Journées

ÉQUATIONS AUX DÉRIVÉES PARTIELLES

Roscoff, 2–6 juin 2014

Scott Armstrong

Uniform Lipschitz estimates in stochastic homogenization

J. É. D. P. (2014), Exposé n° I, 11 p.

<http://jedp.cedram.org/item?id=JEDP_2014_____A1_0>

Journées *Équations aux dérivées partielles*
Roscoff, 2–6 juin 2014
GDR 2434 (CNRS)

Uniform Lipschitz estimates in stochastic homogenization

Scott Armstrong

Abstract

We review some recent results in quantitative stochastic homogenization for divergence-form, quasilinear elliptic equations. In particular, we are interested in obtaining L^∞-type bounds on the gradient of solutions and thus giving a demonstration of the principle that solutions of equations with random coefficients have much better regularity (with overwhelming probability) than a general equation with non-constant coefficients.

1. Introduction

In this note, we review some recent progress in the quantitative theory of stochastic homogenization for uniformly elliptic equations in divergence form.

At the most general level, we consider equations of the form

$$-\nabla\cdot\left(\mathbf{a}\left(\nabla u^\varepsilon(x),\frac{x}{\varepsilon}\right)\right)=0 \quad \text{in } U\subseteq\mathbb{R}^d, \tag{1.1}$$

where $\varepsilon>0$ is a small parameter and $p\mapsto\mathbf{a}(p,x)$ is a uniformly monotone, Lipschitz vector field (and say measurable in x). We further assume that $\mathbf{a}$ is a random object– that is, it is sampled by a stationary-ergodic probability measure $\mathbb{P}$ on the space of all such coefficient fields. Then the essential result of qualitative homogenization is roughly that, subject to a given Dirichlet boundary condition on ∂U, we have

$$\mathbb{P}\left[\limsup_{\varepsilon\to0}\|u^\varepsilon-u_{\mathrm{hom}}\|_{L^2(U)}=0\right]=1, \tag{1.2}$$

where u_{hom} is the solution of a deterministic equation

$$-\nabla\cdot\overline{\mathbf{a}}(\nabla u_{\mathrm{hom}})=0 \quad \text{in } U,$$

where the existence of the non-random coefficients $\overline{\mathbf{a}}$ is part of the assertion of the theorem. This was proved for linear equations by Papanicolaou and Varadhan [14], Kozlov [12] and Yurinskii [15], all independently. The first statement for nonlinear

MSC 2000: 35B27, 60H25, 35J20, 35J62.
Keywords: Stochastic homogenization, Lipschitz regularity, error estimate.

equations was obtained by Dal Maso and Modica [6, 7], who considered minimizers of convex integral functionals which corresponds to (1.1) in the case that $\mathbf{a}(\cdot, x)$ is the gradient of a convex function.

If the qualitative theory of stochastic homogenization is concerned with proving (1.2), the goal of the quantitative theory is then to quantify it, i.e., to find the rate at which $\|u^\varepsilon - u_{\mathrm{hom}}\|_{L^2(U)}$ converges to zero and the rate at which the probability in (1.2) converges to one. The first such statement was obtained by Yurinskii [16], who used probability methods to show that, for linear equations, we have the following statement: for every $p \in [1, \infty)$, there exists $\alpha > 0$, depending on p in addition to the dimension d and the ellipticity, which we hereafter denote by $\Lambda \geq 1$, such that

$$\mathbb{P}\left[\|u^\varepsilon - u_{\mathrm{hom}}\|_{L^2(U)} \geq \varepsilon^\alpha\right] \leq C\varepsilon^p. \tag{1.3}$$

Here the prefactor C depends also on (d, Λ, p). Such a quantitative result is not possible to prove under qualitative assumptions, so Yurinskii naturally had to assume that the coefficients satisfies a mixing condition, that is, he imposed a quantitative ergodicity hypothesis on his probability measure.

