$  emphasizes the dependence on a special random vector  $\xi : \Omega \rightarrow \mathbb{R}^m$ , and a (possibly infinite) sequence of random variables  $\theta$ . Once we fix  $\xi$ ,  $A(x, y)$  exhibits some stationarity in  $y$  (although only weak-stationarity of  $A^{-1}$  is needed for homogenization in one dimension).

Note that functions such as  $\mathbb{E}_\xi \{A_\varepsilon\}$  do not depend on  $y$ , so we write  $\mathbb{E}_\xi \{Q\}(x) = \mathbb{E}_\xi \{Q(x, 0)\}$  for functions  $Q(x, \xi, y, \theta)$  whose conditional expectation does not depend on  $y$ . It is necessary to make some ergodicity assumptions on the process (in  $y$ ). In one dimension, we require mean-ergodicity of  $\mathbb{E}_\xi \{A^{-1}\}$ , which we express through decay of the covariance (10). We also assume  $\|f\|_{L^2} < \infty$ .

#### 3.1 Homogenization

We assume weak stationarity of  $A^{-1}$  and ellipticity of  $A_0$ . In other words, we assume

$$0 < c_1(\xi) \leq \mathbb{E}_\xi \{A^{-1}(x, y)\} = \mathbb{E}_\xi \{A^{-1}(x, 0)\} = A_0(x)^{-1} \leq c_2 < \infty,$$

and that the conditional covariance

$$\text{Cov}_\xi(x_1, x_2, z) := \mathbb{E}_\xi \left\{ \left[ A^{-1}(x_1, y) - \mathbb{E}_\xi \{A^{-1}\}(x_1) \right] \left[ A^{-1}(x_2, y+z) - \mathbb{E}_\xi \{A^{-1}\}(x_2) \right] \right\},$$

is independent of  $y \in [0, \infty)$ . We also assume

$$|\text{Cov}_\xi(x_1, x_2, z)| \leq \Gamma(z) \quad \text{with} \quad \|\Gamma\|_{L^1} =: C_{A^{-1}} < \infty. \quad (9)$$

Note that this implies

$$\varepsilon^2 \int_0^{\varepsilon^{-1}} \int_0^{\varepsilon^{-1}} |\text{Cov}_\xi(x_1, x_2, y - \tilde{y})| \, dy \, d\tilde{y} \leq \varepsilon C_{A^{-1}}. \quad (10)$$

This allows

**Theorem 3.1.**

$$\sqrt{\mathbb{E}_\xi \|u_\varepsilon - u_0\|_{L^2([0,1])}^2} \leq 3\sqrt{\varepsilon} \|f\|_{L^2} \sqrt{C_{A^{-1}}}.$$

Moreover, as  $\varepsilon \rightarrow 0$ , with  $\xi$  fixed,

$$u_\varepsilon(x) - u_0(x) \rightarrow 0 \quad \text{for every } x, \quad P_\xi \quad \text{a.s.}$$

and if the ellipticity bound (8) is independent of  $\varepsilon$ ,

$$\|u_\varepsilon - u_0\|_{L^2} \rightarrow 0, \quad P_\xi \quad \text{a.s.}$$

## 3.2 Gaussian corrector

To quantify the rate at which the random coefficient decorrelates, we introduce the following mixing condition (referred to in the literature as  $\rho$  mixing).

**Assumptions 3.1** (Mixing Condition). For two Borel sets  $A, B \subset \mathbb{R}$ , let  $\mathcal{S}_A, \mathcal{S}_B$  be the sub sigma algebras generated by  $A(x, y)$ ,  $x \in [0, 1]$ ,  $y \in A$  and  $y \in B$  respectively. We assume the existence of non-negative bounded and decreasing  $\varphi : [0, \infty) \rightarrow \mathbb{R}$  such that  $\varphi^{1/3} \in L^1$  that also satisfies

$$\sup \{ |\mathbb{E}[\eta_a \eta_b]| : \eta_a \in \mathcal{S}_A, \eta_b \in \mathcal{S}_B, \mathbb{E}\eta_a^2 = \mathbb{E}\eta_b^2 = 1, \mathbb{E}\eta_a = \mathbb{E}\eta_b = 0 \} \leq \varphi(d(A, B)).$$

After conditioning on  $\xi$  (e.g. fixing one realization of  $a(x, \xi)$  if  $A = a + b$ ) we are able to partially characterize the limiting distribution of  $u_\varepsilon - u_0$ .

**Theorem 3.2.** *If  $A(x, y)$  is stationary in  $y$ , and  $(s_1, s_2) \mapsto \text{Cov}_\xi(s_1, s_2, 0)$  is continuous at  $(x, x)$ ,  $A^{-1}$  satisfies the mixing condition (3.1), and*

$$\mathbb{E}_\xi \left\{ \left( \frac{1}{A(x, y)} - \mathbb{E}_\xi \left\{ \frac{1}{A(x, y)} \right\} \right)^6 \right\} \leq C_\xi < \infty,$$

then

$$\frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon}} \xrightarrow{\text{dist.}} v(x) := \int_0^1 G(x, t) \sigma(t) \, dW_t,$$

where  $W_t$  is a one-dimensional Brownian motion,  $G$  is given by (7), and

$$\sigma^2(t) = \int_{-\infty}^{\infty} \text{Cov}_\xi(t, t, q) \, dq.$$

*Remark 3.1.* The continuity on  $\text{Cov}_\xi$  is equivalent to mean-square continuity of  $x \mapsto A(x, 0)^{-1}$  (see e.g. [18]). This means more-or-less that the media varies slowly with respect to  $x$ . This is where scale-separation comes in.

*Remark 3.2.* The variance of the above expression is then given by

$$\int_0^1 G(x, t)^2 \sigma^2(t) dt. \quad (11)$$

*Remark 3.3.* The lack of a uniform (in  $\omega$ ) ellipticity lower bound is dealt with through the bounds on sixth moments and the mixing conditions.

## 4 Large Deviations

Here we derive a rigorous “pointwise” large-deviations result for  $u_\varepsilon$ , and an approximate rate function.

### 4.1 Moderate and large deviations for our problem

For small enough  $\varepsilon$ , the homogenized solution  $u_0$  captures the bulk of the solution  $u_\varepsilon$ . The corrector attempts to capture some statistics of the term  $u_\varepsilon - u_0 = v_\varepsilon + R_\varepsilon$ . Our corrector result shows that

$$\begin{aligned} \frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon}} &:= \frac{1}{\sqrt{\varepsilon}} \int_0^1 G(x, t) q\left(t, \frac{t}{\varepsilon}\right) + \frac{R_\varepsilon(x)}{\sqrt{\varepsilon}} \\ &\xrightarrow{\text{dist.}} v(x) := \int_0^1 G(x, t) \sigma(t) dW_t. \end{aligned} \quad (12)$$

By definition this means that, for any  $\ell > 0$

$$\mathbb{P}_\xi \left[ \frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon}} \geq \ell \right] \rightarrow \mathbb{P}_\xi[v(x) \geq \ell]. \quad (13)$$

A relevant question is *whether or not*

$$\mathbb{P}_\xi[u_\varepsilon(x) \geq \ell] \approx \mathbb{P}_\xi[u_0(x) + \sqrt{\varepsilon}v(\varepsilon) \geq \ell]. \quad (14)$$

The inequality  $u_\varepsilon \geq \ell$  (for  $\ell > u_0$ ) is a *large deviation* since (12) (or a law of large numbers result) shows that  $u_\varepsilon$  concentrates near  $u_0$ . On the other hand,  $\varepsilon^{-1/2}(u_\varepsilon - u_0) \geq \ell$  is a *moderate deviation*. Generally speaking, (14) *does not* hold. Instead, from (13) we can only rigorously infer something about moderate deviations, namely  $\mathbb{P}_\xi[u_\varepsilon(x) \geq \sqrt{\varepsilon}\ell] \approx \mathbb{P}_\xi[u_0(x) + \sqrt{\varepsilon}v(x) \geq \sqrt{\varepsilon}\ell]$ .

For simplicity, we consider large deviations at only one fixed  $x \in (0, 1)$  and often change the notation to  $u_0, u_\varepsilon, v_\varepsilon$  (dropping the  $x$  dependence).

**Definition 4.1** (Rate functions). A *rate function*  $I$  is a lower semicontinuous mapping (such that for all  $\alpha \in [0, \infty)$ , the sub-level set  $\Psi_I(\alpha) := \{x : I(x) \leq \alpha\}$  is closed)  $I : \mathbb{R}^n \rightarrow [0, \infty]$ . A *good rate function* is a rate function for which all the sub-level sets  $\Psi_I(\alpha)$  are compact.

**Definition 4.2.** We say that a family of random vectors  $Y_\varepsilon \in \mathbb{R}^n$  satisfy a large deviations principle (LDP) with rate function  $I$  if for all  $\Gamma \subset \mathbb{R}^n$

$$-\inf_{y \in \Gamma^\circ} I(y) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log P[Y_\varepsilon \in \Gamma] \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log P[Y_\varepsilon \in \Gamma] \leq -\inf_{y \in \bar{\Gamma}} I(y).$$

Above,  $\Gamma^\circ, \bar{\Gamma}$  denote the interior and closure of  $\Gamma$ .

Considering now our solution  $u_\varepsilon \in \mathbb{R}$ , by choosing  $A = [\ell, \infty)$  we obtain limiting upper and lower bounds on  $P_\xi[u_\varepsilon \geq \ell]$ . We also often have, for  $\ell > \lim \mathbb{E}_\xi u_\varepsilon$ ,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log P[u_\varepsilon \geq \ell] = -I(\ell). \quad (15)$$

We will find a rate function  $I_{u_\varepsilon}$  for  $u_\varepsilon$ . We also find a rate function  $\tilde{I}$  for  $u_0 + v_\varepsilon$ , which proves to be more tractable. Simulations show that  $I_{u_\varepsilon} \approx \tilde{I}$  for  $\ell \sim O(1)$  but not too large (see figures 5, 6, and 9).

The rate-function  $I_\varepsilon$  gives the exponential rate of convergence in the sense that given  $\ell \in \mathbb{R}$ ,  $\delta > 0$ , there exists  $\varepsilon_0$  such that for  $\varepsilon < \varepsilon_0$

$$P_\xi[u_\varepsilon \geq \ell] \leq e^{-(I(\ell) - \delta)/\varepsilon}. \quad (16)$$

As an example, consider the corrector in our 1-d problem,  $\sqrt{\varepsilon}v$ . It is Gaussian with mean zero and variance equal to  $\varepsilon C_c$  for some  $C_c$  (given by (11)). Therefore  $u_\varepsilon \approx u_0 + \sqrt{\varepsilon}v \sim \mathcal{N}(u_0, \varepsilon C_c)$ . Keeping the first term in an asymptotic expansion of the complementary error function, we have that for  $\ell > u_0$ ,

$$P_\xi[u_0 + \sqrt{\varepsilon}v \geq \ell] \sim \frac{\sqrt{C_c \varepsilon}}{\sqrt{2\pi}(\ell - u_0)} e^{-(\ell - u_0)^2 / (2C_c \varepsilon)}.$$

Thus, when

$$\log(C_c \varepsilon / (\ell - u_0)^2) \ll \frac{(\ell - u_0)^2}{C_c \varepsilon}, \quad (17)$$

we will have  $\varepsilon \log P_\xi[u_0 + \sqrt{\varepsilon}v \geq \ell] \approx -(\ell - u_0)^2 / (2C_c)$ . Note that this is a “small  $\varepsilon$  and large  $|\ell - u_0|$ ” condition, as it should be. Comparing this to (16) we see that the Gaussian corrector captures the asymptotic tail behavior when

$$\frac{(\ell - u_0)^2}{2C_c} \approx I(\ell). \quad (18)$$

If (18) does not hold, the corrector cannot capture the tail behavior of  $u_\varepsilon$ . A cautionary note is in order here. The large deviations result captures the exponential rate of decay, and important algebraic factors in  $\varepsilon$  are not captured. So, for finite  $\varepsilon$  the rate function can be used for comparative purposes, not to estimate the true tail. It should be noted that so-called concentration inequalities provide a number of upper bounds on sums of random variables, often in pre-asymptotic regimes. Often these require fewer assumptions but do not claim to be tight. For example, *Chernoff’s bounding method* is given in (20). See [6] for a survey, and [19] for an application to uncertainty quantification. For independent media, we derive rigorous upper bounds for finite  $\varepsilon$  (e.g. (20)). For dependent media similar bounds are available but are more complex than the asymptotic bounds. Another reason for using asymptotic bounds is that a “fair” comparison such as (18) can be made.



## 4.2 Independent sums and basic definitions

It is instructive to start here. Let  $X_n$  be random variables and define

$$\hat{S}_N := \frac{1}{N} \sum_{n=1}^N X_n.$$

Now, for  $\lambda \geq 0$ ,

$$P[\hat{S}_N \geq \ell] = P[N\hat{S}_N \geq N\ell] = \mathbb{E} \left\{ \mathbf{1}_{N\hat{S}_N \geq N\ell} \right\} \leq e^{-N\lambda\ell} \mathbb{E} \left\{ e^{\lambda N\hat{S}_N} \right\}. \quad (19)$$

With vector-valued random variables in mind, we define

**Definition 4.3.** The *logarithmic moment generating function* for a random variable  $Y$  is defined as

$$\Lambda(Y, \lambda) = \Lambda_Y(\lambda) := \log \mathbb{E} \left\{ e^{\lambda \cdot Y} \right\}.$$

(19) allows us to conclude

$$\frac{1}{N} \log P[\hat{S}_N \geq \ell] \leq - \sup_{\lambda \geq 0} \left[ \lambda \ell - \frac{1}{N} \Lambda(N\hat{S}_N, \lambda) \right]. \quad (20)$$

The above (*Chernoff's*) bound holds without any assumptions. Suppose the  $X_n$  are i.i.d. This leads to

$$\mathbb{E} \left\{ e^{\lambda N\hat{S}_N} \right\} = \prod_{n=1}^N \mathbb{E} \left\{ e^{\lambda X_n} \right\},$$

and thus

$$\frac{1}{N} \Lambda(N\hat{S}_N, \lambda N) = \frac{1}{N} \sum_{n=1}^N \log \mathbb{E} \left\{ e^{\lambda X_n} \right\} = \Lambda_X(\lambda). \quad (21)$$

Note that  $\Lambda_X$  is convex since by Hölder's inequality

$$\begin{aligned} \Lambda_X(t\lambda_1 + (1-t)\lambda_2) &= \log \mathbb{E} \left\{ (e^{\lambda_1 X})^t (e^{\lambda_2 X})^{(1-t)} \right\} \\ &\leq \log \left\{ \mathbb{E} \left\{ e^{\lambda_1 X} \right\}^t \mathbb{E} \left\{ e^{\lambda_2 X} \right\}^{(1-t)} \right\} = t\Lambda_X(\lambda_1) + (1-t)\Lambda_X(\lambda_2). \end{aligned}$$

**Definition 4.4.** The *Fenchel-Legendre transform* of  $\Lambda(Y, \cdot)$  is defined by

$$\Lambda^*(Y, \ell) = \Lambda_Y^*(\ell) := \sup_{\lambda \in \mathbb{R}^n} [\lambda \cdot \ell - \Lambda_Y(\lambda)].$$

Inserting (21) back into (20) (taking into account negative  $\lambda$ ) we have a large deviation upper bound with rate function  $\Lambda_X^*(\cdot)$ . The “trick” is to obtain a lower bound and thus show that this upper bound is tight in the limit  $N \rightarrow \infty$ . This indeed is the case and the result is

**Theorem 4.1** (Cramér, [13]). *The sum  $\hat{S}_N$  satisfies the LDP with good convex rate function  $\Lambda_X^*(\cdot)$ . Moreover, (15) holds.*

Since  $\Lambda_X(0) = 0$ , we always have  $\Lambda_X^*(\ell) \geq 0$ , i.e. we never have exponential growth. Jensen's inequality shows that  $\Lambda_X(\lambda) \geq \lambda \mathbb{E}X$ , so for  $\ell = \mathbb{E}X$  we also have  $\lambda \ell - \Lambda_X(\lambda) \leq 0$ . Therefore  $\Lambda_X^*(\mathbb{E}X) = 0$ . This makes sense in view of the law of large numbers.

### 4.3 The Gärtner-Ellis theorem and contraction principle

**Definition 4.5.** For any function  $H$  taking values in  $(-\infty, \infty]$ , we define  $\mathcal{D}_H := \{x : H(x) < \infty\}$ .

The Gärtner-Ellis theorem will be used to prove results for our model problems.

