



**HAL**  
open science

# Optimal quantitative estimates in stochastic homogenization for elliptic equations in nondivergence form

Scott N. Armstrong, Jessica Lin

► **To cite this version:**

Scott N. Armstrong, Jessica Lin. Optimal quantitative estimates in stochastic homogenization for elliptic equations in nondivergence form. 2017. hal-01483482

**HAL Id: hal-01483482**

**<https://hal.science/hal-01483482>**

Preprint submitted on 6 Mar 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# OPTIMAL QUANTITATIVE ESTIMATES IN STOCHASTIC HOMOGENIZATION FOR ELLIPTIC EQUATIONS IN NONDIVERGENCE FORM

SCOTT ARMSTRONG AND JESSICA LIN

ABSTRACT. We prove quantitative estimates for the stochastic homogenization of linear uniformly elliptic equations in nondivergence form. Under strong independence assumptions on the coefficients, we obtain optimal estimates on the subquadratic growth of the correctors with stretched exponential-type bounds in probability. Like the theory of Gloria and Otto [18, 19] for divergence form equations, the arguments rely on nonlinear concentration inequalities combined with certain estimates on the Green's functions and derivative bounds on the correctors. We obtain these analytic estimates by developing a  $C^{1,1}$  regularity theory down to microscopic scale, which is of independent interest and is inspired by the  $C^{0,1}$  theory introduced in the divergence form case by the first author and Smart [5].

## 1. INTRODUCTION

**1.1. Motivation and informal summary of results.** We identify the *optimal* error estimates for the stochastic homogenization of solutions  $u^\varepsilon$  solving:

$$(1.1) \quad \begin{cases} -\operatorname{tr}\left(A\left(\frac{x}{\varepsilon}\right)D^2u^\varepsilon\right) = 0 & \text{in } U, \\ u^\varepsilon(x) = g(x) & \text{on } \partial U. \end{cases}$$

Here  $U$  is a smooth bounded subset of  $\mathbb{R}^d$  with  $d \geq 2$ ,  $D^2v$  is the Hessian of a function  $v$  and  $\operatorname{tr}(M)$  denotes the trace of a symmetric matrix  $M \in \mathbb{S}^d$ . The coefficient matrix  $A(\cdot)$  is assumed to be a stationary random field, with given law  $\mathbb{P}$ , and valued in the subset of symmetric matrices with eigenvalues belonging to the interval  $[\lambda, \Lambda]$  for given ellipticity constants  $0 < \lambda \leq \Lambda$ . The solutions  $u^\varepsilon$  are understood in the *viscosity sense* [13] although in most of the paper the equations can be interpreted classically. We assume that the probability measure  $\mathbb{P}$  has a product-type structure and in particular possesses a *finite range of dependence* (see Section 1.3 for the precise statement). According to the general qualitative theory of stochastic homogenization developed in [23, 24] for nondivergence form elliptic equations (see also the later work [12]), the solutions  $u^\varepsilon$  of (1.1) converge uniformly as  $\varepsilon \rightarrow 0$ ,  $\mathbb{P}$ -almost surely, to that of the homogenized problem

$$(1.2) \quad \begin{cases} -\operatorname{tr}(\bar{A}D^2u) = 0 & \text{in } U, \\ u(x) = g(x) & \text{on } \partial U, \end{cases}$$

---

*Date:* November 18, 2016.

*2010 Mathematics Subject Classification.* 35B27, 35B45, 35J15.

*Key words and phrases.* stochastic homogenization, correctors, error estimate.

for some deterministic, uniformly elliptic matrix  $\bar{A}$ . Our interest in this paper is to study the rate of convergence of  $u^\varepsilon$  to  $u$ .

Error estimates quantifying the speed homogenization of  $u^\varepsilon \rightarrow u$  have been obtained in [24, 25, 11, 4]. The most recent paper [4] was the first to give a general result stating that the typical size of the error is at most algebraic, that is,  $O(\varepsilon^\alpha)$  for some positive exponent  $\alpha$ . The earlier work [25] gave an algebraic error estimate in dimensions  $d > 4$ . The main purpose of this paper is to reveal explicitly the optimal exponent.

Our main quantitative estimates concern the size of certain stationary solutions called the *approximate correctors*. These are defined, for a fixed symmetric matrix  $M \in \mathbb{S}^d$  and  $\varepsilon > 0$ , as the unique solution  $\phi_\varepsilon \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  of the equation

$$(1.3) \quad \varepsilon^2 \phi_\varepsilon - \operatorname{tr} \left( A(x)(M + D^2 \phi_\varepsilon) \right) = 0 \quad \text{in } \mathbb{R}^d.$$

Our main result states roughly that, for every  $x \in \mathbb{R}^d$ ,  $\varepsilon \in (0, \frac{1}{2}]$  and  $t > 0$ ,

$$(1.4) \quad \mathbb{P} \left[ |\varepsilon^2 \phi_\varepsilon(x) - \operatorname{tr}(\bar{A}M)| \geq t \mathcal{E}(\varepsilon) \right] \lesssim \exp \left( -t^{\frac{1}{2}} \right),$$

where the typical size  $\mathcal{E}(\varepsilon)$  of the error depends only on the dimension  $d$  in the following way:

$$(1.5) \quad \mathcal{E}(\varepsilon) := \begin{cases} \varepsilon |\log \varepsilon| & \text{in } d = 2, \\ \varepsilon^{\frac{3}{2}} & \text{in } d = 3, \\ \varepsilon^2 |\log \varepsilon|^{\frac{1}{2}} & \text{in } d = 4, \\ \varepsilon^2 & \text{in } d > 4. \end{cases}$$

Note that the rescaling  $\phi^\varepsilon(x) := \varepsilon^2 \phi_\varepsilon \left( \frac{x}{\varepsilon} \right)$  allows us to write (1.3) in the so-called theatrical scaling as

$$\phi^\varepsilon - \operatorname{tr} \left( A \left( \frac{x}{\varepsilon} \right) (M + D^2 \phi^\varepsilon) \right) = 0 \quad \text{in } \mathbb{R}^d.$$

This is a well-posed problem (it has a unique bounded solution) on  $\mathbb{R}^d$  which homogenizes to the equation

$$\phi - \operatorname{tr} \left( \bar{A} (M + D^2 \phi) \right) = 0 \quad \text{in } \mathbb{R}^d.$$

The solution of the latter is obviously the constant function  $\phi \equiv \operatorname{tr}(\bar{A}M)$  and so the limit

$$(1.6) \quad \varepsilon^2 \phi_\varepsilon(x) = \phi^\varepsilon(\varepsilon x) \rightarrow \operatorname{tr}(\bar{A}M)$$

is a qualitative homogenization statement. Therefore, the estimate (1.4) is a quantitative homogenization result for this particular problem which asserts that the speed of homogenization is  $O(\mathcal{E}(\varepsilon))$ .

Moreover, it is well-known that estimating the speed of homogenization for the Dirichlet problem is essentially equivalent to obtaining estimates on the approximate correctors (see [15, 7, 11, 4]). Indeed, the estimate (1.4) can be transferred without any loss of exponent to an estimate on the speed of

homogenization of the Dirichlet problem. One can see this from the standard two-scale ansatz

$$u^\varepsilon(x) \approx u(x) + \varepsilon^2 \phi_\varepsilon\left(\frac{x}{\varepsilon}\right),$$

which is easy to formalize and quantify in the linear setting since the homogenized solution  $u$  is completely smooth. We remark that since (1.4) is an estimate at a single point  $x$ , an estimate in  $L^\infty$  for the Dirichlet problem will necessarily have an additional logarithmic factor  $|\log \varepsilon|^q$  for some  $q(d) < \infty$ . Since the argument is completely deterministic and essentially the same as in the case of periodic coefficients, we do not give the details here and instead focus on the proof of (1.4).

The estimate (1.4) can also be expressed in terms of an estimate on the subquadratic growth of the correctors  $\phi$ , which are the solutions, for given  $M \in \mathbb{S}^d$ , of the problem

$$(1.7) \quad \begin{cases} -\operatorname{tr}(A(x)(M + D^2\phi)) = -\operatorname{tr}(\bar{A}M) & \text{in } \mathbb{R}^d, \\ \limsup_{R \rightarrow \infty} R^{-2} \operatorname{osc}_{B_R} \phi = 0. \end{cases}$$

Recall that while  $D^2\phi$  exists as a stationary random field,  $\phi$  itself may not be stationary. The estimate (1.4) implies that, for every  $x \in \mathbb{R}^d$ ,

$$\mathbb{P}[|\phi(x) - \phi(0)| \geq t|x|^2 \mathcal{E}(|x|^{-1})] \lesssim \exp\left(-t^{\frac{1}{2}}\right).$$

Notice that in dimensions  $d > 4$ , this implies that the typical size of  $|\phi(x) - \phi(0)|$  stays bounded as  $|x| \rightarrow \infty$ , suggesting that  $\phi$  is a locally bounded, stationary random field. In Section 7, we prove that this is so: the correctors are locally bounded and stationary in dimensions  $d > 4$ .

The above estimates are optimal in the size of the scaling, that is, the function  $\mathcal{E}(\varepsilon)$  cannot be improved in any dimension. This can be observed by considering the simple operator  $-a(x)\Delta$  where  $a(x)$  is a scalar field with a random checkerboard structure. Fix a smooth (deterministic) function  $f \in C_c^\infty(\mathbb{R}^d)$  and consider the problem

$$(1.8) \quad -a\left(\frac{x}{\varepsilon}\right) \Delta u^\varepsilon = f(x) \quad \text{in } \mathbb{R}^d.$$

We expect this to homogenize to a problem of the form

$$-\bar{a}\Delta u = f(x) \quad \text{in } \mathbb{R}^d.$$

In dimension  $d = 2$  (or in higher dimensions, if we wish) we can also consider the Dirichlet problem with zero boundary conditions in a ball much larger than the support of  $f$ . We can then move the randomness to the right side of the equation to have

$$-\Delta u^\varepsilon = a^{-1}\left(\frac{x}{\varepsilon}\right) f(x) \quad \text{in } \mathbb{R}^d$$

and then write a formula for the value of the solution  $u^\varepsilon$  at the origin as a convolution of the random right side against the Green's kernel for the Laplacian operator. The size of the fluctuations of this convolution is easy to determine, since it is essentially a sum (actually a convolution) of i.i.d. random variables,

and it turns out to be precisely of order  $\mathcal{E}(\varepsilon)$  (with a prefactor constant depending on the variance of  $a(\cdot)$  itself). For instance, we show roughly that

$$\text{var}[u^\varepsilon(0)] \simeq \mathcal{E}(\varepsilon)^2.$$

This computation also serves as a motivation for our proof of (1.4), although handling the general case of a random diffusion matrix is of course much more difficult than that of (1.8), in which the randomness is scalar and can be split from the operator.

**1.2. Method of proof and comparison to previous works.** The arguments in this paper are inspired by the methods introduced in the divergence form setting by Gloria and Otto [18, 19] and Gloria, Neukamm and Otto [16] (see also Mourrat [21]). The authors combined certain concentration inequalities and analytic arguments to prove optimal quantitative estimates in stochastic homogenization for linear elliptic equations of the form

$$-\nabla \cdot (A(x)\nabla u) = 0.$$

The concentration inequalities provided a convenient mechanism for transferring quantitative ergodic information from the coefficient field to the solutions themselves, an idea which goes back to an unpublished paper of Naddaf and Spencer [22]. Most of these works rely on some version of the Efron-Stein inequality [14] or the logarithmic Sobolev inequality to control the fluctuations of the solution by estimates on the spatial derivatives of the Green's functions and the solution.

A variant of these concentration inequalities plays an analogous role in this paper (see Proposition 2.2). There are then two main analytic ingredients we need to conclude: first, an estimate on the decay of the Green's function for the heterogeneous operator (note that, in contrast to the divergence form case, there is no useful deterministic bound on the decay of the Green's function); and (ii) a higher-order regularity theory asserting that, with high  $\mathbb{P}$ -probability, solutions of our random equation are more regular than the deterministic regularity theory would predict. We prove each of these estimates by using the sub-optimal (but algebraic) quantitative homogenization result of [4]: we show that, since solutions are close to those of the homogenized equation on large scales, we may “borrow” the estimates from the constant-coefficient equation. This is an idea that was introduced in the context of stochastic homogenization for divergence form equations by the first author and Smart [5] (see also [17, 3]) and goes back to work of Avellaneda and Lin [6, 7] in the case of periodic coefficients.

We remark that, while the scaling of the error  $\mathcal{E}(\varepsilon)$  is optimal, the estimate (1.4) is almost certainly sub-optimal in terms of stochastic integrability. This seems to be one limitation of an approach relying on (nonlinear) concentration inequalities, which so far yield only estimates with exponential moments [17] rather than Gaussian moments [1, 2, 20]. Recently, new approaches based on renormalization-type arguments (rather than nonlinear concentration inequalities) have been introduced in the divergence form setting [1, 2, 20]. It was shown in [2] that this approach yields estimates at the critical scale which are

also optimal in stochastic integrability. It would be very interesting to see whether such arguments could be developed in the nondivergence form case.

**1.3. Assumptions.** In this subsection, we introduce some notation and present the hypotheses. Throughout the paper, we fix ellipticity constants  $0 < \lambda \leq \Lambda$  and the dimension  $d \geq 2$ . The set of  $d$ -by- $d$  matrices is denoted by  $\mathbb{M}^d$ , the set of  $d$ -by- $d$  symmetric matrices is denoted by  $\mathbb{S}^d$ , and  $Id \in \mathbb{S}^d$  is the identity matrix. If  $M, N \in \mathbb{S}^d$ , we write  $M \leq N$  if every eigenvalue of  $N - M$  is nonnegative;  $|M|$  denotes the largest eigenvalue of  $M$ .

**1.3.1. Definition of the probability space  $(\Omega, \mathcal{F})$ .** We begin by giving the structural hypotheses on the equation. We consider matrices  $A \in \mathbb{M}^d$  which are uniformly elliptic:

$$(1.9) \quad \lambda Id \leq A(\cdot) \leq \Lambda Id$$

and Hölder continuous in  $x$ :

$$(1.10) \quad \exists \sigma \in (0, 1] \quad \text{such that} \quad \sup_{x, y \in \mathbb{R}^d} \frac{|A(x) - A(y)|}{|x - y|^\sigma} < \infty,$$

We define the set

$$(1.11) \quad \Omega := \{A : A \text{ satisfies (1.9) and (1.10)}\}.$$

Notice that the assumption (1.10) is a qualitative one. We purposefully do not specify any quantitative information regarding the size of the supremum in (1.10), because none of our estimates depend on this value. We make this assumption in order to ensure that a comparison principle holds for our equations.

We next introduce some  $\sigma$ -algebras on  $\Omega$ . For every Borel subset  $U \subseteq \mathbb{R}^d$  we define  $\mathcal{F}(U)$  to be the  $\sigma$ -algebra on  $\Omega$  generated by the behavior of  $A$  in  $U$ , that is,

$$(1.12) \quad \mathcal{F}(U) := \sigma\text{-algebra on } \Omega \text{ generated by the family of random variables} \\ \{A \mapsto A(x) : x \in U\}.$$

We define  $\mathcal{F} := \mathcal{F}(\mathbb{R}^d)$ .

**1.3.2. Translation action on  $(\Omega, \mathcal{F})$ .** The translation group action on  $\mathbb{R}^d$  naturally pushes forward to  $\Omega$ . We denote this action by  $\{T_y\}_{y \in \mathbb{R}^d}$ , with  $T_y : \Omega \rightarrow \Omega$  given by

$$(T_y A)(x) := A(x + y).$$

The map  $T_y$  is clearly  $\mathcal{F}$ -measurable, and is extended to  $\mathcal{F}$  by setting, for each  $E \in \mathcal{F}$ ,

$$T_y E := \{T_y A : A \in E\}.$$

We also extend the translation action to  $\mathcal{F}$ -measurable random elements  $X : \Omega \rightarrow S$  on  $\Omega$ , with  $S$  an arbitrary set, by defining  $(T_y X)(F) := X(T_y F)$ .

We say that a random field  $f : \mathbb{Z}^d \times \Omega \rightarrow S$  is  $\mathbb{Z}^d$ -stationary provided that  $f(y+z, A) = f(y, T_z A)$  for every  $y, z \in \mathbb{Z}^d$  and  $A \in \Omega$ . Note that an  $\mathcal{F}$ -measurable random element  $X : \Omega \rightarrow S$  may be viewed as a  $\mathbb{Z}^d$ -stationary random field via the identification with  $\tilde{X}(z, A) := X(T_z A)$ .

1.3.3. *Assumptions on the random environment.* Throughout the paper, we fix  $\ell \geq 2\sqrt{d}$  and a probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  which satisfies the following:

(P1)  $\mathbb{P}$  has  $\mathbb{Z}^d$ -stationary statistics: that is, for every  $z \in \mathbb{Z}^d$  and  $E \in \mathcal{F}$ ,

$$\mathbb{P}[E] = \mathbb{P}[T_z E].$$

(P2)  $\mathbb{P}$  has a range of dependence at most  $\ell$ : that is, for all Borel subsets  $U, V$  of  $\mathbb{R}^d$  such that  $\text{dist}(U, V) \geq \ell$ ,

$$\mathcal{F}(U) \text{ and } \mathcal{F}(V) \text{ are } \mathbb{P}\text{-independent.}$$

Some of our main results rely on concentration inequalities (stated in Section 2.2) which require a stronger independence assumption than finite range of dependence (P2) which was used in [4]. Namely, we require that  $\mathbb{P}$  is the pushforward of another probability measure which has a product space structure. We consider a probability space  $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$  and denote

$$(1.13) \quad (\Omega_*, \mathcal{F}_*, \mathbb{P}_*) := (\Omega_0^{\mathbb{Z}^d}, \mathcal{F}_0^{\mathbb{Z}^d}, \mathbb{P}_0^{\mathbb{Z}^d}).$$

We regard an element  $\omega \in \Omega_*$  as a map  $\omega : \mathbb{Z}^d \rightarrow \Omega_0$ . For each  $\Gamma \subseteq \mathbb{Z}^d$ , we denote by  $\mathcal{F}_*(\Gamma)$  the  $\sigma$ -algebra generated by the family  $\{\omega \mapsto \omega(z) : z \in \Gamma\}$  of maps from  $\Omega_*$  to  $\Omega_0$ . We denote the expectation with respect to  $\mathbb{P}_*$  by  $\mathbb{E}_*$ . Abusing notation slightly, we also denote the natural  $\mathbb{Z}^d$ -translation action on  $\Omega_*$  by  $T_z$ , that is,  $T_z : \Omega_* \rightarrow \Omega_*$  is defined by  $(T_z \omega)(y) := \omega(y + z)$ .

We assume that there exists an  $(\mathcal{F}_*, \mathcal{F})$ -measurable map  $\pi : \Omega_* \rightarrow \Omega$  which satisfies the following:

$$(1.14) \quad \mathbb{P}[E] = \mathbb{P}_*[\pi^{-1}(E)] \quad \text{for every } E \in \mathcal{F},$$

$$(1.15) \quad \pi \circ T_z = T_z \circ \pi \quad \text{for every } z \in \mathbb{Z}^d,$$

with the translation operator interpreted on each side in the obvious way, and

$$(1.16) \quad \text{for every Borel subset } U \subseteq \mathbb{R}^d \text{ and } E \in \mathcal{F}(U),$$

$$\pi^{-1}(E) \in \mathcal{F}_* \left( \{z \in \mathbb{Z}^d : \text{dist}(z, U) \leq \ell/2\} \right).$$

We summarize the above conditions as:

(P3) There exists a probability space  $(\Omega_*, \mathcal{F}_*, \mathbb{P}_*)$  of the form (1.13) and a map

$$\pi := \Omega_* \rightarrow \Omega,$$

which is  $(\mathcal{F}, \mathcal{F}_*)$ -measurable and satisfies (1.14), (1.15) and (1.16).

Note that, in view of the product structure, the conditions (1.14) and (1.16) imply (P2) and (1.14) and (1.15) imply (P1). Thus (P1) and (P2) are superseded by (P3).

**1.4. Statement of main result.** We next present the main result concerning quantitative estimates of the approximate correctors.

**Theorem 1.1.** *Assume that  $\mathbb{P}$  is a probability measure on  $(\Omega, \mathcal{F})$  satisfying (P3). Let  $\mathcal{E}(\varepsilon)$  be defined by (1.5). Then there exist  $\delta(d, \lambda, \Lambda) > 0$  and  $C(d, \lambda, \Lambda, \ell) \geq 1$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$ ,*

$$(1.17) \quad \mathbb{E} \left[ \exp \left( \left( \frac{1}{\mathcal{E}(\varepsilon)} \sup_{x \in B_{\sqrt{d}}(0)} |\varepsilon^2 \phi_\varepsilon(x) - \text{tr}(\overline{AM})| \right)^{\frac{1}{2} + \delta} \right) \right] \leq C.$$

The proof of Theorem 1.1 is completed in Section 6.

**1.5. Outline of the paper.** The rest of the paper is organized as follows. In Section 2, we introduce the approximate correctors and the modified Green's functions and give some preliminary results which are needed for our main arguments. In Section 3, we establish a  $C^{1,1}$  regularity theory down to unit scale for solutions. Section 4 contains estimates on the modified Green's functions, which roughly state that these functions have the same rate of decay at infinity as the Green's function for the homogenized equation (i.e, the Laplacian). In this section, we also mention estimates on the invariant measure associated to the linear operator in (1.1). In Section 5, we use results from the previous sections to measure the ‘‘sensitivity’’ of solutions of the approximate corrector equation with respect to the coefficients. Finally, in Section 6, we obtain the optimal rates of decay for the approximate corrector, proving our main result, and, in Section 7, demonstrate the existence of stationary correctors in dimensions  $d > 4$ . In the appendix, we give a proof of the concentration inequality we use in our analysis, which is a stretched exponential version of the Efron-Stein inequality.

## 2. PRELIMINARIES

**2.1. Approximate correctors and modified Green's functions.** For each given  $M \in \mathbb{S}^d$  and  $\varepsilon > 0$ , the approximate corrector equation is

$$\varepsilon^2 \phi_\varepsilon - \text{tr}(A(x)(M + D^2 \phi_\varepsilon)) = 0 \quad \text{in } \mathbb{R}^d.$$

The existence of a unique solution  $\phi_\varepsilon$  belonging to  $C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  is obtained from the usual Perron method and the comparison principle.