It is of great practical and theoretical importance to determine the optimal exponent α in (1.3), especially under the most natural ergodicity hypothesis on the coefficients, which is independence (or to be precise, a finite range of dependence). Progress on this question proved elusive and Yurinskii's result was unsurpassed for more than twenty years until Gloria and Otto [9, 10] obtained, still in the case of linear equations, the optimal scaling for the error. Their first papers were written for discrete equations on $\mathbb{Z}^d$, but their basic methodology extends to the continuum setting without more than routine technical difficulties (as has been shown recently by those authors in [11]). They proved (1.3) with $\alpha = 1$ (up to a logarithmic correction in $d = 2$, which they show is optimal). A corollary of their analysis is the existence of stationary correctors, answering a long-standing theoretical question.

The methodology of Gloria and Otto is based in part on an insights in an unpublished paper of Naddaf and Spencer [13], which is that one should measure the magnitude of random fluctuations in the solutions u^ε by applying appropriate concentration inequalities. The right ones for the task are "spectral gap" inequalities (also called Poincaré inequalities, or Efron-Stein inequalities). A basic version of this family of inequalities states that, if $Z = F(X_1, \ldots, X_n)$ is a real-valued function of independent random variables $X_1, \ldots, X_n$, then

$$\mathrm{var}\left[Z\right] \leq \mathbb{E}\left[\sum_{i=1}^n \left(Z - \mathbb{E}\left[Z \,|\, \{X_j\}_{j \neq i}\right]\right)^2\right]$$

Notice that the random variable $Z - \mathbb{E}\left[Z \,|\, \{X_j\}_{j \neq i}\right]$ measures precisely how sensitive Z is on the value of X_i (informally, it is the "derivative" of Z with respect to X_i) and thus this is a kind of Poincaré inequality. In the context of stochastic homogenization, we think of the *solutions* u^ε (playing the role of Z) as functions of the coefficients $\mathbf{a}$ (playing the role of the $X_1, \ldots, X_n$, here each X_i may represent the value of the coefficients in a unit-sized cube with vertices in $\mathbb{Z}^d$, say).

The key step in the analysis is then to obtain good bounds on how sensitively the solutions depend on the coefficients. But this turns out to reduce to estimating the expected size of the gradient of the solution. Indeed, if we resample the coefficients

in a unit cube, then from the equation we see that the solution should change in proportion to how large its gradient was in that cube.

Unfortunately, it is well-known that there is no general L^∞ estimate for the gradient of solutions with highly oscillating coefficients. The best Hölder regularity in general is known to be $C^{0,\varepsilon}$ (De Giorgi-Nash-Moser estimate) and the best Sobolev regularity is $W^{1,2+\varepsilon}$ (Meyers estimate), while what we need is Lipschitz, i.e., $C^{0,1} = W^{1,\infty}$. Therefore the crucial step can be seen from the point of view of elliptic regularity theory, with a stochastic twist: how can we show that solutions of equations with random coefficients are more regular (at least with high probability) than solutions of general elliptic equations? In the work of Gloria and Otto, this question is tackled in an indirect way: they analyze the gradient of the approximate correctors and that of the Green's function, using the random properties of these functions.

In this note, I will review some recent results of the author in collaborations with Smart [3] and Mourrat [1] in which we answer this question completely. Roughly speaking, what we proved (for variational equations in [3] and recently for general quasilinear equations of the form (1.1) in [1]) is that an arbitrary solution with $U = B_1$ (i.e., we take no boundary condition) of the equation (1.1), under independence assumptions on the coefficients, satisfies, for every $\varepsilon \in (0, 1]$,

$$|\nabla u^\varepsilon(0)| \le \mathcal{X}\left(1 + \|u^\varepsilon\|_{B_1}\right),$$

where $\mathcal{X}$ is a random variable (it is necessarily not bounded above by a constant almost surely) which satisfies the estimate

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] < \infty \quad \text{for every } s < d.$$

The previous estimate is optimal in the sense that it is false for $s = d$, because the probability of sampling any particular set of coefficients in a ball of radius R is roughly $\exp(-cR^d)$. In other words, there is some chance we will get unlucky and select the coefficients which witness the counterexample to higher elliptic regularity and our estimate must factor in this possibility.