**Definition 4.6.** Convex  $\Lambda : \mathbb{R}^n \rightarrow (-\infty, \infty]$  is called *essentially smooth* if  $\mathcal{D}_\Lambda^\circ$  is nonempty,  $\Lambda$  is differentiable throughout  $\mathcal{D}_\Lambda^\circ$ , and  $\Lambda$  is *steep*. Steep means that  $\lim_{j \rightarrow \infty} |\nabla \Lambda(\lambda_j)| = \infty$  whenever  $\{\lambda_j\}$  is a sequence in  $\mathcal{D}_\Lambda^\circ$  converging to a boundary point of  $\mathcal{D}_\Lambda^\circ$ .

**Theorem 4.2** (Gärtner-Ellis [13]). *Suppose*

$$\Lambda(\lambda) := \lim_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{E} \left\{ e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} \right\}$$

*exists as an extended real number. Furthermore suppose that  $\Lambda$  is essentially smooth, lower semicontinuous and that the origin belongs to the interior of  $\mathcal{D}_\Lambda$ . Then  $Z_\varepsilon$  satisfies an LDP with good convex rate function  $\Lambda^*$  defined by*

$$\Lambda^*(\ell) := \sup_{\lambda \in \mathbb{R}^n} [\lambda \cdot \ell - \Lambda(\lambda)].$$

*Remark.* If  $n = 1$  and  $\Lambda^*$  is finite in a neighborhood of  $\ell > a := \lim \mathbb{E} Z_\varepsilon$  then convex  $\Lambda^*$  is non-decreasing on  $[a, \infty)$  and continuous in this neighborhood. Therefore (15) holds.

Notice that the Gärtner-Ellis theorem does not require independence. Indeed, one can use it to prove an LDP for mixing random variables [13, 9].

While the Gärtner-Ellis theorem allows us to obtain an LDP for oscillatory integrals, we need the contraction principle for functions of those integrals such as 5.

**Theorem 4.3** (Contraction principle). *Suppose  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is continuous and  $I : \mathbb{R}^n \rightarrow [0, \infty]$  is a good rate function for the family of random variables  $Z_\varepsilon$  and associated measures  $\mu_\varepsilon$  ( $\mu_\varepsilon(A) = P[Z_\varepsilon \in A]$ ). For  $y \in \mathbb{R}^m$ , define*

$$I'(y) := \inf \{I(x) : x \in \mathbb{R}^n, y = f(x)\}.$$

*Then  $I'$  is a good rate function controlling the LDP associated with the measures  $\mu_\varepsilon \circ f^{-1}$  ( $\mu_\varepsilon \circ f^{-1}(B) = P[f(Z_\varepsilon) \in B]$ ).*

In other words, with  $Y_\varepsilon := f(Z_\varepsilon)$ , the rate at which  $Z_\varepsilon$  concentrates away from  $\ell$  will be determined by the point in  $f^{-1}(\{\ell\})$  holding the most mass.

### 4.4 Large deviations for $u_\varepsilon$

In light of (5), we have

$$\begin{aligned} u_\varepsilon(x) &= g(Z_\varepsilon), & g : \mathbb{R}^2 \times (0, \infty)^2 &\rightarrow \mathbb{R}, & g(z) &= g(z_1, z_2, z_3, z_4) = -z_1 + z_2 \frac{z_3}{z_4}, \\ Z_\varepsilon &= (Z_{\varepsilon 1}, \dots, Z_{\varepsilon 4}), & Z_{\varepsilon i} &= \int_0^1 \frac{H_i(s)}{A_\varepsilon(s)} ds, \\ H_1 &= F \mathbf{1}_{(0,x)}, & H_2 &= F, & H_3 &= \mathbf{1}_{(0,x)}, & H_4 &= 1. \end{aligned} \tag{22}$$

Also set

$$\mathbf{H} := (H_1, H_2, H_3, H_4),$$

so that  $\lambda \cdot Z_\varepsilon = \int \lambda \cdot \mathbf{H} / A_\varepsilon$ .

Using the Gärtner-Ellis theorem we will find rate function  $I_{Z_\varepsilon}$  for  $Z_\varepsilon$ , and then by the contraction principle the rate function for  $u_\varepsilon$  is

$$I_{u_\varepsilon}(\ell) = \inf_{z \in g^{-1}\{\ell\}} I_{Z_\varepsilon}(z). \quad (23)$$

The real work involved here is in characterizing the limiting Cramér functional  $\lim \varepsilon \log \mathbb{E}_\xi \exp\{\varepsilon^{-1} \lambda \cdot Z_\varepsilon\}$ . We will do this for the media types (i) and (ii) mentioned in the introduction.

#### 4.4.1 Parameterized, independent, uniformly elliptic media

We assume here that the high frequency media is piecewise constant and independent:

$$\begin{aligned} A(x, y) &= a(x, \xi) + b(y, \theta), \quad \text{with every realization } a(x, \xi) \text{ continuous,} \\ b(y, \theta) &= \nu_b \sum_{j=1}^{\infty} \theta_j \mathbf{1}_{[n-1, n)}(y), \quad \theta_i \sim_{\text{i.i.d.}} \pi_\theta, \quad |\theta| \leq 1, \\ 0 < \nu_1 &\leq A(x, y) \leq \nu_2. \end{aligned} \quad (24)$$

Considered as a discrete process (at points centered at  $n\varepsilon$ ), the field is stationary and ergodic. It is not truly stationary since the correlation  $\mathbb{E}_\xi A(x, y)A(x, z)$  depends on more than the difference  $y - z$ . Nonetheless we apply our theorems and obtain results that are validated by simulation. The correlation condition is satisfied trivially since  $A(x_1, y_1)$ , and  $A(x_2, y_2)$  are conditionally independent whenever  $|y_1 - y_2| > 1$ . Note also that the low and high frequency parts are in the form of Karhunen-Loève expansions, but the total field is not.

Since this solution (5) involves  $A_\varepsilon^{-1}$ , it is not surprising that we need to define

$$V_\alpha := \frac{1}{\alpha + \nu_b \theta} \sim \pi_{V_\alpha}(v) = \frac{1}{\nu_b} \frac{1}{v^2} \pi_\theta \left( \frac{1}{\nu_b} \left( \frac{1}{v} - \alpha \right) \right), \quad (25)$$

The main result for this media is

**Theorem 4.4.** *With  $g$ ,  $Z_\varepsilon$ ,  $\mathbf{H}(s)$ ,  $V_\alpha$  given by (22), (24), (25), define*

$$\Lambda(\lambda) := \int_0^1 \Lambda(V_{a(s, \xi)}, \lambda \cdot \mathbf{H}(s)) ds,$$

then  $\Lambda \in C^\infty(\mathbb{R}^4)$  is a convex function such that when  $A_\varepsilon$  is defined by (24)

$$\lim \varepsilon \log \mathbb{E}_\xi e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} = \Lambda(\lambda).$$

Moreover, for fixed  $\xi$ ,  $Z_\varepsilon$  satisfies a large deviation principle with good convex rate function

$$\Lambda^*(\ell) := \sup_{\lambda \in \mathbb{R}^4} [\lambda \cdot \ell - \Lambda(\lambda)],$$

and  $u_\varepsilon$  satisfies a large deviation principle with good rate function

$$I_{u_\varepsilon}(\ell) := \inf_{z \in g^{-1}\{\ell\}} \Lambda^*(z).$$

*Proof.* We will condition on  $\xi$  and then approximate  $a(x)$  by a piecewise constant function so that  $a(x)$  appears only as a parameter. Now

$$\mathbb{E}_\xi \left\{ e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} \right\} = \mathbb{E}_\xi \left\{ \exp \left\{ \varepsilon^{-1} \int_0^1 \frac{\lambda \cdot \mathbf{H}(s)}{A_\varepsilon(s)} ds \right\} \right\}.$$

So for  $N < \varepsilon^{-1} < N + 1 \in \mathbb{N}$

$$\begin{aligned} \varepsilon \log \mathbb{E}_\xi \left\{ e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} \right\} &= \varepsilon N \frac{1}{N} \sum_{n=1}^N \log \mathbb{E}_\xi \left\{ e^{\lambda \cdot X_N(n)} \right\} + \varepsilon \log \mathbb{E}_\xi \left\{ e^{\lambda \cdot Y_N} \right\}, \\ X_N(n) &:= \int_{n-1}^n \frac{\mathbf{H}(\frac{s}{N})}{a(\frac{s}{N}) + \nu_b \theta}, \quad Y_N := \int_N^{\varepsilon^{-1}} \frac{\mathbf{H}(\frac{s}{N})}{a(\frac{s}{N}) + \nu_b \theta}. \end{aligned}$$

Since  $\nu_2^{-1} \leq (a + \nu_b \theta)^{-1} \leq \nu_1^{-1}$  we also have upper and lower bounds on  $\mathbb{E}_\xi \{ \exp \{ \lambda \cdot Y_N \} \}$ . Therefore  $\varepsilon \log \mathbb{E}_\xi \{ \exp \{ \lambda \cdot Y_N \} \} \rightarrow 0$ . Using also the fact  $\varepsilon N \rightarrow 1$  we have

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{E}_\xi \left\{ e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} \right\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \log \mathbb{E}_\xi \left\{ e^{\lambda \cdot X_N(n)} \right\}.$$

We henceforth study the limit on the right.

We can approximate

$$\begin{aligned} X_N(n) &\leq \frac{\bar{\mathbf{H}}_N(n)}{\underline{a}_N(n) + \nu_b \theta_n} \sim \bar{\mathbf{H}}_N(n) V_{\underline{a}_N(n)}, \\ \bar{\mathbf{H}}_N(n) &:= \max_{n-1 \leq s < n} \mathbf{H}(\frac{s}{N}), \quad \underline{a}_N(n) := \min_{n-1 \leq s < n} a(\frac{s}{N}). \end{aligned}$$

We thus have

$$\varepsilon \log \mathbb{E} \{ \lambda v_\varepsilon / \varepsilon \} \leq \frac{1}{N} \sum_{n=1}^N \Lambda(V_{\underline{a}_N(n)}, \lambda \cdot \bar{\mathbf{H}}_N(n)). \quad (26)$$

Similarly we can choose  $\underline{\mathbf{H}}_N(n)$  and  $\bar{a}_N(n)$  to provide a lower bound. Together they yield

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \Lambda(V_{\bar{a}_{N-1}(n)}, \lambda \cdot \underline{\mathbf{H}}_{N-1}(n)) &\leq \frac{1}{N} \sum_{n=1}^N \log \mathbb{E}_\xi \left\{ e^{\lambda \cdot X_N(n)} \right\} \\ &\leq \frac{1}{N} \sum_{n=1}^N \Lambda(V_{\underline{a}_N(n)}, \lambda \cdot \bar{\mathbf{H}}_N(n)). \end{aligned}$$

The ellipticity bounds imply that  $(\alpha, \lambda) \mapsto \Lambda(V_\alpha, \lambda)$  is  $C^\infty$  for  $\alpha$  taking values in the closure of  $\text{Range}(a)$ . Therefore,

$$\Lambda(V_{\underline{a}_N(n)}, \lambda \cdot \bar{\mathbf{H}}_N(n)) \rightarrow \Lambda(V_{a(\frac{n}{N})}, \lambda \cdot \mathbf{H}(\frac{n}{N})), \quad N \rightarrow \infty.$$

And thus (after extending  $\underline{a}_N, \overline{\mathbf{H}}_N$  to be piecewise constant), and using the continuity of  $s \mapsto a(s, \xi)$  and  $\mathbf{H}(s)$  we have

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \Lambda(V_{\underline{a}_N(n)}, \lambda \cdot \overline{\mathbf{H}}_N(n)) &= \int_0^1 \Lambda(V_{\underline{a}_N(sN)}, \lambda \cdot \overline{\mathbf{H}}_N(sN)) \, ds \\ &\rightarrow \int_0^1 \Lambda(V_{a(s)}, \lambda \cdot \mathbf{H}(s)) \, ds. \end{aligned}$$

The same holds for the lower bound. We have thus shown

$$\lim \varepsilon \log e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} = \int_0^1 \Lambda(V_{a(s)}, \lambda \cdot \mathbf{H}(s)) \, ds. \quad (27)$$

Since  $\mathbb{E} \{\exp \{\lambda V_\alpha\}\}$  is finite for all  $\alpha$ , the hypothesis of the Gärtner-Ellis theorem are trivially satisfied. Recalling the definition of  $V_\alpha, \Lambda$ , we see that the theorem is proved.  $\square$

#### 4.4.2 Convolved media with no uniform (in $\omega$ ) lower bound

Here we obtain a large deviation principle for dependent media given by a convolution of random variables that, while being positive, have no uniform lower bound. Convolution provides a convenient way to generate dependencies.

To avoid additional technicalities, we restrict our attention to families of functions indexed by  $\varepsilon$  such that  $\varepsilon^{-1} \in \mathbb{N}$ . Starting from this result, using the notion of *exponential equivalence*, and adding an assumption of a finite logarithmic moment generating function, it would be possible to prove an LDP for general  $\varepsilon \in (0, \infty)$ .

Define

$$\begin{aligned} A_\varepsilon(x) &= A\left(\frac{x}{\varepsilon}\right), \quad \text{where} \quad \frac{1}{A(s)} := \sum_{n=1}^{\infty} \mathbf{1}_{[n-1, n)}(s) \gamma_n, \\ \gamma_n &:= \sum_{m=-\infty}^{\infty} h_{n-m} \beta_m, \quad h_n \geq 0, \quad \|h\|_1 := \sum_k h_k < \infty, \end{aligned} \quad (28)$$

and the  $\{\beta_m\}_{m=-\infty}^{\infty}$  are non-negative i.i.d. random variables each depending on the same random vector (parameter)  $\xi$ . The random variables  $\gamma_n$  are well defined so long as the characteristic function  $\phi_M(t) := \prod_{|k| < M} \mathbb{E} \{e^{ith_k \beta}\}$  has a continuous limit  $\phi(t)$ . We take this as an assumption and proceed.

In this case, the large deviation principles for  $Z_\varepsilon, u_\varepsilon$  are a direct result of lemmas 4.1, 4.2 (below), the Gärtner-Ellis theorem, and the contraction principle.

**Theorem 4.5.** *With  $g, Z_\varepsilon$ , and  $\mathbf{H}(s)$  given by (22), and (24), define*

$$\Lambda(\lambda) := \int_0^1 \Lambda_{\beta(\xi)}(\|h\|_1 \lambda \cdot \mathbf{H}(s)) \, ds,$$

*then  $\Lambda \in C^\infty(\mathbb{R}^4)$  is a convex function such that when  $A_\varepsilon$  is defined by (28)*

$$\lim \varepsilon \log \mathbb{E}_\xi e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} = \Lambda(\lambda).$$

If in addition  $\Lambda$  is steep (see proposition 4.1), then for fixed  $\xi$ ,  $Z_\varepsilon$  satisfies a large deviation principle with good convex rate function

$$\Lambda^*(\ell) := \sup_{\lambda \in \mathbb{R}^4} [\lambda \cdot \ell - \Lambda(\lambda)],$$

and  $u_\varepsilon$  satisfies a large deviation principle with good rate function

$$I_{u_\varepsilon}(\ell) := \inf_{z \in g^{-1}\{\ell\}} \Lambda^*(z).$$

**Lemma 4.1.** Let  $\gamma_n := \sum_k h_{n-k} \beta_k$  with non-negative  $\{h_k\} \in \ell^1$  and non-negative  $\beta_k$  i.i.d. with the same law as  $\beta$ . For  $G \in L^\infty([0, 1])$ , define

$$\hat{S}_N := \int_0^1 G(s) \gamma_{\lfloor sN \rfloor} ds.$$

Then for  $\lambda \in \mathbb{R}$  the following limit exists in  $(-\infty, \infty]$

$$\lim \frac{1}{N} \Lambda(N \hat{S}_N, \lambda) = \int_0^1 \Lambda_\beta(\lambda G(s) \|h\|_1) ds,$$

with  $\Lambda_\beta$  as in definition 4.3.

