AN INFORMAL INTRODUCTION TO QUANTITATIVE STOCHASTIC HOMOGENIZATION

J.-C. MOURRAT

ABSTRACT. Divergence-form operators with random coefficients homogenize over large scales. Over the last decade, an intensive research effort focused on turning this asymptotic statement into quantitative estimates. The goal of this note is to review one approach for doing so based on the idea of renormalization. The discussion is highly informal, with pointers to mathematically precise statements.

1. INTRODUCTION

For most physical systems, the model we wish to use for its description depends on the time or space scale of interest. It is then a classical question of mathematical physics to ask whether these different models are consistent with one another as we move from one scale to another.

In this note, we focus on studying the large-scale behavior of divergence-form operators with random coefficients. We expect that, over large scales, an effective differential operator with constant coefficients suffices to capture the coarse properties of the heterogeneous operator. This property can indeed be verified mathematically, in the limit of infinite separation of scales. However, going beyond the verification of this asymptotic consistency and towards explicit rates of convergence, or even towards a precise description of next-order corrections, turns out to be a much more difficult challenge which has been the focus of intensive research over the last ten years. The goal of this note is to provide a gentle introduction to some of the work in this area.

We may choose the language of heat diffusion to describe the equations of interest. For convenience, we study equilibrium problems (that is, without time dependence) posed on a domain of $\mathbb{R}^d$. The equilibrium condition prescribes that the heat flux $\mathbf{j}$ be of null divergence. Moreover, by Fourier's law, we expect the heat flux to be proportional to the gradient of the temperature u, that is, $\mathbf{j} = -\mathbf{a}\nabla u$. We end up with the classical elliptic equation $-\nabla \cdot (\mathbf{a}\nabla u) = 0$.

The point of focus of this note concerns the situation where the medium in which heat diffuses is inhomogeneous: a typical example to have in mind is that of a composite material. That is, the coefficient field $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d \times d}_{\text{sym}}$ depends on the space variable (and takes values in the set of symmetric positive definite matrices). Yet, we do not want to study elliptic equations with arbitrary coefficients. Rather, we think that although the medium is disordered, the statistics of the disorder do not depend on the spatial location. Mathematically, this is encoded in the assumption that the coefficient field $(\mathbf{a}(x), x \in \mathbb{R}^d)$ is *random*, and that its law is invariant under translations. For our purposes, it suffices to assume that this invariance holds under translations along a lattice, which we can fix to be $\mathbb{Z}^d$ by a change of variables.
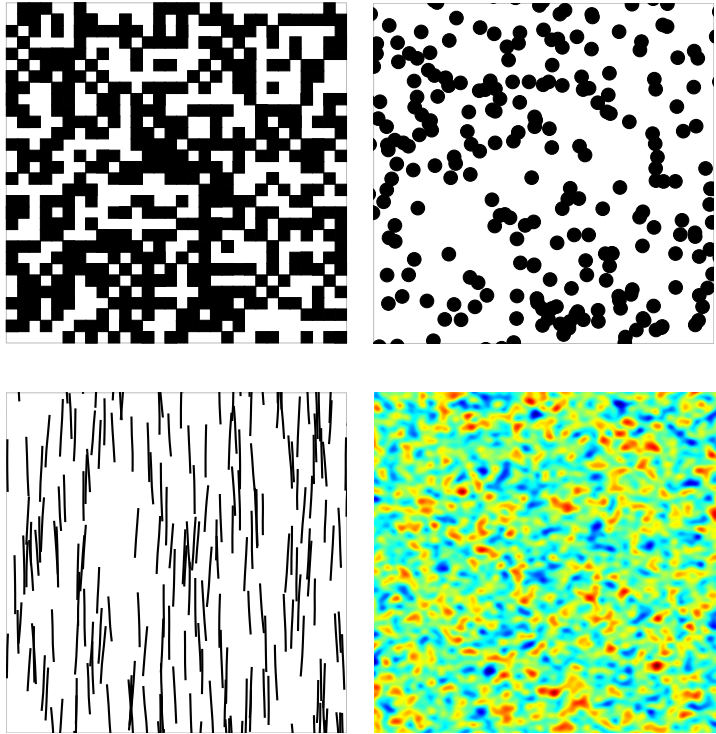
FIGURE 1. Examples of coefficient fields, the color scale indicating the value of $\mathbf{a}(x)$. Top left: each unit square is independently colored black or white with probability 1/2; top right: circles of a fixed radius are drawn around points of a Poisson point process; bottom left: similar to top right but with more complicated shapes; Bottom right: a local function of a white noise field in a color scale.

Besides the assumption that the coefficient field is random, we aim to make the simplest possible assumptions to facilitate the presentation of the main ideas. We will thus assume that the coefficient field is uniformly elliptic: there exists a constant $\Lambda \in [1, \infty)$ such that with probability 1, we have, for every $x \in \mathbb{R}^d$,

$$(1.1) \qquad \forall \xi \in \mathbb{R}^d, \quad \Lambda^{-1}|\xi|^2 \leqslant \xi \cdot \mathbf{a}(x)\xi \leqslant \Lambda|\xi|^2.$$

We also assume that the coefficient field has a unit range of dependence: whenever $U, V \subseteq \mathbb{R}^d$ are separated by a distance at least 1, the families of random variables $(\mathbf{a}(x), x \in U)$ and $(\mathbf{a}(x), x \in V)$ are independent. Examples of coefficient fields satisfying these assumptions are depicted on Figure 1.

In order to instantiate the effect of largely separated scales, we can introduce a small parameter $\varepsilon > 0$ and focus on understanding the solution u_ε of the problem

$$(1.2) \qquad \begin{cases} -\nabla \cdot \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u_\varepsilon = 0 & \text{in } U, \\ u_\varepsilon = f & \text{on } \partial U, \end{cases}$$

where the domain $U \subseteq \mathbb{R}^d$ and the boundary condition f are given and assumed to be sufficiently regular.
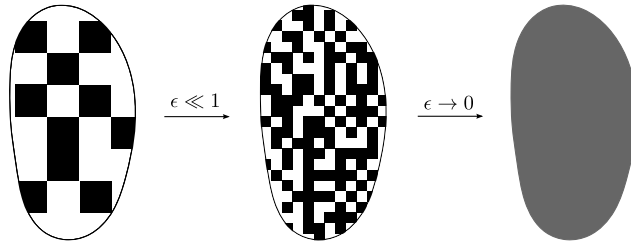
FIGURE 2. As $\varepsilon \to 0$, we expect the medium to "homogenize" and that the solution to (1.2) approaches the solution of an equation with constant coefficients.

As Figure 2 illustrates, as ε tends to 0, it is natural to expect the solution u_ε to resemble the solution of an effective "homogenized" equation with constant coefficients. This is indeed what happens, as has been proved in [23, 29, 28]: there exists a matrix $\bar{\mathbf{a}}$ such that with probability one, the function u_ε converges in $L^2(U)$ to the solution $\bar{u}$ of the equation

$$(1.3) \qquad \begin{cases} -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = 0 & \text{in } U, \\ \bar{u} = f & \text{on } \partial U. \end{cases}$$

The crucial feature of this result is that the matrix $\bar{\mathbf{a}}$ is now constant in space, and deterministic. Also, it depends neither on the domain U nor on the boundary condition f; it is only a function of the law of the original coefficient field.