We also introduce, for each  $\varepsilon \in (0, 1]$  and  $y \in \mathbb{R}^d$ , the ‘‘modified Green's function’’  $G_\varepsilon(\cdot, y; A) = G_\varepsilon(\cdot, y)$ , which is the unique solution of the equation

$$(2.1) \quad \varepsilon^2 G_\varepsilon - \text{tr}(A(x)D^2 G_\varepsilon) = \chi_{B_\ell(y)} \quad \text{in } \mathbb{R}^d.$$

Compared to the usual Green's function, we have smeared out the singularity and added the zeroth-order ‘‘massive’’ term.

To see that (2.1) is well-posed, we build the solution  $G_\varepsilon$  by compactly supported approximations. We first solve the equation in the ball  $B_R$  (for  $R > \ell$ ) with zero Dirichlet boundary data to obtain a function  $G_{\varepsilon,R}(\cdot, y)$ . By the maximum principle,  $G_{\varepsilon,R}(\cdot, y)$  is increasing in  $R$  and we may therefore define its limit as  $R \rightarrow \infty$  to be  $G_\varepsilon(\cdot, y)$ . We show in the following lemma that it is bounded and decays at infinity, and from this it is immediate that



it satisfies (2.1) in  $\mathbb{R}^d$ . The lemma is a simple deterministic estimate which is useful only in the regime  $|x - y| \gtrsim \varepsilon^{-1} |\log \varepsilon|$  and follows from the fact (as demonstrated by a simple test function) that the interaction between the terms on the left of (2.1) give the equation a characteristic length scale of  $\varepsilon^{-1}$ .

**Lemma 2.1.** *There exist  $a(d, \lambda, \Lambda) > 0$  and  $C(d, \lambda, \Lambda) > 1$  such that, for every  $A \in \Omega$ ,  $x, y \in \mathbb{R}^d$  and  $\varepsilon \in (0, 1]$ ,*

$$(2.2) \quad G_\varepsilon(x, y) \leq C\varepsilon^{-2} \exp(-\varepsilon a|x - y|).$$

*Proof.* Without loss of generality, let  $y = 0$ . Let  $\phi(x) := C \exp(-\varepsilon a|x|)$  for  $C, a > 0$  to be selected below. Compute, for every  $x \neq 0$ ,

$$D^2\phi(x) = \phi(x) \left( -\varepsilon a \frac{1}{|x|} \left( I - \frac{x \otimes x}{|x|^2} \right) + \varepsilon^2 a^2 \frac{x \otimes x}{|x|^2} \right).$$

Thus for any  $A \in \Omega$ ,

$$-\operatorname{tr}(A(x)D^2\phi(x)) \geq \phi(x) \left( \frac{1}{|x|} \lambda \varepsilon a(d-1) - \Lambda(\varepsilon^2 a^2) \right) \quad \text{in } \mathbb{R}^d \setminus \{0\}.$$

Set  $a := \frac{1}{\sqrt{2\Lambda}}$ . Then

$$(2.3) \quad \varepsilon^2 \phi - \operatorname{tr}(A(x)D^2\phi(x)) \geq \varepsilon^2 \phi(x) (1 - \Lambda a^2) + \frac{\phi(x)}{|x|} \lambda \varepsilon a(d-1) \geq 0 \quad \text{in } \mathbb{R}^d \setminus \{0\}.$$

Take  $C := \exp(a\ell)$  so that  $\phi(x) \geq 1$  in  $|x| \leq \ell$ . Define  $\tilde{\phi} := \varepsilon^{-2} \min\{1, \phi\}$ . Then  $\tilde{\phi}$  satisfies the inequality

$$\varepsilon^2 \phi(x) - \operatorname{tr}(A(x)D^2\phi(x)) \geq \chi_{B_\ell} \quad \text{in } \mathbb{R}^d.$$

As  $\tilde{\phi} > 0$  on  $\partial B_R$ , the comparison principle yields that  $\tilde{\phi} \geq G_{\varepsilon, R}(\cdot, 0)$  for every  $R > 1$ . Letting  $R \rightarrow \infty$  yields the lemma.  $\square$

**2.2. Spectral gap inequalities.** In this subsection, we state the probabilistic tool we use to obtain the quantitative estimates for the modified correctors. The result here is applied precisely once in the paper, in Section 6, and relies on the stronger independence condition (P3). It is a variation of the Efron-Stein (“spectral gap”) inequality; a proof is given in Appendix A.

**Proposition 2.2.** *Fix  $\beta \in (0, 2)$ . Let  $X$  be a random variable on  $(\Omega_*, \mathcal{F}_*, \mathbb{P}_*)$  and set*

$$X'_z := \mathbb{E}_* [X | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})] \quad \text{and} \quad \mathbb{V}_*[X] := \sum_{z \in \mathbb{Z}^d} (X - X'_z)^2.$$

*Then there exists  $C(\beta) \geq 1$  such that*

$$(2.4) \quad \mathbb{E}_* \left[ \exp(|X - \mathbb{E}[X]|^\beta) \right] \leq C \mathbb{E}_* \left[ \exp \left( (C \mathbb{V}_*[X])^{\frac{\beta}{2-\beta}} \right) \right]^{\frac{2-\beta}{\beta}}.$$

The conditional expectation  $X'_z$  can be identified by resampling the random environment near the point  $z$  (this is explained in depth in Section 5). Therefore, the quantity  $(X - X'_z)$  measures changes in  $X$  with respect to changes in the environment near  $z$ . Following [16], we refer to  $(X - X'_z)$  as the *vertical derivative* of  $X$  at the point  $z$ .

**2.3. Suboptimal error estimate.** We recall the main result of [4], which is the basis for much of the analysis in this paper. We reformulate their result slightly, to put it in a form which is convenient for our use here.

**Proposition 2.3.** *Fix  $\sigma \in (0, 1]$  and  $s \in (0, d)$ . Let  $\mathbb{P}$  satisfy (P1) and (P2). There exists an exponent  $\alpha(\sigma, s, d, \lambda, \Lambda, \ell) > 0$  and a nonnegative random variable  $\mathcal{X}$  on  $(\Omega, \mathcal{F})$ , satisfying*

$$(2.5) \quad \mathbb{E}[\exp(\mathcal{X})] \leq C(\sigma, s, d, \Lambda, \lambda, \ell) < \infty$$

and such that the following holds: for every  $R \geq 1$ ,  $f \in C^{0,\sigma}(B_R)$ ,  $g \in C^{0,\sigma}(\partial B_R)$  and solutions  $u, v \in C(\overline{B}_R)$  of the Dirichlet problems

$$(2.6) \quad \begin{cases} -\operatorname{tr}(A(x)D^2u) = f(x) & \text{in } B_R, \\ u = g & \text{on } \partial B_R, \end{cases}$$

and

$$(2.7) \quad \begin{cases} -\operatorname{tr}(\overline{A}D^2v) = f(x) & \text{in } B_R, \\ v = g & \text{on } \partial B_R, \end{cases}$$

we have, for a constant  $C(\sigma, s, d, \Lambda, \lambda, \ell) \geq 1$ , the estimate

$$R^{-2} \sup_{B_R} |u(x) - v(x)| \leq CR^{-\alpha} (1 + \mathcal{X}R^{-s}) \Gamma_{R,\sigma}(f, g),$$

where

$$\Gamma_{R,\sigma}(f, g) := \sup_{B_R} |f| + R^\sigma [f]_{C^{0,\sigma}(B_R)} + R^{-2} \operatorname{osc}_{\partial B_R} g + R^{-2+\sigma} [g]_{C^{0,\sigma}(\partial B_R)}.$$

We note that, in view of the comparison principle, Proposition 2.3 also give one-sided estimates for subsolutions and supersolutions.

### 3. UNIFORM $C^{1,1}$ ESTIMATES

In this section, we present a large-scale ( $R \gg 1$ ) regularity theory for solutions of the equation

$$(3.1) \quad -\operatorname{tr}(A(x)D^2u) = f(x) \quad \text{in } B_R$$

Recall that according to the Krylov-Safonov Hölder regularity theory [10], there exists an exponent  $\sigma(d, \lambda, \Lambda) \in (0, 1)$  such that  $u$  belongs to  $C^{0,\sigma}(B_{R/2})$  with the estimate

$$(3.2) \quad R^{\sigma-2} [u]_{C^{0,\sigma}(B_{R/2})} \leq C \left( R^{-2} \operatorname{osc}_{B_R} u + \left( \int_{B_R} |f(x)|^d dx \right)^{\frac{1}{d}} \right).$$

This is the best possible estimate, independent of the size of  $R$ , for solutions of general equations of the form (3.1), even if the coefficients are smooth. What we show in this section is that, due to statistical effects, solutions of our random equation are typically much more regular, at least on scales larger than  $\ell$ , the length scale of the correlations of the coefficient field. Indeed, we will show that solutions, with high probability, are essentially  $C^{1,1}$  on scales larger than the unit scale.

The arguments here are motivated by a similar  $C^{0,1}$  regularity theory for equations in divergence form developed in [5] and should be seen as a nondivergence analogue of those estimates. In fact, the arguments here are almost identical to those of [5]. They can also be seen as generalizations to the random setting of the results of Avellaneda and Lin for periodic coefficients, who proved uniform  $C^{0,1}$  estimates for divergence form equations [6] and then  $C^{1,1}$  estimates for the nondivergence case [7]. Note that it is natural to expect estimates in nondivergence form to be one derivative better than those in divergence form (e.g., the Schauder estimates). Note that the “ $C^{1,1}$  estimate” proved in [17] has a different statement which involves correctors; the statement we prove here would be simply false for divergence form equations.

The rough idea, similar to the proof of the classical Schauder estimates, is that, due to homogenization, large-scale solutions of (3.1) can be well-approximated by those of the homogenized equation. Since the latter are harmonic, up to a change of variables, they possess good estimates. If the homogenization is fast enough (for this we need the results of [4], namely Proposition 2.3), then the better regularity of the homogenized equation is inherited by the heterogeneous equation. This is a quantitative version of the idea introduced in the context of periodic homogenization by Avellaneda and Lin [6, 7].

Throughout this section, we let  $\mathcal{Q}$  be the set of polynomials of degree at most two and let  $\mathcal{L}$  denote the set of affine functions. For  $\sigma \in (0, 1]$  and  $U \subseteq \mathbb{R}^d$ , we denote the usual Hölder seminorm by  $[\cdot]_{C^{0,\sigma}(U)}$ .

**Theorem 3.1** ( $C^{1,1}$  regularity). *Assume (P1) and (P2). Fix  $s \in (0, d)$  and  $\sigma \in (0, 1]$ . There exists an  $\mathcal{F}$ -measurable random variable  $\mathcal{X}$  and a constant  $C(s, \sigma, d, \lambda, \Lambda, \ell) \geq 1$  satisfying*

$$(3.3) \quad \mathbb{E}[\exp(\mathcal{X}^s)] \leq C < \infty$$

*such that the following holds: for every  $M \in \mathbb{S}^d$ ,  $R \geq 2\mathcal{X}$ ,  $u \in C(B_R)$  and  $f \in C^{0,\sigma}(B_R)$  satisfying*

$$-\operatorname{tr}(A(x)(M + D^2u)) = f(x) \quad \text{in } B_R$$

*and every  $r \in [\mathcal{X}, \frac{1}{2}R]$ , we have the estimate*

$$(3.4) \quad \frac{1}{r^2} \inf_{l \in \mathcal{L}} \sup_{B_r} |u - l| \leq C \left( |f(0) + \operatorname{tr}(\bar{A}M)| + R^\sigma [f]_{C^{0,\sigma}(B_R)} + \frac{1}{R^2} \inf_{l \in \mathcal{L}} \sup_{B_R} |u - l| \right).$$

It is well-known that any  $L^\infty$  function which can be well-approximated on all scales by smooth functions must be smooth. The proof of Theorem 3.1 is based on a similar idea: any function  $u$  which can be well-approximated on all scales above a fixed scale  $\mathcal{X}$ , which is of the same order as the microscopic scale, by functions  $w$  enjoying an *improvement of quadratic approximation* property must itself have this property. This is formalized in the next proposition, the statement and proof of which are similar to those of [4, Lemma 5.1].

**Proposition 3.2.** *For each  $r > 0$  and  $\theta \in (0, \frac{1}{2})$ , let  $\mathcal{A}(r, \theta)$  denote the subset of  $L^\infty(B_r)$  consisting of  $w$  which satisfy*

$$\frac{1}{(\theta r)^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_{\theta r}} |w(x) - q(x)| \leq \frac{1}{2} \left( \frac{1}{r^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_r} |w(x) - q(x)| \right).$$

*Assume that  $\alpha, \gamma, K, L > 0$ ,  $1 \leq 4r_0 \leq R$  and  $u \in L^\infty(B_R)$  are such that, for every  $r \in [r_0, R/2]$ , there exists  $v \in \mathcal{A}(r, \theta)$  such that*

$$(3.5) \quad \frac{1}{r^2} \sup_{x \in B_r} |u(x) - v(x)| \leq r^{-\alpha} \left( K + \frac{1}{r^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_{2r}} |u(x) - l(x)| \right) + Lr^\gamma.$$

*Then there exist  $\beta(\theta) \in (0, 1]$  and  $C(\alpha, \theta, \gamma) \geq 1$  such that, for every  $s \in [r_0, R/2]$ ,*

$$(3.6) \quad \frac{1}{s^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_s} |u(x) - l(x)| \leq C \left( K + LR^\gamma + \frac{1}{R^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_R} |u(x) - l(x)| \right).$$

*and*

$$(3.7) \quad \frac{1}{s^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_s} |u(x) - q(x)| \leq C \left( \frac{s}{R} \right)^\beta \left( LR^\gamma + \frac{1}{R^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_R} |u(x) - q(x)| \right) + Cs^{-\alpha} \left( K + LR^\gamma + \frac{1}{R^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_R} |u(x) - l(x)| \right).$$

*Proof.* Throughout the proof, we let  $C$  denote a positive constant which depends only on  $(\alpha, \theta, \gamma)$  and may vary in each occurrence. We may suppose without loss of generality that  $\alpha \leq 1$  and that  $\gamma \leq c$  so that  $\theta^\gamma \geq \frac{2}{3}$ .

*Step 1.* We set up the argument. We keep track of the two quantities

$$G(r) := \frac{1}{r^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_r} |u(x) - q(x)| \quad \text{and} \quad H(r) := \frac{1}{r^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_r} |u(x) - l(x)|.$$

By the hypotheses of the proposition and the triangle inequality, we obtain that, for every  $r \in [r_0, R/2]$ ,

$$(3.8) \quad G(\theta r) \leq \frac{1}{2}G(r) + Cr^{-\alpha} (K + H(2r)) + Lr^\gamma.$$

Denote  $s_0 := R$  and  $s_j := \theta^{j-1}R/4$  and let  $m \in \mathbb{N}$  be such that

$$s_m^{-\alpha} \leq \frac{1}{4} \leq s_{m+1}^{-\alpha}.$$

Denote  $G_j := G(s_j)$  and  $H_j := H(s_j)$ . Noting that  $\theta \leq \frac{1}{2}$ , from (3.8) we get, for every  $j \in \{1, \dots, m-1\}$ ,

$$(3.9) \quad G_{j+1} \leq \frac{1}{2}G_j + Cs_j^{-\alpha} (K + H_{j-1}) + Ls_j^\gamma.$$

For each  $j \in \{0, \dots, m-1\}$ , we may select  $q_j \in \mathcal{Q}$  such that

$$G_j = \frac{1}{s_j^2} \sup_{x \in B_{s_j}} |u(x) - q_j(x)|.$$

We denote the Hessian matrix of  $q_j$  by  $Q_j$ . The triangle inequality implies that

$$(3.10) \quad G_j \leq H_j \leq G_j + \frac{1}{s_j^2} \sup_{x \in B_{s_j}} \frac{1}{2}x \cdot Q_j x = G_j + \frac{1}{2}|Q_j|,$$

and

$$\frac{1}{s_j^2} \sup_{x \in B_{s_{j+1}}} |q_{j+1}(x) - q_j(x)| \leq G_j + \theta^2 G_{j+1}.$$

The latter implies

$$|Q_{j+1} - Q_j| \leq \frac{2}{s_{j+1}^2} \sup_{x \in B_{s_{j+1}}} |q_{j+1}(x) - q_j(x)| \leq \frac{2}{\theta^2} G_j + 2G_{j+1}.$$

In particular,

$$(3.11) \quad |Q_{j+1}| \leq |Q_j| + C(G_j + G_{j+1})$$

Similarly, the triangle inequality also gives

$$(3.12) \quad |Q_j| = \frac{2}{s_j^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_{s_j}} |q_j(x) - l(x)| \leq 2G_j + 2H_j \leq 4H_j.$$

Thus, combining (3.11) and (3.12), yields

$$(3.13) \quad |Q_j| \leq |Q_0| + C \sum_{i=0}^j G_i \leq C \left( H_0 + \sum_{i=0}^j G_i \right).$$

Next, combining (3.9) (3.10) and (3.13), we obtain, for every  $j \in \{0, \dots, m-1\}$ ,

$$(3.14) \quad G_{j+1} \leq \frac{1}{2} G_j + C s_j^{-\alpha} \left( K + H_0 + \sum_{i=0}^j G_i \right) + L s_j^\gamma.$$

The rest of the argument involves first iterating (3.14) to obtain bounds on  $G_j$ , which yield bounds on  $|Q_j|$  by (3.13), and finally on  $H_j$  by (3.10).

*Step 2.* We show that, for every  $j \in \{1, \dots, m\}$ ,

$$(3.15) \quad G_j \leq 2^{-j} G_0 + C s_j^{-\alpha} (K + H_0) + CL (s_j^\gamma + R^\gamma s_j^{-\alpha}).$$

We argue by induction. Fix  $A, B \geq 1$  (which are selected below) and suppose that  $k \in \{0, \dots, m-1\}$  is such that, for every  $j \in \{0, \dots, k\}$ ,

$$G_j \leq 2^{-j} G_0 + A s_j^{-\alpha} (K + H_0) + L (A s_j^\gamma + B R^\gamma s_j^{-\alpha}).$$

Using (3.14) and this induction hypothesis (and that  $G_0 \leq H_0$ ), we get

$$\begin{aligned} G_{k+1} &\leq \frac{1}{2} G_k + C s_k^{-\alpha} \left( K + H_0 + \sum_{j=0}^k G_j \right) + L s_k^\gamma \\ &\leq \frac{1}{2} \left( 2^{-k} G_0 + A s_k^{-\alpha} (K + H_0) + L (A s_k^\gamma + B R^\gamma s_k^{-\alpha}) \right) + L s_k^\gamma \\ &\quad + C s_k^{-\alpha} \left( K + H_0 + \sum_{j=0}^k \left( 2^{-j} G_0 + A s_j^{-\alpha} (K + H_0) + L (A s_j^\gamma + B R^\gamma s_j^{-\alpha}) \right) \right) \\ &\leq 2^{-(k+1)} G_0 + s_{k+1}^{-\alpha} (K + H_0) \left( \frac{1}{2} A + C A s_k^{-\alpha} + C \right) \\ &\quad + L s_{k+1}^\gamma \left( \frac{1}{2} \theta^{-\gamma} A + C \right) + L R^\gamma s_{k+1}^{-\alpha} \left( \frac{1}{2} B + C A + C B s_k^{-\alpha} \right). \end{aligned}$$

Now suppose in addition that  $k \leq n$  with  $n$  such that  $Cs_n^{-\alpha} \leq \frac{1}{4}$ . Then using this and  $\theta^\gamma \geq \frac{2}{3}$ , we may select  $A$  large enough so that

$$\frac{1}{2}A + CAs_k^{-\alpha} + C \leq \frac{3}{4}A + C \leq A \quad \text{and} \quad \frac{1}{2}\theta^{-\gamma}A + C \leq A,$$

and then select  $B$  large enough so that

$$\frac{1}{2}B + CA + CBs_k^{-\alpha} \leq B.$$

We obtain

$$G_{k+1} \leq 2^{-(k+1)}G_0 + As_{k+1}^{-\alpha}(K + H_0) + ALs_{k+1}^\gamma + BLR^\gamma s_{k+1}^{-\alpha}.$$

By induction, we now obtain (3.15) for every  $j \leq n \leq m$ . In addition, for every  $j \in \{n+1, \dots, m\}$ , we have that  $1 \leq s_j/s_m \leq C$ . This yields (3.15) for every  $j \in \{0, \dots, n\}$ .

*Step 3.* The bound on  $H_j$  and conclusion. By (3.13), (3.15), we have

$$\begin{aligned} |Q_j| &\leq C \left( H_0 + \sum_{i=0}^j G_i \right) \\ &\leq C \left( H_0 + \sum_{i=0}^j (2^{-i}G_0 + Cs_i^{-\alpha}(K + H_0) + CL(s_i^\gamma + R^\gamma s_i^{-\alpha})) \right) \\ &\leq C (H_0 + G_0 + Cs_j^{-\alpha}(K + H_0) + CLR^\gamma (1 + s_j^{-\alpha})) \\ &\leq CH_0 + CKs_j^{-\alpha} + CLR^\gamma. \end{aligned}$$

Here we also used that  $s_j^{-\alpha} \leq s_m^{-\alpha} \leq C$ . Using the previous inequality, (3.10) and (3.15), we conclude that

$$H_j \leq G_j + \frac{1}{2}|Q_j| \leq CH_0 + CKs_j^{-\alpha} + CLR^\gamma.$$

This is (3.6). Note that (3.15) already implies (3.7) for  $\beta := (\log 2)/|\log \theta|$ .  $\square$

Next, we recall that solutions of the homogenized equation (which are essentially harmonic functions) satisfy the ‘‘improvement of quadratic approximation’’ property.

**Lemma 3.3.** *Let  $r > 0$ . Assume that  $\bar{A}$  satisfies (1.9) and let  $w \in C(B_r)$  be a solution of*

$$(3.16) \quad -\operatorname{tr}(\bar{A}D^2w) = 0 \quad \text{in } B_r.$$

*There exists  $\theta(d, \lambda, \Lambda) \in (0, \frac{1}{2}]$  such that, for every  $r > 0$ ,*

$$\frac{1}{(\theta r)^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_{\theta r}} |w(x) - q(x)| \leq \frac{1}{2} \left( \frac{1}{r^2} \inf_{q \in \mathcal{Q}} \sup_{x \in B_r} |w(x) - q(x)| \right).$$

*Proof.* Since  $\bar{A}$  is constant, the solutions of (3.16) are harmonic (up to a linear change of coordinates). Thus the result of this lemma is classical.  $\square$

Equipped with the above lemmas, we now give the proof of Theorem 3.1.