*Proof of lemma 4.1.* For a result involving sums of convolved stationary random variables see [11]. The key difference is that here we allow the moment generating function to be infinite, and the term  $G(s)$  makes the sum non-stationary. This necessitates a new proof technique that relies on non-negativity of the  $\beta_k$  and the  $h_k$ .

$$\begin{aligned} N \hat{S}_N &= N \int_0^1 G(s) \gamma_{\lfloor sN \rfloor} ds = \sum_{n=1}^N \left[ \int_{n-1}^n G(s/N) ds \right] \gamma_n \\ &= \sum_{n=1}^N \left[ \int_{n-1}^n G(s/N) ds \right] \sum_{k \in \mathbb{Z}} \beta_k h_{n-k}. \end{aligned}$$

Making the substitution  $j = n - k$  we have

$$N \hat{S}_N = \sum_{k \in \mathbb{Z}} \beta_k H_{N,k}, \quad H_{N,k} := \sum_{j=1-k}^{N-k} h_j \int_{j+k-1}^{j+k} G(s/N) ds.$$

This is a sum of independent random variables. Since  $\beta_k \sim \beta$ ,

$$\frac{1}{N} \Lambda(N \hat{S}_N, \lambda) = \frac{1}{N} \sum_{k \in \mathbb{Z}} \Lambda_\beta(\lambda H_{N,k}) = \frac{1}{N} \int_{-\infty}^{\infty} \Lambda_\beta(\lambda H_{N, \lfloor s \rfloor}) ds.$$

Using (43) and changing  $s \mapsto Ns$  we have

$$\frac{1}{N} \Lambda(N \hat{S}_N, \lambda) = \int_{-\infty}^{\infty} \varphi_N(s) ds, \quad \varphi_N(s) := \Lambda_\beta(\lambda H_{N, \lfloor sN \rfloor}).$$

The key to our proof is the fact that  $H_{N,k}$  itself can be written as an expectation

$$\begin{aligned} H_{N,k} &= \int_0^N G(s/N) h_{\lfloor s-k+1 \rfloor} ds = \mathbb{E}_{\pi_k} \{ \|h\|_1 G(\cdot/N) \mathbf{1}_{[0,N]} \}, \\ \mathbb{E}_{\pi_k} \{ f(S) \} &:= \int_{-\infty}^{\infty} f(s) \frac{h_{\lfloor s-k+1 \rfloor}}{\|h\|_1} ds. \end{aligned}$$

Since  $\pi_k$  is a density that concentrates near  $s = k - 1$ , for almost every  $s \in [0, 1]$ ,  $\mathbb{E}_{\pi_{\lfloor sN \rfloor}} \{ \|h\|_1 G(\cdot/N) \mathbf{1}_{[0,N]} \} \rightarrow \|h\|_1 G(s)$  (this is a result on the Lebesgue set for integrable functions, see e.g. [15] theorem 3.20 or [13] theorem C.13). Also,

$$|H_{N, \lfloor sN \rfloor}| \leq \|G\|_{\infty} \sum_{j=\lfloor 1-sN \rfloor}^{\lfloor N+1-sN \rfloor} h_j. \quad (29)$$

Since this tends to zero for  $s \notin [0, 1]$ , we have

$$\varphi_N(s) \rightarrow \varphi(s) := \begin{cases} \Lambda_{\beta}(\lambda \|h\|_1 G(s)), & a.e. s \in [0, 1], \\ 0, & a.e. s \notin [0, 1]. \end{cases} \quad (30)$$

To prove the lemma, we thus have to show  $\int \varphi_N \rightarrow \int \varphi$ .

We first obtain an upper bound in cases where  $\Lambda(\lambda) < \infty$ . When  $\Lambda(\lambda) = \infty$ , the lower bound we derive will be infinite, and thus the upper bound as well.

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_N(s) ds &= \int_{-\infty}^{\infty} \Lambda_{\beta} \left( \mathbb{E}_{\pi_{\lfloor sN \rfloor}} \{ \lambda \|h\|_1 G(\cdot/N) \mathbf{1}_{[0,N]} \} \right) ds \\ &\leq \int_{-\infty}^{\infty} \mathbb{E}_{\pi_{\lfloor sN \rfloor}} \{ \Lambda_{\beta}(\lambda \|h\|_1 G(\cdot/N)) \} ds \\ &= \int_{-\infty}^{\infty} \left[ \int_0^N \Lambda_{\beta}(\lambda \|h\|_1 G(t/N)) \pi_{\lfloor sN \rfloor}(t) dt \right] ds \\ &= \frac{1}{N} \int_{-\infty}^{\infty} \left[ \int_0^N \Lambda_{\beta}(\lambda \|h\|_1 G(t/N)) \pi_{\lfloor s \rfloor}(t) dt \right] ds \\ &= \int_{-\infty}^{\infty} \left[ \int_0^1 \Lambda_{\beta}(\lambda \|h\|_1 G(t)) \pi_{\lfloor s \rfloor}(tN) dt \right] ds \\ &= \int_0^1 \Lambda_{\beta}(\lambda \|h\|_1 G(t)) \left( \int_{-\infty}^{\infty} \pi_{\lfloor s \rfloor}(tN) ds \right) dt \\ &= \int_0^1 \Lambda_{\beta}(\lambda \|h\|_1 G(t)) dt = \int_0^1 \varphi(t) dt. \end{aligned} \quad (31)$$

Above, the inequality is due to Jensen's inequality and the convexity of  $\Lambda_{\beta}$ . The change of integration order is justified by Fubini since  $\varphi \in L^1$ .

We now obtain a lower bound. We will show that  $\varphi_N \geq \psi$ , with  $\psi \in L^1$ . Then a corollary of Fatou's lemma gives us

$$\liminf \int \varphi_N(s) ds \geq \int \varphi(s) ds = \Lambda(\lambda). \quad (32)$$

With  $M(t) := \mathbb{E} \{ e^{t\beta} \}$ , we note that

$$\varphi_N(s) = \log M(\lambda H_{N, \lfloor sN \rfloor}) \geq \log M(-|\lambda H_{N, \lfloor sN \rfloor}|) = -\log 1/M(-|\lambda H_{N, \lfloor sN \rfloor}|). \quad (33)$$

Also, using  $M(-|\lambda H_{N, \lfloor sN \rfloor}|) < 1$  and (29) we have

$$\begin{aligned}
\log 1/M(-|\lambda H_{N, \lfloor sN \rfloor}|) &= \log \left( 1 + \frac{1 - M(-|\lambda H_{N, \lfloor sN \rfloor}|)}{M(-|\lambda H_{N, \lfloor sN \rfloor}|)} \right) \\
&\leq \frac{|1 - M(-|\lambda H_{N, \lfloor sN \rfloor}|)|}{M(-|\lambda H_{N, \lfloor sN \rfloor}|)} \leq \frac{|1 - M(-|\lambda H_{N, \lfloor sN \rfloor}|)|}{e^{\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)}} \\
&\leq \frac{\mathbb{E}|1 - e^{-|\lambda H_{N, \lfloor sN \rfloor}|}|}{e^{\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)}} \leq \frac{|\lambda H_{N, \lfloor sN \rfloor}|}{e^{\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)}} \\
&\leq e^{-\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)} \|G\|_\infty \sum_{j=\lfloor 1-sN \rfloor}^{\lfloor N+1-sN \rfloor} h_j.
\end{aligned} \tag{34}$$

So,

$$\log 1/M(-|\lambda H_{N, \lfloor sN \rfloor}|) \leq e^{-\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)} \|G\|_\infty \|h\|_1. \tag{35}$$

This bound works for all  $s$ . However, when  $s$  is away from  $[0, 1]$  the summation is over the tails of  $h_j$  and we can do better. Specifically there exists  $N_0, N_1$  such that for  $N \geq N_1 \geq N_0$  and  $s \notin [-1, 2]$

$$\sum_{j=\lfloor 1-sN \rfloor}^{\lfloor N+1-sN \rfloor} h_j \leq \sum_{j=\lfloor 1-sN_0 \rfloor}^{\lfloor N_0+1-sN_0 \rfloor} h_j.$$

Therefore, using (34), for  $s \notin [-1, 2]$ ,  $N > N_1$ ,

$$\log 1/M(-|\lambda H_{N, \lfloor sN \rfloor}|) \leq e^{-\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)} \|G\|_\infty \sum_{j=\lfloor 1-sN_0 \rfloor}^{\lfloor N_0+1-sN_0 \rfloor} h_j. \tag{36}$$

So define

$$\psi(s) := \begin{cases} -e^{-\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)} \|G\|_\infty \|h\|_1, & s \in [-1, 2] \\ -e^{-\Lambda_\beta(-\lambda \|h\|_1 \|G\|_\infty)} \|G\|_\infty \sum_{j=\lfloor 1-sN_0 \rfloor}^{\lfloor N_0+1-sN_0 \rfloor} h_j, & s \notin [-1, 2]. \end{cases}$$

Then for  $N > N_1$ , (33), (35), and (36) show that  $\varphi_N \geq \psi$ . To show  $\psi \in L^1$ , we note that

$$\begin{aligned}
\int_{s \notin [-1, 2]} |\psi(s)| ds &\leq C \int_{-\infty}^{\infty} \sum_{j=\lfloor 1-sN_0 \rfloor}^{\lfloor N_0+1-sN_0 \rfloor} h_j ds = \frac{C}{N_0} \int_{-\infty}^{\infty} \sum_{j=\lfloor 1-s \rfloor}^{\lfloor N_0+1-s \rfloor} h_j ds \\
&= C \|h\|_1 < \infty.
\end{aligned}$$

We thus obtain (32) and the proof is complete.  $\square$

Lemma 4.1 allows us to obtain a limiting Cramér functional for the convolved media.

**Lemma 4.2.** *With  $\mathbf{H}$  defined by (22) set*

$$\Lambda(\lambda) := \int_0^1 \Lambda_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s)) ds$$

Then



(i) Restricting attention to  $\varepsilon$  such that  $\varepsilon^{-1} \in \mathbb{N}$ ,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{E}_\xi e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} = \Lambda(\lambda).$$

(ii)  $\Lambda$  is lower semicontinuous.

(iii) Let  $b \in (-\infty, \infty]$  be the number satisfying  $\{\lambda : \mathbb{E} \exp\{\lambda \beta\} < \infty\} = (-\infty, b)$  or  $= (-\infty, b]$ , and let  $\mathbf{H}_\lambda$  be the value of  $\mathbf{H}(s)$  that maximizes  $\lambda \cdot \mathbf{H}(s)$ . Then

$$\{\lambda : \|h\|_1 \lambda \cdot \mathbf{H}_\lambda < b\} \subset \mathcal{D}_\Lambda \subset \{\lambda : \|h\|_1 \lambda \cdot \mathbf{H}_\lambda \leq b\},$$

with  $\Lambda'$  finite throughout  $\mathcal{D}_\Lambda^\circ$ .

(iv) If  $\Lambda$  is steep, then  $\Lambda$  is essentially-smooth.

*Proof of lemma 4.2.* Write

$$\varepsilon \log \mathbb{E}_\xi e^{\varepsilon^{-1} \lambda \cdot Z_\varepsilon} := \varepsilon \log \mathbb{E}_\xi \left\{ \exp \left\{ \varepsilon^{-1} \int_0^1 \frac{\lambda \cdot \mathbf{H}(s)}{A_\varepsilon(s)} ds \right\} \right\}.$$

Applying lemma 4.1 with  $\lambda = 1$  and  $G = \lambda \cdot \mathbf{H}$  we obtain that  $\Lambda$  is indeed the form for the limiting Cramér functional, thus proving (i).

To prove lower semi-continuity, we note that whenever  $\lambda_n \rightarrow \lambda$  and  $\Lambda(\lambda_n) \leq \alpha$  then

$$\begin{aligned} \alpha &\geq \liminf \Lambda(\lambda_n) = \liminf \int_0^1 \Lambda_\beta(\|h\|_1 \lambda_n \cdot \mathbf{H}(s)) ds \\ &\geq \int_0^1 \Lambda_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s)) ds = \Lambda(\lambda), \end{aligned}$$

by a corollary of Fatou's lemma (using the fact that the integrand is bounded below).

To show (iii), suppose first  $\|h\|_1 \lambda \cdot \mathbf{H}_\lambda < b$ . Then  $\Lambda_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s))$  is bounded and differentiable and therefore  $\Lambda(\lambda)$  is too. If on the other hand  $b < \|h\|_1 \lambda \cdot \mathbf{H}_\lambda$ , then since  $\mathbf{H}$  is continuous,  $\Lambda_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s))$  equals  $+\infty$  on a set of positive measure, hence  $\Lambda(\lambda) = \infty$ . We thus have our bounds on  $\mathcal{D}_\Lambda$  and it follows that  $\mathcal{D}_\Lambda^\circ$  is non-empty and  $\Lambda'$  exists in  $\mathcal{D}_\Lambda^\circ$ .

Lastly, (iv) follows from (ii), the assumption of steepness, (iii), and the definition of essential smoothness. □

As hinted at by lemma 4.2, steepness of  $\Lambda$  is a condition that needs extra work to check. We formulate a necessary and sufficient condition below, and then three sufficient conditions that are easy to check.

**Proposition 4.1** (Steepness criteria). *Extend  $\Lambda'_\beta$  to map  $\mathbb{R} \rightarrow [0, \infty]$  by setting  $\Lambda'_\beta(t) = \infty$  whenever  $\Lambda_\beta(t) = \infty$ . Then define  $K_i : \mathbb{R}^2 \rightarrow \mathbb{R}$  by*

$$K_1(\eta) := \int_0^x \Lambda'_\beta(F(s)\eta_1 + \eta_2) ds, \quad K_2(\eta) := \int_x^1 \Lambda'_\beta(F(s)\eta_1 + \eta_2) ds.$$

*Then  $\Lambda$  defined in lemma 4.2 is steep if and only if  $\Lambda_\beta$  is steep and for every  $\eta \in \partial \mathcal{D}_{K_i}^\circ$ ,  $K_i(\eta) = \infty$ ,  $i = 1, 2$ .*

*Moreover, with  $F_M := \max_s F(s)$ ,  $F_m := \min_s F(s)$ ,  $\Lambda$  is steep whenever one of the following sufficient conditions hold:*

1.  $\Lambda_\beta$  is finite everywhere
2.  $F$  is piecewise  $C^2$  and  $\mathcal{D}_{\Lambda_\beta} = (-\infty, b)$ .
3. Let  $\{s_1, \dots, s_n\}$  be the points where  $F(s_i) = F_M$  or  $F(s_i) = F_m$ . Then there exist neighborhoods  $N_i = (s_i - \delta, s_i + \delta) \cap \{s : F_m < F(s) < F_M\}$  such that on  $N_i$ ,  $F$  admits an expansion of the form

$$F(s) := F(s_i) + c(s - s_i)^r + R(s - s_i),$$

$$\frac{R(s - s_i)}{(s - s_i)^r} \rightarrow 0, \quad \frac{R'(s - s_i)}{(s - s_i)^{r-1}} \rightarrow 0, \quad N_i \ni s \rightarrow s_i.$$

Then with  $r := \min\{r_1, \dots, r_n\}$ ,  $\Lambda$ ,  $b = \partial\mathcal{D}_{\Lambda_\beta}$ ,  $\Lambda$  is steep if

$$\int_{b-1}^b \frac{\Lambda'_\beta(t)}{(b-t)^{(r-1)/r}} dt = \infty.$$

*Remark 4.1.* Since  $\beta \geq 0$ , we have  $\Lambda'_\beta \geq 0$ , and therefore  $K_i : [0, \infty]$  are well defined as integrals of functions taking values in  $[0, \infty]$ .

*Proof.* We note that the first condition “ $\Lambda_\beta$  is finite everywhere” trivially implies that  $\Lambda$  is steep. If  $\Lambda_\beta$  is not steep, then it is easy to construct an example showing that  $\Lambda$  is not steep either. So from now on we assume  $\Lambda_\beta$  is steep but  $\mathcal{D}_{\Lambda_\beta} \neq \mathbb{R}$ .

We now show the necessary and sufficient condition involving the  $K_i$ . Recall

$$\mathbf{H} := (F\mathbf{1}_{(0,x)}, F, \mathbf{1}_{(0,x)}, 1).$$

To show steepness we must fix  $\lambda \in \partial\mathcal{D}_\Lambda^\circ$ , let  $\mathcal{D}_\Lambda^\circ \ni \lambda^n \rightarrow \lambda$  and show  $|\nabla\Lambda(\lambda^n)| \rightarrow \infty$ .

Define the function  $\Gamma : \mathbb{R}^4 \rightarrow [0, \infty]$  by

$$\Gamma(\lambda) := \int_0^1 \Lambda'_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s)) ds.$$

We claim that  $\Lambda$  is steep if and only if  $\Gamma(\lambda) = \infty$  for all  $\lambda \in \partial\mathcal{D}_\Lambda^\circ$ . Indeed, if  $\Gamma(\lambda) < \infty$  for some  $\lambda \in \partial\mathcal{D}_\Lambda^\circ$ , then since for  $t \in (0, 1)$   $t \mapsto \Lambda'_\beta(\|h\|_1 t\lambda \cdot \mathbf{H}(s))$  is finite and non-decreasing (by convexity),

$$|\nabla\Lambda(t\lambda)| \leq \|\mathbf{H}\|_{L^\infty} \Gamma(t\lambda) \leq \|\mathbf{H}\|_{L^\infty} \Gamma(\lambda) < \infty,$$

and therefore  $\Lambda$  is not steep. On the other hand suppose  $\Gamma(\lambda) = \infty$  for all boundary points  $\lambda$ , then choose one along with a sequence  $\mathcal{D}_\Lambda^\circ \ni \lambda^n \rightarrow \lambda$ , then by Fatou

$$\liminf |\nabla\Lambda(\lambda^n)| \geq \liminf \Gamma(\lambda^n) \geq \Gamma(\lambda) = \infty.$$

Therefore  $\Lambda$  is steep.

