Qualitative and quantitative results in stochastic homogenization

Antoine Gloria

Mémoire
présenté pour l'obtention du diplôme

d'HABILITATION À DIRIGER DES RECHERCHES
Université Lille 1

Spécialité : Mathématiques appliquées

par

Antoine GLORIA

Sujet : *Qualitative and quantitative results in stochastic homogenization*

Soutenance le 24 février 2012 devant le jury composé de :

Rapporteurs : Robert V. Kohn
Stefano Olla
George C. Papanicolaou

Examinateurs : Christophe Besse
Albert Cohen
Thierry Goudon
Claude Le Bris
Felix Otto
Denis Talay

Remerciements

Je voudrais saisir l'occasion offerte par la rédaction de ce manuscrit d'habilitation pour remercier les personnes qui m'ont accompagné au cours de ma formation de chercheur.

Tout d'abord, je voudrais remercier les rapporteurs, qui m'ont accordé du temps en lisant ce manuscrit. Les travaux de ces trois grands mathématiciens m'ont beaucoup apporté dans ma recherche. C'est un immense honneur que m'a fait George Papanicolaou en acceptant d'être rapporteur de cette habilitation : quel meilleur juge que l'un des pères du domaine, dont les articles fondateurs ont été une grande source d'inspiration ! C'est Jean-Christophe Mourrat qui m'a initié aux travaux de Stefano Olla et ouvert au monde des marches aléatoires en milieu aléatoire — qui a changé en profondeur ma façon d'aborder l'homogénéisation stochastique. Je suis heureux que Stefano Olla ait accepté d'être rapporteur. Bob Kohn m'a fait le plaisir d'être le troisième rapporteur de mon habilitation. En plus d'être un grand analyste inspiré par les applications, Bob Kohn est d'une rare disponibilité. Il ne laisse aucun courrier sans réponse, que ce soit pour une question de mathématiques ou une simple question administrative. Ce sont ces qualités qui font de chercheurs exceptionnels des moteurs de la communauté.

Merci également aux examinateurs, Christophe Besse, Albert Cohen et Denis Talay, de partager avec moi ce moment intense de la soutenance d'habilitation.

En thèse avec Claude Le Bris, j'ai appris à prendre les problèmes à bras le corps, à essayer d'aller vite et bien, à choisir les techniques mathématiques en fonction des problèmes et non l'inverse. La thèse fut une période pleine et intense, je remercie encore Claude d'avoir su me guider de loin avec bienveillance. J'ai ensuite eu la chance de faire un post-doctorat avec Felix Otto. Durant cette année à Bonn, j'ai appris la persévérance et à prendre mon temps ; digérer un problème jusqu'à comprendre comment l'aborder. Depuis ces trois dernières années j'essaie de garder l'équilibre entre savoir aller vite et savoir prendre mon temps ; à chaque problème la juste approche. Ma collaboration avec Felix Otto a bouleversé ma façon d'appréhender les mathématiques appliquées. Felix est brillant et généreux. Il est un catalyseur qui sait extraire l'idée juste d'un raisonnement faux et donner confiance à ses collaborateurs. J'ai la chance incroyable de travailler avec lui depuis quatre ans déjà, et de pouvoir compter sur ses conseils.

Je suis aussi très redevable à Inria, qui m'a donné les moyens de mener ma recherche, en particulier à Max Dauchet et à David Simplot-Ryl, et à Thierry Goudon et toute l'équipe SIMPAF (celle d'hier et celle d'aujourd'hui). Thierry a soutenu ma candidature et m'a accueilli chaleureusement dans son équipe il y a trois ans, je le remercie de m'avoir fait confiance.

Une des plus grandes joies de la recherche est de comprendre ensemble. Je remercie tous mes collaborateurs, avec qui j'ai pris et prends toujours beaucoup de plaisir à travailler, Roberto Alicandro, Marco Cicalese, Andrea Braides, Marco Barchiesi, Jean-François Coulombel, Miguel Fernandez, Jean-Frédéric Gerbeau, encore et toujours Marina Vidrascu, Patrick Le Tallec, François Lequeux, Mathew Penrose, Jean-Christophe Mourrat, Stefan Neukamm, Thierry Goudon, Stella Krell, Maya de Buhan, Zakaria Habibi, Anne-Claire Egloffe, Than Nhan Nguyen... Parler de sa recherche et de ses problèmes mathématiques à ses collègues permet souvent d'avancer. Merci à Guillaume Dujardin, Ingrid Violet, Mathias Rousset, Claire Chainais, Christophe Besse, Emmanuel Creusé, Caterina Calgaro, Stephan de Bièvre, Pauline Lafitte, Alexandre Mouton, Vianney Combet, Gabriel Rivière... de savoir écouter. Merci aussi à Sandrine Catillon pour son aide au quotidien.

Introduction

The issue of establishing the status of nonlinear elasticity theory for rubber with respect to the point of view of polymer physics is at the heart of this manuscript. Our aim is to develop mathematical methods to describe, understand, and solve this multiscale problem.

At the level of the polymer chains, rubber can be described as a network whose nodes represent the cross-links between the polymer chains. This network can be considered as the realization of some stochastic process. Given the free energy of the polymer network, we'd like to derive a continuum model as the characteristic length of the polymer chains vanishes. In mathematical terms, this process can be viewed as a hydrodynamic limit or as a discrete homogenization, depending on the nature of the free energy of the network. In view of the works by Treloar [86], by Flory [34], and by Rubinstein and Colby [79] on polymer physics, and in view of the stochastic nature of the network, stochastic discrete homogenization seems to be the right tool for the analysis. Hence, in order to complete our program we need to understand the stochastic homogenization of discrete systems. Two features make the analysis rich and challenging from a mathematical perspective: the randomness and the nonlinearity of the problem.

The achievement of this manuscript is twofold:

- a complete and sharp quantitative theory for the approximation of homogenized coefficients in stochastic homogenization of discrete linear elliptic equations;
- the first rigorous and global picture on the status of nonlinear elasticity theory with respect to polymer physics, which partially answers the question raised by Ball in his review paper [4] on open problems in elasticity.

Although the emphasis of this manuscript is put on discrete models for rubber, and more generally on the homogenization of discrete elliptic equations, we have also extended most of the results to the case of elliptic partial differential equations — some of the results being even more striking in that case.

Before we turn to the details of our contributions, let us first recall the context of stochastic homogenization.

The first rigorous results in homogenization of linear elliptic equations date back to the seventies with the contributions by De Giorgi and Spagnolo [28], Bensoussan, Lions, and Papanicolaou [8], Murat and Tartar [67, 68] to cite a few. Let A be a periodic matrix and for all $\varepsilon > 0$ let $A_\varepsilon(\cdot) := A(\cdot/\varepsilon)$ be the associated ε-rescaled matrix. Their results ensure that the solution operator $(-\nabla \cdot A_\varepsilon \nabla)^{-1}$ converges as $\varepsilon \to 0$ to the solution operator $(-\nabla \cdot A_{\mathrm{hom}} \nabla)^{-1}$ associated with a constant matrix A_{hom}. As a by-product of their analysis they obtain a characterization of A_{hom} in terms of the solution of an elliptic equation posed on the periodic cell (the so-called cell-problem). In this manuscript we are interested in cases when the corresponding cell-problem is not posed on the periodic cell (also called the unitary cell), but rather on the whole space. This happens in at least two cases:

- in the linear case when the periodicity assumption on A is replaced by the more general assumption of stochastic stationarity;

- in the periodic nonlinear nonconvex case, that is when the linear elliptic equation is replaced by the minimization of an energy functional $u \mapsto \int_D W(x/\varepsilon, \nabla u(x))dx$, where $u : D \to \mathbb{R}^d$ is a deformation, $\nabla u(x)$ is the strain gradient, and $W : \mathbb{R}^d \times \mathcal{M}^d \to \mathbb{R}^+$ is the energy density which is periodic in space and quasiconvex nonconvex in the deformation gradient (in the linear case, W is simply quadratic).

The rigorous derivation of nonlinear elasticity theory from polymer physics combines these two cases (in the setting of discrete equations). Stochastic homogenization of linear elliptic equations has been first studied by Papanicolaou and Varadhan [75] and by Kozlov [47]. The periodic homogenization of nonconvex integral functionals is due to Braides [13] and to Müller [64]. The stochastic homogenization of integral functionals is due to Dal Maso and Modica [26] and Messaoudi and Michaille [61].

This manuscript is organized in four chapters. The first chapter is dedicated to the quantitative analysis of stochastic homogenization of discrete linear elliptic equations. We give a complete error analysis for several approximation methods of the homogenized coefficients. The second chapter is a natural continuation of the first chapter: we generalize some of these results to the case of linear partial differential equations, and focus on the issue of approximating homogenized coefficients in various (general and applied) contexts. The third chapter deals with the stochastic homogenization of integral functionals. The aim of this chapter is to understand how qualitative properties of the heterogeneous integrand W (which are of interest in nonlinear elasticity) can be inherited (or not) by the homogenized integrand during the stochastic homogenization process. The fourth and last chapter of this manuscript is dedicated to a complete and rigorous derivation of nonlinear elasticity theory starting from a model based on polymer physics, in the context of stochastic homogenization of discrete systems.

Chapter 1: Stochastic homogenization of discrete linear elliptic equations

Discrete linear elliptic equations typically model the conduction of electricity in resistance networks (in $\mathbb{Z}^d$, each edge relating two points at distance 1 is typically a conductance). We assume that the conductances are independent and identically distributed (i. i. d.) in some bounded interval isolated from 0. They give rise to a diagonal random conductivity matrix A on $\mathbb{Z}^d$ and to a discrete elliptic operator using finite differences. This problem is a linear and scalar counterpart to the discrete model for rubber we shall present later on.

There are two points of view on this problem: the discrete linear elliptic equation associated with the random conductivity matrix A and the random walk in the random environment characterized by the conductances. Both points of view allow to prove a homogenization result. Within the discrete elliptic equation point of view, Künnemann [51] and Kozlov [48] have proved that the large scale behaviour of the (discrete) solution operator "$(-\nabla \cdot A\nabla)^{-1}$" is almost surely described by the large scale behaviour of the (continuous) solution operator $(-\nabla \cdot A_{\mathrm{hom}}\nabla)^{-1}$ for some deterministic matrix A_{hom}. In terms of random walk in random environments, Kipnis and Varadhan [45] have shown that the homogenization result takes the form of the convergence of the (rescaled) random walk to a Brownian motion of covariance $2A_{\text{hom}}$, in expectation (the almost sure convergence has been shown by Sidoravicius and Sznitman [82] and Mathieu [59]). Both points of view rely on the study of the corrector equation: for any $\xi \in \mathbb{R}^d$:

$$-\nabla \cdot A(\xi + \nabla\phi) \;=\; 0 \text{ in } \mathbb{Z}^d.$$

An equivalent form of this equation can be written in the probability space. The main difficulty of this equation comes from the lack of Poincaré's inequality in the probability space (said differently, with the notation $\langle\cdot\rangle$ for the expectation, there is no Poincaré's inequality of the type $\langle\phi^2\rangle \leq C\,\langle|\nabla\phi|^2\rangle$ for (stationary) functions ϕ such that $\langle\phi\rangle = 0$). This problem is the starting point of troubles in stochastic homogenization. One possibility to study the corrector equation is to introduce a regularization by a zero-order term, and consider the unique stationary solution to

$$T^{-1}\phi_T - \nabla \cdot A(\xi + \nabla\phi_T) \;=\; 0 \text{ in } \mathbb{Z}^d$$

for $T > 0$, and pass to the limit as $T \to \infty$. To be able to obtain quantitative results, we need to understand more thoroughly the corrector equation, or equivalently the dependence of ϕ_T with respect to T.

Our first crucial result is that the combination of a spectral gap estimate, a Caccioppoli inequality in probability, and elliptic regularity theory allows one to derive a "proxy" for the Poincaré inequality (at least in dimension $d > 2$). In particular we have shown that for $d > 2$,

$$\left\langle\phi_T^2\right\rangle \lesssim 1$$

uniformly in T. As a consequence, we obtain a complete existence and uniqueness theory for the corrector equation in the probability space. In other words: there does exist a unique stationary solution to the corrector equation in dimension $d > 2$. This result was rather unexpected. The price to pay for this is to consider i. i. d. conductances (and not only the much more general assumption of ergodicity used in [51], [48], and [45]). This is a first step towards a quantitative stochastic homogenization theory.

In a second contribution we devise new approximation formulas for homogenized coefficients based on the regularized corrector ϕ_T. This is of practical interest since ϕ_T can be accurately approximated on bounded domains whereas ϕ cannot (this is related to the exponential decay of the Green's function of the Helmholtz operator compared to the algebraic decay of the Green's function of the Laplace operator). Papanicolaou and Varadhan [75] and Kipnis and Varadhan [45] have given a spectral representation formula for the homogenized coefficients. We adopt this framework and first define natural approximation formulas $A_{k,T}$ (for $k \in \mathbb{N}$, $T > 0$) in terms of their spectral representations. We only later on re-interpret these formulas in terms of the regularized corrector ϕ_T in physical space. Similar approximations can be obtained directly by a Richardson extrapolation in the physical space. As we shall show, the fundamental property of these approximations is that the convergence rate to zero of $A_{k,T} - A_{\text{hom}}$ in function of T is driven by the values of some spectral exponents related to the spectral representation of the elliptic operator $-\nabla \cdot A\nabla$

in the probability space. Our second crucial result is a sharp estimate of these spectral exponents in any dimension, the proof of which is based on induction using functional calculus, elliptic theory, and the spectral gap estimate.

From a practical point of view one cannot solve the regularized corrector equation in the probability space to compute expectations: instead, we approximate the regularized corrector in physical space, and replace the expectation by spatial averages on large domains $Q_L = (0, L)^d$ — this is where the fact that ϕ_T can be accurately approximated on bounded domains is important. This gives rise to the approximation $A^L_{k,T}$ of $A_{k,T}$. Note that $A^L_{k,T}$ is itself a random variable. The error between the expectation $\langle A^L_{k,T} \rangle = A_{k,T}$ of the computable approximation of A_{hom} and A_{hom} itself has been estimated using the spectral exponents. Remains the statistical error: $A^L_{k,T}$ fluctuates around its expectation $A_{k,T}$. We have essentially proved that the variance of $A^L_{k,T}$ has the scaling of the central limit theorem: $\text{var}\left[A^L_{k,T}\right] \lesssim L^{-d}$ (up to a logarithmic correction for $d = 2$). Again, the proof of this estimate crucially relies on the spectral gap estimate.

The combination of these results then allows us to make a complete and optimal quantitative convergence analysis of several approximation methods of the homogenized coefficients. We prove in particular that the popular "periodization method" yields optimal convergence rates in any dimension.

In the last section of this chapter, we turn to the point of view of the random walk in random environments, and use the convergence of the random walk to a Brownian motion with covariance matrix $2A_{\text{hom}}$ to devise an approximation procedure. The approximation is based on a Monte-Carlo method and consists in computing independent realizations of trajectories of the random walk up to some final time $t > 0$ in independent environments. Using a quantitative version of the Kipnis-Varadhan theorem together with estimates of the spectral exponents, we prove that the error at time t between the expectation of our approximation and the homogenized coefficient is essentially of order $1/t$. We complete this picture by proving large deviation estimates. The results are sharp.

Chapter 2: Quantitative results in homogenization of linear elliptic equations

In this second chapter we generalize some results of Chapter 1 to the case of linear partial differential equations. This time, the crucial assumption on the random symmetric diffusion matrix A is that its correlation-length is bounded (that is, there exists some correlation length $C_L > 0$ such that the random fields $A(x)$ and $A(z)$ are independent if $|x-z| > C_L$). Under this assumption we are able to extend the spectral gap estimate from the discrete setting (in which case the conductances are only countably many) to the continuous setting (in which case the diffusion matrix is a measurable function, and therefore lives in an infinite-dimensional space). Perhaps the most striking result for a specialist in stochastic homogenization is the existence of a stationary solution to the corrector equation in dimension $d > 2$. From the PDE point of view, the extension of the results from the discrete to the continuous setting essentially relies on the De Giorgi-Nash-Moser theory.

Up to now, we have only focused on the approximation of homogenized coefficients. In the second section of this chapter we address a different question. Given a linear elliptic partial differential equation whose diffusion coefficient is perturbed by a stationary noise with correlation-length $\varepsilon > 0$, how does the statistics of the solution depend on ε ? Heuristic arguments give a precise guess on the dependence of some strong and weak norms of the fluctuation upon ε. We present a string of arguments which provides optimal results for small ellipticity ratio only (say when the noise is sufficiently small). This can also be seen as a first step towards a quantitative analysis of the convergence to the solution of the homogenized problem in stochastic homogenization of linear elliptic equations, in the spirit of the work by Yurinskiĭ [91] (who obtains an algebraic convergence rate in ε for $d > 2$) and by Caffarelli and Souganidis [18] (who treat the much more general case of fully nonlinear elliptic equations, but only get a logarithmic convergence in ε).

We come back to the approximation of homogenized coefficients in the third section, and go beyond the case of stationary diffusion coefficients with finite correlation-length. We propose a general approximation method of homogenized coefficients which yields optimal convergence rates both for the stochastic case with finite-correlation length and for the periodic case. This approach combines a filtering method and the approximation formulas devised using the spectral representation of the homogenized coefficients. We show that these approximations are consistent in the general ergodic case using spectral theory, and provide with numerical examples in the periodic and quasiperiodic case which exemplify the interest of the method.

The last section of this chapter is an independent work which illustrates:

- the interest of the homogenization theory for engineering problems;
- the fact that, even in the periodic case, to pass from theoretical to practical results is not always an easy task in homogenization.

The starting point is a coupled system of elliptic/parabolic equations modeling the transport of nuclear waste in a heterogeneous storage device. Due to the nonlinear coupling we consider, even if we start from purely periodic coefficients we end up with a cell problem for the parabolic equation which depends on the space variable through the flux associated with the homogenized elliptic equation. From a practical point of view this is a disaster: we have a priori as many (periodic) cell problems to solve as Gauss points in the domain. A very powerful tool to deal with such parametrized partial differential equations is the reduced basis method. In this section we show how to apply this method to the homogenization problem under consideration, and propose an efficient way to construct the associated linear systems using a fast Fourier transform. Numerical tests are very promising.

Chapter 3: Homogenization of integral functionals

The current existence theory in nonlinear elasticity [5] is based on the minimization of energy functionals of the form $u \mapsto \int_D W(x, \nabla u(x))dx$ rather than on the associated Euler-Lagrange partial differential equations. It is therefore not surprising that the homogenization of integral functionals is used to model composite materials in nonlinear elasticity. When dealing with such energy functionals, natural properties can be expected of the energy density, such as frame-invariance, isotropy, minimality at identity... In terms of convexity properties, if one expects the energy density $W(x, \Lambda)$ to blow up when the determinant of the $d \times d$ matrix Λ approaches zero (that is one cannot shrink the material to a point), $W(x, \cdot)$ cannot be a convex function. This blow up is however compatible with the more general notions of quasiconvexity and polyconvexity. Another subtle property is strong ellipticity (which is equivalent to the convexity of $W(x, \cdot)$ along rank-one connections). Stability of homogeneous deformations as well as short time existence results in elastodynamics require strict strong ellipticity. It is also sometimes convenient to restrict the analysis to the small deformation regime, for which one expects a linear theory to be valid at first order. Natural questions in homogenization of integral functionals are typically:

- under which conditions is the homogenized energy density frame-invariant ? isotropic ?
- is polyconvexity stable by homogenization (that is, if the heterogeneous energy density is polyconvex almost everywhere, is the associated homogenized energy density polyconvex) ? (Quasiconvexity is always stable.)
- is strong ellipticity stable by homogenization ?
- do homogenization and linearization commute in small deformation ?

The homogenized energy density is frame-invariant provided the heterogeneous integrand is frame-invariant almost everywhere. The homogenized energy density can never be isotropic in the periodic case for truly nonlinear and heterogeneous integrands. The other questions have also been solved in the case of *periodic homogenization*. Braides [14] has shown that polyconvexity is not stable by periodic homogenization (see also Barchiesi [6]). Geymonat, Müller, and Triantafyllidis [38] have developed a complete theory establishing conditions under which strict strong ellipticity is conserved or lost by periodic homogenization (both cases indeed occur). More recently, Müller and Neukamm [66] proved the commutability of periodic homogenization and linearization in small deformation (that is, at identity). The proofs of these qualitative results never rely on soft arguments, and often make a crucial use of periodicity. The proof by Braides uses the fundamental construction by Sverak [83]. The theory of [38] is based on the Bloch transform, and therefore on the periodic structure. The proof of the commutability in [66] combines the periodic asymptotic formula for the homogenized integrand with the deep quantitative rigidity estimate by Friesecke, James, and Müller [37].

A large part of the difficulties encountered in these problems originates in the fact that the cell problem is posed on $\mathbb{R}^d$ and not on the unitary cell. It would therefore be very valuable to have an alternative formula for the homogenized integrand depending on a periodic solution on the unitary cell. In view of the contribution [64] by Müller, a natural candidate for the homogenized integrand is the quasiconvex envelope of the cell-integrand (obtained using a periodic solution of the cell-problem on the unitary cell). Our first contribution (Section 3.3) shows however that this natural candidate cannot coincide with the homogenized integrand in general, so that it seems difficult to by-pass the asymptotic character of the homogenization formula.

We've seen so far that periodic homogenization of integral functionals is rather well-understood even if the asymptotic homogenization formula is a difficult object to handle. The picture is quite different for the stochastic homogenization of integral functionals. As a general principle, "homogenization structures" (that is, the assumptions on the heterogeneities which yield homogenization) are the same for linear elliptic equations and integral functionals, so that one expects that if one can prove homogenization for linear elliptic equations one should be able to prove homogenization for general integral functionals. Examples are of course periodicity, and stochastic stationarity. We illustrate this general principle on a homogenization structure introduced by Blanc, Le Bris, and Lions [11], which mixes periodicity and stochastic stationarity in such a way that the obtained structure is neither periodic nor stationary, although it yields homogenization for linear elliptic equations. We show in the first section of this chapter that this homogenization structure is "stationary up to some boundary effects", so that we may still apply the subadditive ergodic theorem following the approach by Dal Maso and Modica [26], and therefore homogenize integral functionals. Regarding the qualitative properties listed above, it is not clear any longer that what holds for periodic structures holds as well for general homogenization structures, and in particular for the stochastic stationary case. Conversely, some properties may hold in some stochastic cases, but not in the periodic case. Isotropy is such an example: the homogenized energy density is isotropic in stochastic homogenization provided the heterogeneous integrand is statistically isotropic. Although periodicity is a particular case of stationarity, the property of statistical isotropy is incompatible with periodicity.

Let us turn to the question of strong ellipticity. The theory by Geymonat, Müller and Triantafyllidis [38] relies on the Bloch transform, so that the periodicity assumption seems to be crucial. In Section 3.5 we show that the crucial assumption is not periodicity but rather stationarity, which allows to define a suitable version of the Bloch transform, and extend the analysis of strong ellipticity to the stochastic stationary case.

Likewise we shall also generalize the commutability result at identity to the case of stochastic homogenization, and more generally to any homogenization structure. Not only this analysis extends the result by Müller and Neukamm [66] but it also implies the weak locality of the Γ-closure at identity. Namely, any homogenized integrand obtained by some homogenization structure can also be obtained by periodic homogenization when restricted to the small deformation regime.

In this chapter we essentially extend all the known qualitative results of periodic homogenization of integral functionals to the case of stochastic homogenization of integral functionals. The following step is to address similar questions for the stochastic homogenization of discrete systems.

Chapter 4: Homogenization of discrete systems and derivation of rubber elasticity

We turn to the main objective of this manuscript: a rigorous derivation of nonlinear elasticity theory from "first principles" in the form of a statistical physics model for rubber. We shall start from polymer physics, derive by heuristic arguments a model for polymer-chain networks which is suitable for the analysis, and then proceed with the rigorous derivation. This chapter involves various areas of mathematics: calculus of variations, ergodic theory, stochastic geometry, Fourier analysis, approximation theory, inverse problems, and scientific computing. Even if all the questions have not been solved yet, we present a rather global picture of the problem.

The starting point is a heuristic derivation of a discrete model based on a full statistical physics description of a polymer-chain network. We shall argue that the free energy of a polymer-chain network under the constraint that the deformation of its boundary is fixed can be approximated by the minimum of an energy functional over a set of admissible deformations described by the positions of the end-to-end vectors of each polymer chain inside the network. The energy functional splits into two terms: the free energies of the polymer chains $\mathbb{R}^+ \ni r \mapsto f(r)$ as if they were isolated, and a volumetric term between polymer chains $\Lambda \mapsto W_{\mathrm{vol}}(\Lambda)$, which would ideally ensure incompressibility.

Under the assumption that the network is the realization of some ergodic stochastic lattice (and provided some control on the growth of f and W_{vol}) we prove in the second section of this chapter that the energy functional Γ-converges, as the characteristic length of the polymer chains vanishes, to some integral functional on some Lebesgue space. This is a counterpart for discrete systems of the result by Dal Maso and Modica [26] for integral functionals. Note however that the randomness is on the network istelf, not on the interactions, contrary to the case of the linear elliptic equations of Chapter 1. The associated homogenized energy density W_{hom} is proved to be homogeneous in space and deterministic (as a consequence of ergodicity). It is given by an asymptotic formula on $\mathbb{R}^d$. The homogenized integrand is frame-invariant, and it is isotropic provided the stochastic lattice is statistically isotropic. Then, if W_{hom} is isotropic, it admits a dilation among its natural states, so that the identity is a natural state after rescaling the reference configuration. This homogenized energy density thus yields a suitable hyperelastic model for rubber.

A question left aside in Section 4.2 is whether statistically isotropic stationary ergodic lattices do exist. This is indeed the case, and we show that the random parking measure (cars are randomly parked in some parking lot or bounded domain, with the constraint that they cannot overlap ; the process ends when no further car can be added) studied by Penrose [76] yields such an isotropic stationary ergodic point set at the thermodynamic limit (that is, when the parking lot tends to the whole space). We also prove that the random parking lattice at the thermodynamic limit can be replaced by the random parking lattice on large (yet bounded) cubes in the asymptotic formula for the homogenized energy density. This ensures the convergence of numerical approximations for which the random parking lattice has to be approximated itself as well.

We turn to numerical approximations in Section 4.4. Unlike the case of the linear discrete elliptic equations of Chapter 1, we first need to approximate the random point set in order to approximate the corrector (and therefore W_{hom}). As pointed out already, the random parking measure on $\mathbb{R}^d$ can be replaced by its approximation on bounded domains. The corrector is then approximated on a bounded domain by solving a minimization problem on a finite-dimensional space (as typical in finite element methods for nonlinear elasticity). The theory of Chapter 1, although it does not directly apply here, gives a hint on the approximation error (statistical fluctuation, systematic error). Numerical approximations allow us to directly compare the homogenized integrand to the mechanical experiments by Treloar on rubber. Numerical results confirm the relevance of the discrete model for rubber and its "thermodynamic" limit to reproduce correctly complex and nonlinear behaviors with a few parameters to fit (unlike phenomenological constitutive laws, only physical parameters appear in the discrete model).

Numerical tests tend to show that the homogenized energy density (or at least its numerical approximation) is strictly strongly elliptic. In view of the work by Geymonat, Müller, and Triantafyllidis [38], it is however not clear a priori whether the homogenized energy density W_{hom} should be strictly strongly elliptic or not. In Section 4.5 we give two examples of discrete homogenization of periodic systems which illustrate both strict strong ellipticity and loss of strong ellipticity of the homogenized integrand. The example which yields a strictly strongly elliptic homogenized integrand is quite interesting from the point of view of polymer physics since the property of the discrete model which ensures the strong ellipticity is satisfied by the free energy f of polymer chains derived in statistical physics. In particular, this function f is a convex and increasing function of the distance r between the end-to-end points of the chain (whereas, and it is useful as well, W_{vol} is assumed to be polyconvex). This is enough to ensure that the periodic discrete model satisfies the Cauchy-Born rule (or affine assumption), so that the homogenized energy density is explicit and can be checked to be strictly strongly elliptic. When the stochastic lattice is not periodic, the Cauchy-Born rule does not hold any longer and no explicit formula exists for the homogenized energy density. However one may still exploit the specific form of the free energy of polymer chains and prove a perturbation result. Namely, the homogenized energy density is strictly strongly elliptic provided the deformation of the polymer network is close to an affine deformation (uniformly in the asymptotic formula). Although even in the scalar linear case this assumption is not known to hold (in Chapter 1 we prove that all the finite moments of the corrector are finite for $d > 2$, but not that the corrector is essentially bounded), this seems to be the case in numerical experiments. This study gives a partial answer to the question of strong ellipticity.

From a conceptual point of view one could be satisfied with this study of the discrete model for rubber: the derivation is rigorous (at least under mild assumptions), the homogenized energy density can be computed numerically, and the results are in good agreement with mechanical experiments. Yet there are two disturbing facts:

- In computational and experimental mechanics, people use rather specific constitutive laws (Mooney-Rivlin, Ciarlet-Geymonat, Ogden...). What is then the link between W_{hom} and these models ?
- If this discrete model were to be used in practice (say as the energy density in a finite element software for nonlinear elasticity) one would face the problem that at each Gauss point of the domain where the Piola stress tensor $\partial_{\Lambda} W_{\mathrm{hom}}(\Lambda)$ has to be evaluated, one

needs to solve a nonlinear elasticity problem (to approximate the asymptotic formula). This is simply not feasible.

Our last contribution to the study of this discrete model for rubber addresses both observations at once. The objective of Section 4.6 is to construct an analytical proxy which fits the homogenized energy density and can be used in standard softwares. Following the know-how of mechanical engineers — and guided by our theoretical analysis — we have chosen to look for an approximation in the class of polyconvex isotropic Ogden materials whose natural state is the identity. From a practical point of view we then have to solve an inverse problem: given a sampling of the (numerical approximation of the) homogenized constitutive law W_{hom}, identify parameters of an Ogden law W_{og} that minimizes some error functional $\mathcal{E}(W_{\mathrm{hom}}, W_{\mathrm{og}})$. Since the energy landscape of the error functional is very complex, deterministic optimization algorithms (such as a Newton algorithm) are likely to get trapped in some local minimum, and we have prefered to use an evolutionary algorithm combined with a splitting procedure to deal with the constraints. The numerical results are very convincing, and show the capability of Ogden laws to accurately approximate the homogenized integrand (which has been itself rigorously derived from the polymer chain model by discrete stochastic homogenization).

The last chapter of this manuscript introduces a complete framework to pass from polymer physics to standard constitutive laws for rubber. Several steps of this program are not rigorous yet and give rise to challenging mathematical problems.

Personal bibliography since my PhD thesis

Qualitative results for the homogenization of multiple integrals

[BG] M. Barchiesi and A. Gloria. New counterexamples to the cell formula in nonconvex homogenization. *Arch. Ration. Mech. Anal.*, 195(3), pp 991-1024, 2010.

[Glo08] A. Gloria. Stochastic diffeomorphisms and homogenization of multiple integrals. *AMRX*, Article ID abn001, 2008.

[Glo11a] A. Gloria. Stochastic homogenization of nonlinear elastic materials, Bloch waves, and strong ellipticity. In preparation.

[GN] A. Gloria and S. Neukamm. Commutability of homogenization and linearization at identity in finite elasticity and applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(6), pp 941-964, 2011.

Qualitative results in homogenization of discrete systems

[ACG] R. Alicandro, M. Cicalese, and A. Gloria. Integral representation results for energies defined on stochastic lattices and application to nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 200(3), pp 881-943, 2011.

[BGLTV] M. de Buhan, A. Gloria, P. Le Tallec, and M. Vidrascu. Reconstruction of a constitutive law for rubber from in silico experiments. In preparation.

[Glo11b] A. Gloria. Strong ellipticity of nonlinear elastic materials and homogenization of periodic and stochastic discrete systems. In preparation.

[GLTV] A. Gloria, P. Le Tallec, and M. Vidrascu. Foundation, analysis, and numerical investigation of a variational network-based model for rubber. Accepted for publication in *Cont. Mech. Thermodynamics*.

[GP] A. Gloria and M. D. Penrose. Random parking, Euclidean functionals, and rubber elasticity. Accepted for publication in *Commun. Math. Physics*.

Quantitative results in stochastic homogenization of discrete elliptic equations

[EGMN] A.-C. Egloffe, A. Gloria., J.-C. Mourrat, and T. N. Nguyen. Theoretical and numerical investigation of convergence rates in stochastic homogenization: corrector equation and random walk in random environments. Preprint.

[Glo12] A. Gloria. Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *M2AN Modél. Math. Anal. Numér*, 46(1), pp 1-38, 2012.

[GMa] A. Gloria and J.-C. Mourrat. Spectral measure and approximation of homogenized coefficients. *Probab. Theory Related Fields*, DOI: 10.1007/s00440-011-0370-7.

[GMb] A. Gloria and J.-C. Mourrat. Quantitative version of the Kipnis-Varadhan theorem and Monte-Carlo approximation of homogenized coefficients. Accepted for publication in *Ann. Appl. Probab.*

[GNOa] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. In preparation.

[GNOb] A. Gloria, S. Neukamm, and F. Otto. Approximation of effective coefficients by periodization in stochastic homogenization. In preparation.

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

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

Quantitative results in stochastic homogenization of linear elliptic PDEs

[Glo11c] A. Gloria. Fluctuation of solutions to linear elliptic equations with noisy diffusion coefficients. Accepted for publication in *Commun. PDE*.

[GOc] A. Gloria and F. Otto. Quantitative estimates in stochastic homogenization of linear elliptic equations. In preparation.

Numerical issues in homogenization

[Glo11d] A. Gloria. Reduction of the resonance error. Part 1: Approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.*, 21(8), pp 1601-1630, 2011.

[Glo11e] A. Gloria. Numerical homogenization: survey, new results, and perspectives. Lecture notes of the CEA-EDF-INRIA winter school "Periodic and stochastic homogenization: theoretical and computational aspects". Esaim Proceedings, 2012.

[GGK] A. Gloria, T. Goudon, and S. Krell. Numerical homogenization of a nonlinearly coupled elliptic-parabolic system, and application to nuclear waste storage. Preprint.

Miscellaneous

[CG] J.-F. Coulombel and A. Gloria. Semigroup stability of finite difference schemes for multidimensional hyperbolic initial boundary value problems. *Math. Comp.*, 80, pp 165-203, 2011.

Notation

- $\mathbb{R}^+ = [0, +\infty)$;
- $d, n \geq 1$ are dimensions;
- $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ is the canonical basis of $\mathbb{R}^d$;
- D is a bounded open Lipschitz domain of $\mathbb{R}^d$;
- $C_0^\infty(D)$ is the set of smooth functions with support in D;
- for all $p \in [1, +\infty]$, $L^p(D)$ is the Lebesgue space of p-integrable functions on D;
- for all $p \in [1, +\infty]$, $W^{1,p}(D)$ is the Sobolev space of p-integrable functions on D whose distributional derivatives are p-integrable functions;
- for all $p \in [1, +\infty]$, $W_0^{1,p}(D)$ is the subset of $W^{1,p}(D)$ whose trace vanishes on the boundary;
- $Q = (0, 1)^d$ is a unitary cell (sometimes we'll take $Q = (-1, 1)^d$), for all $N \in \mathbb{N}$, $Q_N = (0, N)^d$, and for all $R \in \mathbb{R}^+$, $Q_R = (0, R)^d$;
- for all $N \in \mathbb{N}$ and $p \in [1, +\infty)$, $W_\#^{1,p}(Q_N)$ is the closure in $W^{1,p}(Q_N)$ of smooth functions of $\mathbb{R}^d$ which are Q_N-periodic (for $p = +\infty$, it is the set of Q_N-periodic Lipschitz functions);
- $\mathcal{M}^{d \times n}$ is the set of $d \times n$ real matrices, which we endow with the operator norm $|\cdot|$;
- $\mathcal{M}^d$ is the set of $d \times d$ real matrices;
- for all $\Lambda \in \mathcal{M}^d$, we set $\varphi_\Lambda : \mathbb{R}^d \to \mathbb{R}^d, x \mapsto \Lambda x$;
- $\mathcal{M}_+^d$ is the set of $d \times d$ real matrices with positive determinant;
- SO_d is the set of rotations of $\mathbb{R}^d$;
- d_∞ is the distance of the supremum in $\mathbb{R}^d$;
- for any Borel subset D of $\mathbb{R}^d$, $|D|$ denotes its Lebesgue measure;
- for any Borel subset D of $\mathbb{R}^d$, $\fint_D$ denotes $\frac{1}{|D|} \int_D$;
- $\langle \cdot \rangle$ is the ensemble average, or equivalently the expectation in the underlying probability space;
- $\mathrm{var}\,[\cdot]$ is the variance associated with the ensemble average;
- $\mathrm{cov}\,[\cdot; \cdot]$ is the covariance associated with the ensemble average;
- $\lesssim$ and $\gtrsim$ stand for $\leq$ and $\geq$ up to a multiplicative constant which only depends on the dimension d and the ellipticity constants α, β (made precise throughout the text) if not otherwise stated;
- when both $\lesssim$ and $\gtrsim$ hold, we simply write $\sim$;
- we use $\gg$ instead of $\gtrsim$ when the multiplicative constant is (much) larger than 1;

Contents

1

Stochastic homogenization of discrete linear elliptic equations

1.1 Corrector equation and random walk in random environment

This section is an informal introduction to two different points of view on the stochastic homogenization of resistance networks:

- discrete elliptic operator and the corrector equation (PDE),
- random walk in random environment (RWRE).

To be short, the homogenization limit of a discrete linear elliptic equation yields a linear elliptic partial differential equation with constant coefficients A_{hom}, whereas the random walk in random environment leads after rescaling to a Brownian motion with covariance matrix $2A_{\text{hom}}$. The link between these two points of view is similar in spirit to the interpretation of the heat equation by a Brownian motion.

We shall start with the description of the network, then turn to the discrete elliptic point of view, and conclude with the random walk in random environment viewpoint. The aim of this section is to introduce a formalism, and give an intuition on both points of view. This is not a rigorous introduction to stochastic homogenization.

Unlike the other three chapters, we do not only recall here standard results, but we also display the main arguments of their proofs (essentially due to Papanicolaou and Varadhan [75], Kozlov [47], and Kipnis and Varadhan [45]). These are important facts for the understanding of the rest of the chapter.

We present the results in the case of independent and identically distributed conductivities although everything remains valid (in this section) provided the conductivities lie in a compact set of $(0, +\infty)$, are stationary, and ergodic.

1.1.1 Random environment

We say that x, y in $\mathbb{Z}^d$ are neighbors, and write $x \sim y$, whenever $|y - x| = 1$. This relation turns $\mathbb{Z}^d$ into a graph, whose set of (non-oriented) edges is denoted by $\mathbb{B}$. Let us define the associated diffusion coefficients and their statistics.

Definition 1 (environment) *Let $\Omega = [\alpha, \beta]^{\mathbb{B}}$. An element $\omega = (\omega_e)_{e\in\mathbb{B}}$ of Ω is called an* environment. *With any edge $e = (x, y) \in \mathbb{B}$, we associate the* conductance $\omega_{(x,y)} := \omega_e$ *(by construction $\omega_{(x,y)} = \omega_{(y,x)}$). Let ν be a probability measure on $[\alpha, \beta]$. We endow Ω with the product probability measure $\mathbb{P} = \nu^{\otimes\mathbb{B}}$. In other words, if ω is distributed according to the measure $\mathbb{P}$, then $(\omega_e)_{e\in\mathbb{B}}$ are independent random variables of law ν. We denote by $L^2(\Omega)$ the set of real square integrable functions on Ω for the measure $\mathbb{P}$, and write $\langle\cdot\rangle$ for the expectation associated with $\mathbb{P}$.*

We may introduce a notion of stationarity.

Definition 2 (stationarity) *For all $z \in \mathbb{Z}^d$, we let $\theta_z : \Omega \to \Omega$ be such that for all $\omega \in \Omega$ and $(x, y) \in \mathbb{B}$, $(\theta_z\ \omega)_{(x,y)} = \omega_{(x+z,y+z)}$. This defines an additive action group $\{\theta_z\}_{z\in\mathbb{Z}^d}$ on Ω which preserves the measure $\mathbb{P}$, and is ergodic for $\mathbb{P}$.*

We say that a function $f : \Omega \times \mathbb{Z}^d \to \mathbb{R}$ is stationary if and only if for all $x, z \in \mathbb{Z}^d$ and $\mathbb{P}$-almost every $\omega \in \Omega$,

$$f(x + z, \omega) = f(x, \theta_z\ \omega).$$

In particular, with all $f \in L^2(\Omega)$, one may associate the stationary function (still denoted by f) $\mathbb{Z}^d \times \Omega \to \mathbb{R}, (x, \omega) \mapsto f(\theta_x\ \omega)$. In what follows we will not distinguish between $f \in L^2(\Omega)$ and its stationary extension on $\mathbb{Z}^d \times \Omega$.

1.1.2 Corrector equation

We associate with the conductivities on $\mathbb{B}$ a conductivity matrix on $\mathbb{Z}^d$.

Definition 3 (conductivity matrix) *Let Ω, $\mathbb{P}$, and $\{\theta_z\}_{z\in\mathbb{Z}^d}$ be as in Definitions 1 and 2. The stationary diffusion matrix $A : \mathbb{Z}^d \times \Omega \to \mathcal{M}_d(\mathbb{R})$ is defined by*

$$A(x, \omega) = \operatorname{diag}(\omega_{(x,x+\mathbf{e}_1)}, \dots, \omega_{(x,x+\mathbf{e}_d)}).$$

For each $\omega \in \Omega$, we consider the discrete elliptic equation whose operator is

$$L = -\nabla^* \cdot A(\cdot, \omega)\nabla, \tag{1.1}$$

where ∇ and ∇^* are defined for all $u : \mathbb{Z}^d \to \mathbb{R}$ by

$$\nabla u(x) := \begin{bmatrix} u(x + \mathbf{e}_1) - u(x) \\ \vdots \\ u(x + \mathbf{e}_d) - u(x) \end{bmatrix}, \ \nabla^* u(x) := \begin{bmatrix} u(x) - u(x - \mathbf{e}_1) \\ \vdots \\ u(x) - u(x - \mathbf{e}_d) \end{bmatrix}, \tag{1.2}$$

and the backward divergence is denoted by $\nabla^*\cdot$, as usual. In particular, for all $u : \mathbb{Z}^d \to \mathbb{R}$,

$$Lu : \mathbb{Z}^d \to \mathbb{R}, z \mapsto \sum_{z'\sim z} \omega_{(z,z')}(u(z) - u(z')). \tag{1.3}$$

The standard stochastic homogenization theory for such discrete elliptic operators (see for instance [51], [48]) ensures that there exist homogeneous and deterministic coefficients

A_{hom} such that the solution operator of the continuum differential operator $-\nabla\cdot A_{\mathrm{hom}}\nabla$ describes $\mathbb{P}$-almost surely the large scale behavior of the solution operator of the discrete differential operator $-\nabla^*\cdot A(\cdot,\omega)\nabla$. As for the periodic case, the definition of A_{hom} involves the so-called correctors $\phi:\mathbb{Z}^d\times\Omega\to\mathbb{R}$, which are solutions (in a sense made precise below) to the equations

$$-\nabla^*\cdot A(x,\omega)(\xi+\nabla\phi(x,\omega)) \,=\, 0, \qquad x\in\mathbb{Z}^d, \tag{1.4}$$

for all $\xi\in\mathbb{R}^d$. The following lemma gives the existence and uniqueness of the corrector ϕ.

Lemma 1.1 (corrector). *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Then, for all $\xi\in\mathbb{R}^d$, there exists a unique measurable function $\phi:\mathbb{Z}^d\times\Omega\to\mathbb{R}$ such that $\phi(0,\cdot)\equiv 0$, $\nabla\phi$ is stationary, $\langle\nabla\phi\rangle=0$, and ϕ solves* (1.4) *$\mathbb{P}$-almost surely. Moreover, the symmetric homogenized matrix A_{hom} is characterized by*

$$\xi\cdot A_{\mathrm{hom}}\xi \,=\, \langle(\xi+\nabla\phi)\cdot A(\xi+\nabla\phi)\rangle\,. \tag{1.5}$$

The standard proof of Lemma 1.1 makes use of the regularization of (1.4) by a zero-order term $\mu>0$:

$$\mu\phi_\mu(x,\omega)-\nabla^*\cdot A(x,\omega)(\xi+\nabla\phi_\mu(x,\omega)) \,=\, 0, \qquad x\in\mathbb{Z}^d. \tag{1.6}$$

Lemma 1.2 (regularized corrector). *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Then, for all $\mu>0$ and $\xi\in\mathbb{R}^d$, there exists a unique stationary function $\phi_\mu\in L^2(\Omega)$ which solves* (1.6) *$\mathbb{P}$-almost surely.*

***Remark* 1** *Depending on the type of results under consideration we use two different notations for the regularization, namely $\mu>0$ which is meant to be small (and ϕ_μ for the regularized corrector), but also $\mu=T^{-1}$ with T meant to be large (and the slight abuse of notation ϕ_T for the associated regularized corrector).*

There are (at least) two ways to prove Lemma 1.2. The first one (which is maybe more intuitive for a non-probabilist) is to use the Lax-Milgram theorem in the space $\mathcal{H}=\{\psi:\mathbb{Z}^d\times\Omega\to\mathbb{R};\,\psi$ is stationary$\}$, which is a Hilbert space when endowed with the norm

$$\|u\|^2_{\mathcal{H}} \,=\, \lim_{R\to\infty}\int_{\mathbb{Z}^d\cap Q_R}u(x)^2dx,$$

which, by ergodicity, can be rewritten as

$$\|u\|^2_{\mathcal{H}} \,=\, \left\langle u(0)^2\right\rangle,$$

the crucial ingredient being stationarity.

However it is more natural (and a posteriori better) to directly work in $L^2(\Omega)$. Following [75], we introduce difference operators on $L^2(\Omega)$: for all $u\in L^2(\Omega)$, we set

$$\mathrm{D}\,u(\omega) := \begin{bmatrix} u(\theta_{\mathbf{e}_1}\omega)-u(\omega)\\ \vdots\\ u(\theta_{\mathbf{e}_d}\omega)-u(\omega)\end{bmatrix},\ \mathrm{D}^*\,u(\omega) := \begin{bmatrix} u(\omega)-u(\theta_{-\mathbf{e}_1}\omega)\\ \vdots\\ u(\omega)-u(\theta_{-\mathbf{e}_d}\omega)\end{bmatrix}. \tag{1.7}$$

These operators may not be that intuitive, but they play the same roles as the finite differences ∇ and ∇^* — this time for the variable ω.

This allows to define a stochastic counterpart to the operator L defined in (1.1):

Definition 4 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. We define $\mathcal{L}: L^2(\Omega)\to L^2(\Omega)$ by*

$$\begin{aligned}\mathcal{L}u(\omega) &= -\,\mathrm{D}^*\cdot A(\omega)\,\mathrm{D}\,u(\omega)\\ &= \sum_{z\sim 0}\omega_{0,z}(u(\omega)-u(\theta_z\,\omega))\end{aligned}$$

where D *and* D^* *are as in* (1.7).

The fundamental relation between L and $\mathcal{L}$ is the following identity for stationary fields $u:\mathbb{Z}^d\times\Omega\to\mathbb{R}$: for all $z\in\mathbb{Z}^d$ and almost every $\omega\in\Omega$,

$$Lu(z,\omega) \;=\; \mathcal{L}u(\theta_z\omega).$$

In particular, the regularized corrector ϕ_μ is also the unique solution in $L^2(\Omega)$ to the equation

$$\mu\phi_\mu(\omega)-\mathrm{D}^*\cdot A(\omega)(\xi+\mathrm{D}\,\phi_\mu(\omega)) \;=\; 0,\quad \omega\in\Omega.$$

The weak form of this equation reads: find $\phi_\mu\in L^2(\Omega)$ such that for all $\psi\in L^2(\Omega)$,

$$\langle \mu\phi_\mu\psi+\mathrm{D}\,\psi\cdot A(\xi+\mathrm{D}\,\phi_\mu)\rangle \;=\; 0.$$

The second proof of Lemma 1.2 simply relies on the Lax-Milgram theorem in $L^2(\Omega)$.

To obtain Lemma 1.1 from Lemma 1.2, the starting point is the following bounds

$$\big\langle|\nabla\phi_\mu|^2\big\rangle \;=\; \big\langle|\,\mathrm{D}\,\phi_\mu|^2\big\rangle \;\le\; C,\qquad \big\langle\phi_\mu^2\big\rangle \;\le\; C\mu^{-1},$$

for some C independent of μ. This allows to pass to the limit in the weak formulations and obtain the existence of a field $\Phi=(\phi_1,\dots,\phi_d)\in L^2(\Omega,\mathbb{R}^d)$ such that for all $\psi\in L^2(\Omega)$,

$$\langle \mathrm{D}\,\psi\cdot A(\xi+\Phi)\rangle \;=\; 0.$$

Using the following weak Schwarz commutation rule

$$\forall j,k\in\{1,\dots,d\},\quad \langle(\mathrm{D}_j\,\psi)\Phi_k\rangle \;=\; \langle(\mathrm{D}_k\,\psi)\Phi_j\rangle$$

one may define $\phi:\mathbb{Z}^d\times\Omega\to\mathbb{R}$ such that $\nabla\phi$ is stationary, $\Phi=\nabla\phi$, and $\phi(0,\omega)=0$ for almost every $\omega\in\Omega$. By definition this function ϕ is not stationary. It is a priori not clear (and even wrong in dimension $d=1$) whether there exists some function $\psi\in L^2(\Omega)$ such that $\mathrm{D}\,\psi=\Phi$ (this is a major difference with the periodic case).

Another delicate question is the uniqueness of Φ. There are again two proofs of this fact: the first one exploits the observation that ϕ_μ is sublinear at infinity on $\mathbb{Z}^d$, and the second one uses spectral theory in $L^2(\Omega)$. We sketch the second approach. By the spectral decomposition of the unitary group $\{\theta_z\}_{z\in\mathbb{Z}^d}$, there exist spectral projections $U(d\lambda)$ such that for all $z \in \mathbb{Z}^d$,

$$\theta_z \;=\; \int_{[-\pi,\pi)^d} e^{iz\cdot\lambda} U(d\lambda).$$

Let $\Psi \in L^2(\Omega, \mathbb{R}^d)$ be such that $\langle \Psi \rangle = 0$ and such that for all $\psi \in L^2(\Omega)$,

$$\langle \mathrm{D}\,\psi \cdot A\Psi \rangle \;=\; 0, \tag{1.8}$$

and that satisfies the weak Schwarz commutation rule. Proving that $\Psi = 0$ will yield the desired uniqueness of Φ. We'd like to test the weak formulation (1.8) with Ψ instead of $\mathrm{D}\,\psi$. To this aim we construct an approximate Helmholtz projection of Ψ by setting for all $\gamma > 0$,

$$\tilde{\psi}_\gamma(\omega) \;=\; \int_{[-\pi,\pi)^d} \sum_{j=1}^{d} \frac{e^{-i\lambda_j} - 1 - \gamma}{|e^{i\lambda} - 1 - \gamma|^2} U(d\lambda)\Psi_j(\omega),$$

where $|e^{i\lambda} - 1 - \gamma|^2 := \sum_{k=1}^d (e^{-i\lambda_k} - 1 - \gamma)(e^{i\lambda_k} - 1 - \gamma)$. This is an approximate Helmholtz projection in the sense that it satisfies for all $j \in \{1, \ldots, d\}$,

$$(\mathrm{D}_j - \gamma)\tilde{\psi}_\gamma \;=\; \Psi_j,$$

as can be proved using the weak Schwarz commutation rule. We then have by spectral calculus

$$\left\langle |\gamma\tilde{\psi}_\gamma|^2 \right\rangle = \int_{[-\pi,\pi)^d} \gamma^2 \sum_{k,l=1}^{d} \frac{(e^{-i\lambda_k} - 1 - \gamma)(e^{i\lambda_l} - 1 - \gamma)}{|e^{-i\lambda} - 1 - \gamma|^4} \left\langle U(d\lambda)\Psi_j\Psi_l \right\rangle .$$

Hence, by the Lebesgue dominated convergence theorem,

$$\lim_{\gamma\to 0} \left\langle |\gamma\tilde{\psi}_\gamma|^2 \right\rangle = \left\langle U(\{0\})\Psi_j\Psi_l \right\rangle .$$

But $U(\{0\})$ is the projection operator onto the functions invariant by $\{\theta_z\}_{z\in\mathbb{Z}^d}$, which by ergodicity are the constant functions. This implies that this limit vanishes since $\langle \Psi \rangle = 0$. Therefore, testing the equation (1.8) with function $\tilde{\psi}_\gamma$ yields at the limit $\gamma \to 0$,

$$\alpha \left\langle |\Psi|^2 \right\rangle \;\le\; \langle \Psi \cdot A\Psi \rangle \;=\; 0,$$

as desired.

We have presented this proof in detail since spectral theory will play a crucial role throughout this chapter. Note that this argument does not use the symmetry of A. It is also worth noticing that we could have obtained the existence and estimates on the approximate Helmholtz projection by PDE arguments as well.

1.1.3 Random walk in random environment

In this subsection we adopt the point of view of the random walk in the random environment ω. This presentation is primarily meant to non-probabilists.

Random walk in continuous time

We start with a walk in continuous time. Let the environment ω be fixed for a while (that is, we've picked a realization of the conductivities $\omega_e \in [\alpha, \beta]$ on the edges $e \in \mathbb{B}$). We then consider a random process (say in another probability space) associated with a random walker on the graph $\mathbb{B}$ which starts its walk at the origin $z = 0 \in \mathbb{Z}^d$. We denote by $\{X_t\}_{t\in\mathbb{R}^+}$ its trajectory. Let the random walker be at some site $z \in \mathbb{Z}^d$ at time t, and set $p_\omega(z) = \sum_{z'\sim z} \omega_{(z,z')}$. The random walker moves when the clock rings, and jumps to one (of the $2d$) neighboring point z' (with $|z - z'| = 1$) with probability

$$p(z \rightsquigarrow z') = \frac{\omega_{(z,z')}}{p_\omega(z)}.$$

The time $T(z)$ after which the clock rings at site z follows an exponential law of parameter $p_\omega(z)$. This means that for all $s > 0$, the probability that $T(z) > s$ is equal to $\exp(-p_\omega(z)s)$. This choice of the clock makes the random walk a Markov process, since the probability that $T(z) > s_1 + s_2$ knowing that $T(z) > s_1$ is $\exp(-p_\omega(z)s_2) = \exp(-p_\omega(z)(s_1 + s_2))/\exp(-p_\omega(z)s_1)$.

The link between this random walk and the elliptic operators of the previous subsection is as follows. We introduce a semi-group $\{P_t\}_{t\geq 0}$ associated with the random walk, that is for all $f : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded, we set for all $t \geq 0$

$$P_t f(z) = \mathbf{E}_z^\omega[f(X_t)],$$

where $\mathbf{E}_z^\omega$ means the expectation on the random walk starting at site $z \in \mathbb{Z}^d$ (and not at the origin) in the environment ω. We also denote by $\mathbf{P}_z^\omega$ the associated probability measure. This is a semi-group since X_t has the Markov property. The infinitesimal generator of this semi-group coincides with the elliptic operator $-L$ (where L is defined in (1.1)), as we show below. We need to compute

$$\lim_{t\to 0} \left(\frac{P_t f - f}{t} \right).$$

Let t be small. The probability that the clock rings at least once in $[0, t]$ is of order $p_\omega(z)t + O(t^2)$, the probability that it rings twice or more is of order $O(t^2)$, so that the probability that is does not ring if of order $1 - p_\omega(z)t + O(t^2)$. Hence,

$$\begin{aligned} P_t f(z) &= \mathbf{E}_z^\omega[f(X_t)] \\ &= (1 - p_\omega(z)t) f(z) + p_\omega(z)t \left(\sum_{z'\sim z} \frac{\omega_{(z,z')}}{p_\omega(z)} f(z') \right) + O(t^2) \qquad (1.9) \\ &= f(z) + t \sum_{z'\sim z} \omega_{(z,z')}(f(z') - f(z)) + O(t^2), \qquad (1.10) \end{aligned}$$

so that by (1.3),

$$\lim_{t\to 0}\left(\frac{P_t f(z)-f(z)}{t}\right) \;=\; \sum_{z'\sim z}\omega_{(z,z')}(f(z')-f(z)) \;=\; -Lf(z).$$

An important feature of this random walk is its reversibility for the counting measure (that is the measure which weights 1 at each site), which follows from the fact that the jump rates are symmetric: indeed, as can be seen on (1.9), the probability to go from z to z' within a time t is of order $\omega_{(z,z')}t + O(t^2)$, which is symmetric in z and z'. This implies that the measure

$$\overline{\mathbb{P}} = \mathbb{P}\mathbf{P}_0^\omega \tag{1.11}$$

is reversible for the Markov process $t \mapsto \omega(t) := \theta_{X(t)}\omega$ (and therefore the so-called environment viewed by the particle $t \mapsto \omega(t)$ is stationary for $\overline{\mathbb{P}}$). This property is equivalent to the self-adjointness for $\overline{\mathbb{P}}$ of the semi-group associated with $t \mapsto \omega(t)$, which is proved using the translation invariance of $\mathbb{P}$. The reversibility is crucial for the Kipnis-Varadhan argument (see below).

Random walk in discrete time

For the random walk in discrete time, the random walker jumps at every time $t \in \mathbb{N}$. If the random walker Y is at site $z \in \mathbb{Z}^d$ at time t, then the probability that it jumps to site $z' \sim z$ at time $t+1$ is given by

$$p(z \rightsquigarrow z') \;=\; \frac{\omega_{(z,z')}}{p_\omega(z)}.$$

This makes $\{Y_t\}_{t\in\mathbb{N}}$ a Markov chain.

For this Markov chain, the counting measure is not reversible any longer. To find a reversible measure, we look for a measure π on $\mathbb{Z}^d$ which satisfies the detailed balance: for all $z \sim z'$,

$$\pi(z)\frac{\omega_{(z,z')}}{p_\omega(z)} \;=\; \pi(z')\frac{\omega_{(z,z')}}{p_\omega(z')}.$$

The measure $\pi(z) = p_\omega(z)$ satisfies this identity, and one can prove that p_ω is a locally finite measure which is reversible for the random walk. This implies that the measure

$$\overline{\widetilde{\mathbb{P}}}(\omega) \;=\; \frac{\mathbb{P}p_\omega(0)\mathbf{P}_0^\omega}{\langle p_\omega(0)\rangle} \tag{1.12}$$

is reversible for the environment viewed by the particle $t \mapsto \omega(t) = \theta_{Y_t}\omega$.

The Kipnis-Varadhan argument

The aim of this paragraph is to present an argument which shows that the rescaled Markov process $\sqrt{\varepsilon}X_{t/\varepsilon}$ converges in law to a Brownian motion with covariance $2A_{\text{hom}}$, where A_{hom} is as in (1.5).

The idea of Kipnis and Varadhan is to find a decomposition of $\sqrt{\varepsilon}X_{t/\varepsilon}$ as

$$\sqrt{\varepsilon}X_{t/\varepsilon} \;=\; \sqrt{\varepsilon}M_{t/\varepsilon} + \sqrt{\varepsilon}R_{t/\varepsilon}, \tag{1.13}$$

where M_t is a martingale and R_t is some remainder which should be such that $\sqrt{\varepsilon}R_{t/\varepsilon} \to 0$ in $L^2(\overline{\Omega})$ (where $(\overline{\Omega}, \overline{\mathbb{P}})$ is the "global" probability space, with $\overline{\mathbb{P}}$ as in (1.11), and $\overline{\mathbb{E}} = \langle \mathbf{E}_0^\omega \rangle$ the associated expectation). Let $\xi \in \mathbb{R}^d$. The advantage of this decomposition is that general results ensure that if there exists $\sigma^2 \in \mathbb{R}^+$ such that the quadratic variation $[M \cdot \xi, M \cdot \xi]_t$ of $M_t \cdot \xi$, defined by

$$[M \cdot \xi, M \cdot \xi]_t \;:=\; \lim_{k \to \infty} \sum_{j=0}^{k-1} (M_{t(j+1)/k} \cdot \xi - M_{tj/k} \cdot \xi)^2,$$

almost surely satisfies

$$\frac{1}{t}[M \cdot \xi, M \cdot \xi]_t \;\stackrel{t \to \infty}{\longrightarrow}\; \sigma^2,$$

then $\sqrt{\varepsilon}M_{t/\varepsilon} \cdot \xi$ converges in law for the Skorokhod topology to a Brownian motion of variance σ^2 when $\varepsilon \to 0$.

It remains to construct the martingale M_t. We look for a martingale of the form: $M_t = \chi^\omega(X_t)$ for some function χ^ω. We recall that M_t is a martingale for $\mathbf{E}_z^\omega$ if for all $t \geq 0$ and $s \geq 0$,

$$\mathbf{E}_z^\omega[M_{t+s}|\mathcal{F}_t] \;=\; M_t, \tag{1.14}$$

where $\mathcal{F}_t$ is the σ-algebra generated by $\{M_\tau, \tau \in [0,t]\}$. With the choice above a necessary condition for M_t to be a martingale is that for all $z \in \mathbb{Z}^d$,

$$\mathbf{E}_z^\omega[\chi^\omega(X_s)] \;=\; \chi^\omega(z),$$

which we obtain by taking $t = 0$ in (1.14) (recall that $\mathbf{E}_z^\omega$ is the expectation on the random walk starting at z in the environment ω). Since $-L$ is the infinitesimal general of X_s, a Taylor expansion yields

$$\begin{aligned}
\mathbf{E}_z^\omega[\chi^\omega(X_s)] &= P_s \chi^\omega(z) \\
&= (e^{-sL}\chi^\omega)(z) \\
&= \chi^\omega(z) - sL\chi^\omega(z) + O(s^2),
\end{aligned}$$

so that this condition turns into $L\chi^\omega(z) \;=\; 0$. On the other hand, since we want the remainder to be small, we expect χ^ω to be a perturbation of the identity, so that a right choice for χ^ω should be

$$\chi^\omega(z) \;=\; z + \phi(z,\omega) - \phi(0,\omega),$$

where $\phi = (\phi_1, \ldots, \phi_d)$, and for all $i \in \{1, \ldots, d\}$, ϕ_i is the corrector of Definition 1.1 associated with $\xi = \mathbf{e}_i$ (the ith vector of the canonical basis of $\mathbb{R}^d$).

By the following general abstract result, the correctors ϕ are "nice enough" so that $\chi^\omega(X_S)$ indeed defines a martingale: if X_t is a Markov process with generator $-L$, then for all suitable functions $f : \mathbb{Z}^d \to \mathbb{R}^d$, the process $f(X_t) + \int_0^t Lf(X_s)ds$ is a martingale (which we apply to $f = \chi^\omega$). We then consider the decomposition of X_t given by

$$\begin{aligned} M_t &= X_t + \phi(X_t, \omega) - \phi(0, \omega), \\ R_t &= -\phi(X_t, \omega) + \phi(0, \omega). \end{aligned}$$

Since $\nabla\phi$ is stationary, M_t has stationary increments.

In order to conclude one needs to prove two results:

- the convergence of the rescaled quadratic variation of $M_t \cdot \xi$ to $2\xi \cdot A_{\mathrm{hom}}\xi$ for all $\xi \in \mathbb{R}^d$,
- the convergence to zero in $L^2(\Omega)$ of the remainder $\sqrt{\varepsilon}\phi(X_{t/\varepsilon}, \omega)$ as ε vanishes.