*Proof of Theorem 3.1.* Fix  $s \in (0, d)$ . We denote by  $C$  and  $c$  positive constants depending only on  $(s, \sigma, d, \lambda, \Lambda, \ell)$  which may vary in each occurrence. We proceed with the proof of (3.4). Let  $\mathcal{Y}$  be the random variable in the statement of Proposition 2.3, with  $\alpha$  the exponent there. Define  $\mathcal{X} := \mathcal{Y}^{1/s}$  and observe that  $\mathcal{X}$  satisfies (3.3). We take  $\sigma$  to be the smallest of the following: the exponent in (3.2) and half the exponent  $\alpha$  in Proposition 2.3.

We may suppose without loss of generality that  $-\operatorname{tr}(\overline{A}M) = f(0) = 0$ .

*Step 1.* We check that  $u$  satisfies the hypothesis of Proposition 3.2 with  $r_0 = C\mathcal{X}$ . Fix  $r \in [C\mathcal{X}, R/2]$ . We take  $v, w \in C(B_{3r/4})$  to be the solutions of the problems

$$\begin{cases} -\operatorname{tr}(\overline{A}D^2v) = f(x) & \text{in } B_{3r/4}, \\ v = u & \text{on } \partial B_{3r/4}, \end{cases} \quad \begin{cases} -\operatorname{tr}(\overline{A}D^2w) = 0 & \text{in } B_{3r/4}, \\ w = u & \text{on } \partial B_{3r/4}. \end{cases}$$

By the Alexandrov-Bakelman-Pucci estimate [10], we have

$$(3.17) \quad \frac{1}{r^2} \sup_{B_{r/2}} |v - w| \leq C \left( \int_{B_r} |f(x)|^d dx \right)^{\frac{1}{d}} \leq Cr^\sigma [f]_{C^{0,\sigma}(B_r)}.$$

By the Krylov-Safanov Hölder estimate (3.2),

$$(3.18) \quad r^{\sigma-2} [u]_{C^{0,\sigma}(B_{3r/4})} \leq C \left( \frac{1}{r^2} \operatorname{osc}_{B_r} u + \left( \int_{B_r} |f(x)|^d dx \right)^{\frac{1}{d}} \right) \\ \leq C \left( \frac{1}{r^2} \operatorname{osc}_{B_r} u + r^\sigma [f]_{C^{0,\sigma}(B_r)} \right).$$

By the error estimate (Proposition 2.3), we have

$$\frac{1}{r^2} \sup_{B_{r/2}} |u - v| \\ \leq Cr^{-\alpha} (1 + \mathcal{Y}r^{-s}) \left( r^\sigma [f]_{C^{0,\sigma}(B_r)} + \frac{1}{r^2} \operatorname{osc}_{\partial B_{3r/4}} u + r^{\sigma-2} [u]_{C^{0,\sigma}(B_{3r/4})} \right).$$

Using the assumption  $r^s \geq \mathcal{X}^s = \mathcal{Y}$  and (3.18), this gives

$$(3.19) \quad \frac{1}{r^2} \sup_{B_{r/2}} |u - v| \leq Cr^{-\alpha} \left( r^\sigma [f]_{C^{0,\sigma}(B_r)} + \frac{1}{r^2} \operatorname{osc}_{B_r} u \right).$$

Using (3.17) and (3.19), the triangle inequality, and the definition of  $\sigma$ , we get

$$\frac{1}{r^2} \sup_{B_{r/2}} |u - w| \leq Cr^{-\alpha} \left( \frac{1}{r^2} \operatorname{osc}_{B_r} u \right) + Cr^\sigma [f]_{C^{0,\sigma}(B_r)}.$$

By Lemma 3.3,  $w \in \mathcal{A}(r, \theta)$  for some  $\theta \geq c$ .

*Step 2.* We apply Proposition 3.2 to obtain (3.4). The proposition gives, for every  $r \geq r_0 = C\mathcal{X}$ ,

$$\frac{1}{r^2} \inf_{l \in \mathcal{L}} \sup_{x \in \overline{B_r}} |u(x) - l(x)| \leq C \left( R^\sigma [f]_{C^{0,\sigma}(B_R)} + \frac{1}{R^2} \inf_{l \in \mathcal{L}} \sup_{x \in \overline{B_r}} |u(x) - l(x)| \right),$$

which is (3.4).  $\square$

It is convenient to restate the estimate (3.4) in terms of “coarsened” seminorms. Recall that, for  $\phi \in C^\infty(B_1)$ ,

$$\begin{aligned} |D\phi(x_0)| &\simeq \frac{1}{h} \operatorname{osc}_{B_h(x_0)} \phi(x), \\ |D^2\phi(x_0)| &\simeq \frac{1}{h^2} \inf_{p \in \mathbb{R}^d} \operatorname{osc}_{B_h(x_0)} (\phi(x) - p \cdot x), \end{aligned} \quad \text{for } 0 < h \ll 1.$$

This motivates the following definitions: for  $\alpha \in (0, 1]$ ,  $h \geq 0$ ,  $U \subseteq \mathbb{R}^d$  and  $x_0 \in U$ , we define the *pointwise, coarsened  $h$ -scale  $C_h^{0,1}(U)$  and  $C_h^{1,1}(U)$  seminorms* at  $x_0$  by

$$[\phi]_{C_h^{0,\alpha}(x_0,U)} := \sup_{r>h} \frac{1}{r^\alpha} \operatorname{osc}_{B_r(x_0) \cap U} \phi,$$

and

$$[\phi]_{C_h^{1,\alpha}(x_0,U)} := \sup_{r>h} \frac{1}{r^{1+\alpha}} \inf_{l \in \mathcal{L}} \operatorname{osc}_{B_r(x_0) \cap U} (\phi(x) - l(x)).$$

This allows us to write (3.4) in the form

$$\begin{aligned} (3.20) \quad [u]_{C_1^{1,1}(0,B_{R/2})} &\leq C\mathcal{X}^2 \left( |f(0) + \operatorname{tr}(\bar{A}M)| + R^\sigma [f]_{C^{0,\sigma}(B_R)} + \frac{1}{R^2} \inf_{l \in \mathcal{L}} \sup_{x \in B_R} |u - l| \right). \end{aligned}$$

As a simple corollary to Theorem 3.1, we also have  $C_1^{0,1}$  bounds on  $u$ :

**Corollary 3.4.** *Assume the hypotheses and notation of Theorem 3.1. Then,*

$$(3.21) \quad [u]_{C_1^{0,1}(0,B_{R/2})} \leq \mathcal{X} \left( R |f(0) + \operatorname{tr} \bar{A}M| + R^{1+\sigma} [f]_{C^{0,\sigma}(B_R)} + \frac{1}{R} \operatorname{osc}_{B_R} u \right).$$

The proof follows from a simple interpolation inequality, which controls the seminorm  $[\cdot]_{C_h^{0,1}(B_R)}$  in terms of  $[\cdot]_{C_h^{1,1}(B_R)}$  and the oscillation in  $B_R$ :

**Lemma 3.5.** *For any  $R > 0$  and  $\phi \in C(B_R)$ , we have*

$$(3.22) \quad [\phi]_{C_h^{0,1}(0,B_R)} \leq 14 \left( [\phi]_{C_h^{1,1}(0,B_R)} \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}.$$

*Proof.* We must show that, for every  $s \in [h, R]$ ,

$$(3.23) \quad \frac{1}{s} \operatorname{osc}_{B_s} \phi \leq 14 \left( [\phi]_{C_h^{1,1}(0,B_R)} \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}.$$

Set

$$K := \left( [\phi]_{C_h^{1,1}(0,B_R)} \right)^{-\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}$$

and observe that, for every  $s \in [K, R]$ , we have

$$(3.24) \quad \frac{1}{s} \operatorname{osc}_{B_s} \phi \leq K^{-1} \operatorname{osc}_{B_R} \phi = \left( [\phi]_{C_h^{1,1}(0,B_R)} \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}.$$

Thus we need only check (3.23) for  $s \in [h, K]$ .



We next claim that, for every  $s \in [h, R)$ ,

$$(3.25) \quad \frac{2}{s} \operatorname{osc}_{B_{s/2}} \phi \leq 3s [\phi]_{C_h^{1,1}(0, B_R)} + \frac{1}{s} \operatorname{osc}_{B_s} \phi.$$

Fix  $s$  and select  $p \in \mathbb{R}^d$  such that

$$\frac{1}{s^2} \operatorname{osc}_{B_s} (\phi(y) - p \cdot y) \leq [\phi]_{C_h^{1,1}(0, B_R)}.$$

Then

$$|p| = \frac{1}{2s} \operatorname{osc}_{B_s} (-p \cdot y) \leq \frac{1}{2s} \operatorname{osc}_{B_s} (\phi(y) - p \cdot y) + \frac{1}{2s} \operatorname{osc}_{B_s} \phi.$$

Together these yield

$$\begin{aligned} \frac{2}{s} \operatorname{osc}_{B_{s/2}} \phi &\leq \frac{2}{s} \operatorname{osc}_{B_{s/2}} (\phi(y) - p \cdot y) + \frac{2}{s} \operatorname{osc}_{B_{s/2}} (-p \cdot y) \\ &= \frac{2}{s} \operatorname{osc}_{B_{s/2}} (\phi(y) - p \cdot y) + 2|p| \\ &\leq \frac{3}{s} \operatorname{osc}_{B_s} (\phi(y) - p \cdot y) + \frac{1}{s} \operatorname{osc}_{B_s} \phi \\ &\leq 3s [\phi]_{C_h^{1,1}(0, B_R)} + \frac{1}{s} \operatorname{osc}_{B_s} \phi. \end{aligned}$$

This is (3.25).

We now iterate (3.25) to obtain the conclusion for  $s \in [h, K]$ . By induction, we see that for each  $j \in \mathbb{N}$  with  $R_j := 2^{-j}K \geq h$ ,

$$\begin{aligned} R_j^{-1} \operatorname{osc}_{B_{R_j}} \phi &\leq K^{-1} \operatorname{osc}_{B_K} \phi + 3 \left( \sum_{i=0}^{j-1} R_i \right) [\phi]_{C_h^{1,1}(0, B_R)} \\ &\leq K^{-1} \operatorname{osc}_{B_K} \phi + 6K [\phi]_{C_h^{1,1}(0, B_R)}. \end{aligned}$$

Using (3.24), we deduce that for each  $j \in \mathbb{N}$  with  $R_j := 2^{-j}K \geq h$

$$R_j^{-1} \operatorname{osc}_{B_{R_j}} \phi \leq 7 \left( [\phi]_{C_h^{1,1}(0, B_R)} \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}.$$

For general  $s \in [h, R)$  we may find  $j \in \mathbb{N}$  such that  $R_{j+1} \leq s < R_j$  to get

$$s^{-1} \operatorname{osc}_{B_s} \phi \leq R_{j+1}^{-1} \operatorname{osc}_{B_{R_{j+1}}} \phi \leq 2R_j^{-1} \operatorname{osc}_{B_{R_j}} \phi \leq 14 \left( [\phi]_{C_h^{1,1}(0, B_R)} \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} \phi \right)^{\frac{1}{2}}. \quad \square$$

Equipped with this lemma, we now present the simple proof of Corollary 3.4:

*Proof of Corollary 3.4.* By interpolation, we also obtain (3.21). This follows from (3.4) and Lemma 3.5 as follows:

$$\begin{aligned} [u]_{C_h^{1,1}(0, B_R)} &\leq C [u]_{C_h^{1,1}(0, B_R)}^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} u \right)^{\frac{1}{2}} \\ &\leq C \mathcal{X} \left( K_0 + |f(0)| + R^\sigma [f]_{C^{0,\sigma}(B_R)} + R^{-2} \operatorname{osc}_{B_R} u \right)^{\frac{1}{2}} \left( \operatorname{osc}_{B_R} u \right)^{\frac{1}{2}} \\ &\leq C \mathcal{X} \left( K_0 R + R|f(0)| + R^{1+\sigma} [f]_{C^{0,\sigma}(B_R)} + R^{-1} \operatorname{osc}_{B_R} u \right), \end{aligned}$$

where we used (3.22) in the first line, (3.4) to get the second line and Young's inequality in the last line. Redefining  $\mathcal{X}$  to absorb the constant  $C$ , we obtain (3.21).  $\square$

#### 4. GREEN'S FUNCTION ESTIMATES

We will now use a similar argument to the proof of Theorem 3.1 to obtain estimates on the modified Green's functions  $G_\varepsilon(\cdot, 0)$  which are given by the solutions of:

$$(4.1) \quad \varepsilon^2 G_\varepsilon - \operatorname{tr}(A(x)D^2 G_\varepsilon) = \chi_{B_\ell} \quad \text{in } \mathbb{R}^d.$$

**Proposition 4.1.** *Fix  $s \in (0, d)$ . There exist  $a(d, \lambda, \Lambda) > 0$ ,  $\delta(d, \lambda, \Lambda) > 0$  and an  $\mathcal{F}$ -measurable random variable  $\mathcal{X} : \Omega \rightarrow [1, \infty)$  satisfying*

$$(4.2) \quad \mathbb{E}[\exp(\mathcal{X}^s)] \leq C(s, d, \lambda, \Lambda, \ell) < \infty$$

such that, for every  $\varepsilon \in (0, 1]$  and  $x \in \mathbb{R}^d$ ,

$$(4.3) \quad G_\varepsilon(x, 0) \leq \mathcal{X}^{d-1-\delta} \xi_\varepsilon(x)$$

where  $\xi_\varepsilon(x)$  is defined by:

$$(4.4) \quad \xi_\varepsilon(x) := \exp(-a\varepsilon|x|) \cdot \begin{cases} \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right), & \text{in } d = 2, \\ (1+|x|)^{2-d}, & \text{in } d > 2, \end{cases}$$

and

$$(4.5) \quad \operatorname{osc}_{B_1(x)} G_\varepsilon(\cdot, 0) \leq (T_x \mathcal{X}) \mathcal{X}^{d-1-\delta} (1+|x|)^{1-d} \exp(-a\varepsilon|x|).$$

We emphasize that (4.3) is a *random* estimate, and the proof relies on the homogenization process. In contrast to the situation for divergence form equations, there is no *deterministic* estimate for the decay of the Green's functions. Consider that for a general  $A \in \Omega$ , the Green's function  $G(\cdot, 0; A)$  solves

$$-\operatorname{tr}(AD^2 G(\cdot, 0; A)) = \delta_0 \quad \text{in } \mathbb{R}^d.$$

The solution may behave, for  $|x| \gg 1$ , like a multiple of

$$K_\gamma(x) := \begin{cases} |x|^{-\gamma} & \gamma > 0, \\ \log|x| & \gamma = 0, \\ -|x|^{-\gamma} & \gamma < 0, \end{cases}$$

for any exponent  $\gamma$  in the range

$$\frac{d-1}{\Lambda} - 1 \leq \gamma \leq \Lambda(d-1) - 1.$$

In particular, if  $\Lambda$  is large, then  $\gamma$  may be negative and so  $G(\cdot, 0; A)$  may be bounded near the origin. To see that this range for  $\gamma$  is sharp, it suffices to consider, respectively, the diffusion matrices

$$A_1(x) = \Lambda \frac{x \otimes x}{|x|^2} + \left(I - \frac{x \otimes x}{|x|^2}\right) \quad \text{and} \quad A_2(x) = \frac{x \otimes x}{|x|^2} + \Lambda \left(I - \frac{x \otimes x}{|x|^2}\right).$$

Note that  $A_1$  and  $A_2$  can be altered slightly to be smooth at  $x = 0$  without changing the decay of  $G$  at infinity.

Before we discuss the proof of Proposition 4.1, we mention an interesting application to the *invariant measure* associated with (4.1). Recall that the invariant measure is defined to be the solution  $m_\varepsilon$  of the equation in double-divergence form:

$$\varepsilon^2(m_\varepsilon - 1) - \operatorname{div}(D(A(x)m_\varepsilon)) = 0 \quad \text{in } \mathbb{R}^d.$$

By (4.3), we have that for every  $y \in \mathbb{R}^d$ ,

$$\int_{B_\ell(y)} m_\varepsilon(x) dx \leq \int_{\mathbb{R}^d} G_\varepsilon(x, 0) dx \leq \mathcal{X}^{d-1-\delta}.$$

In particular, we deduce that, for some  $\delta > 0$ ,

$$\mathbb{P}\left[\int_{B_\ell(y)} m_\varepsilon(x) dx > t\right] \leq C \exp\left(-t^{\frac{d}{d-1}+\delta}\right).$$

This gives a very strong bound on the location of particles undergoing a diffusion in the random environment.

We now return to the proof of Proposition 4.1. Without loss of generality, we may change variables and assume that the effective operator  $\bar{A} = I$ . The proof of (4.5) is based on the idea of using homogenization to compare the Green's function for the heterogeneous operator to that of the homogenized operator. The algebraic error estimates for homogenization in Proposition 2.3 are just enough information to show that, with overwhelming probability, the ratio of Green's functions must be bounded at infinity. This is demonstrated by comparing the modified Green's function  $G_\varepsilon(\cdot, 0)$  to a family of carefully constructed test functions.

The test functions  $\{\varphi_R\}_{R \geq C}$  will possess the following three properties:

$$(4.6) \quad \inf_{A \in \Omega} -\operatorname{tr}(A(x)D^2\varphi_R) \geq \chi_{B_\ell} \quad \text{in } B_R,$$

$$(4.7) \quad -\Delta\varphi_R(x) \gtrsim |x|^{-d} \quad \text{in } \mathbb{R}^d \setminus B_{R/2},$$

$$(4.8) \quad \varphi_R(x) \lesssim R^{d-1-\delta}(1 + |x|^{2-d}) \quad \text{in } \mathbb{R}^d \setminus B_R.$$

As we will show, these properties imply, for large enough  $R$  (which will be random and depend on the value of  $\mathcal{X}$  from many different applications of Proposition 2.3), that  $G_\varepsilon(\cdot, 0) \leq \varphi_R$  in  $\mathbb{R}^d$ .

The properties of the barrier function  $\varphi_R$  inside and outside of  $B_R$  will be used to compare with  $G_\varepsilon(\cdot, 0)$  in different ways. If  $G_\varepsilon(\cdot, 0) \not\leq \varphi_R$  then, since they both decay at infinity,  $G_\varepsilon(\cdot, 0) - \varphi_R$  must achieve its global maximum somewhere in  $\mathbb{R}^d$ . Since  $\varphi_R$  is a supersolution of (4.6), this point must lie in  $\mathbb{R}^d \setminus B_R$ . As  $\varphi_R$  is a supersolution of the homogenized equation outside  $B_{R/2}$ , this event is very unlikely for  $R \gg 1$ , by Proposition 2.3. Note that there is a trade-off in our selection of the parameter  $R$ : if  $R$  is relatively large, then  $\varphi_R$  is larger and hence the conclusion  $G_\varepsilon(\cdot, 0) \leq \varphi_R$  is weaker, however the probability that the conclusion fails is also much smaller.

Since the Green's function for the Laplacian has different qualitative behavior in dimensions  $d = 2$  and  $d > 2$ , we split the proof of Proposition 4.1 into these two cases, which are handled in the following subsections.

#### 4.1. Proof of Proposition 4.1: Dimensions three and larger.

**Lemma 4.2.** *Let  $s \in (0, d)$ . Then there exist constants  $C, c, \gamma, \beta > 0$ , depending only on  $(s, d, \lambda, \Lambda, \ell)$ , and a family of continuous functions  $\{\varphi_R\}_{R \geq C}$  satisfying the following: (i) for every  $R \geq C$  and  $x \in \mathbb{R}^d$ ,*

$$(4.9) \quad \varphi_R(x) \leq CR^{d-2+\gamma} (1 + |x|)^{2-d},$$

(ii) there exists a smooth function  $\psi_R$  such that

$$(4.10) \quad \begin{cases} -\Delta\psi_R \geq c|x|^{-2-\beta}\psi_R & \text{in } \mathbb{R}^d \setminus B_{R/2}, \\ \varphi_R \leq \psi_R & \text{in } \mathbb{R}^d \setminus B_{R/2}, \end{cases}$$

and (iii) for each  $R \geq C$  and  $A \in \Omega$ , we have

$$(4.11) \quad -\operatorname{tr}(A(x)D^2\varphi_R) \geq \chi_{B_\ell} \quad \text{in } B_R.$$

*Proof.* Throughout, we fix  $s \in (0, d)$  and let  $C$  and  $c$  denote positive constants which depend only on  $(s, d, \lambda, \Lambda, \ell)$  and may vary in each occurrence.

We define  $\varphi_R$ . For each  $R \geq 4\ell$ , we set

$$\varphi_R(x) := \begin{cases} m_R - \frac{h}{\gamma} (\ell^2 + |x|^2)^{\frac{\gamma}{2}}, & 0 \leq |x| \leq R, \\ k_R |x|^{2-d} \exp\left(-\frac{1}{\beta}|x|^{-\beta}\right), & |x| > R, \end{cases}$$

where we define the following constants:

$$2\beta := \alpha(s, d, \lambda, \Lambda, \ell) > 0$$

is the exponent from Proposition 2.3 with  $\sigma = 1$ ,

$$\gamma := \max\left\{\frac{1}{2}, 1 - \frac{\lambda}{2\Lambda}\right\}$$

$$h := \frac{2}{\lambda}(2\ell)^{2-\gamma}$$

$$k_R := h(d - 2 - 2^\beta R^{-\beta})^{-1} R^{d-2+\gamma} \exp\left(\frac{1}{\beta}2^\beta R^{-\beta}\right)$$

$$m_R := \frac{h}{\gamma} (\ell^2 + R^2)^{\frac{\gamma}{2}} + k_R R^{2-d} \exp\left(-\frac{1}{\beta}R^{-\beta}\right).$$

Notice that the choice of  $m_R$  makes  $\varphi_R$  continuous. We next perform some calculations to guarantee that this choice of  $\varphi_R$  satisfies the above claims.