We now show that  $\mathcal{D}_\Gamma^\circ = \mathcal{D}_\Lambda^\circ$ . Note that  $\lambda \in \mathcal{D}_\Gamma^\circ$  implies  $\Lambda'_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s))$  is bounded for all  $s \in (0, 1)$ , and  $\lambda \in (\lambda - \delta, \lambda + \delta)$  for some  $\delta > 0$ . Since  $\mathcal{D}_{\Lambda_\beta}^\circ = \mathcal{D}_{\Lambda'_\beta}^\circ$ , the same holds for  $\Lambda_\beta(\|h\|_1 \lambda \cdot \mathbf{H}(s))$ . Hence  $\mathcal{D}_\Gamma^\circ \subset \mathcal{D}_\Lambda^\circ$ . A similar argument shows  $\mathcal{D}_\Lambda^\circ \subset \mathcal{D}_\Gamma^\circ$ .

We now have  $\Lambda$  is steep if and only if  $\Gamma(\lambda) = \infty$  for all  $\lambda \in \partial\mathcal{D}_\Lambda^\circ = \partial\mathcal{D}_\Gamma^\circ$ . Our next step is to change variables to simplify this boundary. To that end, note that

$\Gamma(\lambda) = K_1(\|h\|_1(\lambda_1 + \lambda_2), \|h\|_1(\lambda_3 + \lambda_4)) + K_2(\|h\|_1(\lambda_2), \|h\|_1(\lambda_4))$ , so by a change of variables  $\eta = (\lambda_1 + \lambda_2, \lambda_3 + \lambda_4, \lambda_2, \lambda_4)$  we have

$$\Gamma(\lambda) = K_1(\|h\|_1\eta_1, \|h\|_1\eta_2) + K_2(\|h\|_1\eta_3, \|h\|_1\eta_4).$$

Changing variables again  $\eta \mapsto \eta/\|h\|_1$  and taking note of the non-negativity of the  $K_i$ , we see that  $\Lambda$  is steep if and only if  $K_i(\eta) = \infty$  for all  $\mathbb{R}^2 \ni \eta \in \partial\mathcal{D}_{K_i}$ .

Having proved the necessary and sufficient condition, we use this to show the three sufficient conditions. The first has already been shown.

As for the third sufficient condition, choose an extremal point  $\bar{s} \in (0, x)$  and assume  $\eta \in \mathcal{D}_{K_1}$ . Without loss of generality, assume we have the given expansion in the open set  $(\bar{s} - \delta, \bar{s})$ . Now  $K_1(\eta) = \infty$  if for all  $\delta > 0$

$$\int_{\bar{s}-\delta}^{\bar{s}} \Lambda'_\beta(F(s)\eta_1 + \eta_2) ds = \infty. \quad (37)$$

We shall reduce this condition to the type stated in the proposition.

Our expansion gives us

$$t(s) := F(s)\eta_1 + \eta_2 = b - c|s - \bar{s}|^r + R(s - \bar{s}),$$

for a new positive constant  $c$  and a new function  $R$  (differing from the old  $R$  by a constant). After possibly shrinking  $\delta$  we can solve for  $s(t)$ .

$$s(t) = \bar{s} - c^{-1/r}(b + R(s - \bar{s}) - t)^{1/r}.$$

Therefore, (37) holds if and only if

$$\int_{\bar{s}-\delta}^{\bar{s}} \Lambda'_\beta(F(s)\eta_1 + \eta_2) ds = \int_{t(\bar{s}-\delta)}^b \Lambda'_\beta(t)s'(t) dt = \infty. \quad (38)$$

Differentiating we have

$$s'(t) \left( 1 - \frac{R'(s - \bar{s})}{c^{1/r}(b + R(s - \bar{s}) - t)^{(r-1)/r}} \right) = \frac{1}{c^{1/r}(b - t)^{(r-1)/r}} \frac{(b - t)^{(r-1)/r}}{(b + R(s - \bar{s}) - t)^{(r-1)/r}}.$$

Noting that  $b + R(s - \bar{s}) - t = c(\bar{s} - s)^r$ , and using our hypothesis on  $R$ , we have positive  $c_1, c_2$  such that

$$\frac{c_1}{(b - t)^{(r-1)/r}} \leq s'(t) \leq \frac{c_2}{(b - t)^{(r-1)/r}}. \quad (39)$$

Due to (39), (38) is equivalent to

$$\int_{b-1}^b \frac{\Lambda'_\beta(t)}{(b - t)^{(r-1)/r}} dt = \infty.$$

Sufficient condition three then follows by considering all possible such points  $\bar{s}$ .

Sufficient condition 2 follows from 3 since given 2 we have the expansion in 3 with  $r \geq 1$  and therefore

$$\begin{aligned} \int_{b-1}^b \frac{\Lambda'_\beta(t)}{(b - t)^{(r-1)/r}} dt &\leq \int_{b-1}^b \Lambda'_\beta(t) dt = \lim_{\delta \nearrow b} \int_{b-1}^\delta \Lambda'_\beta(t) dt \\ &= \lim_{\delta \nearrow b} \Lambda_\beta(\delta) - \Lambda_\beta(b - 1) = \infty, \end{aligned}$$

in light of our assumption  $\mathcal{D}_{\Lambda_\beta} = (-\infty, b)$  and the lower-semicontinuity of  $\Lambda_\beta$ , which follows from Fatou's lemma.  $\square$

## 5 LDP approximations and generalizations

### 5.1 Rate functions for approximate solutions

To compute (naïvely) the rate functions for  $u_\varepsilon$  requires first a four-dimensional (convex) optimization to obtain  $\Lambda^*$ , and then another four dimensional optimization to obtain  $I_{u_\varepsilon}$ . Our corrector theory shows that one can rigorously approximate  $u_\varepsilon = u_0 + v_\varepsilon$  in an  $\sqrt{\varepsilon}$  neighborhood of  $u_0$ . Motivated by this we consider the rate function for  $u_0 + v_\varepsilon$ .

However, a large deviation will necessarily take us outside the  $\sqrt{\varepsilon}$  neighborhood, so more discussion is in order. Looking at terms in the expansion (6)  $u_\varepsilon = u_0 + v_\varepsilon + R_\varepsilon$  we see that when terms of the form

$$\frac{1}{\langle H_k/A_0 \rangle} \int_0^1 \frac{H_i(s)}{A_\varepsilon(s)} ds$$

are not too large, we can approximate  $u_\varepsilon \approx u_0 + v_\varepsilon$ . An exact rate function for  $u_0 + v_\varepsilon$  can then be calculated. We call this our *approximate rate function*  $\tilde{I}(\ell)$ . Note that this is equivalent to approximating  $g$  (from (22)) by some map  $\tilde{g}$  and using the contraction principle. It could be argued then that the inverse images  $g^{-1}\{\ell\} \approx \tilde{g}^{-1}\{\ell\}$  when  $\ell$  is sufficiently small (or some other better conditions). Then continuity of the rate function  $\Lambda^*(z)$  shows that  $I_{u_\varepsilon}(\ell) \approx \tilde{I}(\ell)$ . Since we can also represent  $v_\varepsilon(x)$  (for fixed  $x$ ) in terms of an integral against a single function  $G(x, s)$ , we can obtain a rate function for  $u_0 + v_\varepsilon$  without using the contraction principle. This gives us a more explicit form, and since rate functions are unique (lemma 4.4.1 [13]) these methods give the same result.

We present now the LDP for  $u_0 + v_\varepsilon$ . The proof is a simplified version of the LDP proof for  $u_\varepsilon$ .

**Proposition 5.1.** *With  $G, v_\varepsilon$  given by (7), we have:*

$$\lim \varepsilon \log \mathbb{E}_\varepsilon e^{\varepsilon^{-1} \lambda(u_0 + v_\varepsilon)} = \Lambda(\lambda).$$

Where, when  $A_\varepsilon$  is the parameterized media defined by (24), and  $V_\alpha$  given by (25),

$$\Lambda(\lambda) := \lambda u_0 + \int_0^1 \left[ -\lambda \frac{G(s)}{A_0(s)} + \Lambda(V_{a(s)}, \lambda G(s)) \right] ds,$$

and when  $A_\varepsilon$  is the convolved media defined by (28),

$$\Lambda(\lambda) := \lambda u_0 + \int_0^1 \left[ -\lambda \frac{G(s)}{A_0(s)} ds + \Lambda_\beta(\|h\|_1 \lambda G(s)) \right] ds.$$

In either case  $\Lambda$  is convex and whenever  $\Lambda$  is steep  $u_0 + v_\varepsilon$  satisfies a large deviation principle with good convex rate function

$$\tilde{I}(\ell) := \sup_{\lambda \in \mathbb{R}} [\lambda \ell - \Lambda(\lambda)].$$

Moreover, whenever  $\tilde{I}$  is finite in a neighborhood of  $\ell > \lim \mathbb{E} u_\varepsilon$  then (15) holds.

*Remark 5.1.* The steepness criteria in proposition 4.1 also apply here with  $G(x, s)$  (written  $G(s)$  when we fix  $x$ ) replacing  $F(s)$ .

## 5.2 Generalizations to incompletely characterized media

Gaussian corrector results, e.g. theorem 3.2, require knowledge of the first two moments of the random media (along with other “niceness” assumptions such as mixing and bounds on higher moments). The question arises: How much must be known about the random media for a large deviation result? Here we partially answer this question and leave a thread open for future work.

LDP results for mixing random variables are available [13, 9]. These give existence but not an explicit form for limiting Cramér functional  $\Lambda$ , and hence only existence of an LDP (instead of an explicit form). Restricting our attention to the case of parameterized media (section 4.4.1) or convolved media (section 4.4.2), we ask, “how general can the assumptions on the  $\theta_k$  or  $\beta_k$  be?” In general, we cannot expect the moment generating function of our media to be known. Notice that

$$\tilde{\Lambda} \geq \Lambda \quad \Rightarrow \quad \sup_{\lambda \in \mathbb{R}} [\lambda \ell - \tilde{\Lambda}(\lambda)] \leq \sup_{\lambda \in \mathbb{R}} [\lambda \ell - \Lambda(\lambda)]. \quad (40)$$

In other words, if we can obtain upper bounds for the moment generating function, then we can obtain a lower bound on the rate function. So a strategy would be: First, write the limiting logarithmic moment generating function  $\Lambda(\lambda) := \lim_{\varepsilon \rightarrow 0} \varepsilon \Lambda(\varepsilon^{-1} v_\varepsilon, \lambda)$  in terms of the logarithmic moment generating function of the random media (e.g.  $\Lambda_\beta$  in lemma 4.2). Second, find upper bounds for  $\Lambda_\beta$  using e.g. Bennett’s inequality (Lemma 2.4.1 in [13]) (this bounds the moment generating function of a bounded random variable in terms of its mean and variance). Third, a rigorous limiting upper bound is now available via (40).

## 6 Model Problems

Here we explore two specific examples and give numerical results. In both cases the right hand side  $f(x) \equiv 1$  for  $x \in (0.45, 0.55)$  and  $f(x) \equiv 0$  elsewhere. Hence  $F(s)$  is piecewise smooth. Therefore, using sufficient steepness condition 3 (proposition 4.1) our logarithmic moment generating functions will be steep (since  $s \mapsto G(x, s)$  is piecewise smooth).

### 6.1 Numerical results for parameterized media

Here we use a field that fits into the framework of section 4.4.1. This gives some control over the large deviations.

Let  $\xi = (\xi_0, \dots, \xi_7)$  be  $\sim_{\text{i.i.d.}} \mathcal{U}[-1, 1]$ . We then set

$$a(x, \xi) = \max \left\{ 1 + 2 \frac{1 - 0.75}{0.75} \sum_{m=0}^7 \xi_m r^m \sin[(2m + 1)\pi x], \frac{17}{32} \right\}.$$

Next, let  $\theta = (\theta_1, \dots, \theta_m, \dots)$  be an infinite collection of identically distributed independent rvs  $\theta_i \sim \mathcal{U}[-1, 1]$  (which are also independent of the  $\xi_i$ ). Put

$$b(y, \theta) = \frac{1}{2} \sum_{m=0}^{\infty} \theta_m \mathbf{1}_{m \leq y < m+1}(y).$$

In other words,  $b(y, \theta) = \theta_m$  when  $m < y < m + 1$ .

We are ensured of the ellipticity condition

$$\frac{1}{32} \leq A(x, \frac{x}{\varepsilon}) \leq \frac{7}{2}.$$

The resultant media is pictured in figure 1.

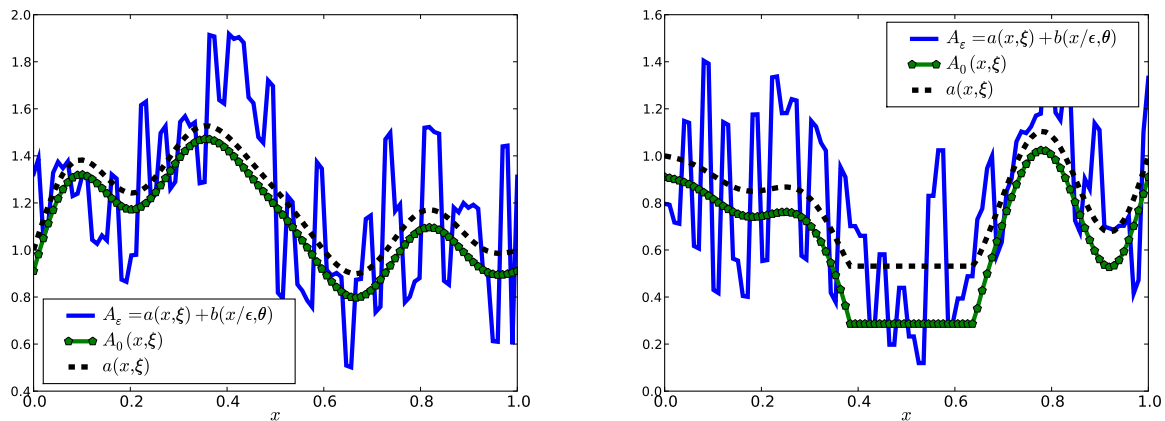


Figure 1: Comparison of two realizations of parameterized diffusion coefficients (section 6.1). The homogenized coefficient differs significantly only when  $a(x, \xi)$  is small. In both cases,  $\varepsilon = 1/50$ . On the left  $a(x, \xi)$  is such that no values of  $\theta_i \sim \mathcal{U}[-1, 1]$  bring  $A_\varepsilon$  close to zero. This realization of  $a(x, \xi)$  is quite typical and is referred to as our “mild” coefficient. On the right we have a less common “wild” coefficient.

Following as in section 4.4.1 we characterize  $V_\alpha := (\alpha + \theta/2)^{-1}$  with  $\theta \sim \mathcal{U}[-1, 1]$ . We have an explicit density for  $V_\alpha$ ,

$$\pi_{V_\alpha}(v) = \mathbf{1}_{\frac{1}{\alpha+1/2} < V_\alpha < \frac{1}{\alpha-1/2}}(v) \frac{1}{v^2}. \quad (41)$$

Hence,

$$\mathbb{E}\{V_\alpha\} := \log\left(\frac{\alpha + 1/2}{\alpha - 1/2}\right), \quad \mathbb{E}\{V_\alpha^2\} := \frac{1}{\alpha^2 - 1/4}.$$

Therefore,

$$\begin{aligned} A_0(x, \xi) &= \mathbb{E}_\xi \left\{ \frac{1}{\alpha + b(0)} \right\}^{-1} \Big|_{\alpha=a(x, \xi)} = \left[ \log\left(\frac{\alpha + 1/2}{\alpha - 1/2}\right) \right]^{-1} \Big|_{\alpha=a(x, \xi)}, \\ \text{Cov}_\alpha(\tau) &= \begin{cases} \mathbb{E}\{V_\alpha^2\} - \mathbb{E}\{V_\alpha\}^2, & 0 \leq \tau < 1 \\ 0, & \text{otherwise,} \end{cases} \\ \sigma^2(t) &= \text{Cov}_\alpha \Big|_{\alpha=a(t, \xi)}. \end{aligned}$$

These calculations are enough to give us explicit integrals defining the homogenized term  $u_0$  and the corrector  $v_\varepsilon$ . Namely,  $u_0$  solves

$$\frac{d}{dx} A_0(x) \frac{d}{dx} u_0 = f(x),$$

and the corrector is given by theorem 3.2 with  $\sigma^2(t)$  as above.