Since the results in [3, 1] apply to general quasilinear equations, they open the way to an optimal quantitative theory for stochastic homogenization of nonlinear elliptic equations, which is work currently in progress. The methods also yield estimates up to the boundary of smooth domains, which paves the way for a quantitative analysis of boundary layers in stochastic homogenization.

The idea for proving an L^∞ estimate of this kind for the gradient of solutions comes from the celebrated work of Avellaneda and Lin [4, 5] in the case of periodic homogenization. They showed that, even though these equations have highly oscillating coefficients, one can "borrow" the higher regularity of the constant coefficient homogenized equation they are limiting to. The method in [4, 5] relies on a compactness argument which is not available in the stochastic case, and therefore the scheme for obtaining the Lipschitz bound, while sharing the philosophy of the work in [4, 5], is not quite the same. It is based on a sub-optimal quantitative homogenization statement, rather than a qualitative statement. We also note that this modified version of the idea of Avellaneda-Lin has shown to be useful even in the case of periodic and almost periodic coefficients, see the recent work of the author and Shen [2].

We conclude by mentioning the recent work of Gloria, Neukamm and Otto [8] who, partially inspired by [3], obtained a quenched Lipschitz estimate for linear equations and systems. In particular, they give the first such estimate for linear equations and systems with nonsymmetric coefficients (as their paper was written after [3] but before [1]), although their techniques do not seem to generalize in a straightforward way to nonlinear equations and do not so far yield estimates with optimal stochastic integrability.

2. The precise statement of the main result

Ee fix the parameters $K_0 \geq 1$ and $\Lambda \geq 1$. We consider coefficient fields $\mathbf{a} = \mathbf{a}(p,x)$ satisfying

$$\mathbf{a} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \quad \text{is Lebesgue measurable,} \tag{2.1}$$

such that $p \mapsto \mathbf{a}(p,x)$ is uniformly monotone and Lipschitz, uniformly in x, that is, for every $x, p_1, p_2 \in \mathbb{R}^d$,

$$|\mathbf{a}(p_1,x) - \mathbf{a}(p_2,x)| \leq \Lambda\,|p_1 - p_2| \tag{2.2}$$

and

$$(\mathbf{a}(p_1,x) - \mathbf{a}(p_2,x)) \cdot (p_1 - p_2) \geq \frac{1}{\Lambda}\,|p_1 - p_2|^2 \tag{2.3}$$

We define the set of all coefficient fields by

$$\Omega := \{\mathbf{a} : \mathbf{a} \text{ satisfies (2.1), (2.2) and (2.3)}\}. \tag{2.4}$$

We endow Ω with the translation group $\{T_y\}_{y\in\mathbb{R}^d}$, which acts on Ω via

$$(T_z\mathbf{a})(p,x) := \mathbf{a}(p, x+z),$$

and the family $\{\mathcal{F}_U\}$ of σ–algebras, with $\mathcal{F}_U$ defined for each Borel $U \subseteq \mathbb{R}^d$ by

$\mathcal{F}_U :=$ σ–algebra on Ω generated by the family of random variables

$$\mathbf{a} \mapsto \int_U \xi \cdot \mathbf{a}(p,x)\varphi(x)\,dx, \quad p,\xi \in \mathbb{R}^d,\ \varphi \in L^1(\mathbb{R}^d).$$

The largest of these σ–algebras is denoted by $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. The translation group also acts naturally on $\mathcal{F}$ via the definition

$$T_z A := \{T_z\mathbf{a} : \mathbf{a} \in A\}, \quad A \in \mathcal{F}.$$

Throughout the paper, we consider a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ which is assumed to satisfy the following three conditions:

