UNIVERSITÉ PARIS-EST

T H È S E

présentée pour obtenir le titre de

Docteur de l'Université Paris-Est
Spécialité : Mathématiques

par Ronan COSTAOUEC

Sujet : Techniques numériques d'homogénéisation application aux matériaux aléatoires

Soutenue le 23 Novembre 2011
devant le jury composé de :

Président :	Annie RAOULT
Rapporteurs :	Josselin GARNIER
	Grigoris PAVLIOTIS
Examinateurs :	Eric BONNETIER
	Gilles FRANCFORT
Directeurs :	Claude LE BRIS
	Frédéric LEGOLL

Sujet : Techniques numériques d'homogénéisation : application aux matériaux aléatoires.

Résumé : Le travail de cette thèse a porté sur le développement de techniques numériques pour l'homogénéisation de matériaux présentant à une petite échelle des hétérogénéités aléatoires. Sous certaines hypothèses, la théorie mathématique de l'homogénéisation stochastique permet d'expliciter les propriétés effectives de tels matériaux. Néanmoins, en pratique, la détermination de ces propriétés demeure difficile. En effet, celle-ci requiert la résolution d'équations aux dérivées partielles stochastiques posées sur l'espace tout entier. Dans cette thèse, cette difficulté est abordée de deux manières différentes. Les méthodes classiques d'approximation conduisent à approcher les propriétés effectives par des quantités aléatoires. Réduire la variance de ces quantités est l'objectif des travaux de la Partie I. On montre ainsi comment adapter au cadre de l'homogénéisation stochastique une technique de réduction de variance déjà éprouvée dans d'autres domaines. Les travaux de la Partie II s'intéressent à des cas pour lesquels le matériau d'intérêt est considéré comme une petite perturbation aléatoire d'un matériau de référence. On montre alors *numériquement* et *théoriquement* que cette simplification de la modélisation permet effectivement une réduction très importante du coût calcul.

Title : Numerical methods for homogenization : applications to random media.

Abstract : In this thesis we investigate numerical methods for the homogenization of materials the structures of which, at fine scales, are characterized by random heterogenities. Under appropriate hypotheses, the effective properties of such materials are given by closed formulas. However, in practice the computation of these properties is a difficult task because it involves solving partial differential equations with stochastic coefficients that are additionally posed on the whole space. In this work, we address this difficulty in two different ways. The standard discretization techniques lead to random approximate effective properties. In Part I, we aim at reducing their variance, using a well-known variance reduction technique that has already been used successfully in other domains. The works of Part II focus on the case when the material can be seen as a small random perturbation of a periodic material. We then show both numerically and theoretically that, in this case, computing the effective properties is much less costly than in the general case.

Sommaire

1 Introduction générale **3**
1.1 Situation du problème . . . 5
1.1.1 Le problème initial . . . 5
1.1.2 La théorie de l'homogénéisation . . . 6
1.1.3 **Cadre de la thèse** . . . 11
1.2 Eléments de théorie dans des cas classiques . . . 12
1.2.1 Cas périodique . . . 12
1.2.2 Cas quasi-périodique . . . 14
1.2.3 Cas stochastique classique . . . 15
1.3 Approximations de la matrice homogénéisée dans le cas stochastique . 17
1.3.1 Approximations classiques . . . 17
1.3.2 **Les contributions de la Partie I** . . . 18
1.4 Problèmes faiblement stochastiques . . . 21
1.4.1 Difféomorphisme aléatoire . . . 22
1.4.2 Cadres perturbatifs . . . 23
1.4.3 **Les contributions de la Partie II** . . . 25
1.5 Perspectives . . . 26
1.5.1 Réduction de variance . . . 27
1.5.2 Cadre perturbatif . . . 33
1.5.3 Perspectives plus générales . . . 35

2 Homogénéisation stochastique numérique : quelques contributions récentes **37**
2.1 Introduction . . . 39
2.2 Basics of stochastic homogenization . . . 40
2.2.1 General and Periodic homogenization . . . 40
2.2.2 Basic theory of stochastic homogenization . . . 46
2.2.3 Bounds and numerical approach . . . 51
2.3 Perturbation of periodic settings and related approaches . . . 58
2.3.1 Perturbation theory . . . 58
2.3.2 A variant of the general setting . . . 73
2.3.3 MsFEM-type approach . . . 80
2.4 Variance issues . . . 85

2.4.1 One-dimensional situation 87
2.4.2 Elements of proof for two simple situations 89
2.4.3 Numerical tests 97
2.5 Trends 101

I Réduction de variance pour l'homogénéisation stochastique 103

3 Etude d'un cas simple 105
3.1 Introduction 107
3.2 Stochastic homogenization theory 108
3.3 Variance reduction 110
3.3.1 Classical Monte Carlo method 110
3.3.2 Antithetic variable for stochastic homogenization 113
3.4 Numerical experiments 116
3.5 Variance reduction for the solution $u^\star$ 119

4 Tests numériques 125
4.1 Introduction 127
4.1.1 Homogenization theoretical setting 128
4.1.2 Numerical approach 129
4.1.3 The technique of antithetic variables 131
4.1.4 A brief summary of our former results 132
4.2 Variance reduction for problems involving correlations or anisotropy . 133
4.2.1 Correlated cases 134
4.2.2 Anisotropic cases 140
4.3 Variance reduction for eigenproblems 142

5 Eléments d'analyse 149
5.1 Introduction 151
5.1.1 Homogenization theoretical setting 152
5.1.2 The questions we consider 154
5.2 A "fully" stochastic case 156
5.2.1 Mathematical setting and statement of our main result 156
5.2.2 Classical results on antithetic variables 162
5.2.3 Monotonicity in periodic homogenization 164
5.2.4 Proof of Proposition 5.2.1 and Corollaries 5.2.1 and 5.2.2 . . . 166
5.3 A "weakly" stochastic case 168
5.3.1 Main result 169
5.3.2 Decomposition of ∇w_p^1 172
5.3.3 Variance of $A_{1,N}^{\star,\mathrm{exact}}$ 177
5.3.4 Extensions 182

II Problèmes faiblement stochastiques 185

6 Approximation numérique d'une classe de problèmes en homogénéisation stochastique 187
6.1 Introduction (in english) . . . 189
6.2 Introduction . . . 190
6.3 Approximation numérique . . . 192
6.4 Résultats numériques . . . 194

7 Estimation de l'erreur dans l'approximation numérique de deux modèles faiblement stochastiques 197
7.1 Introduction . . . 199
7.1.1 Probabilistic setting . . . 201
7.1.2 Stochastic homogenization results . . . 202
7.1.3 Standard numerical approximation . . . 203
7.2 Expansion of the random matrix A_η (Model 1) . . . 204
7.2.1 Assumptions . . . 205
7.2.2 Formal expansion . . . 206
7.2.3 Main result . . . 206
7.2.4 Convergence with respect to h and N . . . 213
7.3 Expansion of stochastic diffeomorphism (Model 2) . . . 215
7.3.1 Hypotheses . . . 216
7.3.2 Formal expansion . . . 216
7.3.3 Main result . . . 217

Bibliographie générale 227

Publications dans des revues à comité de lecture

[P1] R. Costaouec, C. Le Bris et F. Legoll, Approximation numérique d'une classe de problèmes en homogénéisation stochastique (Numerical approximation of a class of problems in stochastic homogenization), *C. R. Acad. Sci. Paris*, Série I, vol. 348 (1-2), 99-103 (2010).

[P2] R. Costaouec, C. Le Bris et F. Legoll, Variance reduction in stochastic homogenization : proof of concept, using antithetic variables, *Boletin Soc. Esp. Mat. Apl.*, vol. 50, 9-27 (2010).

[P3] X. Blanc, R. Costaouec, C. Le Bris et F. Legoll, Variance reduction in stochastic homogenization using antithetic variables, *Markov Processes and Related fields*, à paraître.

[P4] X. Blanc, R. Costaouec, C. Le Bris et F. Legoll, Variance reduction in stochastic homogenization : the technique of antithetic variables, dans Numerical Analysis and Multiscale Computations, *Lect. Notes Comput. Sci. Eng.* , vol. 82, Springer, à paraître.

[P5] A. Anantharaman, R. Costaouec, C. Le Bris, F. Legoll et F. Thomines, Introduction to numerical stochastic homogenization and the related computational challenges : some recent developments, *Multiscale modeling and analysis for materials simulation*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, vol. 22.

[P6] R. Costaouec, Asymptotic expansion of the homogenized matrix in two weakly stochastic homogenization settings, *Applied Mathematics Research eXpress*, à paraître (2011).

Chapitre 1

Introduction générale

Cette thèse a pour objet la proposition et l'analyse de techniques numériques d'homogénéisation avec en vue l'étude du comportement des matériaux aléatoires.