Both results follow from the Kipnis-Varadhan theorem, which relies on spectral theory. Since the operator $\mathcal{L}$ of Definition 4 is a bounded self-adjoint operator on $L^2(\Omega)$ (this is equivalent to the reversibility of the measure $\mathbb{P}$ for $-\mathcal{L}$), $\mathcal{L}$ admits a spectral decomposition in $L^2(\Omega)$. For any $g \in L^2(\Omega)$ we denote by e_g the projection of the spectral measure of $\mathcal{L}$ on g. This defines the following spectral calculus: for any bounded continuous function $\Psi : [0, +\infty) \to \mathbb{R}$,

$$\langle (\Psi(\mathcal{L})g)g \rangle \;=\; \int_{\mathbb{R}^+} \Psi(\lambda) de_g(\lambda).$$

Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed, and define the local drift as $\mathfrak{d} = -\nabla^* \cdot A(0,\omega)\xi = -\mathrm{D}^* \cdot A\xi$. Kipnis and Varadhan have proved that if

$$\int_{\mathbb{R}^+} \frac{1}{\lambda} de_{\mathfrak{d}}(\lambda) \;<\; \infty, \tag{1.15}$$

then

$$\frac{1}{t}\overline{\mathbb{E}}[(R_t \cdot \xi)^2] \;=\; 2\int_{\mathbb{R}^+} \frac{1 - e^{-t\lambda}}{t\lambda^2} de_{\mathfrak{d}}(\lambda) \overset{t\to\infty}{\longrightarrow} 0. \tag{1.16}$$

Let us give the argument for (1.16). We first expand the square

$$\begin{aligned} \overline{\mathbb{E}}[(R_t \cdot \xi)^2] &= \left\langle \mathbf{E}_0^\omega[(R_t \cdot \xi)^2] \right\rangle \\ &= \left\langle \mathbf{E}_0^\omega[(\phi(X_t,\omega) \cdot \xi - \phi(0,\omega) \cdot \xi)^2] \right\rangle \\ &= \left\langle (\phi \cdot \xi)^2 \right\rangle - 2\left\langle \mathbf{E}_0^\omega[(\phi(X_t,\omega) \cdot \xi)(\phi(0,\omega) \cdot \xi)] \right\rangle + \left\langle \mathbf{E}_0^\omega[(\phi(X_t,\omega) \cdot \xi)^2] \right\rangle. \end{aligned}$$

Using that $t \mapsto \omega(t) = \theta_{X(t)}\omega$ is stationary for $\overline{\mathbb{P}}$, and assuming that ϕ is stationary for $\mathbb{P}$ (the argument can be made rigorous by using the regularized corrector ϕ_μ and passing to the limit as $\mu \to 0$), the last term turns into

$$\begin{aligned} \left\langle \mathbf{E}_0^\omega[(\phi(X_t,\omega) \cdot \xi)^2] \right\rangle &= \left\langle \mathbf{E}_0^\omega[(\phi(0,\theta_{X(t)}\omega) \cdot \xi)^2] \right\rangle \\ &= \overline{\mathbb{E}}[(\phi(0,\theta_{X(t)}\omega) \cdot \xi)^2] \\ &= \overline{\mathbb{E}}[(\phi(0,\omega) \cdot \xi)^2] \\ &= \left\langle (\phi \cdot \xi)^2 \right\rangle. \end{aligned}$$

For the second term, using that $\mathbf{E}_0^\omega[\phi(X_t,\omega)] = P_t\phi(0,\omega)$ we obtain

$$\langle \mathbf{E}_0^\omega[\phi(X_t,\omega)\cdot\xi]\phi(0,\omega)\cdot\xi\rangle \;=\; \langle(\phi\cdot\xi)(P_t\phi\cdot\xi)\rangle\,.$$

Hence,

$$\overline{\mathbb{E}}[(R_t\cdot\xi)^2] = 2\left\langle(\phi\cdot\xi)^2\right\rangle - 2\left\langle(\phi\cdot\xi)(P_t\phi\cdot\xi)\right\rangle.$$

Using spectral calculus (the argument can be made rigorous by regularization of the corrector) and the facts that $\mathcal{L}(\phi\cdot\xi) \;=\; \mathfrak{d}$ and that $P_t = e^{-t\mathcal{L}}$ on $L^2(\Omega)$, this identity turns into the desired formula

$$\overline{\mathbb{E}}[(R_t\cdot\xi)^2] = 2\int_{\mathbb{R}^+}\frac{1-e^{-t\lambda}}{\lambda^2}de_{\mathfrak{d}}(\lambda).$$

It is easy to show that this term is finite provided (1.15) holds, as well as the convergence to zero in (1.16).

We sketch now the proof of (1.15), which already appears in [75]. The guideline is to prove that the local drift $\mathfrak{d}$ is in the range of $\mathcal{L}^{1/2}$, that is there exists $h\in L^2(\Omega)$ such that $\mathfrak{d}=\mathcal{L}^{1/2}h$, since (formally)

$$\left\langle|\mathcal{L}^{-1/2}\mathfrak{d}|^2\right\rangle \;=\; \int_{\mathbb{R}^+}\frac{1}{\lambda}de_{\mathfrak{d}}(\lambda).$$

As usual, we proceed by regularization, and for all $\mu>0$ we set $h_\mu := (\mu+\mathcal{L})^{-1/2}\mathfrak{d}$, which is well-defined in $L^2(\Omega)$ by spectral calculus. In addition,

$$\left\langle h_\mu^2\right\rangle \;=\; \left\langle|(\mu+\mathcal{L})^{-1/2}\mathfrak{d}|^2\right\rangle \;=\; \int_{\mathbb{R}^+}\frac{1}{\mu+\lambda}de_{\mathfrak{d}}(\lambda)$$

so that (1.15) follows from the monotone convergence theorem provided we prove the uniform boundedness of h_μ in $L^2(\Omega)$. To this aim we will use the following observation: by integration by parts and Cauchy-Schwarz inequality, for all $g\in L^2(\Omega)$,

$$\begin{aligned}
|\left\langle\mathfrak{d}g\right\rangle| &= |\left\langle g\,\mathrm{D}^*\cdot A\xi\right\rangle| \\
&= |\left\langle\mathrm{D}\,g\cdot A\xi\right\rangle| \\
&\le \left\langle\mathrm{D}\,g\cdot A\,\mathrm{D}\,g\right\rangle^{1/2}\left\langle\xi\cdot A\xi\right\rangle^{1/2} \\
&\le \sqrt{\beta}\left\langle g\mathcal{L}g\right\rangle^{1/2}.
\end{aligned}$$

Applied to the test function $(\mu+\mathcal{L})^{-1/2}g$ this yields:

$$\begin{aligned}
|\left\langle h_\mu g\right\rangle| &= |\left\langle g(\mu+\mathcal{L})^{-1/2}\mathfrak{d}\right\rangle| \\
&= |\left\langle\mathfrak{d}[(\mu+\mathcal{L})^{-1/2}g]\right\rangle| \\
&\le \sqrt{\beta}\left\langle[(\mu+\mathcal{L})^{-1/2}g]\mathcal{L}[(\mu+\mathcal{L})^{-1/2}g]\right\rangle \\
&= \sqrt{\beta}\left\langle(\mathcal{L}(\mu+\mathcal{L})^{-1}g)g\right\rangle^{1/2} \\
&\le \sqrt{\beta}\left\langle g^2\right\rangle^{1/2},
\end{aligned}$$

that is, the uniform boundedness of h_μ, which shows the validity of (1.15).

To conclude this section, we give the argument for the fact that the covariance matrix of the Brownian motion at the limit is $2A_{\rm hom}$. The fact that the rescaled quadratic variation of $M_t \cdot \xi$ converges almost surely to some constant can be proved by a suitable application of the ergodic theorem. We only give the argument to identify this limit. Since $M_t \cdot \xi$ is a martingale with stationary increments, there exists some constant C such that for all $t \geq 0$,

$$\overline{\mathbb{E}}[(M_t \cdot \xi)^2] \;=\; Ct.$$

It is therefore enough to show that

$$\overline{\mathbb{E}}[(M_t \cdot \xi)^2] \;=\; 2t\xi \cdot A_{\rm hom}\xi + O(t^2). \tag{1.17}$$

The starting point is (1.10) for $z = 0$ applied to $f : x \mapsto (x \cdot \xi + \phi(x,\omega) \cdot \xi - \phi(0,\omega) \cdot \xi)^2$, which shows that

$$\begin{aligned}\mathbf{E}_0^\omega[(M_t \cdot \xi)^2] \;=\; \mathbf{E}_0^\omega[f(X_t)] &= f(0) + t\sum_{z'\sim 0} \omega_{(0,z')}\big(z' \cdot \xi + \phi(z',\omega) \cdot \xi - \phi(0) \cdot \xi\big)^2 + O(t^2)\\ &= 2t(\xi + \nabla\phi(0,\omega) \cdot \xi) \cdot A(0,\omega)(\xi + \nabla\phi(0,\omega) \cdot \xi) + O(t^2),\end{aligned}$$

which implies the desired expansion (1.17) by taking the expectation $\langle\cdot\rangle$.

1.2 Existence of stationary correctors for $d > 2$ [GOa]

The aim of this section is to prove the uniform boundedness of the regularized corrector for $d > 2$ in the case when the conductivities are i. i. d. random variables.

Theorem 1 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. For all $q > 0$ there exist $C_{d,q} < \infty$ and $\gamma(q) > 0$ such that for all $T \geq 1$ and $\xi \in \mathbb{R}^d$ with $|\xi| = 1$, the unique stationary solution $\phi_T \in L^2(\Omega)$ to (1.6) with $\mu = T^{-1}$ satisfies*

$$\langle|\phi_T|^q\rangle \;\leq\; \begin{cases} d = 2 : C_{2,q}(\ln T)^{\gamma(q)}, \\ d > 2 : C_{d,q}. \end{cases} \tag{1.18}$$

Remark 2 *In* [GNOa], *we have proved that* (1.18) *holds with* $\gamma(2) = 1$.

In particular this result for $q = 2$ and $d > 2$ provides a uniform bound on $\langle\phi_T^2\rangle$ which allows to obtain the weak convergence of ϕ_T to some $\phi \in L^2(\Omega)$ up to extraction as $T \to \infty$. This function ϕ is a stationary solution to the corrector equation (1.4). This result is rather surprising since in dimension $d = 1$, there cannot exist stationary solutions $\phi \in L^2(\Omega)$ to the corrector equation, as we briefly explain below in the continuous case. The corrector equations then reads:

$$(a(x)(1 + \phi'(x)))' \;=\; 0 \qquad \text{in } \mathbb{R}^d.$$

Integrating this equation twice and using that $\langle\phi'\rangle = 0$, we obtain for all $x \geq 0$,

$$\phi(x)-\phi(0) \;=\; \int_0^x \frac{1}{\langle 1/a\rangle}\frac{dt}{a(t)} \;=\; \sqrt{x}\left[\sqrt{x}\left(\fint_0^x \frac{1}{\langle 1/a\rangle}\frac{dt}{a(t)}-1\right)\right],$$

and therefore

$$(\phi(x)-\phi(0))^2 \;=\; x\left[\sqrt{x}\left(\fint_0^x \frac{1}{\langle 1/a\rangle}\frac{dt}{a(t)}-1\right)\right]^2.$$

We then integrate on $[0,y]$ for $y>0$:

$$\fint_0^y (\phi(x)-\phi(0))^2 dx \;=\; \fint_0^y x\left[\sqrt{x}\left(\fint_0^x \frac{1}{\langle 1/a\rangle}\frac{dt}{a(t)}-1\right)\right]^2 dx.$$

Let us take the expectation. By the central limit theorem, the term into brackets is essentially of order 1, so that the integral on r. h. s. is of order y. On the other hand, if ϕ were stationary, the l. h. s. would be bounded, and there is a contradiction. As Theorem 1 shows, there is a transition between unboundedness and boundedness of the corrector in $L^2(\Omega)$, and dimension $d=2$ is critical (so that a logarithm was to be expected in (1.18)).

The proof of Theorem 1 relies on three ingredients: a spectral gap estimate, Cacciopoli's inequality in probability, and (sharp) bounds on Green's functions. We proceed by induction and only present a simplified string of arguments.

The starting point is the elementary identity for all $m\in\mathbb{N}$

$$\left\langle \phi_T^{2m}\right\rangle \;=\; \langle \phi_T^m\rangle^2+\mathrm{var}\left[\phi_T^m\right], \tag{1.19}$$

which shows that provided we control the variance of moments of ϕ_T, one may bootstrap moments estimates by induction.

The control of the variance relies on the following weak version of a spectral gap estimate:

Lemma 1.3 (variance estimate). *Let $\Omega=[\alpha,\beta]^{\mathbb{B}}$ and $\mathbb{P}$ be as in Definition 1, and let X be a Borel measurable function of $\omega\in\Omega$ (i. e. measurable w. r. t. the smallest σ-algebra on $[\alpha,\beta]^{\mathbb{B}}$ for which all coordinate functions $\Omega\ni\omega\mapsto\omega_e\in[\alpha,\beta]$ with $e\in\mathbb{B}$ are Borel measurable, cf. [46, Definition 14.4]). Then we have*

$$\mathrm{var}\left[X\right] \;\le\; \left\langle \sum_{e\in\mathbb{B}}\sup_{\omega_e}\left|\frac{\partial X}{\partial\omega_e}\right|^2\right\rangle \mathrm{var}\left[\tilde\omega\right], \tag{1.20}$$

where $\tilde\omega$ has the same law as all the ω_e, $e\in\mathbb{B}$.

This lemma can be proved using the Lu-Yau martingale approach (see [54] for related log-Sobolev inequalities).

We shall indeed apply (1.20) to $X=\phi_T^m$. To this aim we need to estimate the susceptibility of the regularized corrector with respect to the conductivities. This is where the Green's function comes into the picture. A formal differentiation of (1.6) with respect to ω_e for $e=(z,z')$, $z'=z+\mathbf{e}_i$ yields for all $x\in\mathbb{Z}^d$

$$T^{-1}\frac{\partial \phi_T}{\partial \omega_e}(x) - \left(\nabla^* \cdot A\nabla \frac{\partial \phi_T}{\partial \omega_e}\right)(x) - (\nabla_i \phi_T(z) + \xi_i)(\delta(x-z) - \delta(x-z')) \;=\; 0.$$

Noting that the Green's function $G_T : \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega \to \mathbb{R}$ is solution to

$$T^{-1}G_T(x,y,\omega) - \nabla_x^* \cdot A(x,\omega)\nabla_x G_T(x,y,\omega) \;=\; \delta(x-y),$$

this identity takes the form

$$\frac{\partial \phi_T}{\partial \omega_e}(x) \;=\; -(\nabla_i \phi_T(z) + \xi_i)(G_T(x,z') - G_T(x,z)) \;=\; -(\nabla_i \phi_T(z) + \xi_i)\nabla_{z_i} G_T(x,z). \quad (1.21)$$

From now on we neglect the supremum in (1.20). Although the rest of this section is therefore only formal, it allows to focus on the core of the argument. By the Leibniz rule, and replacing the sum on the edges of $\mathbb{B}$ by d sums on the sites $\mathbb{Z}^d$, we thus have

$$\mathrm{var}\,[\phi_T(0)^m] \;\lesssim\; \sum_{z\in\mathbb{Z}^d} \left\langle \phi_T(0)^{2(m-1)}(|\nabla \phi_T(z)|^2 + 1)|\nabla_z G_T(0,z)|^2 \right\rangle. \quad (1.22)$$

In view of this estimate, it seems that both sides of (1.19) have the same powers on ϕ_T. Yet the spectral gap estimate (1.22) yields a gradient of ϕ_T, from which we can benefit by the following Cacciopoli inequality in probability: for all $n \in 2\mathbb{N}$,

$$\left\langle \phi_T(0)^n(|\nabla \phi_T(0)|^2 + |\nabla^* \phi_T(0)|^2)\right\rangle \;\lesssim\; \langle \phi_T(0)^n \rangle. \quad (1.23)$$

This inequality follows from testing the regularized corrector equation with function ϕ_T^{2n-1}, integrating by parts, neglecting the non-negative contribution of the zero-order term, and taking the expectation. The strategy is then clear: use deterministic bounds on the Green's function, appeal to Hölder's inequality in probability with exponents $(\frac{m}{m-1}, m)$ in the r. h. s. of (1.22), and use the Cacciopoli inequality on the gradient term.

Let us assume for simplicity that $|\nabla_z G_T(0,z)| \;\lesssim\; (1 + |z|^{1-d})\exp(-c|z|/\sqrt{T})$ (this estimate indeed does not hold pointwise, but survives for squared averages on dyadic annuli by Cacciopoli's inequality in space). We then have by Hölder's inequality

$$\begin{aligned}\mathrm{var}\,[\phi_T(0)^m] &\lesssim \sum_{z\in\mathbb{Z}^d}(1+|z|^{2-2d})\exp(-2c|z|/\sqrt{T})\left\langle \phi_T(0)^{2(m-1)}(|\nabla \phi_T(z)|^2+1)\right\rangle \\ &\lesssim \sum_{z\in\mathbb{Z}^d}(1+|z|^{2-2d})\exp(-2c|z|/\sqrt{T})\left\langle \phi_T(0)^{2m}\right\rangle^{1-1/m}(1+\left\langle |\nabla \phi_T(z)|^{2m}\right\rangle^{1/m}).\end{aligned}$$

By stationarity of $\nabla \phi_T$,

$$\mathrm{var}\,[\phi_T(0)^m] \lesssim \sum_{z\in\mathbb{Z}^d}(1+|z|^{2-2d})\exp(-2c|z|/\sqrt{T})\left\langle \phi_T(0)^{2m}\right\rangle^{1-1/m}(1+\left\langle |\nabla \phi_T(0)|^{2m}\right\rangle^{1/m}).$$

Set $\mu_d(T) = \begin{cases} d=2 : \ln T \\ d>2 : 1 \end{cases}$. Estimating the sum on $\mathbb{Z}^d$ yields

$$\operatorname{var}[\phi_T(0)^m] \lesssim \mu_d(T)\left\langle \phi_T(0)^{2m}\right\rangle^{1-1/m}(1+\left\langle|\nabla\phi_T(0)|^{2m}\right\rangle^{1/m}).$$

Using now the fact that $|\nabla\phi_T(0)| \lesssim |\phi_T(0)| + \sum_{|x|=1}|\phi_T(x)|$, the Cacciopoli inequality (1.23) yields

$$\begin{aligned}\operatorname{var}[\phi_T(0)^m] &\lesssim \mu_d(T)\left\langle \phi_T(0)^{2m}\right\rangle^{1-1/m}\left(1+\left\langle \phi_T(0)^{2(m-1)}(|\nabla^*\phi_T(0)|^2+|\nabla\phi_T(0)|^2)\right\rangle\right)\\ &\lesssim \mu_d(T)\left\langle \phi_T(0)^{2m}\right\rangle^{1-1/m}\left((1+\left\langle \phi_T(0)^{2(m-1)}\right\rangle^{1/m}\right).\end{aligned}$$

We conclude by Hölder's inequality with exponents $(\frac{m}{m-1}, m)$ on the second expectation that

$$\operatorname{var}[\phi_T^m] \lesssim \mu_d(T)(1+\left\langle \phi_T^{2m}\right\rangle^{1-1/m^2}).$$

By Young's inequality, this turns (1.19) into

$$\left\langle \phi_T^{2m}\right\rangle \lesssim \left\langle \phi_T^{m}\right\rangle^2 + \mu_d(T)^{m^2}$$

from which Theorem 1 follows using $\langle \phi_T\rangle = 0$.

To turn this simplified string of arguments into a rigorous proof, the crucial additional ingredient is Meyers' estimate.

1.3 Spectral exponents and approximation formulas for the homogenized coefficient [GOb,GMa,GNOa]

In this section, we fix once and for all some $\xi \in \mathbb{R}^d$ with $|\xi| = 1$.

1.3.1 Motivation

From the practical point of view, in order to compute a numerical approximation of the homogenized matrix A_{hom} defined in (1.5), the natural starting point is the following consequence of ergodicity: almost surely we have

$$\xi\cdot A_{\text{hom}}\xi = \lim_{N\to\infty} \fint_{Q_N\cap\mathbb{Z}^d}(\xi+\nabla\phi(z))\cdot A(z)(\xi+\nabla\phi(z))dz,$$

where $Q_N = (0,N)^d$. We then choose some large $N \in \mathbb{N}$, and consider the (random) approximation A_N of A_{hom}, defined by

$$\xi\cdot A^N(\omega)\xi = \fint_{Q_N\cap\mathbb{Z}^d}(\xi+\nabla\phi(z,\omega))\cdot A(z,\omega)(\xi+\nabla\phi(z,\omega))dz.$$

This approximation is still of no practical interest since the equation (1.4) for ϕ is posed on the whole $\mathbb{Z}^d$. We thus need to find a computable approximation of the corrector ϕ on

Q_N. A natural proxy is the solution of the corrector equation restricted on Q_N. The central question there is the choice of the boundary conditions to impose on ∂Q_N, since the value of ϕ on ∂Q_N is also part of the problem. A possible choice is to consider homogeneous Dirichlet boundary conditions, and replace ϕ by the solution $\phi^N : Q_N \cap \mathbb{Z}^d \to \mathbb{R}$ to

$$\begin{cases} -\nabla^* \cdot A(\xi + \nabla \phi^N) \,=\, 0 \text{ in } Q_N \cap \mathbb{Z}^d, \\ \qquad\qquad \phi^N(z) = 0 \text{ on } \partial Q_N \cap \mathbb{Z}^d. \end{cases}$$

and finally set

$$\xi \cdot A^{N,N}(\omega)\xi \,:=\, \fint_{Q_N \cap \mathbb{Z}^d} (\xi + \nabla \phi^N(z)) \cdot A(z)(\xi + \nabla \phi^N(z)) dz. \tag{1.24}$$

On may then prove that almost surely $A^{N,N}(\omega) \to A_{\mathrm{hom}}$. Yet the convergence rate of $A^{N,N}$ to A_{hom} is likely to be driven by the effect of the boundary conditions, which is expected to scale like a surface effect (that is $1/N$ in any dimension). Indeed, if A were a periodic function, one would have $A^{N,N} - A_{\mathrm{hom}} \sim N^{-1}$. Can we do better than this ?

Instead of ϕ^N, another proxy for ϕ could be the regularized corrector ϕ_T for some large $T > 0$ (with T a function of N). What we gain by considering ϕ_T instead of ϕ is that, due to the exponential decay of the Green's function $G_T(x,y) \leq\, (1+|x-y|)^{2-d} \exp(-c|x-y|/\sqrt{T})$, the solution $\phi_T^R : Q_R \cap \mathbb{Z}^d \to \mathbb{R}$ to

$$\begin{cases} T^{-1}\phi_T^R - \nabla^* \cdot A(\xi + \nabla \phi_T^R) \,=\, 0 \text{ in } Q_R \cap \mathbb{Z}^d, \\ \qquad\qquad\qquad \phi_T^R(z) = 0 \text{ on } \partial Q_R \cap \mathbb{Z}^d, \end{cases}$$

is a very good approximation of ϕ_T on Q_N provided $R - N \gg \sqrt{T}$. Hence, one may consider ϕ_T to a be computable proxy for ϕ, and we choose as approximation of A_{hom} the random quantity

$$\xi \cdot A_T^N(\omega)\xi \,=\, \fint_{Q_N \cap \mathbb{Z}^d} (\xi + \nabla \phi_T(z,\omega)) \cdot A(z,\omega)(\xi + \nabla \phi_T(z,\omega)) dz.$$

By ergodicity,

$$\lim_{N\to\infty} \xi \cdot A_T^N(\omega)\xi \,=\, \langle (\xi + \nabla \phi_T) \cdot A(\xi + \nabla \phi_T) \rangle \,=:\, \xi \cdot A_T \xi,$$

so that one needs to control $A_T - A_{\mathrm{hom}}$ (which we call the systematic error) to study the convergence of $A_T^N(\omega)$ to A_{hom}. A direct computation using the corrector and regularized corrector equations yields

$$\xi \cdot (A_T - A_{\mathrm{hom}})\xi \,=\, \langle (\nabla \phi_T - \nabla \phi) \cdot A(\nabla \phi_T - \nabla \phi) \rangle \,.$$

In particular, this approximation is consistent if and only if $\mathrm{D}\,\phi_T \to \mathrm{D}\,\phi$ in $L^2(\Omega)$. Yet it seems we only know that this convergence is weak in $L^2(\Omega)$ since it has been obtained using the Banach-Alaoglu theorem. This is where spectral theory pops up again.

In order to show that $\mathrm{D}\,\phi_T \to \mathrm{D}\,\phi$ in $L^2(\Omega)$, it is enough to show that $\mathrm{D}\,\phi_T$ is a Cauchy sequence. By ellipticity of A, self-adjointness of $\mathcal{L}$, and by spectral calculus, for all $T_2 \geq T_1 > 0$,

$$\begin{aligned}\left\langle |\nabla\phi_{T_1} - \nabla\phi_{T_2}|^2 \right\rangle &\lesssim \langle (\nabla\phi_{T_1} - \nabla\phi_{T_2}) \cdot A(\nabla\phi_{T_1} - \nabla\phi_{T_2}) \rangle \\ &= \langle (\mathcal{L}(\phi_{T_1} - \phi_{T_2}))(\phi_{T_1} - \phi_{T_2}) \rangle \\ &= \left\langle \left[\mathcal{L}\Big((T_1^{-1} + \mathcal{L})^{-1} - (T_2^{-1} + \mathcal{L})^{-1} \Big)^2 \mathfrak{d} \right] \mathfrak{d} \right\rangle \\ &= \int_{\mathbb{R}^+} \lambda \Big(\frac{1}{T_1^{-1} + \lambda} - \frac{1}{T_2^{-1} + \lambda} \Big)^2 de_{\mathfrak{d}}(\lambda) \\ &= \int_{\mathbb{R}^+} \frac{\lambda (T_2^{-1} - T_1^{-1})^2}{(T_1^{-1} + \lambda)^2 (T_2^{-1} + \lambda)^2} de_{\mathfrak{d}}(\lambda).\end{aligned}$$

Using that the integrand is an increasing function of T_2, we take the limit as $T_2 \to +\infty$ in the r. h. s., which yields for all $T_2 \geq T_1 > 0$,

$$\left\langle |\nabla\phi_{T_1} - \nabla\phi_{T_2}|^2 \right\rangle \lesssim \int_{\mathbb{R}^+} \frac{T_1^{-2}}{(T_1^{-1} + \lambda)^2 \lambda} de_{\mathfrak{d}}(\lambda).$$

In view of (1.15), we conclude by the Lebesgue dominated theorem that $\nabla\phi_T$ is a Cauchy sequence in $L^2(\Omega)$ so that $\nabla\phi_T$ converges to $\nabla\phi$ in $L^2(\Omega)$, and

$$\lim_{T \to \infty} A_T \;=\; A_{\mathrm{hom}}. \tag{1.25}$$

In addition, as a by-product of the analysis we have proved the following identity for all $T > 0$:

$$\xi \cdot (A_T - A_{\mathrm{hom}})\xi \;=\; T^{-2} \int_{\mathbb{R}^+} \frac{1}{(T^{-1} + \lambda)^2 \lambda} de_{\mathfrak{d}}(\lambda). \tag{1.26}$$

We may now put the pieces of the puzzle together. If we knew that $\lambda \mapsto \lambda^{-3}$ were integrable for the measure $de_{\mathfrak{d}}$, then $A_T - A_{\mathrm{hom}}$ would be of order T^{-2}. This we don't know at this stage. However, by spectral calculus, Theorem 1 takes the form for $d > 2$

$$\left\langle \phi_T^2 \right\rangle \;=\; \int_{\mathbb{R}^+} \frac{1}{(T^{-1} + \lambda)^2} de_{\mathfrak{d}}(\lambda) \;\lesssim\; 1,$$

where the bound is uniform in T. This shows in particular by the monotone convergence theorem that

$$\int_{\mathbb{R}^+} \frac{1}{\lambda^2} de_{\mathfrak{d}}(\lambda) \;\lesssim\; 1. \tag{1.27}$$

Hence the qualitative statement $A_T \to A_{\mathrm{hom}}$ above, due to Papanicolaou and Varadhan, can be turned quantitative for $d > 2$:

$$\begin{aligned}\xi\cdot(A_T-A_{\mathrm{hom}})\xi &\lesssim T^{-1}\int_{\mathbb{R}^+}\frac{1}{(T^{-1}+\lambda)\lambda}de_{\mathfrak{d}}(\lambda)\\ &\leq T^{-1}\int_{\mathbb{R}^+}\frac{1}{\lambda^2}de_{\mathfrak{d}}(\lambda)\\ &\lesssim T^{-1},\end{aligned}\tag{1.28}$$

and this gives a first quantitative error estimate for a "computable" approximation of the homogenized coefficients (note that we still have to control the difference between the average on Q_N and the expectation, which we shall do in Section 1.4).

The aim of this section is twofold. We have seen above that bounds on the spectral measure (typically in the form of (1.27)) yield bounds on the error $A_T - A_{\mathrm{hom}}$. Our first goal is to make this statement a bit more systematic and prove optimal bounds on the spectral measure, which in turn will give optimal bounds on $A_T - A_{\mathrm{hom}}$. We shall then see that the bounds on $A_T - A_{\mathrm{hom}}$ depend on the dimension and saturate at T^{-2} in dimension $d = 5$. Our second objective is to devise a family of computable approximations $\{A_{k,T}\}_{k\in\mathbb{N}}$ of A_{hom} for which the error $A_{k,T} - A_{\mathrm{hom}}$ can be made optimal with respect to the bounds on the spectral measure (the order k depending on the dimension d).

1.3.2 Definition of spectral exponents

In order to make the link between (1.27) and (1.28) more systematic, we introduce the following definition of spectral exponents.

Definition 5 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let $\mathcal{L}$ be as in Definition 4, let $g \in L^2(\Omega)$, and let e_g be the projection of the spectral measure of $\mathcal{L}$ on g. We say that the spectral exponents of g are at least $(\gamma, -q)$ for $\gamma > 1$ and $q \geq 0$ if for all $\mu > 0$ we have*

$$\int_0^\mu de_g(\lambda) \;\leq\; \mu^\gamma \ln_+^q(\mu^{-1}),\tag{1.29}$$

where $\ln_+(t) = |\ln t|$.

This definition makes sense since if the spectral exponents of g are at least $(\gamma, -q)$, they are at least $(\gamma', -q')$ for all $1 < \gamma' \leq \gamma$ and $0 \leq q' \leq q$.

The interest of this definition lies in the following two results, which relate the spectral exponents to estimates of the types (1.27) and (1.28), respectively.

Lemma 1.4. *Let e be a non-negative measure. Let $k \in \mathbb{N}$. If there exist $\gamma' < 2k$ and $q \geq 0$ such that for all $\mu > 0$*

$$\int_{\mathbb{R}^+}\frac{1}{(\mu+\lambda)^{2k}}de(\lambda) \;\lesssim\; \mu^{-\gamma'}\ln_+^q(\mu^{-1}),$$

then e satisfies (1.29) with exponents $(\gamma, -q)$, $\gamma = 2k - \gamma'$.

This result directly follows from the inequality: for all $\mu > 0$,

$$\int_0^\mu de(\lambda) \,\leq\, \mu^{2k} \int_{\mathbb{R}^+} \frac{1}{(\mu+\lambda)^{2k}} de(\lambda).$$

On the other hand, we also have:

Lemma 1.5. *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let $\mathcal{L}$ be as in Definition 4, let $g \in L^2(\Omega)$, and let e_g be the projection of the spectral measure of $\mathcal{L}$ on g. If the spectral exponents of g are at least $(\gamma, -q)$ for some $\gamma > 1$ and $q \geq 0$, then for all $k \in \mathbb{N}$,*

$$\mu^{2k} \int_{\mathbb{R}^+} \frac{1}{\lambda(\mu+\lambda)^{2k}} de_g(\lambda) \,\lesssim\, \begin{cases} \mu^{2k} & \text{if } \gamma > 2k+1, \\ \mu^{2k} \ln_+^{1+q}(\mu^{-1}) & \text{if } \gamma = 2k+1, \\ \mu^{\gamma-1} \ln_+^{q}(\mu^{-1}) & \text{if } \gamma < 2k+1. \end{cases} \tag{1.30}$$

The proof of this lemma is slightly more subtle than above. Before we give the argument, we illustrate the use of these lemmas to estimate $A_{\mathrm{hom}} - A_T$ in dimension $d = 2$. The spectral exponents of the local drift $\mathfrak{d} = -\mathrm{D}^* \cdot A\xi$ are $(2, -q)$ by Theorem 1, so that $|A_T - A_{\mathrm{hom}}| \lesssim T^{-1} \ln_+^{1+q} T$ for $d = 2$.

The proof of Lemma 1.5 relies on the following application of the fundamental theorem of calculus and Fubini's theorem: for every differentiable functions $f : \mathbb{R}^+ \to \mathbb{R}$,

$$\int_{\mathbb{R}^+} f(\lambda) de_{\mathfrak{d}}(\lambda) \,=\, -\int_{\lambda=0}^{+\infty} \int_{\delta=\lambda}^{+\infty} f'(\delta) d\delta de_{\mathfrak{d}}(\lambda) \,=\, -\int_{\mathbb{R}^+} f'(\delta) \int_0^\delta de_{\mathfrak{d}}(\lambda) d\delta, \tag{1.31}$$

that we apply to $f(\lambda) = \frac{1}{\lambda(\mu+\lambda)^{2k}}$. We then split the integral into two terms:

$$\int_{\mathbb{R}^+} f'(\delta) \int_0^\delta de_{\mathfrak{d}}(\lambda) d\delta \,=\, \int_0^1 f'(\delta) \int_0^\delta de_{\mathfrak{d}}(\lambda) d\delta + \int_1^\infty f'(\delta) \int_0^\delta de_{\mathfrak{d}}(\lambda) d\delta.$$

For the first term we appeal to the bounds on the spectral exponents, which give the r. h. s. of (1.30). For the second term, we use the non-negativity of the measure, the identity $\int_0^\infty de_{\mathfrak{d}}(\lambda) = \langle \mathfrak{d}^2 \rangle \lesssim 1$, and the integrability of f' on $(1, +\infty)$ uniformly in μ.

1.3.3 Approximation formulas for the homogenized coefficients

In view of Lemmas 1.4 and 1.5, in order for some approximation "$A_{k,T}$" of A_{hom} to benefit from high spectral exponents, the spectral formulation of the error term $A_{k,T} - A_{\mathrm{hom}}$ should involve high powers of λ^{-1} in the integral. There are then two points of view: devise suitable approximations of A_{hom} in spectral space (that is in the form of a spectral integral) and try to translate them back into physical space (that is using the correctors), or devise suitable approximations of A_{hom} in physical space directly and show they yield the desired integrands in spectral space. We quickly present both strategies.

We begin by recalling the spectral representation of A_{hom}: by definition (1.5) of A_{hom} and symmetry of A,

$$\xi \cdot A_{\rm hom}\xi \;=\; \langle (\xi+\nabla\phi)\cdot A(\xi+\nabla\phi)\rangle \;=\; \langle \xi\cdot A\xi\rangle + 2\,\langle \xi\cdot A\nabla\phi\rangle + \langle \nabla\phi\cdot A\nabla\phi\rangle\,.$$

Taking the limit $T\to\infty$ in the weak formulation of the regularized corrector equation (1.6) with test function ϕ_T yields

$$\langle \nabla\phi\cdot A(\xi+\nabla\phi)\rangle = 0,$$

so that by symmetry of A and spectral calculus we end up with

$$\xi\cdot A_{\rm hom}\xi \;=\; \langle \xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\frac{1}{\lambda}de_{\mathfrak{d}}(\lambda). \tag{1.32}$$

The rigorous proof of this fact is made by regularization. Indeed

$$\xi\cdot A_T\xi \;=\; \langle \xi\cdot A\xi\rangle - \langle \nabla\phi_T\cdot A\nabla\phi_T\rangle - 2T^{-1}\left\langle \phi_T^2\right\rangle.$$

so that

$$\xi\cdot A_T\xi \;=\; \langle \xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\left(\frac{\lambda}{(T^{-1}+\lambda)^2} - \frac{2T^{-1}}{(T^{-1}+\lambda)^2}\right)de_{\mathfrak{d}}(\lambda).$$

We then recover (1.32) by taking the limit $T\to\infty$, using (1.25), the Lebesgue dominated convergence theorem, and the monotone convergence theorem.

In view of the spectral formula for $A_{\rm hom}$, we introduce the following family of approximations $\{A_{k,T}\}_{k\in\mathbb{N}\setminus\{0\}}$ of $A_{\rm hom}$, defined by

$$\xi\cdot A_{k,T}\xi \;=\; \langle \xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\frac{P_k(T,\lambda)}{(T^{-1}+\lambda)^2(2T^{-1}+\lambda)^2\dots(2^{k-1}T^{-1}+\lambda)^2}de_{\mathfrak{d}}(\lambda), \tag{1.33}$$

where

$$P_k(T,\lambda) \;=\; \lambda^{-1}\bigg((T^{-1}+\lambda)^2(2T^{-1}+\lambda)^2\dots(2^{k-1}T^{-1}+\lambda)^2 - 2^{k(k-1)}T^{-2k}\bigg).$$

These rather complicated formulas satisfy two properties. First we have

$$|\xi\cdot(A_{k,T}-A_{\rm hom})\xi| \;\lesssim\; T^{-2k}\int_{\mathbb{R}^+}\frac{1}{(T^{-1}+\lambda)^{2k}\lambda}de_{\mathfrak{d}}(\lambda), \tag{1.34}$$

so that whatever the spectral exponents of the local drift $\mathfrak{d} = -\mathrm{D}^*\cdot A\xi$, there is always an integer k which saturates the estimate of Lemma 1.5. The other fundamental property is that there exist (explicit) real coefficients $\eta_{k,i}$ and $\nu_{k,i,j}$ defined by induction such that we have the identity

$$\xi\cdot A_{k,T}\xi \;=\; \langle (\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\rangle + T^{-1}\sum_{i=0}^{k-1}\eta_{k,i}\left\langle \phi_{2^{-i}T}^2\right\rangle + T^{-1}\sum_{i=0}^{k-1}\sum_{j>i}^{k-1}\nu_{k,i,j}\left\langle \phi_{2^{-i}T}\phi_{2^{-j}T}\right\rangle. \tag{1.35}$$

This ensures that the approximations $A_{k,T}$ of A_{hom} are computable in practice (up to replacing the expectations by space averages on Q_N, and the regularized correctors by approximations computed on larger boxes Q_R, $R-N \gg \sqrt{T}$).

So far we have introduced new approximation formulas for A_{hom} based on a spectral analysis and extrapolation at the level of the spectral calculus. The rather tedious part (which is not reproduced here) is to choose the approximation formulas so that they have a suitable counterparts in physical space. Although this family of approximations makes the job, if we only focus on their definitions in physical space, we have essentially no intuition why they behave so well in terms of convergence rate (with respect to the spectral exponents). The interesting feature of this approach is that abstract analysis has been a guide to the choice of efficient computable approximations.

There is however a more intuitive way to obtain similar approximations, using a Richardson extrapolation method. For all $k \geq 1$, we define approximations $\tilde{A}_{k,T}$ of A_{hom} by the following induction rule: $\tilde{A}_{1,T} = A_T$, and for all $k \geq 1$,

$$\xi \cdot \tilde{A}_{k+1,T}\xi \;=\; \frac{1}{2^{k+1}-1}\Big(2^{k+1}\xi \cdot \tilde{A}_{k,T}\xi - \xi \cdot \tilde{A}_{k,T/2}\xi\Big). \tag{1.36}$$

The spectral formulation of these approximations also reveal that one can benefit from large spectral exponents of the local drift $\mathfrak{d}$ provided k is large enough.

In order to complete this analysis, we need to identify the spectral exponents of the local drift $\mathfrak{d}$.

1.3.4 Estimates of the spectral exponents

The central result of this section is the following theorem.

Theorem 2 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let $\mathcal{L}$ be as in Definition 4. The spectral exponents of the local drift $\mathfrak{d} = -\,\mathrm{D}^* \cdot A\xi \in L^2(\Omega)$ are at least*

$$\begin{aligned} d = 2 &: (2, -q), \\ d > 2 &: (d/2+1, 0), \end{aligned}$$

for some $q > 0$ depending only on α, β.

These exponents are expected to be optimal in any dimension but $d = 2$. Note that Mourrat had proved the lower bound $(d/2-2, 0)$ for all $d > 6$ in [63]. Our proof of Theorem 2 is based on an induction procedure combining three main ingredients: the spectral theorem, an a priori estimate for elliptic equations, and the variance estimate of Lemma 1.3 (combined with sharp estimates on Green's functions).

In what follows we highlight the structure of the proof, and skip all the technical details — which are essentially of the same type as for the proof of Theorem 1. Since we have already proved the result for $d = 2$, we focus on the case $d > 2$.

We first introduce a family of functions $\phi_{T,k} \in L^2(\Omega)$ (and their stationary extensions on $\mathbb{Z}^d \times \Omega$) as follows. We set $\phi_{T,1} := \phi_T$ (the regularized corrector), and for all $k \geq 1$, we define $\phi_{T,k+1}$ as the unique stationary solution in $\mathbb{Z}^d$ to

$$T^{-1}\phi_{T,k+1} - \nabla^* \cdot A\nabla\phi_{T,k+1} \;=\; \phi_{T/2,k}. \tag{1.37}$$

By induction, these functions satisfy $\langle \phi_{T,k} \rangle = 0$, and for all $k \geq 1$ there exists $c_k > 0$ (independent of T) such that

$$\phi_{T,k+1} \;=\; c_k T(\phi_{T,k} - \phi_{T/2,k}). \tag{1.38}$$

The main three ingredients are used as follows.

By the spectral theorem,

$$\left\langle \phi_{T,k}^2 \right\rangle \;=\; \int_{\mathbb{R}^+} \frac{1}{(T^{-1}+\lambda)^2(2T^{-1}+\lambda)^2\dots(2^{k-1}T^{-1}+\lambda)^2} de_{\mathfrak{d}}(\lambda), \tag{1.39}$$

so that by Lemma 1.4 bounds on $\left\langle \phi_{T,k}^2 \right\rangle$ yield bounds on the spectral exponents of $\mathfrak{d}$ (and conversely). Likewise,

$$\left\langle |\nabla\phi_{T,k}|^2 \right\rangle \;\lesssim\; \int_{\mathbb{R}^+} \frac{\lambda}{(T^{-1}+\lambda)^2(2T^{-1}+\lambda)^2\dots(2^{k-1}T^{-1}+\lambda)^2} de_{\mathfrak{d}}(\lambda), \tag{1.40}$$

so that bounds on the spectral exponents of $\mathfrak{d}$ yield bounds on $\langle |\nabla\phi_{T,k}|^2 \rangle$ following the proof of Lemma 1.5.

The second ingredient is the variance estimate of Lemma 1.3, which we apply to $\phi_{T,k}$ noting that $\left\langle \phi_{T,k}^2 \right\rangle = \mathrm{var}\,[\phi_{T,k}]$ since $\langle \phi_{T,k} \rangle = 0$. This allows us to prove that there exists some map F such that

$$\mathrm{var}\,[\phi_{T,k}] \;\lesssim\; F(d,T,k,\left\langle |\nabla\phi_T|^2 \right\rangle, \left\langle |\nabla\phi_{T,1}|^2 \right\rangle, \dots, \left\langle |\nabla\phi_{T,k-1}|^2 \right\rangle, \left\langle |\nabla\phi_{T,k}|^2 \right\rangle). \tag{1.41}$$

We now turn to the a priori estimate. Testing (1.37) with the function $\phi_{T,k+1}$, taking the expectation and using Cauchy-Schwarz inequality, we obtain the nonlinear estimate

$$\left\langle |\nabla\phi_{T,k+1}|^2 \right\rangle \;\lesssim\; \left\langle \phi_{T,k}^2 \right\rangle^{1/2} \left\langle \phi_{T,k+1}^2 \right\rangle^{1/2} = \left\langle \phi_{T,k}^2 \right\rangle^{1/2} \mathrm{var}\,[\phi_{T,k+1}]^{1/2}. \tag{1.42}$$

Combined with Young's inequality, this turns (1.41) into

$$\mathrm{var}\,[\phi_{T,k}] \;\lesssim\; \tilde{F}(d,T,k,\left\langle |\nabla\phi_T|^2 \right\rangle, \left\langle |\nabla\phi_{T,1}|^2 \right\rangle, \dots, \left\langle |\nabla\phi_{T,k-1}|^2 \right\rangle, \left\langle \phi_{T,k-1}^2 \right\rangle) \tag{1.43}$$

for some map $\tilde{F}$.

The induction procedure is as follows. By (1.39) and Lemma 1.4, if for some $k \geq 1$, $\left\langle \phi_{T,k}^2 \right\rangle$ satisfies

$$\left\langle \phi_{T,k}^2 \right\rangle \;\lesssim\; \begin{cases} d < 4k-2 : T^{2k-1-d/2}, \\ d = 4k-2 : \ln T, \\ d > 4k-2 : 1, \end{cases} \tag{1.44}$$

then the spectral exponents of $\mathfrak{d}$ are at least as in Theorem 2 up to dimension $4k-2$. More precisely, for all $\mu > 0$,

$$\int_0^\mu de_{\mathfrak{d}}(\lambda) \lesssim \begin{cases} d < 4k-2 : \mu^{d/2+1}, \\ d = 4k-2 : \mu^{2k} \ln_+ \mu, \\ d > 4k-2 : \mu^{2k}. \end{cases} \tag{1.45}$$

As induction assumption, we assume at step $k \geq 1$ that $\langle \phi_{T,k}^2 \rangle$ satisfies (1.44). The boundedness of $\langle \phi_T^2 \rangle$ from Theorem 1 for $d > 2$ initializes the procedure.

Assume now that $\langle \phi_{T,k}^2 \rangle$ satisfies (1.44). We first appeal to the variance estimate in the form of (1.43). To benefit from this inequality, we use the induction assumption (1.44) at step k to bound all the terms of the form $\langle |\nabla \phi_{T,j}|^2 \rangle$ for $j \leq k$ and $\langle \phi_{T,k}^2 \rangle$. This yields (1.44) for $k+1$, and therefore the optimal spectral exponents up to dimension $4(k+1)-2$, as desired.

To conclude this section we combine Theorem 2 with Lemma 1.5 and estimate (1.34) as follows:

Theorem 3 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let A_{hom} be as* (1.5) *and for all $T > 0$ and $k \geq 1$, let $A_{k,T}$ be given by* (1.33). *Then for all $d \geq 2$, there exists $k_d \in \mathbb{N}$ such that for all $k \geq k_d$*

$$|\xi \cdot (A_{k,T} - A_{\mathrm{hom}})\xi| \lesssim \begin{cases} d = 2 : T^{-1} \ln^q T, \\ d > 2 : T^{-d/2}, \end{cases}$$

for some $q > 0$ depending only α and β.

1.4 PDE approximation of homogenized coefficients: a complete error analysis [GOa,Glo12,GNOb,EGMN]

In this section we present a complete error analysis of two approximation methods of the homogenized coefficients A_{hom} defined in (1.5), starting from the corrector equation (1.4). The first method is based on the regularization of the corrector equation through the approximations $A_{k,T}$ defined in (1.33), whereas the second method is based on the periodization of the random field. In both cases, we give optimal convergence rates.

1.4.1 Analysis of the regularization method

Let $T > 0$, and let $R \geq N \in \mathbb{N}$. As already mentioned in the motivation of Subsection 1.3.1, the regularized corrector ϕ_T can be accurately approximated on the box Q_N by ϕ_T^R, the unique solution to

$$\begin{cases} T^{-1}\phi_T^R - \nabla^* \cdot A(\xi + \nabla \phi_T^R) = 0 \ \text{ in } Q_R \cap \mathbb{Z}^d, \\ \phi_T^R = 0 \ \text{ on } x \in \mathbb{Z}^d \setminus Q_R, \end{cases}$$

provided $R - N \gg \sqrt{T}$. More precisely, there exists $c > 0$ depending only the ellipticity constants α, β and on d, such that provided $R - N \sim R$,

$$\fint_{Q_N \cap \mathbb{Z}^d} |\nabla \phi_T^R - \nabla \phi_T|^2 \lesssim T^{3/2} \exp\left(-c\frac{R-N}{\sqrt{T}}\right). \tag{1.46}$$

Hence, for all $\gamma > 1/2$, the error made by replacing ϕ_T by ϕ_T^R on Q_N is of infinite order in T for $R = N + T^\gamma$. In the rest of this section, we will therefore consider that ϕ_T is a computable quantity.

The starting point of this method is the approximation of A_{hom} by the family $\{A_{k,T}\}$ defined in (1.33). We divide this subsection into three paragraphs. We first explain the general strategy on the approximation A_T of A_{hom}, and shall show that the error between A_{hom} and a computable approximation of A_T splits into two parts: the error $|A_T - A_{\mathrm{hom}}|$ (which we have estimated in the previous section), and a term which takes the form of a variance. We estimate this variance in the second paragraph, and present the global error estimate in the last paragraph.

General strategy

Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed. For all $T > 0$ and $L \in \mathbb{N}$, we define a random approximation A_T^N of A_T by

$$\xi \cdot A_T^N(\omega)\xi \,:=\, \fint_{Q_N \cap \mathbb{Z}^d} (\xi + \nabla\phi_T(x,\omega)) \cdot A(x,\omega)(\xi + \nabla\phi_T(x,\omega))dx,$$

where ϕ_T is the regularized corrector of Definition 1.2 (with $\mu = T^{-1}$). By the ergodic theorem, almost surely,

$$\lim_{N\to\infty} \xi \cdot A_T^N(\omega)\xi \,=\, \xi \cdot A_T\xi.$$

We'd like to quantify this statement, and estimate the standard deviation of A_T^N from A_T as a function of N. Noting that $\langle A_T^N \rangle = A_T$, it amounts to estimating $\langle (\xi \cdot A_T^N\xi - \xi \cdot A_T\xi)^2 \rangle = \mathrm{var}\left[\xi \cdot A_T^N\xi\right]$. Unlike the field A, the stationary field $(x,\omega) \mapsto (\xi + \nabla\phi_T(x,\omega)) \cdot A(x,\omega)(\xi + \nabla\phi_T(x,\omega))$ is correlated. If the correlations were small enough, A_T^N would essentially behave as the average of N^d independent and identically distributed random variables. Hence the best estimate we can hope for is

$$\mathrm{var}\left[\xi \cdot A_T^N\xi\right] \,\lesssim\, N^{-d}. \tag{1.47}$$

A direct calculation shows that

$$\left\langle (\xi \cdot A_T^N\xi - \xi \cdot A_{\mathrm{hom}}\xi)^2 \right\rangle \,=\, \mathrm{var}\left[\xi \cdot A_T^N\xi\right] + (\xi \cdot A_T\xi - \xi \cdot A_{\mathrm{hom}}\xi)^2. \tag{1.48}$$

Hence, the combination of (1.47) with the spectral estimates of Theorem 2 yields

$$\left\langle (\xi \cdot A_T^N \xi - \xi \cdot A_{\mathrm{hom}} \xi)^2 \right\rangle \lesssim \begin{cases} d = 2 : N^{-2} + T^{-2} \ln^2 T, \\ d = 3 : N^{-3} + T^{-3}, \\ d = 4 : N^{-4} + T^{-4} \ln^2 T, \\ d > 4 : N^{-d} + T^{-4}. \end{cases}$$

Recall that we need $R - N \gg \sqrt{T}$ to have this estimate. In particular, at first order, $R \sim N \sim \sqrt{T}$, and the systematic error is of higher order than the random fluctuations up to $d = 8$. Then, the systematic error dominates. This is one of the motivations to use the approximations $A_{k,T}$ for $k > 1$.

The main achievement of this section is the quantification of var $\left[\xi \cdot A_{k,T}^N \xi\right]$ in terms of N. Yet the result only holds provided we slightly modify the definition (1.48) of A_T^N, and smooth out the average using a regular mask. Let $\eta_N : \mathbb{Z}^d \to \mathbb{R}^+$ be a function supported in Q_N, and such that $\int_{\mathbb{Z}^d} \eta_N = 1$ and $\sup |\nabla \eta_N| \lesssim N^{-d-1}$. We define A_T^N by

$$\xi \cdot A_T^N(\omega)\xi := \int_{\mathbb{Z}^d} (\xi + \nabla\phi_T(x,\omega)) \cdot A(x,\omega)(\xi + \nabla\phi_T(x,\omega))\eta_N(x)dx, \tag{1.49}$$

and shall prove that (1.47) holds, with however a logarithmic correction for $d = 2$.

Optimal variance estimate

The rigorous version of (1.47) is as follows:

Theorem 4 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed, and for all $T > 0$, let $\phi_T \in L^2(\Omega)$ be the regularized corrector, unique stationary solution to (1.6) with $\mu = T^{-1}$. For all $N \in \mathbb{N}$ let $\eta_N : \mathbb{Z}^d \to \mathbb{R}^+$ be a non-negative mask supported in Q_N such that $\int_{\mathbb{Z}^d} \eta_N = 1$ and $\sup |\nabla \eta_N| \lesssim N^{-d-1}$, and let define the random variable $\overline{A}_T^N$ by*

$$\xi \cdot \overline{A}_T^N(\omega)\xi := \int_{\mathbb{Z}^d} \Big(T^{-1}\phi_T(x,\omega)^2 + (\xi + \nabla\phi_T(x,\omega)) \cdot A(x,\omega)(\xi + \nabla\phi_T(x,\omega)) \Big) \eta_N(x)dx. \tag{1.50}$$

Then there exists $q > 0$ depending only on α, β and d such that the following two variance estimates hold:

$$\mathrm{var}\left[\xi \cdot \overline{A}_T^N \xi\right] \lesssim \begin{cases} d = 2 : N^{-2}(\ln_+ T)^q, \\ d > 2 : N^{-d}, \end{cases} \tag{1.51}$$

$$\mathrm{var}\left[\int_{\mathbb{Z}^d} \phi_T^2 \eta_N\right] \lesssim \begin{cases} d = 2 : (\ln_+ T)^q, \\ d > 2 : N^{2-d}. \end{cases} \tag{1.52}$$

Note that provided $T \geq N$ (which is compatible with the scaling $N \gg \sqrt{T}$) the combination of (1.51) and (1.52) yields (1.47) with a logarithmic correction for $d = 2$.

The proof of Theorem 4 is based on the variance estimate of Lemma 1.3 combined with Theorem 1. The starting point for (1.51) is the following identity for all $e = (z, z + \mathbf{e}_i)$, $z \in \mathbb{Z}^d$, $i \in \{1, \dots, d\}$:

$$\frac{\partial}{\partial\omega_e}\int_{\mathbb{Z}^d}\Big(T^{-1}\phi_T^2+(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\Big)\eta_N$$
$$=2\int_{\mathbb{Z}^d}(\xi_i+\nabla_i\phi_T(z))\nabla_{z_i}G_T(z,x)\left(\sum_{j=1}^d\omega_{(x-\mathbf{e}_j,x)}\nabla_j^*\eta_N(x)(\xi_j+\nabla_j^*\phi_T(x))\right)dx$$
$$+\eta_N(z)(\xi_i+\nabla_i\phi_T(z))^2.$$

To prove this identity we use that the integrand is an energy density (which is the reason why the zero-order term has been added into the definition of $\overline{A}_T^N$), the susceptibility estimate (1.21), and an integration by parts on the weak formulation of regularized corrector equation (1.6).

If we neglect the supremum in the variance estimate of Lemma 1.3, this yields

$$\mathrm{var}\left[\xi\cdot\overline{A}_T^N\xi\right]\lesssim\sum_{e\in\mathbb{B}}\left\langle\left(\frac{\partial}{\partial\omega_e}\int_{\mathbb{Z}^d}\Big(T^{-1}\phi_T^2+(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\Big)\eta_N\right)^2\right\rangle$$
$$\lesssim\int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}\Big\langle(1+|\nabla\phi_T(z)|^2)|\nabla_zG_T(z,x)||\nabla_zG_T(z,x')||\nabla^*\eta_N(x)||\nabla^*\eta_N(x')|$$
$$\times(1+|\nabla^*\phi_T(x)|)(1+|\nabla^*\phi_T(x')|)\Big\rangle dxdx'dz+\int_{\mathbb{Z}^d}\eta_N(z)^2\left\langle(1+|\nabla\phi_T(z)|^4)\right\rangle dz.$$

Let us assume for simplicity that $\phi_T\in L^\infty(\Omega)$ uniformly in T (which is stronger than what we've proved in Theorem 1), so that $\nabla\phi_T\in L^\infty(\Omega)$ as well. We may then bound the first integrand by:

$$\Big\langle(1+|\nabla\phi_T(z)|^2)|\nabla_zG_T(z,x)||\nabla_zG_T(z,x')||\nabla^*\eta_N(x)||\nabla^*\eta_N(x')|$$
$$\times(1+|\nabla^*\phi_T(x)|)(1+|\nabla^*\phi_T(x')|)\Big\rangle$$
$$\lesssim\langle|\nabla_zG_T(z,x)||\nabla_zG_T(z,x')|\rangle\,|\nabla^*\eta_N(x)||\nabla^*\eta_N(x')|$$
$$\le\left\langle|\nabla_zG_T(z,x)|^2\right\rangle^{1/2}\left\langle|\nabla_zG_T(z,x')|^2\right\rangle^{1/2}|\nabla^*\eta_N(x)||\nabla^*\eta_N(x')|.$$

By stationarity of ∇G_T, we have $\langle|\nabla_zG_T(z,x)|^2\rangle=\langle|\nabla_xG_T(z,x)|^2\rangle$, so that the variance estimate turns into

$$\mathrm{var}\left[\xi\cdot\overline{A}_T^N\xi\right]\lesssim\int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}\left\langle|\nabla_xG_T(z,x)|^2\right\rangle^{1/2}\left\langle|\nabla_{x'}G_T(z,x')|^2\right\rangle^{1/2}$$
$$\times|\nabla^*\eta_N(x)||\nabla^*\eta_N(x')|dxdx'dz+\int_{\mathbb{Z}^d}\eta_N(z)^2dz.$$

Combined with the properties of η_N (this is where the assumption that $|\nabla\eta_N|\lesssim N^{-d-1}$ is crucial), we are left with

$$\mathrm{var}\left[\xi\cdot\overline{A}_T^N\xi\right]\lesssim N^{-2(d+1)}\int_{\mathbb{Z}^d}\int_{Q_N\cap\mathbb{Z}^d}\int_{Q_N\cap\mathbb{Z}^d}\left\langle|\nabla_xG_T(z-x)|^2\right\rangle^{1/2}$$
$$\times\left\langle|\nabla_{x'}G_T(z-x')|^2\right\rangle^{1/2}dxdx'dz+N^{-d}.$$

By Caccioppoli's inequality in space (used on dyadic annuli), this integral behaves as if the Green's function satisfied $|\nabla G_T(x,0)| \lesssim (1+|x|)^{1-d}\exp(-c|x|/\sqrt{T})$, so that we end up with the desired estimate

$$\mathrm{var}\left[\xi\cdot\overline{A}_T^N\xi\right] \lesssim N^{-d}.$$

Yet the assumption that ϕ_T is uniformly bounded in $L^\infty(\Omega)$ does not hold, and we only know that ϕ_T is uniformly bounded in $L^q(\Omega)$ for all $1\le q<\infty$ (with the logarithmic correction in dimension $d=2$). To deal with this, we use Meyers' higher integrability results which ensure that $|\nabla G_T|^{2+\gamma}$ still has the optimal decay as above when integrated on dyadic annuli, for some $\gamma>0$ depending only on α,β and d. This allows us to make the argument presented in the case $\phi_T\in L^\infty(\Omega)$ work with the uniform boundedness provided by Theorem 1), at the expense of a logarithmic correction in dimension $d=2$.

To turn this into a rigorous proof, one also has to deal with the supremum in the variance estimate of Lemma 1.3. We refer to [GOa] for details.

The same strategy allows to prove (1.52) as well.

Complete error analysis

The combination of Theorems 3 and 4 with (1.46) yields the following complete error estimate:

Theorem 5 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let A_{hom} be as* (1.5) *and for all $T>0$ and $k\ge 1$, let $A_{k,T}$ be given by* (1.33)*. For all $N\in\mathbb{N}$ let $\eta_N:\mathbb{Z}^d\to\mathbb{R}^+$ be a non-negative mask supported in Q_N such that $\int_{\mathbb{Z}^d}\eta_N=1$ and $\sup|\nabla\eta_N|\lesssim N^{-d-1}$. We define a random approximation $A_{k,T}^N$ of $A_{k,T}$ by replacing the expectation in* (1.33) *by the average on Q_N with mask η_N. Then for all $d\ge 2$, there exists $k_d\in\mathbb{N}$ such that for all $k\ge k_d$, and all $N\ge 1$,*

$$\left\langle(\xi\cdot(A_{k,T}^N-A_{\mathrm{hom}})\xi)^2\right\rangle \lesssim \begin{cases} d=2: (N^{-2}+T^{-2})\ln^2 T, \\ d>2: N^{-d}+T^{-2}N^{2-d}+T^{-d}, \end{cases}$$

for some $q>0$ depending only α,β.

***Remark* 3** *Combined with the requirement $N\gg\sqrt{T}$, Theorem 5 yields the estimate*

$$\left\langle(\xi\cdot(A_{k,T}^N-A_{\mathrm{hom}})\xi)^2\right\rangle \lesssim \begin{cases} d=2: N^{-2}\ln^q T, \\ d>2: N^{-d}, \end{cases}$$

for some $q>0$ depending only on α and β, provided we take $N^2\gg T\gtrsim N$.

1.4.2 Analysis of the periodization method

In this subsection we turn to a rather popular approximation of A_{hom} by means of periodization of the random medium.

Description of the method and error analysis

In the case of i. i. d. conductivities, it may make sense to approximate the homogenized matrix A_{hom} using an approximate corrector obtained by solving the corrector equation on a bounded domain with periodic boundary conditions. For all $N \geq 1$, we denote by $\phi_{\#}^N$ the unique solution to

$$\begin{cases} -\nabla^* \cdot A(x,\omega)(\xi + \nabla \phi_{\#}^N(x,\omega)) \,=\, 0 \text{ in } Q_N, \\ \phi_{\#}^N \text{ is } Q_N - \text{periodic}, \ \int_{Q_N} \phi_{\#}^N \,=\, 0, \end{cases} \tag{1.53}$$

and define the approximation $A_{\#}^N(\omega)$ of A_{hom} by

$$\xi \cdot A_{\#}^N(\omega)\xi \,=\, \fint_{Q_N \cap \mathbb{Z}^d} (\xi + \nabla \phi_{\#}^N(x,\omega)) \cdot A(x,\omega)(\xi + \nabla \phi_{\#}^N(x,\omega))dx. \tag{1.54}$$

As proved by Owhadi [73], $A_{\#}^N$ converges almost surely to A_{hom} as $N \to \infty$ (in the general case of ergodic conductivities).

This qualitative result is made quantitative in the case of i. i. d. conductivities in the following theorem.

Theorem 6 *Let Ω, $\mathbb{P}$, $\{\theta_z\}_{z\in\mathbb{Z}^d}$, and A be as in Definitions 1, 2, and 3. Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$. Let A_{hom} be as* (1.5) *and for all $N \geq 1$, let $A_{\#}^N$ be defined by* (1.54)*. Then for all $d \geq 2$ for and all $N \geq 1$,*

$$\left\langle (\xi \cdot (A_{\#}^N - A_{\mathrm{hom}})\xi)^2 \right\rangle \,\lesssim\, \begin{cases} d = 2 : N^{-2}\_ +^q T, \\ d > 2 : N^{-d}\_, \end{cases}$$

for some $q > 0$ depending only α, β.

As opposed to Theorem 5, the approximation formula does not depend on the dimension for the periodization method, which is quite satisfactory. Yet the analysis rests on the fact that the probability space is a product space. If not, the method has to be carefully adapted and it is not clear whether this can be done in a general way.

Importance of the product structure

We recall that the probability space is obtained by tensorization: $\Omega = [\alpha,\beta]^{\mathbb{B}}$ and $\mathbb{P} = \nu^{\otimes\mathbb{B}}$ where ν is the common law of the i. i. d. conductivities on $[\alpha,\beta]$, and $\mathbb{B}$ is the graph associated with $\mathbb{Z}^d$ for the "nearest neighbor" relation $x \sim z$ if $|x - z| = 1$.

Let $N \in \mathbb{N}$. For the Q_N-periodization of the medium, the natural probability space is given by tensorization as well: $\Omega_N = [\alpha,\beta]^{\mathbb{B}_N}$ and $\mathbb{P} = \nu^{\otimes\mathbb{B}_N}$ where $\mathbb{B}_N$ is the graph associated with $[0,N]^d \cap \mathbb{Z}^d$ for the "N-periodic nearest neighbor" relation $x \sim z$ if $|x-z| = 1 \,\mathrm{Mod}(N)$. We denote by $\langle\cdot\rangle_N$ the associated expectation. The associated translation group is $\Theta_N = \{\theta_z^N\}_{z\in\mathbb{Z}^d}$, defined on Ω_N by: for all $z \in \mathbb{Z}^d$ and $\mathbb{P}_N$-almost every $\omega \in \Omega_N$, $\theta_z^N\omega : e \in \mathbb{B}_N \mapsto \omega_{e+z\,\mathrm{Mod}(N)}$. Given $f \in L^2(\Omega_N)$, we define the N-stationary extension of

f on $\mathbb{Z}^d \times \Omega_N$ by: for all $z \in \mathbb{Z}^d$ and $\mathbb{P}_N$-almost every $\omega \in \Omega_N$, $f(z,\omega) := f(\theta_z^N \omega)$. The product structure is crucially used there.

This allows us to interpret the solution $\phi_\#^N$ to (1.53) as the N-stationary extension of the unique solution in $L^2(\Omega_N)$ with vanishing expectation $\left\langle \phi_\#^N \right\rangle_N = 0$ to

$$-\,\mathrm{D}_N^* \cdot A(\xi + \mathrm{D}_N\, \phi_\#^N) \;=\; 0, \tag{1.55}$$

where D_N and D_N^* are as in (1.7) with $\{\theta_{\mathbf{e}_i}^N\}_{i\in\{1,\dots,d\}}$ in place of $\{\theta_{\mathbf{e}_i}\}_{i\in\{1,\dots,d\}}$. Therefore, the spectral theory developed in $L^2(\Omega)$ can be straightforwardly adapted to $L^2(\Omega_N)$, and the bounds on the associated spectral measures are independent of N.

The product structure of Ω and Ω_N also ensures a suitable coupling of statistics. In particular, any random variable $f \in L^2(\Omega_N)$ can be seen as a random variable in $L^2(\Omega)$, and we have

$$\langle f \rangle_N \;=\; \langle f \rangle\,.$$

In particular, this ensures that $\left\langle A_\#^N \right\rangle_N = \left\langle A_\#^N \right\rangle$, which will be crucial in the proof of Theorem 6.