As a matter of fact, the structure of elliptic equations allows us a much more direct access to gradients and fluxes of solutions than to the solution itself. Even if one is only interested in the convergence of u_ε to $\bar{u}$ stated above, it is thus much more convenient to focus on the convergence of the gradients and fluxes first: we have indeed that, with probability one,

$$(1.4) \qquad \nabla u_\varepsilon \rightharpoonup \nabla \bar{u} \quad \text{and} \quad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla u_\varepsilon \rightharpoonup \bar{\mathbf{a}} \nabla \bar{u} \qquad (\varepsilon \to 0).$$

The precise meaning of the arrow $\rightharpoonup$ is to denote weak convergence in $L^2(U)$. The important point to keep in mind is that $\nabla u_\varepsilon(x)$ will typically *not* be close to $\nabla \bar{u}(x)$: it is only after one takes a local spatial average of ∇u_ε that the result will resemble the local average of $\nabla \bar{u}$. This information on the weak convergence of ∇u_ε to $\nabla \bar{u}$ is more than enough to recover the convergence of u_ε to $\bar{u}$ in $L^2(U)$, see for instance [6, Lemma 1.8].

This result of convergence of u_ε to $\bar{u}$ can be thought of as a law of large numbers: there are many "independent random variables", each contributing little to the overall phenomenon, which are somehow averaged out and together lead to the emergence of an effective deterministic behavior. However, it is crucial to realize that, contrary to what Figure 2 may lead us to believe,

$$(1.5) \qquad \bar{\mathbf{a}} \text{ is } not \text{ the average of } \mathbf{a}(x).$$

This point is easiest to explain in the case of the coefficient field depicted on the bottom left frame of Figure 1, where we assume that $\mathbf{a}(x) = \text{Id}$ in the white region, and $\mathbf{a}(x) = 10^{-5} \text{Id}$ in the black region, say. In words, the black region essentially acts as an insulator. We imagine first trying to carry a flux of heat from the top to the bottom part of the square bounding the picture. While it would certainly be easier to do so in the absence of the black lines, their presence is only a minor

hindrance, since they are mostly aligned in the vertical direction. The situation
changes fundamentally if one wants to propagate a flux of heat from left to right:
in this case, the presence of the black lines forces the flux to "make a lot more
detours", and thus a unit temperature difference between the left and right sides
of the square will be associated with a much smaller heat flux. It follows that we
do not expect the homogenized matrix to be isotropic: denoting by e_1 and e_2 the
unit vectors in the horizontal and vertical directions respectively, we have argued
that $|\bar{\mathbf{a}}e_1| \ll |\bar{\mathbf{a}}e_2|$. In particular, this confirms the statement (1.5) that $\bar{\mathbf{a}}$ is not the
average of $\mathbf{a}(x)$, since the latter is an isotropic matrix.

In fact, this argument suggests that there will not be any simple formula for
calculating the homogenized matrix $\bar{\mathbf{a}}$ as a function of, for instance, the moments
of $\mathbf{a}(x)$. The homogenized matrix must summarize more geometric information
on the law of $\mathbf{a}(x)$ that measures, among other things, "the difficulty of going
around obstacles". (One may point out that a simple formula in fact does exist in
dimension $d = 1$, but there is not much question about how to "go around obstacles"
in one dimension.) The word "homogenization", as opposed to "averaging", is
sometimes used specifically in order to stress that the effective parameters are not
obtained by taking the average of the microscopic parameters.

The phenomenon of homogenization has a natural interpretation in terms of
stochastic processes. Indeed, the differential operator $-\nabla \cdot \mathbf{a}\nabla$ is the infinitesimal
generator of a diffusion process, which we may denote by $(X(t), t \geqslant 0)$. Seen in
this light, the convergence of u_ε to $\overline{u}$ is essentially equivalent to the statement
that the rescaled process $t \mapsto \varepsilon^{-\frac{1}{2}}X(\varepsilon^{-1}t)$ converges in law to a Brownian motion
with covariance matrix $2\bar{\mathbf{a}}$ as ε tends to 0 [28]. In other words, this probabilistic
interpretation recasts the convergence of u_ε to $\overline{u}$ as a central limit theorem for a
diffusion in the random environment given by $\mathbf{a}(x)$. Despite the caveat of (1.5) and
the probabilistic interpretation of homogenization as a central limit theorem, we
will continue thinking about the convergence of u_ε to $\overline{u}$ as a type of law of large
numbers.

The statement of homogenization is interesting for several reasons. For one, it
is among the simplest theoretical models for which the large-scale effective model
is not obtained by a simple average of the microscopic one. This makes it a very
suitable test bed for developing new theoretical ideas and techniques that may
also have some validity in other settings involving separations of scales. From a
computational perspective, the interest is also evident: while the computation of
the solution to (1.2) becomes arbitrarily difficult as ε tends to 0, the solution of
the homogeneous problem (1.3) can be computed very rapidly (provided that one
can calculate the homogenized matrix $\bar{\mathbf{a}}$ effectively). Depending on the quality of
approximation desired, this approximation may or may not be sufficiently precise,
and in the latter case, we may want to devise higher-order approximations of the
solution u_ε. These higher-order approximations may involve additional deterministic
corrections, and also, in the spirit of the central limit theorem, a Gaussian random
field whose statistics are captured by a handful of additional effective parameters.

These considerations motivate the development of a *quantitative* theory of ho-
mogenization. That is, we would like to evaluate the speed at which the solution u_ε
to (1.2) converges to the homogenized solution $\overline{u}$. This endeavor can also be seen

as a first step towards the construction of the higher-order approximations alluded to above.

In analogy with the standard central limit theorem for sums of independent random variables, the fundamental ingredient for obtaining rates of convergence should be encoded in the good mixing properties of the coefficient field: recall that we assume for simplicity that the coefficient field has a unit range of dependence. The difficulty of the problem is that the solution u_ε to (1.2) depends in a very complicated, non-linear and non-local way on the coefficient field $\mathbf{a}(x)$. The question boils down to: how do we "transfer" the very good mixing properties of the coefficient field into information on the solution u_ε? The first breakthrough was obtained in [17, 18, 16, 19], where the corrector (introduced below) was shown to be essentially of bounded growth (see also [30, 27, 24] for earlier sub-optimal results). Inspired by the fundamental insights of [26, 27], one of the main ingredients there is the use of "non-linear" concentration inequalities such as the Efron-Stein inequality. Recall that the Efron-Stein inequality states that, if $X = (X_1, \ldots, X_N)$ are independent random variables and $X' = (X_1', \ldots, X_N')$ is an independent copy of X, then for every function $F : \mathbb{R}^N \to \mathbb{R}$, we have

$$\mathrm{var}\,[F(X)] \leqslant \frac{1}{2} \sum_{i=1}^{N} \mathbb{E}\left[(F(X_1, \ldots, X_{i-1}, X_i', X_{i+1}, \ldots, X_N) - F(X))^2\right].$$

The very useful feature of inequalities such as this one is that they make no assumption whatsoever on the structure of the function F (hence the name "non-linear"). We will not discuss this approach further here, and focus instead on an alternative approach based on renormalization-type arguments. In this alternative approach, we do not try to study solutions of the problem (1.2) directly, and focus instead on energy-type quantities that display a simpler dependence on the coefficients, at least in the large-scale regime.