Les matériaux, ou plus généralement les milieux, qui nous intéressent dans cette thèse possèdent deux propriétés caractéristiques : ils sont *multi-échelles* et *aléatoires*. Multi-échelles, tout d'abord, car ils présentent certaines hétérogénéités à une échelle d'espace, dite *microscopique*, différente de celle, dite *macroscopique*, à laquelle nous souhaitons étudier ou simuler leurs comportements. Aléatoires, enfin, car leurs structures, bien qu'elles témoignent d'une certaine régularité, demeurent en partie mal connues. Une grande variété de matériaux réels relèvent de cette double étiquette.

Le fait pour un matériau d'être à la fois multi-échelle et aléatoire est source de nombreuses difficultés, dès lors que l'on souhaite simuler numériquement son comportement. En effet, la prise en compte de la microstructure requiert, dans le cadre des méthodes classiques de simulation numérique, de type éléments ou volumes finis, l'utilisation d'un maillage adapté à l'échelle des singularités rencontrées. La conséquence est, à système macroscopique de dimensions physiques constantes, une augmentation prohibitive du nombre de degrés de liberté et donc un accroissement du coût de calcul. Par ailleurs, l'intégration du caractère aléatoire du matériau nécessite la considération de nombreuses configurations possibles de ce dernier. A l'aide de celles-ci on peut estimer le comportement moyen du matériau avec une marge d'erreur qui dépend de la quantité d'information disponible. La conséquence est la même que dans le cas précédent : une augmentation déraisonnable du coût calcul. Atténuer ces phénomènes, tout en tenant compte de la nature multi-échelle et aléatoire de ces matériaux, est la démarche au sein de laquelle s'inscrit cette thèse.

Simplifier la simulation du comportement des matériaux composites fut, au cours des décennies précédentes, la source de nombreux travaux universitaires tant dans le domaine des mathématiques que dans celui des sciences physiques. La

théorie de l'*homogénéisation* s'inscrit dans cette famille. Son développement permit, entre autres, de fournir un cadre rigoureux à l'étude des matériaux composites. En termes physiques, on peut affirmer que cette théorie permet de clarifier la *notion de matériau limite*, ou *équivalent*. Elle précise sous quelles hypothèses générales un tel matériau existe. Mais elle fait plus : sous certaines hypothèses relatives à la structure du matériau considéré, elle établit de manière rigoureuse une caractérisation explicite de ce matériau limite, équivalent uni-échelle du matériau de référence. En cela réside l'intérêt majeur et pratique de cette théorie : le matériau limite est l'équivalent simple du matériau initial dans la limite où le rapport entre échelles microscopique et macroscopique, noté ε, tend vers 0. Il porte la trace des microstructures caractérisant le matériau initial, mais ne fait plus intervenir les petites échelles qui compliquaient singulièrement la simulation. Il permet donc de calculer à moindre coût une *approximation* simple, d'autant plus précise que ε est petit, du comportement du matériau d'origine.

La caractérisation du matériau limite n'est pas gratuite. Elle fait dans de nombreux cas intervenir des problèmes auxiliaires qui dépendent du type de microstructure considéré et du contexte physique dans lequel est étudié le matériau. L'efficacité pratique de l'homogénéisation, sa capacité à réduire le coût de simulation, est liée au coût calcul associé à la résolution de ces problèmes auxiliaires. Or, dans le cas de matériaux possédant une *microstructure aléatoire* générale, au sens où ils relèvent du cadre classique de l'homogénéisation stochastique, ce coût est infini. En effet les problèmes auxiliaires sont dans ce cas à la fois stochastiques et posés sur l'espace tout entier. En pratique, on ne peut donc pas calculer les propriétés *exactes* du matériau effectif, et il faut mettre au point des techniques qui permettent de les approximer efficacement. A une première approximation, que constitue en soi l'homogénéisation et qui consiste à considérer que ε tend vers 0, vient donc s'ajouter une seconde approximation, celle des propriétés du matériau *limite* ou *équivalent*.

Ceci motive l'étude de techniques numériques d'homogénéisation dans le cas des matériaux aléatoires. *Le but de cette thèse est précisément l'élaboration et l'analyse de telles techniques.* Dans la suite de cette introduction, nous adoptons le plan suivant. Dans la section 1.1, nous présentons les principes généraux de la théorie de l'homogénéisation et des méthodes numériques qui en découlent plus ou moins directement. La section 1.2 rappelle les éléments de théorie standard qui permettent de caractériser, sous certaines hypothèses précises, les propriétés du matériau limite. Lorsque le matériau initial possède une microstructure aléatoire, cette caractérisation est délicate et nécessite la mise en oeuvre de méthodes d'approximation. La présentation de ces méthodes est l'objet de la section 1.3.1. Celle-ci permet d'introduire la section 1.3.2 qui traite des contributions regroupées au sein de la **Partie I**, relatives à la réduction de variance pour l'homogénéisation stochastique. La section 1.4 contient une présentation des contributions de la **Partie II** ayant toutes trait à l'homogénéisation de problèmes faiblement stochastiques. Enfin, la

section 1.5 présente différentes perspectives ouvertes par les travaux de cette thèse, ainsi que de quelques directions de recherche plus générales.

1.1 Situation du problème

1.1.1 Le problème initial

Dans le cadre de cette thèse, nous nous intéresserons exclusivement à des équations aux dérivées partielles de la forme

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A^\varepsilon \nabla u^\varepsilon\right) &=& f \quad \text{dans} \quad \mathcal{D}, \\ u^\varepsilon &=& 0 \quad \text{sur} \quad \partial\mathcal{D}, \end{array} \right. \tag{1.1}$$

où $\mathcal{D}$ désigne un ouvert borné de $\mathbb{R}^d$, $A^\varepsilon \in \left(L^\infty(\mathcal{D})\right)^{d\times d}$, et $f \in H^{-1}(\mathcal{D})$. La matrice A^ε (de même que son inverse) sera toujours supposée bornée et strictement positive uniformément en ε :

$$\left\{ \begin{array}{l} \exists \alpha > 0,\ \forall \xi \in \mathbb{R}^d,\ \forall \varepsilon > 0,\ \xi^T A^\varepsilon(x)\xi \geq \alpha|\xi|^2 \text{ presque partout}, \\ \exists \beta > 0,\ \forall \xi \in \mathbb{R}^d,\ \forall \varepsilon > 0,\ \xi^T \left(A^\varepsilon(x)\right)^{-1} \xi \geq \beta|\xi|^2 \text{ presque partout}. \end{array} \right.$$

On note en outre $\mathcal{M}(\alpha,\beta)$ l'ensemble des fonctions à valeurs matricielles vérifiant ces conditions.
La matrice A^ε modélise, par exemple, la conductivité thermique d'un matériau. La température u^ε dans le matériau est alors donnée par (1.1), où f représente les sources de chaleur. La conductivité est indexée par le paramètre ε qui représente l'échelle à laquelle le matériau possède des hétérogénéités. Notons qu'en pratique les déterminations de ε et A^ε sont loin d'être simples.

La question à laquelle nous allons à présent tenter de répondre est la suivante : comment calculer u^ε en pratique ? Afin de mesurer la difficulté, supposons par exemple que $\mathcal{D} = (0,1)^2$ et $\varepsilon = 10^{-6}$. Supposons également que nous ayons recours à une méthode d'approximation numérique classique, de type éléments finis par exemple, afin d'approcher u^ε. Il semble raisonnable *a priori* d'utiliser un maillage au moins aussi fin que ε, si nous voulons prendre en compte l'effet des hétérogénéités de A^ε à l'échelle ε. On commet d'ailleurs une erreur dans le cas contraire (cf [62]). Ici, dans le cas d'un maillage régulier et d'éléments finis $\mathbb{P}_1$ de Lagrange, le nombre de degrés de liberté N_{deg} résultant est de l'ordre de 10^{12} ! Il faut trouver une autre manière de procéder.