*Step 1.* We check that for every  $R \geq 4\ell$  and  $x \in \mathbb{R}^d$ , (4.9) holds. Note that  $\beta = \frac{1}{2}\alpha \geq c$  and thus, for every  $R \geq 4\ell$ ,

$$(4.12) \quad c \leq \exp\left(-\frac{1}{\beta}R^{-\beta}\right) \leq 1.$$

For such  $R$ , we also have that since  $d \geq 3$ ,  $(d - 2 - 2^\beta R^{-\beta}) \geq c$ . Moreover, since  $R \geq 4\ell \geq 4$ , this implies that  $(2/R)^\beta \leq 1 - c$ . Using also that  $h \leq C$ , we deduce that for every  $R \geq 4\ell$ ,

$$(4.13) \quad k_R \leq CR^{d-2+\gamma} \quad \text{and} \quad m_R \leq CR^\gamma.$$

For  $|x| > R$ , (4.9) is immediate from the definition of  $\varphi_R$ , (4.12) and (4.13). For  $|x| \leq R$ , we first note that  $\varphi_R$  is a decreasing function in the radial direction and therefore  $\sup_{\mathbb{R}^d} \varphi_R = \varphi_R(0) \leq m_R$ . We then use (4.13) to get, for every  $|x| \leq R$ ,

$$\varphi_R(x) \leq m_R \leq CR^\gamma \leq CR^{d-2+\gamma} (1 + |x|)^{2-d}.$$

This gives (4.9).

*Step 2.* We check that  $\varphi_R$  satisfies

$$(4.14) \quad \varphi_R(x) \leq \psi_R(x) := k_R |x|^{2-d} \exp\left(-\frac{1}{\beta} |x|^{-\beta}\right) \quad \text{in } \mathbb{R}^d \setminus B_{R/2}.$$

Since this holds with equality for  $|x| \geq R$ , we just need to check it in the annulus  $\{R/2 \leq |x| < R\}$ . For this it suffices to show that in this annulus,  $\psi_R - \varphi_R$  is decreasing in the radial direction. Since both  $\psi_R$  and  $\varphi_R$  are decreasing radial functions, we simply need to check that

$$(4.15) \quad |D\varphi_R(x)| < |D\psi_R(x)| \quad \text{for every } x \in B_R \setminus B_{R/2}.$$

We compute, for  $R/2 \leq |x| \leq R$ , since  $\gamma \leq 1$ ,

$$|D\varphi_R(x)| = h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} |x| \leq h|x|^{\gamma-1}$$

and

$$\begin{aligned} |D\psi_R(x)| &= |x|^{-1} (d - 2 - |x|^{-\beta}) \psi_R(x) \\ &= k_R (d - 2 - |x|^{-\beta}) |x|^{1-d} \exp\left(-\frac{1}{\beta} |x|^{-\beta}\right) \\ &\geq k_R (d - 2 - 2^\beta R^{-\beta}) |x|^{1-d} \exp\left(-\frac{1}{\beta} 2^\beta R^{-\beta}\right). \end{aligned}$$

It is now evident that the choice of  $k_R$  ensures that (4.15) holds. This completes the proof of (4.14).

*Step 3.* We check that  $\psi_R$  satisfies

$$(4.16) \quad -\Delta\psi_R(x) \geq c|x|^{-2-\beta}\psi_R(x) \quad \text{in } |x| \geq C.$$

By a direct computation, we find that, for  $x \neq 0$ ,

$$-\Delta\psi_R(x) = |x|^{-2-\beta} (d - 2 + \beta - |x|^{-\beta}) \psi_R(x).$$

This yields (4.16). For future reference, we also note that for every  $|x| > 1$ ,

$$(4.17) \quad |x|^{-2} \operatorname{osc}_{B_{|x|/2}(x)} \psi_R + \sup_{y \in B_{|x|/2}(x)} (|y|^{-1} |D\psi_R(y)|) \leq C|x|^{-2}\psi_R(x).$$

This follows from the computation

$$|D\psi_R(x)| = |x|^{-1}\psi_R(x) (2 - d + |x|^{-\beta}).$$

*Step 4.* We check that (4.11) holds. By a direct computation, we find that for every  $x \in B_R$ ,

$$\begin{aligned} D^2\varphi_R(x) &= -h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} \left( Id + \frac{\gamma-2}{\ell^2 + |x|^2} (x \otimes x) \right) \\ &= -h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} \left( \left( Id - \frac{x \otimes x}{|x|^2} \right) + \frac{\ell^2 - (1-\gamma)|x|^2}{\ell^2 + |x|^2} \left( \frac{x \otimes x}{|x|^2} \right) \right). \end{aligned}$$

Making use of our choice of  $\gamma$ , we see that, for any  $A \in \Omega$  and  $x \in B_R$ ,

$$\begin{aligned} -\operatorname{tr}(A(x)D^2\varphi_R(x)) &\geq h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} ((d-1)\lambda - \Lambda(1-\gamma)(\ell^2 + |x|^2)^{-1}|x|^2) \\ &\geq h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} ((d-1)\lambda - \Lambda(1-\gamma)). \end{aligned}$$

The last expression on the right side is positive since, by the choice of  $\gamma$ ,

$$(d-1)\lambda - \Lambda(1-\gamma) \geq \left(d - \frac{3}{2}\right)\lambda > \lambda > 0,$$

while for  $x \in B_\ell$ , we have, by the choice of  $h$ ,

$$h(\ell^2 + |x|^2)^{\frac{\gamma}{2}-1} ((d-1)\lambda - \Lambda(1-\gamma)) \geq h(2\ell^2)^{\frac{\gamma}{2}-1} \lambda > 1.$$

This completes the proof of (4.11).  $\square$

*Proof of Proposition 4.1 when  $d \geq 3$ .* As before, we fix  $s \in (0, d)$  and let  $C$  and  $c$  denote positive constants which depend only on  $(s, d, \lambda, \Lambda, \ell)$ . We use the notation developed in Lemma 4.2 throughout the proof.

We make one reduction before beginning the main argument. Rather than proving (4.3), it suffices to prove

$$(4.18) \quad \forall x \in \mathbb{R}^d, \quad G_\varepsilon(x, 0) \leq \mathcal{X}^{d-1-\delta} (1 + |x|)^{2-d}.$$

To see this, we notice that

$$G_\varepsilon(x, 0) \leq \left( \sup_{|x| \leq \varepsilon^{-1}} \frac{G_\varepsilon(x, 0)}{(1 + |x|)^{2-d}} \right) \varepsilon^{d-2} \exp(a) \exp(-a\varepsilon|x|) \quad \text{in } \mathbb{R}^d \setminus B_{\varepsilon^{-1}}.$$

Indeed, the right hand side is larger than the left hand side on  $\partial B_{\varepsilon^{-1}}$ , and hence in  $\mathbb{R}^d \setminus B_{\varepsilon^{-1}}$  by the comparison principle and the fact that the right hand side is a supersolution of (2.3) for  $a(d, \lambda, \Lambda) > 0$  (by the proof of Lemma 2.1). We then obtain (4.3) in  $\mathbb{R}^d \setminus B_{\varepsilon^{-1}}$  by replacing  $\mathcal{X}$  by  $C\mathcal{X}$  and  $a$  by  $\frac{1}{2}a$ , using (4.18), and noting that

$$\varepsilon^{d-2} \exp(-a\varepsilon|x|) \leq C|x|^{2-d} \exp\left(-\frac{a}{2}\varepsilon|x|\right) \quad \text{for every } |x| \geq \varepsilon^{-1}.$$

We also get (4.3) in  $B_{\varepsilon^{-1}}$ , with  $\mathcal{X}$  again replaced by  $C\mathcal{X}$ , from (4.18) and the simple inequality

$$\exp(-a\varepsilon|x|) \geq c \quad \text{for every } |x| \leq \varepsilon^{-1}.$$

*Step 1.* We define  $\mathcal{X}$  and check that it has the desired integrability. Let  $\mathcal{Y}$  denote the random variable  $\mathcal{X}$  in the statement of Proposition 2.3 in  $B_R$  with  $s$  as above and  $\sigma = 1$ . Also denote  $\mathcal{Y}_x(A) := \mathcal{Y}(T_x A)$ , which controls the error in balls of radius  $R$  centered at a point  $x \in \mathbb{R}^d$ .

We now define

$$(4.19) \quad \mathcal{X}(A) := \sup \{ |z| : z \in \mathbb{Z}^d, \mathcal{Y}_z(A) \geq 2^d |z|^s \}.$$

The main point is that  $\mathcal{X}$  has the following property by Proposition 2.3: for every  $z \in \mathbb{Z}^d$  with  $|z| > \mathcal{X}$ , and every  $R > \frac{1}{8}|z|$  and  $g \in C^{0,1}(\partial B_R(z))$ , every pair  $u, v \in C(\overline{B_R})$  such that

$$(4.20) \quad \begin{cases} -\operatorname{tr}(A(x)D^2u) \leq 0 \leq -\Delta v & \text{in } B_R(z), \\ u \leq g \leq v & \text{on } \partial B_R(z), \end{cases}$$

must satisfy the estimate

$$(4.21) \quad R^{-2} \sup_{B_R(z)} (u(x) - v(x)) \leq CR^{-\alpha} \left( R^{-2} \operatorname{osc}_{\partial B_R(z)} g + R^{-1} [g]_{C^{0,1}(\partial B_R(z))} \right).$$

Let us check that

$$(4.22) \quad \mathbb{E}[\exp(\mathcal{X}^s)] \leq C(s, d, \lambda, \Lambda, \ell) < \infty.$$

A union bound and stationarity yield, for  $t \geq 1$ ,

$$\begin{aligned} \mathbb{P}[\mathcal{X} > t] &\leq \sum_{z \in \mathbb{Z}^d \setminus B_t} \mathbb{P}[\mathcal{Y}_z \geq 2^d |z|^s] \\ &\leq \sum_{n \in \mathbb{N}, 2^n \geq t} \sum_{z \in B_{2^n} \setminus B_{2^{n-1}}} \mathbb{P}[\mathcal{Y}_z \geq 2^d |z|^s] \\ &\leq C \sum_{n \in \mathbb{N}, 2^n \geq t} 2^{dn} \mathbb{P}[\mathcal{Y} \geq 2^{(n-1)s+d}]. \end{aligned}$$

By Proposition 2.3 and Chebyshev's inequality,

$$\begin{aligned} \sum_{n \in \mathbb{N}, 2^n \geq t} 2^{dn} \mathbb{P}[\mathcal{Y} \geq 2^{(n-1)s+d}] &\leq C \sum_{n \in \mathbb{N}, 2^n \geq t} 2^{dn} \exp(-2^{(n-1)s+d}) \\ &\leq C \exp(-2t^s). \end{aligned}$$

It follows then that

$$\begin{aligned} \mathbb{E}[\exp(\mathcal{X}^s)] &= s \int_0^\infty t^{s-1} \exp(t^s) \mathbb{P}[\mathcal{X} > t] dt \\ &\leq sC \int_0^\infty t^{s-1} \exp(-t^s) dt \leq C. \end{aligned}$$

This yields (4.22).

*Step 2.* We reduce the proposition to the claim that, for every  $R \geq C$ ,

$$(4.23) \quad \left\{ A \in \Omega : \sup_{0 < \varepsilon \leq 1} \sup_{x \in \mathbb{R}^d} (G_\varepsilon(x, 0; A) - \varphi_R(x)) > 0 \right\} \subseteq \{ A \in \Omega : \mathcal{X}(A) > R \}.$$

If (4.23) holds, then by (4.9) we have

$$\begin{aligned} \left\{ A \in \Omega : \sup_{0 < \varepsilon \leq 1} \sup_{x \in \mathbb{R}^d} \left( G_\varepsilon(x, 0; A) - CR^{d-2+\gamma} (1 + |x|)^{2-d} \right) > 0 \right\} \\ \subseteq \{ A \in \Omega : \mathcal{X}(A) > R \}. \end{aligned}$$

However this implies that, for every  $R \geq C$ ,  $0 < \varepsilon \leq 1$  and  $x \in \mathbb{R}^d$ ,

$$G_\varepsilon(x, 0) \leq C \mathcal{X}^{d-2+\gamma} (1 + |x|)^{2-d}.$$

Setting  $\delta := 1 - \gamma \geq c(d, \lambda, \Lambda) > 0$ , we obtain (4.18).

*Step 3.* We prove (4.23). Fix  $A \in \Omega$ ,  $0 < \varepsilon \leq 1$  and  $R \geq 10\sqrt{d}$  for which

$$\sup_{\mathbb{R}^d} (G_\varepsilon(\cdot, 0) - \varphi_R) > 0.$$

The goal is to show that  $\mathcal{X} \geq R$ , at least if  $R \geq C$ . We do this by exhibiting  $|z| > R$  and functions  $u$  and  $v$  satisfying (4.20), but not (4.21).

As the functions  $G_\varepsilon(\cdot, 0)$  and  $\varphi_R$  decay at infinity (c.f. Lemma 2.1), there exists a point  $x_0 \in \mathbb{R}^d$  such that

$$G_\varepsilon(x_0, 0) - \varphi_R(x_0) = \sup_{\mathbb{R}^d} (G_\varepsilon(\cdot, 0) - \varphi_R) > 0.$$

By the maximum principle and (4.11), it must be that  $|x_0| \geq R$ . By (4.14),

$$(4.24) \quad G_\varepsilon(x_0, 0) - \psi_R(x_0) = \sup_{B_{|x_0|/2}(x_0)} (G_\varepsilon(\cdot, 0) - \psi_R).$$

We perturb  $\psi_R$  by setting  $\tilde{\psi}_R(x) := \psi_R(x) + c|x_0|^{-2-\beta}\psi_R(x_0)|x - x_0|^2$  which, in view of (4.16), satisfies

$$-\Delta \tilde{\psi}_R \geq 0 \quad \text{in } B_{|x_0|/2}(x_0).$$

The perturbation improves (4.24) to

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_R(x_0) \geq \sup_{\partial B_{|x_0|/2}(x_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_R) + c|x_0|^{-\beta}\psi_R(x_0).$$

If  $R \geq C$ , then we may take  $z_0 \in \mathbb{Z}^d$  to be the nearest lattice point to  $x_0$  such that  $|z_0| > |x_0|$  and get  $x_0 \in B_{|z_0|/4}(z_0)$ . Since  $\psi(|x|)$  is decreasing in  $|x|$ , this implies

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_R(x_0) \geq \sup_{\partial B_{|z_0|/4}(z_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_R) + c|z_0|^{-\beta}\psi_R(z_0).$$

In view of (4.17), this gives

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_R(x_0) \geq \sup_{\partial B_{|z_0|/4}(z_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_R) + c\Gamma|z_0|^{-\beta},$$

where

$$\Gamma := \operatorname{osc}_{\partial B_{|z_0|/4}(z_0)} \tilde{\psi}_R + |z_0| [\tilde{\psi}_R]_{C^{0,1}(\partial B_{|z_0|/4}(z_0))}.$$

We have thus found functions satisfying (4.20) but in violation of (4.21). That is, we deduce from the definition of  $\mathcal{Y}_z$  that  $\mathcal{Y}_{z_0} \geq c|z_0|^{s+\alpha-\beta} - C$  and, in view of the fact that  $\beta = \frac{1}{2}\alpha < \alpha$  and  $|z_0| > R$ , this implies that  $\mathcal{Y}_{z_0} \geq 2^d|z_0|^s$  provided  $R \geq C$ . Hence  $\mathcal{X} \geq |z_0| > R$ . This completes the proof of (4.23).  $\square$

**4.2. Dimension two.** The argument for (4.3) in two dimensions follows along similar lines as the proof when  $d \geq 3$ , however the normalization of  $G_\varepsilon$  is more tricky since it depends on  $\varepsilon$ .



**Lemma 4.3.** *Let  $s \in (0, 2)$ . Then there exist constants  $C, c, \gamma, \beta > 0$ , depending only on  $(s, d, \lambda, \Lambda, \ell)$ , and a family of continuous functions  $\{\varphi_{R,\varepsilon}\}_{R \geq C, \varepsilon \leq c}$  satisfying the following: (i) for every  $R \geq C$ ,*

$$(4.25) \quad \varphi_{R,\varepsilon}(x) \leq CR^\gamma \log \left( 2 + \frac{1}{\varepsilon(1+|x|)} \right) \exp(-a\varepsilon|x|),$$

(ii) there exists a smooth function  $\psi_{R,\varepsilon}$  such that

$$\begin{cases} -\Delta \psi_{R,\varepsilon} \geq c|x|^{-2-\beta} \psi_{R,\varepsilon} & \text{in } B_{2\varepsilon^{-1}} \setminus B_{R/2}, \\ \varphi_{R,\varepsilon} \leq \psi_{R,\varepsilon} & \text{in } B_{2\varepsilon^{-1}} \setminus B_{R/2}, \end{cases}$$

and (iii) for every  $R \geq C$  and  $A \in \Omega$ ,

$$(4.26) \quad -\operatorname{tr}(A(x)D^2\varphi_{R,\varepsilon}) \geq \chi_{B_\ell} \quad \text{in } B_R \cup (\mathbb{R}^2 \setminus B_{\varepsilon^{-1}}).$$

*Proof.* Throughout we assume  $d = 2$ , we fix  $s \in (0, 2)$  and let  $C$  and  $c$  denote positive constants which depend only on  $(s, \lambda, \Lambda, \ell)$  and may vary in each occurrence. We roughly follow the outline of the proof of (4.3) above in the case of  $d \geq 3$ .

*Step 1.* The definition of  $\varphi_R$ . For  $\varepsilon \in (0, \frac{1}{2}]$ ,  $4\ell \leq R \leq \varepsilon^{-1}$  and  $x \in \mathbb{R}^2$ , we set

$$\varphi_{R,\varepsilon}(x) := \begin{cases} m_{R,\varepsilon} - \frac{h}{\gamma} (\ell^2 + |x|^2)^{\frac{\gamma}{2}}, & 0 \leq |x| \leq R, \\ k_R \left( \frac{1}{a} \exp(a) + |\log \varepsilon| - \log |x| \right) \exp\left(-\frac{1}{\beta}|x|^{-\beta}\right), & R < |x| \leq \frac{1}{\varepsilon}, \\ b_{R,\varepsilon} \exp(-a\varepsilon|x|), & |x| > \frac{1}{\varepsilon}, \end{cases}$$

where the constants are defined as follows:

$$\begin{aligned} a &:= a(\lambda, \Lambda) \quad \text{is the constant from Lemma 2.1,} \\ 2\beta &:= \alpha(s, \lambda, \Lambda, \ell) > 0 \quad \text{is the exponent from Proposition 2.3 with } \sigma = 1, \\ \gamma &:= \max \left\{ \frac{1}{2}, 1 - \frac{\lambda}{2\Lambda} \right\}, \\ h &:= \frac{2}{\lambda} (2\ell)^{2-\gamma}, \\ k_R &:= 2h \exp\left(\frac{1}{\beta} 2^\beta R^{-\beta}\right) R^\gamma, \\ m_{R,\varepsilon} &:= \frac{h}{\gamma} (\ell^2 + R^2)^{\frac{\gamma}{2}} + k_R \left( \frac{1}{a} \exp(a) + |\log \varepsilon| - \log R \right) \exp\left(-\frac{1}{\beta} R^{-\beta}\right), \\ b_{R,\varepsilon} &:= \frac{1}{a} k_R \exp\left(2a - \frac{1}{\beta} \varepsilon^\beta\right) \end{aligned}$$

Observe that

$$(4.27) \quad k_R \leq CR^\gamma, \quad m_{R,\varepsilon} \leq CR^\gamma (1 + |\log \varepsilon| - \log R) \quad \text{and} \quad b_{R,\varepsilon} \leq CR^\gamma.$$

*Step 2.* We show that, for every  $\varepsilon \in (0, \frac{1}{2}]$ ,  $4\ell \leq R \leq \varepsilon^{-1}$  and  $x \in \mathbb{R}^2$ , (4.25) holds. This is relatively easy to check from the definition of  $\varphi_{R,\varepsilon}$ , using (4.12) and (4.27). For  $x \in B_R$ , we use  $\varphi_{R,\varepsilon} \leq m_{R,\varepsilon}$ , (4.27) and  $\exp(-a\varepsilon R) \geq c$  to

immediately obtain (4.25). For  $x \in B_{\varepsilon^{-1}} \setminus B_R$ , the estimate is obtained from the definition of  $\varphi_{R,\varepsilon}$ , the bound for  $k_R$  in (4.27) and (4.12). For  $x \in \mathbb{R}^2 \setminus B_{\varepsilon^{-1}}$ , the logarithm factor on the right side of (4.25) is  $C$  and we get (4.25) from the bound for  $b_{R,\varepsilon}$  in (4.27).