2. Renormalization heuristics for homogenization

In this section, we describe at a highly heuristic level why renormalization ideas should have a powerful take on the understanding of homogenization. We start by recording the following elementary but fundamental observation. Suppose that the coefficient field has very small oscillations, that is, there exists $\delta \ll 1$ such that with probability one, for every $x \in \mathbb{R}^d$,

$$(2.1) \qquad (1-\delta)\mathrm{Id} \leqslant \mathbf{a}(x) \leqslant (1+\delta)\mathrm{Id}.$$

In this case, we have, for some universal constant $C < \infty$,

$$(2.2) \qquad \left|\bar{\mathbf{a}} - \fint \mathbf{a}\right| \leqslant C\delta^2,$$

where we use the informal notation $\fint \mathbf{a}$ to denote the average of $\mathbf{a}(x)$, that is, with B_r denoting the ball of size r centered at the origin and $|B_r|$ denoting its volume,

$$\fint \mathbf{a} := \lim_{r \to \infty} \frac{1}{|B_r|} \int_{B_r} \mathbf{a}(x)\,\mathrm{d}x.$$

Since we assume the law of $\mathbf{a}(x)$ to be invariant under translations by vectors in $\mathbb{Z}^d$, we can also write

$$\fint \mathbf{a} = \mathbb{E}\left[\int_{[0,1]^d} \mathbf{a}(x)\,\mathrm{d}x\right].$$

In words, the inequality (2.2) says that, although (1.5) is indeed true and $\bar{\mathbf{a}}$ is not the average of $\mathbf{a}(x)$, in the regime of small oscillations of the coefficients, the average of $\mathbf{a}(x)$ provides an approximation of $\bar{\mathbf{a}}$ that is valid *up to second order* in the size of these oscillations.

The inequality (2.2) is an immediate consequence of the classical fact that $\bar{\mathbf{a}}$ is bounded above and below by, respectively, the arithmetic and harmonic means of $\mathbf{a}(x)$. That is, we have

(2.3) $$\left(\fint \mathbf{a}^{-1}\right)^{-1} \leqslant \bar{\mathbf{a}} \leqslant \fint \mathbf{a},$$

as will be explained in the next section around (3.6) and (3.8). For now, we take (2.2) for granted and proceed to explain why this should lend support for the possibility of a renormalization-type argument that would uncover the phenomenon of homogenization and its rate of convergence.

Instead of encoding widely separated scales via the introduction of a small parameter ε used to rescale the coefficient field as in (1.2), we adopt the equivalent and more convenient point of view that the coefficient field $\mathbf{a}(x)$ is fixed once and for all, and that we aim to understand problems that are posed over increasingly large scales. The coefficient field $\mathbf{a}(x)$ gives us the correspondence between gradients and fluxes at the microscopic level, while the homogenized matrix $\bar{\mathbf{a}}$ encodes the same correspondence, but in the limit of infinitely large scales. The idea of renormalization is that there should be "something happening inbetween" these two extremes, a sort of "progressive homogenization" that allows us to move smoothly accross scales. That is, for any given length scale $r > 0$, there should exist a coefficient field $\mathbf{a}_r(x)$ encoding the correspondence between gradients and fluxes of solutions "after we forget about small-scale details below scale r". Ideally, we would then have a "renormalization map" telling us how to calculate the coarser coefficient field, say $\mathbf{a}_{2r}(x)$, from the finer information encoded in $\mathbf{a}_r(x)$. An analysis of this renormalization map should then reveal the rate of convergence to the limit.

Assuming that the identification of such "coarsened coefficients" $\mathbf{a}_r(x)$ is indeed possible, we can intuitively guess how they will behave using (2.1)-(2.2). Indeed, as the length scale r increases to infinity, we expect $\mathbf{a}_r(x)$ to become more and more sharply concentrated around its constant limit $\bar{\mathbf{a}}$. In particular, this procedure of progressive homogenization automatically brings the coarsened coefficient field $\mathbf{a}_r(x)$ into a regime of small fluctuations very similar to that assumed in (2.1). The relation (2.2) expresses a relationship between the "infinite-scale" coarsened coefficient field $\bar{\mathbf{a}}$ and the global average of the microscopic coefficient field $\mathbf{a}(x)$. By analogy, we expect that as r becomes very large, the approximation of $\mathbf{a}_{2r}(x)$ by a local average of $\mathbf{a}_r(x)$ should become asymptotically sharp, up to a lower order error. If this is so and this procedure of progressive homogenization indeed becomes closer and closer to a simple local averaging procedure as $r \to \infty$, then we should expect $\mathbf{a}_r(x)$ to have asymptotic fluctuations dictated by the scaling of the

central limit theorem. That is, we expect $\mathbf{a}_r(x) - \bar{\mathbf{a}}$ to have Gaussian fluctuations of typical size $r^{-\frac{d}{2}}$.

Continuing with intuitive arguments, we explore possible consequences of this prediction (and along the way hopefully start to clarify, still on an informal level, a somewhat more precise meaning for the coarsened coefficient field $\mathbf{a}_r(x)$). For simplicity, let us fix $p \in \mathbb{R}^d$ and consider solving for the Dirichlet problem on a very large cube $\square$ with affine boundary condition $x \mapsto p \cdot x$ on $\partial\square$. Clearly, the solution to the corresponding homogenized problem is simply the affine function $x \mapsto p \cdot x$. In order to emphasize this, we write down the solution to the heterogeneous problem in the form $x \mapsto p \cdot x + \phi_p(x)$. That is, we solve for

$$(2.4) \qquad \begin{cases} -\nabla \cdot \mathbf{a}(p + \nabla\phi_p) = 0 & \text{in } \square, \\ \phi_p = 0 & \text{on } \partial\square, \end{cases}$$

and we are particularly interested in showing that $|\phi_p(x)| \ll |x|$ for large values of $|x|$. Although we will not justify this, it is possible to make sense of $\nabla\phi_p$ also in the limit where the size of the cube $\square$ is sent to infinity, and this limiting object is called the gradient of the *corrector* in the direction of p. (Contrarily to its gradient, the function ϕ_p itself may fluctuate unboundedly as the cube is blown up to infinity. In fact, its behavior can be compared with that of a Gaussian free field[1] that is regularized on a unit scale, as will be seen below.)

As was explained around (1.4), although we may be mostly interested in obtaining quantitative estimates on the growth of ϕ_p itself, we will have more direct access to information about its gradient and flux. We thus aim instead to give quantitative estimates on their weak convergence. More precisely, we would like to know the rate at which the spatial average of $\nabla\phi_p$ over a region of size r converges to 0; and similarly for the convergence of spatial averages of $\mathbf{a}(p + \nabla\phi_p)$ towards $\bar{\mathbf{a}}p$.