1.1.2 La théorie de l'homogénéisation

1.1.2.1 Principe

Pour construire une stratégie numérique permettant de traiter (1.1), une façon de faire est d'utiliser la théorie de l'homogénéisation. Elle consiste à considérer la limite de l'équation (1.1) lorsque ε tend vers 0. On verra plus loin que cette limite s'écrit sous la forme

$$\left\{ \begin{array}{rcl} -\mathrm{div}\,(A^\star \nabla u^\star) & = & f \quad \text{dans} \quad \mathcal{D}, \\ u^\star & = & 0 \quad \text{sur} \quad \partial\mathcal{D}, \end{array} \right. \tag{1.2}$$

où $A^\star$ désigne la matrice, dite homogénéisée, reflétant les propriétés du *matériau limite ou effectif*, et $u^\star$ la limite, en un sens à définir, de la suite u^ε. Si le passage de (1.1) à une équation de la forme (1.2) peut sembler relativement naturel, en fait rien n'indique, en général, que la limite $\mathcal{L}^\star$ au sens de l'homogénéisation d'une suite d'opérateurs $\mathcal{L}^\varepsilon \in L$ appartienne à la même "classe" L d'opérateurs. Dans le cas de l'équation (1.1), cette affirmation fait en un sens partie du résultat et il existe des cas où ceci est tout simplement faux (voir [51, 62]).
L'évènement fondamental dans le passage de (1.1) à (1.2) est la disparition de ε, la petite échelle. Si $A^\star$ est connu, nous sommes en mesure de calculer à moindre coût $u^\star$, via une méthode classique de discrétisation, et ainsi d'obtenir une approximation de u^ε pour ε petit.

1.1.2.2 La théorie mathématique

Historique La théorie mathématique de l'homogénéisation est la justification rigoureuse du passage de (1.1) à (1.2). Reprenant ici la distinction proposée dans la première monographie synthétique sur le sujet [16], on peut considérer qu'elle repose sur quatre types principaux d'approches pour homogénéiser le problème (1.1)

- *les méthodes dites d'énergie* ou *des fonctions de test oscillantes*, développées par F. Murat et L. Tartar, qui, historiquement, ont fourni les premières un cadre général et rigoureux pour le passage de (1.1) à (1.2) ;
- *la méthode dite de la convergence à deux échelles*, qui permet d'obtenir le problème limite sous des hypothèses particulières sur A^ε ;
- *les méthodes probabilistes* qui reposent sur le lien entre l'équation (1.1) et un processus aléatoire sous-jacent de diffusion ;
- *les méthodes de décomposition*, plutôt utilisées pour des problèmes dépendant du temps, et qui exploitent dans des contextes particuliers la possibilité de décomposer u^ε en ondes de Bloch.

Cette liste ne prétend pas être exhaustive et il existe d'autres méthodes (la méthode dite d'éclatement par exemple)

On donne maintenant un bref aperçu du développement des différentes méthodes, suivant ici la présentation de [98]. La formalisation mathématique précise du passage de (1.1) à (1.2) pour des coefficients généraux repose historiquement sur les *méthodes dites d'énergie* introduites par F. Murat et L. Tartar dans les années 70. Dans [74] est introduite la notion de H-convergence. Le développement du principe dit de compacité par compensation [75] aboutit ensuite à ce qu'on nomme à présent indistinctement méthode de la fonction test oscillante, de l'énergie ou encore de dualité (voir [94]). Ces travaux conduisent plus tard à la théorie des H-mesures (voir [97] par exemple). Ces outils permettent, en particulier, de traiter le cas où la matrice A^ε n'est *pas* symétrique. Il convient de mentionner que ces travaux ont été menés en parallèle avec ceux de S. Spagnolo qui dès 1968 expose dans [92], une théorie de la G-convergence pour les opérateurs *symétriques* dans les cas elliptiques et paraboliques. Cette notion tient son nom du fait qu'elle traduit la convergence des noyaux de Green des opérateurs $\mathcal{L}^\varepsilon := -\mathrm{div}(A^\varepsilon \nabla)$ associés aux matrices A^ε. Elle même est très étroitement liée à la notion de Γ-convergence initiée par E. De Giorgi et S. Spagnolo (voir [32] pour l'article originel).

Puisque la notion de H-convergence est la plus générale et coïncide avec celle de G-convergence dans le cas d'opérateurs symétriques, c'est d'elle dont nous allons donner une caractérisation. On dit que A^ε H-converge vers $A^\star$ si et seulement si, pour tout $f \in H^{-1}(\mathcal{D})$, on a les convergences faibles

$$\begin{aligned} u^\varepsilon &\rightharpoonup u^\star \text{ dans } H_0^1(\mathcal{D}), \\ A^\varepsilon \nabla u^\varepsilon &\rightharpoonup A^\star \nabla u^\star \text{ dans } \left(L^2(\mathcal{D})\right)^d, \end{aligned}$$

où $u^\star$ désigne la solution du problème (1.2). Deux constats s'imposent. Tout d'abord, on note que la définition précédente fait intervenir les solutions du problème (1.1), et donc *dépend de l'opérateur* $\mathcal{L}^\varepsilon$. Enfin, par définition la matrice homogénéisée $A^\star$ est indépendante du second membre f. Ainsi dès qu'on connaît $A^\star$, on peut résoudre le problème (1.2) pour différentes valeurs de ce second membre. Ceci étant dit, le résultat fondamental de l'homogénéisation peut s'exprimer de la manière suivante : $\mathcal{M}(\alpha, \beta)$ *est relativement compact pour la topologie de la H-convergence.* En d'autre termes, de toute suite de matrices $A^\varepsilon \in \mathcal{M}(\alpha, \beta)$, on peut extraire une sous-suite $A^{\varepsilon'}$ qui H-converge vers une matrice $A^\star \in \mathcal{M}(\alpha, \beta)$.

Comme nous le mentionnions plus haut, dans le cas où A^ε est symétrique, la H-convergence coïncide avec la G-convergence, elle même liée à la notion de Γ-convergence. Cette dernière diffère en fait foncièrement des notions précédentes, en particulier du fait qu'elle ne s'applique pas aux opérateurs différentiels mais aux fonctionnelles dont ils dérivent via le calcul des variations. Ainsi, à l'équation (1.1) (et sous l'hypothèse que A^ε est symétrique) est naturellement associée une fonctionnelle $J^\varepsilon : H_0^1(\mathcal{D}) \to \mathbb{R}$ définie par

$$J_\varepsilon(v) = \frac{1}{2} \int_{\mathcal{D}} A^\varepsilon \nabla v \cdot \nabla v - \int_{\mathcal{D}} f v.$$

Si J^ε Γ-converge vers une fonctionnelle $J^\star$, alors les minimiseurs de J^ε (u^ε ici) convergent vers ceux de $J^\star$ ($u^\star$ ici) et

$$\inf_{v\in H_0^1(\mathcal{D})} J^\varepsilon(v) \longrightarrow_{\varepsilon\to 0} \inf_{v\in H_0^1(\mathcal{D})} J^\star(v).$$

Les deux formalismes n'ont pas été développés dans le même contexte mais coïncident dans le cadre du problème (1.1). On trouvera dans [24, 30] des introductions détaillées à cette notion.

Quelques travaux récents Les outils mathématiques de la théorie de l'homogénéisation ont été initialement développés dans le cadre de l'étude de l'équation (1.1), et d'équations de la mécanique du type

$$\left\{\begin{array}{rcl} -\mathrm{div}\,(A^\varepsilon : \nabla u^\varepsilon) &=& f \quad \text{dans} \quad \mathcal{D}, \\ u^\varepsilon &=& 0 \quad \text{sur} \quad \partial\mathcal{D}, \end{array}\right.$$

où A^ε désigne à présent un tenseur d'ordre 2 et $u^\varepsilon \in (H_0^1(\mathcal{D}))^d$. Néanmoins, ils ont été depuis utilisés dans d'autres contextes et adaptés à une grande variété de problèmes. Pour se faire une idée de la variété d'équations auxquelles s'est intéressée la théorie, on pourra consulter les monographies [16, 27, 58, 85] ainsi que leurs bibliographies. Parmi les travaux théoriques les plus récents, nous pouvons tout d'abord citer ceux de P. Souganidis et P.-L. Lions relatifs à l'homogénéisation des équations d'Hamilton-Jacobi dans le cas aléatoire [67, 90], ainsi que ceux concernant l'homogénéisation des problèmes du second ordre complètement non-linéaires dans les cas aléatoire uniformément elliptique [26] et périodique dégénéré [91]. L'homogénéisation des équations de transport dans les cas déterministes et stochastiques a également fait objet de nombreuses publications (voir par exemple [28, 29]). Dans une autre direction, citons également les travaux récents de G. Francfort, F. Murat et L. Tartar sur l'homogénéisation des opérateurs monotones multivalués sous forme divergence [43]. Enfin, on peut aussi évoquer les travaux de J.-P. Fouque, J. Garnier, G. Papanicolaou, et K. Sølna en lien avec l'équation des ondes [42]. Cette liste est bien entendu très loin d'être exhaustive, mais donne un bref aperçu du dynamisme de la recherche dans ce domaine.