Let us conclude by some remarks on the product structure. The spectral gap estimate of Lemma 1.3 is the only other place in this chapter where the product structure of Ω is used. Yet, the proof also holds if the conductivities are correlated provided the correlation length is finite. It is not clear whether this correlated case can be dealt with by the periodization method. In the particular case when the conductivities are obtained by a deterministic function of i. i. d. random variables and do have finite-correlation length, the right way to perform the periodization method is not via (1.53) but via (1.55) where the conductivites are obtained by the same deterministic function, applied this time to the periodized version of the i. i. d. random variables.

Structure of the proof

The proof of Theorem 6 is based on Theorem 5. The starting point is the following triangle inequality:

$$\begin{aligned}\left\langle |\xi \cdot (A_{\mathrm{hom}} - A_\#^N)\xi|^2 \right\rangle^{1/2} \;\le\; \operatorname{var}\left[A_\#^N\right]^{1/2} + \left|\xi \cdot \left\langle A_\#^N - A_{\#,k,T}^N \right\rangle \xi\right| \\ + \left|\xi \cdot (\left\langle A_{\#,k,T}^N \right\rangle - A_{k,T})\xi\right| + |\xi \cdot (A_{k,T} - A_{\mathrm{hom}})\xi|,\end{aligned} \tag{1.56}$$

where for all $T > 0$, $k, N \ge 1$, $A_{k,T}$ is as in (1.33), and the random matrix $A_{\#,k,T}^N$ is defined as follows. For all $T > 0$ and $N \ge 1$, we let $\phi_{\#,T}^N$ be the unique solution to

$$\begin{cases} T^{-1}\phi_{\#,T}^N(x,\omega) - \nabla^* \cdot A(x,\omega)(\xi + \nabla\phi_{\#,T}^N(x,\omega)) \;=\; 0 \text{ in } Q_N, \\ \phi_{\#,T}^N \text{ is } Q_N - \text{periodic.} \end{cases}$$

For all $k \ge 1$, we then define the random matrix $A_{\#,k,T}^N$ as

$$\xi \cdot A^N_{\#,k,T}\xi = \fint_{Q_N} (\xi + \nabla \phi^N_{\#,T}) \cdot A(\xi + \nabla \phi^N_{\#,T}) + T^{-1} \sum_{i=0}^{k-1} \eta_{k,i} \fint_{Q_N} (\phi^N_{\#,2^{-1}T})^2 + T^{-1} \sum_{i=0}^{k-1} \sum_{j>i}^{k-1} \nu_{k,i,j} \fint_{Q_N} (\phi^N_{\#,2^{-i}T})(\phi^N_{\#,2^{-j}T}),$$

where the coefficients $\eta_{k,i}$ and $\nu_{k,i,j}$ are as in (1.35).

We now estimate each term of the r. h. s. of (1.56), and then optimize in T. The first term is a variance. It can be estimated as in Theorem 4, and we have for all $k, N \geq 1$

$$\operatorname{var}\left[\xi \cdot A^N_{\#}\xi\right] \lesssim \begin{cases} d = 2 : N^{-2}(\ln_+ N)^q, \\ d > 2 : N^{-d}, \end{cases} \tag{1.57}$$

for some $q > 0$ depening only on α, β.

The second term and the last terms of the r. h. s. of (1.56) are systematic errors and can be estimated by Theorem 2 and by a variant of Theorem 2 (where spectral analysis is performed in $(L^2(\Omega_N), \mathbb{P}_N)$) together with Lemma 1.5 and (1.34) (and their "periodized" variants). In particular, this shows that for all $d \geq 2$, there exists $k_d \in \mathbb{N}$ such that for all $k \geq k_d$

$$|\xi \cdot (A_{k,T} - A_{\mathrm{hom}})\xi| + |\xi \cdot \langle A^N_{\#,k,T} - A^N_{\#} \rangle \xi| \lesssim \begin{cases} d = 2 : T^{-1} \ln_+^q T, \\ d > 2 : T^{-d/2}, \end{cases}$$

uniformly in $N \in \mathbb{N}$.

The genuinely new term in the analysis is the third term of the r. h. s. of (1.56). This is the only term with relates T to N and makes the optimization in T nontrivial. The zero-order term has been introduced in order to be able to compare $\phi^N_{\#,T}$ to ϕ_T. Again, the difference between $\phi^N_{\#,T}$ to ϕ_T is due to boundary conditions. We shall proceed as for (1.46). The crucial observation is that by N-stationarity,

$$\left\langle \fint_{Q_N} (\xi + \nabla \phi^N_{\#,T}) \cdot A(\xi + \nabla \phi^N_{\#,T}) \right\rangle_N = \left\langle (\xi + \nabla \phi^N_{\#,T}(0)) \cdot A(0)(\xi + \nabla \phi^N_{\#,T})(0) \right\rangle_N,$$

so that by the coupling of the statistics,

$$\left\langle \fint_{Q_N} (\xi + \nabla \phi^N_{\#,T}) \cdot A(\xi + \nabla \phi^N_{\#,T}) \right\rangle = \left\langle (\xi + \nabla \phi^N_{\#,T}(0)) \cdot A(0)(\xi + \nabla \phi^N_{\#,T})(0) \right\rangle$$

as well. Hence,

$$\left| \left\langle \fint_{Q_N} (\xi + \nabla \phi^N_{\#,T}) \cdot A(\xi + \nabla \phi^N_{\#,T}) \right\rangle - \langle (\xi + \nabla \phi_T) \cdot A(\xi + \nabla \phi_T) \rangle \right| = \left| \left\langle (\xi + \nabla \phi^N_{\#,T}) \cdot A(\xi + \nabla \phi^N_{\#,T}) - (\xi + \nabla \phi_T) \cdot A(\xi + \nabla \phi_T) \right\rangle \right|.$$

Likewise, for all $i, j \in \mathbb{N}$,

$$\left|\left\langle \fint_{Q_N} \phi^N_{\#,2^{-i}T}\phi^N_{\#,2^{-j}T}\right\rangle - \langle \phi_{2^{-i}T}\phi_{2^{-j}T}\rangle\right| = \left|\left\langle \phi^N_{\#,2^{-i}T}\phi^N_{\#,2^{-j}T} - \phi_{2^{-i}T}\phi_{2^{-j}T}\right\rangle\right|.$$

Using the regularized corrector equation, the estimate of the third term thus reduces to the estimate of

$$\left\langle|\nabla\phi^N_{\#,T} - \nabla\phi_T|^2\right\rangle \text{ and } \left\langle(\phi^N_{\#,T} - \phi_T)^2\right\rangle$$

Using deterministic estimates on the Green's function, these quantities satisfy

$$\left\langle|\nabla\phi^N_{\#,T} - \nabla\phi_T|^2\right\rangle^{1/2} \lesssim \begin{cases} d=2: \ln^q_+ T\sqrt{T}\exp\left(-c\frac{N}{\sqrt{T}}\right), \\ d>2: \sqrt{T}\exp\left(-c\frac{N}{\sqrt{T}}\right), \end{cases}$$

$$\left\langle(\phi^N_{\#,T} - \phi_T)^2\right\rangle \lesssim \begin{cases} d=2: \ln^q_+ T T^{3/2}\exp\left(-c\frac{N}{\sqrt{T}}\right), \\ d>2: T^{3/2}\exp\left(-c\frac{N}{\sqrt{T}}\right), \end{cases}$$

where $q>0$ depends only on α,β, and $c>0$ depends additionally on d.

We thus end up with

$$\left\langle|\xi\cdot(A_{\mathrm{hom}} - A^N_{\#})\xi|^2\right\rangle^{1/2} \lesssim \begin{cases} d=2: N^{-1}\ln^q_+ T + T^{-1}\ln_+ T + \sqrt{T}\ln^q_+ T\exp\left(-c\frac{N}{\sqrt{T}}\right), \\ d>2: N^{-d/2} + T^{-d/2} + \sqrt{T}\exp\left(-c\frac{N}{\sqrt{T}}\right), \end{cases} \tag{1.58}$$

which implies Theorem 6 taking $T = N/\ln_+ T$.

Numerical validation

In the numerical tests, $d=2$, and each conductivity of $\mathbb{B}$ takes the value $\alpha=1$ or $\beta=4$ with probability $1/2$. In this simple case, the homogenized matrix is given by Dykhne's formula, namely $A_{\mathrm{hom}} = \sqrt{\alpha\beta}\mathrm{Id} = 2\mathrm{Id}$. We have tested the periodization method and the method with Dirichlet boundary conditions, whose approximations $A^N_{\#}$ and $A^{N,N}$ of A_{hom} are given by (1.54) and (1.24), respectively.

We have plotted on Figure 1.1 an approximation of $\left\langle|A^N_{\#} - A_{\mathrm{hom}}|^2\right\rangle^{1/2}$ and an approximation of $\left\langle|A^{N,N} - A_{\mathrm{hom}}|^2\right\rangle^{1/2}$, for which the expectation is replaced by an empirical average. In both cases, the apparent convergence rate of $N \mapsto \left\langle|A^N_{\#} - A_{\mathrm{hom}}|^2\right\rangle^{1/2}$ and $N \mapsto \left\langle|A^{N,N} - A_{\mathrm{hom}}|^2\right\rangle^{1/2}$ to 0 is -1. This is in agreement with Theorem 6, and with the intuition that the use of Dirichlet boundary conditions induces a surface error which scales as N^{-1} in any dimension. Note that $\left\langle|A^N_{\#} - A_{\mathrm{hom}}|^2\right\rangle^{1/2} \leq \left\langle|A^{N,N} - A_{\mathrm{hom}}|^2\right\rangle^{1/2}$ in the tests.

We have also plotted on Figure 1.2 an approximation of the systematic errors $|\left\langle A^N_{\#}\right\rangle - A_{\mathrm{hom}}|$ and $|\left\langle A^{N,N}\right\rangle - A_{\mathrm{hom}}|$, where the expectation has been replaced by an empirical average using $10N^2$ realizations. The apparent convergence rate for the convergence of $N \mapsto |\left\langle A^N_{\#}\right\rangle - A_{\mathrm{hom}}|$ to zero is close to -2, which is in agreement with the second term of (1.58) for the borderline choice $T = N^2$. This is in contrast with the apparent convergence rate for the convergence of $N \mapsto |\left\langle A^{N,N}\right\rangle - A_{\mathrm{hom}}|$ to zero , which is only -1 — as expected due to the surface effect induced by the boundary conditions.

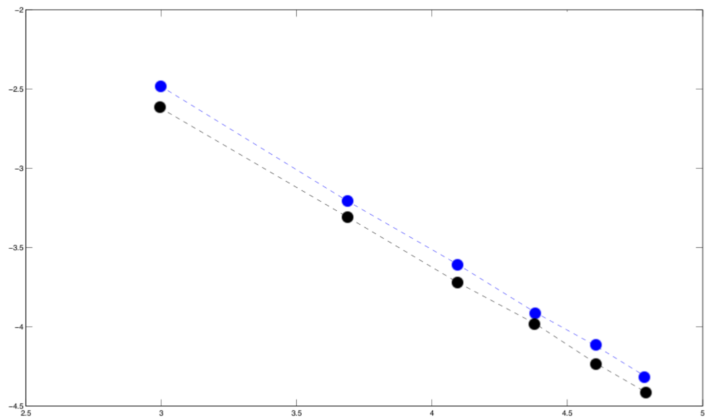

Fig. 1.1. Errors $\left\langle |A^N_\# - A_{\mathrm{hom}}|^2 \right\rangle^{1/2}$ (black) and $\left\langle |A^{N,N} - A_{\mathrm{hom}}|^2 \right\rangle^{1/2}$ (blue) in logarithmic scale

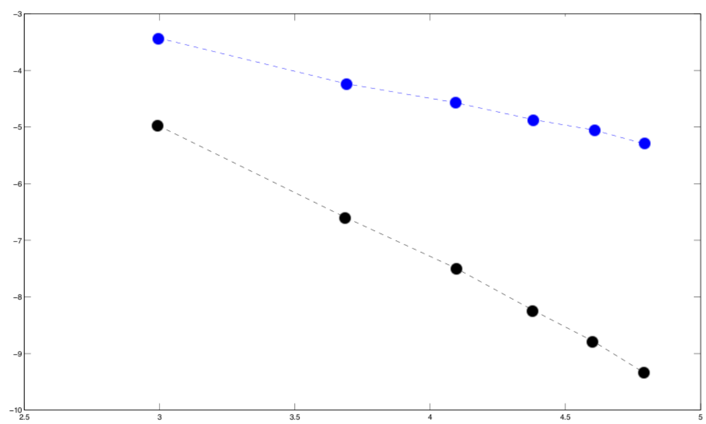

Fig. 1.2. Systematic errors $|\left\langle A^N_\# \right\rangle - A_{\mathrm{hom}}|$ (black) and $|\left\langle A^{N,N} \right\rangle - A_{\mathrm{hom}}|$ (blue) in logarithmic scale

1.5 RWRE approximation of homogenized coefficients: a complete error analysis [GMb,EGMN]

In this section, we come back to the point of view of the random walk in the random environment developed in Subsection 1.1.3, to which we refer for the notation. We first show how information on the spectral exponents of the local drift can be turned into quantitative estimates within the Kipnis-Varadhan theory. We then describe a Monte-Carlo method based on the random walk in discrete time, and apply the quantitative version of the Kipnis-Varadhan theorem to obtain convergence rates on the error between the expectation of the approximation at some time t and the homogenized coefficients. We then complete this analysis by large deviation estimates which quantify the distribution of the approximation at time t around its expectation. Numerical tests confirm the sharpness of the analysis.

1.5.1 Quantitative version of the Kipnis-Varadhan theorem

Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$, $\mathfrak{d} = -\,\mathrm{D}^* \cdot A\xi$ be the local drift, and $e_{\mathfrak{d}}$ be the projection of the spectral measure of $\mathcal{L} = -\,\mathrm{D}^* \cdot A\,\mathrm{D}$ on the local drift. We recall that $\overline{\mathbb{P}}$ is given by (1.11), and that $\overline{\mathbb{E}}$ denotes the associated expectation.

Let X_t be the position at time $t \geq 0$ of a random walker starting at 0 in the random environment with continuous time. By (1.13) one can decompose X_t as the sum $M_t + R_t$ of a martingale M_t with stationary increments and a remainder R_t, which are such that

$$\frac{1}{t}\overline{\mathbb{E}}[(R_t \cdot \xi)^2] \;=\; 2\int_{\mathbb{R}^+} \frac{1-e^{-\lambda t}}{t\lambda^2} de_{\mathfrak{d}}(\lambda) \overset{t\to\infty}{\longrightarrow} 0, \tag{1.59}$$

$$\frac{1}{t}\overline{\mathbb{E}}[(M_t \cdot \xi)^2] \;=\; \sigma^2, \tag{1.60}$$

with $\sigma^2 = 2\xi \cdot A_{\mathrm{hom}}\xi$.

The objective is to show that if the spectral exponents of $\mathfrak{d}$ are at least $(\gamma, -q)$ for some $\gamma > 1$ and $q \geq 0$, then the above results can be turned into

$$\frac{1}{t}\overline{\mathbb{E}}[(R_t \cdot \xi)^2] \lesssim \psi_{\gamma,q}(t) \tag{1.61}$$

$$\left|\frac{1}{t}\overline{\mathbb{E}}[(X_t \cdot \xi)^2] - \sigma^2\right| \lesssim \psi_{\gamma,q}(t), \tag{1.62}$$

where $\psi_{\gamma,q} : (0, +\infty) \to \mathbb{R}^+$ is given by

$$t \mapsto \psi_{\gamma,q}(t) \;=\; \begin{cases} \gamma < 2 : t^{1-\gamma}\ln_+^q(t), \\ \gamma = 2 : t^{-1}\ln_+^{q+1}(t), \\ \gamma > 2 : t^{-1}. \end{cases}$$

Estimate (1.61) directly follows from the spectral formula (1.59) combined with formula (1.31) for $f(\lambda) := \frac{1-e^{-\lambda t}}{\lambda^2}$ and with the assumptions on the spectral exponents.

For the estimate (1.62), we expand the square and appeal to (1.60) and (1.61) to obtain

$$\begin{aligned} \frac{1}{t}\overline{\mathbb{E}}[(X_t \cdot \xi)^2] &= \frac{1}{t}\overline{\mathbb{E}}[(M_t \cdot \xi)^2] + \frac{1}{t}\overline{\mathbb{E}}[(R_t \cdot \xi)^2] + 2\frac{1}{t}\overline{\mathbb{E}}[(M_t \cdot \xi)(R_t \cdot \xi)] \\ &= \sigma^2 + 2\frac{1}{t}\overline{\mathbb{E}}[(M_t \cdot \xi)(R_t \cdot \xi)] + O(\psi_{\gamma,q}(t)). \end{aligned}$$

By reversibility, $(\omega(t-s))_{0\leq s\leq t}$ has the same law under $\overline{\mathbb{P}}$ as $(\omega(s))_{0\leq s\leq t}$. The martingale M_t is unchanged by this time reversal whereas R_t is changed into $-R_t$. Hence, $\overline{\mathbb{E}}[(M_t \cdot \xi)(R_t \cdot \xi)] = 0$ for all $t > 0$, and (1.62) follows. This orthogonality argument already appeared in [29] and [63].

1.5.2 Approximation method

The Monte-Carlo approximation method we have proposed is inspired by [74] and relies on the numerical simulation of the random walk X_t up to some final time t. The combination of (1.62) with the estimates of the spectral exponents of the local drift obtained in Theorem 2 yields

$$\left|\frac{1}{t}\overline{\mathbb{E}}[(X_t\cdot\xi)^2]-2\xi\cdot A_{\mathrm{hom}}\xi\right| \lesssim \begin{cases} d=2: \dfrac{\ln_+^2(t)}{t}, \\ d>2: \dfrac{1}{t}. \end{cases}$$

The numerical simulation of X_t requires the computation of the waiting time associated with the clock. In order to save some computational time we turn to the discrete time random walk Y_t, for which the clock rings at each integer time $t\in\mathbb{N}$. In this case, the Kipnis-Varadhan argument implies that $\sqrt{\varepsilon}Y_{t/\varepsilon}$ converges in law under $\tilde{\overline{\mathbb{P}}}$ (see (1.12)) for the Skorokhod topology to a Brownian motion with covariance $\sigma^2_{\mathrm{disc}}=2\xi\cdot A^{\mathrm{disc}}_{\mathrm{hom}}\xi$, and with $A^{\mathrm{disc}}_{\mathrm{hom}}=\dfrac{A_{\mathrm{hom}}}{\mathbb{E}[p]}$. Since we have better access to $\overline{\mathbb{P}}$ rather than to $\tilde{\overline{\mathbb{P}}}$, we prefer to base our Monter-Carlo on the former, and we consider the computable approximation of $\xi\cdot A^{\mathrm{disc}}_{\mathrm{hom}}\xi$ given for $n\in\mathbb{N}$ and $t>0$ by

$$a_n(t) = \frac{\sum_{k=1}^n p_{\omega^k}(0)(\xi\cdot Y^k_t)^2}{t\sum_{k=1}^n p_{\omega^k}(0)}, \tag{1.63}$$

where $\{\omega^k\}_{k\in\{1,\dots,n\}}$ are independent realizations of the environment, and for all $k\in\{1,\dots,n\}$, Y^k_t is a realization of a random walk in the environment ω^k up to time t. Denoting by $\sigma^2_{t,\mathrm{disc}}$ the expectation of $a_1(t)$, for all $n\geq 1$ we have $\sigma^2_{t,\mathrm{disc}}=\overline{\mathbb{E}}[a_n(t)]$.

There is a quantitative version of the Kipnis-Varadhan theorem in the discrete time setting as well, and we have

$$|\sigma^2_{t,\mathrm{disc}}-\sigma^2_{\mathrm{disc}}| \lesssim \begin{cases} d=2: \dfrac{\ln_+^q(t)}{t}, \\ d>2: \dfrac{1}{t}, \end{cases} \tag{1.64}$$

for some $q>0$ depending only on α,β. The proof of this quantitative version is slightly different than in the continuous time case because M_t and R_t are not orthogonal any longer in the discrete time case. Yet the error to orthogonality can be controlled and the result still holds. Concerning the spectral exponents, it is worth noting that in the discrete time case we work in $L^2(\Omega,\tilde{\mathbb{P}})$ with $\tilde{\mathbb{P}}(\omega)=\frac{p_\omega(0)}{\mathbb{E}[p]}\mathbb{P}(\omega)$. In particular, the elliptic operator $\tilde{\mathcal{L}}$ is now given by

$$\tilde{\mathcal{L}} = -\frac{1}{p_\omega(0)}\,\mathrm{D}^*\cdot A\,\mathrm{D}\,.$$

Likewise the local drift in direction ξ is now given by

$$\tilde{\mathfrak{d}}(\omega) \;=\; -\frac{1}{p_\omega(0)}\,\mathrm{D}^*\cdot A\xi.$$

We have proved that the spectral exponents of $\tilde{\mathcal{L}}$ projected on the local drift $\tilde{\mathfrak{d}}$ are at least

$$\begin{aligned} d=2 &: (2,-q),\\ d>2 &: (2,0), \end{aligned}$$

for some $q>0$ depending only on α,β. This is enough to prove (1.64). To obtain the estimates on the spectral exponents, we may proceed as in Subsection 1.3.4 with however the following modified version of the regularized corrector equation:

$$T^{-1}\phi_T(x,\omega)-\frac{1}{p_\omega(x)}\nabla^*\cdot A(x,\omega)(\xi+\nabla\phi_T(x,\omega)) \;=\; 0,$$

whose dependence with respect to the conductivities ω_e is more complex than in (1.6), so that there are additional terms to deal with compared to the susceptibility estimate (1.21). It is likely that the exponent q is indeed 2 in (1.64), although we have not checked all the details.

Estimate (1.64) is a systematic error which controls the error between the expectation of the computed quantity $a_n(t)$ and its limit $2\xi\cdot A_{\mathrm{hom}}^{\mathrm{disc}}\xi$. As for the methods relying on the corrector equation, one needs to estimate the fluctuations of $a_n(t)$ around its expectation. In this case we are able to prove more than a variance estimate, and we shall indeed give large deviation estimates.

1.5.3 Large deviation estimates

Let us formalize some more the Monte-Carlo method. Let $\omega^1,\omega^2,\dots$ be a countable family of environments, that we denote by $\overline{\omega}$. Let then $Y^1,Y^2,\dots$ be independent random walks evolving in the environments $\omega^1,\omega^2,\dots$ and starting at zero. We write $\mathbf{P}_0^{\overline{\omega}}$ for their joint distribution. The family of environments $\overline{\omega}$ is itself random, and we let $\mathbb{P}^{\otimes}$ be the product distribution with marginal $\mathbb{P}$. We finally set

$$\overline{\mathbb{P}}^{\otimes} \;=\; \mathbb{P}^{\otimes}\mathbf{P}_0^{\overline{\omega}}, \tag{1.65}$$

and denote by $\overline{\mathbb{E}}^{\otimes}$ the associated expectation. Note that $\overline{\mathbb{E}}^{\otimes}[a_n(t)]=\sigma^2_{t,\mathrm{disc}}$.

The fluctuations of $a_n(t)$ around its expectation $\overline{\mathbb{E}}^{\otimes}[a_n(t)]$ is given by

Theorem 7 *Let Ω, $\mathbb{P}$ be as in Definitions 1 and 2, and $\overline{\mathbb{P}}^{\otimes}$ be as in* (1.65)*, and for all $n\geq 1$ and $t>0$ let $a_n(t)$ be given by* (1.63)*. Then there exists $c>0$ such that for all $n\geq 1$ and $t>0$*

$$\overline{\mathbb{P}}^{\otimes}\Big[|a_n(t)-\sigma^2_{t,\mathrm{disc}}|\geq \varepsilon/t\Big] \;\leq\; \exp\Big(-\frac{n\varepsilon^2}{ct^2}\Big).$$

This result is proved using standard tools of large deviation theory (log-Laplace transform) once we are given sharp upper bounds on the transition probabilities of the random walk (which are well-known, see [89, Theorem 14.12]).

The combination of the estimate of the systematic error by the quantitative version of the Kipnis-Varadhan theorem with the large deviation estimates of the fluctuations yields the following complete error estimate for the Monte-Carlo method based on the RWRE.

Theorem 8 *Let Ω, $\mathbb{P}$ be as in Definitions 1 and 2, and $\overline{\mathbb{P}}^{\otimes}$ be as in* (1.65). *For all $n \geq 1$ and $t > 0$, let $a_n(t)$ be given by* (1.63). *There exist $C, c > 0$ such that for all $\varepsilon > 0$ and all t large enough*

$$\overline{\mathbb{P}}^{\otimes}\Big[|a_n(t) - 2\xi \cdot A_{\mathrm{hom}}^{\mathrm{disc}}\xi| \geq (C\mu_d(t) + \varepsilon)/t\Big] \leq \exp\Big(-\frac{n\varepsilon^2}{ct^2}\Big),$$

where $\mu_2(t) = \ln_+^q(t)$ for some $q > 0$ depending only on α, β, and $\mu_d(t) = 1$ for $d > 2$.

1.5.4 Numerical validation

In the numerical tests, each conductivity of $\mathbb{B}$ takes the value $\alpha = 1$ or $\beta = 4$ with probability 1/2. In this simple case, the homogenized matrix is given by Dykhne's formula, namely $A_{\mathrm{hom}} = \sqrt{\alpha\beta}\mathrm{Id} = 2\mathrm{Id}$. For the simulation of the random walk, we generate — and store — the environment along the trajectory of the walk. In particular, this requires to store up to a constant times t data. In terms of computational cost, the expensive part of the computations is the generation of the randomness. In particular, to compute one realization of $a_{t^2}(t)$ costs approximately the generation of $t^2 \times 4t = 4t^3$ random variables. A natural advantage of the method is its full scalability: the t^2 random walks used to calculate a realization of $a_{t^2}(t)$ are completely independent.

We first test the estimate of the systematic error: up to a logarithmic correction, the convergence is proved to be linear in time. In view of Theorem 7, typical fluctuations of $t(a_{n(t)}(t) - \sigma_{t,\mathrm{disc}}^2)$ are of order no greater than $t/\sqrt{n(t)}$, and thus become negligible when compared with the systematic error as soon as the number $n(t)$ of realizations satisfies $n(t) \gg t^2$. We display in Table 1.1 an estimate of the systematic error $|\frac{5}{4}a_{n(t)}(t) - 2|$ obtained with $n(t) = K(t)t^2$ realizations. The systematic error is plotted on Figure 1.3 in function of the time in logarithmic scale. The apparent convergence rate (linear fitting) is $-.85$, which is consistent with (1.64), which predicts -1 and a logarithmic correction.

t	10	20	40	80	160	320	640
$K(t)$	10^5	3000	3000	1000	500	100	20
Systematic error	1.27E-01	7.43E-02	4.17E-02	2.46E-02	1.26E-02	6.96E-03	3.72E-03

Table 1.1. Systematic error in function of the final time t for $K(t)t^2$ realizations.

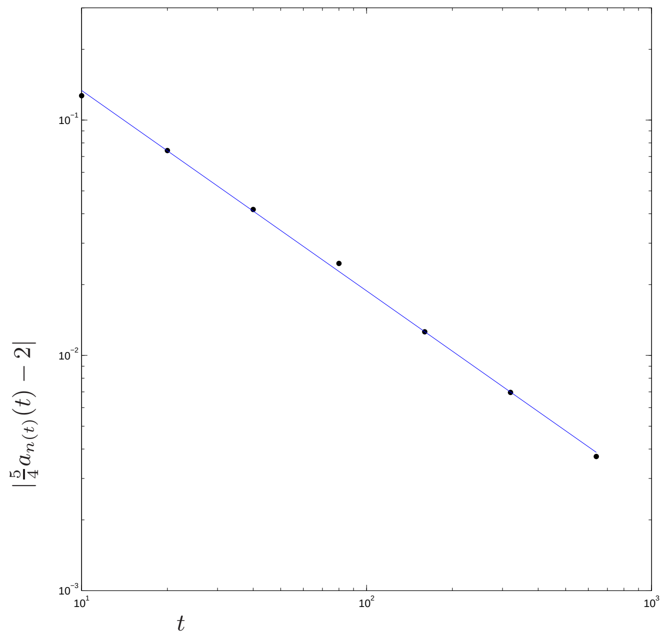

Fig. 1.3. Systematic error in function of the final time t for $n(t) = K(t)t^2$ realizations

We now turn to the random fluctuations of $a_{n(t)}(t)$. Theorem 7 gives us a large deviation estimate which essentially says that the fluctuations of $t(a_{n(t)}(t) - \sigma^2_{t,\mathrm{disc}})$ have a Gaussian tail, measured in units of $t/\sqrt{n(t)}$. The Figures 1.4-1.7 display the histograms of $t(a_{t^2}(t) - \sigma^2_{t,\mathrm{disc}})$ for $t = 10, 20, 40$ and 80 (with 10000 realizations of $a_{t^2}(t)$ in each case). As expected, they look Gaussian.

1.6 Perspectives

First it would be very nice to extend the results of this chapter to the case of systems, since it would shed some light on the approximation methods of Chapter 3 in simplified cases. In [GNOa] we have proposed a proof of Theorem 1 which requires mild estimates on the Green's function. In particular the proof of [GNOa] only relies on Meyers' estimates (which hold for systems as well), and not on the Harnack inequality (which is used in [GOa]). Yet this proof still crucially relies on Cacciopoli's inequality — which does not hold for systems.

A second direction would be to extend the results the case of the random walk on the percolation cluster. There again the proof of Theorem 1 presented in [GNOa] seems easier to generalize since the estimates on the Green's function are brought to a minimum. The difficulty of this generalization comes from the fact that constants in these estimates will depend on the realization of the cluster, so that their control will only be probabilistic (the constants will typically be good with high probability).

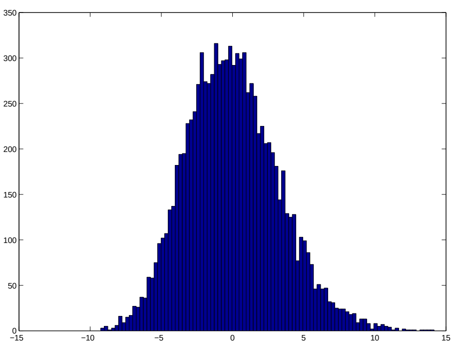

Fig. 1.4. Histogram of the rescaled fluctuations for $t = 10$

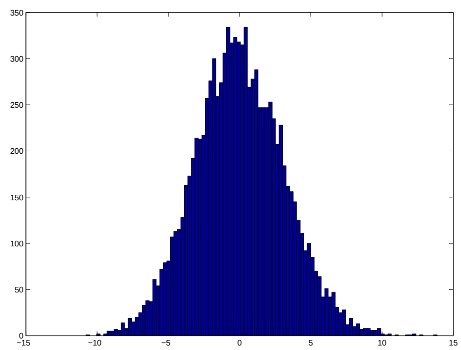

Fig. 1.5. Histogram of the rescaled fluctuations for $t = 20$

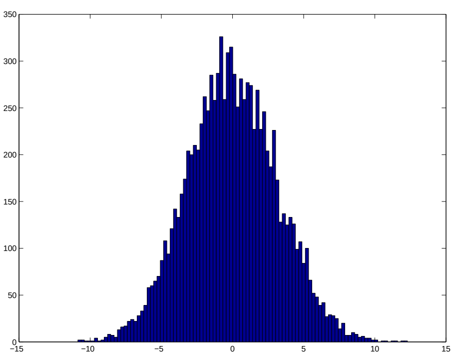

Fig. 1.6. Histogram of the rescaled fluctuations for $t = 40$

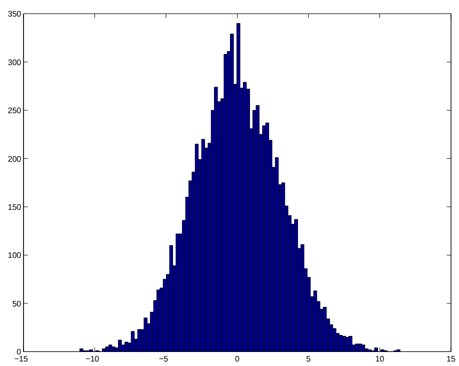

Fig. 1.7. Histogram of the rescaled fluctuations for $t = 80$

A third direction consists in weakening the assumptions on the statistics, namely allow for long-range correlations. A suitable measure of correlation could be the one introduced by Dobrushin and Shlosman in [31, 32] for spin systems since the variance estimate of Lemma 1.3 is inspired by the results of spectral gap, Poincaré, and log-Sobolev inequalities originally developed for spin systems.

A fourth question is concerned with the estimate of the random fluctuations of $A^N_\#$. In Theorem 4 we have estimated the variance. In contrast, for the Monte-Carlo method based on the RWRE, we have proved large deviation estimates. It would be interesting to derive such large deviation estimates for $A^N_\#$ as well. Indeed, numerical experiments in dimension $d = 2$ tend to show that the distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ converges to some Gaussian as $L \to \infty$. We learnt very recently that Nolen successfully solved this question using Chatterjee's approach of Stein's method.

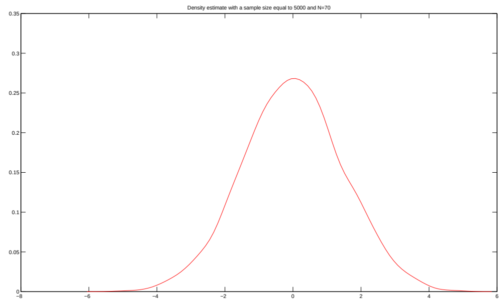

Fig. 1.8. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 70$

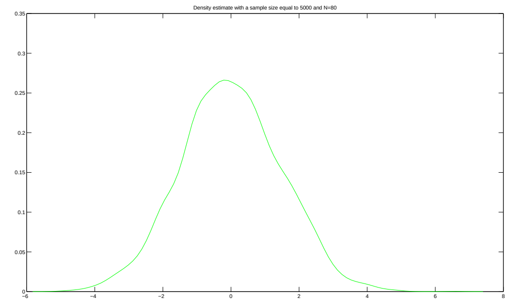

Fig. 1.9. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 80$

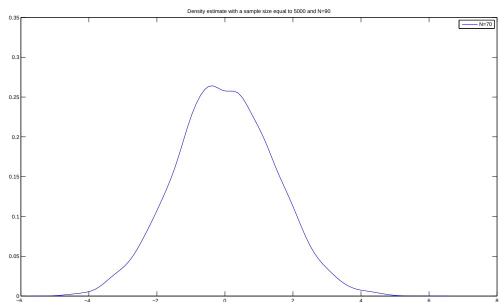

Fig. 1.10. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 90$

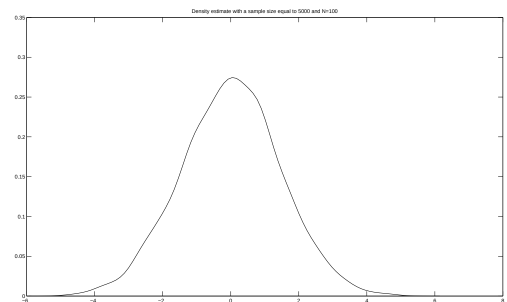

Fig. 1.11. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 100$

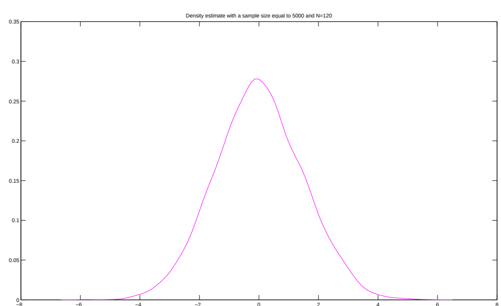

Fig. 1.12. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 120$

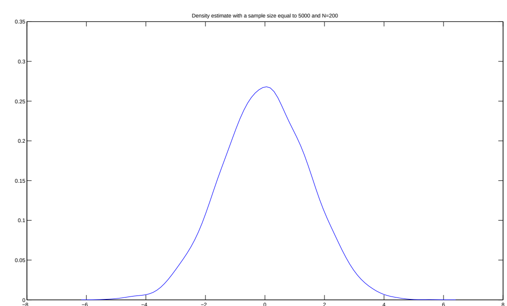

Fig. 1.13. Distribution of $L^{d/2}(A^L_\# - \langle A^L_\# \rangle)$ around its expectation for $L = 200$

2

Quantitative results in homogenization of linear elliptic equations

2.1 Quantitative results in stochastic homogenization [GOc]

In this section, we generalize some results of Chapter 1 to the case of linear elliptic equations. Besides technical adaptations (which are not necessarily obvious), the main achievement of this section is the generalization of the spectral gap estimate of Lemma 1.3 to the continuous setting. In the first subsection we recall standard qualitative results in stochastic homogenization, turn to the spectral gap estimate in the following subsection, then address the existence of stationary correctors in dimension $d > 2$, and conclude by a complete error analysis for the approximation of homogenized coefficients (essentially in dimensions $d = 2, 3$).

The notation and assumptions of this section are fixed in Subsection 2.1.1.

2.1.1 Qualitative stochastic homogenization

In this subsection we recall standard qualitative results in stochastic homogenization of linear elliptic equations. We refer the reader to the original papers [75] by Papanicolaou and Varadhan, and [47] by Kozlov for details (see also the monography [44]).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and denote by $\langle \cdot \rangle$ the associated expectation. We shall say that the family of mappings $(\theta_z)_{z\in\mathbb{R}^d}$ from Ω to Ω is a strongly continuous measure-preserving ergodic translation group if:

- $(\theta_z)_{z\in\mathbb{R}^d}$ has the group property: $\theta_0 = \mathrm{Id}$ (the identity mapping), and for all $x, y \in \mathbb{R}^d$, $\theta_{x+y} = \theta_x \circ \theta_y$;
- $(\theta_z)_{z\in\mathbb{R}^d}$ preserves the measure: for all $x \in \mathbb{R}^d$, and every measurable set $F \in \mathcal{F}$, $\theta_x F$ is measurable and $\mathbb{P}(\theta_x F) = \mathbb{P}(F)$;
- $(\theta_z)_{z\in\mathbb{R}^d}$ is strongly continuous: for any measurable function f on Ω, the function $(\omega, x) \mapsto f(\theta_x \omega)$ defined on $\Omega \times \mathbb{R}^d$ is measurable (with the Lebesgue measure on $\mathbb{R}^d$);
- $(\theta_z)_{z\in\mathbb{R}^d}$ is ergodic: for all $F \in \mathcal{F}$, if for all $x \in \mathbb{R}^d$, $\theta_x F \subset F$, then $\mathbb{P}(F) \in \{0, 1\}$.

Let $\infty > \beta \geq \alpha > 0$, and let $A \in L^2(\Omega, \mathcal{M}^d(\mathbb{R}))$ be such that for $\mathbb{P}$-almost every ω and all $\xi \in \mathbb{R}^d$,

$$\xi \cdot A(\omega)\xi \geq \alpha|\xi|^2, \quad |A(\omega)\xi| \leq \beta|\xi|. \tag{2.1}$$

We define a stationary extension of A (still denoted by A) on $\mathbb{R}^d \times \Omega$ as follows:

$$(x,\omega) \mapsto A(x,\omega) \;=\; A(\theta_x\omega).$$

Homogenization theory ensures that the solution operator associated with $-\nabla \cdot A(x/\varepsilon,\omega)\nabla$ converges as $\varepsilon > 0$ vanishes to the solution operator of $-\nabla \cdot A_{\text{hom}}\nabla$ for $\mathbb{P}$-almost every ω, where A_{hom} is a deterministic elliptic matrix characterized as follows. For all $\xi,\zeta \in \mathbb{R}^d$, and $\mathbb{P}$-almost every ω,

$$\begin{aligned}\xi \cdot A_{\text{hom}}\zeta &= \lim_{R\to\infty} \fint_{Q_R} (\xi + \nabla\phi^\xi(x,\omega)) \cdot A(x,\omega)(\zeta + \nabla\phi^\zeta(x,\omega))dx \\ &= \left\langle (\xi + \nabla\phi^\xi) \cdot A(\zeta + \nabla\phi^\zeta) \right\rangle,\end{aligned}$$

where $\phi^\xi : \mathbb{R}^d \times \Omega \to \mathbb{R}$ is Borel measurable, is such that $\phi^\xi(0,\cdot) \equiv 0$, $\nabla\phi^\xi$ is stationary, and $\phi^\xi(\cdot,\omega) \in H^1_{\text{loc}}(\mathbb{R}^d)$ is almost surely a distributional solution to the corrector equation

$$-\nabla \cdot A(x,\omega)(\xi + \nabla\phi^\xi(x,\omega)) \;=\; 0 \text{ in } \mathbb{R}^d, \tag{2.2}$$

and similarly for ϕ^ζ.

The proof of existence and uniqueness of these correctors is obtained by regularization, and we consider for all $T > 0$ the stationary solution ϕ_T^ξ with zero expectation to the equation

$$T^{-1}\phi_T^\xi(x,\omega) - \nabla \cdot A(x,\omega)(\xi + \nabla\phi_T^\xi(x,\omega)) \;=\; 0 \text{ in } \mathbb{R}^d.$$

This equation has an equivalent form in the probability space, to which we can apply the Lax-Milgram theorem. The functional analysis is however slightly more tedious. In particular, the stochastic counterpart of ∇_i (for $i \in \{1,\dots,d\}$) is denoted by D_i and defined by

$$\mathrm{D}_i\, f(\omega) \;=\; \lim_{h\to 0} \frac{f(\theta_{h\mathbf{e}_i}\omega) - f(\omega)}{h}.$$

These are the infinitesimal generators of the d one-parameter strongly continuous unitary groups on $L^2(\Omega)$ defined by the translations in each of the d directions. These operators commute and are closed and densely defined on $L^2(\Omega)$. We denote by $\mathcal{H}(\Omega)$ the domain of $\mathrm{D} = (\mathrm{D}_1,\dots,\mathrm{D}_d)$. This subset of $L^2(\Omega)$ is a Hilbert space for the norm

$$\|f\|^2_{\mathcal{H}} \;=\; \left\langle |\,\mathrm{D}\, f|^2 \right\rangle + \left\langle f^2 \right\rangle.$$

Since the groups are unitary, the operators are skew-adjoint so that we have the "integration by parts" formula: for all $f,g \in \mathcal{H}(\Omega)$

$$\langle f\, \mathrm{D}_i\, g\rangle \;=\; -\langle g\, \mathrm{D}_i\, f\rangle.$$

The equivalent form of the regularized corrector equation is as follows:

$$T^{-1}\phi_T^\xi - \mathrm{D}\cdot A(\xi + \mathrm{D}\,\phi_T^\xi) \;=\; 0, \tag{2.3}$$

which admits a unique weak solution in $\phi_T^\xi \in \mathcal{H}(\Omega)$, that is such that for all $\psi \in \mathcal{H}(\Omega)$,

$$\left\langle T^{-1}\phi_T^\xi \psi + \mathrm{D}\,\psi \cdot A(\xi + \mathrm{D}\,\phi_T^\xi)\right\rangle \;=\; 0.$$

Following the approach of Subsection 1.1.2, one may prove that $\mathrm{D}\,\phi_T^\xi$ is bounded in $L^2(\Omega)$ and converges weakly in $L^2(\Omega)$ some solution Φ^ξ, which is a gradient. Using then the spectral representation of the translation group we may prove uniqueness of the corrector ϕ^ξ (which is such that $\nabla\phi^\xi = \Phi^\xi$).

Up to here, we have not required A to be symmetric. Let $\mathcal{M}_{\alpha\beta}^d$ denote the set of $d\times d$ real symmetric matrices which satisfy (2.1), and set

$$\mathcal{A} = L^\infty(\mathbb{R}^d, \mathcal{M}_{\alpha\beta}^d). \tag{2.4}$$

In the rest of this section, we shall consider that $A \in L^2(\Omega, \mathcal{M}_{\alpha\beta}^d)$ (although some of the results will hold true for non-symmetric matrices) so that one can appeal to spectral theory. Note that the stationary extension of A belongs to $L^2(\Omega,\mathcal{A})$, and that A_{hom} is also symmetric.

Let $\mathcal{L} = -\,\mathrm{D}\cdot A\,\mathrm{D}$ be the operator defined on $\mathcal{H}(\Omega)$ as a quadratic form. We still denote by $\mathcal{L}$ its Friedrichs extension in $L^2(\Omega)$. This operator is a nonnegative selfadjoint operator. By the spectral theorem it admits a spectral resolution

$$\mathcal{L} \;=\; \int_{\mathbb{R}^+} \lambda G(d\lambda).$$

We shall denote by $e_{\mathfrak{d}}$ the projection of G onto the local drift $\mathfrak{d} \;:=\; -\,\mathrm{D}\cdot A\xi$.

As in the discrete case we have the following useful spectral identities:

$$\xi\cdot A_{\mathrm{hom}}\xi = \langle \xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\frac{1}{\lambda}de_{\mathfrak{d}}(\lambda),$$
$$\left\langle(\phi_T^\xi)^2\right\rangle = \int_{\mathbb{R}^+}\frac{1}{(T^{-1}+\lambda)^2}de_{\mathfrak{d}}(\lambda).$$

2.1.2 The spectral gap estimate in the continuous case

In this subsection we present an extension of the spectral gap estimate of Lemma 1.3 to the continuous setting. To this end we first introduce a specific class of random fields A which generalize the discrete i. i. d. case to the continuous setting.

Let $A \in L^2(\Omega, \mathcal{M}_{\alpha\beta}^d)$. We say that A to $\mathbb{R}^d\times\Omega$ has finite correlation length c_L if for all $x, y \in \mathbb{R}^d$, the random matrices $A(x,\cdot)$ and $A(y,\cdot)$ (understood as the stationary extensions of A) are independent if $|x-y| \geq c_L$. The spectral gap estimate then reads:

Lemma 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\theta_z\}_{z\in\mathbb{R}^d}$ be as in Subsection 2.1.1. Let $A \in L^2(\Omega, \mathcal{M}^d_{\alpha\beta})$ have correlation length $c_L = 1$. Let $X : \mathcal{A} \to \mathbb{R}, M \mapsto X(M)$ be a function — which we shall consider as a random variable when applied to A — with the following regularity property: X can be uniformly approximated by functions $\tilde{X} : \mathcal{A} \to \mathbb{R}$ that depend on M only through the restriction $M_{|\tilde{U}}$ for some compact set $\tilde{U} \subset \mathbb{R}^d$.*
Then we have the following variance estimate

$$\mathrm{var}\left[X(A)\right] \lesssim \left\langle \int_{\mathbb{R}^d} \left(\underset{A_{|Q_3(z)}}{\mathrm{osc}} X(A) \right)^2 dz \right\rangle. \tag{2.5}$$

where $\underset{A_{|Q_3(z)}}{\mathrm{osc}} X(A)$ denotes the oscillation of $X(A)$ with respect to A restricted onto the cube $Q_3(z)$ of lateral size 3 and center at $z \in \mathbb{R}^d$. Note that for some set $U \subset \mathbb{R}^d$, $\underset{A_{|U}}{\mathrm{osc}}\, X(A)$ itself is a random variable:

$$\begin{aligned}\left(\underset{A_{|U}}{\mathrm{osc}}\, X\right)(A) &= \left(\sup_{A_{|U}} X\right)(A) - \left(\inf_{A_{|U}} X\right)(A) \\ &= \sup\left\{X(\tilde{A}) | \tilde{A} \in \mathcal{A},\ \tilde{A}_{|\mathbb{R}^d\setminus U} = A_{|\mathbb{R}^d\setminus U}\right\} \\ &\quad - \inf\left\{X(\tilde{A}) | \tilde{A} \in \mathcal{A},\ \tilde{A}_{|\mathbb{R}^d\setminus U} = A_{|\mathbb{R}^d\setminus U}\right\}. \end{aligned}\tag{2.6}$$

The proof of this spectral gap estimate is similar to the proof of Lemma 1.3. There are two differences between these two lemmas. In the discrete case the regularity assumption on X follows directly from the measurability assumption. Here we have also replaced the Lipschitz control by the oscillation, which is a more general inequality (since the Lipschitz constant controls the oscillation on a bounded domain).

2.1.3 Quantitative results on the corrector equation

The variance estimate of Lemma 2.1 allows us to generalize Theorem 1 to the continuous setting.

Theorem 9 *Let Ω, $\mathbb{P}$, and $\{\theta_z\}_{z\in\mathbb{Z}^d}$ be as in Subsection 2.1.1, and let $A \in L^2(\Omega, \mathcal{M}^d_{\alpha\beta})$ have correlation length $c_L = 1$. For all $q > 0$ there exist $C_{d,q} < \infty$ and $\gamma(q) > 0$ such that for all $T \geq 1$ and $\xi \in \mathbb{R}^d$ with $|\xi| = 1$, the unique stationary solution $\phi_T \in \mathcal{H}(\Omega)$ to (2.3) satisfies*

$$\langle |\phi_T|^q \rangle \leq \begin{cases} d = 2 : C_{2,q}(\ln T)^{\gamma(q)}, \\ d > 2 : C_{d,q}. \end{cases} \tag{2.7}$$

The ingredients of the proof of Theorem 9 are the same of for the proof of Theorem 1: spectral gap estimate, Cacciopoli's inequality in probability, and elliptic regularity theory. Besides technical adaptations due to the fact that we do not have the nice algebraic structure for the susceptility estimate (such as (1.21)), there is a rather important difference in the continuous case. In particular, we do not have the discrete $L^p - L^q$ estimate any longer, which ensures that $\|u\|_{L^p(D\cap\mathbb{Z}^d)} \leq \|u\|_{L^q(D\cap\mathbb{Z}^d)}$ for any domain D of $\mathbb{R}^d$ provided that $p \geq q \geq 1$. Instead we rely on the De Giorgi-Nash-Moser theory and intensively use the Hölder regularity of A-harmonic functions.

As in the discrete case, Theorem 9 implies the existence and uniqueness of a solution $\phi \in \mathcal{H}$ to the corrector equation

$$-\,\mathrm{D}\cdot A(\xi + \mathrm{D}\,\phi) \;=\; 0,$$

whose stationary extension solves (2.2) for almost every $\omega \in \Omega$.

2.1.4 Error analysis for the approximation of homogenized coefficients

In this subsection we only focus on the (random) approximation A_T^L of A_{hom} corresponding to (1.49) in the discrete case, and given for all $L > 0$, $T > 0$, and almost every $\omega \in \Omega$ by

$$\xi \cdot A_T^L(\omega) \;=\; \int_{\mathbb{R}^d} (\xi + \nabla\phi_T(x,\omega)) \cdot A(x,\omega)(\xi + \nabla\phi_T(x,\omega))\eta_L(x)dx,$$

where $\eta_L : \mathbb{R}^d \to \mathbb{R}^+$ is a smooth function supported in Q_L, and such that $\int_{\mathbb{R}^d} \eta_L = 1$ and $\sup|\nabla\eta_L| \lesssim L^{-d-1}$.

As in the discrete case, ϕ_T is a computable approximation of ϕ on Q_L provided the regularized corrector equation is solved on a domain Q_R with $R - L \gg \sqrt{T}$.

Likewise, a direct calculation yields

$$\left\langle \left(\xi \cdot (A_T^L - A_{\mathrm{hom}})\xi\right)^2 \right\rangle \;=\; \mathrm{var}\left[\xi \cdot A_T^L\xi\right] + \left(\xi \cdot (A_T - A_{\mathrm{hom}})\xi\right)^2$$

since $\left\langle \xi \cdot A_T^L\xi \right\rangle = A_T$ by stationarity.

Proceeding as for the discrete case we have proved:

Theorem 10 *Let Ω, $\mathbb{P}$, and $\{\theta_z\}_{z\in\mathbb{Z}^d}$ be as in Subsection 2.1.1, and let $A \in L^2(\Omega, \mathcal{M}^d_{\alpha\beta})$ have correlation length $c_L = 1$. Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed, and for all $T > 0$, let $\phi_T \in \mathcal{H}(\Omega)$ be the regularized corrector, unique weak solution to (2.3). For all $L > 0$ let $\eta_L : \mathbb{R}^d \to \mathbb{R}^+$ be a non-negative mask supported in Q_L such that $\int_{\mathbb{R}^d} \eta_L = 1$ and $\sup|\nabla\eta_L| \lesssim L^{-d-1}$, and let define the random variable $\overline{A}_T^L$ by*

$$\xi\cdot\overline{A}_T^L(\omega)\xi \;:=\; \int_{\mathbb{R}^d} \left(T^{-1}\phi_T(x,\omega)^2 + (\xi+\nabla\phi_T(x,\omega))\cdot A(x,\omega)(\xi+\nabla\phi_T(x,\omega))\right)\eta_L(x)dx. \quad (2.8)$$

Then there exists $q > 0$ depending only on α, β and d such that the following two variance estimates hold:

$$\mathrm{var}\left[\xi \cdot \overline{A}_T^L\xi\right] \lesssim \begin{cases} d = 2 : L^{-2}(\ln_+ T)^q, \\ d > 2 : L^{-d}, \end{cases} \quad (2.9)$$

$$\mathrm{var}\left[\int_{\mathbb{R}^d} \phi_T^2\eta_L\right] \lesssim \begin{cases} d = 2 : (\ln_+ T)^q, \\ d > 2 : L^{2-d}. \end{cases} \quad (2.10)$$

The combination of (2.9) and (2.10) yields

$$\mathrm{var}\left[\xi \cdot A_T^L \xi\right] \lesssim \begin{cases} d=2 : L^{-2}(\ln_+ T)^q, \\ d>2 : L^{-d}, \end{cases} \tag{2.11}$$

provided $T \geq L$.

For the systematic error the starting point is again the spectral formula

$$\xi \cdot (A_T - A_{\mathrm{hom}})\xi \;=\; T^{-2} \int_{\mathbb{R}^+} \frac{1}{(T^{-1}+\lambda)^2 \lambda} de_{\mathfrak{d}}(\lambda),$$

which is also valid in the continuous case. The estimates of the systematic error

$$|\xi \cdot (A_T - A_{\mathrm{hom}})\xi| \lesssim \begin{cases} d=2 : T^{-1}(\ln_+ T)^q, \\ d=3 : T^{-3/2}, \\ d=4 : T^{-2}\ln_+ T, \\ d>4 : T^{-2} \end{cases} \tag{2.12}$$

follow from the combination of a "continuous" version of Lemma 1.5 with:

Theorem 11 *Let Ω, $\mathbb{P}$, and $\{\theta_z\}_{z\in\mathbb{Z}^d}$ be as in Subsection 2.1.1, and let $A \in L^2(\Omega, \mathcal{M}^d_{\alpha\beta})$ have correlation length $c_L = 1$. The spectral exponents of the local drift $\mathfrak{d} = -\,\mathrm{D}^* \cdot A\xi \in L^2(\Omega)$ are at least*

$$\begin{aligned} d=2 &: (2,-q), \\ 2<d\leq 5 &: (d/2+1, 0), \\ d=6 &: (4,-1), \\ d>6 &: (4,0), \end{aligned}$$

for some $q>0$ depending only on α, β.

Theorem 11 is proved using the induction procedure presented in Subsection 1.3.4. Yet we have only checked the first two iterations. Although we have not checked all the details, the induction procedure implemented in its full generality should permit to upgrade these spectral exponents to $(d/2+1,0)$ for all $d>2$, as in the discrete case.

The combination of (2.11) and (2.12) for $T=L$ gives a complete error analysis with an optimal convergence in dimensions $d=2,3$:

$$\left\langle \left(\xi \cdot (A_L^L - A_{\mathrm{hom}})\xi\right)^2 \right\rangle^{1/2} \lesssim \begin{cases} d=2 : L^{-1}(\ln_+ L)^q, \\ d=3 : L^{-2}. \end{cases}$$

Let us conclude this section by a concrete example and by a comment on the periodization method (which we have not analyzed yet in the continuous case).

The simplest example of random diffusion matrix with finite correlation length is as follows. We let $\mathcal{P}$ be a Poisson point process of fixed intensity in $\mathbb{R}^d$. With any realization of $\mathcal{P}$, we associate the diffusion matrix $A \in \mathcal{A}$ defined by: for all $x \in \mathbb{R}^d$, if $\inf\{|x-y|, y \in \mathcal{P}\} < 1/2$ then $A(x) := \alpha\mathrm{Id}$, otherwise $A(x) := \beta\mathrm{Id}$. By definition of the Poisson point process, $A(\mathcal{P})$ has correlation length $c_L = 1$.

Let $R > 0$. In view of the construction of the random diffusion matrix, a natural way to periodize the medium does not consist in taking the periodic replication of $A(\xi)|_{Q_R}$, but rather by periodizing $\xi \cap Q_R$ itself (the periodization of which is denoted by ξ_R) and consider the random diffusion matrix $A(\xi_R)$ — whose realizations are Q_R-periodic by construction.

2.2 Fluctuations of solutions to equations with noisy diffusion coefficients [Glo11c]

2.2.1 Motivation

In this section we contribute to the analysis of a problem posed by Nolen and Papanicolaou in [71]. Let D be a Lipschitz domain, let $f \in L^2(D)$, let A be some deterministic diffusion coefficients, and for all $\varepsilon > 0$ let B_ε be a random diffusion matrix with correlation length ε such that $A + B_\varepsilon \in \mathcal{A}$ (where $\mathcal{A}$ is defined in (2.4)). For all $\varepsilon > 0$, we let $u_\varepsilon \in H_0^1(D)$ be the unique weak solution to

$$-\nabla \cdot (A + B_\varepsilon)\nabla u_\varepsilon \;=\; f. \tag{2.13}$$

The problem consists in estimating the fluctuations of the solution u_ε around its expectation in terms of ε.

In the case when the noise B_ε is replaced by a perturbation b_ε of a zero-order term a such that $a + b_\varepsilon > 0$ almost surely on D, and when $u_\varepsilon \in H_0^1(D)$ is the unique weak solution to

$$(a + b_\varepsilon)u_\varepsilon - \nabla \cdot \nabla u_\varepsilon \;=\; f,$$

the problem of estimating the statistics of u_ε has been addressed in [33] and [3]. In particular, if $\langle b_\varepsilon \rangle = 0$, then for all $x \in D$ the following convergence in distribution holds for $d \leq 3$:

$$\frac{u_\varepsilon(x,\omega) - \langle u_\varepsilon(x,\cdot)\rangle}{\varepsilon^{d/2}} \longrightarrow \mathcal{G}(\sigma(x),\omega), \tag{2.14}$$

where $\mathcal{G}$ is a Gaussian random variable with covariance $\sigma(x)$ (which can be characterized explicitly).

In the discrete setting, a similar problem has been addressed by Conlon and Naddaf [24], who have considered the solution $u_\varepsilon \in L^2(\mathbb{Z}^d)$ to

$$u_\varepsilon - \nabla^* \cdot (\mathrm{Id} + B)\nabla u_\varepsilon \;=\; \varepsilon^2 f_\varepsilon \qquad \text{in } \mathbb{Z}^d,$$

where B is an i. i. d. conductivity function with range in $[\alpha, \beta]$, $\beta \geq \alpha > -1$, and $f_\varepsilon(z) := f(\varepsilon z)$ for some fixed $f \in C_0^\infty(\mathbb{R}^d)$. They have proved two types of results, and estimated both a strong and a weak norm of the fluctuation. The first result is that there exists an exponent $0<\gamma_1\leq 2$, depending only α,β,d and which goes to 2 as the contrast $\frac{1+\beta}{1+\alpha}\to 1$, such that

$$\varepsilon^d\int_{\mathbb{Z}^d}\mathrm{var}\left[u_\varepsilon(z)\right]dz\;\lesssim\;\|f\|^2_{L^2(\mathbb{R}^d)}\varepsilon^{\gamma_1}. \tag{2.15}$$

The second result is a weak measure of the fluctuation (more in the spirit of (2.14)): for all $d\geq 2$ there exists $0<\gamma_2\leq d$, depending only α,β,d and which converges to d as the contrast $\frac{1+\beta}{1+\alpha}\to 1$, such that for all $g\in C_0^\infty(\mathbb{R}^d)$

$$\varepsilon^{2d}\mathrm{var}\left[\int_{\mathbb{Z}^d}u_\varepsilon(z)g_\varepsilon(z)dz\right]\;\lesssim\;\|f\|^2_{L^1(\mathbb{R}^d)}\|g\|^2_{L^\infty(\mathbb{R}^d)}\varepsilon^{\gamma_2}. \tag{2.16}$$

In the regime of small ellipticity ratio, one expects $\gamma_1=2$ and $\gamma_2=d$ (at least for $d>2$). In this section we shall first prove results in the spirit of (2.15) and (2.16) for the solution $u_\varepsilon\in H_0^1(D)$ to (2.13), that is in the continuous setting and on a bounded domain. We will also give improved exponents for (2.15) and (2.16) in the discrete case — which are however not yet optimal.

2.2.2 Main result

We present two results. The first one is a weak and a strong estimate of the fluctuation for solutions on a bounded domain.

Theorem 12 *Let D be a bounded domain of $\mathbb{R}^d$ which satisfies a uniform exterior cone condition. Let $\mathcal{A}$ be as in* (2.4)*. Let A be a conductivity function on D, and B be a stationary random diffusion matrix on $\mathbb{R}^d$ with finite correlation length and such that $A(x)+B(y)\in\mathcal{A}$ for all $x\in D, y\in\mathbb{R}^d$ almost surely. Let $f\in L^2(D)$, and for all $\varepsilon>0$, let $u_\varepsilon\in H_0^1(D)$ denote the unique weak solution to*

$$-\nabla\cdot(A+B_\varepsilon)\nabla u_\varepsilon\;=\;f,\qquad \text{in } D, \tag{2.17}$$

where $B_\varepsilon(\cdot):=B(\cdot/\varepsilon)$.
Then, there exists a Hölder exponent $0<\gamma\leq 1$ depending only on α,β, and d (and which tends to one when $1-\alpha/\beta\to 0$) such that for all $g\in L^\infty(D)$, the fluctuation of u_ε is estimated by:

$$\int_D\mathrm{var}\left[u_\varepsilon(x)\right]dx\lesssim\|f\|^2_{L^2(D)}\begin{cases}d=2:\varepsilon^{2\gamma},\\ d=3:\varepsilon^{1+\min\{1,2\gamma\}},\\ d>3:\varepsilon^2,\end{cases} \tag{2.18}$$

$$\mathrm{var}\left[\int_D g(x)u_\varepsilon(x)dx\right]\lesssim\|f\|^2_{L^2(D)}\|g\|^2_{L^\infty(D)}\,\varepsilon^{d-2(1-\gamma)}. \tag{2.19}$$

Our second result, which is stronger, is an estimate of the fluctuations of the solution to an equation on the whole $\mathbb{R}^d$, for which one can benefit from the stationarity of the Green's function.

Theorem 13 *Let $\mathcal{A}$ be as in (2.4). Let A be a (constant) symmetric matrix, and B be a statianary random diffusion matrix on $\mathbb{R}^d$ with finite correlation-length and such that $A + B(y) \in \mathcal{A}$ almost surely. Let $f \in L^2(\mathbb{R}^d)$, and for all $\varepsilon > 0$, let $u_\varepsilon \in H^1(\mathbb{R}^d)$ denote the unique weak solution to*

$$u_\varepsilon - \nabla \cdot (A + B_\varepsilon)\nabla u_\varepsilon \;=\; f, \qquad \text{in } \mathbb{R}^d,$$

where $B_\varepsilon(\cdot) := B(\cdot/\varepsilon)$.
Then there exist a Meyers exponent $p > 2$ and a Hölder exponent $\gamma > 0$ depending only on α, β and d (and such that $p \to \infty$ and $\gamma \to 1$ when $1 - \beta/\alpha \to 0$), such that the fluctuation of u_ε is estimated by:

$$\int_{\mathbb{R}^d} \mathrm{var}\,[u_\varepsilon(x)]\,dx \lesssim \|f\|^2_{L^2(\mathbb{R}^d)} \begin{cases} d = 2 : \max\{\varepsilon^2|\ln\varepsilon|, \\ \qquad\qquad \varepsilon^{2-(1-\gamma)\max\{0,4-p\}}|\ln\varepsilon|^{\max\{0,4-p\}}\}, \\ d = 3 : \begin{cases} \varepsilon^2\,\varepsilon^{p-3+\frac{\gamma}{1-\gamma}} & \text{for } p < 3 - \frac{\gamma}{1-\gamma}, \\ \varepsilon^2|\ln\varepsilon| & \text{for } p = 3 - \frac{\gamma}{1-\gamma}, \\ \varepsilon^2 & \text{for } p > 3 - \frac{\gamma}{1-\gamma}, \end{cases} \\ d > 3 : \varepsilon^2, \end{cases} \tag{2.20}$$

$$\mathrm{var}\left[\int_{\mathbb{R}^d} g(x)u_\varepsilon(x)dx\right] \lesssim \|f\|^2_{L^2(\mathbb{R}^d)}\|g\|^2_{L^\infty(\mathbb{R}^d)} \begin{cases} d = 2 : \varepsilon^{d-(1-\gamma)\max\{0,4-p\}}|\ln\varepsilon|^{\max\{0,4-p\}} \\ d > 2 : \varepsilon^{d-(1-\gamma)\max\{0,4-p\}}. \end{cases} \tag{2.21}$$

Estimates (2.20) & (2.21) improve estimates (2.18) & (2.19) in the case when A is constant, since $(1-\gamma)\max\{0, 4-p\} < 2(1-\gamma)$. For small ellipticity ratios $\beta/\alpha - 1$, (2.21) is optimal, which is a first step towards the analysis of the disribution of $\varepsilon^{-d/2}\left(\int_{\mathbb{R}^d} g(x)(u_\varepsilon(x) - \langle u_\varepsilon(x)\rangle)dx\right)$, and towards a result of the type (2.14) when the noise is in the diffusion matrix and not in the zero-order term.