(P1) The random field $\mathbf{a}$ is locally uniformly bounded on the support of $\mathbb{P}$ in the sense that, for all $x \in \mathbb{R}^d$,

$$\mathbb{P}\left[|\mathbf{a}(0,x)| \leq K_0\right] = 1$$

(P2) $\mathbb{P}$ is stationary with respect to $\mathbb{Z}^d$–translations: for every $z \in \mathbb{Z}^d$ and $A \in \mathcal{F}$,

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

(P3) $\mathbb{P}$ satisfies a unit range of dependence, that is, for all Borel $U, V \subseteq \mathbb{R}^d$,

$$\mathcal{F}_U \text{ and } \mathcal{F}_V \text{ are independent provided that } \quad \operatorname{dist}(U,V) \geq 1.$$

We now present the statement of the main result.

Theorem 2.1. *Suppose that $\mathbb{P}$ satisfies (P1), (P2), (P3). Fix $p > d$. Then for every $M \geq 1$ and $s \in (0, d)$, there exist $C(d, \Lambda, M, s) \geq 1$ and a random variable $\mathcal{X} \geq 1$ satisfying*

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] \leq C \tag{2.5}$$

such that the following holds: if $R \geq 1$, $u \in H^1(B_R)$ satisfies both

$$K_0 + \frac{1}{R}\left(\fint_{B_R} |u(x)|^2 \, dx\right)^{\frac{1}{2}} \leq M \tag{2.6}$$

and

$$-\nabla \cdot a\left(\nabla u, x\right) \quad \text{in } B_R, \tag{2.7}$$

then

$$\fint_{B_r(0)} |\nabla u(x)|^2 \, dx \leq CM^2 \quad \text{for every} \quad \mathcal{X} \leq r \leq \frac{1}{2}R. \tag{2.8}$$

Remark 2.2. We call Theorem 2.1 a "Lipschitz" estimate because typically R is very large and in this scaling the microscopic scale is $O(1)$ rather than $O(\varepsilon)$. We therefore think of the left side as representing $|\nabla u(0)|$. Notice that (2.8) also implies

$$\fint_{B_1} |\nabla u(x)|^2 \, dx \leq C\mathcal{X}^d \left(1 + \fint_{B_R} |\nabla u(x)|^2 \, dx\right),$$

so we may also think of $\mathcal{X}$ as a random prefactor constant in an estimate with deterministic scales. Under the additional assumption that the coefficients are smooth on the microscopic scale, the left side of (2.8) with $r = \mathcal{X}$ controls the local Lipschitz constant of u in $B_{r/2}$ (by the Schauder estimates) and, in that case, we then have a true (pointwise) Lipschitz estimate. It is better, however, to think of control over the very small scales as a separate issue, since the behavior of solutions on scales smaller than the microscopic scale is outside of the purview of homogenization.

Remark 2.3. An estimate like (2.8) is the best possible for the control of microscopic-scale fluctuations in terms of large scale fluctuations for solutions of (1.1). Indeed, we expect the gradient of a solution to act like "white noise" on the microscopic scale, so in particular there is no hope to obtain a uniform modulus of continuity.

Remark 2.4. Optimal results under much weaker mixing conditions than (P3) appear in [1].

Remark 2.5. The integrability of $\mathcal{X}$ given in (2.5) is optimal, because the probability of obtaining any *particular* coefficient field (for the random checkerboard, say) in a ball of radius R is of order $\exp(-cR^d)$.

Remark 2.6. The parameter M in Theorem 2.1 may be removed in the case that F is positively homogeneous of order two (e.g., in the case that the corresponding PDE is linear), but it is necessary and not an artifact of our method in the general nonlinear case.