1.1.2.3 Le problème du calcul de la matrice homogénéisée $A^\star$.

Les résultats théoriques mentionnés ci-dessus (concernant le passage de (1.1) à (1.2)) nous assurent de l'existence d'une matrice $A^\star$ reflétant les propriétés *effectives* d'un matériau associé à une matrice A^ε très générale. On peut démontrer à partir de cette seule caractérisation *abstraite*, que certaines propriétés de A^ε sont conservées par la H-convergence. Ainsi dans le chapitre 4 de [58], il est montré que la

H-convergence préserve la similarité et l'ordre des matrices symétriques. Mais pour obtenir des informations plus précises, il faut faire des hypothèses supplémentaires sur la forme de la matrice A^ε. Ces informations peuvent être de natures différentes.

Une première famille de travaux cherche à obtenir une définition explicite de la matrice $A^\star$ que ne fournissent pas les résultats généraux énoncés ci-dessus. En effet, dans le cas général, on sait seulement que, pour tout $p \in \mathbb{R}^d$, $A^\star p$ est la limite faible, quand $\varepsilon \to 0$, de la quantité $A^\varepsilon(p + \nabla w^\varepsilon)$ où w^ε est telle que $-\mathrm{div}\,(A^\varepsilon(p + \nabla w^\varepsilon)) \to 0$ dans $H^{-1}(\mathcal{D})$ quand $\varepsilon \to 0$ et w^ε converge faiblement dans $H^1(\mathcal{D})$. La détermination de la classe des matrices A^ε pour lesquelles on dispose d'une caractérisation explicite de la matrice homogénéisée $A^\star$ demeure une question ouverte. En pratique, on se restreint donc à des cas particuliers pour lesquels on dispose d'une caractérisation constructive de $A^\star$.

Une seconde famille de travaux est consacrée à la détermination de propriétés qualitatives de la matrice homogénéisée sous des hypothèses particulières. Lorsque A^ε est symétrique, l'établissement de *bornes* au sens des matrices symétriques est un exemple de telles propriétés. Les bornes obtenues sont très utilisées par les mécaniciens. Les exemples les plus connus sont les bornes de Voigt-Reuss et Hashin-Shtrikman. On renvoie à [P5] pour une introduction sur le sujet. Une question attenante, qui a également donné lieu à un nombre important de travaux est la suivante : étant donnée une classe de fonctions A^ε possédant certaines propriétés, quel est l'ensemble des matériaux limite atteignables et quelles structures réalisent les optimums ? Cette problématique est mieux connue sous l'intitulé anglo-saxon de *G-closure*.

Le *corpus* de travaux qui nous concerne ici plus particulièrement est celui consacré aux contextes dans lesquels on possède une définition explicite de $A^\star$. Nous allons dans les sections suivantes en évoquer plusieurs : cadre périodique, quasi-périodique, stationnaire ergodique. Leur commun dénominateur est de requérir des hypothèses structurelles sur A^ε. Dans les faits, on suppose que $A^\varepsilon = A(x/\varepsilon)$ où A est supposée, selon le contexte, périodique, quasi-périodique ou stochastique. Cette hypothèse, dite de *séparation des échelles*, conduit à des définitions de la matrice homogénéisée de la forme : pour tout $p \in \mathbb{R}^d$,

$$A^\star p = \int_{\mathcal{I}} A(p + \nabla w_p), \tag{1.3}$$

où la fonction w_p est une solution du problème auxiliaire suivant, aussi appelé *problème des correcteurs*,

$$\begin{cases} -\mathrm{div}\,(A\,(p + \nabla w_p)) = 0 \text{ sur } \mathcal{E}, \\ \text{conditions aux limites.} \end{cases} \tag{1.4}$$

Dans (1.3) et (1.4), $\mathcal{I}$ et $\mathcal{E}$ désignent deux sous-ensembles de $\mathbb{R}^d$. A chacun des

cadres sus-cités pour l'homogénéisation correspondent des choix précis de ces sous-ensembles ainsi que des conditions limites intervenant dans l'équation (1.4). On peut dès lors considérer que l'on dispose bien dans chaque cas d'une caractérisation en un sens explicite de la matrice homogénéisée. Mais nous verrons plus loin que dans les cadres quasi-périodique, et plus particulièrement stochastique, le terme "explicite" est équivoque. En effet, dans ces cadres particuliers, $\mathcal{E}$ n'est pas borné. On dispose certes de formules pour $A^\star$ mais cette matrice n'est pas vraiment explicite : il est en effet impossible de résoudre *directement*, via des méthodes numériques standards, les problèmes auxiliaires.

1.1.2.4 Des approches apparentées pour le problème initial

Avant d'aborder en détail les éléments de théorie et les approches numériques qui nous concernent directement dans ce travail, il faut noter que, dans le cadre de la pratique numérique, la voie de l'homogénéisation n'est pas la seule méthode pour approcher u^ε. En effet, il existe plusieurs méthodes numériques ne reposant pas directement sur le calcul de $A^\star$. Elles permettent soit de traiter des cas non couverts par la théorie de l'homogénéisation (lorsque l'existence d'un problème limite quand $\varepsilon \to 0$ n'est pas assurée), soit de s'intéresser à la résolution de (1.1) pour des valeurs de ε qui ne relèvent pas du régime asymptotique $\varepsilon \to 0$ de l'homogénéisation. Néanmoins dans leurs définitions même, ou dès lors qu'il s'agit d'établir des preuves de convergence, elles se réfèrent le plus souvent à cette théorie. On décrit ci-dessous quelques exemples représentatifs.

Les méthodes *Multi Scale Finite Elements* (MsFEM) ont été développées par X.-W. Hu, T. Y. Hou et Y. Efendiev. Initialement proposée dans [55], la méthode fit ensuite l'objet de nombreuses extensions et améliorations, par exemple dans [36,37]. On renvoie à [38] pour une présentation exhaustive des différentes variantes. Le principe général de l'approche peut s'énoncer de la manière suivante. On cherche à approximer directement la fonction u^ε sans hypothèses structurelles sur la matrice A^ε. La seule information dont on dispose, c'est un ordre de grandeur de l'échelle ε à laquelle le matériau considéré présente ses plus petites hétérogénéités. On maille ensuite grossièrement le domaine $\mathcal{D}$ à l'échelle $H \gg \varepsilon$. Puis, au lieu de considérer les fonctions de base classiques définies via des degrés de libertés associés à ce maillage macroscopique (par exemple des fonctions $\mathbb{P}_1$), on définit des fonctions de bases macroscopiques ϕ^ε oscillantes (voir section 2.3.3.1 pour plus de détails), également associées aux degrés de liberté du maillage macroscopique. Ces fonctions sont définies comme les solutions sur chaque élément du maillage macroscopique de problèmes aux limites liés à la théorie de l'homogénéisation. Ces problèmes sont très semblables au problème (1.4) dit des correcteurs, intervenant dans la définition de $A^\star$.
Qu'a-t-on gagné au cours d'une telle procédure ? A maillage macroscopique donné,

la matrice de rigidité à "inverser" *in fine* est de même dimension que celle à laquelle nous aurions eu affaire dans le cadre d'une méthode classique d'approximation ne tenant pas compte de l'échelle ε, mais la précision est bien meilleure. La méthode requiert la résolution de problèmes à l'échelle microscopique. Mais chacun de ces problèmes est posé sur un domaine petit comparé à $\mathcal{D}$. Indépendants, ces problèmes peuvent de plus être traités *en parallèle* et en pré-calcul par rapport au second membre f.

L'approche *Heterogeneous Multiscale Methods* (HMM) est quelque peu différente. Le terme apparaît pour la première fois dans [33]. Son application dans le cadre de l'homogénéisation est détaillée dans [34]. Dans ce contexte, l'objectif initial est d'approximer $u^\star$ et non u^ε. A cet effet, ces méthodes se donnent un maillage macroscopique avec des éléments de taille H. Les fonctions de base associées aux degrés de liberté macroscopiques sont les fonctions classiques. La prise en compte des variations à l'échelle ε s'effectue lors de l'assemblage de la matrice de rigidité. En effet, au lieu de considérer aux points d'intégration, notés ici (x_i), la valeur $A^\varepsilon(x_i)$, on lui substitue une autre quantité qui tient moralement le rôle de $A^\star(x_i)$. Cette valeur est elle même définie via un problème, inspiré à nouveau du problème des correcteurs, posé sur un domaine centré autour du point x_i et maillé à l'échelle $h \ll \varepsilon$. Le système matriciel associé au problème macroscopique est donc une nouvelle fois de dimension raisonnable. Les problèmes microscopiques le sont également, et peuvent à nouveau être traités en parallèle. Pour un exposé détaillé et systématique de ces méthodes, le lecteur pourra se référer à [35].