The large deviations result for  $u_\varepsilon$  is given by theorem 4.4, and our approximate rate function by 5.1. In particular, the limiting Cramér functional (theorem 4.4) is given by

$$\Lambda(\lambda) := \int_0^1 \log \int_{(a(s)+1/2)^{-1}}^{(a(s)-1/2)^{-1}} \frac{e^{\lambda \cdot \mathbf{H}(s)v}}{v^2} dv ds,$$

and for the approximate rate-function,

$$\Lambda(\lambda) := \lambda u_0 + \int_0^1 \left[ -\lambda \frac{G(s)}{A_0(s)} + \log \int_{(a(s)+1/2)^{-1}}^{(a(s)-1/2)^{-1}} \frac{e^{\lambda G(s)v}}{v^2} dv \right] ds.$$

We define empirical rate functions  $\mathcal{E}(u_\varepsilon(0.5), \ell)$ ,  $\mathcal{E}(u_0(0.5) + \sqrt{\varepsilon}v(0.5), \ell)$  (for  $u_\varepsilon(0.5)$  and  $u_0(0.5) + \sqrt{\varepsilon}v(0.5)$  respectively) as follows. With  $\{X_1, \dots, X_N\}$  a set of samples from  $W$  ( $W = u_\varepsilon$  or  $W = u_0 + \sqrt{\varepsilon}v$ ), we set

$$\begin{aligned} \widehat{W} &:= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N X_j, \\ \mathcal{E}(W, \ell) &:= \begin{cases} \varepsilon \log \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{X_j \geq \ell}, & \ell > \widehat{W} \\ \varepsilon \log \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{X_j \leq \ell}, & \ell < \widehat{W}. \end{cases} \end{aligned} \quad (42)$$

Note that  $\lim_{N \rightarrow \infty} \mathcal{E}(u_0 + \sqrt{\varepsilon}v, \ell)$  has an explicit expression, and we use this in place of (42) to compute  $\mathcal{E}(u_0 + \sqrt{\varepsilon}v, \ell)$ . Since  $X_j \geq \ell$  is a rare event we cannot compute  $\mathcal{E}$  by direct sampling. A crude importance sampling technique was used whereby the  $\theta_i \sim \mathcal{U}[-1, 1]$  were replaced by with (scaled and shifted) Bradford random variates,

$$\pi_{Brad}(\theta) = \frac{c}{2 \log(1+c) \left(1 + \frac{c}{2}(\theta + 1)\right)} \mathbf{1}_{|x| < 1}(\theta).$$

As  $c > 0$  increases, the draws  $\theta_i$  are more likely to concentrate near  $-1$ . This gives a smaller diffusion coefficient and hence larger solution. Calculation of  $\mathcal{E}$  must then be re-weighted by the factor  $\pi_{Uniform}(\theta_i)/\pi_{Brad}(\theta_i)$ . See e.g. [12, 21] for an overview of importance sampling.

In figure 1 two realizations of the parameterized media are shown. We will fix the low frequency part  $a(x, \xi)$  and study the behavior of the solution over different realizations of  $b(\frac{x}{\varepsilon}, \theta)$ . The ‘‘mild’’ medium (left) has  $a(x, \xi)$  far from zero, so no matter what  $b(\frac{x}{\varepsilon}, \theta)$  is the solution is small. The ‘‘wild’’ medium has a section of very small  $a(x, \xi)$ . In all cases, one notes that the homogenized coefficient  $A_0(x, \xi)$  differs from the low-frequency coefficient  $a(x, \xi)$  most when the medium is small. In this case,  $A_0 < a$  in an attempt to affect a large jump in the solution  $u_0$  to approximate the often large solution  $u_\varepsilon$  (although  $b(y, \theta)$  is symmetric about 0, the resultant solution  $u_\varepsilon$  is not symmetric about  $u_0$ ).

We verify theorem 3.2 in figures 3 and 4. In figure 3 one can see that the pdf of the corrected solution (at the fixed point  $x = 1/2$ ) agrees well with the true solution so long as  $\varepsilon$  is small enough. The fit is worse for the ‘‘wild’’ medium, and in particular the true pdf shows an asymmetry that the Gaussian corrector cannot have. In figure 4, one sees that when  $\varepsilon \approx 1/50$  (or smaller) the corrector captures the variance quite

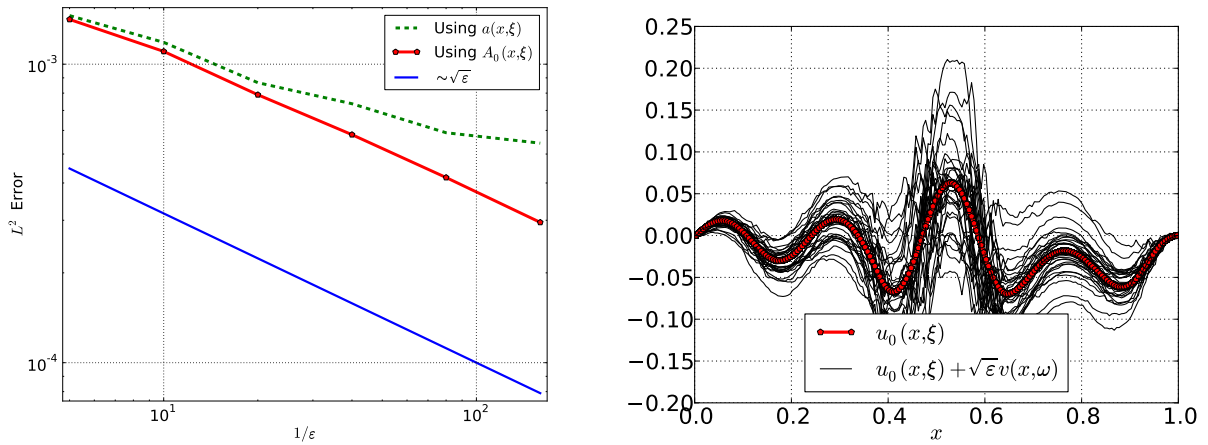


Figure 2: Left:  $L^2$  convergence of  $\mathbb{E}\{\|u_\varepsilon - u_0\|\}$  and  $\mathbb{E}\{\|u_\varepsilon - u_a\|\}$  where  $u_a$  is the result of a truncated coefficient expansion (using  $a(x, \xi)$  only). This verifies theorem 3.1. Right: Homogenized solution  $u_0(x, \xi)$  and many realization of the corrected solution  $u_0(x, \xi) + \sqrt{\varepsilon}v(x, \omega)$ . For all realizations  $\xi$  was fixed at the same value as the wild coefficient (figure 1 right)

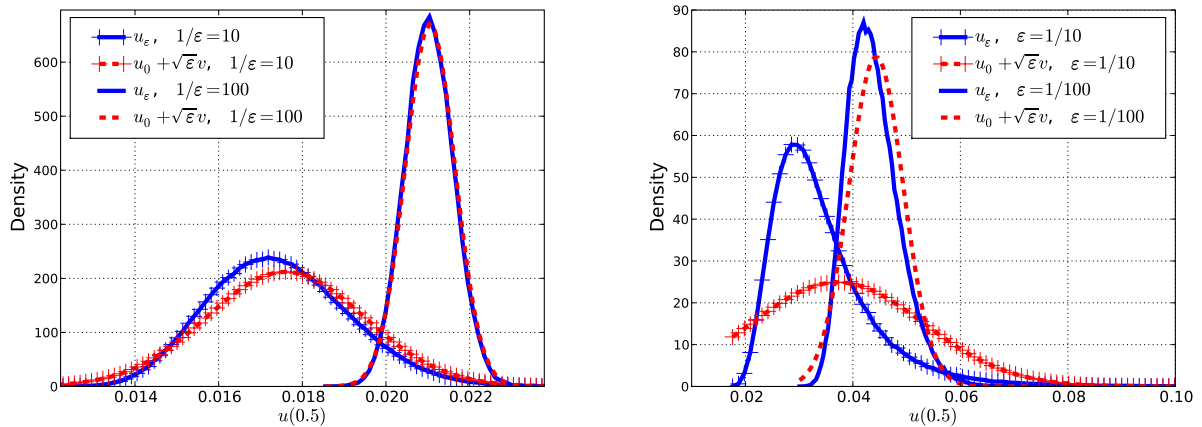


Figure 3: Left: Observed density of  $u(0.5)$  for the mild medium (figure 1 left). Plot shows that the corrector captures the bulk of the variance quite well. Right: The wild medium (figure 1 right) is shown. Results are not as good.



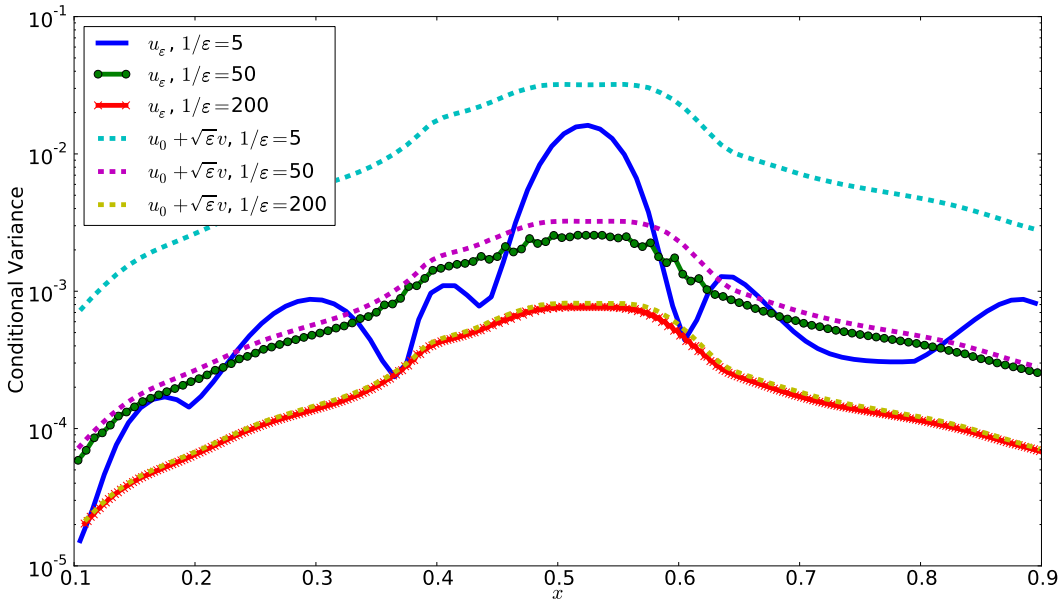


Figure 4: Verification of theorem 3.2. Conditional variance  $\mathbb{E}_\xi \{u_\varepsilon(x, \omega)^2\}$  and  $\mathbb{E}_\xi \{(u_0(x) + \sqrt{\varepsilon}v(x, \omega))^2\}$  for various  $\varepsilon$ . For all realizations  $\xi$  was fixed at the same value as the wild coefficient (figure 1 left). Plots show good agreement when  $1/\varepsilon \geq 50$ . Note also that  $\mathbb{E}_\xi \{(u_0(x) + \sqrt{\varepsilon}v(x))^2\}$  is available explicitly via (11).

well. Rate functions for the mild medium are compared in figure 5. One can see that the corrector and true rate function are almost indistinguishable until the true solution saturates around 0.35. The approximate rate function also works well up until  $u(0.5) \approx 0.3$ . In this case one could use the approximate rate function to see *a-priori* that the corrector stands a chance of capturing the large-deviation behavior well. The case is different for the wild medium LDP results in figure 6. Here one can see that the corrector rate function separates from the true rate function fairly early on. While the fit between the approximate rate function  $\tilde{I}$  and the true rate function  $\mathcal{E}(u_\varepsilon, \cdot)$  is not perfect, one could still tell, using only  $\tilde{I}$ , that the Gaussian corrector stands little chance of capturing the large deviation behavior.

It should be noted that since for the mild medium, the maximum possible value of  $u_\varepsilon(0.5)$  was approximately 0.35, we consider  $\ell = 0.03$  a large deviation. The scale is harder to set with the wild medium since our sampling could not achieve results near the maximum. However, one does note (figure 6) that by  $\ell \approx 0.6$  the empirical rate function differs from a Gaussian rate function by quite a bit. For that reason, we consider  $\ell > 0.6$  to be a large deviation. It is important to note however that the approximate rate function, being based on a linearization, does differ from the true rate function for large enough  $\ell$ .

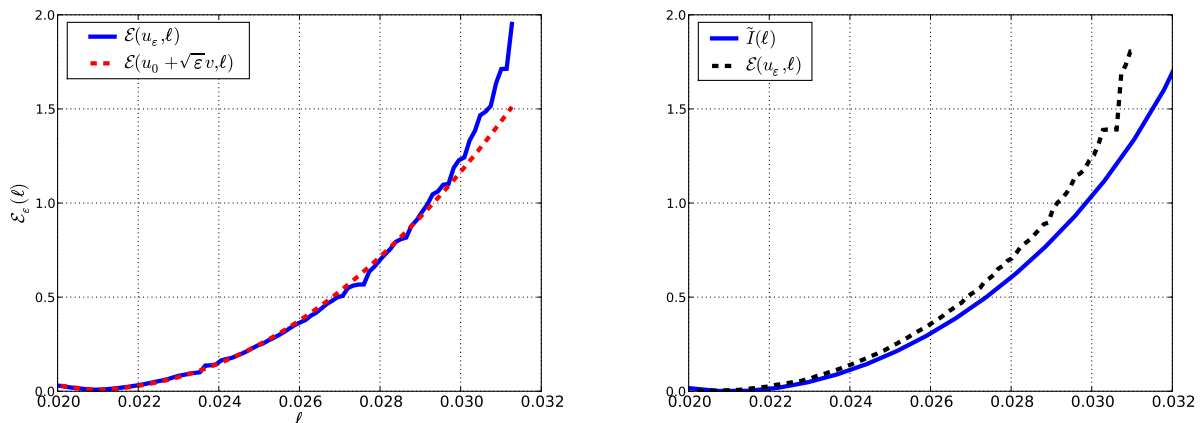


Figure 5: Left: Comparison of empirical rate function for  $u_\varepsilon$ ,  $\mathcal{E}(u_\varepsilon, \ell)$ , with the corrector rate function  $(\ell - u_0)^2/(2C_c)$  in the case of the mild medium (figure 1 left). In this case, the corrector captures the large deviation behavior well up until  $\approx 0.3$  when the true solution nears its theoretical upper bound. Right: Comparison of theoretical approximate rate function  $\tilde{I}(\ell)$  with empirical rate function  $\mathcal{E}$  with the mild medium and  $\varepsilon = 1/100$ . The approximate rate function works quite well for values  $\approx 0.30$ . Since the largest possible solution is  $u_\varepsilon(0.5) \approx 0.35$  we consider  $0.03 \sim O(1)$ .

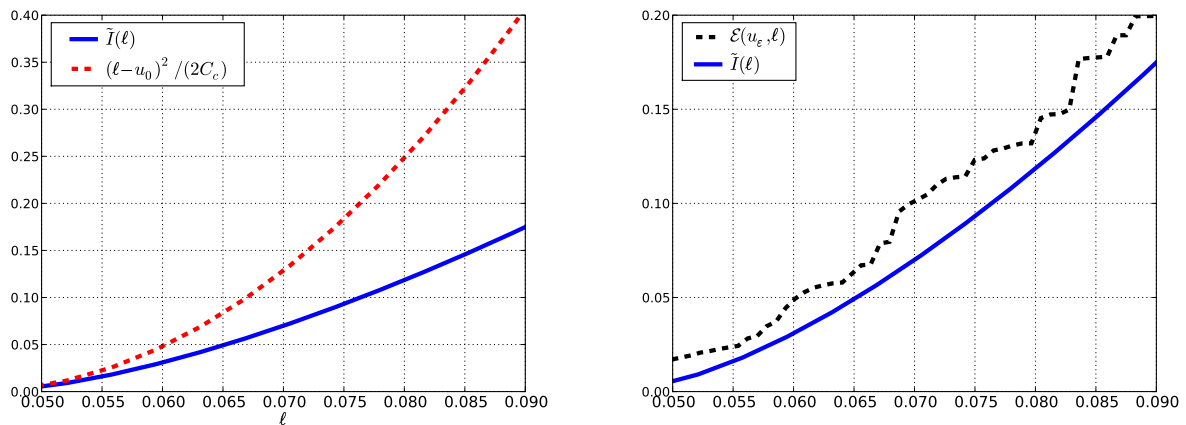


Figure 6: Left: Comparison of empirical rate function for  $u_\varepsilon$ ,  $\mathcal{E}(u_\varepsilon, \ell)$ , with the corrector rate function  $\mathcal{E}(u_0 + \sqrt{\varepsilon}v, \ell)$  in the case of the wild medium (figure 1 right). Right: Comparison of theoretical approximate rate function  $\tilde{I}(\ell)$  with empirical rate function  $\mathcal{E}(u_\varepsilon, \ell)$  with the wild medium and  $\varepsilon = 1/100$ .