2.2.3 Sketch of the proof

We begin with the strong norm of the variance. The starting point is the change of variable $x \mapsto x/\varepsilon$ to make the correlation be of order 1:

$$\int_D \mathrm{var}\,[u_\varepsilon(x)]\,dx \;=\; \varepsilon^d \int_{D/\varepsilon} \mathrm{var}\,[v_\varepsilon(x)]\,dx, \tag{2.22}$$

where v_ε is the weak solution in $H^1_0(D/\varepsilon)$ to

$$-\nabla \cdot C(x)\nabla v_\varepsilon(x) \;=\; \varepsilon^2 f_\varepsilon(x)$$

with $C(x) := A_\varepsilon(x) + B(x)$, and

$$A_\varepsilon(x) = A(\varepsilon x),$$
$$f_\varepsilon(x) = f(\varepsilon x)$$

and is given for all $x \in D/\varepsilon$ by

$$v_\varepsilon(x) \;=\; u_\varepsilon(\varepsilon x).$$

We then use the variance estimate of Lemma 2.1, which yields

$$\mathrm{var}\left[v_\varepsilon(x)\right] \;\lesssim\; \int_{\mathbb{R}^d} \left\langle \underset{B_{|Q(z)}}{\mathrm{osc}^2}\; v_\varepsilon(x) \right\rangle dz. \tag{2.23}$$

We need to estimate the oscillation of v_ε with respect to the coefficient B. Let G_ε be the Green's function associated with $-\nabla\cdot(A_\varepsilon + B)\nabla$ and Dirichlet boundary conditions on D/ε. For $|z-x|\geq 1$, we have

$$\underset{B_{|Q(z)}}{\mathrm{osc}^2}\; v_\varepsilon(x) \;\lesssim\; \int_{Q(z)} |\nabla_y G_\varepsilon(x,y)|^2 dy \int_{Q(z)} |\nabla v_\varepsilon(y)|^2 dy, \tag{2.24}$$

whereas for $|z-x|<1$,

$$\underset{B_{|Q(z)}}{\mathrm{osc}^2}\; v_\varepsilon(x) \;\lesssim\; \int_{Q(z)} |\nabla v_\varepsilon(y)|^2 dy + \varepsilon^4 \int_{Q_4(z)} f_\varepsilon(y)^2 dy. \tag{2.25}$$

The last ingredient is the following Hölder estimate of the gradient of the Green's function: for all $|z-x|>1$,

$$\left(\int_{Q(x)\cap D/\varepsilon} |\nabla_y G_\varepsilon(y,z)|^2 dy\right)^{1/2} \;\lesssim\; \begin{cases} d=2: \dfrac{\ln_+|z-x|}{|z-x|^\gamma}, \\ d>2: \dfrac{1}{|z-x|^{d-2-\gamma}}, \end{cases} \tag{2.26}$$

where γ only depends on α,β and d (see [40, 55]).

The combination of (2.22)–(2.26) then yields

$$\begin{aligned}
\int_D \mathrm{var}\left[u_\varepsilon(x)\right] dx &\lesssim \varepsilon^d \int_{D/\varepsilon} (|\nabla v_\varepsilon(x)|^2 + \varepsilon^4 f_\varepsilon(x)^2) dx \begin{cases} d=2: \varepsilon^{2\gamma-2}, \\ d=3: \varepsilon^{\min\{1,2\gamma\}-1}, \\ d>3: 1, \end{cases} \\
&= \varepsilon^2 \int_D (|\nabla u_\varepsilon(x)|^2 + \varepsilon^2 f(x)^2) dx \begin{cases} d=2: \varepsilon^{2\gamma-2}, \\ d=3: \varepsilon^{\min\{1,2\gamma\}-1}, \\ d>3: 1, \end{cases} \\
&\lesssim \|f\|^2_{L^2(D)} \begin{cases} d=2: \varepsilon^{2\gamma}, \\ d=3: \varepsilon^{1+\min\{1,2\gamma\}}, \\ d>3: \varepsilon^2, \end{cases}
\end{aligned}$$

as desired.

We rewrite the weak norm of the fluctuation as

$$\mathrm{var}\left[\int_D u_\varepsilon(x) g(x) dx\right] = \varepsilon^{2d} \mathrm{var}\left[\int_{D/\varepsilon} v_\varepsilon(x) g_\varepsilon(x) dx\right],$$

where $g_\varepsilon(x) := g(\varepsilon x)$ for all $x \in D$. By the variance estimate of Lemma 2.1,

$$\mathrm{var}\left[\int_D u_\varepsilon(x)g(x)dx\right] \lesssim \varepsilon^{2d}\left\langle \int_{\mathbb{R}^d} \underset{B_{|Q(z)}}{\mathrm{osc}^2}\left(\int_{D/\varepsilon} v_\varepsilon(x)g_\varepsilon(x)dx\right)dz\right\rangle.$$

Since g_ε does not depend on B, we may use the elementary inequality

$$\begin{aligned}\underset{B_{|Q(z)}}{\mathrm{osc}}\left(\int_{D/\varepsilon} v_\varepsilon(x)g_\varepsilon(x)dx\right) &\le \int_{D/\varepsilon}\underset{B_{|Q(z)}}{\mathrm{osc}}\Big(v_\varepsilon(x)g_\varepsilon(x)\Big)dx\\ &\le \int_{D/\varepsilon}\left(\underset{B_{|Q(z)}}{\mathrm{osc}}\; v_\varepsilon(x)\right)|g_\varepsilon(x)|dx,\end{aligned}$$

which turns the variance estimate into

$$\mathrm{var}\left[\int_D u_\varepsilon(x)g(x)dx\right] \lesssim \varepsilon^{2d}\int_{\mathbb{R}^d}\left\langle \int_{D/\varepsilon}\left(\underset{B_{|Q(z)}}{\mathrm{osc}}\; v_\varepsilon(x)\right)|g_\varepsilon(x)|dx\right\rangle^2 dz. \tag{2.27}$$

The estimate (2.19) follows from the combination of (2.27) with (2.24), (2.25), and (2.26).

The proof of Theorem 13 is much more technical than above. The new ingredient there is the stationarity of the Green's function on $\mathbb{R}^d$, that we combine with Meyers' estimates. The starting point is the Green representation formula for smooth r. h. s. f:

$$u_\varepsilon(\varepsilon x) \;=\; \varepsilon^2\int_{\mathbb{R}^d} G_{\varepsilon^2}(x,y)f(\varepsilon y)dy,$$

where G_{ε^2} is the Green's function associated with $(\varepsilon^2 - \nabla\cdot(A+B)\nabla)$ on $\mathbb{R}^d$, from which we deduce

$$\int_{\mathbb{R}^d}\mathrm{var}\left[u_\varepsilon(x)\right]dx \;=\; \varepsilon^{d+4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\varepsilon y)f(\varepsilon y')\mathrm{cov}\left[G_{\varepsilon^2}(x,y);G_{\varepsilon^2}(x,y')\right]dydy'dx,$$

where $\mathrm{cov}\,[\cdot;\cdot]$ denotes the covariance. We shall then make use of the following generalization of the variance estimate: if X and Y satisfy the assumptions of Lemma 2.1, then

$$\mathrm{cov}\left[X(A+B);Y(A+B)\right] \lesssim \int_{\mathbb{R}^d}\left\langle \underset{B_{|Q_3(z)}}{\mathrm{osc}^2}\; X(A+B)\right\rangle^{1/2}\left\langle \underset{B_{|Q_3(z)}}{\mathrm{osc}^2}\; Y(A+B)\right\rangle^{1/2} dz.$$

For simplicity, in the rest of the argument we neglect the singularity of the Green's function (the argument can be made rigorous). We then essentially obtain

$$\begin{aligned}\int_{\mathbb{R}^d}\mathrm{var}\left[u_\varepsilon(x)\right]dx \lesssim \varepsilon^{d+4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\varepsilon y)f(\varepsilon y')\left\langle|\nabla_z G_{\varepsilon^2}(x,z)|^2|\nabla_z G_{\varepsilon^2}(z,y)|^2\right\rangle^{1/2}\\ \left\langle|\nabla_z G_{\varepsilon^2}(x,z)|^2|\nabla_z G_{\varepsilon^2}(z,y')|^2\right\rangle^{1/2}dzdydy'dx.\end{aligned}$$

Let us assume that for all $T > 0$ and $d > 2$, $x \mapsto \nabla_x G_{\varepsilon^2}(x,0)$ has the optimal decay $|x|^{1-d}\exp(-c|x|/\sqrt{T})$ when integrated to power 4 on dyadic annuli — which is a consequence of Meyers' estimate provided $\beta-\alpha$ is small enough. By Cauchy-Schwarz inequality and stationarity of G_{ε^2}, this inequality turns into

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathrm{var}\left[u_\varepsilon(x)\right] dx &\lesssim \varepsilon^{d+4} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\varepsilon y) f(\varepsilon y') \left\langle |\nabla_x G_{\varepsilon^2}(x-z,0)|^4\right\rangle^{1/2} \\
&\qquad \left\langle |\nabla_y G_{\varepsilon^2}(z-y,0)|^4\right\rangle^{1/4} \left\langle |\nabla_{y'} G_{\varepsilon^2}(z-y')|^4\right\rangle^{1/4} dz dy dy' dx \\
&= \varepsilon^{d+4} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left\langle |\nabla_x G_{\varepsilon^2}(x-z,0)|^4\right\rangle^{1/2} \\
&\qquad \left(\int_{\mathbb{R}^d} f(\varepsilon y) \left\langle |\nabla_y G_{\varepsilon^2}(z-y,0)|^4\right\rangle^{1/4} dy\right)^2 dz dx \\
&\le \varepsilon^{d+4} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left\langle |\nabla_x G_{\varepsilon^2}(x-z,0)|^4\right\rangle^{1/2} \\
&\qquad \left(\int_{\mathbb{R}^d} f(\varepsilon y) \left\langle |\nabla_y G_{\varepsilon^2}(z-y,0)|^4\right\rangle^{1/4} dy\right)^2 dz dx.
\end{aligned}$$

We then use the symmetry of the first integrand in x and z, and the Cauchy-Schwarz inequality on the last term to obtain

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathrm{var}\left[u_\varepsilon(x)\right] dx \;&\lesssim\; \varepsilon^4 \left(\int_{\mathbb{R}^d} \left\langle |\nabla_x G_{\varepsilon^2}(x,0)|^4\right\rangle^{1/4} dx\right)^2 \\
&\qquad \times \left(\varepsilon^d \int_{\mathbb{R}^d} f(\varepsilon y)^2 dy\right)\left(\int_{\mathbb{R}^d} \left\langle |\nabla_y G_{\varepsilon^2}(y,0)|^4\right\rangle^{1/2} dy\right).
\end{aligned}$$

From the decay assumption on $x \mapsto \nabla_x G_{\varepsilon^2}(x,0)$ we deduce that the first factor scales as ε^{-2}, the second one as $\|f\|^2_{L^2(D)}$, and the last one as 1. This shows (2.20) for $p \ge 4$ and $d > 2$.

The estimate of the weak norm of the fluctuation can be dealt with in a similar way. In the case when $p < 4$, we cannot proceed exactly this way and combine this approach with the Hölder estimate (2.26).

2.2.4 The discrete case

In the discrete setting, we have:

Theorem 14 *Let a be a constant conductivity function on $\mathbb{B}$, and b be an i. i. d. conductivity function on $\mathbb{B}$ such that the associated conductivity matrices A and B on $\mathbb{Z}^d$ satisfy $A+B \in \mathcal{A}$. Let $f \in C^0(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$ and $g \in C^0_b(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$ (that is, continuous, bounded, and square-integrable), and for all $\varepsilon > 0$, let $u_\varepsilon \in L^2(\varepsilon\mathbb{Z}^d)$ denote the unique solution to*

$$u_\varepsilon - \nabla^*_\varepsilon \cdot (A+B_\varepsilon)\nabla_\varepsilon u_\varepsilon \;=\; f, \qquad \text{in } \varepsilon\mathbb{Z}^d. \tag{2.28}$$

Then there exist a Hölder exponent $0 < \gamma \leq 1$ and a Meyers exponent $p > 2$ depending only on α, β and d (the latter goes to infinity when $1 - \beta/\alpha \to 0$), such that the fluctuation of u_ε is estimated by:

$$\int_{\varepsilon\mathbb{Z}^d} \mathrm{var}\,[u_\varepsilon(x)]\,dx \lesssim \|f\|^2_{L^2(\mathbb{R}^d)} \begin{cases} d = 2: \max\{\varepsilon^2|\ln\varepsilon|, \\ \qquad\qquad \varepsilon^{2-(1-\gamma)\max\{0,4-p\}}|\ln\varepsilon|^{\max\{0,4-p\}}\}, \\ d = 3: \begin{cases} \varepsilon^2\,\varepsilon^{p-3+\frac{\gamma}{1-\gamma}} & \text{for } p < 3 - \frac{\gamma}{1-\gamma}, \\ \varepsilon^2|\ln\varepsilon| & \text{for } p = 3 - \frac{\gamma}{1-\gamma}, \\ \varepsilon^2 & \text{for } p > 3 - \frac{\gamma}{1-\gamma}, \end{cases} \\ d > 3: \varepsilon^2, \end{cases} \tag{2.29}$$

$$\mathrm{var}\left[\int_{\varepsilon\mathbb{Z}^d} g(x)u_\varepsilon(x)dx\right] \lesssim \|f\|^2_{L^2(\mathbb{R}^d)}\|g\|^2_{L^\infty(\mathbb{R}^d)} \begin{cases} d = 2: \varepsilon^{d-(1-\gamma)\max\{0,4-p\}}|\ln\varepsilon|^{\max\{0,4-p\}} \\ d > 2: \varepsilon^{d-(1-\gamma)\max\{0,4-p\}}. \end{cases} \tag{2.30}$$

Theorem 14 improves the results of [24, Theorems 1.2 & 1.3] by Conlon & Naddaf. For the strong measure (2.29) of the fluctuation, we precisely identify the logarithmic correction for $d = 2$, provide with an upper bound independent of the ellipticity ratio for $d = 3$, and prove an optimal estimate for $d > 3$. For the weak measure (2.30) of the fluctuation, the optimal scaling is reached provided p is larger than 4 (whereas the optimal scaling is only met asymptotically in [24, Theorem 1.3]), and for $d > 2$ the estimate gives a non-trivial upper bound uniformly in the ellipticity ratio.

Note that the structure of the proof is completely different than in [24] — which does not rely on the spectral gap estimate.

2.3 About boundary conditions for the "cell-problem" [Glo11d]

2.3.1 Motivation

Let us come back to the numerical approximation of homogenized coefficients in a general framework. The aim of this section is to propose a numerical strategy to approximate homogenized coefficients with the following three features:

- the method should not use the specific structure of the heterogeneities (such as periodicity, independence, etc.);
- the method should converge for "any suitable" structure;
- the method should yield optimal convergence rates for standard cases (at least for periodic structures and stationary structures with finite correlation-length).

Throughout this section the diffusion matrix A is assumed to be symmetric.

2.3.2 General method

Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed. We say that some "heterogeneous" matrix A on $\mathbb{R}^d$ can be homogenized if there exists a distributional solution $\phi \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ to the corrector equation

$$-\nabla \cdot A(\xi + \nabla\phi) \,=\, 0 \qquad \text{in } \mathbb{R}^d, \tag{2.31}$$

and if the following asymptotic formula makes sense:

$$\xi \cdot A_{\mathrm{hom}}\xi \,=\, \mathcal{M}\{(\xi + \nabla\phi) \cdot A(\xi + \nabla\phi)\}, \tag{2.32}$$

where $\mathcal{M}\{\cdot\}$ is the average operator on $\mathbb{R}^d$:

$$\mathcal{M}\{\mathcal{E}\} \,:=\, \lim_{R\to\infty} \frac{1}{|Q_R|} \int_{Q_R} \mathcal{E}(x)dx,$$

and $Q_R := (-R, R)^d$.

To approximate numerically A_{hom}, one needs to approximate the corrector field ϕ and the average operator $\mathcal{M}\{\cdot\}$. These two approximations lead to the so-called resonance error. The simplest way to approximate ϕ and $\mathcal{M}\{\cdot\}$ consists in solving (2.31) on a large domain $Q_R = (-R, R)^d$ (with suitable boundary conditions, say homogeneous Dirichlet), and taking the average of the energy density on Q_R. Doing so, we make at least two errors:

- A geometric error (Q_R is not necessarily a multiple of the unit cell in the periodic case, so that even if the solution on Q_R were the true corrector, the average of the energy density on Q_R would not coincide with its average on a periodic cell);
- An error related to the boundary condition (we do not know a priori what to impose on ∂Q_R, and consequently we make an error on the corrector).

More precisely, A_{hom} is approximated by

$$\xi \cdot A_R\xi \,:=\, \frac{1}{|Q_R|} \int_{Q_R} (\xi + \nabla\phi_R(x)) \cdot A(x)(\xi + \nabla\phi_R(x))dx,$$

where ϕ_R is the unique solution in $H^1_0(Q_R)$ to

$$-\nabla \cdot A(\xi + \nabla\phi_R) \,=\, 0.$$

In the case when A is periodic, the associated error is of the order

$$|A_R - A_{\mathrm{hom}}| \,\sim\, \frac{1}{R}$$

in any dimension. Using oversampling and filtering methods, that is setting

$$\xi \cdot \tilde{A}_R\xi \,:=\, \int_{Q_R} (\xi + \nabla\phi_R(x)) \cdot A(x)(\xi + \nabla\phi_R(x))\eta_R(x)dx,$$

where η_R is typically a smooth non-negative mask (see Definition 6 herafter), we may hope to reduce both sources of the error. However, the overall error is still of order

$$|\tilde{A}_R - A_{\mathrm{hom}}| \,\sim\, \frac{1}{R} \tag{2.33}$$

in any dimension, as already noticed by E & Yue in [90]. Only the prefactor may have been reduced.

In this section, inspired by our analysis of the stochastic case and of the analysis of the periodic case, we propose to treat separately the two sources of error. The geometric error is an error localized at the boundary, and a filtering method with a suitable mask is enough to significantly reduce it. The error we make on the boundary conditions has however non-local effects due to the poor decay of the Green's function of the Laplace operator. To reduce this effect, it is natural as in the stochastic case to add a zero-order term to the equation, which makes the associated Green's function decay exponentially fast. This allows to drastically reduce the spurious effect of the boundary conditions away from a boundary layer. Yet, this modifies the corrector equation and introduces a bias, which has to be quantified (typically using spectral theory). The last task consists in suitably choosing the different parameters at stake.

As a proxy for the corrector field ϕ solution to (2.31), we consider $\phi_{T,R}$, solution to

$$\begin{cases} T^{-1}\phi_{T,R} - \nabla\cdot A(\xi+\nabla\phi_{T,R}) = 0 \text{ in } Q_R, \\ \phi_{T,R} = 0 \text{ on } \partial Q_R, \end{cases} \tag{2.34}$$

where $T>0$ controls the importance of the zero-order term and $R>0$ is the size of the domain Q_R. To deal with the geometric error, we shall make use of high order masks:

Definition 6 *A function $\eta:[-1,1]\to\mathbb{R}^+$ is said to be a filter of order $p\geq 0$ if*

(i) $\eta\in C^p([-1,1])\cap W^{p+1,\infty}((-1,1))$,
(ii) $\int_{-1}^1\eta(x)dx=1$,
(iii) $\eta^{(k)}(-1)=\eta^{(k)}(1)=0$ *for all* $k\in\{0,\dots,p-1\}$.

The associated mask $\eta_L:[-L,L]^d\to\mathbb{R}^+$ in dimension $d\geq 1$ is then defined for all $L>0$ by

$$\eta_L(x) := L^{-d}\prod_{i=1}^d \eta(L^{-1}x_i),$$

where $x=(x_1,\dots,x_d)\in\mathbb{R}^d$.

We then define the family of approximations $\{A^R_{k,T,L}\}_{k\in\mathbb{N}}$ of the homogenized coefficients inspired by (1.36): for $k=0$,

$$\xi\cdot A^R_{0,T,L}\xi := \int_{Q_R}(\xi+\nabla\phi_{T,R}(x))\cdot A(x)(\xi+\nabla\phi_{T,R}(x))\eta_L(x)dx, \tag{2.35}$$

where η_L is as in Definition 6; and for all $k\in\mathbb{N}$,

$$\xi\cdot A^R_{k+1,T,L}\xi := \frac{1}{2^{k+1}-1}(2^{k+1}\xi\cdot A^R_{k,T,L}\xi-\xi\cdot A^R_{k,T/2,L}\xi). \tag{2.36}$$

So defined, the family of approximations does not depend on the particular homogenization structure of A — which makes the numerical method generic, as desired.

2.3.3 Elements of convergence analysis

In this subsection we show that the approximation formulas $A_{k,T,R,L}$ satisfy the last two requirements, namely convergence in general and optimal convergence rates for standard cases.

General convergence result

Let us consider the general case of stationary ergodic media (which covers the typical cases of periodicity, quasi-periodicity, random with finite-correlation length, etc. see [44]). In particular, as recalled in Subsection 2.1.1 it is known in this case that the corrector equation is well-posed, that $\nabla\phi$ is obtained as the (strong) limit of a regularized corrector in the probability space $L^2(\Omega)$, and that we have spectral calculus at our disposal.

By the triangle inequality,

$$|A^R_{k,T,L} - A_{\mathrm{hom}}| \;\le\; |A_{k,T} - A_{\mathrm{hom}}| + |A_{k,T,L} - A_{k,T}| + |A^R_{k,T,L} - A_{k,T,L}|,$$

where the different quantities are defined below. Recall that for all $T>0$ the regularized corrector ϕ_T is solution to

$$T^{-1}\phi_T - \nabla\cdot A(\xi+\nabla\phi_T) \;=\; 0 \qquad \text{in } \mathbb{R}^d.$$

The matrice $A_{k,T}$ are then characterized for all $k\in\mathbb{N}$ by:

$$\begin{aligned}\xi\cdot A_{0,T}\xi &:= \mathcal{M}\{(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\} \;=\; \langle(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\rangle\,,\\ \xi\cdot A_{k+1,T}\xi &:= \frac{1}{2^{k+1}-1}(2^{k+1}\xi\cdot A_{k,T}\xi - \xi\cdot A_{k,T/2}\xi).\end{aligned}$$

Using spectral theory, we recall that we have

$$\begin{aligned}\xi\cdot A_{0,T}\xi &= \langle\xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\frac{1}{T^{-1}+\lambda}de_{\mathfrak{d}}(\lambda),\\ \xi\cdot A_{\mathrm{hom}}\xi &= \langle\xi\cdot A\xi\rangle - \int_{\mathbb{R}^+}\frac{1}{\lambda}de_{\mathfrak{d}}(\lambda),\end{aligned}$$

so that by the Lebesgue dominated convergence theorem, $A_{0,T}$ converges to A_{hom} as $T\to\infty$. Hence, by definition of the family $\{A_{k,T}\}_{k\in\mathbb{N}}$, for all $k\in\mathbb{N}$,

$$\lim_{T\to\infty}\xi\cdot A_{k,T}\xi \;=\; \xi\cdot A_{\mathrm{hom}}\xi. \tag{2.37}$$

Let us turn to the second error term $|A_{k,T}-A_{k,T,L}|$, where for all $k\in\mathbb{N}$,

$$\begin{aligned}\xi\cdot A_{0,T,L}\xi &:= \int_{Q_L}(\xi+\nabla\phi_T(x))\cdot A(x)(\xi+\nabla\phi_T(x))\eta_L(x)dx,\\ \xi\cdot A_{k+1,T,L}\xi &:= \frac{1}{2^{k+1}-1}(2^{k+1}\xi\cdot A_{k,T,L}\xi - \xi\cdot A_{k,T/2,L}\xi).\end{aligned}$$

It is clear by stationarity and definition of η_L that $\langle \xi \cdot A_{k,T,L}\xi\rangle = \xi \cdot A_{k,T}\xi$, so that by the ergodic theorem for all $k \in \mathbb{N}$ and $T > 0$,

$$\lim_{L\to\infty} A_{k,T,L} \;:=\; A_{k,T} \tag{2.38}$$

almost surely.

We finally deal with the last error term $|A^R_{k,T,L} - A_{k,T,L}|$, which we have already estimated in the discrete case in (1.46). Since the arguments only rely on standard deterministic elliptic estimates (which still hold in the continuous setting), they carry over to the general ergodic setting dealt with here, so that we have for all $T > 0, R > L > 0$,

$$|A^R_{k,T,L} - A_{k,T,L}| \;\lesssim\; \left(\frac{R}{L}\right)^{d/2} \left(\frac{R}{R-L}\right)^{d-1/2} T^{3/4} \exp\left(-c\frac{R-L}{\sqrt{T}}\right), \tag{2.39}$$

for some $c > 0$ depending only on α, β and d.

The combination of (2.37)–(2.39) thus yields the following general consistency estimate: for all $k \in \mathbb{N}$,

$$\lim_{T\to\infty} \lim_{(L\to\infty, R-L\to\infty)} A^R_{k,T,L} \;=\; A_{\mathrm{hom}}. \tag{2.40}$$

Convergence rates

We present convergence rates for two standard cases: stochastic with finite-correlation length, and periodic.

Stochastic case with finite correlation length

The convergence analysis in the discrete stochastic case with i. i. d. conductivities has been done in Section 1.4. For the stochastic case with finite correlation-length in the continuous setting, the analysis of convergence of $A^R_{0,T,L}$ has been presented in Subsection 2.1.4. In both cases, provided $R - L \gg \sqrt{T}$ and $R \le T$, we have for all $k \in \mathbb{N}$:

$$\left\langle |A^R_{k,T,L} - A_{\mathrm{hom}}|^2 \right\rangle^{1/2} \;\lesssim\; \begin{cases} d = 2 : L^{-1} \ln_+^q L, \\ d = 3 : L^{-3/2}, \end{cases}$$

for some $q > 0$ depending only on α, β and d. This is the scaling of the central limit theorem, which is the best one can hope for. In higher dimensions, the analysis is only complete in the discrete setting (although we believe this could be adapted to the continuous setting at the expense of technicalities).

Periodic case

The periodic case illustrates quite well the geometric error. The interest of the mask lies in the following lemma:

Lemma 2.2. *Let η be a filter of order $p \geq 0$ according to Definition 6. Let $q > 1$ and let $\phi \in L^q_{\mathrm{loc}}(\mathbb{R}^d)$ be a Q-periodic function. Then, for all $L > 0$, we have*

$$\left| \int_{Q_L} \phi(x)\eta_L(x)dx - \mathcal{M}(\phi) \right| \lesssim L^{-(p+1)} \|\phi\|_{L^q(Q)}, \tag{2.41}$$

where $Q_L = (-L, L)^d$ and $\eta_L(x) := L^{-d} \prod_{i=1}^d \eta(L^{-1}x_i)$, $x = (x_1, \dots, x_d)$. The constant in (2.41) *only depends on p and η.*

This lemma can be proved using a Fourier series expansion, the Parseval identity for $d \geq 2$, and the Hardy-Littlewood inequality for $1 < q < 2$. What Lemma 2.2 shows is that one can approximate at any order of convergence L^{-p} the average of a periodic function by suitable averages on large boxes Q_L, without knowing the period of the function.

The last argument we need to make a complete error analysis of the periodic case is an estimate of the spectral exponents. We recall that in the periodic case, $L^2(\Omega)$ is simply $L^2(Q)$ with Q the period endowed with the Lebesgue measure (or more precisely the torus), and $\mathcal{H}(\Omega) = H^1_\#(Q)$ (the closure in $H^1(Q)$ of Q-periodic smooth functions of $\mathbb{R}^d$ with vanishing average). Let $A \in \mathcal{A}$ be a Q-periodic function. The operator $\mathcal{L}$ is given on $H^1_\#(Q)$ by $-\nabla \cdot A \cdot \nabla$. Let $\xi \in \mathbb{R}^d$ with $|\xi| = 1$ be fixed. Since there is a Poincaré inequality on $H^1_\#(Q)$, the spectral measure has a spectral gap, and the spectral exponents of the local drift $\mathfrak{d} = -\nabla \cdot A\xi \in H^{-1}_\#(Q)$ (the dual space of $H^1_\#(Q)$) are at least $(\gamma, 0)$ for all $\gamma > 1$. Hence, arguing as in Subsection 1.3.3 we have for all $k \geq 0$:

$$|\xi \cdot A_{k,T}\xi - \xi \cdot A_{\mathrm{hom}}\xi| \lesssim T^{-2(k+1)}.$$

The complete error estimate is as follows:

Theorem 15 *Let $d \geq 2$, $A \in \mathcal{A}$ be Q-periodic, η be a filter of order $p \geq 0$, and A_{hom} and $A^R_{k,T,L}$ be the homogenized matrix and its approximation* (2.35)-(2.36) *respectively, for $k \geq 0$, $R^2 \gtrsim T \gtrsim R$, $R \geq L \sim R \sim R - L$. Then, there exists $c > 0$ depending only on α, β and d such that we have*

$$|A^R_{k,T,L} - A_{\mathrm{hom}}| \lesssim L^{-(p+1)} + T^{-2(k+1)} + T^{1/4} \exp\left(-c\frac{R-L}{\sqrt{T}}\right). \tag{2.42}$$

2.3.4 Numerical validation

In this subsection we illustrate Theorem 15 in two regimes: L large, and L "moderately" large (which is the regime where the approximation will be used in practice). In order to reach large values of L with an affordable computational cost, we treat the case of a discrete elliptic equation with periodic coefficients. For moderate values of L, we present a continuous example.

In order to further illustrate the interest of the method, we also display numerical simulations on a simple quasiperiodic example.

Discrete periodic example

The matrix A is $[0,4)^2$-periodic, and sketched on a periodic cell on Figure 2.1. In the example considered, $a(x, x+\mathbf{e}_1)$ and $a(x, x+\mathbf{e}_2)$ represent the conductivities 1 or 100 of the horizontal edge $[x, x+\mathbf{e}_1]$ and the vertical edge $[x, x+\mathbf{e}_2]$ respectively, according to the colors on Figure 2.1. The homogenization theory for such discrete elliptic operators is similar to the continuous case (see for instance [88] in two dimensions, and [1] in the general case). By symmetry arguments, the homogenized matrix associated with A is a multiple of the identity. It can be evaluated numerically (note that we do not make any other error than the machine precision). Its numerical value is $A_{\mathrm{hom}} = 26.240099009901\dots$.

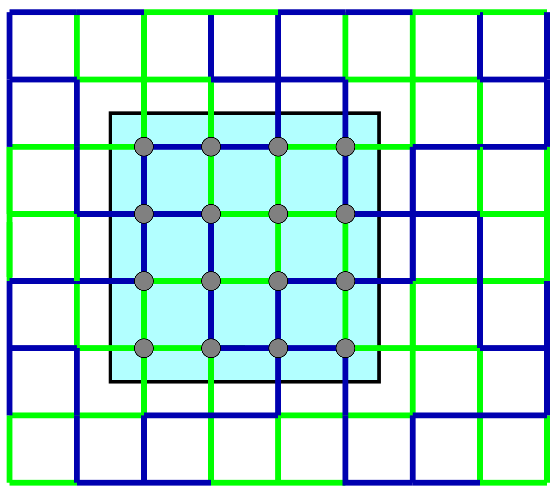

Fig. 2.1. Periodic cell in the discrete case

We have considered the first two approximations formulas $A^R_{0,T,L}$ and $A^R_{1,T,L}$ of A_{hom}. In all the cases treated, we've taken $L = R/3$. For the approximation $A^R_{0,T,L}$, we have tested the following parameters:

- Four values for the zero-order term: $T = \infty$ (no zero-order term), $T \sim R$, $T \sim R^{3/2}$, and $T \sim R^{7/4}$;
- Two different filters: orders $p = 0$ (no filter) and $p = \infty$.

For the approximation $A^R_{1,T,L}$, we have tested the following parameters:

- One value of the zero-order term: $T \sim R^{3/2}$;
- Filter of infinite order $p = \infty$.

The predictions of Theorem 15 in terms of convergence rate of $A^R_{k,T,L}$ to A_{hom} in function of R are gathered and compared to the results of numerical tests in Table 2.1. More details are also given on Figures 2.2–2.5, where the overall error

$$\mathrm{Error}(k,T,R) := |A_{\mathrm{hom}} - A^R_{k,T,L,}|$$

is plotted in log scale in function of R.

Table 2.1. Order of convergence: predictions and numerical results.

	$T=\infty$		$T\sim R$		$T\sim R^{3/2}$		$T\sim R^{7/4}$	
k=0	pred.	test	pred.	test	pred.	test	pred.	test
p = 0	1	1	1	1	1	1	1	1
p= ∞	1	1	2	2	3	3.1	3.5	3.4
k=1					pred.	test		
p= ∞					6	6		

Let us quickly comment on the values of T in Figures 2.2–2.5. For the three dependences of T upon R, we have chosen the prefactors so that their values roughly coincide for $2R = 25$ (that is for 25 periodic cells per dimension):

$$T = 2R/25,$$
$$T = (8R)^{3/2}/1000,$$
$$T = (8R)^{7/4}/5000.$$

The numerical result confirm the analysis, and perfectly illustrate the specific influences of the three parameters k, p and T.

Continuous periodic example

We consider the following matrix A:

$$A(x) = \left(\frac{2+1.8\sin(2\pi x_1)}{2+1.8\cos(2\pi x_2)} + \frac{2+\sin(2\pi x_2)}{2+1.8\cos(2\pi x_1)}\right) \mathrm{Id}, \tag{2.43}$$

used as benchmark tests in [42] and [10]. In this case, $\alpha \simeq 0.35$, $\beta \simeq 20.5$, and $A_{\mathrm{hom}} \simeq 2.75\,\mathrm{Id}$. We take $L = R/3$, $T = R/5$ and a filter of order 2. The global error $|A^R_{0,T,L} - A_{\mathrm{hom}}|$ and the error without zero order term and without filtering are plotted on Figures 2.6 & 2.7. Without zero-order term, the convergence rate is R^{-1} as expected, and the use of a filtering method reduces the prefactor but does not change the rate. With the zero-order term and the filtering method, the apparent convergence rate is R^{-3} (note that the asymptotic theoretical rate R^{-2} is not attained yet), which coincides with the convergence rate associated with filters of order 2 (cf. Lemma 2.2). This is in agreement with the tests in the discrete case, and confirms the analysis.

Continuous quasiperiodic example

The last series of tests is dedicated to a quasiperiodic example. We consider the following coefficients used in [10]:

$$A(x) = \begin{pmatrix} 4+\cos(2\pi(x_1+x_2))+\cos(2\pi\sqrt{2}(x_1+x_2)) & 0 \\ 0 & 6+\sin^2(2\pi x_1)+\sin^2(2\pi\sqrt{2}x_1) \end{pmatrix}. \tag{2.44}$$

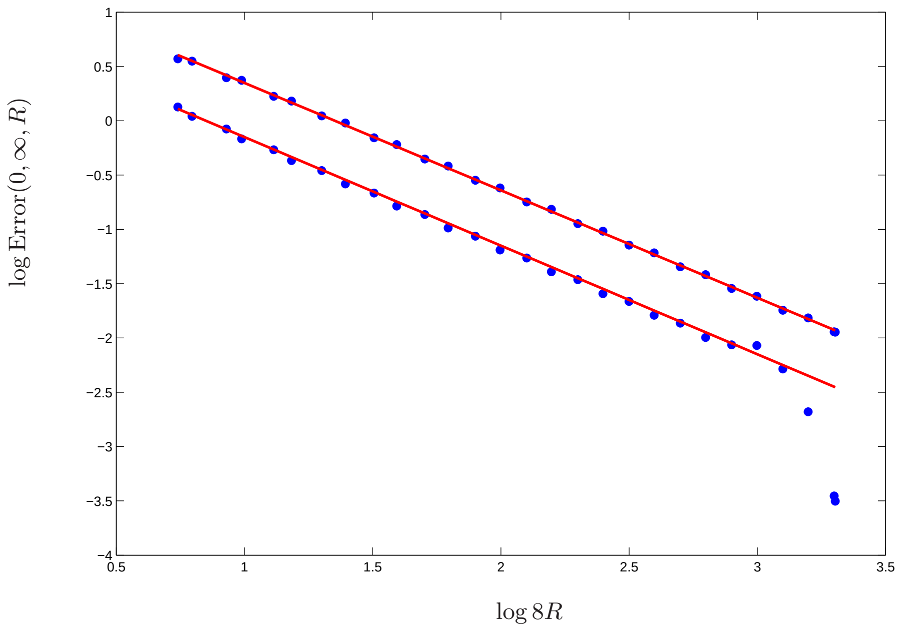

Fig. 2.2. Absolute error in log scale without zero order term, no filter (slope -1), infinite order filter (slope -1, better prefactor).

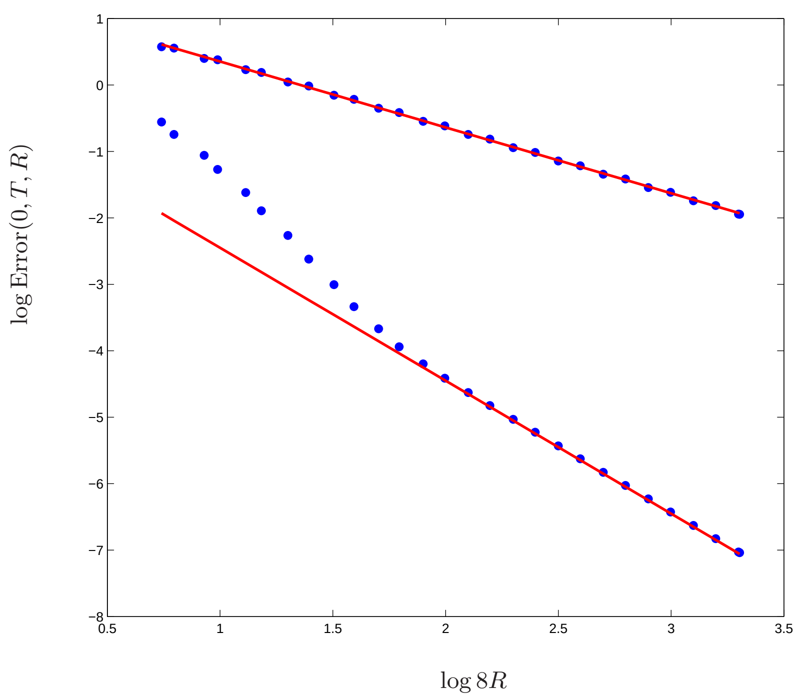

Fig. 2.3. Absolute error in log scale for $T = 2R/25$, no filter (slope -1), infinite order filter (slope -2).

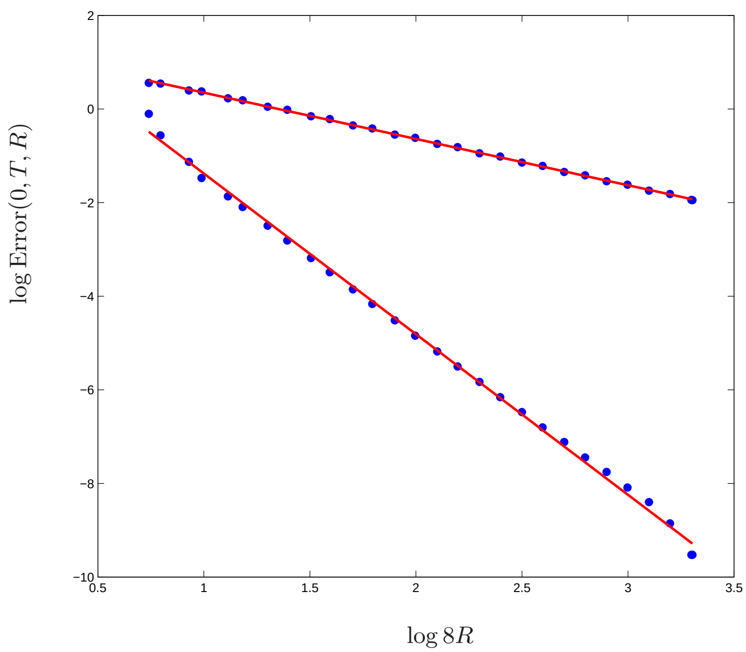

Fig. 2.4. Absolute error in log scale for $T = (8R)^{7/4}/5000$, no filter (slope -1), infinite order filter (slope -3.4).

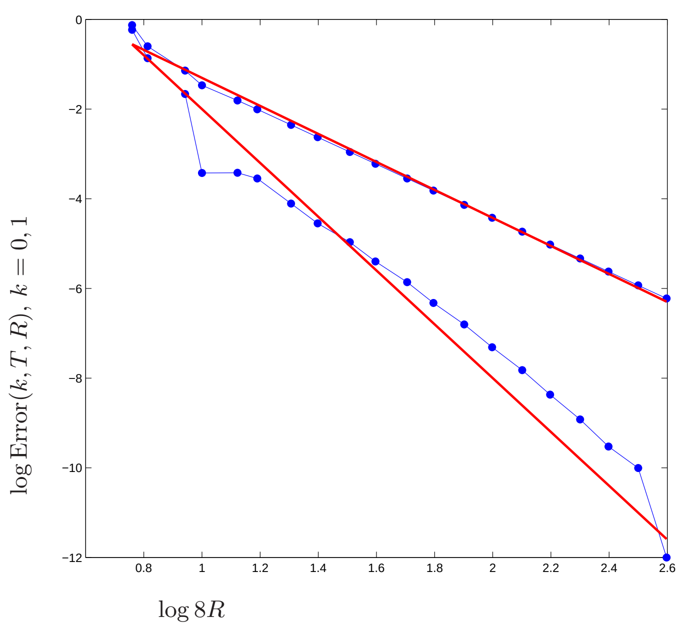

Fig. 2.5. Absolute error in log scale for $T = (8R)^{3/2}/1000$, $A^R_{0,T,L}$ (slope -3.1) and $A^R_{1,T,L}$ (slope -6), filter of infinite order.

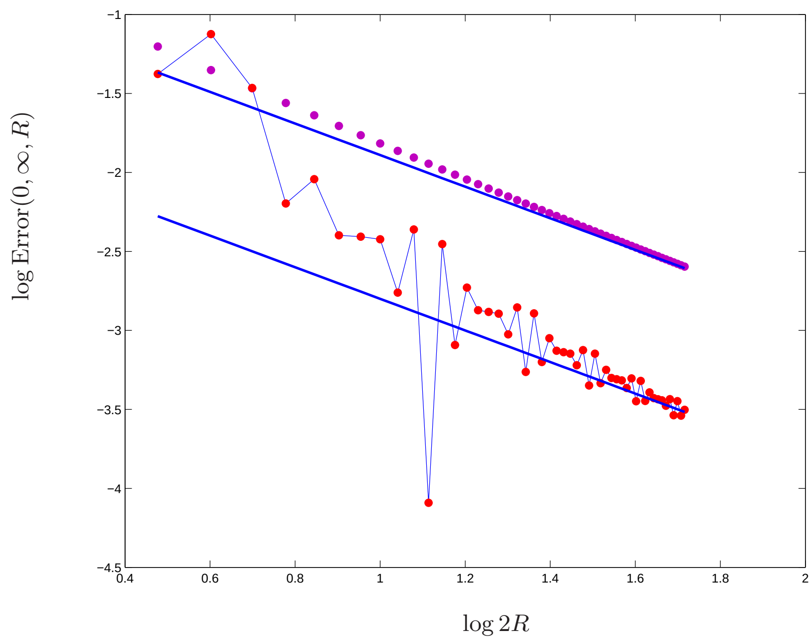

Fig. 2.6. Error in log-log for (2.43) in function of the number of cells per dimension $2R \in [3, 52]$ without zero-order term, with and without filtering: Slope -1 in both cases.

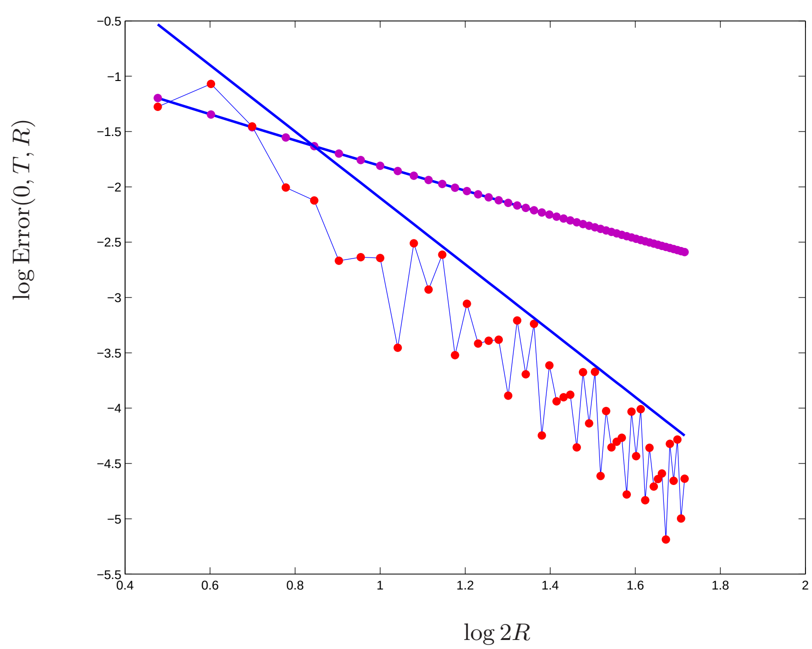

Fig. 2.7. Error in log-log for (2.43) in function of the number of cells per dimension $2R \in [3, 52]$ with a zero-order term $T = R/5$, with and without filtering: Slopes -1 and -3.

In this case, the homogenized coefficients are not easy to compute. They can only be extrapolated. We have taken for the approximation of the homogenized coefficients (that we call coefficient of reference) the output of the computation with $k = 0$, $T = R/50$ and $2R = 52$. Although this may introduce a bias in favour of the proposed strategy, it can be checked a posteriori: the method without zero-order term and without filtering is expected to converge at a rate R^{-1}. This is effectively what we observe on Figure 2.8 using this coefficient of reference. Instead, if we use as a reference the output of the computation for $2R = 52$ without zero-order term nor filtering, then we observe a super-linear convergence which is artificial (see Figure 2.8). With the proposed method, as can be seen on Figure 2.9, the rate of convergence seems to be much better (the slope of the straight line is -5). The reason for this fact is that there should be a spectral gap in this case as well.

2.4 “Real-life” homogenization: a radioactive example [GGK]

In this section we address a problem suggested by the French agency for nuclear waste storage. The problematics there is to understand how nuclear waste can spread in a highly heterogeneous storage device. At first approximation, we consider the medium to be periodic. We shall see that even in this simple setting, practical homogenization remains a challenging problem.

2.4.1 Modeling of a nuclear waste depository

We consider the numerical treatment of a nonlinearly coupled elliptic-parabolic system of equations involving coefficients varying on a fast scale, as encountered in nuclear waste storage devices. Resolving the finest scales induces a prohibitive numerical cost, both in terms of computational time and memory storage. Our goal consists in finding relevant “averaged” models, combined with efficient numerical methods (in particular in order to evaluate the coefficients of the effective equations obtained). A strong motivation is the modeling of radionuclide transport in nuclear waste storage devices. The realization of routine simulations should rely on fast computations, which excludes to resolve the finest scales. Homogenization is the natural tool to derive effective models, which hopefully smooth out in a consistent way the small scale features of the problem. In the case of the nonlinearly coupled system treated here, (periodic) homogenization alone is not enough to drastically reduce the computational cost, since a cell-problem has to be solved at each Gauss point of the computational domain — this could surprise the expert: although diffusion coefficients are assumed to be periodic, and the equations are linear, the nonlinear coupling condition makes the homogenized diffusion matrix depend on the space variable. This is where the reduced basis (RB) method comes into the picture: these cell-problems can be viewed as a d-parameter (in dimension d) family of elliptic equations, which is an ideal setting for the RB method. A further practical issue is related to the dependence of the elliptic operator upon the parameters, which is not affine (according to the terminology of the RB approach) and therefore requires a specific treatment. This section is devoted

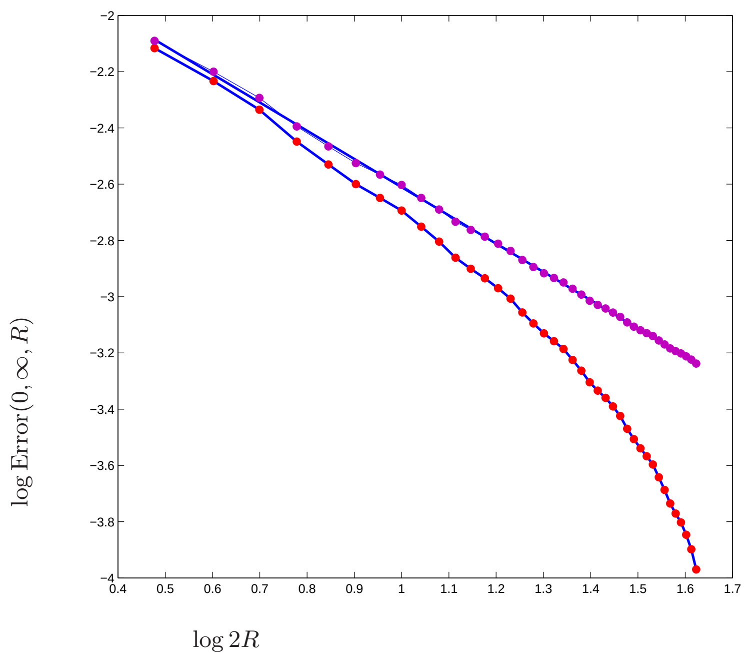

Fig. 2.8. Error in log-log for (2.44) in function of the number of cells per dimension $2R \in [3, 42]$ without zero-order term and without filtering, for the two different coefficients of reference: Slope -1 and artificial super-linear convergence.

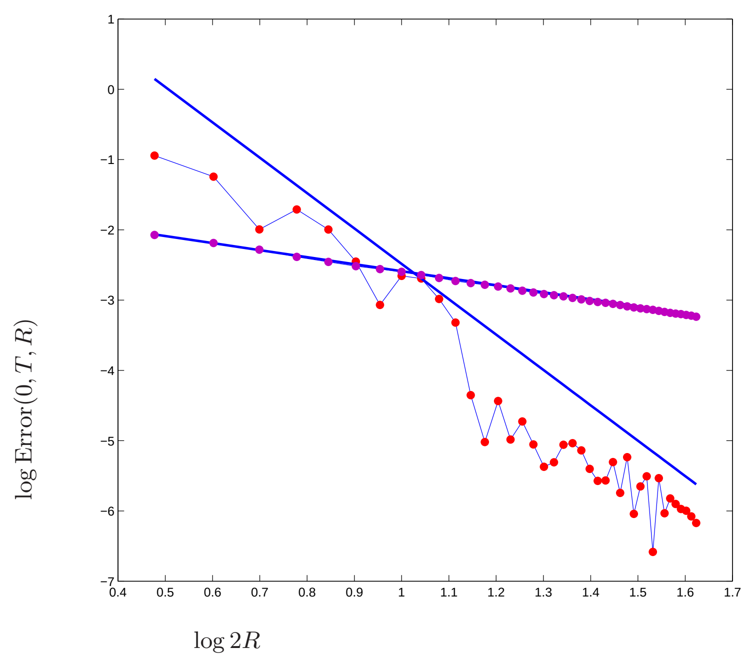

Fig. 2.9. Error in log-log for (2.44) in function of the number of cells per dimension $2R \in [3, 42]$ with a zero-order term $T = R/100$, with and without filtering: Slopes -1 and -5.

to the homogenization and numerical approximation of solute transfer equations driven by diffusion and convection, with a coupling with the Darcy law.

2.4.2 Periodic homogenization of a coupled system of elliptic/parabolic equations

In this section Ω denotes a bounded Lipschitz open domain of $\mathbb{R}^d$ (and not a probability space).

We consider the following weakly coupled system of PDEs:

$$\begin{cases} U = -K\nabla\Theta & \text{in } \Omega, \\ \nabla\cdot U = q & \text{in } \Omega, \\ \partial_t C - \nabla\cdot(D(U)\nabla C - UC) + \lambda C = S & \text{in }]0,T[\times\Omega. \end{cases} \tag{2.45}$$

We let $\lambda > 0$, and for the source terms we let $q \in L^\infty(\Omega)$ and $S \in L^2(0,T;H^{-1}(\Omega))$. The weak coupling condition reads

$$D(U)(x) := D_0(x) + \alpha|U(x)|\mathrm{Id} + \beta\frac{U(x)\otimes U(x)}{|U(x)|}, \tag{2.46}$$

for a. e. $x \in \Omega$, where $\alpha > 0, \beta \geq 0$. The functions $x \mapsto K(x)$ and $x \mapsto D_0(x)$ are matrix-valued; they both satisfy uniform bounds and a strong ellipticity condition, uniformly over Ω: namely, there exists $\Lambda > 0$ such that for a. e. $x \in \Omega$ and all $\xi \in \mathbb{R}^d$

$$\begin{aligned} |K(x)\xi| &\leq \Lambda|\xi|, \ \xi\cdot K(x)\xi \geq \Lambda^{-1}|\xi|^2, \\ |D_0(x)\xi| &\leq \Lambda|\xi|, \xi\cdot D_0(x)\xi \geq \Lambda^{-1}|\xi|^2. \end{aligned}$$

The system (2.45) is completed by boundary conditions and an initial condition, which for simplicity we take as follows (mixed Neumann-Dirichlet boundary conditions could be considered as well)

$$\begin{cases} \Theta = 0 & \text{on } \partial\Omega, \\ C(0,\cdot) = C_{\text{init}} & \text{in } \Omega, \\ C = 0 & \text{on }]0,T[\times\partial\Omega, \end{cases} \tag{2.47}$$

for some $C_{\text{init}} \in L^2(\Omega)$.

We are interested in the case when K is an ε-periodic matrix, and $\varepsilon \to 0$. Before we turn to this problem, we first define a notion of weak solution for the coupled system (2.45)–(2.47), and give an existence and uniqueness result.

Definition 7 *A weak solution of* (2.45)–(2.47) *is a pair* $(\Theta, C) \in H_0^1(\Omega)\times L^2(0,T;H_0^1(\Omega))\cap C^0(0,T;L^2(\Omega))$ *such that* $\partial_t C \in L^2(0,T;H^{-1}(\Omega))$, $\int_0^T\int_\Omega \nabla C\cdot D(U)\nabla C < \infty$ *with* $U = -K\nabla\Theta$, *and which satisfies* (2.45)–(2.47) *in the following sense:*

- Θ *is a weak solution in* $H_0^1(\Omega)$ *to* $(2.45)_{1,2}$ *&* $(2.47)_1$*;*

- *For all* $v \in L^2(0,T;H_0^1(\Omega)) \cap L^2(0,T;L^\infty(\Omega))$ *such that* $\int_0^T \int_\Omega \nabla v \cdot D(U)\nabla v < \infty$, *we have*

$$\int_0^T \langle \partial_t C, v\rangle_{H^{-1},H_0^1} + \int_0^T \int_\Omega \nabla v \cdot D(U)\nabla C + \int_0^T \int_\Omega vU \cdot \nabla C + \int_0^T \int_\Omega Cv(q+\lambda) \;=\; \int_0^T \langle S, v\rangle_{H^{-1},H_0^1} .$$

The following theorem states the existence and uniqueness of such weak solutions.

Theorem 16 *For all* $q \in L^\infty(\Omega)$, $S \in L^2(0,T;H^{-1}(\Omega))$, *and* $C_{\mathrm{init}} \in L^2(\Omega)$, *there exists a unique weak solution to* (2.45)–(2.47) *in the sense of Definition 7.*

We now turn to the periodic homogenization of (2.45)–(2.47). Let K be a $\mathbb{Y} = (0,1)^d$-periodic matrix. For all $\varepsilon > 0$, we consider the coupled system

$$\begin{cases} U_\varepsilon = -K_\varepsilon \nabla \Theta_\varepsilon & \text{in } \ \Omega, \\ \nabla \cdot U_\varepsilon \,=\, q & \text{in } \ \Omega, \\ \partial_t C_\varepsilon - \nabla \cdot (D(U_\varepsilon)\nabla C_\varepsilon - U_\varepsilon C_\varepsilon) + \lambda C_\varepsilon \,=\, S & \text{in } \ ]0,T[\times \Omega, \\ \Theta_\varepsilon \,=\, 0 & \text{on } \ \partial\Omega, \\ C_\varepsilon\,(0,\cdot) \,=\, C_{\mathrm{init}} & \text{in } \ \Omega, \\ C_\varepsilon \,=\, 0 & \text{on } \ ]0,T[\times\partial\Omega, \end{cases} \tag{2.48}$$

where q, S, C_{init} and the function D are as above, and K_ε is defined by $K_\varepsilon(x) := K(x/\varepsilon)$ on Ω. Theorem 16 ensures the existence and uniqueness of a weak solution $(\Theta_\varepsilon, C_\varepsilon)$ of (2.48) for any $\varepsilon > 0$. The following result characterizes the asymptotic behavior of $(\Theta_\varepsilon, C_\varepsilon)$ as $\varepsilon \to 0$.

Theorem 17 *Let* $q \in L^\infty(\Omega)$, $S \in L^2(0,T;H^{-1}(\Omega))$, $C_{\mathrm{init}} \in L^2(\Omega)$, D *be as in* (2.46), *and* K *be a* $\mathbb{Y}$*-periodic bounded and strongly elliptic matrix. For all* $\varepsilon > 0$, *we set* $K_\varepsilon := K(\cdot/\varepsilon)$. *Then the unique weak solution* $(\Theta_\varepsilon, C_\varepsilon)$ *to* (2.48) *converges to some* (Θ_0, C_0) *in the following senses: strongly in* $L^2(\Omega)$ *and* $L^2((0,T)\times\Omega)$, *and weakly in* $H^1(\Omega)$ *and* $L^2(0,T;H^1(\Omega))$. *In addition* C_ε *converges in* $C^0([0,T], L^2(\Omega) -$ *weak) to* C_0, *and* (Θ_0, C_0) *is the unique weak solution to*

$$\begin{cases} U_0 = -K^* \nabla \Theta_0 & \textit{in } \ \Omega, \\ \nabla \cdot U_0 \,=\, q & \textit{in } \ \Omega, \\ \partial_t C_0 - \nabla \cdot (D^* \nabla C_0 - U_0 C_0) + \lambda C_0 \,=\, S & \textit{in } \ ]0,T[\times \Omega, \\ \Theta_0 \,=\, 0 & \textit{on } \ \partial\Omega, \\ C_0\,(0,\cdot) \,=\, C_{\mathrm{init}} & \textit{in } \ \Omega, \\ C_0 \,=\, 0 & \textit{on } \ ]0,T[\times\partial\Omega, \end{cases} \tag{2.49}$$

where K^* *is a constant matrix, and* $D^*(x)$ *is a function of* $\nabla\Theta_0(x)$ *which we define below. To this aim we need to introduce auxiliary quantities. For all* $i \in \{1,\dots,d\}$, *we let* φ_i *denote the unique periodic weak solution in* $H^1_\#(\mathbb{Y})$ *to*

$$-\nabla \cdot K(\mathbf{e}_i + \nabla\varphi_i) \,=\, 0. \tag{2.50}$$

The matrix K^* is strongly elliptic and characterized by: for all $i, j \in \{1, \ldots, d\}$

$$\mathbf{e}_j \cdot K^* \mathbf{e}_i \;=\; \int_{\mathbb{Y}} (\mathbf{e}_j + \nabla \varphi_j) \cdot K(\mathbf{e}_i + \nabla \varphi_i). \tag{2.51}$$

This allows to uniquely define the funtion Θ_0 through the homogenized Darcy equation $(2.49)_{1,2}$. The homogenized drift is then given by

$$U_0 \;=\; -K^* \nabla \Theta_0.$$

It remains to define D^*. Let $\tilde{U}, \tilde{D}$ be defined on $\Omega \times \mathbb{Y}$ as follows:

$$\tilde{U}(x, y) = -K(y)(\mathrm{Id} + \nabla \varphi(y)) \nabla \Theta_0(x), \tag{2.52}$$

$$\tilde{D}(x, y) = D_0(x) + \alpha |\tilde{U}(x, y)| \mathrm{Id} + \beta \frac{\tilde{U}(x, y) \otimes \tilde{U}(x, y)}{|\tilde{U}(x, y)|} = D_0(x) + \mathbb{D}(\tilde{U}(x, y)), \tag{2.53}$$

where $\varphi = (\varphi_1, \ldots, \varphi_d)$. For all $i \in \{1, \ldots, d\}$ and a. e. $x \in \Omega$, we let $\Phi_i(x, \cdot)$ denote the unique periodic weak solution in $H^1_\#(\mathbb{Y})$ to the elliptic equation parametrized by x:

$$-\nabla_y \cdot \tilde{D}(x, y)(\mathbf{e}_i + \nabla_y \Phi_i(x, y)) \;=\; 0.$$

The homogenized coefficients D^* are then characterized by: for all $x \in \Omega$ and all $i, j \in \{1, \ldots, d\}$,

$$\mathbf{e}_j \cdot D^*(x) \mathbf{e}_i \;=\; \int_{\mathbb{Y}} (\mathbf{e}_j + \nabla_y \Phi_j(x, y)) \cdot \tilde{D}(x, y)(\mathbf{e}_i + \nabla_y \Phi_i(x, y))\, dy. \tag{2.54}$$

Although the diffusion D^* is not of the form (2.46), for all $x \in \Omega$, $D^*(x)$ only depends on $\nabla \Theta_0(x)$, and $D^* \in L^2(\Omega)$. Hence existence and uniqueness of weak solutions for the homogenized system can be proved the same way as for Theorem 16. From the homogenization point of view, Theorem 17 is a rather direct application of two-scale convergence and Theorem 16. Although $D(U)$ is unbounded, it is square-integrable and the homogenized system remains elliptic-parabolic (for the homogenization of elliptic equations with unbounded coefficients which are not equi-integrable, nonlocal effects may appear, see [7] and [16]). In the case of strong coupling, homogenization has been proved in [20]. Yet [20] is an overkill for the problem under consideration (uniqueness is not discussed in [20] though), and a more direct proof can be done.

2.4.3 Numerical strategy: reduced basis method

There are essentially three steps to solve numerically (2.49):

1. the computation of K^* and the approximation of Θ_0. The latter is solution of a standard elliptic equation once K^* is known, see $(2.49)_{1,2}$.
2. the approximation of $D^*(x)$ at every Gauss point x of Ω. This requires to solve a family of elliptic equations on the periodic cell $\mathbb{Y}$, parametrized by the Gauss points x via $\nabla \Theta_0(x)$.
3. To find the solution of the advection-diffusion equation $(2.49)_3$.

The bottleneck of the numerical approximation of (2.49) in terms of computational cost is the approximation of D^*.

The homogenization limit

To illustrate Theorem 17, we consider a numerical test suggested by ANDRA[1]. We take $d = 2$ and let $\Omega = (0,2)^2$ be a square domain, and $[0,T]$ be the time interval with $T = 1$. The permeability is defined on the domain $\mathbb{Y} = (0,1)^2$ by:

$$\forall y = (y_1, y_2) \in \mathbb{Y},\ \forall y_1 \in (0,1),\ K(y_1, y_2) = \begin{cases} 4.94064, \text{ if } y_2 \geq 0.5, \\ 0.57816, \text{ if } y_2 < 0.5. \end{cases}$$

We consider mixed Dirichlet-Neumann boundary conditions, and compare approximations of the solution $(\Theta_\varepsilon, C_\varepsilon)$ to the heterogeneous system (2.48) to approximations of the solution (Θ_0, C_0) to the homogenized system (2.49), for several values of ε We display in Table 2.2 the $L^2(\Omega)$ norm of the error $\Theta_0 - \Theta_\varepsilon$ and the $L^2(\Omega \times (0,T))$-norm of the error $C_0 - C_\varepsilon$ for $\varepsilon \in \{0.2, 0.1, 0.05, 0.025\}$. These results are obtained using FreeFem++ (see [36]). The linear systems are solved with a direct solver. We obtain a first order of convergence for both errors. As can be seen on Table 2.2 the apparent convergence rates are of order 1,

ε	$\frac{\lVert\Theta_0 - \Theta_\varepsilon\rVert_{L^2}}{\lVert\Theta_0\rVert_{L^2}}$	Rate	$\frac{\lVert C_0 - C_\varepsilon\rVert_{L^2(L^2)}}{\lVert C_0\rVert_{L^2(L^2)}}$	Rate
0.2	1.667e-3	-	5.528e-4	-
0.1	8.095e-4	1.04	2.525e-4	1.13
0.05	3.992e-4	1.02	1.270e-4	0.99
0.025	1.983e-4	1.01	6.704e-5	0.92

Table 2.2. Error in function of ε

which is consistent with a formal two-scale expansion, and shows the interest of replacing $(\Theta_\varepsilon, C_\varepsilon)$ by its homogenized counterpart (Θ_0, C_0).