De nombreuses autres méthodes numériques ont été proposées afin d'approximer la solution u^ε du problème (1.1) à moindre coût. On peut citer les récents travaux de H. Owhadi en collaboration avec L. Berlyand dans [17] puis avec L. Wang dans [82], basés sur la construction de bases locales d'approximation. Dans [81], J.Nolen, G. Papanicolaou et O. Pironneau ont proposé des méthodes multi-échelles adaptatives basées sur une méthode de projection. Des bases d'ondelettes ont été utilisées par B. Engquist [39]. Dans un cadre stochastique, M. Jardak et R. G. Ghanem ont proposé dans [44] une méthode fondée sur la décomposition spectrale de A^ε.

1.1.3 Cadre de la thèse

Maintenant que nous avons fourni un rapide panorama de la théorie de l'homogénéisation ainsi qu'un aperçu des principales méthodes numériques visant à la résolution du problème (1.1), nous sommes en mesure de situer plus précisément le cadre du présent travail. Les contributions de cette thèse s'inscrivent exclusivement dans la voie de l'homogénéisation aléatoire. Autrement dit, notre objectif est de faciliter, dans le cadre de microstructures aléatoires, l'estimation de $A^\star$ et par conséquent celle de $u^\star$. Dans toute la suite, nous ferons des hypothèses structurelles sur A^ε (par exemple hypothèses de stationarité) permettant d'obtenir une

caractérisation précise de la matrice $A^\star$. Nous présentons dans la suite de ce chapitre les bases de théorie de l'homogénéisation qui permettent d'arriver à de telles caractérisations dans les cas classiques. Enfin nous montrons en quoi la mise en oeuvre pratique de ces caractérisations conduit à des difficultés. Pour faire face à ces difficultés nous présenterons deux manières possibles de procéder, associées aux deux parties principales de cette thèse : traiter le problème dans le cadre général en ayant recours à des outils dédiés (de type réduction de variance) au genre de difficultés rencontrées (**Partie I**), ou restreindre la classe des matériaux considérés, via des hypothèses additionnelles. Celles-ci postulent que le matériau étudié peut être considéré comme une petite perturbation aléatoire d'un matériau de référence dont on sait calculer simplement les propriétés limites. Elles permettent de se placer dans un cadre où la pratique numérique, simplifiée *de facto* peut être rigoureusement analysée (**Partie II**).
Nous nous restreignons ici à l'étude de l'équation (1.1). Il est cependant probable, bien que nous ne l'ayons pas vérifié, que ces deux approches puissent s'étendre à d'autres problèmes.

1.2 Eléments de théorie dans des cas classiques

Sous des hypothèses structurelles sur la matrice A^ε dans (1.1) on peut obtenir des caractérisations explicites de $A^\star$. Nous rappelons dans cette section ces résultats très classiques d'homogénéisation.

1.2.1 Cas périodique

1.2.1.1 Définition de $A^\star$

Le cas le plus simple est celui où la matrice A^ε intervenant dans (1.1) s'écrit sous la forme

$$A^\varepsilon(x) = A_{\text{per}}\left(\frac{x}{\varepsilon}\right),$$

où A_{per} désigne une fonction $\mathbb{Z}^d$-périodique. On dispose alors d'une expression explicite de la matrice homogénéisée qu'on note ici $A^\star_{\text{per}}$. En effet, on a :

$$\left[A^\star_{\text{per}}\right]_{ij} = \int_Q \left(e_i + \nabla w_{e_i}\right)^T A_{\text{per}} \left(e_j + \nabla w_{e_j}\right), \tag{1.5}$$

où $Q = (0,1)^d$, w_p est, pour tout $p \in \mathbb{R}^d$, la solution, définie à une constante près, de l'équation

$$\begin{cases} -\text{div}\left(A_{\text{per}}\left(p + \nabla w_p\right)\right) = 0 \text{ sur } \mathbb{R}^d, \\ w_p \quad Q-\text{périodique.} \end{cases} \tag{1.6}$$

Ce problème dit du correcteur ne fait plus intervenir la variable ε et l'on peut donc approximer sa solution en utilisant par exemple une méthode d'éléments finis

associés à un maillage de finesse raisonnable de la cellule de périodicité Q. On détermine à partir de là une approximation de $A^\star$.

Afin de dériver cette expression de $A^\star$, nous allons utiliser la technique dite du développement à deux échelles. Elle permet d'obtenir formellement et de manière simple le résultat. Nous rappelons ci-dessous les grandes lignes de cette dérivation. Pour une justification rigoureuse de ces résultats, il faut avoir recours à la méthode de la convergence à deux échelles développée par G. Allaire [2] et G. Nguetseng [77].

1.2.1.2 Développement à deux échelles

On commence par supposer que u^ε admet le développement suivant, dit à deux échelles, en ε :

$$u^\varepsilon(x) = u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(x, \frac{x}{\varepsilon}\right) + \cdots, \tag{1.7}$$

où, pour tout $k \in \mathbb{N}$, la fonction u_k dépend de deux variables : une variable macroscopique x, et une variable microscopique $\dfrac{x}{\varepsilon}$. On suppose en outre que chaque u_k est $[0,1]^d$-périodique en la variable $y = \frac{x}{\varepsilon}$, et on insère le développement (1.7) dans l'équation (1.1). On obtient dès lors une cascade d'équations, chacune étant associée à un ordre en ε :

$$\begin{aligned}
& -\mathrm{div}\left(A_{\mathrm{per}}(y)\nabla_x u_\varepsilon(x,y)\right) \\
= \; & -\frac{1}{\varepsilon^2}\mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_y u_0(x,y)\right) \\
& -\frac{1}{\varepsilon}\left[\mathrm{div}_x\left(A_{\mathrm{per}}(y)\nabla_y u_0(x,y)\right) + \mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_x u_0(x,y)\right)\right] \\
& -\frac{1}{\varepsilon}\mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_y u_1(x,y)\right) \\
& -\mathrm{div}_x\left(A_{\mathrm{per}}(y)\nabla_x u_0(x,y)\right) + \mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_x u_1(x,y)\right) \\
& +\mathrm{div}_x\left(A_{\mathrm{per}}(y)\nabla_y u_1(x,y)\right) + \mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_y u_2(x,y)\right) - f(x) \\
& +O(\varepsilon).
\end{aligned}$$

L'équation (1.1) ne fait initialement intervenir que la variable x, car $y = \frac{x}{\varepsilon}$. Dorénavant on suppose qu'on peut désolidariser x et y, considérant de fait que les équations associées aux différents termes du développement sont vérifiées pour tout $(x,y) \in \mathcal{D} \times \mathbb{R}^d$. La nullité du terme en ε^{-2} entraine

$$-\mathrm{div}_y\left(A_{\mathrm{per}}(y)\nabla_y u_0(x,y)\right) = 0.$$

Puisqu'en outre u_0 est supposée périodique en la variable y, le problème (1.8) possède une unique solution définie à une constante dépendante de x près

$$\nabla_y u_0(x,y) = 0.$$

La nullité du terme en ε^{-1} fournit l'équation

$$\begin{cases} -\mathrm{div}_y \left(A_{\mathrm{per}}(y) \left(\nabla_x u_0(x) + \nabla_y u_1(x,y) \right) \right) = 0 \text{ dans } \mathbb{R}^d, \\ u_1(x,y) \text{ est } Q - \text{périodique en la variable } y. \end{cases} \tag{1.8}$$

On constate donc que u_1 dépend linéairement de $\nabla_x u_0(x)$, et peut donc s'écrire

$$u_1(x,y) = \sum_{i=1}^{d} \frac{\partial u_0}{\partial x_i}(x) w_i(y), \tag{1.9}$$

où w_i est le correcteur associé à la direction e_i, satisfaisant l'équation

$$\begin{cases} -\mathrm{div}_y \left(A_{\mathrm{per}}(y) \left(e_i + \nabla w_i(y) \right) \right) = 0 \ \text{ dans } \mathbb{R}^d, \\ w_i \text{ est } Q - \text{périodique.} \end{cases}$$

La fonction u_1 définie comme la solution de l'équation (1.8) est unique à l'addition d'une constante près $v(x)$ dépendant de la seule variable x. Dans le cadre de la dérivation de $A^\star$, il se trouve qu'elle ne joue aucun rôle particulier. Cependant, il y a un contexte dans lequel elle doit être prise en compte et que nous n'avons pas encore évoqué jusqu'à présent : c'est lorsqu'il s'agit d'évaluer précisément en quel sens le développement à deux échelles tronqué est une bonne approximation de u^ε. Pour le moment, il reste à traiter le terme d'ordre 0 qui s'écrit

$$-\mathrm{div}_y \left(A(y) \left(\nabla_x u_1(x,y) + \nabla_y u_2(x,y) \right) \right) = \mathrm{div}_x \left(A(y) \left(\nabla_x u_0(x) + \nabla_y u_1(x,y) \right) \right) + f(x),$$

où $u_2(x,y)$ est supposée périodique en la variable y.