We aim to obtain this information from our prediction on the asymptotic size of the difference $\mathbf{a}_r(x) - \bar{\mathbf{a}}$. In order to do so, we need to "forget about small scale details" of the solution $x \mapsto p \cdot x + \phi_p(x)$ below scale r. For instance, we may take a smooth function $\chi \in C_c^\infty(\mathbb{R}^d)$ with compact support and satisfying $\int_{\mathbb{R}^d} \chi = 1$, and then define, for every $r \geqslant 1$, the convolution

$$\phi_{p,r} := \phi_p * \left(r^{-d} \chi \left(\tfrac{\cdot}{r} \right) \right).$$

The fundamental idea of the coarsened coefficient field $\mathbf{a}_r(x)$ is that it should govern, at least approximately, the behavior of this coarsened solution $\phi_{p,r}$, in the sense that

$$-\nabla \cdot \mathbf{a}_r \left(p + \nabla\phi_{p,r} \right) \simeq 0.$$

Denoting

$$\mathsf{W}_r(x) := \mathbf{a}_r(x) - \bar{\mathbf{a}},$$

we may rearrange the approximate equation above in the form

$$(2.5) \qquad -\nabla \cdot \mathbf{a}_r \nabla\phi_{p,r} \simeq \nabla \cdot \left(\mathsf{W}_r p \right).$$

---

[1] The name "Gaussian free field" has become standard in the probability literature, although it is a bit redundant, since a random field is said to be free if it is Gaussian. The name "masless free field" is usually favored in the theoretical physics literature.
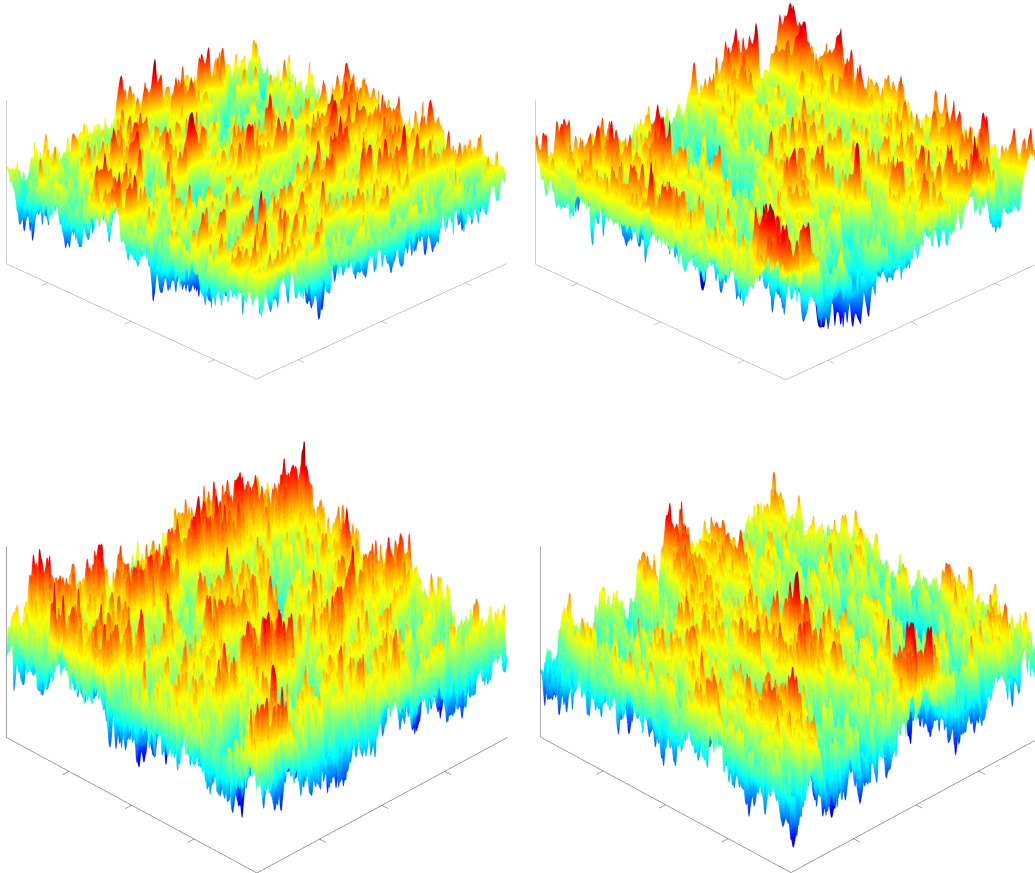
FIGURE 3. Top: correctors ϕ_{e_1} and ϕ_{e_2} in the directions of the two basis vectors e_1 and e_2, for random coefficients as in the top left frame of Figure 1. Bottom: examples of Gaussian random fields with the structure predicted by (2.7), after convolution with a smooth bump function on a unit scale.

Recalling our prediction that the size of $\mathsf{W}_r(x)$ should be of the order of $r^{-\frac{d}{2}}$, a classical testing of the equation above against $\phi_{p,r}$ suggests that

$$(2.6) \qquad\qquad |\nabla \phi_{p,r}| \lesssim r^{-\frac{d}{2}}.$$

This is exactly the information we were looking for: indeed, since gradients and convolutions commute, the estimate above states that the average of $\nabla \phi_p$ over a region of size r decays to 0 like $r^{-\frac{d}{2}}$. Once such information is obtained, one can recover sharp information on the growth of ϕ_p itself: it grows like $\log^{\frac{1}{2}} |x|$ in dimension $d = 2$, and remains essentially bounded in dimension $d \geqslant 3$ (see [6, Theorem 4.1]).

In fact, even finer information can be read off from the approximate equation (2.5). Indeed, since the prediction for the size of W_r was obtained via a central limit theorem argument, the reasoning in fact yields a more refined description of this quantity: namely, we expect it to closely resemble the convolution on scale r of a white noise random field. Moreover, in view of (2.6) and of the fact that $|\mathbf{a}_r - \bar{\mathbf{a}}| \lesssim r^{-\frac{d}{2}}$, we see that replacing $\mathbf{a}_r$ by $\bar{\mathbf{a}}$ in (2.5) produces a lower-order error,

and thus

$$(2.7) \qquad -\nabla \cdot \bar{\mathbf{a}} \nabla \phi_{p,r} \simeq \nabla \cdot (\mathsf{W}_r p).$$

Replacing $\simeq$ by $=$ and W_r by a true white noise, we would obtain the equation defining a Gaussian random field very similar to a Gaussian free field (see [6, Section 5.1]). The approximate equation (2.7) thus suggests that the law of the large-scale convolution of ϕ_p becomes asymptotically very close to the law of the large-scale convolution of this Gaussian random field, a result that can be compared with a central limit theorem; see [6, Theorem 5.24] for a precise statement, and Figure 3 for an illustration.

3. A MATHEMATICAL APPROACH TO RENORMALIZATION FOR HOMOGENIZATION

In my opinion, the intuitive picture described in the previous section is extremely compelling, and strongly calls for a mathematical embodiement. It is the purpose of the present section to review, at a more precise but still highly informal level, the outline of such an approach as developed in [10, 9, 7, 8]; see also [20] for an alternative approach in a similar spirit, and [6] for a book on the topic.