3. The scheme for proving the Lipschitz estimate

The general scheme for proving Lipschitz estimates introduced in [3] is presented in the following proposition. It reduces Theorem 2.1 to Theorem 4.1 (stated in the next section) by a Campanato-type iteration, and most of the effort in the proof of the main theorem is focused on obtaining the latter. The statement here is close to the same as [3, Lemma 5.1], but we have formulated the result using the L^2 norm rather than the L^∞ norm (which makes no difference in the argument but clarifies that the methods work for systems).

We first introduce some notation. We take $\mathcal{L}$ to be the set of affine functions on $\mathbb{R}^d$ and define, for each $\sigma \in (0, \frac{1}{2}]$ and $r > 0$, the set

$$\mathcal{A}(r,\sigma) := \left\{ v \in L^2(B_r) \,:\, \frac{1}{\sigma r} \inf_{l \in \mathcal{L}} \left(\fint_{B_{\sigma r}} |v(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right. \\ \left. \leq \frac{1}{2} \left(\frac{1}{r} \inf_{l \in \mathcal{L}} \left(\fint_{B_r} |v(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right) \right\}.$$

Roughly, the set $\mathcal{A}(r,\sigma)$ consists of those $L^2(B_r)$ functions u which satisfy one step of a $C^{1,\beta}$ Campanato iteration with dilation factor σ: the flatness of u in $B_{\sigma r}$ is improved from its flatness in B_r by a factor of two (note that β can be computed in terms of σ, but this does not matter for our purposes).

Proposition 3.1. *Assume $R \geq 1$, $\alpha, \sigma > 0$, $K \geq 0$, $r_0 \in [1, R/4]$ and $u \in L^2(B_R)$ have the property that, for every $r \in [r_0, R/2]$,*

$$\inf_{v \in \mathcal{A}(r,\sigma)} \frac{1}{r} \left(\fint_{B_r} |u(x) - v(x)|^2 \, dx \right)^{\frac{1}{2}} \leq r^{-\alpha} \left(K + \frac{1}{2r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{2r}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right).$$

Then there exists $C(\alpha,\sigma) \geq 1$ such that, for every $s \in [r_0, R/2]$,

$$\frac{1}{s} \inf_{a \in \mathbb{R}} \left(\fint_{B_s} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \leq C \left(\frac{1}{R} \inf_{a \in \mathbb{R}^d} \left(\fint_{B_R} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} + K \left(\frac{s}{R} \right)^{\alpha} \right). \tag{3.1}$$

Proposition 3.1 asserts that if a function $u \in L^2(B_R)$ (we are thinking of R very large) has the property that, in every ball B_r with radius r between $R/2$ and a "minimal radius" r_0, it can be well-approximated by a function in $\mathcal{A}(r,\sigma)$, then in fact u does not oscillate too much on scales larger than the minimal radius. Its proof is a completely straightforward and elementary argument.

In order to make use of Proposition 3.1, we need to check that local minimizers of the homogenized functional belong to $\mathcal{A}(r,\sigma)$. This is handled by the following simple lemma, which is a reflection of the well-known fact that a family of scale-invariant functions satisfies a $C^{1,\beta}$ estimate if and only if they satisfy an improvement of flatness property.

Lemma 3.2. *Suppose that $\beta \in (0,1]$, $K \geq 0$, and $u \in C^{1,\beta}(B_R)$ has the property that, for every $r \in (0, R/2]$,*

$$[\nabla u]_{C^{0,\beta}(B_r)} \leq K r^{-\beta} \left(\frac{1}{2r} \inf_{l \in \mathcal{L}} \left(\fint_{B_{2r}} |u(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right).$$

Then there exists $\sigma(\beta, K) \in (0, \frac{1}{2}]$ such that $u \in \mathcal{A}(r,\sigma)$ for every $r \in (0, R/2]$.

The previous lemma states that solutions of constant coefficient equations belong to $\mathcal{A}(r,\sigma)$ on all scales. This is what makes Proposition 3.1 useful in our proof of Theorem 2.1.