*Step 3.* We show that, for  $\varepsilon \in (0, c]$  and  $R \geq C$ , we have

$$(4.28) \quad \varphi_{R,\varepsilon}(x) \leq \psi_{R,\varepsilon}(x) := k_R \left( \frac{1}{a} \exp(a) + |\log \varepsilon| - \log |x| \right) \exp \left( -\frac{1}{\beta} |x|^{-\beta} \right) \\ \text{in } \frac{1}{2}R \leq |x| \leq \frac{2}{\varepsilon}.$$

We have equality in (4.28) for  $R \leq |x| \leq \varepsilon^{-1}$  by the definition of  $\varphi_{R,\varepsilon}$ . As  $\varphi_{R,\varepsilon}$  is radial, it therefore suffices to check that the magnitude of the radial derivative of  $\varphi_{R,\varepsilon}$  is less than (respectively, greater than) than that of  $\psi_{R,\varepsilon}$  in the annulus  $\{R/2 \leq |x| \leq R\}$  (respectively,  $\{\varepsilon^{-1} \leq |x| \leq 2\varepsilon^{-1}\}$ ). This is ensured by the definitions of  $k_R$  and  $b_{R,\varepsilon}$ , as the following routine computation verifies: first, in  $x \in B_R \setminus B_{R/2}$ , we have

$$|D\varphi_{R,\varepsilon}(x)| = h \left( \ell^2 + |x|^2 \right)^{\frac{\gamma}{2}-1} |x| < h|x|^{\gamma-1},$$

and thus in  $B_R \setminus B_{R/2}$ , provided  $R \geq C$ , we have that

$$|D\psi_{R,\varepsilon}(x)| = k_R |x|^{-1} \left| -1 + |x|^{-\beta} \left( \frac{1}{a} \exp(a) + |\log \varepsilon| - \log |x| \right) \right| \exp \left( -\frac{1}{\beta} |x|^{-\beta} \right) \\ > \frac{1}{2} k_R |x|^{-1} \exp \left( -\frac{1}{\beta} 2^\beta R^{-\beta} \right) = h R^\gamma |x|^{-1} \geq h|x|^{\gamma-1} > |D\varphi_{R,\varepsilon}(x)|.$$

Next we consider  $x \in B_{2\varepsilon^{-1}} \setminus B_{\varepsilon^{-1}}$  and estimate

$$|D\varphi_{R,\varepsilon}(x)| = a\varepsilon b_{R,\varepsilon} \exp(-a\varepsilon|x|) > a\varepsilon b_{R,\varepsilon} \exp(-2a) = 2\varepsilon k_R \exp \left( -\frac{1}{\beta} \varepsilon^\beta \right)$$

and

$$|D\psi_{R,\varepsilon}(x)| \leq k_R |x|^{-1} \left( 1 + \frac{1}{a} \exp(a) |x|^{-\beta} \right) \exp \left( -\frac{1}{\beta} \frac{\varepsilon^\beta}{2^\beta} \right) \leq 2\varepsilon k_R \exp \left( -\frac{1}{\beta} \varepsilon^\beta \right),$$

the latter holding provided that  $\varepsilon \leq c$ . This completes the proof of (4.28).

*Step 4.* We show that  $\psi_{R,\varepsilon}$  satisfies

$$(4.29) \quad -\Delta \psi_{R,\varepsilon} \geq c|x|^{-2-\beta} \psi_{R,\varepsilon}(x) \quad \text{in } C \leq |x| \leq \frac{2}{\varepsilon}.$$

By a direct computation, for every  $x \in \mathbb{R}^2 \setminus \{0\}$ , we have

$$-\Delta \psi_{R,\varepsilon}(x) = |x|^{-2-\beta} \left( (\beta - |x|^{-\beta}) \psi_{R,\varepsilon}(x) + k_R \left( \frac{1}{|x|^2} + 1 \right) \exp \left( -\frac{1}{\beta} |x|^{-\beta} \right) \right) \\ \geq |x|^{-2-\beta} (\beta - |x|^{-\beta}) \psi_{R,\varepsilon}(x).$$

From  $\beta \geq c$  and the definition of  $\psi_{R,\varepsilon}$ , we see that  $\psi_{R,\varepsilon} > 0$  and  $(\beta - |x|^{-\beta}) \geq c$  for every  $\beta^{-1/\beta} \leq |x| \leq 2\varepsilon^{-1}$ . This yields (4.29).

For future reference, we note that, for every  $|x| \leq 2\varepsilon^{-1}$ ,

$$(4.30) \quad |x|^{-2} \operatorname{osc}_{B_{|x|/2}(x)} \psi_{R,\varepsilon} + \sup_{y \in B_{|x|/2}(x)} (|y|^{-1} D\psi_{R,\varepsilon}(y)) \leq C|x|^{-2} \leq C|x|^{-2} \psi_{R,\varepsilon}(x).$$

*Step 5.* We check (4.26) by checking that for every  $A \in \Omega$ ,

$$(4.31) \quad -\operatorname{tr}\left(A(x)D^2\varphi_{R,\varepsilon}\right) \geq \chi_{B_\ell} \quad \text{in } B_R \cup (\mathbb{R}^2 \setminus B_{\varepsilon^{-1}}).$$

In fact, for  $B_R$ , the computation is identical to the one which established (4.11), since our function  $\varphi_{R,\varepsilon}$  here is the same as  $\varphi_R$  (from the argument in the case  $d > 2$ ) in  $B_R$ , up to a constant. Therefore we refer to Step 5 in Lemma 4.2 for details. In  $\mathbb{R}^2 \setminus B_{\varepsilon^{-1}}$ ,  $\varphi_{R,\varepsilon}$  is a supersolution by the proof of Lemma 2.1 and by the choice of  $a$ .

We remark that in the case that  $R = \varepsilon^{-1}$ , the middle annulus in the definition of  $\varphi_{R,\varepsilon}$  disappears, and we have that  $\varphi_{R,\varepsilon}$  is a global (viscosity) solution of (4.31), that is,

$$(4.32) \quad -\operatorname{tr}\left(A(x)D^2\varphi_{R,\varepsilon}\right) \geq \chi_{B_\ell} \quad \text{in } \mathbb{R}^2 \quad \text{if } R = \varepsilon^{-1}.$$

To see why, by (4.26), we need only check that  $\varphi_{R,\varepsilon}$  is a viscosity supersolution of (4.26) on the sphere  $\partial B_R = \partial B_{\varepsilon^{-1}}$ . However, the function  $\varphi_{R,\varepsilon}$  cannot be touched from below on this sphere, since its inner radial derivative is smaller than its outer radial derivative by the computations in Step 3. Therefore we have (4.32). It follows by comparison that, for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^2$ ,

$$(4.33) \quad \begin{aligned} G_\varepsilon(x, 0) &\leq \varphi_{R,\varepsilon}(x) \quad \text{if } R = \varepsilon^{-1} \\ &\leq C\varepsilon^{-\gamma} \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-a\varepsilon|x|) \end{aligned}$$

□

*Proof of Proposition 4.1 when  $d = 2$ .* The proof follows very similar to the case when  $d \geq 3$ , using the appropriate adaptations for the new test function introduced in Lemma 4.3. As before, we fix  $s \in (0, 2)$  and let  $C$  and  $c$  denote positive constants which depend only on  $(s, \lambda, \Lambda, \ell)$ . We use the notation developed in Lemma 4.3 throughout the proof.

*Step 1.* We define the random variable  $\mathcal{X}$  in exactly the same way as in (4.19), so

$$(4.34) \quad \mathcal{X}(A) := \sup\{|z| : z \in \mathbb{Z}^d, \mathcal{Y}_z(A) \geq 2^d|z|^s\},$$

The argument leading to (4.22) follows exactly as before, so that

$$\mathbb{E}[\exp(\mathcal{X}^s)] \leq C(s, \lambda, \Lambda, \ell) < \infty.$$

*Step 2.* We reduce the proposition to the claim that, for every  $R \geq C$ ,

$$(4.35) \quad \left\{ A \in \Omega : \sup_{0 < \varepsilon < \frac{1}{R}} \sup_{x \in \mathbb{R}^2} (G_\varepsilon(x, 0, A) - \varphi_{R,\varepsilon}(x)) > 0 \right\} \subseteq \{A \in \Omega : \mathcal{X}(A) > R\}.$$

If (4.35) holds, then by (4.25) we have

$$\left\{ A : \sup_{0 < \varepsilon < \frac{1}{R}} \sup_{x \in \mathbb{R}^2} \left( G_\varepsilon(x, 0, A) - CR^\gamma \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-a\varepsilon|x|) \right) > 0 \right\} \subseteq \{A \in \Omega : \mathcal{X}(A) > R\}.$$

From this, we deduce that, for every  $0 < \varepsilon < \mathcal{X}^{-1}$  and  $x \in \mathbb{R}^2$ ,

$$(4.36) \quad G_\varepsilon(x, 0) \leq C\mathcal{X}^\gamma \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-a\varepsilon|x|).$$

Moreover, if  $\varepsilon \in (0, \frac{1}{2}]$  and  $\varepsilon \geq \mathcal{X}^{-1}$ , then by (4.33) we have

$$\begin{aligned} G_\varepsilon(x, 0) &\leq C\varepsilon^{-\gamma} \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-a\varepsilon|x|) \\ &\leq C\mathcal{X}^\gamma \log\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-a\varepsilon|x|). \end{aligned}$$

Thus we have (4.36) for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^2$ . Taking  $\delta := 1 - \gamma \geq c$ , we obtain the desired conclusion (4.3) for  $d = 2$ .

*Step 3.* We prove (4.35). This is almost the same as the last step in the proof of (4.23) for dimensions larger than two. Fix  $A \in \Omega$ ,  $0 < \varepsilon \leq 1$  and  $R \geq 2\ell$  for which

$$\sup_{\mathbb{R}^d} (G_\varepsilon(\cdot, 0) - \varphi_{R,\varepsilon}) > 0.$$

As in the case of dimensions larger than two, we select a point  $x_0 \in \mathbb{R}^d$  such that

$$G_\varepsilon(x_0, 0) - \varphi_{R,\varepsilon}(x_0) = \sup_{\mathbb{R}^d} (G_\varepsilon(\cdot, 0) - \varphi_{R,\varepsilon}) > 0.$$

By the maximum principle and (4.26), it must be that  $R \leq |x_0| \leq \varepsilon^{-1}$ . By (4.28),

$$(4.37) \quad G_\varepsilon(x_0, 0) - \psi_{R,\varepsilon}(x_0) = \sup_{B_{|x_0|/2}(x_0)} (G_\varepsilon(\cdot, 0) - \psi_{R,\varepsilon}).$$

We perturb  $\psi_{R,\varepsilon}$  by setting  $\tilde{\psi}_{R,\varepsilon}(x) := \psi_{R,\varepsilon}(x) + c|x_0|^{-2-\beta}\psi_{R,\varepsilon}(x_0)|x - x_0|^2$  which, in view of (4.29), satisfies

$$-\Delta \tilde{\psi}_{R,\varepsilon} \geq 0 \quad \text{in } B_{|x_0|/2}(x_0).$$

According to (4.24), we have

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_{R,\varepsilon}(x_0) \geq \sup_{\partial B_{|x_0|/2}(z_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_{R,\varepsilon}) + c|x_0|^{-\beta}\psi_{R,\varepsilon}(x_0).$$

Assuming  $R \geq C$ , we may take  $z_0 \in \mathbb{Z}^d$  to be the nearest lattice point to  $x_0$  such that  $|z_0| > |x_0|$  and deduce that  $x_0 \in B_{|z_0|/4}$  as well as

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_{R,\varepsilon}(x_0) \geq \sup_{\partial B_{|z_0|/4}(z_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_{R,\varepsilon}) + c|z_0|^{-\beta}\psi_{R,\varepsilon}(z_0).$$

In view of (4.30), this gives

$$G_\varepsilon(x_0, 0) - \tilde{\psi}_{R,\varepsilon}(x_0) \geq \sup_{\partial B_{|z_0|/4}(z_0)} (G_\varepsilon(\cdot, 0) - \tilde{\psi}_{R,\varepsilon}) + c\Gamma|z_0|^{-\beta},$$

where

$$\Gamma := |z_0|^{-2} \operatorname{osc}_{\partial B_{|z_0|/4}(z_0)} \tilde{\psi}_{R,\varepsilon} + |z_0|^{-1} [\tilde{\psi}_{R,\varepsilon}]_{C^{0,1}(\partial B_{|z_0|/4}(z_0))}.$$

We have thus found functions satisfying (4.20) but in violation of (4.21), that is, we deduce from the definition of  $\mathcal{Y}_z$  that  $\mathcal{Y}_{z_0} \geq c|z_0|^{s+\alpha-\beta} - C$ . In view of the fact that  $\beta = \frac{1}{2}\alpha < \alpha$  and  $|z_0| > R$ , this implies that  $\mathcal{Y}_{z_0} \geq 2^d|z_0|^s$  provided

$R \geq C$ . By the definition of  $\mathcal{X}$ , we obtain  $\mathcal{X} \geq |z_0| > R$ . This completes the proof of (4.35).  $\square$

## 5. SENSITIVITY ESTIMATES

In this section, we present an estimate which uses the Green's function bounds to control the vertical derivatives introduced in the spectral gap inequality (Proposition 2.2). Recall the notation from Proposition 2.2:

$$X'_z := \mathbb{E}_* [X | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})] \quad \text{and} \quad \mathbb{V}_*[X] := \sum_{z \in \mathbb{Z}^d} (X - X'_z)^2.$$

The vertical derivative  $(X - X'_z)$  measures, in a precise sense, the *sensitivity* of  $X$  subject to changes in the environment near  $z$ . We can therefore interpret  $(X - X'_z)$  as a derivative of  $X$  with respect to the coefficients near  $z$ . The goal then will be to understand the vertical derivative  $(X - X'_z)$  when  $X$  is  $\phi_\varepsilon(x)$ , for fixed  $x \in \mathbb{R}^d$ .

The main result of this section is the following proposition which computes  $\phi_\varepsilon(x) - \mathbb{E}_*[\phi_\varepsilon(x) | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})]$  in terms of the random variable introduced in Proposition 4.1. Throughout the rest of the section, we fix  $M \in \mathbb{S}^d$  with  $|M| = 1$  and let  $\xi_\varepsilon$  be defined as in (4.4).

**Proposition 5.1.** *Fix  $s \in (0, d)$ . There exist positive constants  $a(d, \lambda, \Lambda) > 0$ ,  $\delta(d, \lambda, \Lambda) > 0$  and an  $\mathcal{F}_*$ -measurable random variable  $\mathcal{X} : \Omega_* \rightarrow [1, \infty)$  satisfying*

$$(5.1) \quad \mathbb{E}[\exp(\mathcal{X}^s)] \leq C(s, d, \lambda, \Lambda, \ell) < \infty$$

such that, for every  $\varepsilon \in (0, \frac{1}{2}]$ ,  $x \in \mathbb{R}^d$  and  $z \in \mathbb{Z}^d$ ,

$$(5.2) \quad |\phi_\varepsilon(x) - \mathbb{E}_*[\phi_\varepsilon(x) | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})]| \leq (T_z \mathcal{X})^{d+1-\delta} \xi_\varepsilon(x-z).$$

Before beginning the proof of Proposition 5.1, we first provide a heuristic explanation of the main argument. We begin with the observation that we may identify the conditional expectation  $\mathbb{E}_*[X | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})]$  via resampling in the following way. Let  $(\Omega'_*, \mathcal{F}'_*, \mathbb{P}'_*)$  denote an independent copy of  $(\Omega_*, \mathcal{F}_*, \mathbb{P}_*)$  and define, for each  $z \in \mathbb{Z}^d$ , a map

$$\theta'_z : \Omega \times \Omega' \rightarrow \Omega$$

by

$$\theta'_z(\omega, \omega')(y) := \begin{cases} \omega(y) & \text{if } y \neq z, \\ \omega'(z) & \text{if } y = z. \end{cases}$$

It follows that, for every  $\omega \in \Omega$ ,

$$(5.3) \quad \mathbb{E}_*[X | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})](\omega) = \mathbb{E}'_*[X(\theta'_z(\omega, \cdot))].$$

Therefore, we are interested in estimating differences of the form  $X(\omega) - X(\theta'_z(\omega, \omega'))$ , which represent the expected change in  $X$  if we resample the environment at  $z$ . Observe that, by (1.16), if  $\omega, \omega' \in \Omega_*$ ,  $z \in \mathbb{R}^d$ , and  $A := \pi(\omega)$  and  $A' := \pi(\theta'_z(\omega, \omega'))$ , then

$$(5.4) \quad A \equiv A' \quad \text{in } \mathbb{R}^d \setminus B_{\ell/2}(z).$$

Denote by  $\phi_\varepsilon$  and  $\phi'_\varepsilon$  the corresponding approximate correctors with modified Green's functions  $G_\varepsilon$  and  $G'_\varepsilon$ . Let  $w := \phi_\varepsilon - \phi'_\varepsilon$ . Then we have

$$\begin{aligned} \varepsilon^2 w - \operatorname{tr}(A'(x)D^2w) &= \operatorname{tr}((A(x) - A'(x))(M + D^2\phi_\varepsilon)) \\ &\leq d\Lambda(1 + |D^2\phi_\varepsilon(x)|)\chi_{B_{\ell/2}}(x), \end{aligned}$$

where  $\chi_E$  denotes the characteristic function of a set  $E \subseteq \mathbb{R}^d$ . By comparing  $w$  to  $G'_\varepsilon$ , we deduce that, for  $C(d, \lambda, \Lambda) \geq 1$ ,

$$(5.5) \quad \phi_\varepsilon(x) - \phi'_\varepsilon(x) \leq C(1 + [\phi_\varepsilon]_{C^{1,1}(B_{\ell/2}(z))})G'_\varepsilon(x, z).$$

If  $\phi_\varepsilon$  satisfied a  $C^{1,1}$  bound, then by (3.20) and (4.3), we deduce that

$$\phi_\varepsilon(0) - \phi'_\varepsilon(0) \leq C((T_z\mathcal{X})(\omega))^2((T_z\mathcal{Y})(\theta'_z(\omega, \omega')))^{d-1-\delta}\xi_\varepsilon(z),$$

for  $\mathcal{X}$  defined as in (3.20),  $\mathcal{Y}$  defined as in (4.3), and  $\xi_\varepsilon(z)$  defined as in (4.4).

Taking expectations of both sides with respect to  $\mathbb{P}'_*$ , we obtain

$$(5.6) \quad \phi_\varepsilon(0) - \mathbb{E}_*[\phi_\varepsilon(0) | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})] \leq C(T_z\mathcal{X})^2(T_z\mathcal{Y}_*)^{d-1-\delta}\xi_\varepsilon(z),$$

where

$$(5.7) \quad \mathcal{Y}_* := \mathbb{E}'_*[\mathcal{Y}(\theta'_0(\omega, \omega'))^{d-1-\delta}]^{1/(d-1-\delta)}.$$

Jensen's inequality implies that the integrability of  $\mathcal{Y}_*$  is controlled by the integrability of  $\mathcal{Y}$ . First, consider that, for  $s \geq d-1-\delta$ , by the convexity of  $t \mapsto \exp(t^r)$  for  $r \geq 1$ , we have

$$\begin{aligned} \mathbb{E}[\exp(\mathcal{Y}_*^s)] &= \mathbb{E}\left[\exp\left(\mathbb{E}'_*[\mathcal{Y}(\theta'_0(\omega, \omega'))^{d-1-\delta}]^{s/(d-1-\delta)}\right)\right] \\ &\leq \mathbb{E}[\mathbb{E}'_*[\exp(\mathcal{Y}(\theta'_0(\omega, \omega'))^s)]] \\ &= \mathbb{E}[\exp(\mathcal{Y}^s)]. \end{aligned}$$

Integrability of lower moments for  $s \in (0, d-1-\delta)$  follows by the bound

$$\mathbb{E}[\exp(\mathcal{Y}_*^s)] = \mathbb{E}\left[\exp(\mathcal{Y}_*^{d-1-\delta})\right] \sup|\exp(-\mathcal{Y}_*^{d-1-\delta} + \mathcal{Y}_*^s)| \leq \mathbb{E}\left[\exp(\mathcal{Y}_*^{d-1-\delta})\right]$$

by the monotonicity of the map  $p \mapsto x^p$  for  $p \geq 0$  and  $x \geq 1$ , which we can take without loss of generality by letting  $\mathcal{Y}_* = \mathcal{Y}_* + 1$ . We may now redefine  $\mathcal{X}$  to be  $\mathcal{X} + \mathcal{Y}_*$  to get one side of the desired bound (5.2). The analogous bound from below is obtained by exchanging  $M$  for  $-M$  in the equation for  $\phi_\varepsilon$ , or by repeating the above argument and comparing  $w$  to  $-G'_\varepsilon$ .

The main reason that this argument fails to be rigorous is technical: the quantity  $[\phi_\varepsilon]_{C^{1,1}(B_{\ell/2}(z))}$  is not actually controlled by Theorem 3.1, rather we have control only over the coarsened quantity  $[\phi_\varepsilon]_{C^{1,1}(B_{\ell/2}(z))}$ . Most of the work in the proof of Proposition 5.1 is therefore to fix this glitch by proving that (5.5) still holds if we replace the Hölder seminorm on the right side by the appropriate coarsened seminorm. This is handled by the following lemma, which we write in a rather general form:

**Lemma 5.2.** *Let  $\varepsilon \in (0, \frac{1}{\ell}]$  and  $z \in \mathbb{Z}^d$  and suppose  $A, A' \in \Omega$  satisfy (5.4). Also fix  $f, f' \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  and let  $u, u' \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$  be the solutions of*

$$\begin{cases} \varepsilon^2 u - \operatorname{tr}(A(x)D^2 u) = f & \text{in } \mathbb{R}^d, \\ \varepsilon^2 u' - \operatorname{tr}(A'(x)D^2 u') = f' & \text{in } \mathbb{R}^d. \end{cases}$$

*Then there exists a constant  $C(d, \lambda, \Lambda, \ell) \geq 1$  such that, for every  $x \in \mathbb{R}^d$  and  $\delta > 0$ ,*

$$(5.8) \quad u(x) - u'(x) \leq C \left( \delta + [u]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} + \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (f(y') - f(y) - \delta \varepsilon^2 |y' - z|^2) \right) G'_\varepsilon(x, z) + \sum_{y \in \mathbb{Z}^d} G_\varepsilon(x, y) \sup_{B_{\ell/2}(y)} (f - f').$$

*Here  $G'_\varepsilon$  denotes the modified Green's function for  $A'$ .*

*Proof.* We may assume without loss of generality that  $z = 0$ . By replacing  $u(x)$  by the function

$$u''(x) := u(x) - \sum_{y \in \mathbb{Z}^d} G_\varepsilon(x, y) \sup_{B_{\ell/2}(y)} (f - f'),$$

we may assume furthermore that  $f' \geq f$  in  $\mathbb{R}^d$ . Fix  $\varepsilon \in (0, \frac{1}{2}]$ . We will show that

$$\sup_{x \in \mathbb{R}^d} (u(x) - u'(x) - K G'_\varepsilon(x, 0)) > 0$$

implies a contradiction for

$$K > C \left( 1 + [u]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} + \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (f(y') - f(y) - \varepsilon^2 |y' - z|^2) \right) G'_\varepsilon(x, z)$$

and  $C = C(d, \lambda, \Lambda, \ell)$  chosen sufficiently large.