The most fundamental task is to identify a candidate for the notion of coarsened coefficients discussed in the previous section. Borrowing an idea of [11, 12], we introduce, for each bounded domain $U \subseteq \mathbb{R}^d$ and $p \in \mathbb{R}^d$, the energy-type quantity

$$(3.1) \qquad \nu(U,p) \coloneqq \inf_{u \in \ell_p + H_0^1(U)} \fint_U \frac{1}{2} \nabla v \cdot \mathbf{a} \nabla v,$$

where $\ell_p(x) \coloneqq p \cdot x$, we use the notation $\fint_U \coloneqq |U|^{-1} \int_U$ for the integral normalized by the volume of U, and $H_0^1(U)$ denotes the set of functions in $L^2(U)$ with gradient in $L^2(U)$ and which vanish on ∂U. It is classical to verify that the minimizer of this problem exists and is unique, and is simply the $\mathbf{a}$-harmonic function with boundary condition given by $p \cdot x$. (We say that a function v is $\mathbf{a}$-harmonic if it satisfies $-\nabla \cdot \mathbf{a} \nabla v = 0$.)

The quantity $\nu(U,p)$ satisfies a number of remarkable properties. In particular, the mapping $U \mapsto \nu(U,p)$ is subadditive. In order to verify this, consider a partition of the domain U into subdomains $U_1, \ldots, U_k$. Then the minimizers $v_1, \ldots, v_k$, defined on $U_1, \ldots, U_k$ respectively, share a consistent boundary condition on the interfaces between subdomains. As a consequence, we can glue them together and obtain a candidate function for the minimization in the larger domain U. It thus follows that

$$(3.2) \qquad \nu(U,p) \leqslant \sum_{i=1}^{k} \frac{|U_i|}{|U|} \nu(U_i, p).$$

Another elementary but very useful property of $\nu(U,p)$ is that it depends only on the coefficient field inside U. Finally, for convenience, we also note that the mapping $p \mapsto \nu(U,p)$ is quadratic. This follows from the fact that the solution of the Dirichlet problem with boundary condition $p \cdot x$ depends linearly on p. Hence, there exists a symmetric matrix $\mathbf{a}(U)$ such that for every $p \in \mathbb{R}^d$,

$$(3.3) \qquad \nu(U,p) = \frac{1}{2} p \cdot \mathbf{a}(U) p.$$

This matrix $\mathbf{a}(U)$ is our candidate "homogenized matrix on the scale of the domain U". It encodes the (spatially averaged) energy of solutions of Dirichlet problems on the domain U with affine boundary condition. For every $r > 0$, we denote the cube of side length r centered at the origin by

$$\square_r := \left(-\frac{r}{2}, \frac{r}{2}\right)^d.$$

The subadditivity property (3.2) implies that with probability one, the quantity $\nu(\square_r, p)$ converges to some limit as r tends to infinity. This can be obtained by an application of the very general subadditive ergodic theorem [1], or more directly using the stronger independence assumptions at our disposal, see [6, Proposition 1.3]. Since the mapping $p \mapsto \nu(\square_r, p)$ is quadratic, the limit will also be quadratic, and we *define* the matrix $\bar{\mathbf{a}}$ so that

$$(3.4) \qquad\qquad \lim_{r \to \infty} \nu(\square_r, p) = \frac{1}{2} p \cdot \bar{\mathbf{a}} p.$$

It is then our task to show that homogenization happens with this definition of the matrix $\bar{\mathbf{a}}$, and to quantify the rate of convergence.

The fact that $\mathbf{a}(U)$ is defined as the spatially averaged energy of certain solutions is a first indication that it may serve as a suitable notion of coarsened coefficient field. We can also witness that this matrix also encodes a correspondence between coarsened versions of the gradients and fluxes of solutions. Indeed, denote by $v(\cdot, U, p)$ the minimizer in the definition (3.1) of $\nu(U, p)$. Since $v \in \ell_p + H_0^1(U)$, we have

$$(3.5) \qquad\qquad \fint_U \nabla v(\cdot, U, p) = p.$$

Moreover, for every $p, p' \in \mathbb{R}^d$, we have

$$p' \cdot \mathbf{a}(U) p = \fint_U \nabla v(\cdot, U, p') \cdot \mathbf{a} \nabla v(\cdot, U, p).$$

Indeed, this identity with $p' = p$ boils down to definitions; the more general version then follows using elementary symmetry properties of the bilinear forms. Recalling that $v(\cdot, U, p') \in \ell_{p'} + H_0^1(U)$, and that $v(\cdot, U, p)$ is an $\mathbf{a}$-harmonic function, we deduce that

$$\fint_U \nabla \left(v(\cdot, U, p') - \ell_{p'} \right) \cdot \mathbf{a} \nabla v(\cdot, U, p) = 0,$$

and thus

$$p' \cdot \mathbf{a}(U) p = \fint_U p' \cdot \mathbf{a} \nabla v(\cdot, U, p),$$

that is,

$$\mathbf{a}(U) p = \fint_U \mathbf{a} \nabla v(\cdot, U, p).$$

Comparing this last identity with (3.5), we thus obtain that $\mathbf{a}(U)$ transfers information about the spatial average of the gradient of a solution into information about the spatial average of its flux.

From the definition of $\bar{\mathbf{a}}$ in (3.4), we can already justify the rightmost inequality appearing in (2.3). Indeed, by choosing the affine function ℓ_p as a candidate minimizer in (3.1), we immediately get that

$$(3.6) \qquad\qquad \mathbf{a}(U) \leqslant \fint_U \mathbf{a}.$$

The desired inequality between $\bar{\mathbf{a}}$ and the average of $\mathbf{a}(x)$ follows by choosing $U = \Box_r$ and then letting r tend to infinity. For the leftmost inequality in (2.3), one can first show that

$$\nu(U, p) = \sup_{\mathbf{g}} \fint_U \left(-\frac{1}{2} \mathbf{g} \cdot \mathbf{a}^{-1} \mathbf{g} + p \cdot \mathbf{g} \right),$$

where the supremum is taken over every divergence-free vector field $\mathbf{g} \in L^2(U)$ (see [6, (2.8)]). This is a "dual" formulation of the Dirichlet problem that puts emphasis on fluxes rather than gradients of solutions. Choosing $\mathbf{g} = q$ for a fixed $q \in \mathbb{R}^d$ (constant flux instead of constant gradient as in the previous argument) then yields that

$$(3.7) \qquad \nu(U, p) \geqslant p \cdot q - \frac{1}{2} q \cdot \left(\fint_U \mathbf{a}^{-1} \right) q.$$

Optimizing over q, we obtain that

$$(3.8) \qquad \nu(U, p) \geqslant \frac{1}{2} p \cdot \left(\fint_U \mathbf{a}^{-1} \right)^{-1} p,$$

and this implies the leftmost inequality in (2.3).

This reasoning suggests the possible relevance to study not only solutions of Dirichlet problems with affine boundary conditions, but also Neumann problems. We thus introduce a corresponding energy: for every bounded domain U and $q \in \mathbb{R}^d$, we set

$$\nu^*(U, q) := \sup_{u \in H^1(U)} \fint_U \left(-\frac{1}{2} \nabla u \cdot \mathbf{a} \nabla u + q \cdot \nabla u \right).$$

The maximizer of this problem is unique up to the addition of a constant, and is the $\mathbf{a}$-harmonic function such that $(\mathbf{a} \nabla u - q) \cdot \mathbf{n}$ vanishes on ∂U, where $\mathbf{n}$ denotes the outer normal to ∂U. The mapping $U \mapsto \nu^*(U, q)$ is also subadditive, for reasons "opposite" to ν: the maximizer in the definition of $\nu^*(U, q)$ restricted to subdomains gives us candidate functions on the subdomains.