4. The error estimate for the Dirichlet problem

The main analytic step in the proof of Theorem 2.1, which is an interesting result in its own right, is a "quenched" error estimate for the Dirichlet problem. Thus the Lipschitz estimate, which we explained in the introduction is the key step to obtaining optimal error estimates, is itself a consequence of a *sub*-optimal error estimate.

The estimate, given in Theorem 4.1 below, is roughly speaking the same sort of estimate as Yurinskii's result (1.3) mentioned in the introduction. Both are algebraic yet sub-optimal in the scaling of the error. The one we present improves on Yurinskii's result in several respects: first, and this is also our reason for calling it "quenched," is that it gives an estimate for the error which is independent of the boundary condition (which is necessary in our strategy, based on Proposition 3.1, for proving Theorem 2.1). Secondly, it is also optimal in stochastic integrability (which is where the optimal integrability of $\mathcal{X}$ in Theorem 2.1 comes from) and in particular gives a scaling exponent $\alpha > 0$ which is independent of the parameter p in (1.3). Finally, we mention another important advantage of our approach is that it applies to general nonlinear equations. In fact, the versions of the following theorem proved in [3, 1] are the first quantitative results in stochastic homogenization for nonlinear equations in divergence form.

Theorem 4.1. *Consider a bounded Lipschitz domain $U_0 \subseteq \mathbb{R}^d$, $M \geq 1$ and exponents $\delta > 0$ and $s < d$. There exist $\alpha(d,\Lambda,\delta,s) > 0$, $C(d,\Lambda,M,\delta,U_0,s) \geq 1$ and a random variable $\mathcal{Y}$ satisfying*

$$\mathbb{E}\left[\exp\left(\mathcal{Y}\right)\right] \leq C \tag{4.1}$$

such that the following holds: for every $R \geq 1$, $U := RU_0$ and $f \in W^{1,2+\delta}(U)$ satisfying

$$K_0 + \left(\fint_U |\nabla f(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}} \leq M,$$

the unique solutions $u, u_{\mathrm{hom}} \in f + H_0^1(U)$ of the equations

$$-\nabla \cdot \mathbf{a}\left(\nabla u, x\right) = 0 \quad \text{and} \quad -\nabla \cdot \overline{\mathbf{a}}\left(\nabla u_{\mathrm{hom}}, x\right) = 0 \quad \text{in } U \tag{4.2}$$

satisfy the estimate

$$R^{-2} \fint_U |u(x) - u_{\mathrm{hom}}(x)|^2 \, dx \leq CM^2 \left(1 + \mathcal{Y} R^{-s}\right) R^{-\alpha}. \tag{4.3}$$

The conclusion (4.3) may appear to be put in a strange form, but Chebyshev's inequality and (4.1) immediately puts it into a form like (1.3), but with stronger exponential-type bounds on the right side.

We next show that the combination of Proposition 3.1 and Theorem 4.1 does indeed imply our main result, Theorem 2.1.

Proof of Theorem 2.1. Take $\delta(d,\Lambda) > 0$ to be the exponent δ_0 in the statement of the interior Meyers estimate (see [1] for a statement). We may suppose without loss

of generality that $R = 2^k$ for some $k \in \mathbb{N}$. For each $r \in [1, R/2]$, let $u_{\mathrm{hom},r} \in u + H_0^1(B_r)$ denote the (unique) solution of

$$-\nabla \cdot \mathbf{a}\,(\nabla u_{\mathrm{hom},r}) = 0 \quad \text{in } B_r.$$

We first define the random minimal radius $\mathcal{X} \geq 1$. With $\mathcal{Y}$ as in the second statement of Theorem 4.1, we define

$$\mathcal{X} := \mathcal{Y}^{\frac{1}{s}}.$$

Then the conclusions of the second statement of Theorem 4.1 give the bound

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] \leq C$$

and imply that, for every $r \in \left[\mathcal{X}, \frac{1}{2}R\right]$, we have

$$\frac{1}{r^2} \fint_{B_r} |u(x) - u_{\mathrm{hom},r}(x)|^2 \, dx \leq CM^2 r^{-\alpha}$$

provided that r satisfies

$$K_0 + \frac{1}{4r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{4r}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \leq cC'M, \tag{4.4}$$

where the exponent $\alpha > 0$ in (4) depends on the appropriate quantities.