En moyennant en $y \in Q$, on voit qu'une condition nécessaire (et qui se trouve être suffisante) à l'existence d'une fonction périodique u_2 solution de l'équation précédente est que

$$-\mathrm{div}_x \left(\int_Q A(y) \left(\nabla_x u_0(x) + \nabla_y u_1(x,y) \right) dy \right) = f(x).$$

En utilisant (1.9), on déduit de cette équation que

$$-\mathrm{div}_x \left(A^\star \nabla_x u_0(x) \right) = f(x),$$

où $A^\star$ est constante et définie par (1.5).

1.2.2 Cas quasi-périodique

Supposons à présent que la fonction A^ε intervenant dans (1.1) s'écrit sous la forme

$$A^\varepsilon(x) = A_{\mathrm{q-per}} \left(\frac{x}{\varepsilon} \right),$$

où $A_{\mathrm{q-per}}$ désigne une fonction non plus périodique mais quasi-périodique au sens où $A \in \left(C^0(\mathbb{R}^d)\right)^{d\times d}$ et il existe un $m \in \mathbb{N}^*$ tel que

$$\exists F \text{ périodique} \in \left(C^0(\mathbb{R}^{dm})\right)^{d\times d},\ A(x) = F(x,\cdots,x).$$

Alors la caractérisation de la matrice homogénéisée, ici notée $A^\star_{\mathrm{q-per}}$ change. En effet, on a

$$\left[A^\star_{\mathrm{q-per}}\right]_{ij} = \lim_{N\to+\infty} \frac{1}{N^d} \int_{[0,N]^d} (e_i + \nabla w_{e_i})^T A_{\mathrm{q-per}} \left(e_j + \nabla w_{e_j}\right), \tag{1.10}$$

où w_p est définie pour tout $p \in \mathbb{R}^d$ comme la solution, unique à une constante près, de l'équation

$$\begin{cases} -\mathrm{div}\left(A_{\mathrm{q-per}}\left(p + \nabla w_p\right)\right) = 0 \text{ dans } \mathbb{R}^d, \\ w_p \quad \text{quasi-périodique.} \end{cases} \tag{1.11}$$

La définition de $A^\star_{\mathrm{q-per}}$ fait intervenir les solutions d'un problème dont ε a disparu. Mais le problème (1.11) est posé sur $\mathbb{R}^d$ tout entier et la formule (1.10) ne peut bien sûr pas être évaluée directement. Si l'on dispose d'une formule explicite de $A^\star_{\mathrm{q-per}}$, le caractère explicite de celle-ci doit être relativisé. Il convient alors d'étudier comment approcher efficacement $A^\star_{\mathrm{q-per}}$. On renvoie par exemple à [18] pour une étude récente sur ce sujet.

1.2.3 Cas stochastique classique

Qu'ils soient d'origine naturelle ou industrielle, les *matériaux réels* peuvent ne pas être périodiques, ni même quasi-périodiques. En pratique, nous avons une connaissance imparfaite de leurs microstructures. Identifier les objets mathématiques adéquats à ces dernières est un domaine de recherche en soi. Dans la suite, nous allons faire des hypothèses fortes sur les propriétés statistiques des microstructures.

1.2.3.1 Définitions

On commence par se donner un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et on note

$$\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$$

l'espérance de la variable aléatoire $X \in L^1(\Omega, d\mathbb{P})$ relativement à la probabilité $\mathbb{P}$. On suppose ensuite que le groupe $(\mathbb{Z}^d, +)$ agit sur Ω via l'action de groupe $(\tau_k)_{k\in\mathbb{Z}^d}$. On suppose en outre que cette action préserve la mesure $\mathbb{P}$, i.e.

$$\forall k \in \mathbb{Z}^d,\ \ \forall A \in \mathcal{F},\ \ \mathbb{P}(\tau_k A) = \mathbb{P}(A).$$

On suppose également τ *ergodique*, au sens où

$$\forall A \in \mathcal{F},\ \ \left(\forall k \in \mathbb{Z}^d,\ \ \tau_k A = A\right) \Rightarrow \mathbb{P}(A) = 0 \text{ ou } 1.$$

Ceci nous permet de définir la notion de stationnarité discrète (voir [19]) : une fonction $F \in L^1_{\text{loc}}\left(\mathbb{R}^d, L^1(\Omega)\right)$ est dite *stationnaire* si, pour tout $k \in \mathbb{Z}^d$,

$$F(x+k,\omega) = F(x,\tau_k\omega), \text{ presque partout et presque sûrement.} \tag{1.12}$$

Cette définition abstraite implique qu'en deux points de $\mathbb{R}^d$ distants d'un nombre entier, les propriétés statistiques de F sont identiques. Ainsi, $\mathbb{E}\left[F(x,\cdot)\right]$ est $\mathbb{Z}^d$-périodique.

1.2.3.2 Définition de $A^\star$

On suppose à présent que la matrice A^ε apparaissant dans (1.1) est aléatoire et s'écrit

$$A^\varepsilon(x) = A\left(\frac{x}{\varepsilon},\omega\right), \tag{1.13}$$

où A est une fonction aléatoire *stationnaire* au sens de (1.12). Les propriétés du matériau sont donc statistiquement les mêmes en deux points dont les coordonnées diffèrent d'un nombre entier de périodes. L'équation (1.1) devient

$$\left\{\begin{array}{rcl} -\text{div}\left(A\left(\frac{x}{\varepsilon},\omega\right)\nabla u^\varepsilon(x,\omega)\right) &=& f \quad \text{dans} \quad \mathcal{D}, \\ u^\varepsilon(x,\omega) &=& 0 \quad \text{sur} \quad \partial\mathcal{D}. \end{array}\right. \tag{1.14}$$

Dans ce cadre, on sait que $A^\star$ est donnée par

$$[A^\star]_{ij} = \mathbb{E}\left[\int_Q \left(e_i + \nabla w_{e_i}\right)^T A\left(e_j + \nabla w_{e_j}\right)\right], \tag{1.15}$$

où, pour tout $p \in \mathbb{R}^d$, w_p est l'unique fonction, définie à une constante aléatoire près, dans $\left\{w \in L^2_{\text{loc}}(\mathbb{R}^d, L^2(\Omega)), \quad \nabla w \in L^2_{\text{unif}}(\mathbb{R}^d, L^2(\Omega))\right\}$, qui vérifie

$$\left\{\begin{array}{l} -\text{div}\left(A\left(p+\nabla w_p\right)\right) = 0 \text{ sur } \mathbb{R}^d, \\ \nabla w_p \text{ est stationnaire au sens de (1.12)}, \\ \mathbb{E}\left[\int_Q \nabla w_p\right] = 0. \end{array}\right. \tag{1.16}$$

La démonstration de ce résultat sous des hypothèses minimales est due à G. Papanicolaou et S.R.S Varadhan dans [83]. Des travaux similaires tels que [57, 60] obtiennent un résultat identique sous des hypothèses supplémentaires relatives à A (conditions de mélange ....). Enfin il est aussi possible de démontrer ce résultat dans le cadre d'un développement à deux échelles stochastique introduit dans [21].

Commençons par noter une propriété remarquable de la matrice $A^\star$: bien qu'exprimant les propriétés effectives d'un composite doté d'une microstructure *aléatoire*, elle est *déterministe*. En fait, ceci n'a rien d'étonnant et résulte

directement des propriétés de stationnarité et d'ergodicité. Plus précisément, c'est une conséquence du théorème ergodique. Il est possible de construire des matériaux pour lesquels un tel effet de moyennisation n'a pas lieu. C'est en particulier le cas lorsque l'aléa ne porte pas uniquement sur la microstructure mais est également attaché à l'échelle macroscopique. La matrice homogénéisée est alors aléatoire, mais cet aléa n'a pas de lien avec la procédure d'homogénéisation.

1.3 Approximations de la matrice homogénéisée dans le cas stochastique

L'avantage que représente le caractère déterministe de $A^\star$ est à mettre en regard d'une difficulté du même type (mais plus grande encore) que celle rencontrée dans le cadre de l'homogénéisation quasi-périodique : l'impossibilité de calculer directement $A^\star$. En effet, la définition de $A^\star$ implique la connaissance préalable des fonctions w_p. Or la résolution de (1.14) pose deux difficultés en pratique : il s'agit d'une équation aux dérivées partielles posée sur l'espace $\mathbb{R}^d$ *tout entier*, et *aléatoire*.