Recalling (3.5) and using the minimizer $v(\cdot, U, p)$ of $\nu(U, p)$ as a candidate for this problem, we obtain that for every $p, q \in \mathbb{R}^d$,

$$(3.9) \qquad \nu^*(U, q) \geqslant p \cdot q - \nu(U, p).$$

Notice the similarity between this inequality and that in (3.7). In words, the mapping $q \mapsto \nu^*(U, q)$ is an upper bound for the convex dual of the mapping $p \mapsto \nu(U, p)$. Since we expect homogenization to happen in the large-scale limit, it is plausible that over large scales, the maximizer in the definition of $\nu^*(U, q)$ becomes very close to some function $v(\cdot, U, p)$ for the appropriate choice of the slope p. That is, we expect $\nu(U, \cdot)$ and $\nu^*(U, \cdot)$ to become approximately convex dual to one another in the large-scale limit. This suggests in particular that

$$(3.10) \qquad \lim_{r \to \infty} \nu^*(\Box_r, q) = \frac{1}{2} q \cdot \bar{\mathbf{a}}^{-1} q.$$

The rough outline allowing to obtain the predicted rate of convergence in the limit (3.4) can be described as follows.

(1) We show the existence of a possibly very small exponent $\alpha > 0$ such that

$$(3.11) \qquad \left| \nu(\Box, p) - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \right| \lesssim |\Box|^{-\alpha}.$$

(2) By (3.11), the coarsened coefficients on the scale of a given cube $\square$ fluctuate by approximately $|\square|^{-\alpha}$. In view of (2.1)-(2.2), we thus expect that

$$(3.12) \qquad \left| \nu(\square, p) - 2^{-d} \sum_{z \in \mathcal{Z}} \nu(z + \square', p) \right| \lesssim |\square|^{-2\alpha},$$

where we decomposed the cube $\square$ into 2^d subcubes of half the side length, which we denote by $(z + \square')_{z \in \mathcal{Z}}$.

(3) By (3.12), the quantity ν is almost additive, and thus, using independence,

$$(3.13) \qquad |\nu(\square, p) - \mathbb{E}[\nu(\square, p)]| \lesssim |\square|^{-((2\alpha) \wedge \frac{1}{2})}.$$

In (3.13), we use the notation $(2\alpha) \wedge \frac{1}{2}$ to denote the minimum between 2α and $\frac{1}{2}$. From (3.12), we also infer that the difference between the expectation of $\nu(\square, p)$ and its limit $\frac{1}{2} p \cdot \bar{\mathbf{a}} p$ is bounded by $C|\square|^{-2\alpha}$. In this rough outline, we thus end up with the conclusion that we can in fact replace the exponent α in (3.11) by the exponent $(2\alpha) \wedge \frac{1}{2}$. In principle, the reasoning could then be iterated until all statements are known with α replaced by the exponent $\frac{1}{2}$, in agreement with the prediction of the previous section.

In fact, we already know from (3.2) that the quantity between absolute values in (3.12) is nonpositive. The control of this quantity from the other side should resemble the proof of the leftmost inequality in (2.3), but at the level of coarsened coefficients. The quantity $\nu^*(U, q)$ is a fundamental tool for this purpose.

Part (3) of the outline above is the simplest to understand. Indeed, if we pretend for a moment that the additivity of $U \mapsto \nu(U, p)$ is exact (that is, that the left side of (3.12) is actually zero), then the calculation of $\nu(\square, p)$ really boils down to taking the average of the cubes at a slightly smaller scale. We can thus apply classical techniques to prove central limit theorems for sums of independent random variables, and obtain fluctuations of the order of $|\square|^{-\frac{1}{2}}$. In practice, the result can only be as good as the error in the additivity statement (3.12), and this explains the exponent on the right side of (3.13).

We now say a few words on how to obtain the statement (3.11) for some possibly very small exponent $\alpha > 0$. The real difficulty is to control the rate of convergence of the expectation of $\nu(\square_r, p)$; fluctuations can then be recovered using subadditivity (to be precise, subadditivity of ν and ν^* allow to control their respective upper fluctuations, and we can relate the lower fluctuations of ν with the upper fluctuations of ν^*). In order to obtain this small exponent in the rate of convergence of the quantity $\mathbb{E}[\nu(\square_r, p)]$ towards its limit, it would be sufficient to assert the existence of a constant $C < \infty$ such that for every $r \geq 1$,

$$(3.14) \qquad \mathbb{E}[\nu(\square_r, p)] - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \leq C \left(\mathbb{E}[\nu(\square_r, p)] - \mathbb{E}[\nu(\square_{2r}, p)] \right).$$

(Subadditivity and stationarity imply that both sides are nonnegative). Indeed, a rearrangement would then give that

$$\mathbb{E}[\nu(\square_{2r}, p)] - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \leq \frac{C}{C-1} \left(\mathbb{E}[\nu(\square_r, p)] - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \right),$$

which could be iterated to produce the desired decay in r, at least along powers of 2. It seems however very difficult to show an inequality such as (3.14) directly, in particular because $\bar{\mathbf{a}}$ is defined as an infinite-volume limit, while we have more

direct access to finite-volume quantities. The idea of [10] is to try instead to prove a "finite-volume" statement of the form

$$(3.15) \qquad \mathbb{E}\left[\nu(\Box_r, p)\right] - \sup_{q \in \mathbb{R}^d}\left(p \cdot q - \mathbb{E}\left[\nu^*(\Box_r, q)\right]\right) \leqslant C\tau(r),$$

where similarly to the right side of (3.14), the quantity $\tau(r)$ is meant to capture the "additivity defect" for the quantities ν and ν^*, for instance

$$\tau(r) := \sup_{|p| \leqslant 1}\left(\mathbb{E}\left[\nu(\Box_r, p)\right] - \mathbb{E}\left[\nu(\Box_{2r}, p)\right]\right) + \sup_{|q| \leqslant 1}\left(\mathbb{E}\left[\nu^*(\Box_r, q)\right] - \mathbb{E}\left[\nu^*(\Box_{2r}, q)\right]\right).$$

In view of (3.10), we expect the supremum over q on the left side of (3.15) to approach $\frac{1}{2}p \cdot \bar{\mathbf{a}}p$ as r tends to infinity. In words, the inequality (3.15) states that the "convex duality defect" between ν and ν^*, see (3.9), can be controlled by the additivity defect between two scales. Let us argue very briefly on the plausibility of the idea that a small additivity defect should imply a small convex duality defect. Subadditivity of ν and ν^* is obtained by comparing the optimizer on the larger scale with a patching of essentially independent and identically distributed pieces on the smaller scale. The smallness of the additivity defect should indicate that these two objects are relatively close. This in turn should imply that the large-scale maximizer remains relatively close to an affine function. Finally, if we know that the maximizer for $\nu^*(\Box_r, q)$ is close to an affine function of slope p, say, then we would expect it to be close to the function $v(\cdot, \Box_r, p)$ as well (indeed, if $v(\cdot, \Box_r, p)$ is close to some affine function, then this must be the affine function with slope p, by (3.5)). From this, we would then deduce that the inequality (3.9) is in fact almost an equality, as desired. A more accurate discussion would involve quantifying closeness to affine functions in $L^2(\Box_r)$, and then using the Caccioppoli inequality to deduce closeness of the optimizers in $H^1(\Box_r)$. Precise statements and proofs can be found in [6, Chapter 2].