*Step 1.* We find a touching point  $x_0 \in B_{\ell/2}$ . Consider the auxiliary function

$$\xi(x) := u(x) - u'(x) - K G'_\varepsilon(x, 0).$$

By (5.4) and using that  $f' \geq f$ , we see that  $\xi$  satisfies

$$\varepsilon^2 \xi - \operatorname{tr}(A(x)D^2 \xi) \leq 0 \quad \text{in } \mathbb{R}^d \setminus B_{\ell/2}.$$

By the maximum principle and the hypothesis,  $\sup_{\mathbb{R}^d} \xi = \sup_{B_{\ell/2}} \xi > 0$ . Select  $x_0 \in B_{\ell/2}$  such that

$$(5.9) \quad \xi(x_0) = \sup_{\mathbb{R}^d} \xi.$$

*Step 2.* We replace  $u$  by a quadratic approximation in  $B_\ell$  and get a new touching point. Select  $p \in \mathbb{R}^d$  such that

$$(5.10) \quad \sup_{x \in B_\ell} |u(x) - u(x_0) - p \cdot (x - x_0)| \leq 4\ell^2 [u]_{C_1^{1,1}(0, B_{1/\varepsilon})}.$$

Fix  $\nu \geq 1$  to be chosen below and define the function

$$\psi(x) := u(x_0) + p \cdot (x - x_0) - \nu |x - x_0|^2 - u'(x) - K G'_\varepsilon(x, 0),$$

The claim is that

$$(5.11) \quad x \mapsto \psi(x) \quad \text{has a local maximum in } B_\ell.$$

To verify (5.11), we check that  $\psi(x_0) > \sup_{\partial B_\ell} \psi$ . For  $y \in \partial B_\ell$ , we compute

$$\begin{aligned} \psi(x_0) &= u(x_0) - u'(x_0) - KG'_\varepsilon(x_0, 0) \\ &\geq u(y) - u'(y) - KG'_\varepsilon(y, 0) && \text{(by (5.9))} \\ &\geq u(x_0) + p \cdot (y - x_0) - 8\ell^2 [u]_{C_1^{1,1}(0, B_{1/\varepsilon})} - u'(y) - KG'_\varepsilon(y, 0) && \text{(by (5.10))} \\ &= \psi(y) + \nu|y - x_0|^2 - 8\ell^2 [u]_{C_1^{1,1}(0, B_{1/\varepsilon})} \\ &\geq \psi(y) + \ell^2 \nu - 8\ell^2 [u]_{C_1^{1,1}(0, B_{1/\varepsilon})}, \end{aligned}$$

where in the last line we used  $|y - x_0| \geq \frac{\ell}{2}$ . Therefore, for every

$$\nu > 8 [u]_{C_1^{1,1}(0, B_{1/\varepsilon})},$$

the claim (5.11) is satisfied.

*Step 3.* We show that, for every  $x \in \mathbb{R}^d$ ,

$$(5.12) \quad u(x) \leq C\delta\varepsilon^{-2} + |x|^2 + \sup_{y \in \mathbb{R}^d} (\varepsilon^{-2} f(y) - |y|^2)$$

Define

$$w(x) := u(x) - (|x|^2 + L), \quad \text{for } L := \sup_{x \in \mathbb{R}^d} (\varepsilon^{-2} f(x) - \delta|x|^2) + 2d\Lambda\varepsilon^{-2}\delta.$$

Using the equation for  $u$ , we find that

$$\varepsilon^2 w - \text{tr}(A(x)D^2 w) \leq f - \varepsilon^2(|x|^2 + L) + 2d\Lambda\delta.$$

Using the definition of  $L$ , we deduce that

$$\varepsilon^2 w - \text{tr}(A(x)D^2 w) \leq 0 \quad \text{in } \mathbb{R}^d.$$

Since  $w(x) \rightarrow -\infty$  as  $|x| \rightarrow \infty$ , we deduce from the maximum principle that  $w \leq 0$  in  $\mathbb{R}^d$ . This yields (5.12).

*Step 4.* We conclude by obtaining a contradiction to (5.11) for an appropriate choice of  $K$ . Observe that, in  $B_\ell$ , the function  $\psi$  satisfies

$$\begin{aligned} \varepsilon^2 \psi - \text{tr}(A'(x)D^2 \psi) &\leq \varepsilon^2(u(x_0) + p \cdot (x - x_0) - \nu|x - x_0|^2) + C\nu - f'(x) - K \\ &\leq \varepsilon^2 u(x) + C\nu - f(x) - K \\ &\leq C\nu + \sup_{y \in \mathbb{R}^d} (\delta + f(y) - f(x) - \varepsilon^2|y|^2) - K. \end{aligned}$$

Thus (5.11) violates the maximum principle provided that

$$K > C(\nu + \delta) + \sup_{x \in B_\ell, y \in \mathbb{R}^d} (f(y) - f(x) - \varepsilon^2|y|^2).$$

This completes the proof.  $\square$

We now use the previous lemma and the estimates in Sections 3 and 4 to prove the sensitivity estimates (5.2).



*Proof of Proposition 5.1.* Fix  $s \in (0, d)$ . Throughout,  $C$  and  $c$  will denote positive constants depending only on  $(s, d, \lambda, \Lambda, \ell)$  which may vary in each occurrence, and  $\mathcal{X}$  denotes an  $\mathcal{F}_*$ -measurable random variable on  $\Omega_*$  satisfying

$$\mathbb{E}[\exp(\mathcal{X}^p)] \leq C \quad \text{for every } p < s,$$

which we also allow to vary in each occurrence.

We fix  $z \in \mathbb{Z}^d$  and identify the conditional expectation with respect to  $\mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})$  via resampling, as in (5.3). By the discussion following the statement of the proposition, it suffices to prove the bound (5.6) with  $\mathcal{Y}_*$  defined by (5.7) for some random variable  $\mathcal{Y} \leq \mathcal{X}$ . To that end, fix  $\omega \in \Omega$ ,  $\omega' \in \Omega'$ , and denote  $\tilde{\omega} := \theta_z(\omega, \omega')$  as well as  $A = \pi(\omega)$  and  $A' = \pi(\tilde{\omega})$ . Note that (5.4) holds. Also let  $\phi_\varepsilon$  and  $\phi'_\varepsilon$  denote the approximate correctors and  $G_\varepsilon$  and  $G'_\varepsilon$  the Green's functions.

*Step 1.* We use Theorem 3.1 to estimate  $[\phi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))}$ . In preparation, we rewrite the equation for  $\phi_\varepsilon$  in terms of

$$w_\varepsilon(x) := \frac{1}{2}x \cdot Mx + \phi_\varepsilon(x),$$

which satisfies

$$-\operatorname{tr}(A(x)D^2w_\varepsilon) = -\varepsilon^2\phi_\varepsilon(x).$$

In view of the fact that the constant functions  $\pm \sup_{x \in \mathbb{R}^d} |\operatorname{tr}(A(x)M)|$  are super/subsolutions, we have

$$(5.13) \quad \sup_{x \in \mathbb{R}^d} \varepsilon^2 |\phi_\varepsilon(x)| \leq \sup_{x \in \mathbb{R}^d} |\operatorname{tr}(A(x)M)| \leq d\Lambda \leq C,$$

and this yields

$$(5.14) \quad \varepsilon^2 \operatorname{osc}_{B_{4/\varepsilon}} w_\varepsilon \leq \varepsilon^2 \operatorname{osc}_{B_{4/\varepsilon}} \frac{1}{2}x \cdot Mx + \varepsilon^2 \operatorname{osc}_{B_{4/\varepsilon}} \phi_\varepsilon \leq C.$$

By the Krylov-Safonov Holder estimate (3.2) applied to  $w_\varepsilon$  in  $B_R$  with  $R = 4\varepsilon^{-1}$  yields

$$(5.15) \quad \varepsilon^{2-\beta} [w_\varepsilon]_{C^{0,\beta}(B_{2/\varepsilon})} \leq C.$$

Letting  $Q(x) := \frac{1}{2}x \cdot Mx$ , we have

$$(5.16) \quad \varepsilon^2 [\phi_\varepsilon]_{C^{0,\beta}(B_{2/\varepsilon})} \leq \varepsilon^2 [w_\varepsilon]_{C^{0,\beta}(B_{2/\varepsilon})} + \varepsilon^2 [Q]_{C^{0,\beta}(B_{2/\varepsilon})} \leq C\varepsilon^\beta + C|M|\varepsilon^\beta \leq C\varepsilon^\beta.$$

We now apply Theorem 3.1 (specifically (3.20)) to  $w_\varepsilon$  with  $R = 2\varepsilon^{-1}$  to obtain

$$\begin{aligned} [w_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} &\leq (T_z \mathcal{X})^2 \left( \sup_{B_{2/\varepsilon}} \varepsilon^2 \phi + \varepsilon^{-\beta} [\varepsilon^2 \phi_\varepsilon]_{C^{0,\beta}(B_{2/\varepsilon})} + \varepsilon^2 \operatorname{osc}_{B_{2/\varepsilon}} w_\varepsilon \right) \\ &\leq C(T_z \mathcal{X})^2. \end{aligned}$$

As  $w_\varepsilon$  and  $\phi_\varepsilon$  differ by the quadratic  $Q$ , we obtain

$$(5.17) \quad [\phi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} \leq C((T_z \mathcal{X})^2 + |M|) \leq C(T_z \mathcal{X})^2.$$

*Step 2.* We estimate  $G'_\varepsilon(\cdot, z)$  and complete the argument for (5.6). By Proposition 4.1, we have

$$(5.18) \quad G'_\varepsilon(x, z) \leq (T_z \mathcal{X})^{d-1-\delta} \xi_\varepsilon(x-z)$$

for  $\xi_\varepsilon(x-z)$  defined as in (4.4). Lemma 5.2 yields, for every  $x \in \mathbb{R}^d$ ,

$$|\phi_\varepsilon(x) - \phi'_\varepsilon(x)| \leq C \left( 1 + [\phi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} \right) G'_\varepsilon(x, z) + 2d\Lambda G_\varepsilon(x, z).$$

Inserting (5.17) and (5.18) gives

$$(5.19) \quad |\phi_\varepsilon(x) - \phi'_\varepsilon(x)| \leq C (T_z \mathcal{X})^2 (T_z \mathcal{X})^{d-1-\delta} \xi_\varepsilon(x-z).$$

This is (5.6).  $\square$

## 6. OPTIMAL SCALING FOR THE APPROXIMATE CORRECTORS

We complete the rate computation for the approximate correctors  $\phi_\varepsilon$ . We think of breaking up the decay of  $\varepsilon^2 \phi_\varepsilon(0) - \text{tr}(\overline{AM})$  into two main contributions of error:

$$\varepsilon^2 \phi_\varepsilon(0) - \text{tr}(\overline{AM}) = \underbrace{\varepsilon^2 \phi_\varepsilon(0) - \mathbb{E}[\varepsilon^2 \phi_\varepsilon(0)]}_{\text{“random error”}} + \underbrace{\mathbb{E}[\varepsilon^2 \phi_\varepsilon(0)] - \text{tr}(\overline{AM})}_{\text{“deterministic error”}}.$$

The “random error” will be controlled by the concentration inequalities established in Section 5. We will show that the “deterministic error” is controlled by the random error, and this will yield a rate for  $\varepsilon^2 \phi_\varepsilon(0) + \text{tr}(\overline{AM})$ .

First, we control the random error using Proposition 2.2 and the estimates from the previous three sections.

**Proposition 6.1.** *There exist  $\delta(d, \lambda, \Lambda) > 0$  and  $C(d, \lambda, \Lambda, \ell) \geq 1$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$ , and  $x \in \mathbb{R}^d$ ,*

$$(6.1) \quad \mathbb{E} \left[ \exp \left( \left( \frac{1}{\mathcal{E}(\varepsilon)} |\varepsilon^2 \phi_\varepsilon(x) - \mathbb{E}[\varepsilon^2 \phi_\varepsilon(x)]| \right)^{\frac{1}{2}+\delta} \right) \right] \leq C.$$

*Proof.* For readability, we prove (6.1) for  $x = 0$ . The argument for general  $x \in \mathbb{R}^d$  is almost the same. Define

$$\xi_\varepsilon(x) := \exp(-a\varepsilon|x|) \cdot \begin{cases} \log \left( 2 + \frac{1}{\varepsilon(1+|x|)} \right) & \text{if } d = 2, \\ (1+|x|)^{2-d} & \text{if } d \geq 3. \end{cases}$$

According to Proposition 5.1, for  $\beta > 0$ ,

$$\begin{aligned} & \exp \left( C \left( \mathbb{V}_* \left[ \frac{\varepsilon^2 \phi_\varepsilon(0)}{\mathcal{E}(\varepsilon)} \right] \right)^\beta \right) \\ &= \exp \left( C \left( \frac{\varepsilon^4}{\mathcal{E}(\varepsilon)^2} \sum_{z \in \mathbb{Z}^d} (\phi_\varepsilon(0) - \mathbb{E}_*[\phi_\varepsilon(0) | \mathcal{F}_*(\mathbb{Z}^d \setminus \{z\})])^2 \right)^\beta \right) \\ &\leq \exp \left( C \left( \frac{\varepsilon^4}{\mathcal{E}(\varepsilon)^2} \sum_{z \in \mathbb{Z}^d} (T_z \mathcal{X})^{2d+2-2\delta} \xi_\varepsilon(z)^2 \right)^\beta \right). \end{aligned}$$

We claim (and prove below) that

$$(6.2) \quad \sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2 \leq C \varepsilon^{-4} \mathcal{E}(\varepsilon)^2.$$

Assuming (6.2) and applying Jensen's inequality for discrete sums, we have

$$\begin{aligned} & \exp\left(C \left(\frac{\varepsilon^4}{\mathcal{E}(\varepsilon)^2} \sum_{z \in \mathbb{Z}^d} (T_z \mathcal{X})^{2d+2-2\delta} \xi_\varepsilon(z)^2\right)^\beta\right) \\ & \leq \exp\left(C \left(\frac{\sum_{z \in \mathbb{Z}^d} (T_z \mathcal{X})^{2d+2-2\delta} \xi(z)^2}{\sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2}\right)^\beta\right) \\ & \leq \frac{\sum_{z \in \mathbb{Z}^d} \xi(z)^2 \exp\left(C (T_z \mathcal{X})^{(2d+2-2\delta)\beta}\right)}{\sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2}. \end{aligned}$$

Select  $\beta := d/(2d+2-3\delta)$ . Taking expectations, using stationarity, and applying Proposition 5.1, we obtain

$$\mathbb{E}_* \left[ \exp\left(C \left(\mathbb{V}_* \left[\frac{\varepsilon^2 \phi_\varepsilon(0)}{\mathcal{E}(\varepsilon)}\right]\right)^\beta\right) \right] \leq C.$$

Finally, an application of Proposition 2.2 gives, for  $\gamma := 2\beta/(1+\beta) \in (0, 2)$ ,

$$\mathbb{E} \left[ \exp\left(\left|\frac{\varepsilon^2 \phi_\varepsilon(0)}{\mathcal{E}(\varepsilon)} - \mathbb{E} \left[\frac{\varepsilon^2 \phi_\varepsilon(0)}{\mathcal{E}(\varepsilon)}\right]\right|^\gamma\right) \right] \leq C.$$

This completes the proof of the proposition, subject to the verification of (6.2), which is a straightforward computation. In dimension  $d \geq 3$ , we have

$$\begin{aligned} \sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2 & \leq C \int_{\mathbb{R}^d} (1+|x|)^{4-2d} \exp(-2a\varepsilon|x|) dx \\ & = C \varepsilon^{d-4} \int_{\mathbb{R}^d} (\varepsilon+|y|)^{4-2d} \exp(-2a|y|) dy \\ & \leq C \varepsilon^{d-4} \left( \int_{\mathbb{R}^d \setminus B_\varepsilon} |y|^{4-2d} \exp(-2a|y|) dy + \int_{B_\varepsilon} \varepsilon^{4-2d} dy \right) \\ & = C \cdot \begin{cases} 1 + \varepsilon^{d-4} & \text{in } d \neq 4, \\ 1 + |\log \varepsilon| & \text{in } d = 4, \end{cases} \\ & = C \varepsilon^{-4} \mathcal{E}(\varepsilon)^2. \end{aligned}$$

In dimension  $d = 2$ , we have

$$\begin{aligned} \sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2 & \leq C \int_{\mathbb{R}^d} \log^2\left(2 + \frac{1}{\varepsilon(1+|x|)}\right) \exp(-2a\varepsilon|x|) dx \\ & \leq C \varepsilon^{-2} \int_{\mathbb{R}^d} \log^2\left(2 + \frac{1}{\varepsilon+|y|}\right) \exp(-2a|y|) dy \\ & \leq C \varepsilon^{-2} \left( \int_{B_\varepsilon} \log^2\left(2 + \frac{1}{\varepsilon}\right) \exp(-2a|y|) dy \right. \\ & \quad \left. + \int_{\mathbb{R}^d \setminus B_\varepsilon} \log^2\left(2 + \frac{1}{|y|}\right) \exp(-2a|y|) dy \right). \end{aligned}$$

We estimate the two integrals on the right as follows:

$$\int_{B_\varepsilon} \log^2 \left( 2 + \frac{1}{\varepsilon} \right) \exp(-2a|y|) dy \leq |B_\varepsilon| \log^2 \left( 2 + \frac{1}{\varepsilon} \right) \leq C\varepsilon^2 \log^2 \left( 2 + \frac{1}{\varepsilon} \right)$$

and

$$\begin{aligned} \int_{\mathbb{R}^d \setminus B_\varepsilon} \log^2 \left( 2 + \frac{1}{|y|} \right) \exp(-2a|y|) dy \\ \leq |B_1| \log^2 \left( 2 + \frac{1}{\varepsilon} \right) + C \int_{\mathbb{R}^d \setminus B_1} \exp(-2a|y|) dy \leq C \log^2 \left( 2 + \frac{1}{\varepsilon} \right). \end{aligned}$$

Assembling the last three sets of inequalities yields in  $d = 2$  that

$$\sum_{z \in \mathbb{Z}^d} \xi_\varepsilon(z)^2 \leq C\varepsilon^{-2} \log^2 \left( 2 + \frac{1}{\varepsilon} \right) \leq C\varepsilon^{-4} \mathcal{E}(\varepsilon)^2. \quad \square$$

Next, we show that the deterministic error is controlled from above by the random error. The basic idea is to argue that if the deterministic error is larger than the typical size of the random error, then this is inconsistent with the homogenization. The argument must of course be quantitative, so it is natural that we will apply Proposition 2.3. Note that if we possessed the bound  $\sup_{x \in \mathbb{R}^d} |\phi_\varepsilon(x) - \mathbb{E}[\phi_\varepsilon(0)]| \lesssim \varepsilon^{-2} \mathcal{E}(\varepsilon)$ , then our proof here would be much simpler. However, this bound is too strong—we do not have, and of course cannot expect, such a uniform estimate on the fluctuations to hold—and therefore we need to cut off larger fluctuations and argue by approximation. This is done by using the Alexandrov-Bakelman-Pucci estimate and (6.1) in a straightforward way.

**Proposition 6.2.** *There exists  $C(d, \lambda, \Lambda, \ell) \geq 1$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^d$ ,*

$$(6.3) \quad |\mathbb{E}[\varepsilon^2 \phi_\varepsilon(x)] - \text{tr}(\overline{AM})| \leq C\mathcal{E}(\varepsilon).$$

*Proof.* By symmetry, it suffices to prove the following one-sided bound: for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^d$ ,

$$(6.4) \quad \text{tr}(\overline{AM}) - \mathbb{E}[\varepsilon^2 \phi_\varepsilon(x)] \geq -C\mathcal{E}(\varepsilon).$$

The proof of (6.4) will be broken down into several steps.

*Step 1.* We show that

$$(6.5) \quad \mathbb{E} \left[ \exp \left( \left( \frac{1}{\varepsilon^{-2} \mathcal{E}(\varepsilon)} \left( \text{osc}_{B_{\sqrt{d}}} \phi_\varepsilon \right) \right)^{\frac{1}{2} + \delta} \right) \right] \leq C.$$

Let  $k \in \mathcal{L}$  be the affine function satisfying

$$\sup_{x \in B_{\sqrt{d}}} |\phi_\varepsilon - k(x)| = \inf_{l \in \mathcal{L}} \sup_{x \in B_{\sqrt{d}}} |\phi_\varepsilon - l(x)|.$$

According to (5.17),

$$(6.6) \quad \sup_{x \in B_{\sqrt{d}}} |\phi_\varepsilon - k(x)| \leq C\mathcal{X}^2.$$

Since  $k$  is affine, its slope can be estimated by its oscillation on  $B_{\sqrt{d}} \cap \mathbb{Z}^d$ :

$$|\nabla k| \leq C \text{osc}_{B_{\sqrt{d}} \cap \mathbb{Z}^d} k.$$

The previous line and (6.6) yield that

$$|\nabla k| \leq C \operatorname{osc}_{B_{\sqrt{d}} \cap \mathbb{Z}^d} \phi_\varepsilon + C \mathcal{X}^2.$$

By stationarity and (6.1), we get

$$(6.7) \quad \mathbb{E} \left[ \exp \left( \left( \frac{1}{\varepsilon^{-2} \mathcal{E}(\varepsilon)} \left( \operatorname{osc}_{B_{\sqrt{d}} \cap \mathbb{Z}^d} \phi_\varepsilon \right) \right)^{\frac{1}{2} + \delta} \right) \right] \leq C$$

Therefore,

$$\mathbb{E} \left[ \exp \left( \left( \frac{1}{\varepsilon^{-2} \mathcal{E}(\varepsilon)} |\nabla k| \right)^{\frac{1}{2} + \delta} \right) \right] \leq C$$

The triangle inequality, (6.6) and (6.7) imply (6.5).