1.3.1 Approximations classiques

En pratique, une manière de procéder consiste à considérer l'équation (1.14) sur un domaine fini, munie de conditions aux limites adéquates. On introduit par exemple le problème des correcteurs tronqué

$$\begin{cases} -\mathrm{div}\left(A_N(\cdot,\omega)\left(p+\nabla w_p^N(\cdot,\omega)\right)\right)=0 \quad \text{sur} \quad \mathbb{R}^d, \\ w_p^N \quad Q_N\text{-périodique}, \end{cases} \tag{1.17}$$

où $Q_N = (-N,N)^d$ et A_N désigne la Q_N-périodisation de A, liée au choix arbitraire de conditions aux limites périodiques. Ces correcteurs approchés permettent de calculer

$$\left[A_N^\star\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^N(y,\omega)\right)^T A(y,\omega)\left(e_j+\nabla w_{e_j}^N(y,\omega)\right)dy. \tag{1.18}$$

On montre alors (voir [23]) que $A_N^\star \to A^\star$ presque sûrement lorsque $N \to +\infty$. L'approximation (1.18) est donc bien consistante. Il a également été montré dans [23] que l'on peut remplacer les conditions aux limites périodiques dans (1.17) par des conditions de Dirichlet ou encore de Neumann.

Sans hypothèse supplémentaire, on ne sait *a priori* rien dire quant à la vitesse de convergence de $A_N^\star$ vers $A^\star$. Pour obtenir des résultats dans ce sens, il faut faire des hypothèses sur la vitesse de décorrélation du champ stationnaire initial $A(y,\omega)$: on parle de condition de *mixing*. A. Bourgeat et A. Piatnitski montrent dans [23] que

$$\exists \gamma > 0, \quad \mathbb{E}\left(|A_N^\star - A^\star|\right) \leq CN^{-\gamma},$$

où C désigne une constante positive qui dépend de la dimension d et d'un paramètre associé à la vitesse de décorrélation. La preuve utilise de manière décisive une estimée tirée de [103]. A notre connaissance, ce résultat est le seul donnant une vitesse de convergence et il est peu explicite. De nombreux travaux ont ces dernières années tenté de l'améliorer. A. Gloria et F. Otto ont proposé dans [49, 50] des estimateurs différents de $A^\star$ dans le cas d'opérateurs *discrets* analogues. Ces estimateurs font intervenir deux paramètres supplémentaires T et L liés respectivement à l'ajout d'un terme d'ordre zéro dans l'équation définissant le correcteur approché, et à l'emploi de filtres lors de l'étape d'intégration. En s'inspirant des travaux de A. Naddaf et T. Spencer dans [76], A. Gloria et F. Otto ont en effet obtenu des vitesses précises de convergence pour ces estimateurs.

Dans le cas d'opérateurs *continus*, un angle d'attaque de cette question difficile a consisté en un examen du cas monodimensionnel. C'est l'esprit de l'article [13] de G. Bal, J. Garnier, S. Motsch et V. Perrier, basé sur les formules explicites du coefficient homogénéisé et de la fonction u^ε. Ceci leur permet d'établir des convergences en loi du type :

$$\frac{u^\varepsilon(x,\cdot) - u^\star(x)}{\varepsilon^\alpha} \underset{\varepsilon\to 0}{\longrightarrow} K(x),$$

où la valeur de α et la nature du processus aléatoire K dépendent des propiétés précises du processus $a(x,\omega)$. En dimensions supérieures, des résultats analogues ont été obtenus dans [14] dans le cas d'équations de la forme

$$-\Delta u_\varepsilon + V\left(\frac{x}{\varepsilon},\omega\right) u_\varepsilon = f.$$

Mais la question reste *a priori* ouverte dans le cas des opérateurs $\mathcal{L}^\varepsilon = -\mathrm{div}(A^\varepsilon\nabla)$ qui nous concernent plus spécifiquement ici.

1.3.2 Les contributions de la Partie I

Dans le cadre de cette thèse, notre approche est tout à fait différente. A N fixé, on peut décomposer l'erreur *aléatoire* entre la matrice homogénéisée exacte $A^\star$ et son approximation $A_N^\star$ fournie par (1.18), de la manière suivante

$$A^\star - A_N^\star(\omega) = A^\star - \mathbb{E}(A_N^\star) + \mathbb{E}(A_N^\star) - A_N^\star(\omega).$$

Le terme $A^\star - \mathbb{E}(A_N^\star)$ correspond à un biais. Lorsque N est suffisamment grand, on le supposera négligeable. En d'autres termes on peut supposer que $\mathbb{E}(A_N^\star) \approx A^\star$. Dès lors la source principale de l'erreur est le terme $\mathbb{E}(A_N^\star) - A_N^\star(\omega)$: c'est l'*erreur statistique* liée à l'estimation de $\mathbb{E}(A_N^\star)$. En pratique, pour approcher cette valeur de $\mathbb{E}(A_N^\star)$, on a recours à une méthode de type Monte Carlo. On définit l'estimateur

$$\mu_{2M}(A_N^\star) = \frac{1}{2M}\sum_{m=1}^{2M} A_{N,m}^\star(\omega),$$

où $\left(A^\star_{N,m}\right)_{1\leq m\leq 2M}$ désigne une suite de $2M$ copies indépendantes et identiquement distribuées de la variable aléatoire $A^\star_N$. On sait alors que, pour $1 \leq i,j \leq d$, on a

$$\left|\mathbb{E}([A^\star_N]_{ij}) - [\mu_{2M}(A^\star_N)]_{ij}\right| \leq 1.96\frac{\sqrt{\mathbb{V}\mathrm{ar}([A^\star_N]_{ij})}}{\sqrt{2M}}. \tag{1.19}$$

La précision de l'estimation de $\mathbb{E}([A^\star_N]_{ij})$ dépend donc directement de la variance de $A^\star_N$. Réduire cette variance constitue l'objectif des travaux regroupés dans la **Partie I**. Nous y proposons des estimateurs $\widetilde{A}^\star_N$ de $A^\star$, tels que $\mathbb{E}\left(\widetilde{A}^\star_N\right) = \mathbb{E}(A^\star_N)$ (on ne veut pas risquer d'augmenter le biais supposé négligeable), et qui possèdent une variance plus faible que $A^\star_N$.

Le **Chapitre 3** s'intéresse à l'utilisation dans le cadre de l'homogénéisation stochastique de la technique des variables antithétiques [69]. Nous proposons une nouvelle approximation $\widetilde{A}^\star_N$ de $A^\star_N$, basée sur la construction d'une matrice homogénéisée antithétique notée $B^\star_N$. Cette variable est définie comme la matrice homogénéisée associée à un champ hétérogène B, obtenu par transformation *déterministe* du champ A. A titre d'exemple, on suppose la matrice initiale A de la forme $A = a\mathbb{I}_d$ avec

$$a(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x)a_k(\omega), \tag{1.20}$$

où Q désigne toujours la cellule unité de $\mathbb{R}^d$ et $(a_k(\omega))_{k\in\mathbb{Z}^d}$ désigne une famille de variables aléatoires indépendantes et identiquement distribuées selon la loi uniforme $\mathcal{U}([\alpha,\beta])$ (il s'agit d'un des cas étudiés dans [P2]). On définit tout d'abord le champs antithétique $B = b\mathbb{I}_d$ où

$$b(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x)\left(\alpha+\beta-a_k(\omega)\right).$$

On note que sur chaque cellule $Q+k$ on a remplacé la variable aléatoire $a_k(\omega)$ par sa variable antithétique $\alpha+\beta-a_k(\omega)$. On obtient la matrice aléatoire $B^\star_N(\omega)$ par homogénéisation Q_N-périodique de la matrice B. En d'autres termes on substitue B à A dans (1.17-1.18). La nouvelle approximation de $A^\star$ est alors définie par

$$\widetilde{A}^\star_N(\omega) = \frac{1}{2}\left(A^\star_N(\omega) + B^\star_N(\omega)\right). \tag{1.21}$$

Une réalisation de $\widetilde{A}^\star_N$ implique la résolution de deux problèmes des correcteurs contre un seul pour $A^\star_N$. Ainsi un estimateur équivalent à $\mu_{2M}(A^\star_N)$ en terme de coût calcul est

$$\mu_M\left(\widetilde{A}^\star_N\right) = \frac{1}{M}\sum_{m=1}^{M}\widetilde{A}^\star_{N,m}(\omega),$$

où $\left(\widetilde{A}^\star_{N,m}\right)_{1\leq m\leq M}$ désigne une suite de M copies indépendantes et identiquement distribuées de la variable $\widetilde{A}^\star_N$ correspondant à $2M$ réalisations du problème des correcteurs sous-jacent. Gardant en mémoire l'estimation (1.19), nous choisissons donc dans [P2] d'estimer pour différentes valeurs de N le gain en terme de ratio renormalisé des variances

$$R\left([A^\star_N]_{11}\right) = \frac{\mathbb{V}\mathrm{ar}\left([A^\star_N]_{11}\right)}{2\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}^\star_N\right]_{11}\right)}.$$

Ce ratio est homogène à un gain en termes de coût calcul à précision fixée. Sa racine carrée correspond au ratio des largeurs des intervalles de confiance associés aux estimateurs de $A^\star_N$ et $\widetilde{A}^\star_N$, à coût calcul fixé.