The reduced basis method

The reduced basis method was introduced for the accurate online evaluation of (outputs of) solutions to a parameter-dependent family of elliptic PDEs. The basis of the method and further references can be found in [57]. The application to the homogenization of elliptic equations is discussed in [12]. Abstractly, it can be viewed as a method to determine a “good” N-dimensional space $\mathcal{S}_N$ to be used in approximating the elements of a set $\mathcal{F} = \left\{ \left(\overline{\Phi}_1(\xi), ..., \overline{\Phi}_d(\xi)\right),\ \xi \in \mathscr{P} \right\}$ of parametrized elements lying in a Hilbert space $\mathcal{S}$, the parameter ξ ranging a certain subset $\mathscr{P} \subset \mathbb{R}^n$.

Let us describe how the computation of the effective coefficients we are concerned with enters such a framework. First of all, the auxiliary function Θ_0 is simply determined by solving the problem $(2.49)_{1,2}$, with effective coefficients obtained by solving the cell equations (2.50). There is no difficulty in this step and $\nabla_x \Theta_0$ can be considered as given in this

[1] Agence nationale pour la gestion des déchets radioactifs — http://www.andra.fr

discussion. Then, we write the coefficient (2.53) for the concentration equation $(2.49)_3$ as follows

$$\tilde{D}(x,y) = \widehat{\mathscr{D}}(\nabla_x \Theta_0(x))(y)$$

where $\xi \in \mathbb{R}^d \mapsto \widehat{\mathscr{D}}(\xi) \in L^\infty(\mathbb{Y}, \mathcal{M}^d)$ is defined by

$$\begin{aligned} &\widehat{\mathscr{D}}(\xi)(y) \;=\; D_0 + \alpha|M(y)\xi|\mathrm{Id} + \beta\frac{M(y)\xi \otimes M(y)\xi}{|M(y)\xi|} = D_0 + \mathbb{D}(M(y)\xi),\\ &M(y) = K(y)(\mathrm{Id} + \nabla\varphi(y)),\\ &\varphi = (\varphi_1, \ldots, \varphi_d) \quad \text{solutions of (2.50).}\end{aligned} \tag{2.55}$$

We recall that $\alpha, \beta > 0$, and D_0 is a positive-definite symmetric matrix while $M : \mathbb{Y} \to \mathcal{M}^d$ is a square-integrable function. We are interested in the solution $\overline{\varPhi}_k(\xi)$ to the problem: for all $\overline{\varPsi} \in H^1_\#(\mathbb{Y})$,

$$\int_{\mathbb{Y}} \nabla\overline{\varPsi}(y) \cdot \widehat{\mathscr{D}}(\xi)(y)(\mathbf{e}_k + \nabla\overline{\varPhi}_k(\xi)(y))\, dy \;=\; 0.$$

In the present context, $\mathcal{S} = H^1_\#(\mathbb{Y})$ and we wish to find a convenient finite dimensional approximation space $\mathcal{S}_N$ which allows to describe those solutions. The working plan faces the following technical difficulties:

- Here the parameter ξ ranges over the whole $\mathbb{R}^d$ while the method is designed to deal with parameters lying in a compact set.
- The method simplifies significantly when the dependence of $\mathscr{D}$ upon ξ is affine, which it is not here.
- The matrix M is square-integrable only and not essentially bounded, and the results of the literature do not apply to this case.

To construct the N-finite dimensional space $\mathcal{S}_N$, we proceed by induction using a greedy algorithm. We first choose a finite dimensional sampling $\mathcal{K}$ of $\mathscr{P}$, and construct the N-finite dimensional space $\mathcal{S}_N$ using $\mathcal{F}_{\mathcal{K}} = \{(\overline{\varPhi}_1(\xi), \ldots, \overline{\varPhi}_d(\xi)), \xi \in \mathcal{K}\}$. In order to proceed we need a distance on $\mathcal{F}_{\mathcal{K}}$, also called an estimator in this context.

Rewriting of the problem: from $\mathbb{R}^d$ to the unit ball

The starting point to rewrite the problem is the following fact: for all $\xi \in \mathbb{R}^d$ and all $k \in \{1, \ldots, d\}$, the corrector $\overline{\varPhi}_k(\xi) \in H^1_\#(\mathbb{Y})$ is solution to

$$-\nabla \cdot \frac{\widehat{\mathscr{D}}(\xi)}{1 + |\xi|}(\mathbf{e}_k + \nabla\overline{\varPhi}_k(\xi)) \;=\; 0. \tag{2.56}$$

Let S^{d-1} denote the unit hypersphere in dimension d and let $\mathscr{P} = [0,1] \times S^{d-1}$ (which is nothing but the closed unit ball). Define

$$\begin{aligned} \mathscr{D} : [0,1] \times S^{d-1} &\longrightarrow L^2(\mathbb{Y}, \mathcal{M}^d)\\ (\rho, X) &\longmapsto \mathscr{D}(\rho, X)\end{aligned}$$

by

$$\mathscr{D}(\rho, X) :\ y \mapsto\ (1-\rho)D_0 + \rho\left(\alpha|M(y)X|\mathrm{Id} + \beta\frac{M(y)X \otimes M(y)X}{|M(y)X|}\right), \tag{2.57}$$

For all $(\rho, X) \in [0,1] \times S^{d-1}$ and $k \in \{1, \dots, d\}$, we define $\overline{\Phi}_k(\rho, X)$ as the unique weak solution in $H^1_\#(\mathbb{Y})$ to

$$-\nabla \cdot \mathscr{D}(\rho, X)(\mathbf{e}_k + \nabla\overline{\Phi}_k(\rho, X))\ =\ 0.$$

Let $\xi \in \mathbb{R}^d$, and set

$$\rho = \frac{|\xi|}{1+|\xi|}, \qquad X = \frac{\xi}{|\xi|}$$

so that

$$\frac{\widehat{\mathscr{D}}(\xi)}{1+|\xi|}\ =\ \mathscr{D}(\rho, X);$$

the identity (2.56) implies that

$$\overline{\Phi}_k(\xi)\ \equiv\ \overline{\Phi}_k(\rho, X)$$

by uniqueness of correctors. In particular, this shows that

$$\left\{\overline{\Phi}_k(\xi), \xi \in \mathbb{R}^d, k \in \{1, \dots, d\}\right\} = \left\{\overline{\Phi}_k(\rho, X), (\rho, X) \in [0,1) \times S^{d-1}, k \in \{1, \dots, d\}\right\}.$$

What we gain by applying the reduced basis method on this new formulation is that the parameters now belong to the compact set $\mathscr{P} := [0,1] \times S^{d-1}$.

To complete the description of the RB method, we need to choose an estimator. Let $j \in \mathbb{N}$ and let $\mathcal{V}_j$ be a subspace of $H^1_\#(\mathbb{Y})$ of dimension j. Set for all $(\rho, X) \in [0,1] \times S^{d-1}$ and $k \in \{1, \dots, d\}$,

$$\overline{\mathcal{E}}^j(\rho, X, k)\ =\ \sqrt{\frac{|\mathbf{e}_k \cdot (\overline{D}^*(\rho, X) - \overline{D}^{*,j}(\rho, X))\mathbf{e}_k|}{\mathbf{e}_k \cdot \overline{D}^*(\rho, X)\mathbf{e}_k}}, \tag{2.58}$$

where, denoting by $\overline{\Phi}^j_k(\rho, X)$ the approximation of $\overline{\Phi}_k(\rho, X)$ in $\mathcal{V}_j$, we have

$$\begin{aligned}
\mathbf{e}_k \cdot \overline{D}^*(\rho, X)\mathbf{e}_k &= \int_{\mathbb{Y}} (\mathbf{e}_k + \nabla\overline{\Phi}_k(\rho, X)) \cdot \mathscr{D}(\rho, X)(\mathbf{e}_k + \nabla\overline{\Phi}_k(\rho, X))\ dy, \qquad (2.59)\\
\mathbf{e}_k \cdot \overline{D}^{*,j}(\rho, X)\mathbf{e}_k &= \int_{\mathbb{Y}} (\mathbf{e}_k + \nabla\overline{\Phi}^j_k(\rho, X)) \cdot \mathscr{D}(\rho, X)(\mathbf{e}_k + \nabla\overline{\Phi}^j_k(\rho, X))\ dy.
\end{aligned}$$

Since we also have for all $\xi \in \mathbb{R}^d$

$$\begin{aligned}
\overline{D}^*(\rho, X) &= \frac{1}{1+|\xi|}D^*(\xi),\\
\overline{D}^{*,j}(\rho, X) &= \frac{1}{1+|\xi|}D^{*,j}(\xi),
\end{aligned}$$

for $\rho = \frac{|\xi|}{1+|\xi|}$ and $X = \frac{\xi}{|\xi|}$, it is equivalent to approximate the desired homogenized matrix D^* (see (2.54)) and $\overline{D}^*$. We will focus on the latter in what follows.

Before we turn to fast-assembly, let us make a comment of the RB method used here. There exists $C_1 > 0$ such that for all $j \in \mathbb{N}$, $(\rho, X) \in [0,1] \times S^{d-1}$, and $k \in \{1, \dots, d\}$, the estimator (2.58) satisfies

$$C_1 \|\nabla \overline{\Phi}_k(\rho, X) - \nabla \overline{\Phi}_k^j(\rho, X)\|_{L^2(\mathbb{Y})} \le \overline{\mathcal{E}}^j(\rho, X, k).$$

Yet the converse inequality only holds in a weaker sense. In particular, using that $\overline{D}^*(\rho, X)$ and $\overline{D}^{*,j}(\rho, X)$ can be defined as

$$\mathbf{e}_k \cdot \overline{D}^*(\rho, X)\mathbf{e}_k = \int_{\mathbb{Y}} \mathbf{e}_k \cdot \mathscr{D}(\rho, X)(\mathbf{e}_k + \nabla \overline{\Phi}_k(\rho, X))\, dy,$$
$$\mathbf{e}_k \cdot \overline{D}^{*,j}(\rho, X)\mathbf{e}_k = \int_{\mathbb{Y}} \mathbf{e}_k \cdot \mathscr{D}(\rho, X)(\mathbf{e}_k + \nabla \overline{\Phi}_k^j(\rho, X))\, dy,$$

if $M \in L^2(\mathbb{Y}, \mathcal{M}^d)$ is square-integrable but not *essentially unbounded*, we end up with

$$C_2 \overline{\mathcal{E}}^j(\rho, X, k) \le \|\nabla \overline{\Phi}_k(\rho, X) - \nabla \overline{\Phi}_k^j(\rho, X)\|_{L^2(\mathbb{Y})}^{1/2},$$

for some $C_1 > 0$. As a consequence, the analysis of the reduced basis method and of the greedy algorithm in this case does not follow from [9,17,22,23]. Filling the gap for analyzing the convergence of the RB method when dealing with such unbounded coefficients does not seem to be an easy task. Nevertheless, the numerical experiments show the efficiency of the algorithm to treat this case.

Fast-assembly procedure

In order for the reduced basis method to be fast, it is desirable to have a fast-assembly procedure, which allows to quickly construct the linear system in $\mathcal{S}_N$. Fast-assembly procedures usually rely on the affine dependence of the diffusion matrix on the parameters.

For notational convience we take $d = 2$. In dimension 2, the unit sphere S^1 is parametrized by $[0, 2\pi]$, so that from now on, we write the element of S^1 as

$$X = \mathbf{e}(\theta) = \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2, \tag{2.60}$$

and consider $\mathscr{D}$ as a function of ρ and θ (instead of ρ and X). The diffusion matrix $\mathscr{D} : [0,1] \times [0, 2\pi] \to L^2(\mathbb{Y}, \mathcal{M}^d)$ given by (2.57), that is

$$\mathscr{D}(\rho, \theta) :\ y \mapsto\ (1-\rho)D_0 + \rho\left(\alpha |M(y)\mathbf{e}(\theta)| \mathrm{Id} + \beta \frac{M(y)\mathbf{e}(\theta) \otimes M(y)\mathbf{e}(\theta)}{|M(y)\mathbf{e}(\theta)|}\right),$$

is affine with respect to ρ, but not with respect to $\theta \in [0, 2\pi]$.

To circumvent this difficulty we use a partial Fourier series expansion in the θ-variable, and write:

$$\mathscr{D}(\rho,\theta)(y) \;=\; (1-\rho)D_0 + \rho\left(\frac{a_0(y)}{2} + \sum_{n=1}^{\infty}(a_n(y)\cos(n\theta) + b_n(y)\sin(n\theta))\right),$$

where the functions $y \mapsto a_n(y)$ and $y \mapsto b_n(y)$ are matrices which depend only on $y \mapsto M(y)$.

Given a finite-dimensional space $\mathcal{V}_N = \operatorname{span}\{\Psi_1,\dots,\Psi_N\}$ of dimension $N \geq 1$, and some parameters $(\rho,\theta) \in [0,1]\times[0,2\pi]$ and $k \in \{1,\dots,d\}$, in order to approximate the corrector $\overline{\Phi}_k$ in $\mathcal{V}_N$, it is enough to solve the linear system

$$\mathbf{M}(\rho,\theta)U \;=\; B(\rho,\theta,k),$$

where $\mathbf{M}(\rho,\theta,k)$ is the $N\times N$-matrix given for all $1 \leq j_1, j_2 \leq N$ by

$$\begin{aligned}\mathbf{M}(\rho,\theta)_{j_1j_2} \;=\; &(1-\rho)\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot D_0\nabla\Psi_{j_2}\;dy + \rho\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot\frac{a_0(y)}{2}\nabla\Psi_{j_2}\;dy\\ &+\sum_{n=1}^{\infty}\rho\cos(n\theta)\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot a_n(y)\nabla\Psi_{j_2}\;dy + \sum_{n=1}^{\infty}\rho\sin(n\theta)\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot b_n(y)\nabla\Psi_{j_2}\;dy,\end{aligned}$$

and the r. h. s. is the N-vector given for all $1 \leq j \leq N$ by

$$\begin{aligned}B(\rho,\theta,k)_j \;=\; &-(1-\rho)\int_{\mathbb{Y}}\nabla\Psi_j\cdot D_0\mathbf{e}_k\;dy - \rho\int_{\mathbb{Y}}\nabla\Psi_j\cdot\frac{a_0(y)}{2}\mathbf{e}_k\;dy\\ &-\sum_{n=1}^{\infty}\rho\cos(n\theta)\int_{\mathbb{Y}}\nabla\Psi_j\cdot a_n(y)\mathbf{e}_k\;dy + \sum_{n=1}^{\infty}\rho\sin(n\theta)\int_{\mathbb{Y}}\nabla\Psi_j\cdot b_n(y)\mathbf{e}_k\;dy.\end{aligned}$$

In particular, provided we truncate the Fourier series expansion up to some order $L \in \mathbb{N}$, a fast assembly procedure can be devised if the $2(L+1)$ following matrices of order N and $2Lk(L+1)$ following vectors of order N are stored:

$$\begin{aligned}&\left(\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot D_0\nabla\Psi_{j_2}\;dy\right)_{j_1,j_2}, \quad \left(\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot\frac{a_0(y)}{2}\nabla\Psi_{j_2}\;dy\right)_{j_1,j_2},\\ &\quad\left(\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot a_n(y)\nabla\Psi_{j_2}\;dy\right)_{j_1,j_2}, \quad \left(\int_{\mathbb{Y}}\nabla\Psi_{j_1}\cdot b_n(y)\nabla\Psi_{j_2}\;dy\right)_{j_1,j_2} \text{ for } n \in \{1,\dots,L\},\end{aligned} \tag{2.61}$$

and for $\in \{1,\dots,d\}$,

$$\begin{aligned}&\left(\int_{\mathbb{Y}}\nabla\Psi_j\cdot D_0\mathbf{e}_k\;dy\right)_j, \quad \left(\int_{\mathbb{Y}}\nabla\Psi_j\cdot\frac{a_0(y)}{2}\mathbf{e}_k\;dy\right)_j,\\ &\quad\left(\int_{\mathbb{Y}}\nabla\Psi_j\cdot a_n(y)\mathbf{e}_k\;dy\right)_j, \quad \left(\int_{\mathbb{Y}}\nabla\Psi_j\cdot b_n(y)\mathbf{e}_k\;dy\right)_j \text{ for } n \in \{1,\dots,L\}.\end{aligned} \tag{2.62}$$

Note that the number of real numbers to be stored for the fast-assembly only depends on L and N. In particular, if the reduced basis vectors Ψ_j are approximated in a finite-dimensional subspace of $H^1_\#(\mathbb{Y})$, this number is independent of the size of that subspace, as desired.

In practice, once we are given the reduced basis $\{\Psi_1, \dots, \Psi_N\}$, the matrices (2.61) and vectors (2.62) can be obtained by performing a fast Fourier transform of

$$\theta \mapsto \alpha |M(y)\mathbf{e}(\theta)| \mathrm{Id} + \beta \frac{M(y)\mathbf{e}(\theta) \otimes M(y)\mathbf{e}(\theta)}{|M(y)\mathbf{e}(\theta)|}$$

at each Gauss point $y \in \mathbb{Y}$ to evaluate the values of $a_n(y)$ and $b_n(y)$.

2.4.4 Numerical results

Let $d = 2$, $\mathrm{T}_{\mathbb{Y},\mathsf{h}_1}, \mathrm{T}_{\mathbb{Y},\overline{h}_1}$ be regular tessellations of $\mathbb{Y}$ of meshsize $\mathsf{h}_1, \overline{\mathsf{h}}_1 > 0$, and $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}, \mathcal{V}^1_{\mathbb{Y},\overline{\mathsf{h}}_1}$ be the subspaces of $H^1_\#(\mathbb{Y})$ made of $\mathbb{P}_1$-periodic finite elements associated with $\mathrm{T}_{\mathbb{Y},\mathsf{h}_1}$ and $\mathrm{T}_{\mathbb{Y},\overline{h}_1}$, respectively. The diffusion matrix $M \in L^2(\mathbb{Y}, \mathcal{M}^d)$ is defined by

$$M(y) \;=\; K(y)(\mathrm{Id} + \nabla \varphi^{\overline{\mathsf{h}}_1}(y)),$$

where K is a standard checkerboard: for all $y = (y_1, y_2) \in \mathbb{Y}$,

$$K(y_1, y_2) = \begin{cases} 4.94064, \text{ if } \{y_1 \geq 0.5,\ y_2 \geq 0.5\} \text{ or } \{y_1 < 0.5,\ y_2 < 0.5\}, \\ 0.57816, \text{ elsewhere,} \end{cases}$$

and $\varphi^{\overline{\mathsf{h}}_1} = (\varphi_1^{\overline{\mathsf{h}}_1}, \dots, \varphi_d^{\overline{\mathsf{h}}_1})$ is the approximation of the correctors of (2.50). In the actual computations, we take $\overline{\mathsf{h}}_1 \in \{1/10, 1/20, 1/40\}$ so that $\dim \mathcal{V}^1_{\mathbb{Y},\overline{\mathsf{h}}_1} \sim 100, 400, 1600$. In the rest of this paragraph, we assume that the corrector equations are solved in $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}$, so that the reduced basis will be a subspace of $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}$ as well.

For the reduced basis method we replace the compact space $\mathscr{P} = [0,1] \times [0, 2\pi]$ by the finite set $\big\{(\rho_i, \theta_j),\ (i,j) \in \{1, \dots, p\} \times \{1, \dots, p-1\}\big\}$, with $p \geq 2$, $\theta_j = (j-1)\frac{2\pi}{p-1}$, and $\rho_i = (i-1)\frac{1}{p-1}$. Let us denote by $\mathscr{D}_L$ the diffusion matrix obtained by a truncation of the Fourier series expansion of $\mathscr{D}$ at order L, and let $\overline{D}^*$ denote the homogenized coefficients defined in (2.59) (where the correctors $\overline{\Phi}_k(\rho, X)$ is in fact approximated in $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}$, and with X related to θ through (2.60)), and let $\overline{D}^*_L$ be defined by

$$\mathbf{e}_k \cdot \overline{D}^*_L(\rho, \theta)\mathbf{e}_k \;=\; \int_{\mathbb{Y}} (\mathbf{e}_k + \nabla \overline{\Phi}_k(\rho, \theta)) \cdot \mathscr{D}_L(\rho, \theta)(\mathbf{e}_k + \nabla \overline{\Phi}_k(\rho, \theta))\, dy.$$

We choose L such that

$$\sup_{i,j \in \{1, \dots, p\}} \frac{|\overline{D}^*(\rho_i, \theta_j) - \overline{D}^*_L(\rho_i, \theta_j)|}{|\overline{D}^*(\rho_i, \theta_j)|} \;\leq\; 10^{-6}.$$

Numerical tests show that L depends both on $\overline{\mathsf{h}}_1$ and on p, but not on h_1. As can be expected, the smaller $\overline{\mathsf{h}}_1$, the finer the approximation $\varphi^{\overline{\mathsf{h}}_1}$ of the correctors φ of the Darcy equation, and therefore the more complex $\mathscr{D}$ (it should however stabilize as $\overline{\mathsf{h}}_1 \to 0$). We display the results of the numerical tests on L in Table 2.3.

p \ $\overline{\mathsf{h}}_1$	1/10	1/20	1/40
11	41	61	61
21	41	61	61
41	49	95	175

Table 2.3. Dependence of L upon $\overline{\mathsf{h}}_1$ and p

p \ $\overline{\mathsf{h}}_1$	1/10	1/20	1/40
11	21	25	25
21	23	38	44
41	24	47	60

Table 2.4. Dependence of N upon $\overline{\mathsf{h}}_1$ and p

For all $N \leq p^2$, we denote by $\mathcal{V}_N$ the RB space of dimension N. We then choose N such that

$$\sup_{\mathscr{P}} \left(\mathcal{E}_L^N(\rho,\theta)\right)^2 \leq 10^{-6},$$

where $\mathcal{E}_L^N$ is the estimator associated with $\mathscr{D}_L$ and the space $\mathcal{V}_N$, when the equations are solved in $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}$. As expected, N depends both on $\overline{\mathsf{h}}_1$ and on p, but not on $\mathsf{h}_1 \in \{1/10, 1/20, 1/40, 1/80, 1/160, 1/320\}$ (which is the desired scaling property). The dimension N of the reduced basis in fonction of $\overline{\mathsf{h}}_1$ and on p is displayed in Table 2.4.

In order to check a posteriori the efficiency of the method (both in terms of L and N), we have picked at random a set $\widetilde{\mathscr{P}}$ of 100 pairs of parameters $(\rho,\theta) \in [0,1]\times[0,2\pi]$, computed the corresponding approximations $\overline{D}^*(\rho,\theta)$ of the homogenized coefficients in $\mathcal{V}^1_{\mathbb{Y},\mathsf{h}_1}$, and compared them to the approximations $\overline{D}_L^{*,N}(\rho,\theta)$ using the reduced basis method of order N and a Fourier series expansion of $\mathscr{D}$ truncated at order L. The numerical tests show that this error

$$\sup_{\widetilde{\mathscr{P}}} \frac{|\overline{D}^*(\rho,\theta) - \overline{D}_L^{*,N}(\rho,\theta)|}{|\overline{D}^*(\rho,\theta)|}$$

does not depend on h_1 but depends again on $\overline{\mathsf{h}}_1$ and p. We've chosen $p \in \{11, 21, 41\}$ so that the sample sets are included in one another, which ensures that the error due to the RB method decreases as p increases, as can be checked on Table 2.5. Note also that the error increases as $\overline{\mathsf{h}}_1 \to 0$.

A last comment is as follows. For $p = 41$ and $\overline{\mathsf{h}}_1 = 1/40$, the error is not reduced much with respect to $p = 21$. On Figure 2.10 the points chosen by the greedy algorithm are plotted for $p = 41$ and $\overline{\mathsf{h}}_1 = 1/40$ (circles denote points for the corrector in the direction $\mathbf{e}_1$ and crosses denote points for the corrector in the direction $\mathbf{e}_2$). This figure shows that most of the information for the RB lies in the region ρ close to 1 and θ in $[0, \pi]$ (this latter fact is indeed a consequence of the identity $\mathscr{D}(\rho, \theta) = \mathscr{D}(\rho, \pi - \theta)$). This motivates us to put more points in this region rather than in the rest of $\mathscr{P}$, and allows us to focus on the right region of the parameters. Taking for instance 5×168 points in the region $[0.9, 1] \times [0, \pi]$ and 10×30 in $[0, 1] \times [0, \pi]$, that is a total of 1140 points (to be compared to the $41 \times 40 = 1640$ uniformly chosen points in $\mathscr{P}$), the reduced basis has dimension $N = 68$ for $\overline{\mathsf{h}}_1 = 1/40$, $L = 177$, and the error on the 100 random points of $\widetilde{\mathscr{P}}$ is reduced to $4.1e-05$ (instead of $2.0e-04$).

p \ $\overline{\mathsf{h}}_1$	1/10	1/20	1/40
11	1.2e-04	9.0e-04	1.6e-03
21	3.4e-05	1.8e-04	2.6e-04
41	7.4e-06	3.5e-05	2.0e-04

Table 2.5. Dependence of the RB error upon $\overline{\mathsf{h}}_1$ and p on a random sampling of 100 points

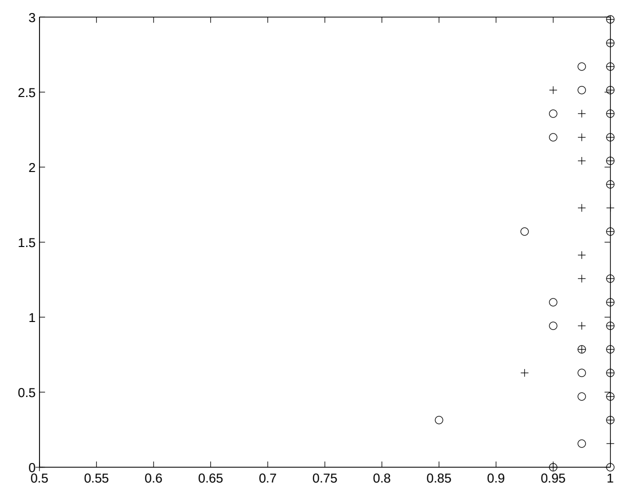

Fig. 2.10. Points chosen by the greedy algorithm for $p = 41$ and $\overline{\mathsf{h}}_1 = 1/40$ (all the points chosen in $[0, 1] \times [0, 2\pi]$ lie in $[0.5, 1] \times [0, \pi]$).

In conclusion, these tests widely confirm the efficiency of the method.

2.5 Perspectives

There are numerous perspectives related to the results of this chapter.

Quantitative stochastic homogenization theory

In the field of quantitative estimates in stochastic homogenization, an interesting problem is to understand which type of correlations ensures the validity of the spectral gap estimate.

As in the discrete case, this is a necessary step to extend the quantitative analysis to nonlinear problems, say to integral functionals with convex integrands.

A second direction is the extension to systems. This requires a proxy for the Cacciopoli inequality which is crucially used in the proof of Theorem 9.

Concerning the estimate of the fluctuations obtained in Section 2.2, it is not yet clear how to get rid off the restriction on the ellipticity ratio and obtain the optimal scalings in the general case. In another direction, for small ellipticity ratio, the weak norm of the fluctuation has the central limit theorem scaling, and it would be of interest to identify whether the rescaled process converges to a Gaussian random variable.

In Section 2.2 we have estimated the quantitity $\int_{\mathbb{R}^d} \langle (u_\varepsilon - \langle u_\varepsilon \rangle)^2 \rangle$. In order to quantify the convergence rate of u_ε to its homogenized limit u_{hom}, it remains to estimate the term $\int_{\mathbb{R}^d} (\langle u_\varepsilon \rangle - u_{\mathrm{hom}})^2$. This would hopefully improve the results by Yurinskiĭ [91] and by Caffarelli-Souganidis [18] in the case of linear elliptic equations whose random diffusion matrix has finite correlation length.

A completely different direction — which is of interest for applications — would not to perturb the diffusion matrix but rather its law in general, and try to understand how to quantify the related uncertainties.

About homogenization structures

In a series of works [69, 70] Nguetseng has introduced so-called homogenization structures, which are meant to be the largest class of structures which ensure homogenization. It would be satisfying to be able to prove that the general method introduced in Section 2.3 to approximate homogenized coefficients is also consistent in the class of homogenization structures.

Approximation of homogenized coefficients beyond the linear case

Numerous methods have been proposed in the litetature to perform "upscaling" or "numerical homogenization" of linear elliptic equations. These approaches can usually be combined with the method of Section 2.3 to reduce the effect of boundary conditions.

When the elliptic equation is nonlinear, the (possibly expensive) approximation of a homogenized matrix is not enough, since the homogenized elliptic equation is also nonlinear itself. Hence we cannot so easily rely on "precomputations". This explains why these multiscale methods remain prohibitive in terms of computational cost for nonlinear problems. In the case of integral functionals, it becomes reasonable to try to approximate the homogenized energy density. This raises quite interesting questions in approximation theory and inverse problems: how to reconstruct a convex function from partial measurements, etc. This approach will be used in the last section of this manuscript in the framework of nonlinear elasticity — although its analysis is currently out of reach.

3

Homogenization of integral functionals

In this chapter we first recall standard facts on homogenization of multiple integrals, addressing both periodic and stochastic homogenization. We illustrate the role of the "homogenization structure" in the first section. The homogenization structure is the assumption which ensures that the homogenized integrand does not depend on the space variable (whence "homogenization"). Typical such assumptions are periodicity and stochastic stationarity. We address here a variant which mixes periodicity and stochastic stationarity.

The aim of this chapter is to go beyond the standard homogenization result and understand how important qualitative properties are inherited or not by the homogenized integrand. This task is made difficult even in the periodic case because of the asymptotic character of the homogenization formula (see (3.4)). A possibility to get around this difficulty would be to derive a more tractable homogenization formula. In the periodic case, a natural candidate is the quasiconvex envelope of the cell formula. Unfortunetaly, we provide a counter-example which shows that the quasiconvex envelope of the cell formula does not coincide in general with the homogenization formula. Hence it seems difficult to by-pass the asymptotic character of the homogenization formula.

A natural approach to obtain qualitative properties on a nonlinear problem is to linearize around an equilibrium. We consider integrands which model standard hyperelastic materials (frame-invariant, minimal at identity, and which admits a Taylor expansion at identity). We prove that homogenization and linearization commute at identity in this class in general — whatever the structure ensuring homogenization. This yields information on the Taylor expansion of the homogenized integrand at identity. The key ingredient which allows to have a uniform control in the asymptotic homogenization formula is the celebrated quantitative rigidity estimate.

Other qualitative properties are however truly nonlinear. This is the case of strong ellipticity. Strong ellipticity is a property which guarantees that minimizers are isolated (and thus stable), and which allows to obtain short time existence theory in elastodynamics. In the periodic case, a complete theory was developed in the 90's using Bloch waves. This has allowed to identify conditions which ensure either that the homogenized integrand remains strongly elliptic or loses strong ellipticity. We generalize this approach to a more general homogenization structure: the stochastic stationary case.

3.1 Well-known facts

In this section, we recall classical results of periodic and stochastic homogenization of multiple integrals. For general results, we refer the reader to the monographies [15] and [44]. A suitable notion to study homogenization of integral functionals is Γ-convergence.

Definition 8 *Let $\mathcal{U}$ be a metric space. We say that $I : \mathcal{U} \to [-\infty, +\infty]$ is the Γ-limit of a sequence $I_k : \mathcal{U} \to [-\infty, +\infty]$, or that I_k Γ-converges to I, if for every $u \in \mathcal{U}$ the following conditions are satisfied:*

i) Liminf inequality*: for every sequence u_k in $\mathcal{U}$ such that $u_k \to u$,*

$$I(u) \leq \liminf_{k \to +\infty} I_k(u_k);$$

ii) Recovery sequence*: there exists a sequence u_k in $\mathcal{U}$ such that $u_k \to u$ and*

$$I(u) = \lim_{k \to +\infty} I_k(u_k).$$

This definition is compatible with minimization problems due to the following result.

Lemma 3.1. *Let $\mathcal{U}$ be a metric space, let I_k be a sequence of equi-mildly coercive functions on $\mathcal{U}$ (i. e. there exists a compact set K such that $\inf_{\mathcal{U}} I_k = \inf_K I_k$ for all $k \in \mathbb{N}$), and let $I = \Gamma - \lim_k I_k$; then*

$$\exists \min_{\mathcal{U}} I = \lim_{k \to +\infty} \inf_{\mathcal{U}} I_k.$$

Moreover, if u_k is a converging sequence such that $\lim_k I_k(u_k) = \lim_k \inf_{\mathcal{U}} I_k$, then its limit is a minimum point for I.

Let $d, n \geq 1$ be dimensions, and D be a bounded Lispchitz domain of $\mathbb{R}^d$. We focus on Γ-convergence for periodic and stochastic homogenization of integral functionals on the normed space $L^p(D, \mathbb{R}^n)$, $p \in (1, +\infty)$.

We start with the definition of integral functionals, and more specifically with the definition of admissible integrands.

Definition 9 *A function $W : \mathbb{R}^d \times \mathcal{M}^{d\times n} \to \mathbb{R}$ is a Carathéodory function if for every $\Lambda \in \mathcal{M}^{d\times n}$, $W(\cdot, \Lambda)$ is measurable and if for almost all $x \in \mathbb{R}^d$, $W(x, \cdot)$ is continuous.*

A Carathéodory function on $\mathbb{R}^d \times \mathcal{M}^{d\times n}$ is equivalent to a Borel function on $\mathbb{R}^d \times \mathcal{M}^{d\times n}$, so that for every $u \in W^{1,1}(D, \mathbb{R}^n)$, the function $x \mapsto W(x, \nabla u(x))$ is measurable on D.

Let $C > 0$ and $1 < p < +\infty$. We say that a Carathéodory function W satisfies a standard growth condition of order p on D if for all $\Lambda \in \mathcal{M}^{d\times n}$ and almost all $x \in D$,

$$\frac{1}{C}|\Lambda|^p - C \leq W(x, \Lambda) \leq C(1 + |\Lambda|^p). \tag{3.1}$$

Let $Q = (0, 1)^d$. The periodicity assumption is as follows.

Hypothesis 1 *The function $W : \mathbb{R}^d \times \mathcal{M}^{d\times n} \to [0,+\infty)$ is a Carathéodory function, which is Q-periodic in the first variable, and satisfies a standard growth condition.*

Under Hypothesis 1, we consider for any $\varepsilon > 0$ the integral functional $I_\varepsilon : L^p(D,\mathbb{R}^n) \to [0,+\infty]$ defined by

$$I_\varepsilon(u) := \begin{cases} \displaystyle\int_D W\big(\frac{x}{\varepsilon}, \nabla u(x)\big)dx & \text{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases}$$

Note that if for almost every $y \in Q$, $W(y,\cdot)$ is a quasiconvex function on $\mathcal{M}^{d\times n}$, then the functional I_ε is lower-semicontinuous for the weak convergence in $W^{1,p}(D,\mathbb{R}^n)$. Before we go further, let us recall that a function $f : \mathcal{M}^{d\times n} \to [0,+\infty)$ is rank-one convex if it is convex along rank-one connections (i. e. the function $t \mapsto W(\Lambda + ta\otimes b)$ is convex for all $\Lambda \in \mathcal{M}^{d\times n}$ and $a \in \mathbb{R}^d, b \in \mathbb{R}^n$), it is quasiconvex if for all $\Lambda \in \mathcal{M}^{d\times n}$

$$f(\Lambda) \;=\; \inf\left\{ \fint_Q f(\Lambda + \nabla u(x))dx \;:\; u \in W_0^{1,\infty}(\Lambda) \right\},$$

and f is polyconvex if there exists a convex function $\tilde{f}$ such that $f(\Lambda) = \tilde{f}(m_1(\Lambda),\ldots,m_k(\Lambda))$ where $m_1(\Lambda),\ldots,m_k(\Lambda)$ are the minors of the matrix Λ. We have the series of implications:

$$\text{convex} \quad\Longrightarrow\quad \text{polyconvex} \quad\Longrightarrow\quad \text{quasiconvex} \quad\Longrightarrow\quad \text{rank-one convex.} \tag{3.2}$$

Given a function $f : \mathcal{M}^{d\times n} \to [0,+\infty)$, its rank-one convex (resp. quasiconvex, polyconvex, convex) envelope $\mathcal{R}f$ (resp. $\mathcal{Q}f$, $\mathcal{P}f$, $\mathcal{C}f$) is the largest rank-one convex (resp. quasiconvex, polyconvex, convex) function lower or equal to f.

We now define two quantities related to the homogenization of W.

Definition 10 *We call* cell integrand *related to W the function $W_{\text{cell}} : \mathcal{M}^{d\times n} \to [0,+\infty)$ defined by*

$$W_{\text{cell}}(\Lambda) := \inf\Big\{ \fint_Q W\big(y, \Lambda + \nabla\phi(y)\big)dy \;:\; \phi \in W^{1,p}_{\#}(Q,\mathbb{R}^n)\Big\}. \tag{3.3}$$

We call homogenized integrand *related to W the function $W_{\text{hom}} : \mathcal{M}^{d\times n} \to [0,+\infty)$ defined by*

$$W_{\text{hom}}(\Lambda) := \lim_{N\to\infty} \frac{1}{|Q_N|} \inf\Big\{ \int_{Q_N} W\big(y, \Lambda + \nabla\phi(y)\big)dy \;:\; \phi \in W^{1,p}_{\#}(Q_N,\mathbb{R}^n)\Big\}, \tag{3.4}$$

where for all $N \in \mathbb{N}$, $Q_N = (0,N)^d$.

The periodic homogenization theorem is as follows:

Theorem 18 *[13, 64] Assume that W satisfies Hypothesis 1. Then the asymptotic formula (3.4) is well-defined, the homogenized integrand $W_{\rm hom}$ is a quasiconvex function satisfying (3.1), and for any $\varepsilon_k \searrow 0^+$ the sequence I_{ε_k} Γ-converges to the functional $I_{\rm hom} : L^p(D, \mathbb{R}^n) \to [0, +\infty]$ defined by*

$$I_{\rm hom}(u) := \begin{cases} \displaystyle\int_D W_{\rm hom}\big(\nabla u(x)\big)dx & \text{if} \quad u \in W^{1,p}(D, \mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases}$$

In addition, if $W(y, \cdot)$ is convex for a.e. $y \in Q$, then $W_{\rm hom}$ is convex and coincides with the cell integrand $W_{\rm cell}$ related to W.

We turn to stochastic homogenization, and start with some definitions. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We shall say that the family of mappings $(\tau_z)_{z\in\mathbb{R}^d}$ from Ω to Ω is a strongly continuous measure-preserving ergodic translation group if:

- $(\tau_z)_{z\in\mathbb{R}^d}$ has the group property: $\tau_0 = \mathrm{Id}$ (the identity mapping), and for all $x, y \in \mathbb{R}^d$, $\tau_{x+y} = \tau_x \circ \tau_y$;
- $(\tau_z)_{z\in\mathbb{R}^d}$ preserves the measure: for all $z \in \mathbb{R}^d$, and every measurable set $F \in \mathcal{F}$, $\tau_z F$ is measurable and $\mathbb{P}(\tau_z F) = \mathbb{P}(F)$;
- $(\tau_z)_{z\in\mathbb{R}^d}$ is strongly continuous: for any measurable function f on Ω, the function $(\omega, z) \mapsto f(\tau_z\omega)$ defined on $\Omega \times \mathbb{R}^d$ is measurable (with the Lebesgue measure on $\mathbb{R}^d$);
- $(\tau_z)_{z\in\mathbb{R}^d}$ is ergodic: for all $F \in \mathcal{F}$, if for all $z \in \mathbb{R}^d$, $\tau_z F \subset F$, then $\mathbb{P}(F) \in \{0, 1\}$.

The assumptions on W in the stochastic case are as follows:

Hypothesis 2 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\tau_z)_{z\in\mathbb{R}^d}$ be a strongly continuous measure-preserving ergodic translation group, and let $C > 0$ and $1 < p < +\infty$. The function $W : \mathbb{R}^d \times \mathcal{M}^{d\times n} \times \Omega \to [0, +\infty)$ is such that:*

- *for almost every $y \in \mathbb{R}^d$ and for all $\Lambda \in \mathcal{M}^{d\times n}$, $W(x, \Lambda, \cdot)$ is measurable,*
- *for $\mathbb{P}$-almost $\omega \in \Omega$, $W(\cdot, \cdot, \omega)$ is a Carathéodory function satisfying the standard growth condition (3.1).*

The function W is stationary for $(\tau_z)_{z\in\mathbb{R}^d}$: for $\mathbb{P}$-almost every $\omega \in \Omega$, almost every $x \in \mathbb{R}^d$, every $\Lambda \in \mathcal{M}^{d\times n}$ and every $z \in \mathbb{R}^d$

$$W(x + z, \Lambda, \omega) \;=\; W(x, \Lambda, \tau_z\omega). \tag{3.5}$$

Under Hypothesis 2, we consider for any $\varepsilon > 0$ the random integral functional $I_\varepsilon : L^p(D, \mathbb{R}^n) \times \Omega \to [0, +\infty]$ defined by

$$I_\varepsilon(u, \omega) := \begin{cases} \displaystyle\int_D W\big(\frac{x}{\varepsilon}, \nabla u(x), \omega\big)dx & \text{if} \quad u \in W^{1,p}(D, \mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases}$$

We may define a homogenized integrand:

Definition 11 *We call* homogenized integrand *related to* W *the function* $W_{\text{hom}} : \mathcal{M}^{d\times n} \to [0,+\infty)$ *defined by*

$$W_{\text{hom}}(\Lambda) := \left\langle \lim_{R\to\infty} \frac{1}{|Q_R|} \inf\Big\{ \int_{Q_R} W\big(y, \Lambda + \nabla\phi(y), \omega\big)dy \;:\; \phi \in W_0^{1,p}(Q_R, \mathbb{R}^d) \Big\} \right\rangle, \quad (3.6)$$

where for all $R > 0$, $Q_R = (0,R)^d$, *and* $\langle\cdot\rangle$ *denotes the expectation.*

The stochastic homogenization theorem is as follows:

Theorem 19 *[26,61] Assume that* W *satisfies Hypothesis 2. Then the asymptotic formula* (3.6) *is well-defined, the homogenized integrand* W_{hom} *is a quasiconvex function satisfying* (3.1)*, and for any* $\varepsilon_k \searrow 0^+$ *the sequence* $I_{\varepsilon_k}(\cdot,\omega)$ *Γ-converges for* $\mathbb{P}$*-almost every* $\omega \in \Omega$ *to the functional* $I_{\text{hom}} : L^p(D,\mathbb{R}^n) \to [0,+\infty]$ *defined by*

$$I_{\text{hom}}(u) := \begin{cases} \displaystyle\int_D W_{\text{hom}}\big(\nabla u(x)\big)dx & \text{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases}$$

In addition, if $W(y,\cdot)$ *is convex for a.e.* $y \in Q$*, then* W_{hom} *is convex.*

The proof of this result heavily relies on the subadditive ergodic theorem, which allows to by-pass the existence of correctors (and therefore to deal with nonconvex problems).

***Remark* 4** *We may consider a group parametrized by* $\mathbb{Z}^d$ *in place of* $\mathbb{R}^d$*. Let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space. We shall say that the family of mappings* $(\tau_z)_{z\in\mathbb{Z}^d}$ *from* Ω *to* Ω *is a measure-preserving ergodic translation group if:*

- $(\tau_z)_{z\in\mathbb{Z}^d}$ *has the group property:* $\tau_0 = \mathrm{Id}$ *(the identity mapping), and for all* $x,y \in \mathbb{Z}^d$, $\tau_{x+y} = \tau_x \circ \tau_y$;
- $(\tau_z)_{z\in\mathbb{Z}^d}$ *preserves the measure: for all* $z \in \mathbb{Z}^d$*, and every measurable set* $F \in \mathcal{F}$, $\tau_z F$ *is measurable and* $\mathbb{P}(\tau_z F) = \mathbb{P}(F)$;
- $(\tau_z)_{z\in\mathbb{Z}^d}$ *is ergodic: for all* $F \in \mathcal{F}$*, if for all* $z \in \mathbb{Z}^d$, $\tau_z F \subset F$*, then* $\mathbb{P}(F) \in \{0,1\}$.

We define stationarity by restricting the identity (3.5) *to all* $z \in \mathbb{Z}^d$*: a measurable function* $f : \mathbb{R}^d \times \Omega$ *is stationary if for* $\mathbb{P}$*-almost every* $\omega \in \Omega$*, almost every* $x \in \mathbb{R}^d$*, and every* $z \in \mathbb{Z}^d$

$$f(x+z,\omega) \;=\; f(x,\tau_z\omega). \quad (3.7)$$

The homogenization result of Theorem 19 holds as well in this case.

3.2 Playing with homogenization structures [Glo08]

In this section, we prove a result corresponding to Theorems 18 and 19 when the integrand is neither periodic nor stationary, but mixes both notions. This extends results by Blanc, Le Bris, and Lions [11] to the case of integral functionals.

3.2.1 Mixing periodicity and stationarity

The idea of Blanc, Le Bris, and Lions is to deform a periodic structure by a stationary stochastic diffeomorphism defined as follows.

Definition 12 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\tau_z)_{z\in\mathbb{Z}^d}$ from Ω to Ω be a (discrete) measure-preserving ergodic translation group. An application $\Phi : \mathbb{R}^d \times \Omega \to \mathbb{R}^d$, which is continuous in the first variable and measurable in the second variable, is said to be a stationary stochastic diffeomorphism if*

- *for $\mathbb{P}$-almost all ω, $\Phi(\cdot, \omega)$ is a diffeormorphism from $\mathbb{R}^d$ onto itself,*
- *$\nabla\Phi$ is stationary in the sense of (3.7),*

if its Jacobian is uniformly bounded from below

$$\operatorname*{ess\,inf}_{\omega\in\Omega,\ x\in\mathbb{R}^d} \det(\nabla\Phi(x,\omega)) \geq \nu > 0 \tag{3.8}$$

and if its gradient is uniformly bounded from above

$$\operatorname*{ess\,sup}_{\omega\in\Omega,\ x\in\mathbb{R}^d} |\nabla\Phi(x,\omega)| \leq M < \infty. \tag{3.9}$$

Given an integrand W satisfying Hypothesis 1 or Hypothesis 2 (with however a discrete ergodic group as in Remark 4), we consider the new integrand $(x, \Lambda, \omega) \mapsto W(\Phi^{-1}(x,\omega), \Lambda, \omega)$ (in the periodic case there is of course no stochastic dependence besides the change of variables). Under these assumptions, we consider for any $\varepsilon > 0$ the random integral functional $I_\varepsilon : L^p(D, \mathbb{R}^n) \times \Omega \to [0, +\infty]$ defined by

$$I_\varepsilon(u,\omega) := \begin{cases} W\Big(\Phi^{-1}\Big(\dfrac{x}{\varepsilon},\omega\Big), \nabla u(x), \omega\Big)dx & \text{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases} \tag{3.10}$$

Note that since Φ is not assumed to be stationary, Φ^{-1} is not stationary either, so this case does not enter the assumptions of Theorem 19.

We may define a homogenized integrand:

Definition 13 *We call* homogenized integrand *related to W and Φ the function $W_{\text{hom}} : \mathcal{M}^{d\times n} \to [0, +\infty)$ defined by*

$$\begin{aligned} W_{\text{hom}}(\Lambda) &= \left\langle \lim_{N\to\infty} \frac{1}{|Q_N|} \inf\left\{ \int_{Q_N} W(\Phi^{-1}(y,\cdot), \Lambda + \nabla v, \cdot)dy, v \in W^{1,p}_0(Q_N, \mathbb{R}^d) \right\} \right\rangle \\ &= \left\langle \lim_{N\to\infty} \frac{1}{|Q_N|} \inf\left\{ \int_{Q_N} W(y, (\nabla\Phi(y,\cdot))^{-1}(\Lambda + \nabla v), \cdot) \det\big(\nabla\Phi(y,\cdot)\big)dy, \right.\right. \\ &\qquad\qquad \left.\left. v \in W^{1,p}_0(Q_N, \mathbb{R}^d) \right\} \right\rangle \det\left(\left\langle \int_Q \nabla\Phi(z,\cdot)dz \right\rangle \right)^{-1}, \end{aligned} \tag{3.11}$$

where $\langle\cdot\rangle$ denotes the expectation.

The homogenization result is as follows:

Theorem 20 *Let Φ be a stationary stochastic diffeomorphism, and let W satisfy Hypothesis 1 or Hypothesis 2 (with a discrete ergodic group as in Remark 4) for some $C > 0$ and $1 < p < +\infty$. Then the asymptotic formula* (3.11) *is well-defined, the homogenized integrand W_{hom} is a quasiconvex function satisfying* (3.1)*, and for any $\varepsilon_k \searrow 0^+$ the sequence $I_{\varepsilon_k}(\cdot,\omega)$ defined in* (3.13) *Γ-converges for $\mathbb{P}$-almost every $\omega \in \Omega$ to the functional $I_{\mathrm{hom}} : L^p(D,\mathbb{R}^n) \to [0,+\infty]$ defined by*

$$I_{\mathrm{hom}}(u) := \begin{cases} \displaystyle\int_D W_{\mathrm{hom}}\big(\nabla u(x)\big)dx & \textit{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ \\ +\infty & \textit{otherwise.} \end{cases}$$

For $x \mapsto \Phi(x,\cdot) \equiv x$, we recover the classical result by Braides [13] and by Müller [64] in the periodic case, and the result by Dal Maso and Modica [26] and by Messaoudi and Michaille [61] in the stationary stochastic case.

3.2.2 Sketch of proof

The key observation (and in fact the key motivation for the deformation through Φ^{-1} and not Φ) is the following change of variable: $x \rightsquigarrow y = \varepsilon\Phi^{-1}(\frac{x}{\varepsilon},\omega)$. This yields for all $u \in W^{1,p}(D)$,

$$\begin{aligned} I_\varepsilon(u,\omega) &= \int_D W\left(\Phi^{-1}\left(\frac{x}{\varepsilon},\omega\right),\nabla u(x),\omega\right)dx \\ &= \int_{\varepsilon\Phi^{-1}\left(\frac{1}{\varepsilon}D,\omega\right)} W\left(y/\varepsilon,\left(\nabla\Phi\left(y/\varepsilon,\omega\right)\right)^{-1}\nabla\tilde{u}(y),\omega\right)\det\left(\nabla\Phi\left(y/\varepsilon,\omega\right)\right)dy, \end{aligned} \tag{3.12}$$

where $\tilde{u} : \varepsilon\Phi^{-1}\left(\dfrac{1}{\varepsilon}D,\omega\right) \to \mathbb{R}^n, y \mapsto u(\varepsilon\Phi(y/\varepsilon,\omega))$.

What we've gained in this formulation is that the integrand is now stationary (provided we forget that $\tilde{u}$ depends on ω as well). Hence, for any bounded Lipschitz domain A, the functional $F_\varepsilon^A : L^p(A) \times \Omega \to [0,+\infty]$ defined by

$$F_\varepsilon^A(v,\omega) := \begin{cases} \displaystyle\int_A W\left(y/\varepsilon,\left(\nabla\Phi\left(y/\varepsilon,\omega\right)\right)^{-1}\nabla v(y),\omega\right)\det\left(\nabla\Phi\left(y/\varepsilon,\omega\right)\right)dy, & \text{if} \quad v \in W^{1,p}(A,\mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases} \tag{3.13}$$

satisfies Hypothesis 2 (for a discrete translation group), and can therefore be homogenized.

It remains to prove that the Γ-convergence of F_ε^A implies the Γ-convergence of I_ε. Let us proceed formally. By Theorem 2 there exists some integrand $W^* : \mathcal{M}^{d\times n} \to [0,+\infty)$

such that for every bounded Lipschitz domain A, the sequence F_ε^A Γ-converges $\mathbb{P}$-almost surely to $F_*^A : L^p(A) \to [0,+\infty]$ defined by

$$F_*^A(v) := \begin{cases} \displaystyle\int_A W^*(x)dx, & \text{if} \quad v \in W^{1,p}(A,\mathbb{R}^n), \\ \\ +\infty & \text{otherwise.} \end{cases}$$

The argument relies on two further observations:

- for all $\varepsilon > 0$ and $\mathbb{P}$-almost every $\omega \in \Omega$,

$$u \in W_0^{1,p}(A,\mathbb{R}^n) \quad \Longleftrightarrow \quad \tilde{u} \in W_0^{1,p}\left(\Phi^{-1}\left(\frac{1}{\varepsilon}A,\omega\right),\mathbb{R}^n\right);$$

- there exists an invertible matrix $L \in \mathcal{M}^d$ such that for all $x \in \mathbb{R}^d$ and for $\mathbb{P}$-almost every $\omega \in \Omega$,

$$\lim_{\varepsilon\to 0} \varepsilon\Phi^{-1}(x/\varepsilon) \;=\; L^{-1}x.$$

  In addition, this convergence holds locally uniformly.

The combination of both observations yields after some work that for any $\varepsilon_k \searrow 0^+$ and any sequence $u_k : D \to \mathbb{R}^n$ which converges to some u in $L^p(D,\mathbb{R}^n)$, for $\mathbb{P}$-almost every $\omega \in \Omega$ the sequence $\overline{u}_k : \varepsilon_k\Phi^{-1}(D/\varepsilon_k,\omega) \to \mathbb{R}^n, x \mapsto u_k(\varepsilon_k\Phi(x/\varepsilon_k,\omega))$ converges to $\overline{u} = u \circ L$ in $L^p(A)$ for all A strictly contained in $L^{-1}(D)$. Likewise, for $\mathbb{P}$-almost every $\omega \in \Omega$, $\varepsilon_k\Phi^{-1}(D/\varepsilon_k,\omega) \to L^{-1}(D)$ (in the sense of that the associated characteristic functions converge in $L^1(A)$).

We may now try to pass formally to the Γ-limit in the r. h. s. of (3.13): the domain of integration converges to $L^{-1}(D)$, whereas the integrand converges to $W^*(\nabla(u \circ L))$, so that

$$I_{\mathrm{hom}}(u) \;=\; \int_{L^{-1}(D)} W^*(L\nabla u(Ly))dy.$$

The change of variables $y \rightsquigarrow x = Ly$ then yields

$$I_{\mathrm{hom}}(u) \;=\; \int_D \det(L^{-1})W^*(\nabla u(x))dx.$$

This formally proves that $W_{\mathrm{hom}} \equiv \det(L^{-1})W^*$.

The last ingredient to turn this into a rigorous argument is the extensive use of the growth condition. As one may guess, the linear map L is simply given by

$$L \;=\; \left\langle \int_Q \nabla\Phi(z,\cdot)dz \right\rangle.$$

3.3 On the asymptotic periodic homogenization formula [BG]

In this section we come back to the periodic case. We first recall the two-dimensional example by Müller [64] which shows that in general, for nonconvex integrands, there may exist $\Lambda \in \mathcal{M}^d$ such that

$$W_{\mathrm{hom}}(\Lambda) \;<\; W_{\mathrm{cell}}(\Lambda).$$

As we shall show, the integrand W_{cell} of Müller's example is not a quasiconvex function, so that it cannot coincide with W_{hom}. A natural question is whether quasiconvexifying W_{cell} yields W_{hom}. This question

- makes sense because W_{hom} is a quasiconvex function lower or equal to W_{cell} (so that it could indeed be its quasiconvex envelope),
- is of interest because it is technically much easier to deal with a quasiconvexification than with an asymptotic formula.

Müller's argument does not allow to decide whether the quasiconvex envelope $\mathcal{Q}W_{\mathrm{cell}}$ of W_{cell} coincides with W_{hom} or not. We then provide with an example for which we are able to prove that for some $\Lambda \in \mathcal{M}^d$

$$W_{\mathrm{hom}}(\Lambda) \;<\; \mathcal{Q}W_{\mathrm{cell}}(\Lambda).$$

3.3.1 Müller's counter-example

The energy under consideration $W^\eta : \mathbb{R}^2 \times \mathcal{M}^2 \to [0, +\infty), (x, \Lambda) \mapsto \chi^\eta(x) W_0(\Lambda)$ models a two-dimensional laminate composite, made of a strong material and a soft material. The coefficient χ^η is the Q-periodic extension on $\mathbb{R}^2$ of

$$\chi^\eta(x) := \begin{cases} 1 & \text{if } x_1 \in (0, 1/2) \\ \eta & \text{if } x_1 \in [1/2, 1) \end{cases},$$

where $Q \ni x = (x_1, x_2)$ and $\eta > 0$. The energy density $W_0 : \mathcal{M}^2 \to [0, +\infty)$ is given by $W_0(\Lambda) = |\Lambda|^4 + f(\det \Lambda)$ where

$$f(z) := \begin{cases} \dfrac{8(1+a)^2}{z+a} - 8(1+a) - 4 & \text{if } z > 0 \\[2ex] \dfrac{8(1+a)^2}{a} - 8(1+a) - 4 - \dfrac{8(1+a)^2}{a^2} z & \text{if } z \le 0 \end{cases}$$

for some $a \in (0, 1)$.

In particular, $W^\eta(x, \cdot)$ is a polyconvex function satisfying a standard growth condition (3.1) of order $p = 4$. Its zero level-set $(W^\eta(x, \cdot))^{-1}$ is SO_2 for almost every $x \in Q$.

We denote by W^η_{cell} and W^η_{hom} the cell integrand and the homogenized integrand associated with W^η through (3.3) and (3.4), respectively.

Using the one-well rigidity (Liouville theorem) on the unitary cell Q and using 'buckling-like' test-functions on several periodic cells (see Figure 3.3.1), Müller obtained the following result.

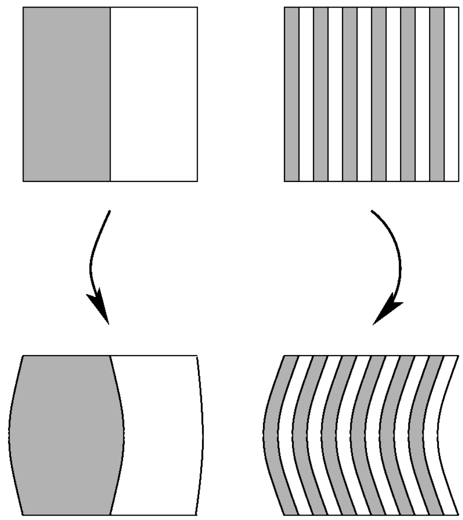

Fig. 3.1. Compression of one periodic cell and buckling of several periodic cells

Lemma 3.2. *[64] For all $\lambda \in (0,1)$, there exist $c_1, c_2 > 0$ independent of η, such that*

$$W^\eta_{\mathrm{hom}}(\Lambda) \leq \eta\, c_1, \quad W^\eta_{\mathrm{cell}}(\Lambda) \geq c_2, \tag{3.14}$$

where $\Lambda := \mathrm{diag}(1,\lambda)$, hence proving that the strict inequality $W^\eta_{\mathrm{cell}}(\Lambda) > W^\eta_{\mathrm{hom}}(\Lambda)$ holds provided η is small enough.

One may now refine Lemma 3.2. Indeed we have proved that W^η_{cell} is not even a rank-one convex function (so that, by (3.2) and the fact that W_{hom} is polyconvex, W_{cell} cannot coincide with W_{hom}):

Lemma 3.3. *For all $\lambda \in (0,1)$, there exists $c > 0$ independent of η such that*

$$\mathcal{R}W^\eta_{\mathrm{cell}}(\Lambda) \leq c\eta, \tag{3.15}$$

where $\Lambda := \mathrm{diag}(1,\lambda)$.

The proof of this lemma consists in the explicit construction of an upper bound for $\mathcal{R}W^\eta_{\mathrm{cell}}(\Lambda)$.

In view of Lemma 3.3, Lemma 3.2 does not allow to conclude whether the inequality $\mathcal{Q}W^\eta_{\mathrm{cell}}(\Lambda) \geq W^\eta_{\mathrm{hom}}(\Lambda)$ is strict or not. This is rather surprising since this example shows that buckling can be captured by rank-one convexification of the cell-formula (for which there is no buckling). In dimension $d \geq 3$, rank-one convexification is not enough, but quasiconvexification is. Hence it is not clear whether microscopic relaxation (such as the buckling) can be recovered by macroscopic relaxation (quasiconvexification of W_{cell}).

3.3.2 A counter-example from solid-solid phase transformations

The aim of this subsection is to exhibit an example with $d = n = 2$ for which one can prove that $\mathcal{Q}W_{\mathrm{cell}}(\Lambda) > W_{\mathrm{hom}}(\Lambda)$ for some $\Lambda \in \mathcal{M}^2$.

The strategy is as follows: we wish to construct an example such that two-well rigidity implies that W_{cell} is bounded from below by a strictly positive constant, whereas W_{hom} vanishes for some $\Lambda \in \mathcal{M}^2$.

The construction is a little technical, and we need the following ingredients:

- Matrices in $\mathcal{M}^2$

$$A_1 := \operatorname{diag}(1,1),\ A_2 := \operatorname{diag}(4,3),\ B_1 := \operatorname{diag}(1,3),\ B_2 := \operatorname{diag}(4,1),$$
$$C := \frac{1}{2}\operatorname{diag}(5,4),\ R := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

- Compact sets in $\mathcal{M}^2$

$$\begin{aligned} K_1 &:= SO_2A_1 \cup SO_2A_2, & K_2 &:= SO_2B_1 \cup SO_2B_2, \\ H_1 &:= K_1R, & H_2 &:= (K_2R)^{pc}, \end{aligned}$$

  where $(K_2R)^{pc}$ denotes the polyconvex hull of the set K_2R (that is the intersection of the zero-levelsets $f^{-1}(0)$ of all the non-negative polyconvex functions $f : \mathcal{M}^2 \to [0,+\infty)$ vanishing on K_2R).
- Subsets of $\mathbb{R}^2$

$$\begin{aligned} T_1 &:= \{x \in Q \ :\ x_2 \geq x_1 + 1\}, & T_2 &:= \{x \in Q \ :\ x_2 \geq -x_1 + 1\}, \\ T_3 &:= \{x \in Q \ :\ x_2 \leq -x_1 - 1\}, & T_4 &:= \{x \in Q \ :\ x_2 \leq x_1 - 1\}, \\ U_1 &:= \textstyle\bigcup_{i=1}^4 T_i, & U_2 &:= Q \setminus U_1. \end{aligned}$$

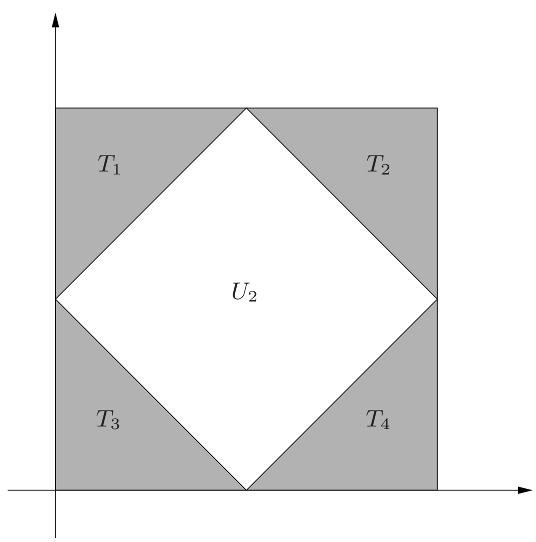

Fig. 3.2. Geometry.

The example is as follows:

Theorem 21 *Let $W^1, W^2 : \mathcal{M}^2 \to [0, +\infty)$ be two quasiconvex functions satisfying a standard growth condition* (3.1) *and such that*

$$(W^1)^{-1}(0) = H_1 \quad \textit{and} \quad (W^2)^{-1}(0) = H_2. \tag{3.16}$$

Consider the energy density $W : \mathbb{R}^2 \times \mathcal{M}^2 \to [0, +\infty)$ defined by

$$W(x, \Lambda) = \chi(x) W^1(\Lambda) + (1 - \chi(x)) W^2(\Lambda),$$

where χ is defined by $\chi = \chi_{U_1}$ in Q and extended by periodicity to the whole $\mathbb{R}^2$. The following properties hold:

1) the cell integrand W_{cell} related to W is bounded from below by a constant $c > 0$;
2) the matrix CR belongs to the zero level set of the homogenized integrand W_{hom} related to W.

Therefore $\mathcal{Q}W_{\mathrm{cell}}(CR) \geq c > 0 = W_{\mathrm{hom}}(CR)$.