## 6.2 Numerical results for convolved media

Here we implement a particular case of the media described in section 4.4.2. With  $\varepsilon^{-1} \in \mathbb{N}$ , we define

$$A_\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right), \quad \text{where} \quad \frac{1}{A(s)} := \sum_{n=1}^{\infty} \mathbf{1}_{[n-1, n)}(s) \gamma_n,$$

$$\gamma_n := \sum_{m=-\infty}^{\infty} h_{n-m} \beta_m,$$

$$h_n \geq 0, \quad \|h\|_1 := \sum_k h_k < \infty,$$

and the  $\{\beta_m\}_{m=-\infty}^{\infty}$  are i.i.d. chi-squared random variables with  $\xi$  degrees of freedom. This means

$$\beta_1 \sim \pi_\beta(\beta | \xi) = \frac{\beta^{\xi/2-1} e^{-\beta/2}}{\Gamma(\xi/2) 2^{\xi/2}}.$$

The moment generating and characteristic functions of every  $\beta_n$  are

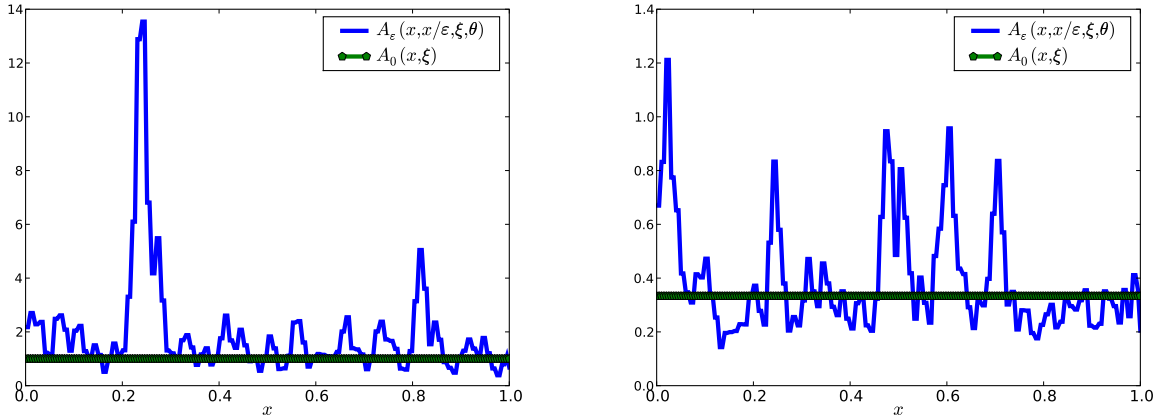


Figure 7: Typical realization of random media when  $\xi = 1$  (left) and when  $\xi = 3$  (right).

$$\mathbb{E}_\xi \left\{ e^{\lambda\beta} \right\} = \frac{1}{(1 - 2\lambda)^{\xi/2}}, \quad \mathbb{E}_\xi \left\{ e^{it\beta} \right\} = \frac{1}{(1 - 2it)^{\xi/2}}. \quad (43)$$

The random variables  $\gamma_n$  are well defined since the characteristic function  $\phi_M(t) := \prod_{|k|<M} \mathbb{E} \left\{ e^{ith_k\beta} \right\}$  has a continuous limit  $\phi(t)$ . Indeed,

$$\log \phi_M(t) = -\frac{\xi}{2} \sum_{|k|<M} \log(1 - 2ith_k).$$

This converges absolutely as can be seen using  $|\log(1 - 2ith_k)| \leq C|2th_k|$  and  $\|h\|_1 < \infty$ .

Note that  $\mathbb{E}_\xi \beta_n = \xi$ ,  $\mathbb{E}_\xi (\beta_n - \xi)^2 = 2\xi$ , so

$$\begin{aligned} \frac{1}{A_0(s)} &:= \mathbb{E}_\xi \frac{1}{A(s)} = \sum_{n=1}^{\infty} \mathbf{1}_{[n-1, n)}(s) \mathbb{E}_\xi \gamma_n \equiv \xi \|h\|_1 \\ \text{Cov}_\xi(\tau) &= \mathbb{E}_\xi \left\{ \left( \frac{1}{A(0)} - \mathbb{E} \left\{ \frac{1}{A(0)} \right\} \right) \left( \frac{1}{A(\tau)} - \mathbb{E} \left\{ \frac{1}{A(\tau)} \right\} \right) \right\} \\ &= 2\xi \sum_{k \in \mathbb{Z}} h_{-k} h_{|\tau|-k}. \end{aligned} \quad (44)$$

So the single random variable  $\xi$  defines the coarse-scale randomness. From realization to realization  $\xi$  varies with a geometric distribution (with parameter  $1/5$ ) i.e.

$$\pi_\xi(\xi) = \frac{1}{5} \left( \frac{4}{5} \right)^{\xi-1}.$$

Truncation has no meaning in this context, so we consider homogenization. We have

$$A_0^{-1} = \mathbb{E} \{ A_\varepsilon^{-1} \} (x) \equiv \xi \|h\|_1 = \text{const.}, \quad (45)$$

and then  $u_0$  is the solution to

$$-\xi \|h\|_1 \frac{d^2}{dx^2} u_0 = f(x), \quad u_0(0) = u_0(1) = 0.$$

Using (44) we obtain  $\sigma^2(t) = 2\xi \|h\|_1^2$ . The Gaussian corrector is then given by theorem 3.2. The large deviations result and rate function is given by theorem 4.5

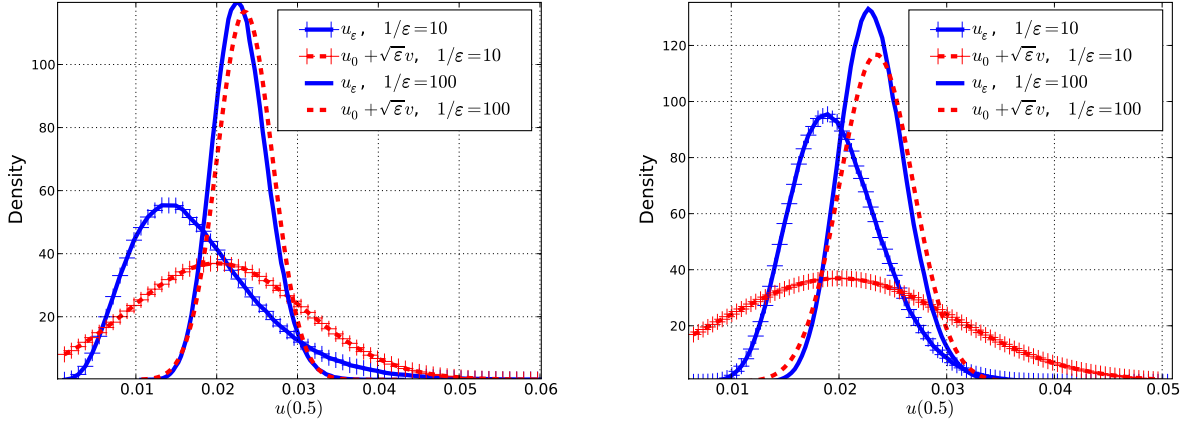


Figure 8: Pdf of  $u_\varepsilon$  vs. that of  $u_0 + \sqrt{\varepsilon}v$  for two values of  $\varepsilon$  and the convolved media. On the left,  $\kappa = 1$ , and on the right  $\kappa = 10$ . In all cases  $\xi = 1$ . Shows good agreement once  $\varepsilon$  is small enough, although agreement is worse when  $\kappa = 10$ .

with

$$\Lambda(\lambda) := \int_0^1 \frac{\xi}{2} \log \frac{1}{1 - 2\|h\|_1 \lambda \cdot \mathbf{H}(s)} ds$$

and for the approximate LDP (proposition 5.1)

$$\Lambda(\lambda) := \lambda u_0 + \int_0^1 \left[ -\lambda \frac{G(s)}{A_0(s)} + \frac{\xi}{2} \log \frac{1}{1 - 2\|h\|_1 \lambda G(s)} \right] ds.$$

Figure 7 shows typical realizations of the diffusion coefficient when the media “building block”  $\beta$  has  $\xi = 1$  or  $\xi = 3$  degrees of freedom. More degrees of freedom means larger  $u_\varepsilon$ . The behavior however is not analogous to the “mild/wild” comparison of section 6.1. In particular, the corrector captures the bulk of the distribution (moderate deviations) for all values of  $\xi$  so long as  $\varepsilon$  is small enough. For this reason we only picture pdfs for  $\xi = 1$  (figure 8). This is expected since a  $\chi_\xi^2$  random variable behaves similar to a Gaussian random variable  $\mathcal{N}(\xi, 2\xi)$  when  $\xi$  is large. The media correlation length  $\kappa$  does effect the results, and in particular  $\kappa$  must be much smaller than  $\varepsilon^{-1}$  for the corrector theory to work well. This is expected since sums of highly correlated random variables tend to a Gaussian at a slower rate than independent ones. Figures 9

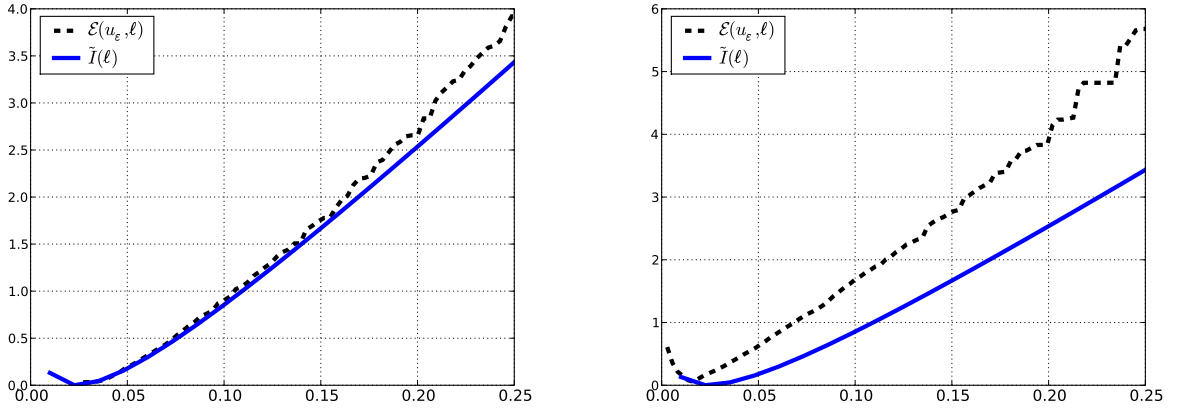


Figure 9: LDP results for the convolved media. Comparison of empirical rate function  $\mathcal{E}(u_\varepsilon, \ell)$  with theoretical (approximate) rate function  $\tilde{I}(\ell)$ . Here  $\xi = 1$ ,  $\|h\|_1 = 1$ ,  $\kappa = 1$  and  $\varepsilon = 1/100$  (left) or  $\varepsilon = 1/10$  (right).

and 10 show that the approximate rate function captures the large-deviation behavior well (for  $\ell$  not too large). As in the case of the wild medium from section 6.1 we consider  $\ell$  a large deviation if the empirical rate function  $\mathcal{E}(u_\varepsilon, \cdot)$  differs significantly from the Gaussian rate function at that point. In all cases,  $\varepsilon$  must be small enough, but when it is the match is good. Figure 11 shows that the corrector cannot capture the large deviation behavior of this media. As is the case with the parameterized media of section 6.1 the approximate rate function does a much better job than the corrector (once  $\varepsilon$  is small enough).

## 7 Proof of Homogenization/Gaussian Corrector Results

Here we prove extensions of known results.

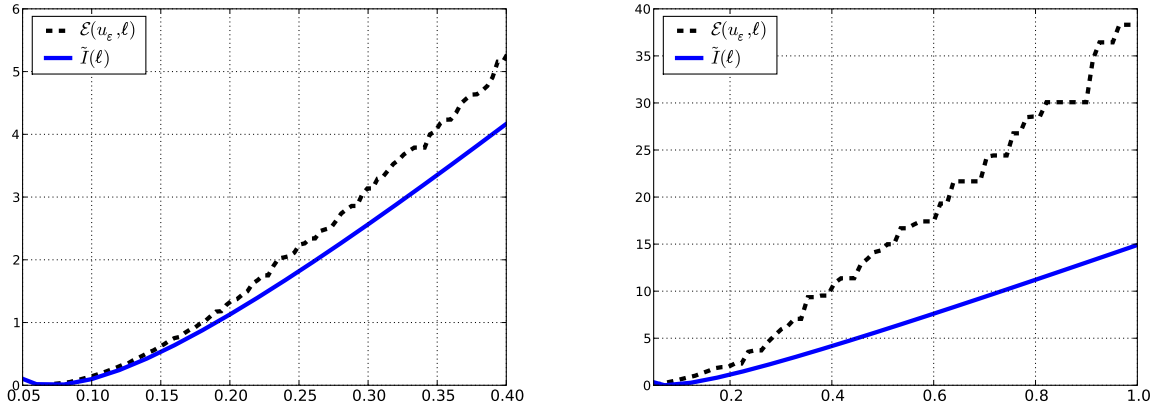


Figure 10: LDP results for the convolved media. Comparison of empirical rate function  $\mathcal{E}(u_\varepsilon, \ell)$  with theoretical (approximate) rate function  $\tilde{I}(\ell)$ . Here  $\xi = 3$ ,  $\|h\|_1 = 1$ ,  $\kappa = 1$  and  $\varepsilon = 1/100$  (left) or  $\varepsilon = 1/10$  (right). Note that the  $\varepsilon = 1/100$  plot only has  $\ell \in (0, 0.4)$  due to the difficulty in sampling the extremely rare event  $u_\varepsilon \geq 0.5$ .

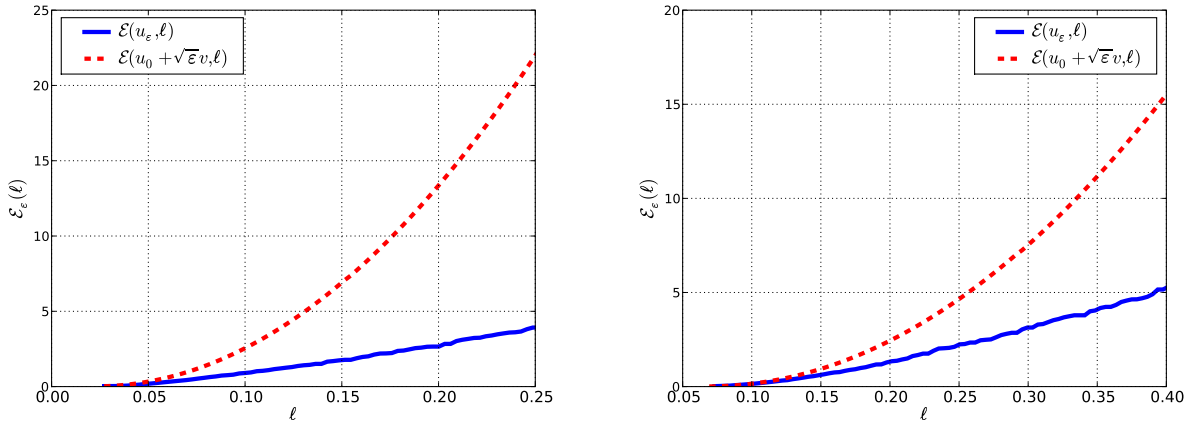


Figure 11: LDP results for the convolved media. Comparison of empirical rate function  $\mathcal{E}(u_\varepsilon, \ell)$  with empirical corrector rate function  $\mathcal{E}(u_0 + \varepsilon v, \ell)$ . Here  $\|h\|_1 = 1$ ,  $\kappa = 1$ ,  $\varepsilon = 1/100$ , and  $\xi = 1$  (left) or  $\xi = 3$  (right). In both cases the true rate function is much different than that of the corrector. The corrector's rate function performs better when  $\xi = 3$  as can be expected by comparison of the pdf for  $\chi_3^2$  and  $\chi_1^2$  random variables.

## 7.1 Proof of theorem 3.1

Here we prove one-dimensional homogenization results for our media, which has no uniform (in  $\omega$ ) ellipticity lower bound and is stationary only in the second variable.