According to the conclusion of Theorem 4.1, for every $r \in \left[\mathcal{X}, \frac{1}{2}R\right]$,

$$\frac{1}{r} \left(\fint_{B_r} |u(x) - u_{\mathrm{hom},r}(x)|^2 \, dx \right)^{\frac{1}{2}} \leq Cr^{-\beta} \left(K_0 + \left(\fint_{B_{2r}} |u(x)|^2 \, dx \right)^{\frac{1}{2}} \right). \tag{4.5}$$

We claim that, for every $s \in \left[\mathcal{X}, \frac{1}{2}R\right]$,

$$\frac{1}{s} \inf_{a \in \mathbb{R}^d} \left(\fint_{B_s} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \leq C \left(K_0 + \frac{1}{R} \inf_{a \in \mathbb{R}} \left(\fint_{B_R} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right). \tag{4.6}$$

It suffices to prove the estimate for $s \in \left[\mathcal{X}, \frac{1}{2}R\right] \cap \{2^{-k-1}R : k \in \mathbb{N}\}$. We argue by induction: suppose that the estimate holds for every radius $s = 2^{-j}$ with $j = 1, \ldots, k$ such that $2^{-k} > 2\mathcal{X}$, and with $C'' \geq 1$ in place of C. Then we have (4.4) for $r = 2^{-k}$, provided we choose C' to be large enough multiple of C'', and so an application of Proposition 3.1 yields

$$K_0 + \frac{1}{2^{-k-1}R} \inf_{a \in \mathbb{R}^d} \left(\fint_{B_{2^{-k-1}R}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \leq CM \leq C''M,$$

provided $C'' = C$ is chosen large enough. By induction, we thus obtain (4.6) for all $s = 2^{-k}R \geq \mathcal{X}$.

We conclude by observing that, by the Caccioppoli inequality, (4.6) and the Poincaré inequality, give us, for every $r \in \left[\mathcal{X}, \frac{1}{2}R\right]$,

$$\begin{aligned}
\fint_{B_r} |\nabla u(x)|^2 \, dx &\leq C \inf_{a \in \mathbb{R}} \fint_{B_{2r}} |u(x) - a|^2 \, dx \\
&\leq C \left(K_0 + \inf_{a \in \mathbb{R}} \frac{1}{R} \left(\fint_{B_R} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right) \\
&\leq C \left(K_0 + \left(\fint_{B_R} |\nabla u(x)|^2 \, dx \right)^{\frac{1}{2}} \right) \\
&\leq CM.
\end{aligned}$$

This completes the proof of the theorem. □

5. A vague overview of the proof of Theorem 4.1

Most of the analysis in [3, 1] is focused on the proof of Theorem 4.1 and it is too involved to give the detailed argument here. We mention only some of the important ideas and for simplicity we restrict ourselves to the variational setting (i.e., we assume that $\mathbf{a} = \nabla L$ for a uniformly convex Lagrangian L). The detailed arguments can be found in [3].

There are two components in the proof of Theorem 4.1: the first is the most important, and is a probabilistic statement that asserts that the energy of a minimizer in a large cube with prescribed affine boundary data with slope say p converges in expectation to $\overline{L}(p)$, where $\overline{L}$ is the effective Lagrangian, with a rate which is a power of the side length of the cube. This is [3, Theorem 3.1]. The second, presented in [3, Proposition 4.1], is a deterministic statement which asserts the error in homogenization for a general Dirichlet problem is controlled by the spatial average of the differences between the energy in mesoscopic cubes and the homogenized limit, effectively reducing Theorem 4.1 to the probabilistic statement.