The proof of this theorem uses the following facts:

i) The compact set K_1 is polyconvex and rigid, *i.e.*, if $U \subseteq \mathbb{R}^2$ is an open connected set and $\psi : U \to \mathbb{R}^2$ is a Lipschitz function such that

$$\nabla \psi(x) \in K_1 \quad \text{for } \mathcal{L}^2\text{-}a.e.\ x \in U,$$

then ψ is affine. We refer to [84, Theorem 2] and [65, Theorem 4.11] for the proofs. Since R is a rotation, the same properties hold for H_1.

ii) $H_1 \cap H_2 = \emptyset$, because

$$H_2 \subseteq \left\{ \Lambda \in \mathcal{M}^2 \ : \ \det \Lambda \in [3, 4] \right\}$$

and $\det \Lambda \in \{1, 12\}$ if $\Lambda \in H_1$. This inclusion is a consequence of the definition of polyconvexity: the set $\{\det \Lambda : \Lambda \in H_2\}$ is included in the convex hull of $\{\det \Lambda : \Lambda \in K_2 R\}$.

iii) A_1 is rank-one connected to B_1 and B_2, and A_2 to B_1 and B_2 also. More precisely, if we denote by $\{\mathbf{e}_1, \mathbf{e}_2\}$ the canonical basis in $\mathbb{R}^2$,

$$\begin{aligned} A_1 - B_1 &= -2\mathbf{e}_2 \otimes \mathbf{e}_2 \\ A_1 - B_2 &= -3\mathbf{e}_1 \otimes \mathbf{e}_1 \\ A_2 - B_1 &= 3\mathbf{e}_1 \otimes \mathbf{e}_1 \\ A_2 - B_2 &= 2\mathbf{e}_2 \otimes \mathbf{e}_2. \end{aligned}$$

The proof that W_{cell} is bounded from below is as follows. First we note that W_{cell} is a continuous function which satisfies a standard growth condition, so that it is infinite at infinity and therefore attains its minimum on $\mathcal{M}^2$. We then assume that there exists $\Lambda \in \mathcal{M}^2$ such that $W_{\mathrm{cell}}(\Lambda) = 0$, and let $\phi \in W^{1,p}_{\#}(Q)$ be a minimizer for (3.3). By exploiting the rigidity of H_1, we learn that $\nabla \phi + \Lambda \in H_1$ is constant on each triangle T_i. By Q-periodicity these constant must coincide. We denote by $D \in H_1$ this matrix. To conclude, we observe that $x \mapsto (\Lambda - D)x + \phi(x) \in W^{1,\infty}_0(U_2)$ so that by quasiconvexity,

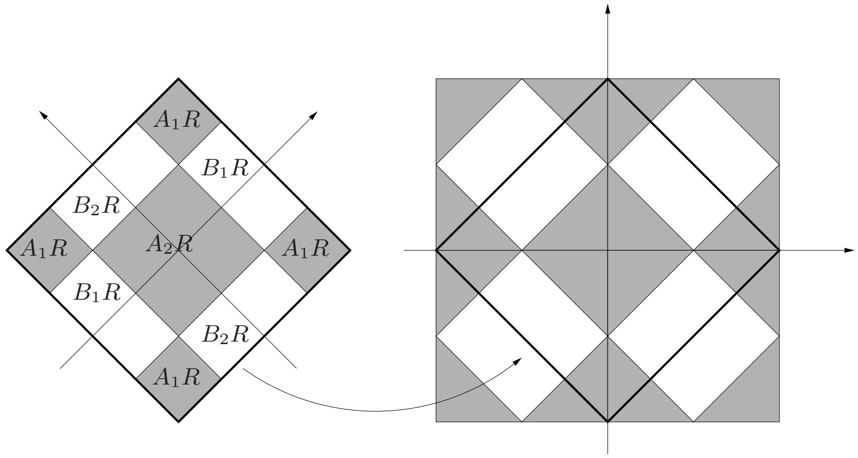

Fig. 3.3. The values of $CR+\nabla\phi$ in $R^{-1}(-1/\sqrt{2},1/\sqrt{2})^2$. On the left the axis are oriented in the directions $R^{-1}\mathbf{e}_1$ and $R^{-1}\mathbf{e}_2$.

$$|U_2|W_2(D) \leq \int_{U_2} W_2(D+(\Lambda-D)+\nabla\phi(x))dx \leq |Q|W_{\text{cell}}(\Lambda) = 0.$$

Hence, $D \in H_1 \cap H_2 = \emptyset$ and we obtain a contradiction.

For the second claim of the theorem, it is enough to prove that $W_2(CR) = 0$, where $W_2 : \mathcal{M}^2 \to [0,\infty)$ is defined by

$$\Lambda \mapsto W_2(\Lambda) = \inf\left\{ \fint_{Q_2} W(x,\Lambda+\nabla u(x))dx \,:\, u \in W^{1,p}_{\#}(Q_2,\mathbb{R}^2)\right\},$$

since by the asymptotic formula (3.4) and non-negativity of W, $0 \leq W_{\text{hom}} \leq W_2$. To this aim, one needs to construct some $\phi \in W^{1,p}_{\#}(Q_2,\mathbb{R}^2)$ such that

$$\int_{Q_2} W(x,CR+\nabla\phi(x))dx = 0.$$

Such a construction is illustrated on Figure 3.3.

Compared to Müller's example, the periodicity constraint is not lost in the x_1-direction. In particular the Q-periodicity constraint combined with the geometry implies that the four values in T_1, T_2, T_3, T_4 coincide. This is no longer the case for Q_2-periodic functions (as can be seen on Figure 3.3).

***Remark* 5** *In the case of periodic homogenization of discrete systems, Meunier, Pantz and Raoult [62] have given a nontrivial example for which $W_{\text{hom}} \equiv \mathcal{Q}W_1 \not\equiv W_1$.*

3.4 Linearization of the asymptotic homogenization formula at identity [GN]

In this section we address the question of the commutativity of homogenization and linearization. Their commutativity would indeed drastically simplify the linearization of the homogenized integrand, since it would coincide with the homogenization of the “linearized” integrand (which is a quadratic integrand). The difficulty comes again from the asymptotic homogenization formula. In the case of a periodic integrand, if we linearize first and then homogenize, the cell formula is enough. On the contrary, if we homogenize first, we need to linearize an asymptotic formula. The key ingredient to prove the commutativity is the quantitative rigidity estimate [37].

Let $p \geq 2$. We consider a sequence of integral functionals $I_{\varepsilon_k} : L^p(D) \to [0, +\infty]$ defined by

$$I_{\varepsilon_k}(u) := \begin{cases} \displaystyle\int_D W_{\varepsilon_k}\big(x, \nabla u(x)\big)dx & \text{if} \quad u \in W^{1,p}(D, \mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.17}$$

where $W_{\varepsilon_k} : D \times \mathcal{M}^d \to [0, +\infty]$ is a Borel function. As it is common in finite elasticity, we assume that W_{ε_k} is frame indifferent and minimal at identity. Moreover, we assume that W_{ε_k} is non-degenerate and admits a quadratic expansion at identity with quadratic term Q_{ε_k}; as a consequence, in situations when the deformation is close to a rigid-body motion, say when $|\nabla u - \mathrm{Id}\,| \sim h \ll 1$, we can accurately describe the functional I_{ε_k} (after scaling by h^{-2}) by the quadratic functional $E_{\varepsilon_k} : L^2(D) \to [0, +\infty]$ defined by

$$E_{\varepsilon_k}(u) := \begin{cases} \displaystyle\int_D Q_{\varepsilon_k}\big(x, \nabla g(x)\big)dx & \text{if} \quad g(x) := h^{-1}(u(x) - x), \quad \text{and } u \in H^1(D, \mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.18}$$

Since $Q_{\varepsilon_k}(\cdot, \Lambda)$ genuinely only depends on the symmetric part of the strain gradient Λ, the energy E_{ε_k} corresponds to linear elasticity. On the other hand, if I_{ε_k} has some specific structure in space rescaled by ε_k (think of periodicity for instance), we may expect a homogenization property to hold as ε_k vanishes, which justifies to replace the nonlinear oscillating-in-space energy density $(x, \Lambda) \mapsto W_{\varepsilon_k}(x, \Lambda)$ by a nonlinear homogeneous-in-space energy density $\Lambda \mapsto W_{\mathrm{hom}}(\Lambda)$ (or more generally by an energy density $(x, \Lambda) \mapsto W^*(x, \Lambda)$ whose oscillations in x are independent of ε_k).

The case when W_{ε_k} is $\varepsilon_k Q$-periodic has been successfully treated by Müller and Neukamm in [66]. In this section, we extend the commutation result to the general “compact” case, and give an application of this result to the problem of the determination of the Γ-closure.

3.4.1 Commutation of linearization and homogenization

We say that ρ is a modulus approximation if it is an increasing function from $\mathbb{R}^+$ to $[0, +\infty]$ such that $\lim_{h\to 0} \rho(h) = 0$. We first precisely define the class of nonlinear integrands we shall consider.

Definition 14 *For all $\alpha > 0$ and every modulus of approximation ρ, we denote by $\mathcal{W}_{\alpha,\rho}$ the set of Borel functions $W : \mathcal{M}^d \to [0, +\infty]$ which satisfy the following three properties:*

W is frame indifferent, i.e. (W1)

$$W(R\Lambda) = W(\Lambda) \qquad \textit{for all } \Lambda \in \mathcal{M}^d,\ R \in SO_d;$$

W is non degenerate, i.e. (W2)

$$W(\Lambda) \geq \tfrac{1}{\alpha}\,\mathrm{dist}^2(\Lambda, SO_d) \qquad \textit{for all } \Lambda \in \mathcal{M}^d;$$

W is minimal at Id *and admits the following quadratic expansion at* Id*:* (W3)

$$\sup_{0<|G|\leq\delta} \frac{|W(\mathrm{Id} + G) - Q(G)|}{|G|^2} \leq \rho(\delta) \qquad \textit{for all } \delta > 0,$$

where $Q : \mathcal{M}^d \to [0, \infty)$ is a quadratic form satisfying

$$0 \leq Q(G) \leq \alpha |G|^2 \qquad \textit{for all } G \in \mathcal{M}^d.$$

For all $1 < p < +\infty$, we further denote by $\mathcal{W}^p_{\alpha,\rho}$ the subset of the integrands of $\mathcal{W}_{\alpha,\rho}$ which satisfy in addition a standard growth condition (3.1). This set is not empty provided $p \geq 2$.

In the following two remarks we make important observations on the link between the assumptions on the nonlinear integrands and linear elasticity.

Remark 3.4. The energy densities of class $\mathcal{W}_{\alpha,\rho}$ describe elastic materials with a single, quadratic energy well at SO_d. The minimality condition in (W3) implies that the reference state $\Lambda = \mathrm{Id}$ is stress free. The combination of (W2) and (W3) might be interpreted as a generalization of Hooke's law to geometrically nonlinear material laws: for infinitesimal small strains we expect a linear stress-strain relation. Indeed, in view of condition (W3) the material law is sufficiently smooth to allow a linearization around the reference state.

Remark 3.5. Let $W \in \mathcal{W}_{\alpha,\rho}$ and let Q denote the quadratic form associated with W through (W3). Because of (W1) – (W3) the quadratic form Q generically satisfies conditions that are common in linear elasticity; namely, the growth and ellipticity condition

$$\forall G \in \mathcal{M}^d \ :\ \tfrac{1}{\alpha'}\,|\mathrm{sym}G|^2 \leq Q(G) \leq \alpha'\,|G|^2 \tag{Q1}$$

for some positive constant α' that only depends on α, and

$$\forall G \in \mathcal{M}^d \ :\ Q(\mathrm{skw}G) = 0, \tag{Q2}$$

where $G = \mathrm{sym}G + \mathrm{skw}G$ is the decomposition of G into its symmetric $1/2(G + G^T)$ and skew symmetric $1/2(G - G^T)$ parts. The property (Q2) follows from a Taylor expansion of W at identity using (W3) and the fact that $W(\Lambda)$ depends only on $\Lambda^T\Lambda$ by (W1). The non-degeneracy condition (Q1) on the quadratic form is inherited from the non-degeneracy condition (W2).

We may now define a class of linearized integrands:

Definition 15 *We denote by $\mathcal{Q}_{\alpha'}$ the set of non-negative quadratic forms $Q : \mathcal{M}^d \to \mathbb{R}^+$ satisfying* (Q1) *and* (Q2)*.*

We denote by $\mathcal{W}^p_{\alpha,\rho}(D \times \mathcal{M}^d)$ and $\mathcal{Q}_\alpha(D \times \mathcal{M}^d)$ the sets of Carathéodory integrands such that for almost every $x \in D$, $W(x,\cdot) \in \mathcal{W}^p_{\alpha,\rho}$ and $Q(x,\cdot) \in \mathcal{Q}_\alpha$, respectively. Given a sequence $\{W_{\varepsilon_k}\}_k$ of integrands in $\mathcal{W}^p_{\alpha,\rho}(D \times \mathcal{M}^d)$, we consider the sequence of integral functionals $I_{\varepsilon_k} : L^p(D) \to [0,+\infty]$ defined by (3.17), and the associated sequence $E_{\varepsilon_k} : L^2(D) \to [0,+\infty]$ of linearized functionals defined by (3.18). Likewise, for any Carathéodory integrand W^*, we consider the integral functional $I^* : L^p(D) \to [0,+\infty]$ defined by

$$I^*(u) := \begin{cases} \displaystyle\int_D W^*\big(x, \nabla u(x)\big)dx & \text{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.19}$$

If in addition W^* admits a quadratic expansion $Q^* \in \mathcal{Q}_\alpha(D \times \mathcal{M}^d)$, we consider the linearized functional $E^* : L^2(D) \to [0,+\infty]$ defined by

$$E^*(u) := \begin{cases} \displaystyle\int_D Q^*\big(x, \nabla g(x)\big)dx & \text{if} \quad g(x) := h^{-1}(u(x) - x), \quad \text{and } u \in H^1(D,\mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.20}$$

The commutation result is then as follows:

Theorem 3.6. *Let $2 \le p < +\infty$, $\{W_{\varepsilon_k}\}_k$ be a sequence of integrands in $\mathcal{W}^p_{\alpha,\rho}(D \times \mathcal{M}^d)$, and W^* be a Carathéodory integrand. Assume that the sequence I_{ε_k} defined in* (3.17) *Γ-converges in $L^p(D)$ to I^* defined in* (3.19)*. Then there exist positive constants α', α'', a modulus of approximation ρ' (all only depending on α and ρ), and $Q^* \in \mathcal{Q}_{\alpha''}(D \times \mathcal{M}^d)$ such that the following properties hold:*

(a) $W^* \in \mathcal{W}^p_{\alpha',\rho'}(D \times \mathcal{M}^d)$ and the expansion

$$W^*(x, \mathrm{Id} + G) = Q^*(x,G) + o(|G|^2)$$

holds for almost every $x \in D$ and for all $G \in \mathcal{M}^d$;

(b) the sequence E_{ε_k} defined in (3.18) *Γ-converges in $L^2(D)$ to E^* defined in* (3.20)*;*

(c) the following diagram commutes

$$\begin{array}{ccc} \mathcal{G}_{h,\varepsilon_k} & \xrightarrow{(1)} & E_{\varepsilon_k} \\ {\scriptstyle(2)}\downarrow & & \downarrow{\scriptstyle(3)} \\ \mathcal{G}^*_h & \xrightarrow[(4)]{} & E^* \end{array}$$

where $\mathcal{G}_{h,\varepsilon_k}$ and $\mathcal{G}^*_h$ denote the functionals from $H^1_0(D)$ to $[0,+\infty]$ defined as

$$\mathcal{G}_{h,\varepsilon}(g) := \frac{1}{h^2} I_\varepsilon(\varphi_{\mathrm{Id}} + hg), \qquad \mathcal{G}^*_h(g) := \frac{1}{h^2} I^*(\varphi_{\mathrm{Id}} + hg);$$

and (1),(4), and (2),(3) mean Γ-convergence in $H^1_0(D)$ with respect to the strong topology of $L^2(D)$ as $h \to 0$ and $\varepsilon_k \to 0$, respectively. Moreover, the families (I_ε) and (E_ε) are equi-coercive w. r. t. weak convergence in $H^1_0(D)$.

Note that the assumptions of this theorem are not restrictive in terms of Γ-convergence, since any sequence I_{ε_k} Γ-converges to some integral functional I^* up to extraction (by "Γ-compactness").

The proof of Theorem 3.6 mainly relies on an expansion result around Id for the function

$$\mathcal{M}^d \ni \Lambda \mapsto W_D(\Lambda) := \lim_{k\to\infty} \left\{ \inf_{v\in H^1_0(D)} I_{\varepsilon_k}(\varphi_\Lambda + v) \right\} \qquad (\text{where } \varphi_\Lambda(x) := \Lambda x).$$

In particular, we have proved under the assumptions and notation of Theorem 3.6 that

$$\frac{\left|W_D(\mathrm{Id} + G) - \inf_{v\in H^1_0(D)} \mathcal{E}^*(\varphi_G + v)\right|}{|G|^2} \leq |D|\rho'(|G|) \tag{3.21}$$

for all $G \in \mathcal{M}^d$. The rest of the proof is standard and makes use of the characterization of integral functionals by their minima [25]. The core of the work is the proof of (3.21). Let us first proceed formally and fix $G \in \mathcal{M}^d$. We assume that there exists a doubly indexed sequence $v_{\varepsilon_k,h}$ which is bounded in $W^{1,\infty}(D)$ and satisfies for all ε_k and h:

$$I_{\varepsilon_k}(\varphi_{\mathrm{Id}+hG} + hv_{\varepsilon_k,h}) \;=\; \inf_{v\in H^1_0(D)} I_{\varepsilon_k}(\varphi_{\mathrm{Id}+hG} + v).$$

We then make a Taylor expansion of the l. h. s.:

$$\begin{aligned} I_{\varepsilon_k}(\varphi_{\mathrm{Id}+hG} + hv_{\varepsilon_k,h}) &= h^2 E_{\varepsilon_k}(\varphi_G + v_{\varepsilon_k,h}) + o(h^2) \\ &\geq h^2 \inf_{v\in H^1_0(D)} E_{\varepsilon_k}(\varphi_G + v) + o(h^2) \end{aligned}$$

where the remainder $o(h^2)$ is uniform in ε_k. Passing to the limit $\varepsilon_k \to 0$, and using the convergence of infimum problems together with the Γ-compactness of quadratic integral functionals, this yields

$$W_D(\mathrm{Id} + hG) \geq h^2 \inf_{v\in H^1_0(D)} E^*(\varphi_G + v) + o(h^2).$$

Conversely, let assume there exists a doubly indexed sequence $u_{\varepsilon_k,h}$ which is bounded in $W^{1,\infty}(D)$ and satisfies for all ε_k and h:

$$E_{\varepsilon_k}(\varphi_G + u_{\varepsilon_k,h}) \;=\; \inf_{v\in H^1_0(D)} E_{\varepsilon_k}(\varphi_G + v).$$

We then have by definition of $u_{\varepsilon_k,h}$ and $v_{\varepsilon_k,h}$, and by a Taylor expansion,

$$\begin{aligned} I_{\varepsilon_k}(\varphi_{\mathrm{Id}+hG} + hv_{\varepsilon_k,h}) &\leq I_{\varepsilon_k}(\varphi_{\mathrm{Id}+hG} + hu_{\varepsilon_k,h}) \\ &= h^2 E_{\varepsilon_k}(\varphi_G + u_{\varepsilon_k,h}) + o(h^2) \\ &= h^2 \inf_{v\in H^1_0(D)} E_{\varepsilon_k}(\varphi_G + v) + o(h^2). \end{aligned}$$

Taking the limit as $\varepsilon_k \to 0$ then yields

$$W_D(\mathrm{Id} + hG) \leq h^2 \inf_{v \in H_0^1(D)} E^*(\varphi_G + v) + o(h^2).$$

The combination of the two estimates shows that W_D admits the desired Taylor expansion at $\Lambda = \mathrm{Id}$. To turn this into a rigorous argument we need to show that the remainders are uniform in ε_k although the sequence $v_{\varepsilon_k,h}$ is not bounded in $W^{1,\infty}(D)$. The proof makes use of three arguments:

- the non-degeneracy of W_{ε_k} combined with the rigidity estimate of [37] gives a uniform control of $v_{\varepsilon_k,h}$ in $H_0^1(D)$;
- Meyers' estimates imply that $u_{\varepsilon_k,h}$ is uniformly bounded in $W^{1,q}(D)$ for some $q > 2$;
- one may replace $u_{\varepsilon_k,h}$ by a Lipschitz function by a truncation argument of [37], and quantitatively control the error on the integral functional using Meyers' estimates.

3.4.2 Locality of the Γ-closure at identity

The problem of Γ-closure consists in characterizing all the energy densities which can be reached by Γ-convergence starting from a composite made of a finite number of constituents with prescribed volume fraction. In particular, the Γ-closure is said to be local in some class of integrands if and only if any such "homogenized" energy density is the pointwise limit of a sequence of homogenized energy densities obtained by *periodic* homogenization. In the linear case, this property has been proved independently by Tartar in [85] and Lurie and Cherkaev in [56]. The corresponding locality property of the G-closure for monotone operators is due to Raitums in [78] (generalizing an unpublished work by Dal Maso and Kohn). Related results of locality of the Γ-closure in the class of convex integrands can be found in [2]. Yet, the local character of the Γ-closure is an open question in the class of quasiconvex nonconvex integrands satisfying standard growth conditions. We focus here on a smaller class. In particular, we consider energy densities which are frame indifferent, non-degenerate, minimal at identity, admit a quadratic Taylor expansion at identity, and satisfy standard growth conditions. Then, we show that for any $F \mapsto W^*(F)$ in the Γ-closure of this set, there exists a sequence of periodic energy densities whose homogenized energy densities have quadratic Taylor expansions arbitrary close to the Taylor expansion of W^* at identity. This can be seen as a weak version of the local character of the Γ-closure in this set at identity. Although quite restricted, this is the first such result for quasiconvex nonconvex energy densities.

The notion of locality of the Γ-closure at identity is made precise in the following three definitions.

Definition 16 *Let* $1 < p < \infty$, $\{W_i\}_{i \in \{1,\dots,k\}} \in \mathcal{W}_\alpha^p$, *and* $\theta \in [0,1]^k$ *be such that* $\sum_{i=1}^k \theta_i = 1$. *We define the set of periodic homogenized energy densities associated with* $\{W_i, \theta_i\}_{i \in \{1,\dots,k\}}$ *as*

$$\mathcal{P}_\theta = \Bigg\{ (W_\chi)_{\text{hom}} : \mathcal{M}^d \to [0,+\infty) \, : \, \exists \chi \in L^\infty(\mathbb{R}^d, \{0,1\}^k) \text{ such that}$$
$$\chi \text{ is } Q\text{-periodic with } \int_Q \chi_i dy = \theta_i$$
$$\text{and } (W_\chi)_{\text{hom}} \text{ is associated with } W_\chi : (y, \Lambda) \mapsto \sum_{i=1}^k W_i(\Lambda)\chi_i(y) \text{ through (3.4)} \Bigg\},$$

and its closure for the pointwise convergence by

$$\mathcal{G}_\theta = \big\{ W^* : \mathcal{M}^d \to [0,+\infty) \, : \, \text{there exists a sequence } \{(W_{\chi^l})_{\text{hom}}\}_{l\in\mathbb{N}} \text{ of } \mathcal{P}_\theta$$
$$\text{such that } (W_{\chi^l})_{\text{hom}} \to W^* \text{ pointwise} \big\}.$$

Definition 17 *Let $2 \le p < \infty$, $\{W_i\}_{i\in\{1,\dots,k\}} \in \mathcal{W}^p_{\alpha,\rho}$, and $\theta \in [0,1]^k$ such that $\sum_{i=1}^k \theta_i = 1$. We define the set of periodic homogenized energy densities associated with $\{W_i, \theta_i\}_{i\in\{1,\dots,k\}}$ at identity as*

$$\mathcal{P}^{\text{Id}}_\theta = \Bigg\{ W^* : \mathcal{M}^d \to \mathbb{R} \, : \, \exists (W_\chi)_{\text{hom}} \in \mathcal{P}_\theta$$
$$\text{such that } \Big| W^*(\text{Id} + G) - (W_\chi)_{\text{hom}}(\text{Id} + G) \Big| = o(|G|^2) \Bigg\},$$

and its closure:

$$\mathcal{G}^{\text{Id}}_\theta = \Bigg\{ W^* : \mathcal{M}^d \to \mathbb{R} \, : \, \text{there exists a sequence } \{(W_{\chi^l})_{\text{hom}}\}_{l\in\mathbb{N}} \text{ of } \mathcal{P}_\theta$$
$$\text{such that } \Big| W^*(\text{Id} + G) - \lim_{l\to\infty} (W_{\chi^l})_{\text{hom}}(\text{Id} + G) \Big| = o(|G|^2) \Bigg\}.$$

Definition 18 *Let $2 \le p < \infty$, $\{W_i\}_{i\in\{1,\dots,k\}} \in \mathcal{W}^p_{\alpha,\rho}$. We say that the Γ-closure of $\{W_i\}_{i\in\{1,\dots,k\}}$ is local at identity if and only if for every sequence $\{\chi^l\}_{l\in\mathbb{N}}$ of $L^\infty(D,\{0,1\}^k)$ with $\sum_{i=1}^k \chi^l_i \equiv 1$ and such that*

- *there exists $\theta \in L^\infty(D,[0,1]^k)$ such that χ^l converges weakly-* to θ in $L^\infty(D,[0,1]^k)$,*
- *the functional $I_{\chi^l} : L^p(D) \to [0,+\infty]$ defined by*

$$I_{\chi^l}(u) := \begin{cases} \displaystyle\int_D \sum_{i=1^k} W_i\big(\nabla u(x)\big)\chi^l_i(x)dx & \text{if} \quad u \in W^{1,p}(D,\mathbb{R}^n), \\ +\infty & \text{otherwise.} \end{cases}$$

Γ-converges to the functional $I^* : L^p(D) \to [0,+\infty]$ defined by (3.19),

one has

$$W^*(x,\cdot) \in \mathcal{G}^{\mathrm{Id}}_{\theta(x)}$$

for almost every $x \in D$.

The above definition is a weakened version of the standard definition of the locality of the Γ-closure, which we have obtained by restricting the property of approximation by periodic homogenized energy densities to a neighborhood of identity via a Taylor expansion. We have:

Theorem 3.7. *Let* $2 \le p < \infty$ *and* $\{W_i\}_{i\in\{1,\dots,k\}} \in \mathcal{W}^p_{\alpha,\rho}$, *then the* Γ*-closure of* $\{W_i\}_{i\in\{1,\dots,k\}}$ *is local at identity.*

The proof of this result relies on the combination of the following three arguments:

- By an argument by Babadjian and Barchiesi [2] it is enough to prove the result in the homogeneous case (that is when W^* does not depend on the space variable);
- Theorem 3.6 ensures that W_{ε_k} and W^* admit quadratic expansions with a quantitative control of the error which is uniform in ε_k;
- The Γ-closure is local for quadratic functionals.

The locality of the Γ-closure for quadratic functionals then implies the locality of the Γ-closure at identity for the nonlinear functionals by the uniform control of the validity of the Taylor expansion.

3.5 Strong ellipticity and stochastic homogenization [Glo11a]

In Section 3.4 we have shown that linearization and stochastic homogenization commute at identity for a specific class of integrands. As a by-product, this implies the strong ellipticity of the homogenized integrand at identity. Indeed, the homogenized integrand admits a Taylor expansion at identity whose quadratic part satisfies (Q1) & (Q2) by Theorem 3.6, so that the quadratic form is strictly strongly elliptic (with ellipticity constant "$1/\alpha''$").

To complement Section 3.4 we'd like to address the question of strong ellipticity at any deformation gradient $\Lambda \in \mathcal{M}^d$. The new difficulty lies in the fact that (Q1) is not necessarily satisfied any longer (which implies that the associated quadratic functional is not necessarily convex). Our aim is to extend the analysis by Geymonat, Müller and Triantafyllidis to the stochastic setting. In [38] crucial use is made of the Bloch transform — which is a powerful tool for periodic problems. To treat the stochastic case, we need to use a stochastic version of the Bloch transform which replaces periodicity by stationarity. The stochastic Bloch transform we introduce is inspired by the work [75] by Papanicolaou and Varadhan on explicit formulas in stochastic homogenization of linear elliptic equations.

The results and proofs are rather technical, and we've chosen to insist on the differences between the stochastic and periodic Bloch transforms rather than introduce the reader to the (very nice) arguments of [38].

3.5.1 General framework

We consider an energy density W satisfying Hypothesis 2 for a *discrete* group of translations, cf. Remark 4. Theorem 19 ensures the existence of a deterministic homogenized integrand W_{hom}. As opposed to the linear case, this homogenized integrand is defined through a sequence of minimization problems (3.6) and not by a corrector (whose existence is not known). Let us assume yet that such a corrector exists. Let $\Lambda \in \mathcal{M}^d$, and let $\nabla\phi \in L^\infty(\mathbb{R}^d \times \Omega)$ be a stationary field such that almost surely

$$
\begin{aligned}
W_{\mathrm{hom}}(\Lambda) &= \lim_{R\to\infty} \fint_{Q_R} W(x, \Lambda + \nabla\phi(x,\omega), \omega)dx, \\
&= \left\langle \int_Q W(x, \nabla\phi(x,\cdot), \cdot)dx \right\rangle.
\end{aligned}
$$

In Section 3.4 we have proved that stochastic homogenization and linearization commute at identity in some specific class of integrands. We will not prove such a definite result here, but shall rather make assumptions which in turn imply the commutativity result. Yet the validity of these assumptions remains open. Provided linearization and homogenization commute, it makes sense to study the homogenization of the quadratic energy associated with the fourth-order tensor $\mathbb{L}(x,\omega) := \dfrac{\partial^2 W}{\partial \Lambda^2}(x, \Lambda + \nabla\phi(x,\omega), \omega)$, which is a stationary field by assumption.

Although $\mathbb{L}$ is almost surely strictly strongly elliptic almost everywhere, it is not clear whether, for any general open bounded domain D of $\mathbb{R}^d$, the associated energy functional

$$
u \in H^1_0(D) \mapsto \int_D \nabla u(x) \cdot \mathbb{L}(x,\omega)\nabla u(x)dx
$$

is coercive on $H^1_0(D)$ or not. This is in contrast to the analysis of Section 3.4. In order for the homogenization property to hold, a necessary and sufficient condition is the non-negativity of the following ellipticity constant:

$$
\lambda = \left\langle \inf_{v\in C_0^\infty(\mathbb{R}^d)} \frac{\int_{\mathbb{R}^d} \nabla v(x) \cdot \mathbb{L}(x,\cdot)\nabla v(x)dx}{\int_{\mathbb{R}^d} |\nabla v(x)|^2 dx} \right\rangle. \tag{3.22}
$$

By stationarity of $\mathbb{L}$, the quantity $\inf\limits_{v\in C_0^\infty(\mathbb{R}^d)} \dfrac{\int_{\mathbb{R}^d} \nabla v(x) \cdot \mathbb{L}(x,\cdot)\nabla v(x)dx}{\int_{\mathbb{R}^d} |\nabla v(x)|^2 dx}$ is invariant by the translation group $\{\tau_z\}_{z\in\mathbb{Z}^d}$ so that it takes almost surely the value λ by ergodicity. The proof that $\mathbb{L}$ can be homogenized provided $\lambda \geq 0$ can be done following the argument of [38] (for $\lambda > 0$, one may proceed classically, focus on the PDE, define a corrector and use compensated compactness to prove homogenization, whereas for $\lambda = 0$ we proceed by regularization and state the homogenization property solely as a Γ-convergence result).

In the next subsection we introduce a few other ellipticity constants which allow to decide whether the homogenized fourth-order tensor $\mathbb{L}_{\mathrm{hom}}$ is strictly strongly elliptic or not.

3.5.2 Stochastic Bloch transform and definition of ellipticity constants

In this subsection we assume that $\mathbb{L} : \mathbb{R}^d \times \Omega \to (\mathcal{M}^d)^2$ is a bounded random field of fourth order tensors which is stationary for the discrete ergodic translation group $\{\tau_z\}_{z\in\mathbb{Z}^d}$ (in short, $\mathbb{Z}^d$-stationary). In order to study the "homogenizability" of integral functionals associated with $\mathbb{L}$, we focus on various ellipticity constants, next to the constant λ defined in (3.22). To this aim we introduce some functional spaces. For every integer $N \geq 1$, we define the Hilbert space of N-stationary complex functions of $H^1_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{C}^d)$ as:

$$\mathcal{H}^1_N := \left\{ v \in H^1_{\mathrm{loc}}(\mathbb{R}^d, L^2(\Omega, \mathbb{C}^d)) \,\middle|\, v(x + Nz, \omega) = v(x, \tau_{Nz}\omega) \text{ for all } x \in \mathbb{R}^d, z \in \mathbb{Z}^d \right\},$$

endowed with the norm

$$\|v\|^2_{\mathcal{H}^1_N} = \left\langle \int_{Q_N} \left(v(x,\cdot)^2 + |\nabla v(x,\cdot)|^2\right) dx \right\rangle.$$

We extend the Euclidian scalar product to the Hermitian scalar product, that we still denote by "$\cdot$". The link between stochastic Bloch waves and λ is encoded (as we shall see below) in the following ellipticity constants

$$\lambda_1 = \inf \left\{ \frac{\left\langle \int_Q \nabla v \cdot \mathbb{L} \nabla v \right\rangle}{\left\langle \int_Q |\nabla v|^2 \right\rangle} \,\middle|\, v(x,\omega) = e^{i\gamma\cdot x} p(x,\omega), \gamma \in [0, 2\pi)^d, p \in \mathcal{H}^1_1 \right\},$$

$$\lambda_2 = \inf \left\{ \frac{\left\langle \int_{Q_N} \nabla v \cdot \mathbb{L} \nabla v \right\rangle}{\left\langle \int_{Q_N} |\nabla v|^2 \right\rangle} \,\middle|\, N \geq 1, v(x,\omega) = e^{i\gamma\cdot x} p(x,\omega), \gamma \in \mathbb{R}^d, p \in \mathcal{H}^1_N \right\},$$

$$\lambda_3 = \inf \left\{ \frac{\left\langle \int_{Q_N} \nabla v \cdot \mathbb{L} \nabla v \right\rangle}{\left\langle \int_{Q_N} |\nabla v|^2 \right\rangle} \,\middle|\, N \geq 1, v \in \mathcal{H}^1_N \right\},$$

which indeed satisfy

$$\lambda = \lambda_1 = \lambda_2 = \lambda_3. \tag{3.23}$$

The definition of these ellipticity constants makes use of Bloch waves, which are random fields of the form $(x,\omega) \mapsto e^{i\gamma\cdot x} p(x,\omega)$ with p stationary. The last three ellipticity contants we need are defined as

$$\lambda_4 = \inf \left\{ \frac{\left\langle \int_Q (a \otimes b + \nabla v) \cdot \mathbb{L}(a \otimes b + \nabla v) \right\rangle}{\left\langle \int_Q |a \otimes b + \nabla v|^2 \right\rangle} \,\middle|\, a \in \mathbb{C}^d, b \in \mathbb{R}^d, v \in \mathcal{H}^1_1 \right\},$$

$$\lambda_5 = \lim_{\gamma\to 0} \quad \inf \left\{ \frac{\left\langle \int_Q \nabla v \cdot \mathbb{L} \nabla v \right\rangle}{\left\langle \int_Q |\nabla v|^2 \right\rangle} \,\middle|\, v(x,\omega) = e^{i\gamma\cdot x} p(x,\omega), p \in \mathcal{H}^1_N \right\},$$

$$\lambda_6 = \inf \left\{ \frac{\left\langle \int_Q \nabla v \cdot \mathbb{L} \nabla v \right\rangle}{\left\langle \int_Q |\nabla v|^2 \right\rangle} \,\middle|\, v \in \mathcal{H}^1_1 \right\},$$

and satisfy:

$$\lambda_4 \;=\; \lambda_5. \tag{3.24}$$

The combination of (3.23) and (3.24) with the definitions of the ellipticity constants then yields

$$\lambda \;=\; \lambda_1 \;=\; \lambda_2 \;=\; \lambda_3 \;\le\; \lambda_4 \;=\; \lambda_5 \;\le\; \lambda_6$$

since the two inequalities $\lambda_4 \le \lambda_6$ and $\lambda_5 \ge \lambda_2$ are obvious.

***Remark* 6** *In the periodic case, the corresponding ellipticity constants are recovered by erasing the expectations, and replacing $\mathcal{H}^1_N$ by $H^1_\#(Q_N, \mathbb{C}^d)$.*

Before we define the stochastic Bloch transform and turn to some elements of proof of (3.23), let us give an interpretation of these ellipticity constants:

- λ measures the global coercivity of the nonhomogeneous tensor $\mathbb{L}$. It can be computed using smooth functions or equivalently using Bloch waves via λ_1.
- $\lambda_4 = \lambda_5$ measures coercivity with respect to long-wavelength ($\gamma \to 0$) perturbations or equivalently with respect to shearing deformations (both modulo $\mathbb{Z}^d$-stationary contributions).
- λ_6 measures coercivity with respect to $\mathbb{Z}^d$-stationary, possibly highly localized deformations.

The main result of the analysis is the following: let λ_{hom} be the best ellipticity constant for the homogenized integrand $\mathbb{L}_{\mathrm{hom}}$ (which exists whenever $\lambda \ge 0$). Then, $\lambda_{\mathrm{hom}} \ge \lambda_4$ and $\lambda_{\mathrm{hom}} = 0$ if $\lambda_4 = 0$.

The structure of the proof of these results follows the string of arguments developed in [38] provided we define a suitable Bloch wave transform. We illustrate this point by proving the inequality $\lambda \ge \lambda_1$ in detail. This proof exemplifies very well the differences between the periodic and stochastic cases, and the adaptations to pass from one to the other.

Let D be an open bounded Lipschitz domain. We denote by $C^\infty_D(\mathbb{R}^d, \mathbb{R}^d)$ the set of smooth functions from $\mathbb{R}^d$ to $\mathbb{R}^d$ with support in D. For all $v \in L^2(\Omega, C^\infty_D(\mathbb{R}^d, \mathbb{R}^d))$ we define the Bloch transform of v as follows: for all $\gamma \in [0, 2\pi)^d$, $\tilde{v}_\gamma : \Omega \times \mathbb{R}^d \to \mathbb{C}^d$ is given by

$$\tilde{v}_\gamma(x, \omega) \;:=\; \int_{z \in \mathbb{Z}^d} e^{-i\gamma\cdot z} v(x + z, \tau_{-z}\omega).$$

Since $x \mapsto v(x, \omega)$ has support in D almost surely, the random field $\tilde{v}_\gamma$ is well-defined. The interest of the Bloch transform is that it maps fields with compact support onto stationary fields (up to a phase). Indeed, for all $x \in \mathbb{R}^d$, $y \in \mathbb{Z}^d$, $\gamma \in [0, 2\pi)^d$, and almost every $\omega \in \Omega$, we have using the group property of $\{\tau_z\}_{z\in\mathbb{Z}^d}$

$$\begin{aligned}
\tilde{v}_\gamma(x + y, \omega) &= \int_{z\in\mathbb{Z}^d} e^{-i\gamma\cdot z} v(x + y + z, \tau_{-z}\omega) \\
&= e^{i\gamma\cdot y} \int_{z\in\mathbb{Z}^d} e^{-i\gamma\cdot(y+z)} v(x + y + z, \tau_{-(y+z)}\tau_y\omega) \\
&= e^{i\gamma\cdot y} \tilde{v}_\gamma(x, \tau_y\omega),
\end{aligned}$$

so that $\hat{v}_\gamma : (x,\omega) \mapsto e^{-i\gamma\cdot x}\tilde{v}_\gamma(x,\omega)$ is a stationary field. This transform is a natural tool to relate λ to λ_1.

We have all the ingredients to prove that $\lambda \geq \lambda_1$. The starting point and crucial observation is that ergodicity allows us to use stationary functions to bound λ: for all $r > 0$, there exists a stationary random field $v_r \in L^2(\Omega, C_0^\infty(\mathbb{R}^d, \mathbb{R}^d))$ such that

$$\frac{\left\langle \int_{\mathbb{R}^d} \nabla v_r \cdot \mathbb{L} \nabla v_r \right\rangle}{\left\langle \int_{\mathbb{R}^d} |\nabla v_r|^2 \right\rangle} \leq \lambda + 1/r. \tag{3.25}$$

Here comes the argument. For all $L > 0$, we let $C^\infty_{0,L}(\mathbb{R}^d, \mathbb{R}^d)$ be the set of smooth functions from $\mathbb{R}^d$ to $\mathbb{R}^d$ whose support has a diameter at most $2L$ (i. e. the support is contained in some ball of radius L). Choose a sequence $r_{k_1} \to 0$ and a sequence $L_{k_2} \to \infty$. Since the random variable

$$\inf_{v \in C^\infty_{0,L}(\mathbb{R}^d,\mathbb{R}^d)} \frac{\int_{\mathbb{R}^d} \nabla v(x) \cdot \mathbb{L}(x,\cdot) \nabla v(x) dx}{\int_{\mathbb{R}^d} |\nabla v(x)|^2 dx}$$

is $\mathbb{Z}^d$-invariant, it is deterministic (up to a negligible set of events), and there exists a set of events of full measure Ω_1 such that for all k_1 and k_2 there exists a stationary field $w_{k_1,k_2} \in L^2(\Omega, C^\infty_{0,L_{k_2}}(\mathbb{R}^d, \mathbb{R}^d))$ such that for all $\omega \in \Omega_1$

$$\inf_{v \in C^\infty_{0,L_{k_2}}(\mathbb{R}^d,\mathbb{R}^d)} \frac{\int_{\mathbb{R}^d} \nabla v(x) \cdot \mathbb{L}(x,\omega) \nabla v(x) dx}{\int_{\mathbb{R}^d} |\nabla v(x)|^2 dx} \geq \left\langle \frac{\int_{\mathbb{R}^d} \nabla w_{k_1,k_2} \cdot \mathbb{L} \nabla w_{k_1,k_2}}{\int_{\mathbb{R}^d} |\nabla w_{k_1,k_2}|^2} \right\rangle - r_{k_1}/2.$$

Likewise, there exists a set of events of full measure Ω_2 such that for all k_1 and all $\omega \in \Omega_2$ there exists $w_{k_1} \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$ such that

$$\frac{\int_{\mathbb{R}^d} \nabla w_{k_1}(x) \cdot \mathbb{L}(x,\omega) \nabla w_{k_1}(x) dx}{\int_{\mathbb{R}^d} |\nabla w_{k_1}(x)|^2 dx} \leq \lambda + r_{k_1}/2.$$

Since $\Omega_1 \cap \Omega_2$ has full measure, there exists $\omega \in \Omega$ such that for all k_1, if we choose k_2 large enough such that $w_{k_1} \in C^\infty_{0,L_{k_2}}(\mathbb{R}^d, \mathbb{R}^d)$. We then have

$$\begin{aligned} \lambda + r_{k_1}/2 &\geq \frac{\int_{\mathbb{R}^d} \nabla w_{k_1}(x) \cdot \mathbb{L}(x,\omega) \nabla w_{k_1}(x) dx}{\int_{\mathbb{R}^d} |\nabla w_{k_1}(x)|^2 dx} \\ &\geq \inf_{v \in C^\infty_{0,L_{k_2}}(\mathbb{R}^d,\mathbb{R}^d)} \frac{\int_{\mathbb{R}^d} \nabla v(x) \cdot \mathbb{L}(x,\omega) \nabla v(x) dx}{\int_{\mathbb{R}^d} |\nabla v(x)|^2 dx} \\ &\geq \left\langle \frac{\int_{\mathbb{R}^d} \nabla w_{k_1,k_2} \cdot \mathbb{L} \nabla w_{k_1,k_2}}{\int_{\mathbb{R}^d} |\nabla w_{k_1,k_2}|^2} \right\rangle - r_{k_1}/2, \end{aligned}$$

which yields the desired estimate (3.25). Without loss of generality we assume that $\left\langle \int_{\mathbb{R}^d} |\nabla v_r|^2 \right\rangle = 1$.

By ergodicity, (3.25) implies that almost surely

$$\frac{\int_{\mathbb{R}^d} \nabla v_r \cdot \mathbb{L} \nabla v_r}{\int_{\mathbb{R}^d} |\nabla v_r|^2} \leq \lambda + 1/r.$$

In order to use the Bloch transform we need the support of v_r to be included in a single open bounded set for almost every $\omega \in \Omega$. This is incompatible with stationarity, and we have to truncate the field v_r: for all $R > 0$ let $Q_R = (-R/2, R/2)^d$, and define $v_{r,R} \in L^2(\Omega, C^\infty_{Q_R}(\mathbb{R}^d, \mathbb{R}^d))$ by

$$v_{r,R}(x,\omega) \;:=\; v_r(x,\omega) \times \rho(x/R),$$

where ρ is a smooth cut-off function with support in $(-1/2, 1/2)^d$. For all $\gamma \in [0, 2\pi)^d$, we consider the Bloch transform $\tilde{v}_{r,R,\gamma}$ of $v_{r,R}$. Let $Q^* = [0, 2\pi)^d$. We then have for almost every $\omega \in \Omega$:

$$\begin{aligned}
&\int_Q \int_{Q^*} \nabla \tilde{v}_{r,R,\gamma}(x,\omega) \cdot \mathbb{L}(x,\omega) \nabla \tilde{v}_{r,R,\gamma}(x,\omega) d\gamma dx \\
&= \int_Q \int_{z \in \mathbb{Z}^d} \int_{z' \in \mathbb{Z}^d} \int_{Q^*} e^{-i\gamma\cdot(z-z')} \nabla v_{r,R}(x+z, \tau_{-z}\omega) \cdot \mathbb{L}(x,\omega) \nabla v_{r,R}(x+z', \tau_{-z'}\omega) d\gamma dx \\
&= (2\pi)^d \int_Q \int_{z \in \mathbb{Z}^d} \nabla v_{r,R}(x+z, \tau_{-z}\omega) \cdot \mathbb{L}(x,\omega) \nabla v_{r,R}(x+z, \tau_{-z}\omega) dx \\
&= (2\pi)^d \int_Q \int_{z \in \mathbb{Z}^d} \nabla v_{r,R}(x+z, \tau_{-z}\omega) \cdot \mathbb{L}(x+z, \tau_{-z}\omega) \nabla v_{r,R}(x+z, \tau_{-z}\omega) dx,
\end{aligned}$$

using that $\int_{Q^*} e^{-i\gamma\cdot(z-z')} d\gamma = (2\pi)^d \delta_{zz'}$, and the stationarity of $\mathbb{L}$. Since the translation group is measure preserving, the expectation of this identity turns into

$$\begin{aligned}
&\left\langle \int_Q \int_{Q^*} \nabla \tilde{v}_{r,R,\gamma}(x,\omega) \cdot \mathbb{L}(x,\omega) \nabla \tilde{v}_{r,R,\gamma}(x,\omega) d\gamma dx \right\rangle \\
&= (2\pi)^d \left\langle \int_Q \int_{z \in \mathbb{Z}^d} \nabla v_{r,R}(x+z,\omega) \cdot \mathbb{L}(x+z,\omega) \nabla v_{r,R}(x+z,\omega) dx \right\rangle \\
&= (2\pi)^d \left\langle \int_{\mathbb{R}^d} \nabla v_{r,R} \cdot \mathbb{L} \nabla v_{r,R} \right\rangle.
\end{aligned}$$

Likewise,

$$\left\langle \int_Q \int_{Q^*} |\nabla \tilde{v}_{r,R,\gamma}(x,\omega)|^2 d\gamma dx \right\rangle \;=\; (2\pi)^d \left\langle \int_{\mathbb{R}^d} |\nabla v_{r,R}|^2 \right\rangle.$$

By definition of λ_1 and Fubini's theorem, since for all $\gamma \in [0, 2\pi)^d$, $\tilde{v}_{r,R,\gamma}(x,\omega) = e^{i\gamma\cdot x} \hat{v}_{r,R,\gamma}(x,\omega)$ with $\hat{v}_{r,R,\gamma} \in \mathcal{H}^1_1$, we have

$$\int_{Q^*} \left\langle \int_Q \nabla \tilde{v}_{r,R,\gamma}(x,\omega) \cdot \mathbb{L}(x,\omega) \nabla \tilde{v}_{r,R,\gamma}(x,\omega) dx \right\rangle d\gamma \;\geq\; \lambda_1 \int_{Q^*} \left\langle \int_Q |\nabla \tilde{v}_{r,R,\gamma}(x,\omega)|^2 dx \right\rangle d\gamma,$$

so that

$$\left\langle \int_{\mathbb{R}^d} \nabla v_{r,R} \cdot \mathbb{L} \nabla v_{r,R} \right\rangle \;\geq\; \lambda_1 \left\langle \int_{\mathbb{R}^d} |\nabla v_{r,R}|^2 \right\rangle.$$

For almost every $\omega \in \Omega$, $v_r(\cdot,\omega)$ has compact support in $\mathbb{R}^d$ and $\mathbb{L}$ is bounded. Hence the Lebesgue dominated convergence theorem ensures that

$$\lim_{R\to\infty}\int_{\mathbb{R}^d}\nabla v_{r,R}\cdot\mathbb{L}\nabla v_{r,R} = \int_{\mathbb{R}^d}\nabla v_r\cdot\mathbb{L}\nabla v_r$$
$$\lim_{R\to\infty}\int_{\mathbb{R}^d}|\nabla v_{r,R}|^2 = \int_{\mathbb{R}^d}|\nabla v_r|^2$$

almost surely. Since $\int_{\mathbb{R}^d}|\nabla v_r|^2 = 1$ and $\mathbb{L}$ is bounded, $\int_{\mathbb{R}^d}\nabla v_{r,R}\cdot\mathbb{L}\nabla v_{r,R}$ and $\int_{\mathbb{R}^d}|\nabla v_{r,R}|^2$ are bounded uniformly in R and ω. Another use of the Lebesgue dominated convergence theorem finally yields

$$\left\langle\int_{\mathbb{R}^d}\nabla v_r\cdot\mathbb{L}\nabla v_r\right\rangle \geq \lambda_1\left\langle\int_{\mathbb{R}^d}|\nabla v_r|^2\right\rangle,$$

from which we deduce

$$\lambda + 1/r \geq \lambda_1,$$

and therefore the desired inequality $\lambda \geq \lambda_1$ by the arbitrariness of r.

3.5.3 Application to nonlinear problems

Recall that W satisfies Hypothesis 2 for a *discrete* group of translation (see Remark 4), so that Theorem 19 ensures the existence of a deterministic homogenized integrand W_{hom}. In this subsection we'd like to apply the analysis sketched above for linear problems to this nonlinear case. To this aim we need linearization and homogenization to commute, which we ensure by the following assumption. Let $\Lambda \in \mathcal{M}^d$ be fixed.

Hypothesis 3 *There exists $t_0 > 0$ such that for all $0 \leq t \leq t_0$ and all $H = a\times b \in \mathcal{M}^d$ with $|H| = 1$, there exists a stationary field $\nabla\phi_{\Lambda+tH} \in L^\infty(\mathbb{R}^d\times\Omega)$ which satisfies*

$$W_{\mathrm{hom}}(\Lambda + tH) = \lim_{R\to\infty}⨍_{Q_R} W(x, \Lambda + tH + \nabla\phi_{\Lambda+tH}(x,\omega),\omega)dx \tag{3.26}$$

almost surely, and such that

$$\|\nabla\phi_{\Lambda+tH} - \nabla\phi_{\Lambda+tH}\|_{L^\infty(\mathbb{R}^d\times\Omega)} \leq r(t) \tag{3.27}$$

where $r(t)\to 0$ as $t\to 0$.

Equation (3.26) states that there exists a corrector whose gradient is stationary (which is not known a priori since W_{hom} is only defined using the subadditive ergodic theorem, and not using the corrector equation), while (3.27) requires that no discontinuous bifurcation of correctors occurs. To prove the validity of these assumptions remains an open problem in any framework which makes sense (even in the case of stochastic homogenization of discrete linear elliptic equations with i. i. d. coefficients !).

Under this assumption, the linearized elasticity tensor is defined by

$$\mathbb{L}_\Lambda(x,\omega) := \frac{\partial^2 W}{\partial \Lambda^2}(x, \Lambda + \nabla\phi_\Lambda(x,\omega), \omega),$$

and the homogenized Piola stress tensor by

$$\Pi_{\mathrm{hom}}(\Lambda) := \lim_{R\to\infty} \fint_{Q_R} \frac{\partial W}{\partial \Lambda}(x, \Lambda + tH + \nabla\phi_{\Lambda+tH}(x,\omega), \omega)dx.$$

Note that the limit exists almost surely and is deterministic by the ergodic theorem since the integrand is $\mathbb{Z}^d$-stationary. We may then define the associated ellipticity constants $\lambda(\Lambda)$, $\lambda_4(\Lambda)$, and $\lambda_6(\Lambda)$. Provided $\lambda(\Lambda) \geq 0$, $\mathbb{L}_\Lambda$ can be homogenized and we denote by $\mathbb{L}_{\mathrm{hom}}(\Lambda)$ the homogenized elasticity tensor.

The main result of this section is the following (partial) characterization of the strong ellipticity of W_{hom} at Λ:

Theorem 22 *Let $1 < p < +\infty$. Assume that W satisfies Hypotheses 2 (for a discrete translation group) and 3, that for almost every $x \in \mathbb{R}^d$ and $\omega \in \Omega$ the function $F \mapsto W(x, F, \omega)$ is three times continuously differentiable on $\mathcal{M}^d$ and that it satisfies for almost every $x \in \mathbb{R}^d, \omega \in \Omega$ and all $F \in \mathcal{M}^d$,*

$$\left|\frac{\partial W}{\partial F}(x, F, \omega)\right| \leq C(1 + |F|^{p-1}),$$
$$\left|\frac{\partial W}{\partial F}(x, F, \omega)\right| \leq h(F),$$

for some constant $C > 0$ and some locally bounded function h.

(i) If $\lambda_4(\Lambda) > 0$, then for all $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$,

$$W_{\mathrm{hom}}(\Lambda + tH) = W_{\mathrm{hom}}(\Lambda) + \Pi_{\mathrm{hom}}(\Lambda) \cdot Ht + \frac{t^2}{2} H \cdot \mathbb{L}_{\mathrm{hom}}(\Lambda)H + o(t^2)$$

as $t \to 0$ (and $H \cdot \mathbb{L}_{\mathrm{hom}} H \geq \lambda_4(\Lambda)$).

(ii) If $\lambda_4(\Lambda) = 0$ and $\lambda_6(\Lambda) > 0$, then there exists $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$ such that

$$W_{\mathrm{hom}}(\Lambda + tH) = W_{\mathrm{hom}}(\Lambda) + \Pi_{\mathrm{hom}}(\Lambda) \cdot Ht + o(t^2)$$

as $t \to 0$.

This theorem shows that W_{hom} is strictly strongly elliptic at Λ in the case *(i)*, and loses strong ellipticity in the case *(ii)*. As in [38], we also have results when $\lambda_6(\Lambda) = 0$, but they are less complete. The proof of this theorem is a direct adaptation of the proof of [38] provided we have the stochastic linear theory of Subsection 3.5.2.

3.6 Perspectives

The following three problems are in the continuation of the results of this section:

- are there nontrivial cases for which the cell integrand coincides with the homogenized integrand in the periodic case ?
- homogenization of one-dimensional elasto-dynamics;
- homogenization of polyconvex energy densities which do not satisfy the standard growth condition from above.

Again, the difficulty to deal with periodic homogenization in nonlinear elasticity partly comes from the asymptotic character of the homogenization formula (3.4). We've seen that, in general, the quasiconvex envelope of the cell integrand does not coincide with the homogenized integrand. This does not preclude the existence of interesting cases for which $W_{\mathrm{hom}} \equiv W_{\mathrm{cell}}$. The asymptotic formula does not reduce to the unitary cell formula in general due to nonconvexity. The simplest nontrivial such energy density we can consider splits into two parts: a homogeneous nonconvex part, and a heterogeneous convex part. For $d = n = 2$, a typical example is

$$W(x, \Lambda) := a(x)|\Lambda|^4 + f(\det \Lambda),$$

with f convex, non-negative, and at most quadratic at infinity, and $0 < \alpha \leq a(x) \leq \beta < \infty$ a Q-periodic function. The precise question is: prove or disprove that $W_{\mathrm{cell}} \equiv W_{\mathrm{hom}}$ in this case. Note that all the known counterexamples to the equality $W_{\mathrm{cell}} = W_{\mathrm{hom}}$ (Section 3.3 and [2, 64]) crucially rely on the nonconvexity of the heterogeneous part of the energy density. A related question will appear Subsection 4.5.3.

The homogenization of scalar hyperbolic equations has been treated by Dalibard [27] using a kinetic formulation. Since there also exists a kinetic formulation for one-dimensional elastodynamics [77], one may hope to successfully combine the two approaches to homogenize the system of one-dimensional elasto-dynamics of hyperelastic materials.

The last mentioned open question is related to the treatment of the volumetric term in nonlinear elasticity. A standard requirement in hyperelasticity is the blow up of the energy density when the determinant of the deformation gradient tends to zero. This models the fact that one has to pay an infinite amount of energy to compress a piece of material to a point. This behavior is incompatible with the standard growth condition from above in (3.1). In his fundamental contribution [5] Ball has shown that integral functionals whose integrands are polyconvex are lower semi-continuous for the weak topology of $W^{1,p}(D)$ and may satisfy the desired physical behavior in compression, which has allowed him to obtain general existence results in nonlinear elasticity. Yet polyconvexity can be lost by homogenization [6, 14] (even if $W(x, \cdot)$ is polyconvex for almost every $x \in Q$, the associated homogenized integrand W_{hom} may be quasiconvex but not polyconvex), so that it is not clear whether polyconvexity is the best notion to start with at this stage. There are at least two approaches to prove the periodic homogenization result of Theorem 18:

the abstract approach by Braides [13] which relies on the growth condition (it uses integral representation results whose proofs crucially rely on the growth condition), and a more direct approach by Müller [64] for which the growth condition from above is not that relevant (in particular Müller extends homogenization results to the case of convex integrands which do not satisfy the standard growth condition from above). The difficulty to prove the homogenization result for general polyconvex integrands is the construction of the recovery sequence. A crucial density result was obtained recently [43] for $d = n = 2$ by Iwaniec, Kovalev, and Onninen. This result could be used to complete the proof of the homogenization of general polyconvex integrands for $d = n = 2$ when Dirichlet boundary conditions are considered.

4

Homogenization of discrete systems and derivation of rubber elasticity

In his review paper on open problems in elasticity [4], Ball mentions the issue of establishing the status of nonlinear elasticity theory for rubber with respect to the point of view of polymer physics. The aim of this chapter is to fill (part of) this gap.

We start with the derivation of a model at the level of the polymer chain network from a full statistical mechanics description. This model is derived using heuristic arguments only. A crucial geometric object is the network, which we assume to be given by an ergodic random point set.

In the second section, we turn to the core of the analysis: we let the typical size of the polymer chains go to zero, and rigorously derive a continuous model using a discrete stochastic homogenization process. The integral functional obtained at the limit is proved to satisfy several typical requirements of nonlinear elasticity models (such as hyperelasticity, frame-invariance, isotropy, minimality at identity, etc.). The homogenized integrand is given by an asymptotic homogenization formula.

In the third section we make the model more specific by considering a particular example of random point set, the so-called random parking measure. This point set has the property to be ergodic and statistically isotropic (which implies the isotropy of the homogenized model).

In the fourth section we turn to the numerical approximation of the homogenized model. We propose and test a numerical method to approximate the random parking measure, and the asymptotic homogenization formula. This allows to compare this homogenized model to the popular mechanical experiments by Treloar, and show they are in very good agreement.

In the fifth section we address the issue of strong ellipticity. Rubber is usually a strictly strongly elliptic material. It is therefore to be expected that the discrete homogenization process yields such a property. We show that the specific form of the polymer chain free energies plays an important role for the strong ellipticity of the homogenized integrand.

The aim of the last section is to construct an analytical approximation of the homogenized energy density. We devise a numerical procedure to approximate the homogenized energy density in a subclass of Ogden materials. The excellent adequacy between the homogenized energy density and its analytical approximation shows the capability of standard energy densities used in the mechanical community to correctly fit the models obtained by discrete homogenization.

In this chapter we go from polymer physics to Ogden laws, using modeling, analysis, and numerical simulations.

4.1 A discrete model for rubber [GLTV]

We consider a macroscopic sample of natural rubber D, whose boundary is linearly deformed through the map $x \mapsto \Lambda \cdot x$, $\Lambda \in \mathcal{M}^3_+$. The sample is made of a network of cross-linked polymer chains. The cross-links are assumed to be permanent. In this first (rough) model, we neglect entanglements, that is, we neglect topological constraints (this will be made clear in the definition of the network). Each polymer chain is itself made of a given number of monomers: the energy of a chain for a given configuration is obtained through the probability density of a random walk (see for instance [50], [86]). We assume that each monomer is surrounded by a fixed volume (from which other monomers are excluded), and that the network of chains is packed and almost incompressible. This assumption adds a volumetric term to the energy which depends on the configuration of the network. This volumetric term accounts for the interaction between the chains (which does not appear in the energy of one single chain). Note that the relevant scale associated with this contribution is much smaller than the one corresponding to the contribution associated with the random walk variable.

A polymer chain network is parametrized as follows: we denote by u the positions of the cross-links, and by $s = \{s_i\}$ the positions of the monomers of the chain i. The Hamiltonian of the system can be split into two parts:

$$H(u, s) = H_{\mathrm{vol}}(u, s) + \sum_i H_i(u, s_i).$$

The first part $H_{\mathrm{vol}}(u, s)$ is the volumetric energy of the network, which models the interactions between the chains, whereas the second part $H_i(u, s_i)$ is the energy of each chain as if it were isolated (and for which u prescribes the end-to-end vector, and s_i describes the positions of the monomers constituting the chain).

At finite temperature $\beta = \frac{1}{k_B T}$, the Gibbs distribution yields the following formula for the free energy of a given deformed network:

$$\begin{aligned} F(\Lambda, D) &= -\frac{1}{\beta} \ln Z \\ &= -\frac{1}{\beta} \ln \left[\int_U \int_{\prod S_i(u)} \exp\left(-\beta H_{\mathrm{vol}}(u, s) - \sum_i \beta H_i(u, s_i) \right) du \prod_i ds_i \right], \end{aligned}$$

where Z is the partition function, U is the set of admissible positions of the cross-links (satisfying the constraint on the boundary), and $S_i(u)$ denotes the set of admissible positions of the monomers composing the chain i whose head and tail are prescribed by u.

This free energy is far from being explicit. However, it is possible to further simplify the problem and still capture some interesting features. We present a heuristic reasoning which leads to the decoupling of the s_i variables. We first assume that $H_{\rm vol}(u,s) = H_{\rm vol}(u)$ only depends on u and not on s, which amounts to replacing the excluded volume constraint around monomers by an excluded volume constraint between cross-links. Note that this is a rather strong assumption whose effect is to make chains interact via their cross-links only: this decouples the variables s_i from one another. We may then rewrite the free energy as follows:

$$F(\Lambda,D) = -\frac{1}{\beta}\ln\Bigg[\int_U \exp\Bigg(-\beta H_{\rm vol}(u) \\ +\beta\sum_i \frac{1}{\beta}\ln\Big[\int_{S_i(u)} \exp\Big(-\beta H_i(u,s_i)\Big)ds_i\Big]\Bigg)du\Bigg].$$

We thus have the following effective Hamiltonian:

$$H_\Lambda(u,\beta) \;:=\; H_{\rm vol}(u) - \sum_i \frac{1}{\beta}\ln\Big[\int_{S_i(u)} \exp\Big(-\beta H_i(u,s_i)\Big)ds_i\Big]. \tag{4.1}$$

We then make the strong assumption that one can replace the integration on U by taking the infimum. This amounts to treating the cross-links at zero temperature and all the other monomers at finite temperature. We are thus lead to

$$\frac{F(\Lambda,D)}{|D|} \;\simeq\; \frac{\inf_u H_\Lambda(u,\beta)}{|D|}. \tag{4.2}$$

In terms of orders of magnitude, recall that polymer chains are typically $100nm$ long whereas the macroscopic sample is of the order of the cm, which yields a factor 10^5. Hence, provided D is a macroscopic sample, (4.2) will be close to the "thermodynamic limit"

$$W_{\rm V}(\Lambda) \;:=\; \lim_{|D|\to\infty} \frac{F(\Lambda,D)}{|D|}, \tag{4.3}$$

where D properly invades $\mathbb{R}^3$ (see for instance [80]). For such a limit to exist, the network of polymer chains should have some ergodic property: either the network has some periodic structure (yet we are not dealing with crystals), or the network should yield spatial decorrelations (in a statistical or stochastic framework) — although other less physically relevant properties could also be considered stricto sensu.

Such a limiting process is addressed in the following section.

4.2 Homogenization of discrete systems on stochastic lattices [ACG]

In this section we introduce a general framework which covers the discrete model for rubber and which is suitable for the analysis. In particular, we first make precise what a discrete network is by the use of stochastic lattices — or more precisely random point sets. We then introduce a "discrete" functional, and prove a homogenization result in the spirit of Theorem 19. We then study the properties satisfied by the homogenized integrand associated with the network-based model for rubber. The details on how to make the network-based model enter this framework will be given in Section 4.4.

4.2.1 Stochastic lattices

Definition 19 *Let $\Sigma \subset \mathbb{R}^d$ be a locally finite set of points. We say that Σ is* general *if no $d+1$ points lie in the same hyperplane and if no $d+2$ points lie in the same hypersphere.*

Definition 20 *Let $0 < \rho_1 \leq \rho_2 \leq \infty$. Given $D \subset \mathbb{R}^d$, suppose Σ is a subset of D. Then Σ is said to be (ρ_1, ρ_2)*-admissible *in D iff for all $x \neq y \in \Sigma$, $|x-y| \geq \rho_1$, and for all $z \in D$, $B(z, \rho_2) \cap \Sigma \neq \emptyset$, where $B(z, \rho_2)$ denotes the open ball centred in z of radius ρ_2 (and $B(z, \infty) := \mathbb{R}^d$)). The set of (ρ_1, ρ_2)-admissible point sets in $\mathbb{R}^d$ is denoted by $\mathcal{A}_{\rho_1,\rho_2}$.*

In other words, an admissible point set in D is one which satisfies the hard-core and non-empty space conditions. We shall sometimes write simply 'admissible' for 'admissible in $\mathbb{R}^d$'.