### 7.1.1 $L^2$ convergence

Solving (1) with  $A_\varepsilon = A_\varepsilon$  and again with  $A_\varepsilon = A_0$  obtain solutions such as (5). Subtracting them we have

$$\begin{aligned} u_0(x) - u_\varepsilon(x) &= \int_0^x F(s) \left[ \frac{1}{A_\varepsilon(s)} - \frac{1}{A_0(s)} \right] ds \\ &\quad + \int_0^x \left[ \frac{\langle F/A_0 \rangle}{\langle 1/A_0 \rangle} \frac{1}{A_0(s)} - \frac{\langle F/A_\varepsilon \rangle}{\langle 1/A_\varepsilon \rangle} \frac{1}{A_\varepsilon(s)} \right] ds \\ &= I_1(x) + I_2(x). \end{aligned} \tag{46}$$

We can re-write  $I_1$  as

$$I_1(x) = \langle F \mathbf{1}_{[0,x]} / A_\varepsilon \rangle - \langle F \mathbf{1}_{[0,x]} / A_0 \rangle.$$

We are thus motivated to prove the following lemma

**Lemma 7.1.** *Let  $H \in L^\infty[0, 1]$  be deterministic, then*

$$\mathbb{E}_\xi (\langle H/A_\varepsilon \rangle - \langle H/A_0 \rangle)^2 \leq \varepsilon \|H\|_{L^\infty}^2 C_{A^{-1}}.$$

*Proof.* Write

$$\begin{aligned} &\mathbb{E}_\xi (\langle H/A_\varepsilon \rangle - \langle H/A_0 \rangle)^2 \\ &= \varepsilon^2 \int_0^{\varepsilon^{-1}} \int_0^{\varepsilon^{-1}} dt ds H(s) H(t) \mathbb{E}_\xi \left\{ \left( \frac{1}{A(s\varepsilon, s)} - \mathbb{E}_\xi \{1/A\}(s\varepsilon) \right) \left( \frac{1}{A(t\varepsilon, t)} - \mathbb{E}_\xi \{1/A\}(t\varepsilon) \right) \right\} \\ &= \varepsilon^2 \int_0^{\varepsilon^{-1}} \int_0^{\varepsilon^{-1}} H(s) H(t) \text{Cov}_\xi(s\varepsilon, t\varepsilon, s-t) dt ds \leq \varepsilon \|H\|_{L^\infty}^2 C_{A^{-1}}. \end{aligned}$$

□

Since  $\|F\|_{L^\infty} \leq \|f\|_{L^2}$ , lemma 7.1 gives us

$$\mathbb{E}_\xi \|I_1\|_{L^2}^2 \leq \varepsilon \|f\|_{L^2}^2 C_{A^{-1}}, \tag{47}$$

We now consider  $I_2$ . After repeated use of the equality  $ab - \tilde{a}\tilde{b} = (a - \tilde{a})b + \tilde{a}(b - \tilde{b})$  we obtain

$$\begin{aligned} I_2(x) &= (\langle F/A_0 \rangle - \langle F/A_\varepsilon \rangle) \left( \int_0^x \frac{1}{A_0(s) \langle 1/A_0 \rangle} ds \right) \\ &\quad + (\langle 1/A_\varepsilon \rangle - \langle 1/A_0 \rangle) \left( \frac{1}{\langle 1/A_0 \rangle \langle 1/A_\varepsilon \rangle} \int_0^x \frac{\langle F/A_\varepsilon \rangle}{A_0} ds \right) \\ &\quad + \left( \int_0^x \left[ \frac{1}{A_0(s)} - \frac{1}{A_\varepsilon(s)} \right] ds \right) \left( \frac{\langle F/A_\varepsilon \rangle}{\langle 1/A_\varepsilon \rangle} \right) \\ &= I_{21}(x) + I_{22}(x) + I_{23}(x). \end{aligned}$$

Each term is the product of a term similar to  $I_1$  and a bounded random variable (*a priori* bounds obtained using the positivity of  $A$ ). We thus obtain

$$\begin{aligned} \sqrt{\mathbb{E}_\xi \|I_2\|_{L^2}^2} &\leq \sqrt{\mathbb{E}_\xi \|I_{21}\|_{L^2}^2} + \sqrt{\mathbb{E}_\xi \|I_{22}\|_{L^2}^2} + \sqrt{\mathbb{E}_\xi \|I_{23}\|_{L^2}^2} \\ &\leq 3\sqrt{\varepsilon} \|f\|_{L^2} \sqrt{C_{A^{-1}}}. \end{aligned} \tag{48}$$

The inequality in theorem 3.1 thus follows from (46), (47), and (48).

### 7.1.2 *a.s.* convergence

Here we use the same decomposition,

$$u_\varepsilon(x) - u_0(x) = I_1(x) + I_{21}(x) + I_{22}(x) + I_{23}(x),$$

and show that each term goes to zero *a.s.* To that end, we notice that every term is the product of a bounded (sometimes random) variable, and a term like  $\langle H/A_\varepsilon \rangle - \langle H/A_0 \rangle$  (with  $H$  depending on  $x$ ). It will thus suffice to prove *a.s.* convergence of this latter term. We thus obtain pointwise (in  $x$ ) *a.s.* convergence. The *a.s.* norm convergence then follows from an *a-priori* bound on every realization of  $u_\varepsilon - u_0$  and the bounded convergence theorem. The *a-priori* bound follows from (8) (which we assume here is independent of  $\varepsilon$ ) and (46). In other words, *a.s.* convergence in theorem 3.1 is a corollary of the following lemma.

**Lemma 7.2.** *Suppose  $H \in L^\infty$  is deterministic, then as  $\tau \rightarrow \infty$ , we have (almost surely  $P_\xi$ )*

$$\langle H/A_{\tau^{-1}} \rangle - \langle H/A_0 \rangle \rightarrow 0.$$

*Proof.* The proof is more-or-less a standard trick where we show *a.s.* convergence on a sequence of  $\tau$  values as well as the difference between the sequence values and “nearby” values. See section 37.7 in [18].

We handle the first convergent term first. Defining

$$\begin{aligned} Y(\tau) &:= \langle H/A_{\tau^{-1}} \rangle - \langle H/A_0 \rangle \\ X(t, s) &:= H(t) \left( \frac{1}{A(t, s)} - \mathbb{E}_\xi \{1/A\}(t) \right), \end{aligned}$$

we have

$$\begin{aligned} Y(\tau) &= \int_0^1 H(s) \left( \frac{1}{A(s, s\tau)} - \mathbb{E}_\xi \{1/A\}(s) \right) ds \\ &= \frac{1}{\tau} \int_0^\tau H(s/\tau) \left( \frac{1}{A(s/\tau, s)} - \mathbb{E}_\xi \{1/A\}(s/\tau) \right) ds = \frac{1}{\tau} \int_0^\tau X(s/\tau, s) ds. \end{aligned}$$

Now for  $m \in \mathbb{N}$ ,  $a > 0$ , and  $m^a \leq \tau < (m+1)^a$ ,

$$\frac{\tau}{m^a} Y(\tau) = Y(m^a) + Z(m^a, \tau), \quad Z(m^a, \tau) := \frac{1}{m^a} \int_{m^a}^\tau X(s/\tau, s) ds.$$

Directly from lemma 7.1 we have

$$\mathbb{E}_\xi \{Y(m^a)^2\} \leq \|H\|_{L^\infty}^2 C_{A^{-1}} m^{-a},$$

and then Chebyshev inequality gives us, for any  $\delta > 0$ ,

$$P_\xi[|Y(m^a)| > \delta] \leq \frac{\mathbb{E}_\xi \{Y(m^a)^2\}}{\delta^2} \leq \frac{\|H\|_{L^\infty}^2 C_{A^{-1}} m^{-a}}{\delta^2}.$$



We choose  $a > 1$  and then

$$\sum_{m=1}^{\infty} \mathbb{P}_{\xi} [|Y(m^a)| > \delta] < \infty.$$

So by the Borel-Cantelli lemma,  $Y(m^a) \rightarrow 0 \mathbb{P}_{\xi}$  *a.s.*

As for  $Z(m^a, \tau)$ , we set

$$U(m^a) := \sup_{m^a \leq \tau < (m+1)^a} |Z(m^a, \tau)| \leq \frac{1}{m^a} \int_{m^a}^{(m+1)^a} |X(s/\tau, s)| ds,$$

and note that

$$\begin{aligned} \mathbb{E}_{\xi} |U(m^a)|^2 &\leq \frac{\|H\|_{\infty}^2}{m^{2a}} \int_{m^a}^{(m+1)^a} \int_{m^a}^{(m+1)^a} C(s-s') ds ds' \\ &\leq \|H\|_{\infty}^2 C_{A^{-1}} \frac{(m+1)^a - m^a}{m^a} \leq \|H\|_{\infty}^2 C_{A^{-1}} \frac{[a+1]!}{m}. \end{aligned}$$

Therefore  $\mathbb{E}_{\xi} \{U(m^a)^2\} \lesssim m^{-2}$ , and by Chebyshev's inequality and the Borel-Cantelli lemma we obtain  $U(m^a) \rightarrow 0 \mathbb{P}_{\xi}$  *a.s.* The same conclusion thus holds for  $Y(\tau)$ ,  $m^a \leq \tau < (m+1)^a$ , and therefore for  $Y(\tau) = \langle H/A_{\tau^{-1}} \rangle - \langle H/A_0 \rangle$ .  $\square$

## 7.2 Proof of theorem 3.2

Here we prove one-dimensional corrector results for our media, which has no uniform (in  $\omega$ ) ellipticity lower bound and is stationary only in the second variable.

### 7.2.1 One dimensional oscillatory integral

Here we study the integral

$$v_{\varepsilon}(x) := \int_0^1 G(x, t) q_{\varepsilon}(t) dt, \quad q_{\varepsilon}(t) := \frac{1}{A(t, \frac{t}{\varepsilon})} - \mathbb{E}_{\xi} \{1/A\}(t),$$

where  $G(x, s)$  is deterministic, piecewise continuous in  $s$ , and uniformly (in  $s$ ) Lipschitz in  $x$ .

First note that

$$\begin{aligned} \mathbb{E}_{\xi} \left( \frac{v_{\varepsilon}(x)}{\sqrt{\varepsilon}} \right)^2 &= \int_0^1 \int_0^1 G(x, t) G(x, s) \frac{\text{Cov}_{\xi}(t, s, \frac{t-s}{\varepsilon})}{\varepsilon} ds dt \\ &= \int_0^1 G(x, s) \left[ \int_{-s/\varepsilon}^{(1-s)/\varepsilon} G(x, \varepsilon t + s) \text{Cov}_{\xi}(\varepsilon t + s, s, t) dt \right] ds. \end{aligned}$$

Using (9) and dominated convergence, we therefore have

$$\mathbb{E}_{\xi} \left( \frac{v_{\varepsilon}}{\sqrt{\varepsilon}} \right)^2 \rightarrow \int_0^1 G(x, t)^2 \sigma^2(t) dt, \quad \sigma^2(t) := \int_{-\infty}^{\infty} \text{Cov}_{\xi}(t, t, q) dq. \quad (49)$$

The scaling (in  $\varepsilon$ ) and the fact that  $q_{\varepsilon}$  is mean zero indicate that a central-limit type result should show convergence to a Gaussian random variable. This is indeed the case.

**Lemma 7.3.** *If  $G(x, s)$  is deterministic, piecewise continuous in  $s$ , and uniformly (in  $s$ ) Lipschitz in  $x$ , then*

$$\frac{1}{\sqrt{\varepsilon}}v_\varepsilon(x) = \frac{1}{\sqrt{\varepsilon}} \int_0^1 G(x, t)q_\varepsilon(t) dt \xrightarrow{dist.} \int_0^1 G(x, t)\sigma(t) dW_t,$$

where  $W_t$  is a standard Brownian motion.

The following result allows us to reduce the problem of proving convergence (of a stochastic process) to one of studying finite dimensional distributions [5].

**Proposition 7.1.** *Suppose  $(Z_n; 1 \leq n \leq \infty)$  are random variables with values in the space of continuous functions  $C([0, 1])$ . Then  $Z_n$  converges in distribution to  $Z_\infty$  provided that:*

- (a) *any finite-dimensional joint distribution  $(Z_n(x_1), \dots, Z_n(x_k))$  converges to the joint distribution  $(Z_\infty(x_1), \dots, Z_\infty(x_k))$  as  $n \rightarrow \infty$ .*
- (b)  *$(Z_n)$  is a tight sequence of random variables. A sufficient condition for tightness of  $(Z_n)$  is the following Kolmogorov criterion: There exist positive constants  $\nu$ ,  $\beta$  and  $\delta$  such that*
  - (i)  $\sup_{n \geq 1} \mathbb{E} \{|Z_n(t)|^\nu\} < \infty$ , for some  $t \in [0, 1]$ ,
  - (ii)  $\mathbb{E} \{|Z_n(s) - Z_n(t)|^\beta\} \lesssim |t - s|^{1+\delta}$ ,*uniformly in  $n \geq 1$  and  $t, s \in [0, 1]$ .*

Tightness is easily verified. Indeed, (9) and (49) show that

$$\begin{aligned} \mathbb{E}_\xi |v_\varepsilon(x) - v_\varepsilon(y)|^2 &= \mathbb{E}_\xi \left( \int_0^1 [G(x, t) - G(y, t)]q_\varepsilon(t) dt \right)^2 \\ &\rightarrow \int_0^1 [G(x, t) - G(y, t)]^2 \sigma^2(t) dt \leq C|x - y|^2. \end{aligned}$$

so condition (b-i) is met with  $\nu = 2$ , and (b-ii) is met with  $\beta = 2$  and  $\delta = 1$ .

This means that to prove the theorem we simply need to fix  $(x_1, \dots, x_n)$  and show

$$\begin{aligned} \varepsilon^{-1/2}\vec{v}_\varepsilon &:= \varepsilon^{-1/2}(v_\varepsilon(x_1), \dots, v_\varepsilon(x_n)) \xrightarrow{dist.} \vec{v} := (v(x_1), \dots, v(x_n)), \\ v(x) &:= \int_0^1 G(x, t)\sigma(t) dW_t. \end{aligned} \tag{50}$$

*Proof.* Any finite dimensional distribution  $\varepsilon^{-1/2}\vec{v}_\varepsilon$  has characteristic function

$$\Phi_\varepsilon(k) := \mathbb{E}_\xi \left\{ e^{i \sum_{j=1}^n k_j \varepsilon^{-1/2} v_\varepsilon(x_j)} \right\}, \quad k = (k_1, \dots, k_n).$$

The above characteristic function may be recast as

$$\Phi_\varepsilon(k) = \mathbb{E}_\xi \left\{ e^{i \int_0^1 m(s) \frac{1}{\sqrt{\varepsilon}} q_\varepsilon(s) ds} \right\}, \quad m(s) := \sum_{j=1}^n k_j G(x_j, s).$$

As a consequence, convergence of the finite dimensional distributions will be proved if we can show convergence of

$$I_\varepsilon := \int_0^1 m(s) \frac{1}{\sqrt{\varepsilon}} q_\varepsilon(s) ds \xrightarrow{dist.} N_1 := \int_0^1 m(s) \sigma(s) dW_s, \quad (51)$$

for piecewise continuous  $m$ .

Proceeding, we now show that  $I_\varepsilon \xrightarrow{dist.} N_1$ . We first consider the case of constant  $m \equiv 1$  and  $q_\varepsilon(s) = q(s, \frac{s}{\varepsilon}) = q(\frac{s}{\varepsilon})$ . To that end, set

$$q_k := \int_{(k-1)\varepsilon}^{k\varepsilon} \frac{1}{\varepsilon} q\left(\frac{s}{\varepsilon}\right) ds = \int_{k-1}^k q(s) ds, \quad k = 1, \dots, \lfloor \varepsilon^{-1} \rfloor.$$

Then  $I_\varepsilon(x) = \sum_{k=1}^{\lfloor \varepsilon^{-1} \rfloor} q_k + R(\varepsilon)$ , where  $R(\varepsilon) \rightarrow 0$  in  $L^2$  (and therefore in probability), and hence may be ignored. We therefore consider the limit

$$\sqrt{\varepsilon} \sum_{k=1}^{\lfloor \varepsilon^{-1} \rfloor} q_k = \frac{1}{\sqrt{\lfloor \varepsilon^{-1} \rfloor}} \sum_{k=1}^{\lfloor \varepsilon^{-1} \rfloor} q_k + R'(\varepsilon), \quad \|R'(\varepsilon)\|_{L^2(\Omega)} \lesssim \varepsilon.$$

So we ignore  $R'$  and consider the limit of the summation. Following theorem 19.2 in [5], we define the sigma fields  $\mathcal{F}_n, \mathcal{F}^n$  generated by  $\{q_k : k \leq n\}, \{q_k : k \geq n\}$  respectively and set

$$\rho_n := \sup \{ |\mathbb{E} \{ \eta_1 \eta_2 \} | : \eta_1 \in \mathcal{F}_1, \eta_2 \in \mathcal{F}^n, \mathbb{E} \{ \eta_i \} = 0, \mathbb{E} \{ \eta_i^2 \} = 1, i = 1, 2 \}.$$

Then, so long as  $\sum_{n=1}^{\infty} \rho_n < \infty$ ,

$$\frac{1}{\sqrt{\lfloor \varepsilon^{-1} \rfloor}} \sum_{k=1}^{\lfloor \varepsilon^{-1} \rfloor} q_k \xrightarrow{dist.} \mathcal{N}(0, \nu^2), \quad \nu^2 := \mathbb{E} \{ q_1^2 \} + 2 \sum_{n=2}^{\infty} \mathbb{E} \{ q_1 q_n \}.$$

Since  $\rho_n \leq \varphi(n-1)$ , the summability condition on  $\rho_n$  is implied by assumptions 3.1. We now show that  $\nu^2 = \sigma^2$ . Indeed,

$$\mathbb{E} \{ q_1^2 \} = \mathbb{E} \left\{ \left[ \int_0^1 q(s) ds \right]^2 \right\} = \int_0^1 \int_0^1 \text{Cov}_\xi(s-s') ds ds'.$$

Also, using the symmetry of  $\text{Cov}_\xi$ , we have

$$\sum_{n=2}^{\infty} \mathbb{E} \{ q_1 q_n \} = \int_0^1 \int_1^\infty \text{Cov}_\xi(s-s') ds ds' = \int_0^1 \int_{-\infty}^0 \text{Cov}_\xi(s-s') ds ds'.$$

Therefore,

$$\nu^2 = \int_0^1 \int_{-\infty}^\infty \text{Cov}_\xi(s-s') ds ds' = \sigma^2.$$

This shows (51).