We now discuss the meaning of the symbols $\lesssim$ in (3.11)-(3.13). They are not only meant to hide multiplicative constants. Indeed, in any finite volume, there is some chance for the coefficient field to be completely atypical, and in this scenario, the left side of (3.11) will *not* be small. The inequalities (3.11)-(3.13) should thus be interpreted as statements that hold with high probability. That is, for any random variable X and $\theta > 0$, the notation $X \lesssim \theta$ should encode that the random variable $\theta^{-1}X$ remains tightly controlled. Since we appeal to the central limit theorem in the derivation of (3.13), a good notion for what "tightly controlled" could mean is that the random variable has Gaussian tails. More generally, for any $s > 0$, we use the notation

$$(3.16) \qquad X \leqslant \mathcal{O}_s(\theta) \quad \Longleftrightarrow \quad \mathbb{E}\left[\exp\left((\theta^{-1}X)_+^s\right)\right] \leqslant 2,$$

where $x_+ := \max(x, 0)$ denotes the positive part of $x \in \mathbb{R}$. The case $s = 2$ corresponds to a Gaussian tail behavior, see [6, Appendix A]. This notation helps to clarify an important difficulty in the inductive argument sketched above: if we have (3.11) with right side of the form $\mathcal{O}_s\left(C|\Box|^{-\alpha}\right)$, then the reasoning in step (2) leads to an inequality (3.12) with right side of the form $\left(\mathcal{O}_s(C|\Box|^{-\alpha})\right)^2$, that is, $\mathcal{O}_{s/2}\left(C^2|\Box|^{-2\alpha}\right)$. While the size of the error itself has improved, the exponent s of stochastic integrability has been divided by a factor of 2. In practice, this means that the probability of large deviations is no longer estimated sharply. We

circumvent this difficulty by always relying on our initial estimate (3.11) with very small exponent α, for which we briefly sketched the argument above, in order to estimate these large deviation events. Technically, it then becomes more convenient to simply improve α by a small fixed amount at each step of the iteration, with the "moderate deviations" being improved by squaring, while keeping the seed "large deviation" estimate intact from the onset. As a consequence, it is fundamental that the initial estimate (3.11) with very small exponent $\alpha > 0$ gives us very strong control on the probability of rare events. The initial estimate should thus give us in particular a statement of the following form: that for any fixed $\varepsilon > 0$, one can find $\alpha > 0$ and a constant $C < \infty$ such that for every cube $\square$,

$$\mathbb{P}\left[\left|\nu(\square, p) - \frac{1}{2}p \cdot \bar{\mathbf{a}}p\right| \leqslant |\square|^{-\alpha}\right] \leqslant C \exp\left(-|\square|^{(1-\varepsilon)}\right).$$

It is easy to see that this estimate is essentially best possible. Indeed, taking the "random cherckerboard" depicted in the top left frame of Figure 1 as an example of coefficient field, we see that the probability to observe only white unit squares in a given large cube $\square$ is of the order of $\exp(-c|\square|)$, and in this case, the difference $\nu(\square, p) - \frac{1}{2}p \cdot \bar{\mathbf{a}}p$ will be of order one.

The final, and technically most problematic, caveat to the outline described in the steps (1-3) above is that in fact, the optimizers for ν and ν^* cannot be close to one another at very high precision, since they will be at a distance of order 1 from one another at least in a unit-scale neighborhood of the boundary. The presence of this boundary layer means that it is not possible to establish the inequality (3.12) for a right side which is smaller than $|\square|^{-\frac{1}{d}}$, a suboptimal conclusion that was the main result of [7]. Going past this boundary phenomenon requires to redefine the quantities ν and ν^* in a "smoother way", and was the main achievement of [8]. The whole strategy is exposed in details in [6, Chapter 4].

A core ingredient entering the formalization of this strategy, first introduced for this purpose in [10], is the notion of large-scale regularity of $\mathbf{a}$-harmonic functions. It is simplest to explain this idea by analogy with the more classical Schauder estimates. These estimates state that if the coefficient field $\mathbf{a}$ is α-Hölder continuous, then any $\mathbf{a}$-harmonic function u in the unit ball B_1 is Lipschitz continuous (in fact, $C^{1,\alpha}$) in the interior of the ball—see [6, (3.8)] for a quantitative statement. The idea of the proof is to compare, over a series of balls with geometrically decreasing radius, the true solution with the solution of the equation with "frozen", constant coefficients given by the value of the true coefficient field at the center of the ball. This allows to transfer the higher regularity properties of solutions of equations with constant coefficients onto the solution of the heterogeneous equation; see [6, (3.9)-(3.10)] and [22, Chapter 3] for a more precise discussion.

For equations with random coefficients, we use a similar idea, except that it is over the *largest* scales that $\mathbf{a}$-harmonic functions become closer and closer to solutions of constant-coefficient equations. That is, once the validity of (3.11) for a small exponent $\alpha > 0$ is established, one can deduce an error estimate for homogenization which one then uses to replace the argument involving a local "freezing" of the coefficients in the classical Schauder estimate. The conclusion of the argument is, roughly speaking, that $\mathbf{a}$-harmonic functions are essentially "Lipschitz continuous from the largest scale down to the unit scale". (Naturally, homogenization does not

inform us about regularity on scales smaller than the unit scale.) That is, for every $r \geqslant 1$ and every $\mathbf{a}$-harmonic function u in the ball B_r, we have

$$\fint_{B_1} |\nabla u|^2 \lesssim \fint_{B_r} |\nabla u|^2,$$

where the implicit constant is independent of the radius r (and $\lesssim$ still hides probabilistic quantifiers). Using the Caccioppoli inequality, see [6, Lemma C.2], and invoking the result above with r replaced by $r/2$, we obtain that, for every $r \geqslant 1$,

(3.17) $$\fint_{B_1} |\nabla u|^2 \lesssim r^{-1} \fint_{B_r} u^2.$$

As a striking illustration of the power of this estimate, we now observe that it immediately implies gradient estimates for Green functions. Restricting ourselves to dimensions $d \geqslant 3$ for simplicity, we denote by $(G(x,y), x, y \in \mathbb{R}^d)$ the Green function associated with the operator $-\nabla \cdot \mathbf{a}\nabla$ on $\mathbb{R}^d$. It is well-known (and can be deduced from the similar parabolic construction recalled in [6, Appendix E]) that for arbitrary coefficient fields satisfying the uniform ellipticity condition (1.1), the Green function exists and satisfies the bound

$$G(x,y) \leqslant \frac{C}{|x-y|^{d-2}}.$$

Since the function $x \mapsto G(x,0)$ is $\mathbf{a}$-harmonic in the ball $B_{|x|/2}(x)$ of radius $\frac{|x|}{2}$ centered at x, we can apply the estimate (3.17) in this ball and conclude that

$$\left(\fint_{B_1(x)} |\nabla_x G(\cdot, 0)|^2 \right)^{\frac{1}{2}} \lesssim \frac{C}{|x|^{d-1}}.$$

The precise formulation and proof of the regularity estimate (3.17), as well as higher-order versions thereof, are explained in [6, Chapter 3]. The consequences of these estimates for parabolic and elliptic Green functions (in any dimension) are worked out in [6, Chapter 8], and more refined information on the homogenization of the Green function is obtained in [6, Section 9.2]. The latter results may be of particular interest to probabilists, since they can be interpreted as a quantitative quenched local central limit theorem for the diffusion process.