Definition 21 *A random subset $\mathcal{L}$ of $\mathbb{R}^d$ is called a* point process*. We say that a point process $\mathcal{L}$ in $\mathbb{R}^d$ is* isotropic *if the distribution of $\mathcal{L}$ and of $\mathbf{R}\mathcal{L}$ are the same for every rotation $\mathbf{R} \in SO_d$.*

Let $0 < \rho_1$. Suppose $\Sigma \in \mathcal{A}_{\rho_1,\infty}$. If Σ is general and its convex hull has strictly positive d-Lebesgue measure, then there is a unique Delaunay triangulation of the convex hull of Σ, by simplices with edges given by the edges of the Delaunay graph of Σ (see [30] for Delaunay triangulations of $\mathbb{R}^d$, and for instance [35] for Delaunay triangulations of a bounded domain). If Σ is not general, there exists a possibly non-unique Delaunay triangulation of the convex hull of Σ.

4.2.2 Discrete homogenization result

Recall that d and n are dimensions. Let $0 < \rho_1 < \rho_2 < \infty$. Suppose $\Sigma \in \mathcal{A}_{\rho_1,\rho_2}$. Let $\mathcal{T}(\Sigma)$ denote a Delaunay triangulation of the convex hull of Σ, chosen by some deterministic rule if the Delaunay triangulation is not unique. We denote by $\mathcal{N}(\Sigma)$ the associated neighbour pairs, that is, those unordered pairs of points $\{x, y\}$ such that (x, y) is an edge of $\mathcal{T} := \mathcal{T}(\Sigma)$. For all $\varepsilon > 0$ and for every Lipschitz bounded domain D of $\mathbb{R}^d$, this allows us to uniquely define a space of continuous piecewise-affine functions $\mathfrak{S}^D_\varepsilon(\Sigma)$ on $D \cap \varepsilon\Sigma$:

$$\mathfrak{S}^D_\varepsilon(\Sigma) := \{u \in C^0(D, \mathbb{R}^n) \,\big|\, \forall T \in \mathcal{T}(\Sigma), \text{ with } \varepsilon T \cap D \neq \emptyset,\, u_{|\varepsilon T \cap D} \text{ is affine}\}, \tag{4.4}$$

From now on, we identify $u : \varepsilon\mathcal{L} \cap D \to \mathbb{R}^n$ with its class of piecewise-affine interpolations (still denoted by u) in $\mathfrak{S}^D_\varepsilon(\Sigma) \subset W^{1,\infty}(D, \mathbb{R}^n)$. Note that the extension of $u : \varepsilon\mathcal{L} \cap D \to \mathbb{R}^n$ to $D \setminus \cup_{T \in \mathcal{T}, T \subset \overline{D}} T$ is not uniquely defined — as we shall see, the energy under consideration does not depend on the extension. In order to define an energy functional on the set $\mathfrak{S}^D_\varepsilon(\Sigma)$, we first introduce the following energy functions:

Definition 22 *Let $p > 1$. We denote by $\mathcal{U}_p$ the subset of functions f_{nn} of $\mathcal{C}^0(\mathbb{R}^d \times \mathbb{R}^n, [0, +\infty))$ for which there exists $C > 0$ such that, for all $z \in \mathbb{R}^d$ and $s \in \mathbb{R}^n$,*

$$\frac{1}{C}|s|^p - C \leq f_{nn}(z, s) \leq C(|s|^p + 1). \tag{4.5}$$

We denote by $\mathcal{V}_p$ the subset of functions W_{vol} of $C^0(\mathcal{M}^{d\times n}, [0, +\infty))$ for which there exists $C > 0$ such that for all $\Lambda \in \mathcal{M}^{d\times n}$,

$$W_{\mathrm{vol}}(\Lambda) \leq C(|\Lambda|^p + 1). \tag{4.6}$$

Let $p > 1$ and $f_{nn} \in \mathcal{U}_p, W_{\mathrm{vol}} \in \mathcal{V}_p$. For all $u \in L^p(D, \mathbb{R}^n)$ we define the energy of the network defined by u by

$$F_\varepsilon^D(\Sigma, u) := \begin{cases} F_{nn,\varepsilon}^D(\Sigma, u) + F_{\mathrm{vol},\varepsilon}^D(\Sigma, u) & \text{if } u \in \mathfrak{S}_\varepsilon^D(\Sigma), \\ +\infty & \text{otherwise}, \end{cases} \tag{4.7}$$

where (with $\overline{D}$ denoting the closure of D) we have set

$$F_{nn,\varepsilon}^D(\Sigma, u) = \sum_{\substack{(x,y) \in \mathcal{N}(\Sigma) \\ [\varepsilon x, \varepsilon y] \subset \overline{D}}} \varepsilon^d f_{nn}\left(y - x, \frac{u(\varepsilon y) - u(\varepsilon x)}{\varepsilon|y - x|}\right), \tag{4.8}$$

and

$$F_{\mathrm{vol},\varepsilon}^D(\Sigma, u) = \sum_{\substack{T \in \mathcal{T}(\Sigma) \\ T \subset \overline{D}_{\varepsilon^{-1}}}} \varepsilon^d |T| W_{\mathrm{vol}}(\nabla u_{|\varepsilon T}). \tag{4.9}$$

As announced, if $u_1, u_2 \in \mathfrak{S}_\varepsilon^D(\Sigma)$ are such that $u_1 = u_2$ on $\cup_{T\in\mathcal{T}, T\subset\varepsilon^{-1}\overline{D}}\varepsilon T$, then $F_\varepsilon^D(\Sigma, u_1) = F_\varepsilon^D(\Sigma, u_2)$.

Recall d_∞ denotes the distance of the supremum in $\mathbb{R}^d$. The following Γ-convergence (or discrete homogenization) result holds

Theorem 23 *Let $0 < \rho_1 < \rho_2 < \infty$, let $1 < p < \infty$, and let D be a bounded Lipschitz domain of $\mathbb{R}^d$. Let $\mathcal{L}$ be a stationary and ergodic point process in $\mathcal{A}_{\rho_1,\rho_2}$, which is almost surely general. Let f_{nn} and W_{vol} be of class $\mathcal{U}_p$ and $\mathcal{V}_p$, respectively. Let $F_\varepsilon^D(\mathcal{L})$ be the energy functional given by (4.7). Then for every sequence $\varepsilon_k \to 0$, the functionals $F_{\varepsilon_k}^D(\mathcal{L})$ Γ-converge to the deterministic integral functional $F_{\mathrm{hom}}^D : L^p(D, \mathbb{R}^n) \to [0, +\infty]$ defined by*

$$F_{\mathrm{hom}}^D(u) := \begin{cases} \displaystyle\int_D W_{\mathrm{hom}}(\nabla u(x))dx & \textit{if} \quad u \in W^{1,p}(D, \mathbb{R}^n), \\ +\infty & \textit{otherwise}. \end{cases}$$

where $W_{\rm hom} : \mathcal{M}^{d\times n} \to [0,+\infty)$ *is a deterministic quasiconvex function which depends only on* f_{nn}*,* $W_{\rm vol}$*, and on the point process, and which satisfies a standard growth condition* (3.1) *of order p. In addition it satisfies the following asymptotic homogenization formula almost surely:*

$$W_{\rm hom}(\Lambda) = \lim_{R\to\infty} \frac{1}{|Q_R|} \inf_u \left\{ F_1^{Q_R}(\mathcal{L},u) \;\middle|\; u \in \mathfrak{S}_1^{Q_R}(\mathcal{L}) \text{ such that } u(x) = \Lambda \cdot x \right.$$
$$\left. \text{if } x \in \mathcal{L}\cap Q_R \text{ and } d_\infty(x,\partial Q_R) \le 2\rho_2 \right\}. \quad (4.10)$$

Theorem 23 is a particular case of a more general theorem (see [ACG]) which covers not only nearest-neighbour interactions but also long-range (yet integrable) interactions. The structure of the proof of this result is standard and follows the abstract approach used by Marcellini [58], by Braides [13, 15], and by Dal Maso and Modica [26] to prove homogenization results for integral functionals. This approach has been adapted by Alicandro and Cicalese [1] to treat the homogenization of discrete systems on *periodic* lattices (i. e. $\Sigma = \mathbb{Z}^d$).

In the present stochastic setting it consists in four steps:

Step 1 Prove a so-called Γ-compactness result for functionals defined on sets (say Lipschitz domains) and functions;
Step 2 Prove that any Γ-limit is the restriction of a Borel function by the De Giorgi-Letta criteria;
Step 3 Using the Buttazzo-Dal Maso characterization of integral functionals, prove that any Γ-limit is an integral functional (on some Sobolev space) associated with a quasiconvex integrand;
Step 4 Conclude that the integrand is deterministic and homogeneous in space by a suitable use of the subadditive ergodic theorem and the characterization of integral functionals by their minima.

The structure of the proof is quite clear and essentially dates back to Dal Maso and Modica [26]. The central achievement of the proof of Theorem 23 lies in its level of technicality.

4.2.3 Qualitative properties

Once we have proved the existence of a homogenized model, one may try to identify the (mechanical) properties satisfied by the homogenized integrand. We now restrict the analysis to nonlinear elasticity and set $d = n$. Properties of interest include:

- hyperelasticity,
- frame-invariance,
- isotropy,
- blow up of the integrand $W_{\rm hom}(\Lambda_k)$ when $\det \Lambda_k \to 0^+$,

- strong ellipticity of W_{hom}.

In this section we address the first three properties. The blow up property will be touched upon in Section 4.7 whereas Section 4.5 is precisely dedicated to the issue of strong ellipticity.

Hyperelasticity is obvious since the "homogenized material" is chacterized by an energy which is the integral of an energy density depending locally on the strain gradient.

We also have the following two results:

Theorem 24 *In addition to the assumptions of Theorem 23, let us assume that there exists* $\tilde{f}_{nn} : \mathbb{R}^d \times [0, +\infty) \to [0, +\infty)$ *such that*

$$\begin{aligned} &\text{for all } z_1, z_2 \in \mathbb{R}^d, \qquad f_{nn}(z_1, z_2) = \tilde{f}_{nn}(z_1, |z_2|), \\ &\text{for all } \Lambda \in \mathcal{M}^d, \mathbf{R} \in SO_d, W_{\text{vol}}(\mathbf{R}\Lambda) = W_{\text{vol}}(\Lambda). \end{aligned}$$

Then the energy density W_{hom} *is frame-invariant: for all* $\Lambda \in \mathcal{M}^d$ *and* $\mathbf{R} \in SO_d$*,*

$$W_{\text{hom}}(\mathbf{R}\Lambda) = W_{\text{hom}}(\Lambda).$$

Theorem 25 *In addition to the assumptions of Theorem 23, let us assume that the point process* $\mathcal{L}$ *is isotropic and that there exists* $\bar{f}_{nn} : [0, +\infty) \times \mathbb{R}^d \to [0, +\infty)$ *such that*

$$\begin{aligned} &\text{for all } z_1, z_2 \in \mathbb{R}^d, \qquad f_{nn}(z_1, z_2) = \bar{f}_{nn}(|z_1|, z_2), \\ &\text{for all } \Lambda \in \mathcal{M}^d, \mathbf{R} \in SO_d, W_{\text{vol}}(\Lambda\mathbf{R}) = W_{\text{vol}}(\Lambda). \end{aligned}$$

Then the energy density W_{hom} *is isotropic: for all* $\Lambda \in \mathcal{M}^d$ *and* $\mathbf{R} \in SO_d$*,*

$$W_{\text{hom}}(\Lambda\mathbf{R}) = W_{\text{hom}}(\Lambda).$$

Note that if $\mathcal{L} = \mathbb{Z}^d$, the associated homogenized integrand cannot be isotropic. What is not clear a priori is whether stationary ergodic isotropic point processes in $\mathcal{A}_{\rho_1,\rho_2}$ do indeed exist !

4.3 Existence of isotropic stochastic lattices and approximation result [GP]

In this section we show that the so-called random parking measure studied by Penrose [76] is a stationary ergodic isotropic point process in $\mathcal{A}_{\rho_1,\rho_2}$ (for some suitable $0 < \rho_1 < \rho_2$). In addition this point process can be easily approximated on bounded domains Q_R, and these approximations can be used in (4.10) in place of $\mathcal{L}$ (the convergence holding almost surely as well).

4.3.1 Random parking measure

Rényi's model of random parking (also known as random sequential adsorption or random sequential packing), is defined in d dimensions as follows. A parameter $\rho_0 > 0$ is specified, and open balls of radius ρ_0 (say $B_{1,R}, B_{2,R}, \dots$) arrive sequentially and uniformly at random in the d-dimensional cube Q_R, $R > 2\rho_0$. The first ball is packed. And recursively, an incoming ball is accepted if it does not overlap the balls that are already packed. Since Q_R is bounded, this process stops after a finite number of arrivals almost surely. We denote by ξ^R the point process associated with the centers of the packed balls in Q_R. Penrose [76] proved that the random parking measure ξ^R in Q_R weakly converges in the sense of measures to a measure ξ in $\mathbb{R}^d$, called the random parking measure in $\mathbb{R}^d$.

We first quickly recall the graphical construction of the random parking measure ξ^A in some Borel set $A \subset \mathbb{R}^d$, as introduced by Penrose [76]. Let $\mathcal{P}$ be a homogeneous Poisson process of unit intensity in $\mathbb{R}^d \times \mathbb{R}^+$. An *oriented graph* is a special kind of directed graph in which there is no pair of vertices $\{x, y\}$ for which both (x, y) and (y, x) are included as directed edges. We shall say that x is a *parent* of y and y is an *offspring* of x if there is an oriented edge from x to y. By a *root* of an oriented graph we mean a vertex with no parent.

The graphical construction goes as follows. Let $\rho_0 > 0$ and let B denote the Euclidean ball in $\mathbb{R}^d$ of radius ρ_0 centred at the origin. Make the points of the Poisson process $\mathcal{P}$ on $\mathbb{R}^d \times \mathbb{R}^+$ into the vertices of an infinite oriented graph, denoted by $\mathcal{G}$, by putting in an oriented edge $(X, T) \to (X', T')$ whenever $(X' + B) \cap (X + B) \neq \emptyset$ and $T < T'$. For completeness we also put an edge $(X, T) \to (X', T')$ whenever $(X'+B) \cap (X+B) \neq \emptyset$, $T = T'$, and X precedes X' in the lexicographical order — although in practice the probability that $\mathcal{P}$ generates such an edge is zero. It can be useful to think of the oriented graph as representing the spread of an "epidemic" through space over time; each time an individual is "born" at a Poisson point in space-time, it becomes (and stays) infected if there is an earlier infected point nearby in space (in the sense that the translates of B centred at the two points overlap). This graph determines which items have to be accepted.

For $(X, T) \in \mathcal{P}$, let $C_{(X,T)}$ (the "cluster at (X, T)") be the (random) set of ancestors of (X, T), that is, the set of $(Y, U) \in \mathcal{P}$ such that there is an oriented path in $\mathcal{G}$ from (Y, U) to (X, T). As shown in [76, Corollary 3.1], the "cluster" $C_{(X,T)}$ is finite for $(X, T) \in \mathcal{P}$ with probability 1. It represents the set of all items that can potentially affect the acceptance status of the incoming particle represented by the Poisson point (X, T). The method of recontructing the set of accepted items from the graph $\mathcal{G}$ goes as follows. Let $A \subset \mathbb{R}^d$ be a (possibly unbounded) Borel set, and let $\mathcal{P}_A$ denote the set $\mathcal{P} \cap (A \times \mathbb{R}^+)$, i. e. the set of Poisson points that lie in $A \times \mathbb{R}^+$. Let $\mathcal{G}_{|A}$ denote the restriction of $\mathcal{G}$ to the vertex set $\mathcal{P}_A$. Recursively define subsets $F_i(A)$, $G_i(A)$, $H_i(A)$ of A, $i = 1, 2, 3, \dots$ as follows. Let $F_1(A)$ be the set of roots of the oriented graph $\mathcal{G}_{|A}$, and let $G_1(A)$ be the set of offspring of roots. Set $H_1(A) = F_1(A) \cup G_1(A)$. For the next step, remove the set $H_1(A)$ from the vertex set, and define $F_2(A)$ and $G_2(A)$ the same way; so $F_2(A)$ is the set of roots of the restriction of $\mathcal{G}$ to vertices in $\mathcal{P}_A \setminus H_1(A)$, and $G_2(A)$ is the set of vertices in $\mathcal{P}_A \setminus H_1(A)$ which are offsprings of

$F_2(A)$. Set $H_2(A) = F_2(A) \cup G_2(A)$, remove the set $H_2(A)$ from $\mathcal{P}_A \setminus H_1(A)$, and repeat the process to obtain $F_3(A)$, $G_3(A)$, $H_3(A)$. Continuing ad infinitum gives us subsets $F_i(A)$, $G_i(A)$ of $\mathcal{P}_A$ defined for $i = 1, 2, 3, \ldots$. These sets are disjoint by construction. In the case when $A = \mathbb{R}^d$, we drop the reference to A and use the abbreviation F_i and G_i for $F_i(\mathbb{R}^d)$ and $G_i(\mathbb{R}^d)$.

As proved in [76, Lemma 3.2], for every bounded nonnull Borel set A in $\mathbb{R}^d$, the sets $F_1(A)$, $G_1(A)$, $F_2(A)$, $G_2(A), \ldots$ form a partition of $\mathcal{P}_A$, and F_1, G_1, F_2, $G_2, \ldots$ form a partition of $\mathcal{P}$ with probability 1. In addition, the random parking measure ξ^A in A is given by the projection of the union $\cup_{i=1}^{\infty} F_i(A)$ on $\mathbb{R}^d$. Likewise, the random parking measure ξ in $\mathbb{R}^d$ is given by the projection of the union $\cup_{i=1}^{\infty} F_i$ on $\mathbb{R}^d$. As shown in [76, Theorem 2.2], ξ^A weakly converges in the sense of measures to ξ as A tends to $\mathbb{R}^d$.

The random parking measure ξ satisfies the following properties:

Proposition 4.1. *The random measure ξ is stationary (under real shifts), ergodic, isotropic, and almost surely general.*

These properties are essentially inherited from the associated space-time Poisson process $\mathcal{P}$, as we quickly show below.

Proof. Step 1. Stationarity.
By definition, the Poisson point process $\mathcal{P}$ on $\mathbb{R}^d \times \mathbb{R}^+$ and its translation $(x_1, \ldots, x_d, 0) + \mathcal{P}$ have the same distribution for all $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. Hence the graphical construction is stationary, and therefore also the random parking measure ξ.

Step 2. Isotropy.
The proof of the isotropy of ξ is similar to the proof of the stationarity. For all $\mathbf{R} \in SO_d$, the Poisson point process $\mathcal{P}$ and its rotated version $\mathbf{R}\mathcal{P} := \{(\mathbf{R}x, t) : x \in \mathbb{R}^d, t \in \mathbb{R}^+, (x, t) \in \mathcal{P}\}$ have the same distribution, which implies that the random parking measure ξ is isotropic.

Step 3. General position.
For all $t > 0$ we set $\overline{\mathcal{P}}_t = \{x \in \mathbb{R}^d : \exists \tau \in [0, t], (x, \tau) \in \mathcal{P}\}$. Since $\xi \subset \cup_{n \in \mathbb{N}} \overline{\mathcal{P}}_n$, and since $\overline{\mathcal{P}}_n \subset \overline{\mathcal{P}}_{n+1}$ for all $n \in \mathbb{N}$, the event that ξ is not in general position is contained in the union over $n \in \mathbb{N}$ of the events that $\overline{\mathcal{P}}_n$ is not in general position. Since $\overline{\mathcal{P}}_n$ is a Poisson process of intensity n in $\mathbb{R}^d$, the probability that $\overline{\mathcal{P}}_n$ is not in general position is zero, so that the union of this countable set of events has also probability zero, and ξ is almost surely general.

Step 4. Ergodicity.
Let us view ξ and $\mathcal{P}$ as elements of $\mathcal{A}_{lf}^d$ and $\mathcal{A}_{lf}^{d+1}$ respectively. Also let us extend $\mathcal{P}$ to a homogeneous Poisson process of unit intensity on the whole of $\mathbb{R}^{d+1}$ (also denoted $\mathcal{P}$). Let T_x denote translation by an element x of $\mathbb{R}^d$ (acting either on $\mathcal{A}_{lf}^d$ or $\mathcal{A}_{lf}^{d+1}$ according to context). Then as described earlier in this section, ξ is the image of $\mathcal{P}$ under a certain mapping h from $\mathcal{A}_{lf}^{d+1}$ to $\mathcal{A}_{lf}^d$, which commutes with T_x for any $x \in \mathbb{R}^d$, that is $T_x \circ h = h \circ T_x$.

Suppose A is a measurable subset of $\mathcal{A}_{lf}^d$ which is shift-invariant, meaning $T_x(A) = A$ for all $x \in \mathbb{R}^d$. Then for $x \in \mathbb{R}^d$,

$$h^{-1}(A) = h^{-1}(T_x(A)) = T_x(h^{-1}A)$$

so $h^{-1}(A)$ is invariant under the mapping T_x acting on $\mathcal{A}_{lf}^{d+1}$. Now, $\mathcal{P}$ is ergodic under translations, that is if $B \subset \mathcal{A}_{lf}^{d+1}$ satisfies $T_x(B) = B$ for some non-zero $x \in \mathbb{R}^{d+1}$, then $P[B] \in \{0, 1\}$. See for example the proof of Proposition 2.6 of [60]. Therefore with A as above,

$$P[\xi \in A] = P[\mathcal{P} \in h^{-1}(A)] \in \{0, 1\},$$

for any $x \in \mathbb{R}^d$. Thus ξ is ergodic.

This answers the question left open in the last section.

4.3.2 Approximation result

Another question of practical interest concerns the approximation of the homogenized integrand W_{hom} associated with the random parking measure ξ. Since the convergence in (4.10) does hold almost surely, a possible approximation of W_{hom} is given for all $\Lambda \in \mathcal{M}^d$ by

$$W_{\mathrm{hom}}^R(\Lambda) = \frac{1}{|Q_R|} \inf_u \left\{ F_1^{Q_R}(\xi, u) \;\middle|\; u \in \mathfrak{S}_1^{Q_R}(\xi) \text{ such that } u(x) = \Lambda \cdot x \right. \\ \left. \text{if } x \in \xi \cap Q_R \text{ and } d_\infty(x, \partial Q_R) \le 2\rho_2 \right\} \quad (4.11)$$

for some R large enough. By Theorem 23, $\lim_{R\to\infty} W_{\mathrm{hom}}^R(\Lambda) = W_{\mathrm{hom}}(\Lambda)$ almost surely. The question we address in this paragraph is whether one may "replace" ξ by ξ^R (the random parking measure in Q_R) in (4.11), and still have the almost sure convergence to the homogenized integrand.

We solve this problem in two steps. First we show that for local functionals of bounded domains and point sets, provided they satisfy some averaging property and they are "insensitive" to boundary effects, ξ can be replaced by ξ^R. Second, we show that although the discrete model for rubber does not yield a local functional (it depends on the Delaunay triangulation which is itself slightly nonlocal), it can be approximated by a local functional.

Let $0 < \rho_1 < \rho_2$. Let $\mathcal{O}(\mathbb{R}^d)$ be the set of bounded Lipschitz domains of $\mathbb{R}^d$. Given $D \in \mathcal{O}(\mathbb{R}^d)$ and $R > 0$, let $D_R := \{Rx, x \in D\}$ be the dilation of D by a factor $R > 0$. For $r > 0$ set $D_{R,r} := \{x \in D_R : d_\infty(x, \partial D_R) \ge r\}$. Let $\mathcal{D}(D) := \{D_{R,r} : R > 0, r \ge 0, D_{R,r} \ne \emptyset\}$. We now define functions parametrized by point sets (or restrictions of point sets on bounded domains).

Definition 23 *Let $D \in \mathcal{O}(\mathbb{R}^d)$. A measurable function $S : \mathcal{O}(\mathbb{R}^d) \times \mathcal{A}_{\rho_1,\rho_2} \to \mathbb{R}$ is said:*

- *to be* local on $\mathcal{D}(D)$ *if for all* $\hat{D} \in \mathcal{D}(D)$*, and all* $\zeta, \tilde{\zeta} \in \mathcal{A}_{\rho_1,\rho_2}$ *such that* $\zeta \cap \hat{D} = \tilde{\zeta} \cap \hat{D}$*, we have*
$$S(\hat{D}, \zeta) \;=\; S(\hat{D}, \tilde{\zeta});$$
- *to be* insensitive to boundary effects on $\mathcal{D}(D)$ *if there exists* $0 < \alpha < 1$ *such that we have*
$$\limsup_{R\to+\infty} \sup_{\zeta\in\mathcal{A}_{\rho_1,\rho_2}} \left\{ \frac{|S(D_R,\zeta) - S(D_{R,R^\alpha},\zeta)|}{R^d} \right\} \;=\; 0. \tag{4.12}$$
- *to have the* averaging property *on* $\mathcal{A}_{\rho_1,\rho_2}$ *with respect to* D *if for any stationary point set* ζ *whose realization almost surely belongs to* $\mathcal{A}_{\rho_1,\rho_2}$*, there exists* $\overline{S} \in \mathbb{R}$ *such that almost surely*
$$\lim_{R\to+\infty} \frac{S(D_R,\zeta)}{|D_R|} \;=\; \overline{S}. \tag{4.13}$$

For such functions S, we then have:

Theorem 4.2. *Let* $0 < \rho_1 < \rho_0 < \rho_2 < \infty$*. Let* $D \in \mathcal{O}(\mathbb{R}^d)$*, and for all* $R > 0$*, let* $\xi^R = \xi^{D_R}$ *denote the random parking measure on* $D_R := \{Rx, x \in D\}$*, and let* ξ *be the random parking measure on* $\mathbb{R}^d$ *with parameter* ρ_0*. If the measurable function* $S : \mathcal{O}(\mathbb{R}^d) \times \mathcal{A}_{\rho_1,\rho_2} \to \mathbb{R}^+$ *is local on* $\mathcal{D}(D)$*, insensitive to boundary effects on* $\mathcal{D}(D)$*, and has the averaging property on* $\mathcal{A}_{\rho_1,\rho_2}$ *with respect to* D*, then with* $\overline{S}$ *given by* (4.13)*, almost surely*
$$\lim_{R\to+\infty} \frac{S(D_R,\xi^R)}{|D_R|} \;=\; \lim_{R\to+\infty} \frac{S(D_R,\xi)}{|D_R|} \;=\; \overline{S}. \tag{4.14}$$

This theorem directly follows from the fact that for all $0 < \alpha < 1$, there exists an almost-surely finite random variable R_0 such that for all $R \geq R_0$,
$$S(D_{R,R^\alpha}, \xi^{D_R}) \;=\; S(D_{R,R^\alpha}, \xi).$$

This is a consequence of the stabilization properties of the random parking measure proved by Penrose, Schreiber and Yukich [81], combined with the Borel-Cantelli lemma.

It remains to apply Theorem 4.2 to the homogenized integrand. For all $R > 0$ large enough, we define $\zeta^R = (\xi^R \cap Q_{R-d\rho_1}) \cup \mathcal{S}^R$, where $\mathcal{S}^R$ is a deterministic point set on ∂Q_R such that the convex hull of $\mathcal{S}^R$ is Q_R and such that ζ^R is $(\rho_1, 2d\rho_2)$-admissible. Hence, the Delaunay triangulation of ζ^R is a triangulation of Q_R, and we may define
$$\widetilde{W}^R_{\mathrm{hom}}(\Lambda) \;=\; \frac{1}{|Q_R|} \inf_u \left\{ F_1^{Q_R}(\zeta^R, u) \;\middle|\; u \in \mathfrak{S}_1^{Q_R}(\zeta^R) \text{ such that } u(x) = \Lambda\cdot x \text{ if } x \in \partial Q_R \right\}, \tag{4.15}$$
which is a computable quantity. Using a variant of Theorem 4.2 and results of [ACG] one may indeed prove that almost surely,
$$\lim_{R\to\infty} \widetilde{W}^R_{\mathrm{hom}}(\Lambda) \;=\; W_{\mathrm{hom}}(\Lambda).$$

4.4 Numerical approximation of the homogenized energy density [GLTV]

4.4.1 The discrete model for rubber revisited

Let us come back to the discrete model for rubber introduced in Section 4.1. The "coarse-grained" Hamiltonian of the polymer-chain network at deformation u is given by (4.1), and we have formally argued that the free energy of a macroscopic sample whose boundary is deformed by the linear map $x \mapsto \Lambda x$ for some $\Lambda \in \mathcal{M}^3_+$ is given by (4.2).

In order to make use of this model in practice, one needs to make precise the structure of the network and derive some formula for $H_\Lambda(u, \beta)$. The strongest assumption we shall make is that the polymer-chain network is a (Delaunay) triangulation of D into tetrahedra whose edges are the polymer chains themselves. This allows us to see u as a continuous and piecewise affine function. For the volumetric energy, we may then consider the standard Helmholtz energy density $W_{\mathrm{Helm}} : \mathcal{M}^3_+ \to [0, +\infty)$, given by:

$$W_{\mathrm{Helm}}(\Lambda) = K\Big(\det(\Lambda)^2 - 1 - 2\log(\det(\Lambda))\Big), \tag{4.16}$$

for some $K \geq 0$, so that

$$H_{\mathrm{vol}}(u) \;=\; \int_D W_{\mathrm{Helm}}(\nabla u(x))dx.$$

The second part of the Hamiltonian corresponds to the sum of the free energies of the polymer chains at temperature β if they were isolated and their end-to-end vectors were given by u. This free energy does have an analytic form derived by Kuhn and Grün in [50] under a non-Gaussian assumption: each segment of the chain obeys a non-Gaussian random walk. We refer to [34] for details. In particular, given a polymer chain made of N rigid segments of length l at absolute temperature $\beta = \frac{1}{k_B T}$, with a chain density n, the free energy (of entropic origin) for a chain of length r_c can be modeled by

$$W_c(r_c, N) = \frac{n}{\beta} N \left(\frac{r_c}{Nl}\theta\left(\frac{r_c}{Nl}\right) + \log \frac{\theta\left(\frac{r_c}{Nl}\right)}{\sinh\theta\left(\frac{r_c}{Nl}\right)} \right) - \frac{c}{\beta}, \tag{4.17}$$

where c is a constant and θ the inverse of the Langevin function $t \mapsto \coth t - \frac{1}{t}$.

Note that this free energy is infinite as soon as $r_c > Nl$, the total length of the chain. For discrete to continuum derivations, θ is usually replaced by the first terms of its series expansion:

$$\theta(r) = 3r + \frac{9}{5}r^3 + \frac{297}{175}r^5 + \frac{1539}{875}r^7 + \frac{672}{359}r^9 + O(r^{11}), \tag{4.18}$$

although this simplification is not essential for our discussion (e.g. Padé approximations behave better close to the finite extensibility limit). A series expansion of W_c then reads:

$$\begin{aligned} W_c(r_c, N) = \frac{n}{\beta} N \Bigg[& \frac{3}{2}\left(\frac{r_c}{Nl}\right)^2 + \frac{9}{20}\left(\frac{r_c}{Nl}\right)^4 + \frac{9}{350}\left(\frac{r_c}{Nl}\right)^6 + \frac{81}{7000}\left(\frac{r_c}{Nl}\right)^8 \\ & + \frac{243}{673750}\left(\frac{r_c}{Nl}\right)^{10} \Bigg] + O\left(\left(\frac{r_c}{Nl}\right)^{12}\right). \end{aligned} \tag{4.19}$$

The behavior of the polynomial approximation at infinity satisfies the classical coercivity assumption on hyperelastic materials at infinity. Replacing the inverse of the Langevin function by the first terms of a series expansion is a rather good modeling at high temperature (see [53]). A remarkable property of such an energy is $W_c(0) = 0$ and $W_c(1) > 0$. In particular the prefered configuration of a polymer chain satisfies $r_c = 0$.

When N is fixed, we simply write $W_c(r_c)$ instead of $W_c(r_c, N)$.

We may now put (4.1) into the form of (4.7). Let $\varepsilon_0 > 0$ be the intrinsic lengthscale of the polymer network, that is, the length of a monomer, also denoted by l. We denote by Σ_{ε_0} and $\mathcal{T}_{\varepsilon_0}$ the point set and the Delaunay triangulation associated with the polymer-chain network. An edge e of the polymer chain network is then supposed to be made of

$$N_e \simeq \left(\frac{|e|}{l}\right)^2 = \left(\frac{|e|}{\varepsilon_0}\right)^2$$

segments (or monomers) — $\sqrt{N_e} l$ is indeed the average distance of the random walker from the origin after N_e jumps. Hence, if the edge e has length L after deformation, its free energy is given by

$$\begin{aligned} W_c(L, N_e) &= \frac{n}{\beta} N_e \left(\frac{L}{N_e l} \theta\left(\frac{L}{N_e l}\right) + \log \frac{\theta\left(\frac{L}{N_e l}\right)}{\sinh \theta\left(\frac{L}{N_e l}\right)} \right) \\ &= \frac{n}{\beta}\left(\frac{|e|}{\varepsilon_0}\right)^2 \left(\frac{L\varepsilon_0}{|e|^2} \theta\left(\frac{L\varepsilon_0}{|e|^2}\right) + \log \frac{\theta\left(\frac{L\varepsilon_0}{|e|^2}\right)}{\sinh \theta\left(\frac{L\varepsilon_0}{|e|^2}\right)} \right). \end{aligned}$$

This formula allows us to properly define the energy function $f_{nn} : \mathbb{R}^d \times \mathbb{R}^d \to [0, +\infty]$ used in (4.7), and we set

$$\begin{aligned} f_{nn}(e, \lambda) &= \frac{n}{\beta\varepsilon_0^3}|e|^2 \left(\frac{|\lambda||e|}{|e|^2} \theta\left(\frac{|\lambda||e|}{|e|^2}\right) + \log \frac{\theta\left(\frac{|\lambda||e|}{|e|^2}\right)}{\sinh \theta\left(\frac{|\lambda||e|}{|e|^2}\right)} \right) \\ &= \frac{n}{\beta\varepsilon_0^3}|e|^2 \left(\frac{|\lambda|}{|e|} \theta\left(\frac{|\lambda|}{|e|}\right) + \log \frac{\theta\left(\frac{|\lambda|}{|e|}\right)}{\sinh \theta\left(\frac{|\lambda|}{|e|}\right)} \right) \end{aligned} \tag{4.20}$$

whereas $W_{\mathrm{vol}} : \mathcal{M}^3 \to [0, +\infty]$ is given by

$$W_{\mathrm{vol}}(\Lambda) = \begin{cases} W_{\mathrm{Helm}}(\Lambda) & \text{if } \Lambda \in \mathcal{M}^3_+, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.21}$$

With these definitions, we then have

$$H_\Lambda(u, \beta) = F^D_{\varepsilon_0}(\Sigma, u),$$

for $\Sigma = \varepsilon_0^{-1}\Sigma_{\varepsilon_0}$ (which we assume to be in $\mathcal{A}_{\rho_1,\rho_2}$).

Provided f_{nn} and W_{vol} are of class $\mathcal{U}_p$ and $\mathcal{V}_p$ for some $1 < p < +\infty$, Theorem 23 and (a variant of) Lemma 3.1 imply that

$$|D|W_{\mathrm{hom}}(\Lambda) = \lim_{\varepsilon_k\to 0}\inf_u \left\{ F^D_{\varepsilon_k}(\Sigma,u) \;\middle|\; u \in \mathfrak{S}^D_{\varepsilon_k}(\Sigma) \text{ such that } u(x) = \Lambda\cdot x \right.$$
$$\left. \text{if } x \in \varepsilon_k\Sigma\cap D \text{ and } d_\infty(x,\partial D) \le 2\varepsilon_k\rho_2 \right\}$$

or equivalently

$$\lim_{R\to\infty} \frac{F(\Lambda, D_R)}{|D_R|} = W_{\mathrm{hom}}(\Lambda),$$

which is the rigorous version of (4.3).

Note however that W_{vol} and f_{nn} are not of class $\mathcal{U}_p$ and $\mathcal{V}_p$ for any $p > 1$ because they blow up on a bounded set. The blow up of W_{vol} is related to the issue mentioned in the first chapter. The second blow up models finite extensibility of the chains. In any case, one may at first use a cut-off procedure in order to apply Theorem 23. Note that the estimate from above is satisfied by f_{nn} provided we consider any order of the Taylor expansion of the inverse of the Langevin function (for instance (4.19), in which case $p = 10$).

4.4.2 Numerical method

The starting point for the numerical approximation of W_{hom} is (4.15), and we proceed in two steps.

- We first generate the deterministic set of points on ∂Q_R and a realization of the random parking measure of parameter 1/2 in Q_{R-1}. We then construct the associated Delaunay triangulation of Q_R.
- In a second step, we solve the minimization problem associated with (4.15) for R finite and the Delaunay triangulation of Q_R (well-defined as the minimization of a smooth coercive function on a finite-dimensional space) by a Newton algorithm, as it is classical in nonlinear elasticity (see for instance [52], and [87]) provided one adds the energy of the edges (which are “non-standard” one-dimensional elements). Continuation methods are also used to ensure the convergence of the Newton algorithm.

In practice, we run several independent realizations of the point process and make an empirical average. This enhances the convergence with respect to the randomness. The analysis of numerical methods to approximate homogenized coefficients in stochastic homogenization of discrete linear elliptic equations has been addressed in Chapter 1. Although the analysis does not cover this nonlinear vector case, it may serve as a guide to understand the convergence properties involved.

At the end of the minimization algorithm, the homogenized energy W_{hom} is approximated by the spatial average over Q_R of the energy density of the minimizer which has been numerically obtained, and the Piola-Kirchhoff stress tensor is given by the spatial average on Q_R of the associated local Piola-Kirchhoff stress tensor (provided the minimizer is isolated, and the local Hessian is strongly elliptic, see [39, Section 4.2] for related arguments in the continuous case).

We have conducted two series of tests with two different strain gradients Λ_1 and Λ_2, the first in small deformation ($\sim 25\%$), and the second in large deformation ($\sim 300\%$):

$$\Lambda_1 = \begin{pmatrix} 1.1 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & 0 & 25/33 \end{pmatrix}, \qquad \Lambda_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1/6 \end{pmatrix}.$$

Both deformations are isochoric: $\det \Lambda_1 = \det \Lambda_2 = 1$.

We focus on the first Piola-Kirchhoff stress tensor $\frac{\partial W_{\mathrm{hom}}(\Lambda)}{\partial \Lambda}$. We expect the stress tensors $\Pi_1(N_R, K_R) = [\Pi_1(N_R, K_R)]_{ij}$ and $\Pi_2(N_R, K_R) = [\Pi_2(N_R, K_R)]_{ij}$, associated with Λ_1 and Λ_2 respectiveley, to be diagonal. We therefore focus on the principal stresses. Each effective stress tensor is obtained as the empirical average over K_R realizations of a point process in Q_R (with approximately N_R edges). These numbers are gathered in Table 4.1.

N_R	130	1 600	6 000	16 000	33 000	59 500
K_R	2160	270	80	34	18	10

Table 4.1. Number of edges and associated number of realizations.

The (square root of the) variance of each diagonal term is plotted in function of N_R on Figure 4.1 for Λ_1 and Figure 4.2 for Λ_2, in log-log scale. The straight lines are linear fittings. Their slopes are approximately $-1/2$ (between $-.45$ and $-.5$), as expected (see Chapter 1 for related results).

To conclude, we show on Figures 4.3–4.5 and 4.6–4.8 the convergences of the diagonal terms of the Piola stress tensor (together with their standard deviations) in function of N_R, for Λ_1 and Λ_2, respectively. As can be seen, the global error may largely dominate the standard deviation (see in particular Figure 4.8) so that we expect the systematic error to be the dominant error (recall that in the linear case, the numerical method used here is similar to (1.24), which is expected to yield a systematic error of order $N_R^{-1/3}$ and a standard deviation of order $(K_R N_R)^{-1/2}$ for $d = 3$). Throughout this chapter, we shall *consider* that the approximation has converged for $N_R \sim 100\,000$ and $K_R \sim 10$.

4.4.3 Comparison to mechanical experiments

In this subsection we compare the homogenized limit of the discrete model for rubber to the popular mechanical experiments by Treloar. The plots 4.9, 4.10, and 4.11 display the engineering stresses (or nominal stresses) associated with the mechanical experiments in

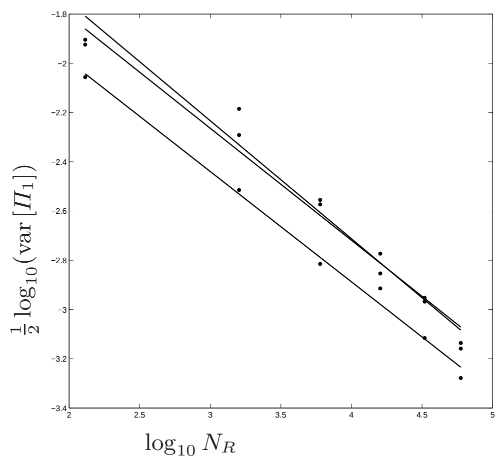

Fig. 4.1. Variance of the diagonal terms of the Piola stress tensor for Λ_1

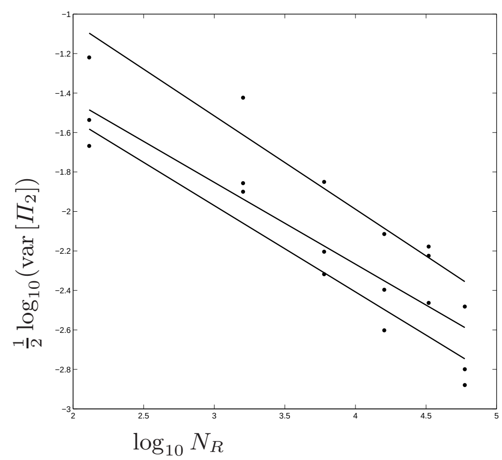

Fig. 4.2. Variance of the diagonal terms of the Piola stress tensor for Λ_2

uniaxial traction, biaxial traction, and in planar tension, respectively. We quickly recall the definitions of the engineering stress for these sollicitations. Uniaxial traction corresponds to a deformation gradient Λ of the form

$$\Lambda = \mathrm{diag}(\lambda, \lambda^{-1/2}, \lambda^{-1/2})$$

for some $\lambda \geq 1$. The quantity $\lambda - 1$ is called the engineering strain. The associated Cauchy stress tensor σ is diagonal, and the associated engineering stress is given by

$$\lambda - 1 \mapsto S_{\mathrm{uniaxial}} = \frac{\sigma_{11} - \sigma_{22}}{\lambda}.$$

For biaxial traction the deformation gradient Λ is of the form

$$\Lambda = \mathrm{diag}(\lambda^{-2}, \lambda, \lambda)$$

for some $\lambda \geq 1$. The quantity $\lambda^{-2} - 1$ is called the engineering strain. The associated Cauchy stress tensor σ is diagonal, and the associated engineering stress is given by

$$\lambda^{-2} - 1 \mapsto S_{\mathrm{biaxial}} = \frac{\sigma_{11} - \sigma_{22}}{\lambda^{-2}}.$$

Finally, the planar tension experiment corresponds to a deformation gradient Λ of the form

$$\Lambda = \mathrm{diag}(1, \lambda, \lambda^{-1})$$

for some $\lambda \geq 1$. The quantity $\lambda - 1$ is called the engineering strain. The associated Cauchy stress tensor σ is diagonal, and the associated engineering stress is given by

$$\lambda - 1 \mapsto S_{\mathrm{planar}} = \frac{\sigma_{22} - \sigma_{11}}{\lambda}.$$

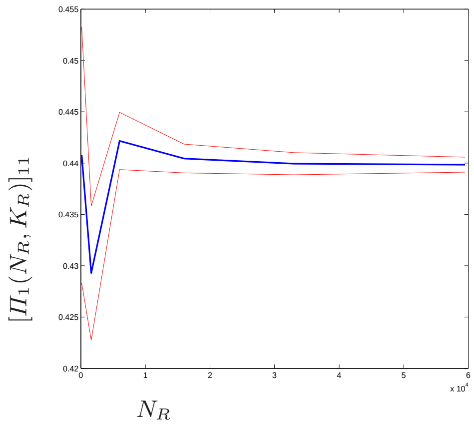

Fig. 4.3. Convergence of the diagonal part 1 of the Piola stress tensor for Λ_1

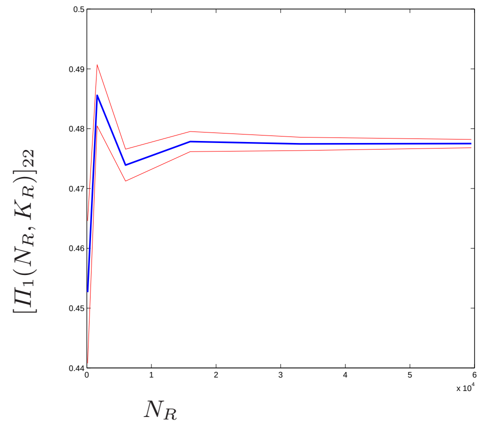

Fig. 4.4. Convergence of the diagonal part 2 of the Piola stress tensor for Λ_1

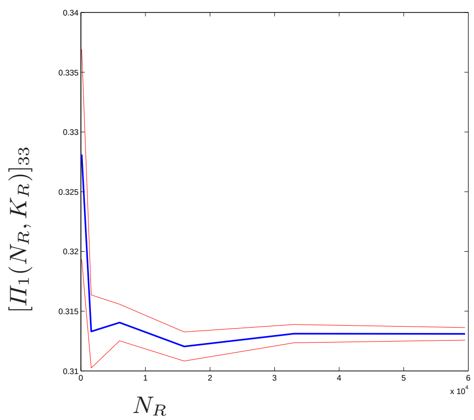

Fig. 4.5. Convergence of the diagonal part 3 of the Piola stress tensor for Λ_1

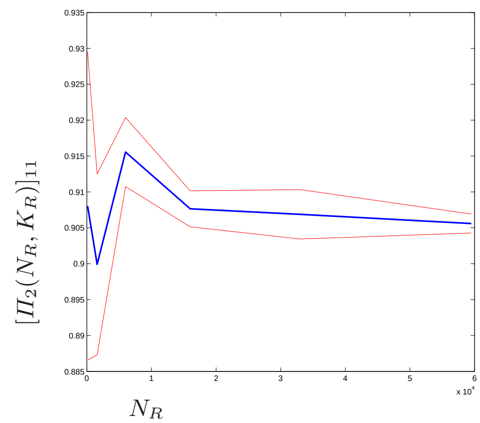

Fig. 4.6. Convergence of the diagonal part 1 of the Piola stress tensor for Λ_2

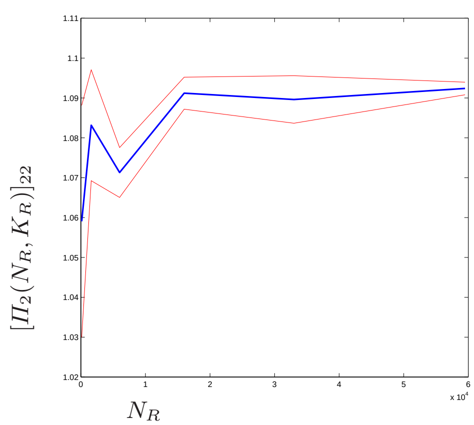

Fig. 4.7. Convergence of the diagonal part 2 of the Piola stress tensor for Λ_2

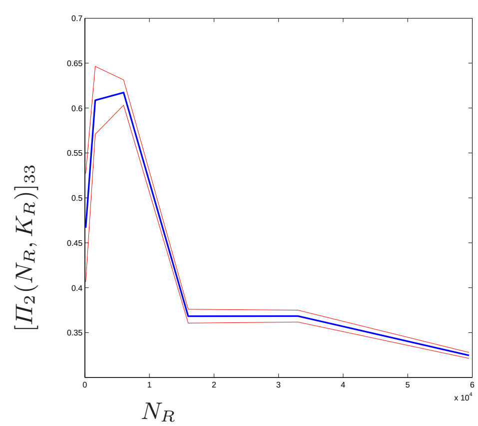

Fig. 4.8. Convergence of the diagonal part 3 of the Piola stress tensor for Λ_2

In addition to the homogenized limit of the discrete model, we also display the results for two other models: Treloar's model and the popular Arruda-Boyce model (also based on polymer-chain statistics) with its standard coefficients to fit Treloar's data, namely $N = 26.5$ and $\frac{n}{\beta} = 0.27$. We have not tried to optimize the parameters for the homogenized limit, and have taken $N = 26.5$ and $\frac{n}{\beta} = 0.27$ as well, and set $K = 5$. As can be seen, the results are promising. The mismatch in large deformations is due to the (bad) approximation of the Langevin function by its Taylor expansion.

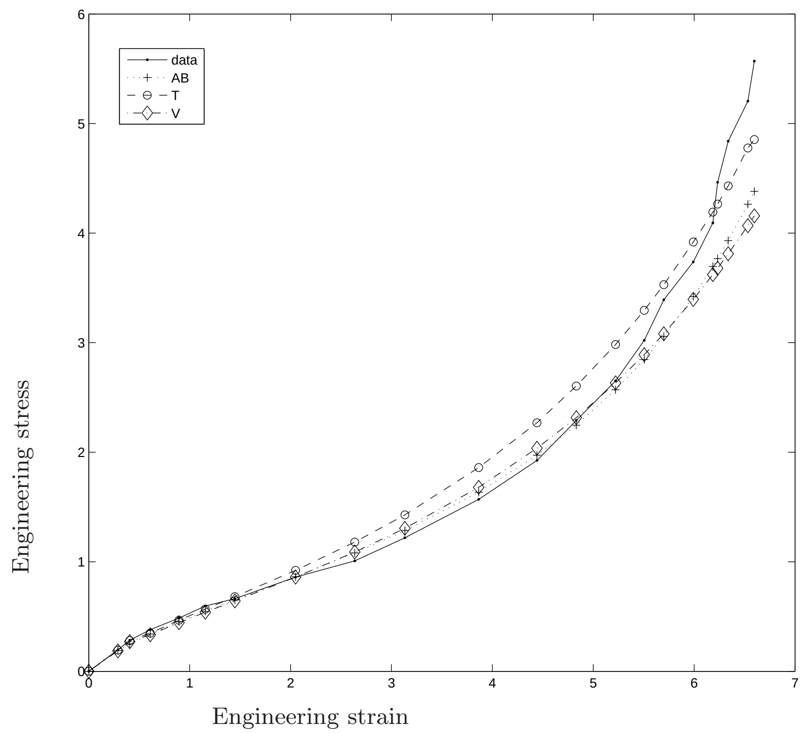

Fig. 4.9. Uniaxial traction — Treloar's experiments (AB: Arruda-Boyce, T: Treloar, V: variational — or homogenized)

4.5 Strong ellipticity and discrete homogenization [Glo11b]

In this section we address the question of the strong ellipticity of homogenized energy densities obtained by discrete homogenization. Using discrete Bloch waves it is possible to adapt the theory by Geymonat, Müller and Triantafyllidis [38] to the periodic homogenization of discrete systems. Instead of presenting the general theory, we prefer to give two meaningful examples which yield both a material which fails being strictly strongly elliptic in compression, and a material which is strictly strongly elliptic at any deformation gradient.

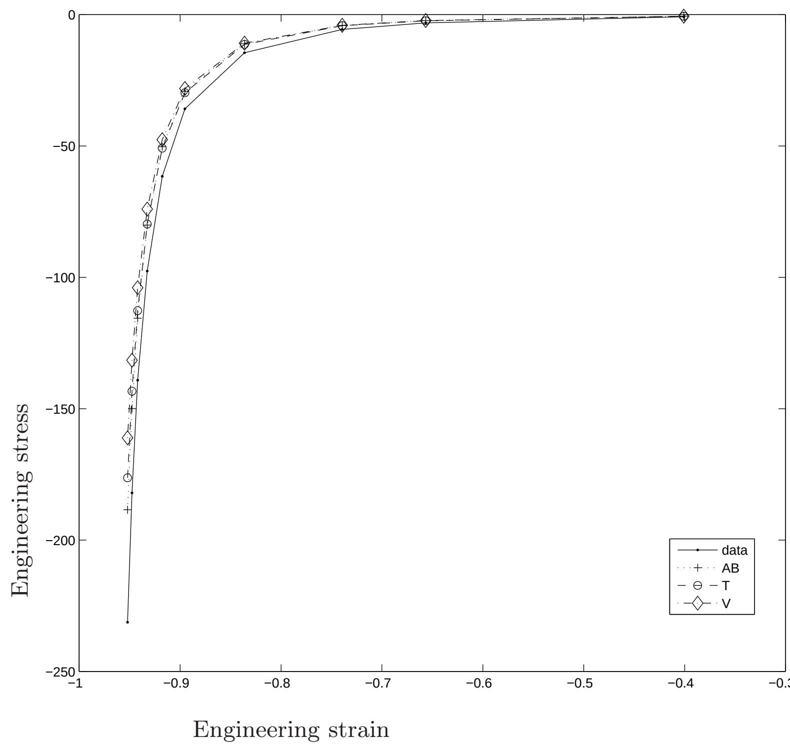

Fig. 4.10. Biaxial tension — Treloar's experiments (AB: Arruda-Boyce, T: Treloar, V: variational — or homogenized)

The second example exploits the specific structure of the energy functions of Section 4.1 derived in statistical mechanics. We show in the last subsection that this specific structure also implies strict strong ellipticity in the case of discrete stochastic homogenization — at least in a perturbation regime. This result sheds some light on the possible microscopic origin of macroscopic strong ellipticity of rubber-like materials.

4.5.1 Loss of strong ellipticity in the periodic discrete case

Before we present this example, let us quickly recall some results on periodic homogenization of discrete systems. Let $\mathcal{T}$ be a $Q = (0,1)^d$-periodic triangulation of $\mathbb{R}^d$ with vertices in $\mathbb{Z}^d$. For every integer $N \geq 1$ we denote by $\mathcal{S}(Q_N, \mathbb{R}^d)$ the set of continuous functions from Q_N to $\mathbb{R}^d$ which are piecewise affine on $\mathcal{T}$, and by $\mathcal{S}_\#(Q_N, \mathbb{R}^d)$ the corresponding subset of Q_N-periodic functions. Any functional $F(\cdot, Q)$ from $\mathcal{S}(Q_N, \mathbb{R}^d)$ to $\mathbb{R}$ can be extended by periodicity to a functional $F(\cdot, Q_N)$ on $\mathcal{S}(Q_N, \mathbb{R}^d)$ by setting for all $v \in \mathcal{S}(Q_N, \mathbb{R}^d)$

$$F(v, Q_N) := \sum_{k \in [0,N)^d \cap \mathbb{Z}^d} F(\tau_z v, Q),$$

where $\tau_z v(x) = v(x+z)$ for all $x \in Q$ and $z \in \mathbb{Z}^d$.

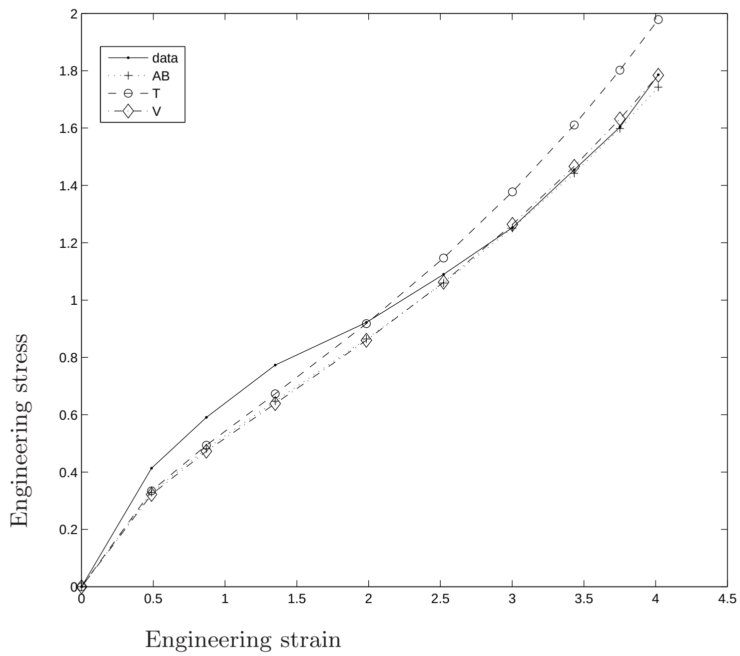

Fig. 4.11. Planar tension— Treloar's experiments (AB: Arruda-Boyce, T: Treloar, V: variational — or homogenized)

The model is written on $\mathbb{Z}^2$ and the geometry corresponds to Figure 4.12, with

$$\begin{aligned}
&z_1 = (0,0), z_2 = (1,0), z_3 = (1,1), z_4 = (0,1),\\
&\mathcal{NN} = \{(z_1,z_2),(z_2,z_3),(z_3,z_4),(z_4,z_1)\},\\
&\mathcal{NNN} = \{(z_2,z_4),(z_1,z_3)\},
\end{aligned}$$

and $\mathcal{T} \cap Q = \{T_1, T_2\}$ with

$$\begin{aligned}
T_1 &= \{x = (x_1,x_2) \in Q : x_2 > x_1\},\\
T_2 &= \{x = (x_1,x_2) \in Q : x_1 > x_2\}.
\end{aligned}$$

It is characterized by the following energy on $\mathcal{S}(Q,\mathbb{R}^2)$:

$$F^\eta(v,Q) = \sum_{(i,j)\in\mathcal{NN}} (|v_i - v_j| - 1)^2 + \eta \sum_{(i,j)\in\mathcal{NNN}} (|v_i - v_j| - \sqrt{2})^2,$$

where $v_i := v(z_i)$, and $\eta > 0$ is a physical parameter, which may be seen as the strength of the soft matrix (and should be small with respect to 1).

This discrete model loses strong ellipticity in compression, and we therefore consider strain gradients of the form

$$\Lambda_\delta := \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix}, \qquad \delta \in (0,1).$$

Recall that for all $\Lambda \in \mathcal{M}^2$, we set $\varphi_\Lambda : x \mapsto \Lambda x$. We make the following assumption:

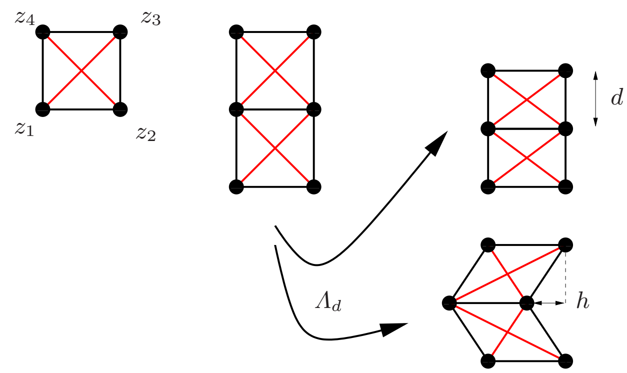

Fig. 4.12. Geometry, 1-periodic and 2-periodic deformations.

Hypothesis 4 *For all $\delta \in (0,1)$ and every integer $N \geq 1$, there exists a minimizer ϕ of $F^\eta(\varphi_{\Lambda_\delta} + \cdot, Q_N)$ on $\mathcal{S}_\#(Q_N, \mathbb{R}^2)$ which satisfies $\phi(z+e_1) = \phi(z)$ for all $z \in \mathbb{Z}^2$.*

This assumption is strong: it implies that the minimizer $\phi : Q_N \to \mathbb{R}^2, (x_1, x_2) \mapsto \phi(x_1, x_2)$ only depends on x_2 (and not on x_1), which is reasonable physically. The same type of assumption is made in [38, Section 6], where the lateral stress is assumed to vanish (see in particular [38, p. 271]).

We now define the cell integrand W_1^η by

$$\begin{aligned} W_1^\eta : \mathcal{M}^2 &\to \mathbb{R}^+ \\ \Lambda &\mapsto \inf\{F^\eta(\varphi_\Lambda + \phi, Q), \phi \in \mathcal{S}_\#(Q, \mathbb{R}^2)\}, \end{aligned}$$

and we define the homogenized integrand W_{hom}^η by

$$\begin{aligned} W_{\mathrm{hom}}^\eta : \mathcal{M}^2 &\to \mathbb{R}^+ \\ \Lambda &\mapsto \lim_{N\to\infty} \frac{1}{|Q_N|} \inf\{F^\eta(\varphi_\Lambda + \phi_N, Q_N), \phi \in \mathcal{S}_\#(Q_N, \mathbb{R}^2)\}, \end{aligned}$$

which is well-defined quasiconvex function (see [1]).

Note that in this case, since the only elements ϕ of $\mathcal{S}_\#(Q, \mathbb{R}^2)$ are constant functions $\phi \equiv c \in \mathbb{R}$, we have

$$W_{\mathrm{CB}}^\eta(\Lambda) := F^\eta(\varphi_\Lambda, Q) = \inf\{F^\eta(\varphi_\Lambda + \phi, Q), \phi \in \mathcal{S}_\#(Q, \mathbb{R}^2)\} = W_1^\eta(\Lambda).$$

The model is said to satisfy the strict Cauchy-Born rule at $\Lambda \in \mathcal{M}^2$ if

$$W_{\mathrm{hom}}^\eta(\Lambda) = W_{\mathrm{CB}}^\eta(\Lambda).$$

We have the following lemma.

Lemma 4.3. *Under Hypothesis 4, for all $\delta \in (0,1)$ we have*

$$W_{\mathrm{hom}}^\eta(\Lambda_\delta) = \mathcal{Q}W_1^\eta(\Lambda_\delta) = \mathcal{R}W_1^\eta(\Lambda_\delta), \tag{4.22}$$

where $\mathcal{Q}W_1^\eta$ and $\mathcal{R}W_1^\eta$ denote the quasiconvex and rank-one convex envelopes of W_1^η, respectively.

Lemma 4.3 is a discrete and stronger version of Lemma 3.3: in this specific case, the homogenized energy density is precisely given by the quasiconvex envelope of the cell energy density, whereas in the continuous case, we only know that $\mathcal{Q}W_1^\eta$ and W_{hom}^η are bounded from above by some constant times η. In a recent work [62] it is actually shown that, in the two-dimensional case above, the identity $\mathcal{Q}W_1^\eta(\Lambda) = W_{\text{hom}}^\eta(\Lambda)$ holds for all $\Lambda \in \mathcal{M}^2$, independently of Hypotheses 4. Yet this does not imply the conclusion of Lemma 4.3 for the rank-one convex envelope, so that it seems that Hypotheses 4 cannot be skipped.

From this lemma, we deduce the main result of this paragraph.

Proposition 1 *Under Hypothesis 4, there exists $\eta^* > 1$ such that for all $\eta < \eta^*$, there exists $\delta^* \in (0,1)$ such that*

$$\begin{aligned} &e_i \otimes e_j \cdot \mathrm{D}^2_{\Lambda^2} W_1^\eta(\Lambda_\delta) e_i \otimes e_j \,>\, 0 \quad \forall i,j \in \{1,2\} \text{ and all } \delta \in (\delta^*, 1), \\ &e_1 \otimes e_2 \cdot \mathrm{D}^2_{\Lambda^2} W_1^\eta(\Lambda_{\delta^*}) e_1 \otimes e_2 \,=\, 0, \end{aligned} \tag{4.23}$$

and

$$\begin{aligned} W_{\text{hom}}^\eta(\Lambda_\delta) &= W_1^\eta(\Lambda_\delta) \text{ for all } \delta \in [\delta^*, 1), \\ W_{\text{hom}}^\eta(\Lambda_\delta) &< W_1^\eta(\Lambda_\delta) \text{ for all } \delta \in (0, \delta^*). \end{aligned} \tag{4.24}$$

Hence, the homogenized material loses strong ellipticity at Λ_{δ^*}.

Remark 7 *Note that the condition on η in Proposition 1 essentially encodes the fact that the matrix (or the diagonal springs) is weaker than the fiber (or the vertical springs).*

4.5.2 The case of periodic polymer networks

The model is written on $\mathbb{Z}^2$ and the geometry corresponds to Figure 4.13, with

$$\begin{aligned} &z_1 = (0,0), z_2 = (1,0), z_3 = (1,1), z_4 = (0,1), z_5 = (1/2,1/2), \\ &\mathcal{NN}_1 = \{(z_1,z_2),(z_2,z_3),(z_3,z_4),(z_4,z_1)\}, \\ &\mathcal{NN}_2 = \{(z_5,z_1),(z_5,z_2),(z_5,z_3),(z_5,z_4)\}, \\ &\mathcal{NN} = \mathcal{NN}_1 \cup \mathcal{NN}_2, \end{aligned}$$

and $\mathcal{T} \cap Q = \{T_1, T_2, T_3, T_4\}$ with

$$\begin{aligned} T_1 &= \{x = (x_1,x_2) \in Q : x_2 > x_1, x_2 < 1 - x_1\}, \\ T_2 &= \{x = (x_1,x_2) \in Q : x_2 > x_1, x_2 > 1 - x_1\}, \\ T_3 &= \{x = (x_1,x_2) \in Q : x_1 > x_2, x_2 > 1 - x_1\}, \\ T_4 &= \{x = (x_1,x_2) \in Q : x_1 > x_2, x_2 < 1 - x_1\}. \end{aligned}$$

It is characterized by the following energy on $\mathcal{S}(Q, \mathbb{R}^2)$:

$$F(v,Q) = \sum_{(i,j)\in\mathcal{NN}} |v_i - v_j|^2(|v_i - v_j|^2 + 1) + \int_Q W_{\text{vol}}(\det \nabla v),$$

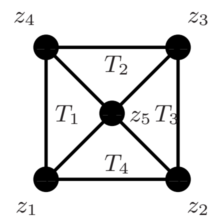

Fig. 4.13. Geometry

where $v_i := v(z_i)$, and $W_{\text{vol}} : \mathcal{M}^2 \to \mathbb{R}^+, \Lambda \mapsto (\det \Lambda - 1)^2$ is a volumetric energy. This model satisfies a standard growth condition of order $p = 4$, and can be homogenized. We denote by W_1 andf W_{hom} the associated cell integrand and homogenized integrand, respectively.