*Step 2.* Consider the function

$$f_\varepsilon(x) := \left( -\varepsilon^2 \phi_\varepsilon(x) + \mathbb{E}[\varepsilon^2 \phi_\varepsilon(0)] \right)_+.$$

We claim that, for every  $R \geq 1$ ,

$$(6.8) \quad \mathbb{E} \left[ \left( \int_{B_R} |f_\varepsilon(x)|^d dx \right)^{\frac{1}{d}} \right] \leq C \mathcal{E}(\varepsilon).$$

Indeed, by Jensen's inequality, (6.1) and (6.5),

$$\mathbb{E} \left[ \left( \int_{B_R} |f_\varepsilon(x)|^d dx \right)^{\frac{1}{d}} \right] \leq \left( \int_{B_R} \mathbb{E}[|f_\varepsilon(x)|^d] dx \right)^{\frac{1}{d}} \leq C \mathcal{E}(\varepsilon).$$

*Step 3.* We prepare the comparison. Define

$$\widehat{f}_\varepsilon(x) := \min \{ -\varepsilon^2 \phi_\varepsilon(x), \mathbb{E}[-\varepsilon^2 \phi_\varepsilon(0)] \} = -\varepsilon^2 \phi_\varepsilon(x) - f_\varepsilon(x),$$

fix  $R \geq 1$  (we will send  $R \rightarrow \infty$  below) and denote by  $\widehat{\phi}_\varepsilon$ , the solution of

$$\begin{cases} -\operatorname{tr}(A(x)(M + D^2 \widehat{\phi}_\varepsilon)) = \widehat{f}_\varepsilon & \text{in } B_R, \\ \widehat{\phi}_\varepsilon = d\Lambda \varepsilon^{-2} |M| & \text{on } \partial B_R. \end{cases}$$

Note that the boundary condition was chosen so that  $\phi_\varepsilon \leq \widehat{\phi}_\varepsilon$  on  $\partial B_R$ . Thus the Alexandrov-Backelman-Pucci estimate and (6.8) yield

$$(6.9) \quad \mathbb{E} \left[ R^{-2} \sup_{B_R} (\phi_\varepsilon - \widehat{\phi}_\varepsilon) \right] \leq C \mathcal{E}(\varepsilon).$$

*Step 4.* Let  $\widehat{\phi}$  denote the solution to

$$(6.10) \quad \begin{cases} -\operatorname{tr}(\overline{A}(M + D^2 \widehat{\phi})) = \mathbb{E}[-\varepsilon^2 \phi_\varepsilon(0)] & \text{in } B_R, \\ \widehat{\phi} = d\Lambda \varepsilon^{-2} |M| & \text{on } \partial B_R. \end{cases}$$

Notice that the right hand side and boundary condition for (6.10) are chosen to be constant. Moreover, we can solve for  $\widehat{\phi}$  explicitly: for  $x \in B_R$ , we have

$$(6.11) \quad \widehat{\phi}(x) = d\Lambda \varepsilon^{-2} |M| - \frac{|x|^2 - R^2}{2 \operatorname{tr} \overline{A}} \left( \operatorname{tr}(\overline{A}M) - \mathbb{E}[\varepsilon^2 \phi_\varepsilon(0)] \right).$$

We point out that since  $\widehat{f}_\varepsilon \leq \mathbb{E}[-\varepsilon^2\phi_\varepsilon(0)]$ , we have that  $\widehat{\phi}_\varepsilon$  satisfies

$$\begin{cases} -\operatorname{tr}(A(x)(M + D^2\widehat{\phi}_\varepsilon)) \leq \mathbb{E}[-\varepsilon^2\phi_\varepsilon(0)] & \text{in } B_R, \\ \widehat{\phi}_\varepsilon = d\Lambda\varepsilon^{-2}|M| & \text{on } \partial B_R. \end{cases}$$

It follows then that we may apply Proposition 2.3 to the pair  $\frac{1}{2}x \cdot Mx + \widehat{\phi}_\varepsilon(x)$  and  $\frac{1}{2}x \cdot Mx + \widehat{\phi}(x)$ , which gives

$$(6.12) \quad \mathbb{E} \left[ R^{-2} \sup_{x \in B_R} (\widehat{\phi}_\varepsilon - \widehat{\phi}) \right] \leq CR^{-\alpha}.$$

*Step 5.* The conclusion. We have, by (6.9), (6.12) and (6.11),

$$\begin{aligned} -d\Lambda\varepsilon^{-2}|M| &\leq \mathbb{E}[\phi_\varepsilon(0)] \\ &\leq \mathbb{E}[\widehat{\phi}_\varepsilon(0)] + C\mathcal{E}(\varepsilon)R^2 \\ &\leq \widehat{\phi}(0) + CR^{2-\alpha} + C\mathcal{E}(\varepsilon)R^2 \\ &= d\Lambda\varepsilon^{-2}|M| + \frac{R^2}{2\operatorname{tr}A} (\operatorname{tr}(\overline{AM}) - \varepsilon^2\mathbb{E}[\phi_\varepsilon(0)]) + CR^{2-\alpha} + C\mathcal{E}(\varepsilon)R^2. \end{aligned}$$

Rearranging, we obtain

$$\operatorname{tr}(\overline{AM}) - \varepsilon^2\mathbb{E}[\phi_\varepsilon(0)] \geq C[-2d\Lambda|M|\varepsilon^{-2}R^{-2} - CR^{-\alpha} - C\mathcal{E}(\varepsilon)].$$

Sending  $R \rightarrow \infty$  yields

$$\operatorname{tr}(\overline{AM}) - \mathbb{E}[\varepsilon^2\phi_\varepsilon(0)] \geq -C\mathcal{E}(\varepsilon).$$

Since (6.5) and stationarity implies that, for every  $x \in \mathbb{R}^d$ ,

$$|\mathbb{E}[\varepsilon^2\phi_\varepsilon(x)] - \mathbb{E}[\varepsilon^2\phi_\varepsilon(0)]| \leq C\mathcal{E}(\varepsilon),$$

the proof of (6.4) is complete.  $\square$

The proof of Theorem 1.1 is now complete, as it follows immediately from Propositions 6.1, 6.2 and (6.5).

## 7. EXISTENCE OF STATIONARY CORRECTORS IN $d > 4$

In this section we prove the following result concerning the existence of stationary correctors in dimensions larger than four.

**Theorem 7.1.** *Suppose  $d > 4$  and fix  $M \in \mathbb{S}^d$ ,  $|M| = 1$ . Then there exists a constant  $C(d, \lambda, \Lambda, \ell) \geq 1$  and a stationary function  $\phi$  belonging  $\mathbb{P}$ -almost surely to  $C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ , satisfying*

$$(7.1) \quad -\operatorname{tr}(A(x)(M + D^2\phi)) = -\operatorname{tr}(\overline{AM}) \quad \text{in } \mathbb{R}^d$$

and, for each  $x \in \mathbb{R}^d$  and  $t \geq 1$ , the estimate

$$(7.2) \quad \mathbb{P}[|\phi(x)| > t] \leq C \exp\left(-t^{\frac{1}{2}}\right).$$

To prove Theorem 7.1, we argue that, after subtracting an appropriate constant,  $\phi_\varepsilon$  has an almost sure limit as  $\varepsilon \rightarrow 0$  to a stationary function  $\phi$ . We introduce the functions

$$\widehat{\phi}_\varepsilon := \phi_\varepsilon - \frac{1}{\varepsilon^2} \operatorname{tr}(\overline{AM})$$

Observe that

$$(7.3) \quad \varepsilon^2 \widehat{\phi}_\varepsilon - \operatorname{tr}(A(x)D^2 \widehat{\phi}_\varepsilon) = -\operatorname{tr}(\overline{AM}).$$

To show that  $\widehat{\phi}_\varepsilon$  has an almost sure limit as  $\varepsilon \rightarrow 0$ , we introduce the functions

$$\psi_\varepsilon := \phi_\varepsilon - \phi_{2\varepsilon}.$$

Then

$$\widehat{\phi}_\varepsilon - \widehat{\phi}_{2\varepsilon} = \psi_\varepsilon - \frac{3}{4\varepsilon^2} \operatorname{tr}(\overline{AM}),$$

and the goal will be to prove bounds on  $\widehat{\phi}_\varepsilon - \widehat{\phi}_\varepsilon$  which are summable over the sequence  $\varepsilon_n := 2^{-n}$ . We proceed as in the previous section: we first estimate the fluctuations of  $\psi_\varepsilon$  using a sensitivity estimate and a suitable version of the Efron-Stein inequality. We then use this fluctuation estimate to obtain bounds on its expectation using a variation of the argument in the proof of Proposition 6.2.

We begin with controlling the fluctuations.

**Lemma 7.2.** *For every  $p \in [1, \infty)$  and  $\gamma > 0$ , there exists  $C(p, \gamma, d, \lambda, \Lambda, \ell) < \infty$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^d$ ,*

$$(7.4) \quad \mathbb{E}[|\psi_\varepsilon(x) - \mathbb{E}[\psi_\varepsilon(x)]|^p]^{1/p} \leq C\varepsilon^{(\frac{d-4}{2}\wedge 2)-\gamma}.$$

*Proof.* In view of the Efron-Stein inequality for  $p$ th moments (cf. (A.2)), it suffices to show that

$$(7.5) \quad \mathbb{E}\left[\mathbb{V}_*[\psi_\varepsilon(x)]^{p/2}\right]^{1/p} \leq C\varepsilon^{(\frac{d-4}{2}\wedge 2)-\gamma}.$$

We start from the observation that  $\psi_\varepsilon$  satisfies the equation

$$(7.6) \quad \varepsilon^2 \psi_\varepsilon - \operatorname{tr}(A(x)D^2 \psi_\varepsilon) = 3\varepsilon^2 \phi_{2\varepsilon} \quad \text{in } \mathbb{R}^d.$$

Denote the right-hand side by  $h_\varepsilon := 3\varepsilon^2 \phi_{2\varepsilon}$ .

*Step 1.* We outline the proof of (7.5). Fix  $z \in \mathbb{R}^d$ . We use the notation from the previous section, letting  $A' := \pi(\theta'_z(\omega, \omega'))$  denote a resampling of the coefficients at  $z$ . We let  $\psi'_\varepsilon, \phi'_\varepsilon, G'_\varepsilon$ , etc, denote the corresponding functions defined with respect to  $A'$ . Applying Lemma 5.2 with  $\delta = \varepsilon^2$ , in view of (7.6), we find that

$$\begin{aligned} & \psi_\varepsilon(x) - \psi'_\varepsilon(x) \\ & \leq C \left( \varepsilon^2 + [\psi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} + \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (h_\varepsilon(y') - h_\varepsilon(y) - \varepsilon^4 |y' - z|^2) \right) G'_\varepsilon(x, z) \\ & \quad + \sum_{y \in \mathbb{Z}^d} G_\varepsilon(x, y) \sup_{B_{\ell/2}(y)} (h_\varepsilon - h'_\varepsilon) \\ & =: CK(z) \xi_\varepsilon(x - z) + C \sum_{y \in \mathbb{Z}^d} H(y, z) \xi_\varepsilon(z - y) \xi_\varepsilon(x - y). \end{aligned}$$

Here we have defined

$$K(z) := \xi_\varepsilon(x-z)^{-1} \left( \varepsilon^2 + [\psi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} + \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (h_\varepsilon(y') - h_\varepsilon(y) - \varepsilon^4 |y' - z|^2) \right) G'_\varepsilon(x, z)$$

and

$$H(y, z) := (\xi_\varepsilon(z-y)\xi_\varepsilon(x-y))^{-1} G_\varepsilon(x, y) \sup_{B_{\ell/2}(y)} (h_\varepsilon - h'_\varepsilon).$$

These are random variables on the probability space  $\Omega \times \Omega'$  with respect to the probability measure  $\tilde{\mathbb{P}} := \mathbb{P}_* \times \mathbb{P}'_*$ . Below we will check that, for each  $p \in [1, \infty)$  and  $\gamma > 0$ , there exists  $C(p, \gamma, d, \lambda, \Lambda, \ell) < \infty$  such that

$$(7.7) \quad \tilde{\mathbb{E}}[K(z)^p]^{\frac{1}{p}} + \tilde{\mathbb{E}}[H(y, z)^p]^{\frac{1}{p}} \leq C\varepsilon^{2-\gamma}.$$

We first complete the proof of (7.5) assuming that (7.7) holds. In view of the discussion in Section 5, we compute:

$$\begin{aligned} & \mathbb{E} \left[ \mathbb{V}_* [\psi_\varepsilon(x)]^{\frac{p}{2}} \right] \\ & \leq \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} \left( CK(z)\xi_\varepsilon(x-z) + C \sum_{y \in \mathbb{Z}^d} H(y, z)\xi_\varepsilon(z-y)\xi_\varepsilon(x-y) \right)^2 \right)^{\frac{p}{2}} \right] \\ & \leq C \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} [K(z)\xi_\varepsilon(x-z)]^2 \right)^{\frac{p}{2}} \right] \\ & \quad + C \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} H(y, z)\xi_\varepsilon(z-y)\xi_\varepsilon(x-y) \right)^2 \right)^{\frac{p}{2}} \right]. \end{aligned}$$

By Jensen's inequality, (6.2) and (7.7),

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} [K(z)\xi_\varepsilon(x-z)]^2 \right)^{\frac{p}{2}} \right] & \leq \left( \sum_{z \in \mathbb{Z}^d} \xi_\varepsilon^2(x-z) \right)^{\frac{p}{2}-1} \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} K(z)^p \xi_\varepsilon^2(x-z) \right) \right] \\ & \leq C\varepsilon^{(2-\gamma)p} \left( \sum_{z \in \mathbb{Z}^d} \xi_\varepsilon^2(x-z) \right)^{\frac{p}{2}} \leq C\varepsilon^{(2-\gamma)p} \end{aligned}$$



and by Jensen's inequality and (7.7),

$$\begin{aligned}
& \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} H(y, z) \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \right)^2 \right)^{\frac{p}{2}} \right] \\
&= \mathbb{E} \left[ \left( \sum_{z \in \mathbb{Z}^d} \sum_{y, y' \in \mathbb{Z}^d} H(y, z) H(y', z) \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \xi_\varepsilon(z-y') \xi_\varepsilon(x-y') \right)^{\frac{p}{2}} \right] \\
&\leq \left( \sum_{z, y, y' \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \xi_\varepsilon(z-y') \xi_\varepsilon(x-y') \right)^{\frac{p}{2}-1} \\
&\quad \times \mathbb{E} \left[ \sum_{z, y, y' \in \mathbb{Z}^d} H(y, z)^{\frac{p}{2}} H(y', z)^{\frac{p}{2}} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \xi_\varepsilon(z-y') \xi_\varepsilon(x-y') \right] \\
&= \left( \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \right)^2 \right)^{\frac{p}{2}-1} \\
&\quad \times \sum_{z, y, y' \in \mathbb{Z}^d} \mathbb{E} [H(y, z)^{\frac{p}{2}} H(y', z)^{\frac{p}{2}}] \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \xi_\varepsilon(z-y') \xi_\varepsilon(x-y') \\
&\leq \left( \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \right)^2 \right)^{\frac{p}{2}} C \varepsilon^{(2-\gamma)p}.
\end{aligned}$$

In view of the inequality

$$(7.8) \quad \left( \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \right)^2 \right)^{\frac{1}{2}} \leq C (1 + \varepsilon^{-1})^{4-\frac{d}{2}}.$$

which we also will prove below, the demonstration of (7.5) is complete.

To complete the proof, it remains to prove (7.7) and (7.8).

*Step 2.* Proof of (7.8). We first show that, for every  $x, z \in \mathbb{R}^d$ ,

$$(7.9) \quad \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \leq C \exp(-c\varepsilon|x-z|) \left( (1+|x-z|)^{4-d} + \varepsilon^{d-4} \right).$$

Recall that in dimensions  $d > 4$ ,  $\xi_\varepsilon(x) = \exp(-a\varepsilon|x|)(1+|x|)^{2-d}$ . Denote  $r := |x-z|$ . We estimate the sum by an integral and then split the sum into five pieces:

$$\begin{aligned}
& \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) \\
& \leq C \int_{\mathbb{R}^d} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \leq C \int_{B_{r/4}(x)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy + C \int_{B_{r/4}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \quad + C \int_{B_{2r}(z) \setminus (B_{r/4}(z) \cup B_{r/4}(x))} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \quad + C \int_{B_{1/\varepsilon}(z) \setminus B_{2r}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \quad + C \int_{\mathbb{R}^d \setminus B_{1/\varepsilon}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy.
\end{aligned}$$

We now estimate each of the above terms. Observe first that

$$\begin{aligned}
& \int_{B_{r/4}(x)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy + \int_{B_{r/4}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \leq C \exp(-c\varepsilon r) \int_{B_r(0)} (1+r)^{2-d}(1+|y|)^{2-d} dy = C \exp(-c\varepsilon r)(1+r)^{4-d}.
\end{aligned}$$

Next, we estimate

$$\begin{aligned}
& \int_{B_{2r}(z) \setminus (B_{r/4}(z) \cup B_{r/4}(x))} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \leq C \exp(-c\varepsilon r) \int_{B_{2r}(z)} (1+|y|)^{2(2-d)} dy = C \exp(-c\varepsilon r)(1+r)^{4-d}
\end{aligned}$$

and

$$\begin{aligned}
& \int_{B_{1/\varepsilon}(z) \setminus B_{2r}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \leq C \int_{B_{1/\varepsilon} \setminus B_{2r}} \exp(-c\varepsilon r)(1+|y|)^{4-2d} dy = C \exp(-c\varepsilon r)(1+r)^{4-d}.
\end{aligned}$$

Finally, since  $d > 4$ ,

$$\begin{aligned}
& \int_{\mathbb{R}^d \setminus B_{1/\varepsilon}(z)} \xi_\varepsilon(z-y)\xi_\varepsilon(x-y) dy \\
& \leq C \exp(-c\varepsilon r) \int_{\mathbb{R}^d \setminus B_{1/\varepsilon}} \exp(-c\varepsilon|y|)(1+|y|)^{4-2d} dy \\
& = C \exp(-c\varepsilon r) \varepsilon^{d-4} \int_{\mathbb{R}^d \setminus B_1} \exp(-2a|y|)(\varepsilon+|y|)^{4-2d} dy \\
& = C \exp(-c\varepsilon r) \varepsilon^{d-4}.
\end{aligned}$$

Combining the above inequalities yields (7.9).

To obtain (7.8), we square (7.9) and sum it over  $z \in \mathbb{Z}^d$  to find that

$$\begin{aligned} & \sum_{z \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} \xi_\varepsilon(z-y) \xi_\varepsilon(x-y) \right)^2 \\ & \leq \int_{\mathbb{R}^d} C \exp(-c\varepsilon|x-z|) \left( (1+|x-z|)^{8-2d} + \varepsilon^{2d-8} \right) dz = C + C\varepsilon^{d-8}. \end{aligned}$$

*Step 3.* The estimate of the first term on the left side of (7.7). Notice that according to Proposition 4.1, we have that

$$(7.10) \quad |K(z)| \leq (T_z \mathcal{X}(\theta'_z(\omega, \omega')))^{d-1-\delta} \left( \varepsilon^2 + [\psi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} \right. \\ \left. + \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (h_\varepsilon(y') - h_\varepsilon(y) - \varepsilon^4|y' - z|^2) \right).$$

We control each part individually. First, we claim that for every  $\gamma \in (0, 1)$ , for every  $p \in (1, \infty)$ , there exists  $C(\gamma, \lambda, \Lambda, d, \ell, p)$  such that

$$(7.11) \quad \mathbb{E} \left[ \left( [\psi_\varepsilon]_{C_1^{1,1}(z, B_{1/\varepsilon}(z))} \right)^p \right]^{\frac{1}{p}} \leq C\varepsilon^{2-\gamma}.$$

Observe that  $\psi_\varepsilon$  is a solution of

$$(7.12) \quad -\operatorname{tr}(A(x)D^2\psi_\varepsilon) = -\varepsilon^2\phi_\varepsilon + 4\varepsilon^2\phi_{2\varepsilon} \quad \text{in } \mathbb{R}^d.$$

Denote the right side by  $f_\varepsilon := -\varepsilon^2\phi_\varepsilon + 4\varepsilon^2\phi_{2\varepsilon} = -\varepsilon^2\psi_\varepsilon + 3\varepsilon^2\phi_{2\varepsilon}$ .

We show that, for every  $\gamma > 0$  and  $p \in [1, \infty)$ , there exists  $C(\gamma, p, d, \lambda, \Lambda, \ell) < \infty$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$ ,

$$(7.13) \quad \mathbb{E} \left[ \left( \|f_\varepsilon\|_{L^\infty(B_{1/\varepsilon})} + \varepsilon^{-\beta} [f_\varepsilon]_{C^{0,\beta}(B_{1/\varepsilon})} \right)^p \right]^{\frac{1}{p}} \leq C\varepsilon^{2-\gamma}.$$

We first observe that (6.1), (6.3), and (6.5) imply that

$$\mathbb{E} \left[ \exp \left( \left( \frac{1}{\mathcal{E}(\varepsilon)} \sup_{B_{\sqrt{d}}} |\varepsilon^2\phi_\varepsilon - \operatorname{tr}(\overline{AM})| \right)^{\frac{1}{2}+\delta} \right) \right] \leq C.$$

A union bound and stationarity then give, for every  $\gamma > 0$  and  $p \in [1, \infty)$ ,

$$(7.14) \quad \mathbb{E} \left[ \left\| \varepsilon^2\phi_\varepsilon - \operatorname{tr}(\overline{AM}) \right\|_{L^\infty(B_{4/\varepsilon})}^p \right]^{\frac{1}{p}} \leq C\varepsilon^{-\gamma}\mathcal{E}(\varepsilon).$$

where  $C = C(\gamma, p, d, \lambda, \Lambda, \ell) < \infty$ . The Krylov-Safonov estimate yields

$$\mathbb{E} \left[ \left( \varepsilon^{-\beta} [\phi_\varepsilon]_{C^{0,\beta}(B_{2/\varepsilon})} \right)^p \right]^{\frac{1}{p}} \leq \varepsilon^{-2}\mathcal{E}(\varepsilon).$$

The previous two displays and the triangle inequality yield the claim (7.13). Now (7.11) follows from Theorem 3.1, (7.13), (7.14) and the Hölder inequality.