To sum up, one can implement a rigorous version of the ideas exposed in the previous section and in the rough outline above, subject to some caveats, the most important of which being the problem caused by boundary layers. Once this is achieved and a sharp estimate on the fluctuations of the coarsened coefficients is known, one can essentially reproduce the argument leading to (2.6) and obtain optimal estimates on the large-scale behavior of the correctors ϕ_p. This can be considered as the most fundamental result of [6], and is achieved in Chapter 4 therein. Following the outline leading to (2.7), one can refine this information further and show that the corrector rescales to a variant of the Gaussian free field. This is the purpose of [6, Chapter 5] (which can also serve as a good reference for the definition and basic properties of white noise and Gaussian free fields).

Once optimal estimates on the corrector are obtained, one can deduce optimal error estimates for general homogenization problems such as the convergence of the solution to (1.2) as ε tends to zero. Intuitively, this is made possible by the fact that the limit homogenized solution is smooth; on scales intermediate between

the microscopic scale of oscillation of the coefficients and the macroscopic scale of the domain, the homogenized solution is thus well approximated by an affine function, and therefore the solution to (1.2) is well approximated by a corrector function with slope given by the local value of the gradient of the homogenized solution. Formally, this is encoded in a comparison between the solution of the heterogeneous equation and a "two-scale expansion" which consists in attaching the oscillations of correctors to the solution of the homogeneous problem. Optimal error estimates for such problems are proved in [6, Chapter 6]. In particular, it is shown that for sufficiently smooth domain U and boundary condition f, and for u_ε and $\overline{u}$ solution to (1.2) and (1.3) respectively, we have

$$(3.18) \qquad \|u_\varepsilon - \overline{u}\|_{L^2(U)} \lesssim \begin{cases} \varepsilon \, |\log \varepsilon|^{\frac{1}{2}} & \text{if } d = 2, \\ \varepsilon & \text{if } d \geqslant 3. \end{cases}$$

These estimates are sharp. (In dimension $d = 1$, it is easy to see that the difference between u_ε and $\overline{u}$ looks like a random walk trajectory with steps of size ε, and in particular $\|u_\varepsilon - \overline{u}\|_{L^2(U)}$ is of the order of $\sqrt{\varepsilon}$.) A precise version of the estimates in (3.18) can be found in [6, Theorem 6.14].

4. Conclusion and perspectives

Summarizing, renormalization ideas provide a powerful guiding principle towards the establishment of optimal quantitative estimates in the homogenization of problems of the form of (1.2). The driving mechanism is a change of focus away from a direct analysis of solutions, aiming instead at a "progressive homogenization" of the equation itself, or more precisely, of the associated energy.

Beyond the interest of the specific problem of finding rates of convergence for the homogenization of equations of the form of (1.2), the mathematical implementation of this program is an opportunity to develop a rich set of ideas and tools that can be of interest for a variety of related questions. For example, the observation (2.1)-(2.2) was influential for the development of new algorithms for the computation of the homogenized matrix $\overline{\mathbf{a}}$, see [25]. Perhaps surprisingly, the homogenized operator can also be used for the efficient computation of the solution to (1.2) at arbitrarily high precision. Indeed, the computational avatar of the idea of renormalization is the multigrid method. This method does not perform well for highly oscillating coefficients, but it was realized in [5, 21] that the homogenized operator can be used as a suitable coarse operator in such methods to restore high computational efficiency.

The results presented here have already been extended to a variety of other settings. For instance, elliptic equations with non-symmetric coefficients and parabolic equations with fixed or time-dependent coefficients can be treated with similar techniques, see [6, Chapters 10 and 8] and [2]. Some significant steps have been taken towards very general non-linear equations, see [6, Chapter 11] and [4]. The uniform ellipticity assumption (1.1) can also be relaxed: indeed, results similar to those presented here have been obtained in the case of discrete finite-difference equations posed on a percolation cluster [3, 13]. Some results have also been obtained for problems involving differential forms [14], as well as for a probability model of interfaces known as the Ginzburg-Landau or "$\nabla\phi$" model [15]. Future

developments will hopefully include the analysis of certain hypoelliptic equations and interacting particle systems.

References

[1] M. A. Akcoglu and U. Krengel. Ergodic theorems for superadditive processes. *J. Reine Angew. Math.*, 323:53–67, 1981.

[2] S. Armstrong, A. Bordas, and J.-C. Mourrat. Quantitative stochastic homogenization and regularity theory of parabolic equations. *Anal. PDE*, 11(8):1945–2014, 2018.

[3] S. Armstrong and P. Dario. Elliptic regularity and quantitative homogenization on percolation clusters. *Comm. Pure Appl. Math.*, 71(9):1717–1849, 2018.

[4] S. Armstrong, S. Ferguson, and T. Kuusi. Homogenization, linearization and large-scale regularity for nonlinear elliptic equations, preprint, arXiv:1805.00467.

[5] S. Armstrong, A. Hannukainen, T. Kuusi, and J. C. Mourrat. An iterative method for elliptic problems with rapidly oscillating coefficients, preprint, arXiv:1803.03551.

[6] S. Armstrong, T. Kuusi, and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity.* Preliminary version available at arXiv:1705.05300.

[7] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, 347(2):315–361, 2016.

[8] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, 208(3):999–1154, 2017.

[9] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.

[10] 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.

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

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

[13] P. Dario. Optimal corrector estimates on percolation clusters, preprint, arXiv:1805.00902.

[14] P. Dario. Quantitative homogenization of differential forms, preprint, arXiv:1806.07760.

[15] P. Dario. Quantitative homogenization of the disordered $\nabla\phi$ model, preprint, arXiv:1810.06428.

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

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

[19] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc. (JEMS)*, 19(11):3489–3548, 2017.

[20] A. Gloria and F. Otto. The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations, preprint, arXiv:1510.08290.

[21] C. Gu. Uniform estimate of an iterative method for elliptic problems with rapidly oscillating coefficients, preprint, arXiv:1807.06565.

[22] Q. Han and F. Lin. *Elliptic partial differential equations*, volume 1 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1997.

[23] S. M. Kozlov. Averaging of differential operators with almost periodic rapidly oscillating coefficients. *Mat. Sb. (N.S.)*, 107(149)(2):199–217, 317, 1978.

[24] 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.

[25] J.-C. Mourrat. Efficient methods for the estimation of homogenized coefficients. *Found. Comput. Math.*, in press.

[26] A. Naddaf and T. Spencer. On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, 183(1):55–84, 1997.

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

[28] 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.

[29] V. V. Yurinskiĭ. On a Dirichlet problem with random coefficients. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 344–353. Springer, Berlin-New York, 1980.

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

(J.-C. Mourrat) DMA, Ecole normale supérieure, CNRS, PSL University, Paris, France

E-mail address: mourrat@dma.ens.fr