Proposition 2 *For all* $\Lambda \in \mathcal{M}^2$,

$$W_{\text{hom}}(\Lambda) = W_1(\Lambda) = W_{\text{CB}}(\Lambda), \tag{4.25}$$

and $e_i \otimes e_j \cdot \mathrm{D}^2_{\Lambda^2} W_{\text{hom}}(\Lambda) e_i \otimes e_j \geq 4$ *for all* $i, j \in \{1, 2\}$. *Hence, the homogenized material is strictly strongly elliptic for all* $\Lambda \in \mathcal{M}^2$.

Equality (4.25) is a consequence of the decomposition of the energy sketched on Figure 4.14. The first contribution is given by

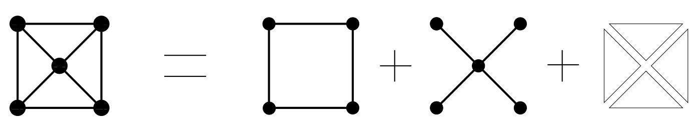

Fig. 4.14. Decomposition of the energy

$$F_1(v, Q) = \sum_{(i,j) \in \mathcal{NN}_1} |v_i - v_j|^2 (|v_i - v_j|^2 + 1),$$

the second by

$$F_2(v, Q) = \sum_{(i,j) \in \mathcal{NN}_2} |v_i - v_j|^2 (|v_i - v_j|^2 + 1),$$

and the third one by

$$F_3(v, Q) = \int_Q W_{\text{vol}}(\nabla v).$$

We claim that for each contribution, the Cauchy-Born rule holds. In particular, the first two terms F_1 and F_2 are the combination of convex potentials in independent directions, for which Jensen's inequality implies that the linear deformation has least energy. The same conclusion holds for the term F_3 by definition of quasiconvexity (recall that W_{vol} is polyconvex). The three contributions are therefore minimized independently at the same Cauchy-Born deformation, from which we deduce (4.25).

The strict strong ellipticity follows from an elementary calculation.

4.5.3 Perturbation result for stochastic polymer networks

In this subsection we turn to the discrete model of rubber elasticity introduced in Section 4.1. The energy functional splits into two parts: a convex energy (the contributions of the edges), and a nonconvex energy (the volumetric energy). The nonconvex part is homogeneous whereas the convex part is heterogeneous. Under the affine assumption this model yields a strictly strongly elliptic energy density at the "homogenization" limit as shown above in the periodic case. The aim of this section is to prove that this argument is stable under suitable perturbations. Our strategy relies on the specific form of the energy functional. Assuming that the minimizer in the asymptotic homogenization formula is (uniformly) close to the affine deformation, we shall prove that the convex part of the energy functional compensates for the variations of the nonconvex part at the level of the quadratic expansions, so that the sum remains strictly strongly elliptic.

To implement this strategy, we need to introduce several notions and quantities. Let $\mathcal{T}$ be a Delaunay triangulation of $\mathbb{R}^d$, and let $\mathcal{N}$ be the associated set of (undirected) edges. We say that $\mathcal{T}$ is n_S-admissible for some $n_S \in \mathbb{N}$ if any edge of $\mathcal{T}$ is shared by at most n_S simplices.

We start with some assumptions on the energy functions which are compatible with the discrete model for rubber:

Hypothesis 5 *The functions $f_{nn} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ and $W_{\mathrm{vol}} : \mathcal{M}^d \to \mathbb{R}^+$ are of class $\mathcal{U}_p$ and $\mathcal{V}_p$ for some $1 < p < \infty$. In addition,*

- *there exists a continuous function $\tilde{f}_{nn} : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^+$ which is three times continuously differentiable, convex, increasing in its second variable, satisfies $\tilde{f}'_{nn}(z, 0) > 0$ locally uniformly in $z \in \mathbb{R}^d$, and such that for all $x, \lambda \in \mathbb{R}^d$, $f_{nn}(x, \lambda) = \tilde{f}_{nn}(x, |\lambda|^2)$;*
- *W_{vol} is a three times continuously differentiable quasiconvex function.*

Under these assumptions, we define two maps $A : \mathcal{M}^d \times \mathcal{T} \to (\mathcal{M}^d)^2$ and $B : \mathcal{M}^d \times \mathcal{M}^d \times \mathcal{T} \to (\mathcal{M}^d)^2$ by: for all $G, \Lambda \in \mathcal{M}^d$ and every simplex $T \in \mathcal{T}$

$$
\begin{aligned}
A(G,T) = \frac{1}{n_S} \sum_{(x,y)\in\mathcal{N}\cap T} \Bigg[& 2\tilde{f}'_{nn}\Big(x-y, |G\cdot\frac{x-y}{|x-y|}|^2\Big) \sum_{j=1}^{d} \frac{x-y}{|x-y|} \otimes \mathbf{e}_j \otimes \frac{x-y}{|x-y|} \otimes \mathbf{e}_j \\
& +4\tilde{f}''_{nn}\Big(x-y, |G\cdot\frac{x-y}{|x-y|}|^2\Big) \frac{x-y}{|x-y|} \otimes (G\cdot\frac{x-y}{|x-y|}) \otimes \frac{x-y}{|x-y|} \otimes (G\cdot\frac{x-y}{|x-y|}) \Bigg],
\end{aligned}
$$

$$
B(\Lambda, G, T) = |T| \Big(\frac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(\Lambda) - \frac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(G) \Big),
$$

where $\tilde{f}'_{nn}$ and $\tilde{f}''_{nn}$ denote the first and second derivatives of $\tilde{f}_{nn}$ with respect to its second variable. These maps take values in the set of symmetric fourth-order tensors and are essentially parts of the local elasticity tensor associated with the energy functional (recall however that this is a discrete energy).

Let $\Lambda \in \mathcal{M}^d$, and $\theta > 0$. We say that an open set $\mathcal{O}$ of $\mathcal{M}^d$ is (Λ, θ)-admissible for $\mathcal{T}$, f_{nn} and W_{vol} if

- $\Lambda \in \mathcal{O}$;
- for all $T \in \mathcal{T}$ and all $G \in \mathcal{O}$,

$$(1-\theta)A(G,T) - B(\Lambda, G, T) \ > \ 0$$

  in the sense of symmetric fourth order tensors.

Let $\mathcal{L}$ be an admissible stochastic lattice and $\mathcal{T}$ be the associated Delaunay triangulation of $\mathbb{R}^d$. The definition above makes sense for $\mathcal{L}$ and $\mathcal{T}$ for $d = 2$ and $d = 3$ since

- the hard-core and the non-empty space conditions yield a deterministic bound on n_S;
- for all $\Lambda \in \mathcal{M}^d$, (Λ, θ)-admissible subsets of $\mathcal{M}^d$ exist for $\theta > 0$ small enough.

We focus on the second statement. Since W_{vol} is three-times continuously differentiable, $B(\Lambda, G, T)|T|^{-1}$ is arbitrary small provided $\mathcal{O}$ is included in some sufficiently small ball centered at Λ. Note that $\mathcal{T}$ has uniformly bounded edge lengths because of the non-empty space condition, so that $|T|$ is uniformly bounded. If T was a simplex with d edges oriented in the d canonical directions, the first term in $A(G,T)$ would yield a coercive fourth-order tensor with coercivity constant $\frac{2}{n_S} \inf_{\mathcal{O}} \inf_{(x,y)\in\mathcal{N}} \tilde{f}'_{nn}\Big(x-y, |G \cdot \frac{x-y}{|x-y|}|^2\Big)$, which is positive uniformly with respect to $\mathcal{O}$, provided $\mathcal{O}$ is included in some ball centered at Λ and of arbitrary (yet fixed) radius, since $x-y$ lies in a compact set of $\mathbb{R}^d$ for the random parking measure. It remains to treat general simplices. In dimension $d = 2$, the hard-core and non-empty space conditions together with the properties of the Delaunay triangulation imply a uniform lower bound on the angles of T, so that $A(G,T)$ is a coercive fourth-order tensor uniformly in $T \in \mathcal{T}$. In dimension $d = 3$, this is not enough, and nothing prevents the vertices of the tetrahedron T to lie “almost” on the same equatorial plane of its circumscribed sphere. Denoting by h_T the smallest height of T, one may prove that $(h_T)^{-1}A(G,T)$ is a coercive fourth-order tensor uniformly in $T \in \mathcal{T}$. Noticing that $(h_T)^{-1} \sim |T|^{-1}$ since the edge lengths are uniformly bounded from above and below, the additional factor $|T|$ in the definition of $B(\Lambda, G, T)$ compensates for the term $(h_T)^{-1}$ needed to have the uniformity of the coercivity constant of A. It is now elementary to deduce the existence of (Λ, θ)-admissible sets of $\mathcal{M}^d$ for $\theta > 0$ small enough, for $d = 2, 3$.

We have all the ingredients to prove the perturbation result. The perturbation assumption is as follows:

Hypothesis 6 *Let $\Lambda \in \mathcal{M}^d$, $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$, and $t_0 > 0$. In addition to the assumptions of Theorem 23, we let f_{nn} and W_{vol} satisfy Hypothesis 5. For all $0 \le t \le t_0$, we set $\varphi_{\Lambda+tH} : x \mapsto (\Lambda + tH)x$ and let $\phi^N_{\Lambda+tH} \in W^{1,\infty}_0(Q_N, \mathbb{R}^d)$ be a sequence of functions such that almost surely*

$$F_1^{Q_N}(\mathcal{L}, \phi^N_{\Lambda+tH} + \varphi_{\Lambda+tH}) \ = \ \inf_u \Big\{ F_1^{Q_N}(\mathcal{L}, u) \ \Big| \ u \in \mathfrak{S}_1^{Q_N}(\mathcal{L}) \text{ such that } u(x) = (\Lambda + tH)\cdot x \text{ if } x \in \mathcal{L} \cap Q_N \text{ and } d_\infty(x, \partial Q_N) \le 2\rho_2 \Big\}.$$

We assume that there exists $r(t)$ with $r(t) \to 0$ as $t \to 0$ such that for all $N \geq 1$,

$$\|\nabla \phi^N_{\Lambda+tH} - \nabla \phi^N_{\Lambda}\|_{L^\infty(Q_N)} \leq r(t)$$

almost surely, and that there exist $\theta > 0$ and a (Λ, θ)-admissible open subset $\mathcal{O}$ of $\mathcal{M}^d$ such that for all N large enough $\Lambda + \nabla \phi^N_\Lambda \in \mathcal{O}$ pointwise almost surely.

The first assumption is similar to Hypothesis 3, and ensures the commutation of linearization and homogenization. The second part of Hypothesis 6 is the perturbation assumption, since it implies that $\nabla \phi^N_\Lambda$ is small. The theorem then reads:

Theorem 26 *Let $d = 2$ or $d = 3$, let $\Lambda \in \mathcal{M}^d$, and let f_{nn} and $W_{\rm vol}$ be as in Hypothesis 6. If for all $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$, Hypothesis 6 holds, then there exists a deterministic matrix $\Pi_{\rm hom} \in \mathcal{M}^d$, and a strictly strongly elliptic elasticity tensor $\mathbb{L}_{\rm hom}$ such that the homogenized energy density $W_{\rm hom}$ of Theorem 23 satisfies the Taylor expansion*

$$W_{\rm hom}(\Lambda + tH) = W_{\rm hom}(\Lambda) + t\Pi_{\rm hom} \cdot H + \frac{t^2}{2} H \cdot \mathbb{L}_{\rm hom} H + o(t^2)$$

for all $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$. Hence $W_{\rm hom}$ is strictly strongly elliptic at Λ.

There are two steps in the proof of this theorem. First we prove the strict strong ellipticity of $W_N(\Lambda) = F_1^{Q_N}(\mathcal{L}, \phi^N_{\Lambda+tH} + \varphi_{\Lambda+tH})$ uniformly with respect to N by introducing an elasticity tensor $\mathbb{L}_N$. Then we show that one can take the limit as N goes to infinity using the perturbation assumption. We only display the argument for the first step.

For all $N \geq 1$, we define the elasticity tensor $\mathbb{L}_N$ by: for all $G \in \mathcal{M}^d$,

$$G \cdot \mathbb{L}_N G := \frac{1}{|Q_N|} \inf \Big\{ \psi \cdot \mathrm{D}^2 F_1^{Q_N}(\mathcal{L}, \varphi_\Lambda + \phi^N_\Lambda)\psi \;\Big|\; \psi \in \mathfrak{S}_1^{Q_N}(\mathcal{L}) \text{ such that } \psi(x) = G \cdot x$$
$$\text{if } x \in \mathcal{L} \cap Q_N \text{ and } d_\infty(x, \partial Q_N) \leq 2\rho_2 \Big\},$$

where $\psi \cdot \mathrm{D}^2 F_1^{Q_N}(\mathcal{L}, \varphi_\Lambda + \phi^N_\Lambda)\psi$ is a notation for the second variation of $F_1^{Q_N}(\mathcal{L}, \varphi_\Lambda + \phi^N_\Lambda)$ in the direction ψ. We denote by ψ^N_G an associated minimizer.

Recall that $\nabla \psi^N_G$ and $\nabla \phi^N_\Lambda$ are piecewise constant on $\mathcal{T}$. By definition of A and B, we have:

$$\begin{aligned}
|Q_N| G \cdot \mathbb{L}_N G \geq & \sum_{T \in \mathcal{T}, T \subset Q_N} \nabla \psi^N_G \cdot A(\Lambda + \nabla \phi^N_\Lambda, T) \nabla \psi^N_G \\
& + \sum_{T \in \mathcal{T}, T \subset Q_N} |T| \nabla \psi^N_G \cdot \frac{\partial^2 W_{\rm vol}}{\partial \Lambda^2}(\Lambda + \nabla \phi^N_\Lambda) \nabla \psi^N_G, \\
= & \sum_{T \in \mathcal{T}, T \subset Q_N} \Big(\nabla \psi^N_G \cdot A(\Lambda + \nabla \phi^N_\Lambda, T) \nabla \psi^N_G - \nabla \psi^N_G \cdot B(\Lambda, \Lambda + \nabla \phi^N_\Lambda, T) \nabla \psi^N_G \Big) \\
& + \sum_{T \in \mathcal{T}, T \subset Q_N} |T| \nabla \psi^N_G \cdot \frac{\partial^2 W_{\rm vol}}{\partial \Lambda^2}(\Lambda) \nabla \psi^N_G,
\end{aligned}$$

so that by the perturbation assumption,

$$|Q_N| G \cdot \mathbb{L}_N G \;\geq\; \theta \sum_{T \in \mathcal{T}, T \subset Q_N} \nabla \psi_G^N \cdot A(\Lambda + \nabla \phi_\Lambda^N, T) \nabla \psi_G^N \\ + \sum_{T \in \mathcal{T}, T \subset Q_N} |T| \nabla \psi_G^N \cdot \frac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(\Lambda) \nabla \psi_G^N.$$

Let now G be a rank-one matrix, that we denote by $H = a \otimes b \in \mathcal{M}^d$ with $|H| = 1$. Then, since $\psi_H^N - \varphi_H \in H_0^1(\cup_{T \subset Q_N} T)$,

$$\sum_{T \in \mathcal{T}, T \subset Q_N} |T| \nabla \psi_H^N \cdot \frac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(\Lambda) \nabla \psi_H^N \;\geq\; \sum_{T \in \mathcal{T}, T \subset Q_N} |T| H \cdot \frac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(\Lambda) H \;\geq\; 0$$

by rank-one convexity of the elasticity tensor $\dfrac{\partial^2 W_{\mathrm{vol}}}{\partial \Lambda^2}(\Lambda)$ (see for instance [19]). Hence

$$|Q_N| H \cdot \mathbb{L}_N H \;\geq\; \theta \sum_{T \in \mathcal{T}, T \subset Q_N} \nabla \psi_H^N \cdot A(\Lambda + \nabla \phi_\Lambda^N, T) \nabla \psi_H^N.$$

To conclude we recall that due to the admissibility of $\mathcal{L}$, $|T|^{-1} A(\Lambda, T)$ is coercive uniformly in $T \in \mathcal{T}$ with constant c. The same holds for $|T|^{-1} A(\Lambda + \nabla \phi_\Lambda^N, T)$ using in addition Hypothesis 5. Hence

$$\begin{aligned} |Q_N| H \cdot \mathbb{L}_N H &\geq \theta \sum_{T \in \mathcal{T}, T \subset Q_N} |T| \nabla \psi_H^N \cdot |T|^{-1} A(\Lambda + \nabla \phi_\Lambda^N, T) \nabla \psi_H^N \\ &\geq \theta \sum_{T \in \mathcal{T}, T \subset Q_N} c \|\nabla \psi_H^N\|_{L^2(T)}^2 \\ &= c\theta \|\nabla \psi_H^N\|_{L^2(\cup_{T \subset Q_N} T)}^2, \end{aligned}$$

so that

$$H \cdot \mathbb{L}_N H \;\geq\; c\theta |H|^2 = \theta c$$

by convexity of the L^2-norm using again the fact that $\psi_H^N - \varphi_H \in H_0^1(\cup_{T \subset Q_N} T)$. This ellipticity constant θc is uniform in N, as desired.

4.6 Towards an analytical formula for the homogenized energy density [BGLTV]

In view of the previous sections one could be satisfied with the analysis and numerical simulation of the discrete model for rubber. Yet, in order to use this model in computational mechanics softwares, the evaluation of the Piola-Kirchoff stress tensor should be fast (this routine is called at each Gauss point of the quadrature rule). Unfortunately this is a very expensive computation. To approximate the stress tensor at one single deformation using the method of Subsection 4.4.2, one needs to solve 10 nonlinear problems with 10^5 degrees of freedom. This is too prohibitive and one cannot afford to include it into a nonlinear elasticity software. (This problematics is similar to the one treated in Subsection 2.4.3 in another context.)

The aim of this section is to construct an analytical proxy for W_{hom} using a set of data generated by the method of Subsection 4.4.2. Data assimilation in rubber elasticity may be an ill-posed problem because the sets of data which are available are often too partial (engineering stress for uniaxial and biaxial tractions for instance). In particular all the regimes cannot be tested by mechanical experiments. On the contrary, for the discrete model for rubber any strain gradient can be considered. In particular we have at our disposal an arbitrary amount of data at arbitrary values of the strain gradient. This opens the door to the use of reliable and efficient data assimilation techniques. In addition, the analysis of the model and of its thermodynamic limit is also a very good guide to restrict the class of admissible energy densities in which to solve the inverse problem.

4.6.1 Choice of a parametrization: Ogden's laws

From the analysis of Section 4.2 we learn that the homogenized integrand W_{hom} (after a change of reference configuration) associated with the discrete model for rubber is quasiconvex, frame-invariant, isotropic, minimal at identity, and hopefully strictly strongly elliptic. Ideally we'd like to use a characterization of this class of functions to devise a parametric data assimilation method. Yet there does not exist any tractable characterization of quasiconvexity [49]. This is a serious practical handicap.

As seen in Section 4.5, the specific form of the discrete model for rubber is very "close" to polyconvexity. In particular if the affine assumption held, the minimum in (4.10) associated with the energy functions (4.20) and (4.21) would be attained for $u(x) = \Lambda \cdot x$, and W_{hom} would be a polyconvex function. Yet, the affine assumption does not hold. Furthermore, unlike quasiconvexity, polyconvexity is not preserved by homogenization (see for instance [2, 14]). Although it is not clear whether replacing quasiconvexity by polyconvexity is justified here, polyconvexity can be handled whereas quasiconvexity cannot and we shall look for an analytical approximation of W_{hom} in the class of polyconvex, isotropic, and frame-invariant, strictly strongly elliptic energy densities, which admit the identity as unique natural state.

To proceed we need a characterization of this manifold (that we shall denote by $\mathcal{P}$). Using the density of convex polynomials in the set of convex functions on bounded domains for the norm of the supremum, we directly deduce the density of polyconvex polynomials in the set of polyconvex functions. Again, it is not an easy task to characterize the set of polyconvex polynomials in terms of their coefficients, and this remains a handicap for the numerical practice of parameter identification. This is where Ogden's laws come into the picture.

As a first practical example of polyconvex functions, Ball considers in his seminal paper [5] the case of the Ogden laws introduced in [72] to model frame-invariant isotropic rubber materials. Relying on the Rivlin-Eriksen representation theorem, Ogden has proposed a restricted class of energy densities:

$$W_{\mathrm{og}}(\Lambda) = \sum_{i=1}^{k_1} a_i(\lambda_1^{\alpha_i} + \lambda_2^{\alpha_i} + \lambda_3^{\alpha_i}) + \sum_{j=1}^{k_2} b_j\big((\lambda_1\lambda_2)^{\beta_j} + (\lambda_2\lambda_3)^{\beta_j} + (\lambda_3\lambda_1)^{\beta_j}\big) + W_3(\lambda_1\lambda_2\lambda_3), \tag{4.26}$$

where λ_1, λ_2 and λ_3 are the square-root of the eigenvalues of the Cauchy-Green strain tensor $\Lambda^T\Lambda$ (or singular values of Λ), $k_1, k_2 \in \mathbb{N}$, $a_i, b_i, \alpha_i, \beta_i \in \mathbb{R}$, and W_3 is a convex function. This class of constitutive laws is rather large, although it is not clear whether its intersection with $\mathcal{P}$ is dense in $\mathcal{P}$ for the topology of local uniform convergence. The interest of Ogden's laws is the following: Ball has obtained in [5] (see also [21, Theorem 4.9-2]) a rather simple set of conditions which ensures the polyconvexity of Ogden's laws:

$$\forall\, i \in \{1, \dots, k_1\},\ a_i > 0, \alpha_i \geq 1, \quad \forall\, j \in \{1, \dots, k_2\},\ b_j > 0, \beta_j \geq 1. \tag{4.27}$$

This is of utmost interest to solve the identification problem with the constraint of polyconvexity (at least in this subclass of polyconvex Ogden's laws).

The second constraint requires that W_{og} be minimal at identity:

$$W_{\mathrm{og}}(\mathrm{Id}) \;=\; \inf W_{\mathrm{og}}. \tag{4.28}$$

The class of functions in which we shall approximate W_{hom} is the following: Ogden's laws (4.26) satisfying the conditions (4.27) and (4.28), with $W_3 : (0, +\infty) \to \mathbb{R}$ characterized by

$$W_3(t) \;=\; K_1 t^2 - 2K_2 \log t$$

for some $K_1, K_2 \geq 0$, which is a variant of the Helmholtz energy density (replacing a single constant K by two constants K_1 and K_2 ensures that one can impose the identity to be a natural state of (4.26)). This manifold is rather complex. Let k_1 and k_2 be implicitly fixed, and set $n = 2(k_1 + k_2 + 1)$. Given $p = (p_1, p_2) \in \mathbb{R}^n$ with $p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2})$ and $p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2})$, we denote by W_{og}^p the associated Ogden law.

4.6.2 Choice of a cost-function

To find the Odgen constitutive law (4.26) which best approximates W_{vol}, we have to identify the corresponding set of parameters $p = (p_1, p_2) \in \mathbb{R}^n$. This is done by minimizing a cost function $\mathcal{E}$ depending on the difference between quantities associated with the energy density W_{og}^p (seen as a function of p) and the corresponding quantities obtained by "in silico experiments" (namely numerical approximations of W_{hom}).

Since we are interested in boundary value problems, the important quantity is not the energy density W_{hom} itself, but rather its derivatives. The numerical method of Section 4.4

which allows us to evaluate $W_{\mathrm{hom}}(\Lambda)$ at any deformation gradient Λ relies on a Newton algorithm. In particular this method provides as outputs approximations of:

$$\Pi_{\mathrm{hom}}(\Lambda) = \frac{\partial W_{\mathrm{hom}}}{\partial \Lambda}, \qquad \text{the first Piola-Kirchhoff stress tensor and}$$

$$\mathbb{L}_{\mathrm{hom}}(\Lambda) = \frac{\partial^2 W_{\mathrm{hom}}}{\partial \Lambda^2}, \qquad \text{the Hessian tensor.}$$

These approximations are the observations for the inverse problem. Since the material is isotropic and frame-invariant, W_{hom} is characterized by its values on diagonal matrices $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ and we only consider such deformation gradients. In particular, the associated Piola-Kirchhoff stress tensor is diagonal and we denote by $\{\tilde{\Pi}_i\}_{1\le i\le 3}$ its diagonal entries. We also set $\{\tilde{H}_{ij}\}_{1\le i,j\le 3}$ the entries $\dfrac{\partial^2 W_{\mathrm{hom}}}{\partial \Lambda_{ij}^2}$ of the Hessian tensor (which is positive for strongly elliptic materials).

With this preliminary, we are in position to introduce the cost function we shall consider:

$$c(p;\lambda_1,\lambda_2,\lambda_3) := \frac{\sum_{i=1}^3 \left(\Pi_i^p - \tilde{\Pi}_i\right)^2}{\sum_{i=1}^3 \tilde{\Pi}_i^2} + \eta \sum_{i=1}^3 \sum_{j=1}^3 \left(\frac{H_{ij}^p - \tilde{H}_{ij}}{\tilde{H}_{ij}}\right)^2, \tag{4.29}$$

where $\eta \ge 0$ is a small regularization parameter, and Π^p and H^p are used for the first Piola-Kirchhoff stress tensor and Hessian tensor associated with the Ogden law W_{og}^p of parameter p.

In the cost function, we would like to restrain the values of the admissible deformations $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ to entries in the interval $(1/6, 6)$ (deformation up to 500%), and to give a bigger weight to strain gradients in small deformations (λ_i close to 1). In order to do so, we introduce the following importance measure on $(1, 6)$:

$$\mu_\kappa(x) = K_\kappa (x-1)^\kappa, \tag{4.30}$$

where $\kappa > -1$ and K_κ is a normalization factor such that :

$$\int_1^6 \mu_\kappa(x)dx = 1 \iff K_\kappa = \frac{1+\kappa}{5^{1+\kappa}}. \tag{4.31}$$

The parameter κ can be chosen such that, for some fixed $x^0 \in (1,6)$, we have

$$\int_1^{x^0} \mu_\kappa(x)dx = \int_{x^0}^6 \mu_\kappa(x)dx \iff \kappa = \frac{\log(2)}{\log(5) - \log(x^0 - 1)} - 1, \tag{4.32}$$

that is the weights given to the intervals $(1, x^0)$ and $(x^0, 6)$ are the same. This is a convenient way to give more importance to the small deformation regime. In Figure 4.15, the weight function $(x-1)^\kappa$ is plotted for different values of κ. The formula (4.30) defines μ_κ on $(1,6)$

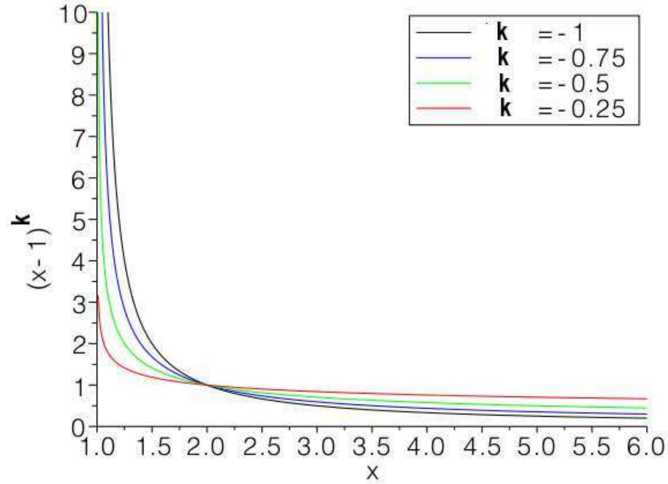

Fig. 4.15. Jacobi weight function $x \mapsto (x-1)^\kappa$ (for $\kappa \in \{-1, -0.75, -0.5, -0.25\}$)

and we extend this function to $(1/6, 1)$ by setting $\mu_\kappa(1/x) = \mu_\kappa(x)$ for all $x \in (1, 6)$. Thus, the measure μ_κ gives equal weight to compression and extension.

In the case of quasi-incompressible materials, the nonlinear constraint $\det \Lambda = \lambda_1\lambda_2\lambda_3 \simeq 1$ has to be taken into account. To this end, we consider the reduced principal strains

$$J = \lambda_1\lambda_2\lambda_3, \qquad \nu_1 = \frac{\lambda_1}{J^{1/3}}, \quad \nu_2 = \frac{\lambda_2}{J^{1/3}}, \quad \text{and} \quad \nu_3 = \frac{\lambda_3}{J^{1/3}}.$$

As primary variables we take J, ν_1 and ν_2, and restrict J to $(1-\delta, 1+\delta)$, for some small $\delta > 0$. Finally, we define the following global cost function:

$$\mathcal{F}(p) = \left(\frac{\displaystyle\int_1^6 \int_1^6 \int_{1-\delta}^{1+\delta} c\left(p; J^{1/3}\frac{1}{\nu_1}, J^{1/3}\nu_2, J^{1/3}\frac{\nu_1}{\nu_2}\right) \mu_\kappa(\nu_1)\mu_\kappa(\nu_2) dJ d\nu_1 d\nu_2}{\displaystyle\int_1^6 \int_1^6 \int_{1-\delta}^{1+\delta} \mu_\kappa(\nu_1)\mu_\kappa(\nu_2) dJ d\nu_1 d\nu_2} \right)^{1/2}. \tag{4.33}$$

Note that by symmetry of c, we take into account all the possible deformation gradients $\Lambda = J^{1/3}\mathrm{diag}(\nu_1, \nu_2, \nu_3)$ with $\nu_1, \nu_2, \nu_3 \in (1/6, 6)$ such that $\nu_1\nu_2\nu_3 = 1$. The identification problem consists in finding $p \in \mathbb{R}^n$ that minimizes $\mathcal{F}$ with the constraints (4.27) and (4.28).

We then approximate the integral in J with a standard three points integration rule and the integrals in ν_1 and ν_2 by the Jacobi integration rule of order $m \in \mathbb{N}$, that is:

$$\int_1^6 f(x)(x-1)^\kappa dx \simeq \sum_{k=1}^m \omega_k f(x_k), \tag{4.34}$$

where x_k, for $k \in \{1, \dots, m\}$, are the roots of the $(0, \kappa)$-Jacobi polynomial of degree m (after a mapping from $(-1, 1)$ to $(1, 6)$) and ω_k are the corresponding weights.

The minimization problem we are considering is as follows:

$$\inf \Bigg\{ \mathcal{F}(p) \,\Bigg|\, p = (p_1, p_2) \in \mathbb{R}^n, p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2}),$$
$$p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2}) \;:$$
$$\alpha_i, \beta_i \geq 1, a_i, b_i, K_1, K_2 \geq 0,\; W^p_{\mathrm{og}}(\mathrm{Id}) = \inf W^p_{\mathrm{og}} \Bigg\}.$$

4.6.3 Numerical method

The minimization problem considered here has the following three properties:

- the cost function $\mathcal{F}$ is not convex and defines a rather complex energy landscape with numerous local minima;
- the cost function $\mathcal{F}$ is twice-differentiable;
- the set of parameters is a nonlinear implicitly defined manifold.

Due to the complexity of the energy landscape, Newton-type algorithms — or more generally deterministic algorithms — are very likely to get trapped into local infima.

In order to circumvent (or at least reduce) this difficulty, we shall appeal to a stochastic optimization procedure — namely, an evolutionary algorithm. The fact that $\mathcal{F}$ is twice-differentiable does not help finding the right region of the energy landscape.

The general procedure is as follows. Assume momentarily that we minimize $\mathcal{F}$ without the constraints (4.27) and (4.28), and that the minimizer $\bar{p}$ of $\mathcal{F}$ on $\mathbb{R}^n$ is unique. Instead of directly looking for $\bar{p}$ in $\mathbb{R}^n$, we look for a probability measure $\bar{\mu}$ which minimizes $\mu \mapsto \int_{\mathbb{R}^n} \mathcal{F}(p) d\mu(p)$. Of course, by uniqueness of the minimizer, $\bar{\mu} = \delta_{\bar{p}}$, the direct mass at $\bar{p}$. The strategy is now to approximate $\bar{\mu}$ by a sequence of Gaussian measures G_k, characterized by their means $m_k \in \mathbb{R}^n$, their covariance matrices $C_k \in SO_n(\mathbb{R})$, and their standard deviation $\sigma_k \in \mathbb{R}^+$. The sequence G_k is an approximation of $\bar{\mu}$ if $\lim_{k\to\infty} G_k = \bar{\mu}$ weakly in the sense of measures.

The evolutionary algorithm is characterized by its updating procedure, that is the construction of G_{k+1} knowing G_k. Let $S > s > 0$ be integers. Given a sampling $p_1^k, \dots, p_S^k$ of G_k, we select the s best search points p_i^k (that is those s points among the S search points which yield the s minimal values of $\mathcal{F}$). The mean m_{k+1} of G_{k+1} is then obtained by taking a (suitable) weighted average of the s best search points, the covariance matrix C_{k+1} is chosen so that the s search points are a suitable sampling of a Gaussian measure with this covariance matrix C_{k+1}. It remains to choose the standard deviation σ_{k+1}. The larger σ_{k+1}, the more regions of the energy landscape will be potentially visited. Yet, $G_k \to \bar{\mu}$ requires $\sigma_k \to 0$. The choice of σ_{k+1} is therefore crucial, and case-dependent. Once G_{k+1} is defined, one generates randomly S samples p_i^{k+1}. We shall use the Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) algorithm. For the precise update of G_k, we refer to [41].

Let us now describe how the CMA-ES algorithm is used in our context. We first neglect the constraint (4.28) that the Odgen energy density should be minimal at identity. The cost function $\mathcal{F}$ has a very specific structure, and the parameters $p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2})$ and $p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2})$ do have different roles. In particular, since the dependence of $\mathcal{F}$ upon p_2 is quadratic (after taking $\mathcal{F}$ to the square), it makes sense to consider the reduced cost function

$$\mathcal{F}_r(p_1) \;:=\; \inf_{p_2 \geq 0} \mathcal{F}(p_1, p_2).$$

The infimum can indeed be effectively computed by deterministic methods (recall that we have neglected the constraint (4.28) that $W_{\rm og}^p$ is minimal at identity). Since

$$\inf_{p \in \mathbb{R}^n \text{ s. t. } (4.27)} \mathcal{F}(p) \;=\; \inf_{p_1 \text{ s. t. } (4.27)} \mathcal{F}_r(p_1),$$

one may either apply the CMA-ES algorithm to $\mathcal{F}$ or $\mathcal{F}_r$. There are two main differences: the nonlinearity of the functionals ($\mathcal{F}_r$ is much more nonlinear than $\mathcal{F}$ since the minimizers p_2 are themselves nonlinear functions of p_1) and the dimension of the parameters ($k_1 + k_2$ for $\mathcal{F}_r$, $2(k_1+k_2+1)$ for $\mathcal{F}$). In both cases we impose the constraint (4.27) on the parameters p_1 by penalization so that the search space remains the linear space $\mathbb{R}^{k_1+k_2}$, and not $\{p_1 \in \mathbb{R}^{k_1+k_2}$ such that $(4.27)\}$.

This picture would be complete if we did not have to deal with the constraint (4.28) that $W_{\rm og}^p$ should be minimal at identity. In order to take this constraint into account, we use a splitting method, and add a projection step to the algorithm. The idea is to proceed by prediction-correction to take into account the constraint (4.28). This amounts to minimizing a different functional $\overline{\mathcal{F}}_r$ defined as follows. Given p_1, we let $\tilde{p}_2$ be a minimizer of $\mathcal{F}(p_1, \cdot)$ on $[0, +\infty)^{k_1+k_2+2}$, and set $\tilde{p} = (p_1, \tilde{p}_2) \in \mathbb{R}^n$. If $W_{\rm og}^{\tilde{p}}$ satisfies (4.28), we set $\overline{\mathcal{F}}_r(p_1) := \mathcal{F}_r(p_1)$. Otherwise, we "project" $W_{\rm og}^{\tilde{p}}$ on the set of Ogden laws satisfying (4.28). To this aim, we let $\gamma > 0$ be the unique minimizer of $t \mapsto W_{\rm og}^{\tilde{p}}(t\mathrm{Id})$ on $\mathbb{R}$, which we may compute by a Newton algorithm (the problem is strictly convex), and finally define

$$\overline{\mathcal{F}}_r(p_1) \;:=\; \mathcal{F}(p_1, p_2),$$

where $p_2 := (\gamma^6 K_1, K_2, \gamma^{\alpha_1} a_1, \cdots, \gamma^{\alpha_{k_1}} a_{k_1}, \gamma^{2\beta_1} b_1, \cdots, \gamma^{2\beta_{k_2}} b_{k_2})$. Setting $p = (p_1, p_2)$, this ensures that $W_{\rm og}^p$ satisfies the minimality condition at identity (4.28).

The splitting method consists in minimizing the functional $p_1 \mapsto \overline{\mathcal{F}}_r(p_1)$ on $[1, +\infty)^{k_1+k_2}$ by the evolutionary algorithm. For all $p_1 \in \mathbb{R}^{k_1+k_2}$, $\overline{\mathcal{F}}_r(p_1) \geq \mathcal{F}_r(p_1)$ by definition. Hence it is not clear whether minimizing $\overline{\mathcal{F}}_r$ is equivalent to minimizing $\mathcal{F}_r$. This is the case if any minimizer p_1 of $\mathcal{F}_r$ satisfies the identity

$$\inf_{p_2 \text{ s.t. } (4.27)\&(4.28)} \mathcal{F}(p_1, p_2) \;=\; \inf_{p_2 \text{ s.t. } (4.27)} \mathcal{F}(p_1, p_2). \tag{4.35}$$

When this condition does not hold, the splitting procedure only gives an approximation of the minimizer.

4.6.4 Numerical results

We present two series of tests based on the minimization of $\overline{\mathcal{F}}_r$ by the evolutionary algorithm. In the numerical tests we have not used the stabilization by the Hessian since it does not provide any significant improvement. It seems from the numerical tests that $m = 3$ is enough for the Jacobi integration rule.

In the first series of tests, we start with an Ogden law and try to recover its coefficients. In the second series of tests, the data are generated by the discrete model for rubber. Note that in the first case, the property (4.35) holds. In the second case, it is not known whether (4.35) holds or not.

In the case of the Ogden law, results are gathered in Table 4.2. As can be seen, the coefficients can be recovered quite precisely. In order to check the robustness of the algorithm,

error $\mathcal{F}$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1	b_2	β_2
exact	2	1	5.86	10.00	0.15	5.10	4.09	1.80			0.03	2.30		
10^{-1}	1	0	6.20	8.24	1.05	3.89								
10^{-1}	0	1	2.32	13.98							9.00	1.30		
10^{-1}	1	1	4.26	12.61	0.35	4.63					7.55	1.00		
10^{-3}	2	0	5.90	10.01	0.15	5.10	4.12	1.80						
10^{-7}	2	1	5.86	10.00	0.15	5.10	4.09	1.80			0.03	2.30		
10^{-7}	2	2	5.86	10.00	0.15	5.10	4.09	1.80			0.03	2.30	10^{-7}	6.57
10^{-7}	3	1	5.86	10.00	0.15	5.10	4.09	1.80	10^{-8}	5.43	0.03	2.30		

Table 4.2. Recovering process for different k_1 and k_2 starting from an Ogden law.

we have performed the same tests with the addition of noise on the data (by independent and identically distributed random variables on each data, of the order of a few percents). The results for typical realizations are diplayed in Table 4.3. These tests show that the algorithm is rather robust, and that the problem is stable (the solution seems to vary rather smoothly).

noise	error $\mathcal{F}$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
		2	1	5.86	10.00	0.15	5.10	4.09	1.80	0.03	2.30
0.00	10^{-7}	2	1	5.86	10.00	0.15	5.10	4.09	1.80	0.03	2.30
0.01	10^{-3}	2	1	5.83	10.00	0.15	5.10	4.14	1.79	0.03	2.30
0.02	10^{-3}	2	1	5.81	10.00	0.15	5.09	4.20	1.78	0.03	2.31
0.05	10^{-2}	2	1	5.73	10.00	0.16	5.08	4.39	1.74	0.02	2.35
0.10	10^{-2}	2	1	6.02	10.24	0.16	5.09	3.98	1.88	0.03	2.21

Table 4.3. Recovering process with noisy data.

We turn now to the core of this section: the analytical proxy for the discrete model for rubber. The results are gathered in Table 4.4. Note that in this case, the law we try to reconstruct does not belong a priori to the search space. In addition, the data we have

error	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1	b_2	β_2
10^{-2}	1	0	4.76	5.01	0.22	2.22								
10^{-2}	1	1	4.75	5.02	0.21	2.26					0.04	1.00		
10^{-3}	2	1	4.78	5.05	0.26	2.03	0.001	4.45			0.002	1.81		
10^{-3}	2	2	4.78	5.05	0.25	2.05	0.001	4.58			0.01	1.00	10^{-5}	3.43
10^{-3}	3	1	4.78	5.05	0.21	2.16	0.07	1.16	0.0002	5.16	0.002	1.97		
10^{-3}	3	2	4.78	5.05	0.21	2.15	0.07	1.16	0.0002	5.16	0.002	1.97	10^{-10}	8.26

Table 4.4. Recovering process for the data generated with the discrete model for rubber.

are only approximations of the homogenized model (cf. Section 4.4), so that a relative error of order 10^{-3} can be considered as a very good result. To complete this picture, we have plotted on Figure 4.16 the comparison between the engineering stress obtained by a direct simulation of the discrete model to the engineering stress given by the reconstructed analytical law on the Treloar experiments (uniaxial compression, uniaxial traction, planar tension). This illustrates the capability of the reconstructed law to reproduce the behavior of the homogenized integrand W_{hom} in various regimes.

4.7 Perspectives

There are several perspectives to the theory for rubber developed in this chapter.

From the analysis point of view, the discrete homogenization result could be extended to the case of volumetric energies which blow up as the determinant of the deformation gradient vanishes, at least in dimensions $d = n = 2$ and for Dirichlet boundary conditions (see Section 3.6).

From the modeling point of view, the connectivity of the network we have considered (Delaunay triangulation associated with the random parking measure) is too high, typically around 20. In practice, polymer chain networks have a connectivity between 3 and 4. A possible solution to reduce the connectivity of the network is to delete edges at random (this procedure is illustrated on Figure 4.18). Yet it is not clear whether the energy functional remains coercive. Instead of deleting edges one may also multiply the values of their energies by a parameter that we slowly set to zero, combined with a continuation method. In the case of uniaxial traction, this yields the results of Figure 4.17, which illustrates the influence of the connectivity on the engineering stress.

From the physical point of view, the model is rather simplistic at the polymer chain level. In particular no entanglements or topological constraints are taken into account. It would therefore be interesting to check the validity of the discrete model at the scale of the polymer chain network itself. This could be possible using results obtained by F. Boué's group at CEA Saclay. This group of physicists has indeed been able to plot the Fourier transform of the length distribution of the polymer chains in a (two-dimensional) deformed network, by using optical measurements. Our numerical procedure to approximate the homogenized integrand W_{hom} also allows us to compute (as output) the Fourier transform of the length distribution of the deformed polymer chains. These two plots could definitely be compared.

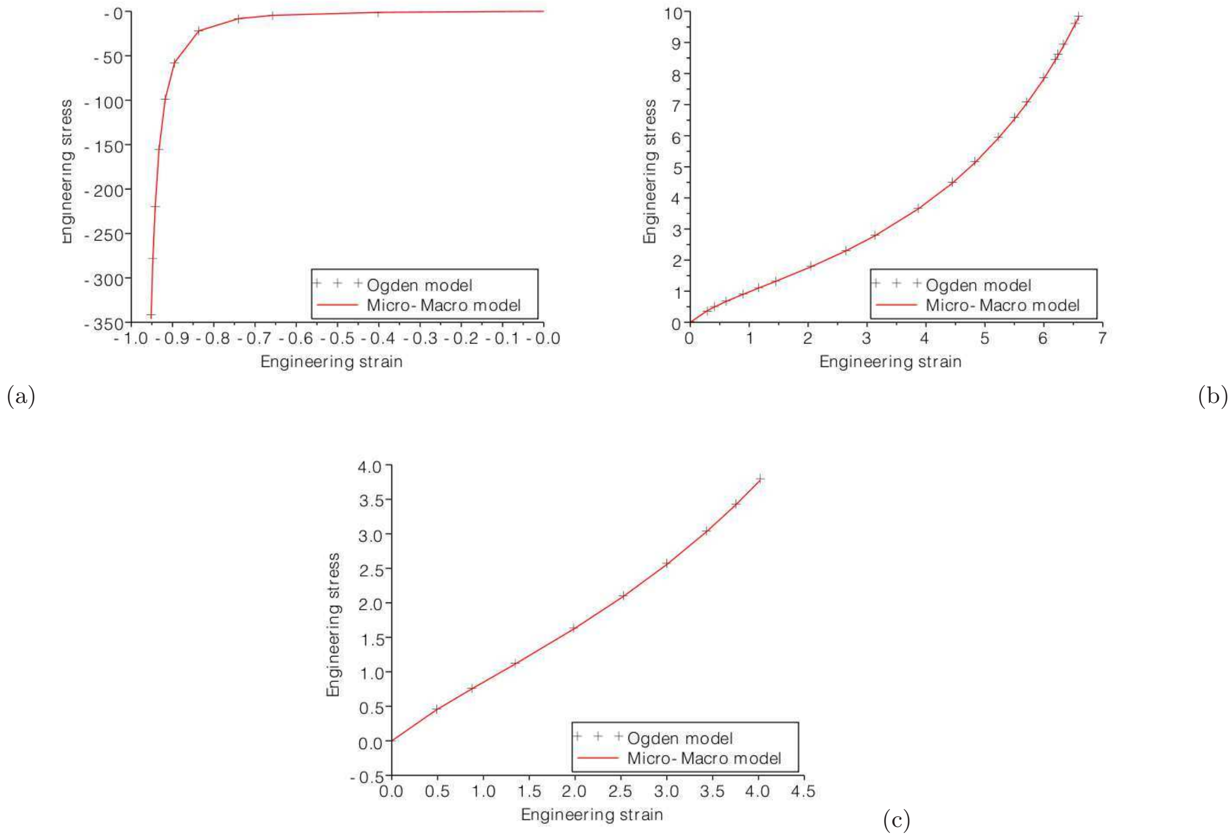

Fig. 4.16. Comparison between the discrete model for rubber and the reconstructed Odgen law. (a) uniaxial compression (b) uniaxial traction $\xi = diag(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}})$ and (c) planar tension $\xi = diag(\lambda, \frac{1}{\lambda}, 1)$. Evolution of the engineering stress with respect to the engineering strain $\lambda - 1$.

From the numerical point of view, our numerical approximation method of the homogenized integrand W_{hom} is rather crude. As we have seen in Chapter 1, an efficient way to approximate homogenized coefficients in stochastic homogenization of linear elliptic equations on $\mathbb{Z}^d$ is the periodization method. In the present case, the method would be as follows. Let $\rho_1 > 0$ be fixed and let $R > 0$. Pick a point $x_0 \in Q_R$, and set $X_0 = x_0 + R\mathbb{Z}^d$. The point set X_0 is accepted, and we set $Y_0 = X_0$. Pick at random another point $x_1 \in Q_R$. If the points of $X_1 = x_1 + R\mathbb{Z}^d$ are at distance at least ρ_1 from the points of Y_0, X_1 is accepted and $Y_1 = Y_0 \cup X_1$. Otherwise $Y_1 = Y_0$. We continue this process untill Q_R (and in fact $\mathbb{R}^d$) is packed by Y_N for some large but (almost surely) finite N. The point set Y_N is Q_R-periodic. One can associate with Y_N a Q_R-periodic Delaunay triangulation of $\mathbb{R}^d$, and then approximate W_{hom} by the infimum of the discrete energy on Q_R with periodic boundary conditions. In view of the analysis of Chapter 1, the convergence rate to W_{hom} is expected to be better than the method used in this chapter. The practical implementation of this method is however much more involved.

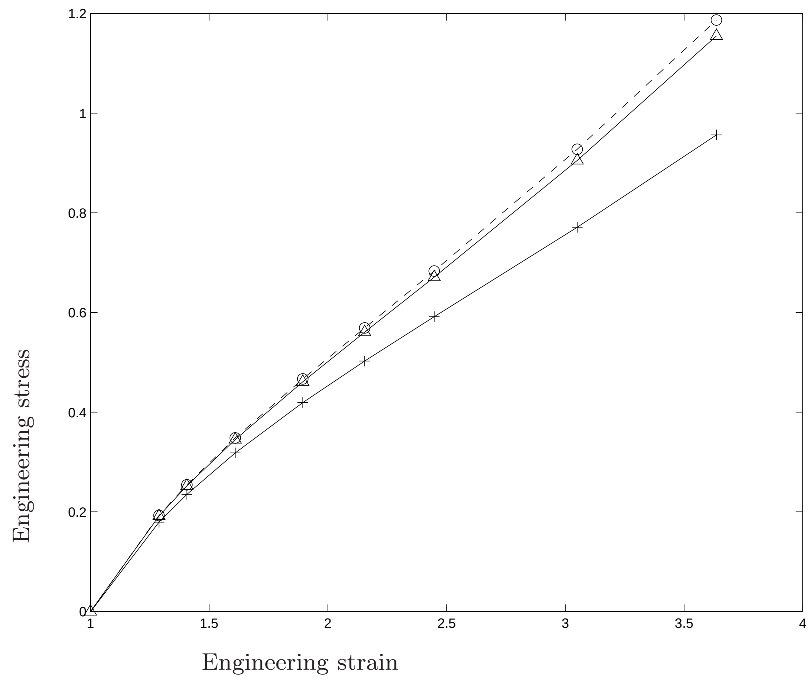

Fig. 4.17. Uniaxial traction — affine assumption (dashed line), variational model for $K = 50$ and connectivities 20 and 4 (from top to bottom)

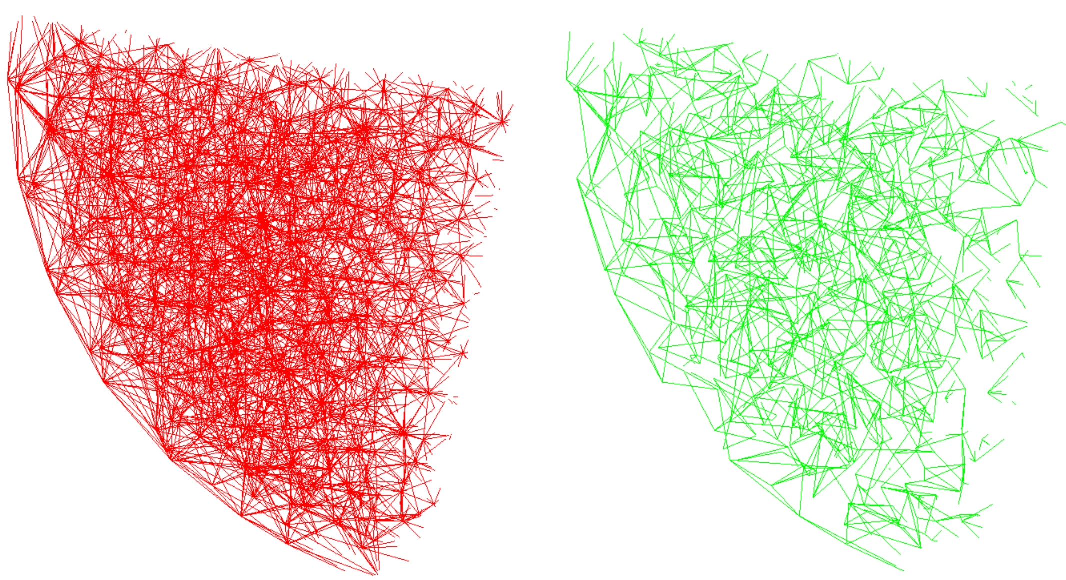

Fig. 4.18. Networks with connectivities 20 and 4

References

1. R. Alicandro and M. Cicalese. A general integral representation result for the continuum limits of discrete energies with superlinear growth. *SIAM J. Math. Anal.*, 36(1):1–37, 2004.
2. J.-F. Babadjian and M. Barchiesi. A variational approach to the local character of G-closure: the convex case. *Ann. I. H. Poincaré – AN*, (26):351–373, 2009.
3. G. Bal. Central limits and homogenization in random media. *Multiscale Model. Simul.*, 7:677–702, 2008.
4. J. M. Ball. Some open problems in elasticity. In *Geometry, mechanics, and dynamics*, pages 3–59. Springer, New York, 2002.
5. J.M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, 63:337–403, 1977.
6. M. Barchiesi. Loss of polyconvexity by homogenization: a new example. *Calc. Var. Partial Differential Equations*, 30(2):215–230, 2007.
7. M. Bellieud and G. Bouchitté. Homogenization of elliptic problems in a fiber reinforced structure. Nonlocal effects. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 26(3):407–436, 1998.
8. A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1978.
9. P. Binev, A. Cohen, W. Dahmen, R. Devore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. Preprint.
10. X. Blanc and C. Le Bris. Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings. *Netw. Heterog. Media*, 5(1):1–29, 2010.
11. X. Blanc, C. Le Bris, and P.L. Lions. Une variante de la théorie de l'homogénéisation stochastique des opérateurs elliptiques. *C. R. Acad. Sci. Paris, Série I*, 343:717–724, 2006.
12. S. Boyaval. Reduced-basis approach for homogenization beyond the periodic setting. *Multiscale Model. Simul.*, 7(1):466–494, 2008.
13. A. Braides. Homogenization of some almost periodic functionals. *Rend. Accad. Naz. Sci. XL*, 103:261–281, 1985.
14. A. Braides. Loss of polyconvexity by homogenization. *Arch. Rat. Mech. Anal.*, 127:183–190, 1994.
15. A. Braides and A. Defranceschi. *Homogenization of Multiple Integrals*, volume 12 of *Oxford Lecture Series in Mathematics and Its Applications*. Oxford University Press, 1998.
16. M. Briane. Homogenization of non-uniformly bounded operators: critical barrier for nonlocal effects. *Arch. Ration. Mech. Anal.*, 164(1):73–101, 2002.
17. A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis. Preprint.
18. 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.
19. Y.-C. Chen. On strong ellipticity and the Legendre-Hadamard condition. *Arch. Ration. Mech. Anal.*, 113:165–175, 1991.
20. C. Choquet and A. Sili. Homogenization of a model of displacement with unbounded viscosity. *Netw. Heterog. Media*, 4(4):649–666, 2009.
21. P.G. Ciarlet. *Mathematical elasticity. Volume I: three-dimensional elasticity*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988.
22. A. Cohen, R. DeVore, and C. Schwab. Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.

23. A. Cohen, R. Devore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Anal. Appl. (Singap.)*, 9(1):11–47, 2011.
24. J.G. Conlon and A. Naddaf. On homogenization of elliptic equations with random coefficients. *Elec. Journal of Probability*, 5:1–58, 2000. Paper no. 9.
25. G. Dal Maso and L. Modica. Integral functionals determined by their minima. *Rend. Semin. Mat. Univ.*, 76:255–267, 1986.
26. G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.
27. A.-L. Dalibard. Homogenization of non-linear scalar conservation laws. *Arch. Ration. Mech. Anal.*, 192(1):117–164, 2009.
28. E. De Giorgi and S. Spagnolo. Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine. *Boll. Un. Mat. Ital. (4)*, 8:391–411, 1973.
29. A. De Masi, P.A. Ferrari, S. Goldstein, and W.D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.*, 55(3/4):787–855, 1988.
30. B. Delaunay. Sur la sphère vide. *Bull. Acad. Sci. URSS*, VII.(6):793–800, 1934.
31. R. L. Dobrushin and S. B. Shlosman. Completely analytical Gibbs fields. In *Statistical physics and dynamical systems (Köszeg, 1984)*, volume 10 of *Progr. Phys.*, pages 371–403. Birkhäuser Boston, Boston, MA, 1985.
32. R. L. Dobrushin and S. B. Shlosman. Constructive criterion for the uniqueness of Gibbs field. In *Statistical physics and dynamical systems (Köszeg, 1984)*, volume 10 of *Progr. Phys.*, pages 347–370. Birkhäuser Boston, Boston, MA, 1985.
33. R. Figari, E. Orlandi, and G. Papanicolaou. Mean field and Gaussian approximation for partial differential equations with random coefficients. *SIAM J. Appl. Math.*, 42:1069–1077, 1982.
34. P.J. Flory. *Statistical mechanics of chain molecules*. Interscience Publishers, New York, 1969.
35. S. Fortune. Voronoĭ diagrams and Delaunay triangulations. In *Computing in Euclidean geometry*, volume 4 of *Lecture Notes Ser. Comput.*, pages 225–265. World Scientific Publishing Co Pte Ltd, Singapore, 1995.
36. FreeFEM. http://www.freefem.org/.
37. G. Friesecke, R.D. James, and S. Müller. A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity. *Comm. Pure Appl. Math.*, 55(11):1461–1506, 2002.
38. G. Geymonat, S. Müller, and N. Triantafyllidis. Homogenization of nonlinearly elastic materials, microscopic bifurcation and macroscopic loss of rank-one convexity. *Arch. Rat. Mech. Anal.*, 122:231–290, 1993.
39. A. Gloria. A direct approach to numerical homogenization in finite elasticity. *Netw. Heterog. Media.*, 1:109–141, 2006.
40. M. Grüter and K.-O. Widman. The Green function for uniformly elliptic equations. *Manuscripta Math.*, 37:303–342, 1982.
41. N. Hansen. The CMA evolution strategy: a comparing review. In J.A. Lozano, P. Larranaga, I. Inza, and E. Bengoetxea, editors, *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pages 75–102. Springer, 2006.
42. T.Y. Hou. Numerical approximations to multiscale solutions in partial differential equations. In Craig A.W. Blowey, J.F. and T. Shardlow, editors, *Frontier in Numerical Analysis*, pages 241–302. Springer Publications, 2003.
43. T. Iwaniec, L. V. Kovalev, and J. Onninen. Diffeomorphic approximation of Sobolev homeomorphisms. *Arch. Rat. Mech. Anal.*, 201:1047–1067, 2011.
44. V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, Berlin, 1994.
45. C. Kipnis and S.R.S. Varadhan. Central limit theorem for additive functional of reversible Markov processes and applications to simple exclusion. *Commun. Math. Phys.*, 104:1–19, 1986.
46. K. Klenke. *Wahrscheinlichkeitstheorie*. Springer-Verlag, Heidelberg-Berlin, 2006. English version appeared as: Probability theory. A comprehensive course, Universitext, Springer-Verlag London, Ltd., London, 2008.
47. S.M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.
48. S.M. Kozlov. Averaging of difference schemes. *Math. USSR Sbornik*, 57(2):351–369, 1987.
49. J. Kristensen. On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16(1):1–13, 1999.
50. W. Kuhn and F. Grün. Beziehung zwischen elastische Konstanten und Dehnungsdoppelberechnung Eigenschaften hochpolymerer Stoffe. *Kolloid-Zeitschrift*, 101:248–271, 1942.
51. R. Künnemann. The diffusion limit for reversible jump processes on $\mathbb{Z}^d$ with ergodic random bond conductivities. *Commun. Math. Phys.*, 90:27–68, 1983.

52. P. Le Tallec. Numerical methods for nonlinear three-dimensional elasticity. In *Handbook of numerical analysis, Vol. III*, pages 465–622. North-Holland, 1994.
53. P. Le Tallec. *Mécanique des Milieux Continus*. Editions de l'École polytechnique, 2006.
54. M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer-Verlag, Berlin, 2001.
55. W. Littman, G. Stampacchia, and H. F. Weinberger. Regular points for elliptic equations with discontinuous coefficients. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:43–77, 1963.
56. K. A. Lurie and A. V. Cherkaev. Exact estimates of conductivity of composites formed by two isotropically conducting media taken in prescribed proportion. *Proc. Roy. Soc. Edinburgh Sect. A*, 99(1-2):71–87, 1984.
57. Y. Maday, A. T. Patera, and G. Turinici. Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations. *C. R. Math. Acad. Sci. Paris*, 335(3):289–294, 2002.
58. P. Marcellini. Periodic solutions and homogenization of nonlinear variational problems. *Annali Mat. Pura Appl.*, 117:139–152, 1978.
59. P. Mathieu. Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.*, 130(5):1025–1046, 2008.
60. R. Meester and R. Roy. *Continuum percolation*, volume 119 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.
61. K. Messaoudi and G. Michaille. Stochastic homogenization of nonconvex integral functionals. *RAIRO Modél. Math. Anal. Numér.*, 28(3):329–356, 1994.
62. N. Meunier, Pantz O., and A. Raoult. Eslastic limit of square lattices with three point interactions. Preprint hal-00589926, 2011.
63. J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(11):294–327, 2011.
64. S. Müller. Homogenization of nonconvex integral functionals and cellular elastic materials. *Arch. Rat. Mech. Anal.*, 99:189–212, 1987.
65. S. Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.
66. S. Müller and S. Neukamm. On the commutability of homogenization and linearization in finite elasticity. *Arch. Rat. Mech. Anal.*, 201:465–500, 2011.
67. F. Murat. H-convergence. Séminaire d'Analyse fonctionnelle et numérique, Univ. Alger, multigraphié, 1978.
68. F. Murat and L. Tartar. H-convergence. In A.V. Cherkaev and R.V. Kohn, editors, *Topics in the Mathematical Modelling of Composites Materials*, volume 31 of *Progress in nonlinear differential equations and their applications*, pages 21–44. Birkhäuser, 1997.
69. G. Nguetseng. Homogenization structures and applications. I. *Z. Anal. Anwendungen*, 22(1):73–107, 2003.
70. G. Nguetseng. Homogenization structures and applications. II. *Z. Anal. Anwendungen*, 23(3):483–508, 2004.
71. J. Nolen and G. Papanicolaou. Fine scale uncertainty in parameter estimation for elliptic equations. *Inv. Problems*, 2010.
72. R. W. Ogden. Large deformation isotropic elasticity: on the correlation of theory and experiment for incompressible rubberlike solids. *Proc. R. Soc. Lond. A*, 326:565–584, 1972.
73. H. Owhadi. Approximation of the effective conductivity of ergodic media by periodization. *Probab. Theory Relat. Fields*, 125:225–258, 2003.
74. G.C. Papanicolaou. Diffusions and random walks in random media. In *The mathematics and physics of disordered media (Minneapolis, Minn., 1983)*, volume 1035 of *Lecture Notes in Math.*, pages 391–399. Springer, Berlin, 1983.
75. 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.
76. M.D. Penrose. Random parking, sequential adsorption, and the jamming limit. *Commun. Math. Phys.*, 218:153–176, 2001.
77. B. Perthame and A. E. Tzavaras. Kinetic formulation for systems of two conservation laws and elastodynamics. *Arch. Ration. Mech. Anal.*, 155(1):1–48, 2000.
78. U. Raitums. On the local representation of G-closure. *Arch. Ration. Mech. Anal.*, 158:213–234, 2001.
79. R. H. Rubinstein, M.and Colby. *Polymer physics*. Oxford University Press, 2003.
80. D. Ruelle. *Statistical mechanics. Rigorous results.* World Scientific Publishing Co., Inc., River Edge, NJ; Imperial College Press, London, 1999. Reprint of the 1989 edition.

81. T. Schreiber, M.D. Penrose, and J.E. Yukich. Gaussian limits for multidimensional random sequential packing at saturation. *Comm. Math. Phys.*, 272(1):167–183, 2007.
82. V. Sidoravicius and A.-S. Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields*, 129(2):219–244, 2004.
83. V. Sverák. New examples of quasiconvex functions. *Arch. Rational Mech. Anal.*, 119:293–300, 1992.
84. V. Sverák. On the problem of two wells. 54:183–189, 1993.
85. L. Tartar. Estimations fines de coefficients homogénéisés. In P. Krée, editor, *Ennio De Giorgi Colloquium*, volume 125 of *Reseach Notes in Mathematics*, pages 168–187, London, 1985. Pitman.
86. L.R.G Treloar. *The Physics of Rubber Elasticity.* Oxford at the Clarendon Press, Oxford, 1949.
87. M. Vidrascu. Solution of non-linear elasticity problems using the *continu* software. *Inria Report Research (http://www.inria.fr/rrrt/rr-4128.html)*, 2001.
88. M. Vogelius. A homogenization result for planar, polygonal networks. *RAIRO Modél. Math. Anal. Numér.*, 25(4):483–514, 1991.
89. W. Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 2000.
90. X. Yue and W. E. The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys.*, 222(2):556–572, 2007.
91. V. V. Yurinskiĭ. Averaging of symmetric diffusion in random medium. *Sibirskii Matematicheskii Zhurnal*, 27(4):167–180, 1986.