The proof of the latter is based on *subadditive* arguments. Let us denote the (normalized) energy of a minimizer in a bounded domain U with given affine boundary data by

$$\nu(U,p) := \inf_{v \in H^1_0(U)} \fint_U L(p + \nabla v(x), x) \, dx.$$

This quantity is naturally monotone in the sense that, if U is the (interior of the closure of the) disjoint union of subdomains $U_1, \ldots, U_n$, then

$$\nu(U,p) \leq \sum_{j=1}^{n} \frac{|U_j|}{|U|} \nu(U_j,p).$$

This is because we can construct a candidate for achieving the infimum in the definition of $\nu(U,p)$ by gluing together the minimizers in each subdomain U_j. Since we assume that the statistics of $\mathbb{P}$ are stationary, this means that, for every $n, m \in \mathbb{N}$ with $n \leq m$, we have

$$\mathbb{E}\left[\mu(Q_m,p)\right] \leq \mathbb{E}\left[\mu(Q_n,p)\right],$$

where we denote dyadic cubes by $Q_n := [0, 2^n]^d$. The key step in the analysis is to quantify the monotone limit

$$\lim_{n\to\infty} \mathbb{E}\left[\mu(Q_n, p)\right] = \inf_{n\in\mathbb{N}} \mathbb{E}\left[\mu(Q_n, p)\right] =: \overline{L}(p).$$

In [3, Theorem 3.1] it is shown that there exists $\alpha > 0$ such that, for every $t \geq 1$ and $n \in \mathbb{N}$,

$$\mathbb{P}\left[\left|\nu(Q_n, p) - \overline{L}(p)\right| \geq C(2^n)^{-\alpha} t\right] \leq C \exp\left(-c(2^n)^s t\right),$$

for any $s < d$. To prove this, we need to introduce the *dual* (in the sense of convex analysis), *superadditive* quantity μ defined by

$$\mu(U, q) := \inf_{u\in H^1(U)} \fint_U \left(L(\nabla u(x), x) - q \cdot \nabla u(x)\right) dx$$

Note that the there is no boundary condition in the definition of μ, which gives it its superadditivity (by restriction of minimizers on larger domains to smaller ones). Together, ν and μ give the problem monotonicity "from both sides", that is, they reveal the additive structure of the problem, rendering the analysis of the randomness more feasible. What is shown in Section 3 of [3] is that, for very large cubes, the minimizers of μ (which we recall do not necessarily have affine traces on the boundary) are nevertheless, with high probability, very flat. This allows them to be compared to solutions with affine boundary data, thereby showing that $\mu(q)$ is close to $\nu(p)$ when p is properly dual to q, and that the limit of μ of $-\overline{L}^*(q)$, where $\overline{L}^*$ is the Legendre transform of $\overline{L}$. This is the justification that μ is "dual in the homogenized limit" to ν, and what allows the quantitative argument to succeed.

References

[1] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. Preprint.

[2] S. N. Armstrong and Z. Shen. Lipschitz estimates in almost-periodic homogenization. Preprint, arXiv:1409.2094.

[3] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. Preprint, arXiv:1406.0996.

[4] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.

[5] M. Avellaneda and F.-H. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8-9):897–910, 1991.

[6] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.

[7] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.

[8] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators, Preprint, arXiv:1409.2678.

[9] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.

[10] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.

[11] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization, Preprint.

[12] S. M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.

[13] A Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, Unpublished preprint.

[14] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[15] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27(4):167–180, 215, 1986.

[16] V. V. Yurinskiĭ. Homogenization error estimates for random elliptic operators. In *Mathematics of random media (Blacksburg, VA, 1989)*, volume 27 of *Lectures in Appl. Math.*, pages 285–291. Amer. Math. Soc., Providence, RI, 1991.

Ceremade (UMR CNRS 7534)
Université Paris-Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16, France
armstrong@ceremade.dauphine.fr