To prove the theorem in the case of non-constant  $m(s)$ , and  $q = q(s, \frac{s}{\varepsilon})$  we note that if  $m$  is replaced by  $m_h$ , and  $q$  is replaced by  $q_h$ , giving us  $I_\varepsilon^h$ , then

$$\begin{aligned} \mathbb{E}_\xi |I_\varepsilon^h(x) - I_\varepsilon(x)|^2 &= \frac{1}{\varepsilon} \mathbb{E}_\xi \int_0^1 \int_0^1 [m(s)q(s, \frac{s}{\varepsilon}) - m_h(s)q_h(s, \frac{s}{\varepsilon})] \\ &\quad \times [m(t)q(t, \frac{t}{\varepsilon}) - m_h(t)q_h(t, \frac{t}{\varepsilon})] ds dt. \end{aligned} \quad (52)$$

We will choose  $m_h, q_h$  such that the above expectation vanishes in the limit. Split  $[0, 1]$  into subintervals of size  $h$  (with  $1/h \in \mathbb{N}$ ) with endpoints  $t_j := jh, j = 1, \dots, 1/h$ . Now set

$$m_h(t) := m(t_j), \text{ and } q_h(t, \frac{t}{\varepsilon}) := q(t_j, \frac{t}{\varepsilon}), \text{ for } t \in [t_j, t_{j+1}).$$

Taking expectation inside of the integral (52), changing  $t \mapsto \varepsilon t$ , then  $t \mapsto t + s/\varepsilon$  we have (similar to (49))

$$\mathbb{E}_\xi |I_\varepsilon^h - I_\varepsilon|^2 = \int_0^1 \int_{-\infty}^{\infty} Q_{\varepsilon, h}(s, t) dt ds,$$

where  $|Q_{\varepsilon, h}(s, t)| \leq C\Gamma(t)$  (with  $\Gamma$  from (9)). Using the continuity (on the diagonal) of  $\text{Cov}_\xi$ , and the piecewise continuity of  $G$ , we also have  $Q_{\varepsilon, h} \rightarrow 0$  as  $(\varepsilon, h) \rightarrow (0, 0)$  (in any way whatsoever, for *a.e.*  $(s, t)$ ). Therefore, it suffices to prove the lemma 7.3 with  $m_h, q_h$  replacing  $m, q$ .

We proceed to split the integral defining  $I_\varepsilon$  up.

$$I_\varepsilon := \int_0^1 m_h(t)q_h(t, \frac{t}{\varepsilon}) dt = \sum_{j=0}^{1/h-1} \int_{[t_j, t_{j+1})} m(t_j)q(t_j, \frac{t}{\varepsilon}) dt.$$

Each subintegral is handled exactly as before, yielding

$$m(t_j) \int_{[t_j, t_{j+1})} q(t_j, \frac{t}{\varepsilon}) dt \xrightarrow{\text{dist.}} \int_{[t_j, t_{j+1})} m(t_j)\sigma(t_j) dW_t.$$

It remains then to show that the limiting Gaussians (above) are independent. We do this in the case where  $I_\varepsilon = I_{\varepsilon, 1} + I_{\varepsilon, 2}$  is split into two intervals, the general case following by induction. To that end, we show that for all  $k_1, k_2 \in \mathbb{R}$

$$\mathcal{E} := \left| \mathbb{E}_\xi \left\{ e^{i(k_1 I_{1, \varepsilon} + k_2 I_{2, \varepsilon})} \right\} - \mathbb{E}_\xi \left\{ e^{ik_1 I_{1, \varepsilon}} \right\} \mathbb{E}_\xi \left\{ e^{ik_2 I_{2, \varepsilon}} \right\} \right| \rightarrow 0. \quad (53)$$

Define

$$\begin{aligned} P_{\varepsilon, 1}^\eta &:= \int_\eta^{x-\eta} \frac{1}{\sqrt{\varepsilon}} q\left(\frac{t}{\varepsilon}\right) dt, & Q_{\varepsilon, 1}^\eta &:= I_{\varepsilon, 1} - P_{\varepsilon, 1}^\eta, \\ P_{\varepsilon, 2}^\eta &:= \int_{x+\eta}^{y-\eta} \frac{1}{\sqrt{\varepsilon}} q\left(\frac{t}{\varepsilon}\right) dt, & Q_{\varepsilon, 2}^\eta &:= I_{\varepsilon, 2} - P_{\varepsilon, 2}^\eta. \end{aligned}$$

Then the first term in  $\mathcal{E}$  decomposes as

$$\begin{aligned} \mathbb{E}_\xi \left\{ e^{i(k_1 I_{\varepsilon, 1} + k_2 I_{\varepsilon, 2})} \right\} &= \mathbb{E}_\xi \left\{ (e^{ik_1 Q_{\varepsilon, 1}^\eta} - 1) e^{ik_1 P_{\varepsilon, 1}^\eta + ik_2 I_{\varepsilon, 2}} \right\} \\ &\quad + \mathbb{E}_\xi \left\{ e^{ik_1 P_{\varepsilon, 1}^\eta + ik_2 I_{\varepsilon, 2}} \right\}. \end{aligned}$$

The first term in the decomposition is small since

$$\left| \mathbb{E}_\xi \left\{ (e^{ik_1 Q_{\varepsilon,1}^\eta} - 1) e^{ik_1 P_{\varepsilon,1}^\eta + ik_2 I_{\varepsilon,2}} \right\} \right| \leq \left( \mathbb{E}_\xi |e^{ik_1 Q_{\varepsilon,1}^\eta} - 1|^2 \right)^{1/2} \lesssim \eta^{1/2}. \quad (54)$$

As for the second term,

$$\begin{aligned} & \mathbb{E}_\xi \left\{ e^{ik_1 P_{\varepsilon,1}^\eta + ik_2 I_{\varepsilon,2}} \right\} - \mathbb{E}_\xi \left\{ e^{ik_1 I_{\varepsilon,1}} \right\} \mathbb{E}_\xi \left\{ e^{ik_2 I_{\varepsilon,2}} \right\} \\ &= \left[ \mathbb{E}_\xi \left\{ e^{ik_1 P_{\varepsilon,1}^\eta + ik_2 I_{\varepsilon,2}} \right\} - \mathbb{E}_\xi \left\{ e^{ik_1 P_{\varepsilon,1}^\eta} \right\} \mathbb{E}_\xi \left\{ e^{ik_2 I_{\varepsilon,2}} \right\} \right] \\ & \quad - \left[ \mathbb{E}_\xi \left\{ (e^{ik_1 Q_{\varepsilon,1}^\eta} - 1) e^{ik_1 P_{\varepsilon,1}^\eta} \right\} \mathbb{E}_\xi \left\{ e^{ik_2 I_{\varepsilon,2}} \right\} \right]. \end{aligned}$$

The first bracketed term is  $\lesssim \varphi(\eta/\varepsilon)$  by our mixing condition (3.1), and the second is  $\lesssim \eta^{1/2}$  in a manner similar to (54). Therefore,

$$\mathcal{E} \lesssim \left( \varphi \left( \frac{2\eta}{\varepsilon} \right) + \eta^{1/2} \right) \rightarrow 0, \text{ as } \varepsilon \rightarrow 0$$

for say  $\eta = \varepsilon^{1/2}$ .

We have thus shown

$$I_\varepsilon \xrightarrow{\text{dist.}} \int_0^1 m_h(t) \sigma_h(t) dW_t,$$

and this completes the proof.  $\square$

We also prove a complementary lemma that allows us to deal with the lack of a uniform (in  $\omega$ ) ellipticity lower bound.

**Lemma 7.4.** *Suppose  $A$  satisfies the hypothesis of theorem 3.2. Then with*

$$q\left(s, \frac{s}{\varepsilon}\right) := \frac{1}{A\left(s, \frac{s}{\varepsilon}\right)} - \frac{1}{A_0(s)},$$

we have for any  $H \in L^\infty([0, 1])$ ,

$$\begin{aligned} \mathcal{E} &:= \mathbb{E}_\xi \left\{ \left( \int_0^1 H(s) q\left(s, \frac{s}{\varepsilon}\right) ds \right)^6 \right\} \leq C_\xi \left( \int_0^1 \int_0^1 H(s) H(t) \rho^{1/3} \left( \frac{s-t}{\varepsilon} \right) ds dt \right)^3 \\ &\leq \varepsilon^3 C_\xi \|H\|_\infty^6 \|\rho^{1/3}\|_{L^1}. \end{aligned}$$

*Proof.*

$$\mathcal{E} \leq \int_{[0,1]^6} |H(s_1) \cdots H(s_6)| \left| \mathbb{E}_\xi \left\{ q\left(s_1, \frac{s_1}{\varepsilon}\right) \cdots q\left(s_6, \frac{s_6}{\varepsilon}\right) \right\} \right| ds.$$

The expectation can be broken up using Cauchy-Schwartz into a product that looks like

$$\left( \mathbb{E}_\xi \left\{ X_1^3 X_2^3 \right\} \mathbb{E}_\xi \left\{ X_3^3 X_4^3 \right\} \mathbb{E}_\xi \left\{ X_5^3 X_6^3 \right\} \right)^{1/3}.$$

For each term  $\mathbb{E}_\xi \left\{ X_i^3 X_{i+1}^3 \right\}$ , the mixing condition, and our bound on  $6^{th}$  moments gives

$$\mathbb{E}_\xi \left\{ q\left(s, \frac{s}{\varepsilon}\right)^3 q\left(t, \frac{t}{\varepsilon}\right)^3 \right\} \leq \rho\left(\frac{s-t}{\varepsilon}\right) C_\xi.$$

The result follows.  $\square$

## 7.2.2 Asymptotic expansion of the solution $u_\varepsilon$

We now complete the proof of theorem 3.2. Starting from (6), (7) we write

$$\frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon}} = \frac{1}{\sqrt{\varepsilon}}v_\varepsilon(x) + \frac{1}{\sqrt{\varepsilon}}R_\varepsilon(x).$$

The convergence of  $\varepsilon^{-1/2}v_\varepsilon$  is assured by lemma 7.3. We now show that  $\varepsilon^{-1/2}R_\varepsilon$  is tight and converges pointwise to zero. Then, proposition 7.1 shows that  $\varepsilon^{-1/2}R_\varepsilon \xrightarrow{dist.} 0$  in the space of continuous paths.

To that end we write  $\varepsilon^{-1/2}R(x)$  as a sum of terms of the form

$$Z \int_0^x B_\varepsilon(s) ds, \quad B_\varepsilon(s) := \frac{1}{\sqrt{\varepsilon}} \left( \frac{1}{A_\varepsilon(s)} - \frac{1}{A_0(s)} \right).$$

for a constant random variable  $Z$ , with either

$$|Z| \leq C'_\xi (X_1^\varepsilon)^2, \quad |Z| \leq C'_\xi |X_1^\varepsilon Y_1^\varepsilon|, \quad |Z| \leq C'_\xi |X_1^\varepsilon| \langle 1/A_\varepsilon \rangle, \quad \text{or} \quad |Z| \leq C'_\xi Y_1^\varepsilon.$$

We first show that  $\mathbb{E}_\xi \{ (\varepsilon^{-1/2}(R_\varepsilon(y) - R_\varepsilon(x)))^3 \} \lesssim \varepsilon |y - x|^2$ , meeting condition (b-ii) of proposition 7.1 with  $\beta = 3$ ,  $\delta = 1$ . Then choosing  $x = 0$ , we have (since  $R_\varepsilon(0) = 0$ )  $\varepsilon^{-1/2}R_\varepsilon(y) \rightarrow 0$  in  $L^3(\Omega, \mathbb{P}_\xi)$ . This meets condition (b-i) of proposition 7.1 with  $\nu = 3$  and also shows that all finite dimensional distributions converge to the zero vector as well (meeting condition (a) the proposition).

Using Hölder, with  $1/p + 1/p' = 1$ , and  $r > 0$

$$\mathbb{E}_\xi \left\{ Z \int_x^y \frac{1}{B_\varepsilon(s)} ds \right\}^r \leq (\mathbb{E}_\xi \{|Z|^{rp}\})^{1/p} \left( \mathbb{E}_\xi \left\{ \left| \int_x^y \frac{1}{B_\varepsilon(x)} ds \right|^{rp'} \right\} \right)^{1/p'}.$$

We choose  $r = 2$ ,  $p = 3/2$ ,  $p' = 3$  to get

$$\mathbb{E}_\xi \left\{ Z \int_x^y \frac{1}{B_\varepsilon(s)} ds \right\}^2 \leq (\mathbb{E}_\xi \{|Z|^3\})^{2/3} \left( \mathbb{E}_\xi \left\{ \left| \int_x^y \frac{1}{B_\varepsilon(x)} ds \right|^6 \right\} \right)^{1/3}.$$

Now due to lemma 7.4,

$$\left( \mathbb{E}_\xi \left\{ \left| \int_x^y \frac{1}{B_\varepsilon(x)} ds \right|^6 \right\} \right)^{1/3} \leq C_\xi^{1/3} \|\rho\|_\infty |y - x|^2.$$

Also, using Hölder and lemma 7.4,  $\mathbb{E}_\xi \{|Z|^3\}$  is  $\lesssim \varepsilon$  in all cases of  $Z$ . We therefore have

$$\mathbb{E}_\xi \left\{ \left| \frac{R_\varepsilon(y) - R_\varepsilon(x)}{\sqrt{\varepsilon}} \right|^3 \right\} \leq \varepsilon C_\xi'' |y - x|^2.$$

## Acknowledgment

The authors would like to thank Mark Adams and Yu Gu for many helpful discussions. This work was supported in part by NSF grant 0904746 and NSF RTG grant DMS-0602235. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## References

- [1] G. Bal. Central limits and homogenization in random media. *Multiscale Model. Simul.*, 7(2):677–702, 2008.
- [2] G. Bal, J. Garnier, S. Motsch, and V. Perrier. Random integrals and correctors in homogenization. *Asymptot. Anal.*, 59:1–26, 2008.
- [3] G. Bal and W. Jing. Corrector theory for msfem and hmm in random media. *submitted*, 2010.
- [4] G. Bal and W. Jing. Homogenization and corrector theory for linear transport in random media. *To appear: Disc. Cont. Dyn. Syst. B*, 2010.
- [5] P. Billingsley. *Convergence of Probability Measures*. John Wiley and Sons, New York, 1999.
- [6] Bousquet O. Boucheron, S. and G. Lugosi. Concentration inequalities. *Advanced lectures in machine learning*, pages 208–240, 2004.
- [7] A. Bourgeat and A. Piatnitski. Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator. *Asympt. Anal.*, 21:303–315, 1999.
- [8] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. *Ann. I. H. Poincaré*, 40:153–165, 2004.
- [9] W. Bryc. On large deviations for uniformly strong mixing sequences. *Stochastic processes and their applications*, 41:191–202, 1992.
- [10] J. Bucklew. *Introduction to rare event simulation*. Springer, 2004.
- [11] R.M. Burton and H. Dehling. Large deviations for some weakly dependent random variables. *Statistics & probability letters*, 9:397–401, 1990.
- [12] R. E. Caflisch. Monte carlo and quasi-monte carlo methods. *Acta Numerica*, pages 1–49, 1998.
- [13] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*. Applications of mathematics. Springer, 1998.
- [14] A.B. Dieker and M. Mandjes. On asymptotically efficient simulation of large deviation probabilities. *Adv. Appl. Prob.*, 37:539–552, 2005.
- [15] G. Folland. *Real analysis, modern techniques and their applications*. John Wiley & Sons, 1984.
- [16] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.
- [17] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, New York, 1994.
- [18] M. Loève. *Probability Theory II*. Springer, fourth edition, 1978.
- [19] Owhadi H. Lucas, L.J. and M. Ortiz. Rigorous verification, validation, uncertainty quantification and certification through concentration-of-measure inequalities. *Comput. Methods Appl. Mech. Engrg.*, 197:4591–4609, 2008.
- [20] G. Pavliotis and A. Stuart. *Multiscale Methods: Averaging and Homogenization*. Springer, 2008.
- [21] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.