We next show that for every  $\gamma \in (0, 1)$  and for every  $p \in (1, \infty)$ ,

$$(7.15) \quad \mathbb{E} \left[ \left( \sup_{y \in B_\ell(z), y' \in \mathbb{R}^d} (h_\varepsilon(y') - h_\varepsilon(y) - \varepsilon^4|y' - z|^2) \right)^p \right]^{\frac{1}{p}} \leq C\varepsilon^{2-\gamma}.$$

By a union bound, we find that, for every  $t > 0$ ,

$$\begin{aligned} & \mathbb{P} \left[ \sup_{y' \in \mathbb{R}^d} (h_\varepsilon(y') - \varepsilon^4 |y' - z|^2) > t \right] \\ & \leq \sum_{n=0}^{\infty} \mathbb{P} \left[ \sup_{y' \in B_{2^n/\varepsilon}(z)} h_\varepsilon(y') > c2^{2n}\varepsilon^2 + t \right] \\ & \leq \sum_{n=0}^{n(t)} \mathbb{P} \left[ \sup_{y' \in B_{2^n/\varepsilon}(z)} h_\varepsilon(y') > t \right] + \sum_{n=n(t)+1}^{\infty} \mathbb{P} \left[ \sup_{y' \in B_{2^n/\varepsilon}(z)} h_\varepsilon(y') > c2^{2n}\varepsilon^2 \right] \end{aligned}$$

where  $n(t)$  is the largest positive integer satisfying  $2^{2n(t)} \leq t\varepsilon^{-2}$ . By (7.14),

$$\begin{aligned} \sum_{n=0}^{n(t)} \mathbb{P} \left[ \sup_{y' \in B_{2^n/\varepsilon}(z)} h_\varepsilon(y') > t \right] & \leq (n(t) + 1) \mathbb{P} \left[ \sup_{y' \in B_{2^{n(t)}/\varepsilon}(z)} h_\varepsilon(y') > t \right] \\ & \leq C(n(t) + 1) 2^{dn(t)} \mathbb{P} \left[ \sup_{y' \in B_{1/\varepsilon}} h_\varepsilon(y') > t \right] \\ & \leq C(n(t) + 1) 2^{dn(t)} (\varepsilon^{\gamma-2} t)^{-p} \\ & \leq C(\log t \varepsilon^{-2}) (t^{-1} \varepsilon^{(2-\gamma)})^{p-\frac{d}{2}} \varepsilon^{\frac{\gamma d}{2}} \end{aligned}$$

and

$$\begin{aligned} \sum_{n=n(t)+1}^{\infty} \mathbb{P} \left[ \sup_{y' \in B_{2^n/\varepsilon}(z)} h_\varepsilon(y') > c2^{2n}\varepsilon^2 \right] & \leq \sum_{n=n(t)+1}^{\infty} C2^{dn} \mathbb{P} \left[ \sup_{y' \in B_{1/\varepsilon}} h_\varepsilon(y') > c2^{2n}\varepsilon^2 \right] \\ & \leq C\varepsilon^{(2-\gamma)p} \sum_{n=n(t)+1}^{\infty} 2^{dn} 2^{-2np} \varepsilon^{-2p} \\ & \leq C(\varepsilon^{(2-\gamma)} t^{-1})^{p-\frac{d}{2}}. \end{aligned}$$

Combining the above, taking  $p$  sufficiently large, integrating over  $t$ , and shrinking  $\gamma$  and redefining  $p$  yields (7.15).

A combination of (6.7), (7.10), (7.11), (7.15) and the Hölder inequality yields the desired bound for the first term on the left side of (7.7).

*Step 4.* The estimate of the second term on the left side of (7.7). According to (5.19) and Proposition 4.1, for  $\mathcal{X}$  and  $\delta(d, \lambda, \Lambda) > 0$  as in Proposition 4.1,

$$\begin{aligned} G_\varepsilon(x, y) \sup_{B_{\ell/2}(y)} (h_\varepsilon - h'_\varepsilon) & \leq C\varepsilon^2 (T_y \mathcal{X})^{d-1-\delta} \xi_\varepsilon(y-x) (T_z \mathcal{X})^{d+1-\delta} \xi_{2\varepsilon}(y-z) \\ & \leq C\varepsilon^2 (T_y \mathcal{X})^{d-1-\delta} (T_z \mathcal{X})^{d+1-\delta} \xi_\varepsilon(y-x) \xi_\varepsilon(y-z). \end{aligned}$$

Therefore,

$$H(y, z) \leq C\varepsilon^2 (T_y \mathcal{X})^{d-1-\delta} (T_z \mathcal{X})^{d+1-\delta}.$$

Thus Hölder's inequality yields that, for every  $p \in (1, \infty)$ ,

$$\widetilde{\mathbb{E}}[H(y, z)^p] \leq C\varepsilon^{2p}.$$

This completes the proof of (7.7).  $\square$

We next control the expectation of  $\widehat{\phi}_\varepsilon - \widehat{\phi}_{2\varepsilon} = \psi_\varepsilon(x) - \frac{3}{4\varepsilon^2} \text{tr}(\overline{AM})$ .

**Lemma 7.3.** *For every  $p \in [1, \infty)$  and  $\gamma > 0$ , there exists  $C(p, \gamma, d, \lambda, \Lambda, \ell) < \infty$  such that, for every  $\varepsilon \in (0, \frac{1}{2}]$  and  $x \in \mathbb{R}^d$ ,*

$$(7.16) \quad \left| \mathbb{E}[\psi_\varepsilon(x)] - \frac{3}{4\varepsilon^2} \operatorname{tr}(\overline{AM}) \right| \leq C\varepsilon^{(\frac{d-4}{2}\wedge 2) - \gamma}.$$

*Proof.* The main step in the argument is to show that

$$(7.17) \quad |\mathbb{E}[\psi_\varepsilon(0)] - 4\mathbb{E}[\psi_{2\varepsilon}(0)]| \leq C\varepsilon^{\frac{d-4}{2}\wedge 2}.$$

Let us assume (7.17) for the moment and see how to obtain (7.16) from it. First, it follows from (7.17) that, for every  $\varepsilon$  and  $m, n \in \mathbb{N}$  with  $m \leq n$ ,

$$|(2^{-m}\varepsilon)^2 \mathbb{E}[\psi_{2^{-m}\varepsilon}(0)] - (2^{-n}\varepsilon)^2 \mathbb{E}[\psi_{2^{-n}\varepsilon}(0)]| \leq C(2^{-m}\varepsilon)^{\frac{d}{2}\wedge 4}.$$

Thus, the sequence  $\{(2^{-n}\varepsilon)^2 \mathbb{E}[\psi_{2^{-n}\varepsilon}(0)]\}_{n \in \mathbb{N}}$  is Cauchy and there exists  $L \in \mathbb{R}$  with

$$|(2^{-m}\varepsilon)^2 \mathbb{E}[\psi_{2^{-m}\varepsilon}(0)] - L| \leq C(2^{-m}\varepsilon)^{\frac{d}{2}\wedge 4}.$$

Taking  $m = 0$  and dividing by  $\varepsilon^2$ , this yields

$$\left| \mathbb{E}[\psi_\varepsilon(0)] - \frac{L}{\varepsilon^2} \right| \leq C\varepsilon^{\frac{d-4}{2}\wedge 2}.$$

But in view of (6.3), we have that  $L = \frac{3}{4} \operatorname{tr}(\overline{AM})$ . This completes the proof of the lemma, subject to the verification of (7.17).

We denote  $h_\varepsilon(x) := \psi_\varepsilon(x) - 4\psi_{2\varepsilon}(x)$  so that we may rewrite (7.17) as

$$(7.18) \quad \mathbb{E}[h_\varepsilon(0)] \leq C\varepsilon^{\frac{d-4}{2}\wedge 2}.$$

We next introduce the function

$$(7.19) \quad \eta_\varepsilon(x) := \psi_\varepsilon - \frac{1}{4}\psi_{2\varepsilon}$$

and observe that  $\eta_\varepsilon$  is a solution of

$$(7.20) \quad -\operatorname{tr}(A(x)D^2\eta_\varepsilon) = -\varepsilon^2 h_\varepsilon \quad \text{in } \mathbb{R}^d.$$

In the first step, we show that  $\psi_\varepsilon$  has small oscillations in balls of radius  $\varepsilon^{-1}$ , and therefore so do  $h_\varepsilon$  and  $\eta_\varepsilon$ . This will allow us to show in the second step that (7.20) is in violation of the maximum principle unless the mean of  $h_\varepsilon$  is close to zero.

*Step 1.* The oscillation bound for  $\psi_\varepsilon$ . The claim is that, for every  $\gamma \in (0, 1)$  and  $p \in [1, \infty)$ , there exists  $C(p, \gamma, \lambda, \Lambda, d, \ell)$  such that

$$(7.21) \quad \mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} |\psi_\varepsilon(x) - \mathbb{E}[\psi_\varepsilon(0)]|^p \right]^{\frac{1}{p}} \leq C\varepsilon^{\frac{d-4}{2}\wedge 2}.$$

By the equation (7.12) for  $\psi_\varepsilon$ , the Krylov-Safonov estimate (3.2) and the bounds (7.13), we have that (taking  $\gamma$  sufficiently small),

$$(7.22) \quad \mathbb{E} \left[ \left( \sup_{x \in B_{1/2\varepsilon}} |\psi_\varepsilon(x) - \psi_\varepsilon([x])| \right)^p \right]^{\frac{1}{p}} \leq C\varepsilon^\sigma \varepsilon^{2-\gamma} \leq C\varepsilon^2.$$

Here  $[x]$  denotes the nearest point of  $\mathbb{Z}^d$  to  $x \in \mathbb{R}^d$ . By the fluctuation estimate (7.4), stationarity and a union bound, we have, for every  $\gamma > 0$  and  $p \in (1, \infty)$ ,

$$\mathbb{E} \left[ \sup_{z \in B_{1/\varepsilon} \cap \mathbb{Z}^d} |\psi_\varepsilon(z) - \mathbb{E}[\psi_\varepsilon(0)]|^p \right]^{\frac{1}{p}} \leq C \varepsilon^{\frac{d-4}{2} \wedge 2 - \gamma}.$$

The previous two lines complete the proof of (7.21).

*Step 2.* We prove something stronger than (7.18) by showing that

$$\mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} |h_\varepsilon(x)| \right] \leq C \varepsilon^{\frac{d-4}{2} \wedge 2}.$$

By (7.21), it suffices to show that

$$\mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} h_\varepsilon(x) \right] \geq -C \varepsilon^{\frac{d-4}{2} \wedge 2} \quad \text{and} \quad \mathbb{E} \left[ \inf_{x \in B_{1/\varepsilon}} h_\varepsilon(x) \right] \leq C \varepsilon^{\frac{d-4}{2} \wedge 2}.$$

We will give only the argument for the second inequality in the display above since the proof of the first one is similar. Define the random variable

$$\kappa := \varepsilon^2 \sup_{x \in B_{1/\varepsilon}} |\psi_\varepsilon(x) - \mathbb{E}[\psi_\varepsilon(0)]|.$$

Observe that the function

$$x \mapsto \psi_\varepsilon(x) - 8\kappa|x|^2 \quad \text{has a local maximum at some point } x_0 \in B_{1/2\varepsilon}.$$

The equation (7.12) for  $\psi_\varepsilon$  implies that

$$-\varepsilon^2 h_\varepsilon(x_0) \geq -16\Lambda d \kappa \geq -C \kappa.$$

Thus

$$\inf_{x \in B_{1/2\varepsilon}} h_\varepsilon(x) \leq h_\varepsilon(x_0) \leq C \kappa = C \sup_{x \in B_{1/\varepsilon}} |\psi_\varepsilon(x) - \mathbb{E}[\psi_\varepsilon(0)]|.$$

Taking expectations and applying (7.21) yields the claim.  $\square$

We now complete the proof of Theorem 7.1.

*Proof.* Proof of Theorem 7.1 According to (7.4), (7.21), (7.16) and a union bound, we have

$$(7.23) \quad \mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} \left| \psi_\varepsilon(x) - \frac{3}{4\varepsilon^2} \text{tr}(\overline{A}M) \right| \right] \leq C \varepsilon^{(\frac{d-4}{2} \wedge 2) - \gamma}.$$

From this we deduce that

$$(7.24) \quad \mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} |\widehat{\phi}_\varepsilon(x) - \widehat{\phi}_{2\varepsilon}(x)| \right] \leq C \varepsilon^{(\frac{d-4}{2} \wedge 2) - \gamma}.$$

We deduce the existence of a stationary function  $\phi$  satisfying

$$(7.25) \quad \mathbb{E} \left[ \sup_{x \in B_{1/\varepsilon}} |\widehat{\phi}_\varepsilon(x) - \phi(x)| \right] \leq C \varepsilon^{(\frac{d-4}{2} \wedge 2) - \gamma}.$$

Passing to the limit  $\varepsilon \rightarrow 0$  in (7.3) and using the stability of solutions under uniform convergence, we obtain that  $\phi$  is a solution of (7.1). The estimates (7.2) are immediate from (1.17). This completes the proof of the theorem.  $\square$

APPENDIX A. PROOF OF THE STRETCHED EXPONENTIAL SPECTRAL GAP  
INEQUALITY

We give the proof of Proposition 2.2, the *Efron-Stein-type* inequality for stretched exponential moments. We first recall the classical Efron-Stein (often called the “spectral gap”) inequality. Given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a sequence  $\mathcal{F}_k \subseteq \mathcal{F}$  of independent  $\sigma$ -algebras, let  $X$  denote a random variable which is measurable with respect to  $\mathcal{F} := \sigma(\mathcal{F}_1, \dots, \mathcal{F}_n)$ . The classical Efron-Stein inequality states that

$$(A.1) \quad \text{var}[X] = \mathbb{E}[|X - \mathbb{E}[X]|^2] \leq \mathbb{E} \left[ \sum_{i=1}^n (X - X'_i)^2 \right]$$

where

$$\begin{aligned} \mathcal{F}'_i &:= \sigma(\mathcal{F}_1, \dots, \mathcal{F}_{i-1}, \mathcal{F}_{i+1}, \dots, \mathcal{F}_n) \\ X'_i &:= \mathbb{E}[X \mid \mathcal{F}'_i]. \end{aligned}$$

Therefore, we see that the variance is controlled by the  $\ell^2$ -norm of the vertical derivative

$$\mathbb{V}[X] := \sum_{i=1}^n (X - X'_i)^2.$$

If we have control of higher moments of  $\mathbb{V}[X]$ , then we can obtain estimates on the moments of  $|X - \mathbb{E}[X]|$ . Indeed, a result of Boucheron, Lugosi, and Massart which can be found in [8, 9] states that for every  $p \geq 2$ ,

$$(A.2) \quad \mathbb{E}[|X - \mathbb{E}[X]|^p] \leq Cp^{p/2} \mathbb{E}[\mathbb{V}[X]^{p/2}],$$

where we may take  $C = 1.271$ . The same authors were also able to obtain similar estimates on exponential moments of  $X - \mathbb{E}[X]$ . Their result [8, 9] states that

$$(A.3) \quad \mathbb{E}[\exp(|X - \mathbb{E}[X]|)] \leq C \mathbb{E}[\exp(C\mathbb{V}[X])].$$

We now give the proof of Proposition 2.2, which is obtained by writing a power series formula for the stretched exponential and then using (A.2) to estimate each term. We thank J. C. Mourrat for pointing out this simple argument and allowing us to include it here.

*Proof of Proposition 2.2.* We show that for every  $\beta \in (0, 2)$ ,

$$\mathbb{E}[\exp(|X - \mathbb{E}[X]|^\beta)] \leq C \mathbb{E} \left[ \exp \left( (C\mathbb{V}[X])^{\frac{\beta}{2-\beta}} \right)^{\frac{2-\beta}{\beta}} \right].$$

We may assume without loss of generality that  $\mathbb{E}[X] = 0$ . Fix  $\beta \in (0, 2)$ . We estimate the power series

$$\mathbb{E}[\exp(|X|^\beta)] = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbb{E}[|X|^{\beta n}]$$

by splitting up the sum into two pieces and estimating each of them separately as follows. First, we consider the terms in which the power of  $|X|$  is less than 2,

and apply (A.1) to get

$$\begin{aligned} \sum_{n=0}^{\lfloor 2/\beta \rfloor} \frac{1}{n!} \mathbb{E}[|X|^{\beta n}] &\leq \sum_{n=0}^{\lfloor 2/\beta \rfloor} \frac{1}{n!} \mathbb{E}[|X|^2]^{n\beta/2} \leq \sum_{n=0}^{\lfloor 2/\beta \rfloor} \frac{1}{n!} \mathbb{E}[\mathbb{V}[X]]^{n\beta/2} \\ &\leq \exp(\mathbb{E}[\mathbb{V}[X]]^{\beta/2}). \end{aligned}$$

For the other terms, we apply (A.2) (with  $\kappa$  in place of  $C$ ) and the discrete Holder inequality to obtain, for  $\alpha > 0$  to be selected below,

$$\begin{aligned} \sum_{n=\lfloor 2/\beta \rfloor+1}^{\infty} \frac{1}{n!} \mathbb{E}[|X|^{\beta n}] &\leq \sum_{n=1}^{\infty} \frac{1}{n!} (\kappa n \beta)^{n\beta/2} \mathbb{E}[(\mathbb{V}[X])^{n\beta/2}] \\ &\leq \left( \sum_{n=1}^{\infty} \frac{1}{n!} \left( \frac{\kappa n \beta}{\alpha} \right)^n \right)^{\beta/2} \left( \sum_{n=1}^{\infty} \frac{1}{n!} \mathbb{E}[(\alpha \mathbb{V}[X])^{n\beta/2}]^{\frac{2}{2-\beta}} \right)^{\frac{2-\beta}{2}}. \end{aligned}$$

We estimate the first factor on the right hand side by using the classical inequality (related to Stirling's approximation) which states that, for every  $n \in \mathbb{N}$ ,

$$n! \geq (2\pi)^{1/2} n^{n+1/2} \exp(-n).$$

This yields that for every  $\alpha > e\kappa\beta$ ,

$$\sum_{n=1}^{\infty} \frac{1}{n!} \left( \frac{\kappa n \beta}{\alpha} \right)^n \leq \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{\infty} n^{-1/2} \left( \frac{e\kappa\beta}{\alpha} \right)^n \leq \frac{\alpha}{\alpha - e\kappa\beta}.$$

Combining this with our previous estimate, we obtain

$$\sum_{n=\lfloor 2/\beta \rfloor+1}^{\infty} \frac{1}{n!} \mathbb{E}[|X|^{\beta n}] \leq \left( \frac{\alpha}{\alpha - e\kappa\beta} \right)^{\beta/2} \left( \sum_{n=1}^{\infty} \frac{1}{n!} \mathbb{E}[(\alpha \mathbb{V}[X])^{n\beta/2}]^{\frac{2}{2-\beta}} \right)^{\frac{2-\beta}{2}}.$$

Observe that

$$\sum_{n=1}^{\infty} \frac{1}{n!} \mathbb{E}[(\alpha \mathbb{V}[X])^{n\beta/2}]^{\frac{2}{2-\beta}} \leq \sum_{n=1}^{\infty} \frac{1}{n!} \mathbb{E}[(\alpha \mathbb{V}[X])^{\frac{n\beta}{2-\beta}}] = \mathbb{E}[\exp((\alpha \mathbb{V}[X])^{\frac{\beta}{2-\beta}})],$$

and this implies that

$$\sum_{n=\lfloor 2/\beta \rfloor+1}^{\infty} \frac{1}{n!} \mathbb{E}[|X|^{\beta n}] \leq \left( \frac{\alpha}{\alpha - e\kappa\beta} \right)^{\beta/2} \left( \mathbb{E}[\exp((\alpha \mathbb{V}[X])^{\frac{\beta}{2-\beta}})] \right)^{\frac{2-\beta}{2}}.$$

Combining all of the previous estimates yields that

$$\mathbb{E}[\exp(|X|^\beta)] \leq \exp(\mathbb{E}[\mathbb{V}[X]]^{\beta/2}) + C \mathbb{E}[\exp((C\mathbb{V}[X])^{\frac{\beta}{2-\beta}})]^{\frac{2-\beta}{2}}.$$

Since  $\beta \leq 2$  and we can take  $\alpha = 20$ , the constant  $C$  is universal. This completes the proof.  $\square$

**Acknowledgements.** The second author was partially supported by NSF Grant DMS-1147523.



## REFERENCES

- [1] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, 347(2):315–361, 2016.
- [2] S. N. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, to appear, arXiv:1602.00512.
- [3] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.
- [4] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of elliptic equations in nondivergence form. *Arch. Ration. Mech. Anal.*, 214(3):867–911, 2014.
- [5] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(2):423–481, 2016.
- [6] M. Avellaneda and F. H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.
- [7] M. Avellaneda and F.H. Lin. Compactness methods in the theory of homogenization. II. Equations in nondivergence form. *Comm. Pure Appl. Math.*, 42(2):139–172, 1989.
- [8] S. Boucheron, G. Lugosi, and P. Massart. Concentration inequalities using the entropy method. *Ann. Probab.*, 31(3):1583–1614, 2003.
- [9] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.
- [10] L. A. Caffarelli and X. Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.
- [11] L. A. Caffarelli and P. E. Souganidis. Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media. *Invent. Math.*, 180(2):301–360, 2010.
- [12] L. A. Caffarelli, P. E. Souganidis, and L. Wang. Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media. *Comm. Pure Appl. Math.*, 58(3):319–361, 2005.
- [13] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [14] B. Efron and C. Stein. The jackknife estimate of variance. *Ann. Statist.*, 9(3):586–596, 1981.
- [15] L. C. Evans. Periodic homogenisation of certain fully nonlinear partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(3-4):245–265, 1992.
- [16] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.
- [17] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators, preprint, arXiv:1409.2678.
- [18] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.
- [19] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.
- [20] A. Gloria and F. Otto. The corrector in stochastic homogenization: Near-optimal rates with optimal stochastic integrability, preprint, arXiv:1510.08290.
- [21] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):294–327, 2011.
- [22] A Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, Unpublished preprint.
- [23] G. C. Papanicolaou and S. R. S. Varadhan. Diffusions with random coefficients. In *Statistics and probability: essays in honor of C. R. Rao*, pages 547–552. North-Holland, Amsterdam, 1982.

- [24] V. V. Yurinskiĭ. Averaging of second-order nondivergent equations with random coefficients. *Sibirsk. Mat. Zh.*, 23(2):176–188, 217, 1982.
- [25] V. V. Yurinskiĭ. On the error of averaging of multidimensional diffusions. *Teor. Veroyatnost. i Primenen.*, 33(1):14–24, 1988. Eng. transl. in *Theory Probab. Appl.* 33 (1988), no. 1, 11–21.

(S. Armstrong) UNIVERSITÉ PARIS-DAUPHINE, PSL RESEARCH UNIVERSITY, CNRS, UMR [7534], CEREMADE, PARIS, FRANCE

*Current address:* Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York 10012

*E-mail address:* `scotta@cims.nyu.edu`

(J. Lin) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON

*E-mail address:* `jessica@math.wisc.edu`