Les résultats numériques obtenus dans [P2] montrent que l'application de la technique des variables antithétiques réduit *systématiquement* la variance des termes diagonaux de la matrice $A^\star_N$ lorsque $A = a\mathbb{I}_d$ avec a définie par (1.20), i.e. pour différents choix de la loi de la variable a_0. Dans ces cas *a priori* très simples, la variance de $u^\star_N$ solution de

$$\left\{\begin{array}{rcl} -\mathrm{div}\left(A^\star_N\nabla u^\star_N\right) &=& f \quad \text{dans} \quad \mathcal{D}, \\ u^\star_N &=& 0 \quad \text{sur} \quad \partial\mathcal{D}, \end{array}\right. \tag{1.22}$$

est également réduite de manière systématique. C'est là encore un résultat de [P2]. Le **Chapitre 4** a pour objectif de généraliser l'usage des variables antithétiques au cas où la matrice $A^\star_N$ résulte d'un champ inital A possédant des propriétés plus complexes : anisotropie et présence de corrélation entre les valeurs de A. Ce chapitre s'intéresse aussi à la réduction de variance obtenue pour des quantités d'intérêt autres que les termes diagonaux de la matrice $A^\star_N$: valeurs propres de la matrice $A^\star_N$, valeurs propres de l'opérateur $-\mathrm{div}\left(A^\star_N\nabla\right)$, termes non diagonaux de $A^\star_N$. Ces résultats numériques confirment l'efficacité systématique de la méthode.

Ces résultats numériques illustrent les résultats théoriques donnés dans le **Chapitre 5**. Lorsque $\widetilde{A}^\star_N$ est définie par (1.21), une condition nécessaire et suffisante pour que soit réduite la variance du terme diagonal $[A^\star_N]_{11}$ par exemple est

$$\mathbb{C}\mathrm{ov}\left([A^\star_N]_{11}, [B^\star_N]_{11}\right) \leq 0.$$

Nous démontrons dans [P3] que cette condition est vérifiée dès lors qu'on peut écrire le champ initial sous la forme

$$\forall x \in Q_N,\ A(x,\omega) = \mathcal{A}\left(X_1(\omega), \cdots, X_n(\omega)\right),$$

où $n \in \mathbb{N}^*$, $(X_k(\omega))_{1\leq k\leq n}$ est une famille de variables i.i.d distribuées selon une loi uniforme $\mathcal{U}([0,1])$ et $\mathcal{A} : \mathbb{R}^n \to \mathbb{S}_N$ est une fonction soit croissante en toutes ses variables, soit décroissante en toutes ses variables (au sens des matrices symétriques).

Notons que cette hypothèse de structure n'est pas contraignante et qu'elle est satisfaite par beaucoup d'exemples en pratique. Elle est compatible avec des propriétés telles que l'anisotropie ou la corrélation. Elle permet de généraliser aux dimensions supérieures la preuve monodimensionnelle déjà présente dans [P2]. En outre, la preuve proposée dans [P3] explique la réduction de variance non seulement pour les termes diagonaux de $A_N^\star$, mais aussi pour toute quantité $\mathcal{F}(A_N^\star)$ avec $\mathcal{F} : \mathbb{S}_N \to \mathbb{R}$ monotone au sens des matrices symétriques. Notons que les résultats de [P3], s'ils expliquent *qualitativement* pourquoi la technique des variables antithétiques s'applique, dans un cadre précis, avec succès aux problèmes d'homogénéisation stochastique, *ne fournissent pas d'estimation quantitative* de la réduction de variance. Cette évaluation quantitative de l'efficacité motive les investigations numériques du **Chapitre 4**. Notons enfin que dans le cas monodimensionnel, il est possible d'obtenir des quantifications explicites de la réduction de variance. De même, la seconde section du **Chapitre 5** montre que dans un cas particulier multidimensionnel, dit cas pertubatif, on peut exprimer directement la variance d'une approximation de $A_N^\star$ à partir de celles des variables $(X_k(\omega))_{1\leq k\leq n}$ intervenant dans la définition du champ initial.

1.4 Problèmes faiblement stochastiques

Dans le cadre stochastique classique $A^\varepsilon(x) = A\left(x/\varepsilon, \omega\right)$ avec A stationnaire, nous avons vu que le calcul des propriétés du matériau effectif *déterministe*, i.e. de la matrice homogénéisée, pose systématiquement problème car elle demande la résolution d'un problème aléatoire posé sur $\mathbb{R}^d$ tout entier. Puisque ces approximations sont aléatoires, nous avons été conduits plus haut à chercher des méthodes qui réduisent systématiquement ce caractère et donc l'incertitude quant au comportement du matériau effectif.

Dans les faits, bien que les matériaux réels ne soient par exemple pas exactement périodiques, leurs microstructures attestent parfois tout de même d'une certaine régularité. Les hypothèses relatives à la forme de la matrice A^ε dans le cadre aléatoire présenté ci-dessus sont donc en quelque sorte trop générales. Elles intègrent une trop grande part d'aléa par rapport aux besoins. Ce constat nous conduit à considérer dans la suite des modèles perturbatifs. Dans tous ces modèles, le matériau est défini comme une petite perturbation d'un matériau dont on sait calculer aisément les propriétés effectives. Nous verrons que ceci a pour conséquence une réduction très importante du coût de calcul des propriétés homogénéisées.

Les premiers pas dans ce sens remontent au travail de L. Tartar [96]. Ils portent sur un cas déterministe. La matrice associée aux propriétés du matériau y est supposée de la forme

$$A_\eta^\varepsilon(x) = \left(a_0(x) + \eta b^\varepsilon(x) + \eta^2 c^\varepsilon(x)\right) \mathbb{I}_d,$$

où a_0 est indépendante du paramètre ε, $b^\varepsilon \rightharpoonup b^0$, $c^\varepsilon \rightharpoonup c^0$ et $(b^\varepsilon - b^0)/\varepsilon^2 \rightharpoonup \sigma$ dans $L^\infty(\mathcal{D})$ faible *. L'auteur démontre dans ce contexte que la matrice homogénéisée est de la forme :

$$A^\star_\eta(x) = a_0(x)\mathbb{I}_d + \eta b^0(x)\mathbb{I}_d + \eta^2(c^0(x)\mathbb{I}_d - \frac{1}{a_0(x)}M(x)) + o(\eta^3),$$

où M est une matrice symétrique positive et telle que $\mathrm{Tr}(M) = \sigma^2$. Si on sait par ailleurs que $A^\star_\eta$ est isotrope on a donc $M(x) = \sigma^2(x)/d\mathbb{I}_d$. En conséquence, on peut calculer le développement à l'ordre 2 en η de $A^\star_\eta$ à partir des seules limites faibles (moyennes ici) des champs initiaux. Dans cette même direction et sous des hypothèses plus spécifiques, on pourra également se référer aux travaux de M. Briane, D. Manceau et G.W Milton [25, 70, 71]. Enfin, il convient également de citer, toujours dans un cadre déterministe, le travail de G. Allaire et S. Gutiérrez [4].

Avant de passer à l'extension de ces idées au cas de l'homogénéisation stochastique, nous décrivons une variante du cadre classique pour l'homogénéisation stochastique. Ce modèle aléatoire peut ensuite être décliné en une version faiblement stochastique que nous avons étudiée.

1.4.1 Difféomorphisme aléatoire

La théorie de l'homogénéisation stochastique a été développée dans le cadre standard exposé dans la section précédente. Dans [19, 20], X. Blanc, C. Le Bris et P.L-Lions ont étudié un modèle alternatif. Dans l'équation (1.1), la matrice A^ε est maintenant définie par

$$A^\varepsilon(x) = A_{\mathrm{per}}\left(\Phi^{-1}\left(\frac{x}{\varepsilon}, \omega\right)\right), \tag{1.23}$$

où A_{per} désigne une matrice périodique et Φ est un difféomorphisme aléatoire dont le gradient est supposé stationnaire au sens de (1.12). L'idée est donc de représenter un matériau parfait (périodique) deformé de manière aléatoire. Comme exemple d'application, on peut penser à un matériau composite, dont les fibres occupent des positions aléatoires.

On remarque que Φ n'étant pas lui-même supposé stationnaire, il n'y a aucune raison pour que A^ε ne le soit. Néanmoins, pour cette famille de champs aléatoires, les auteurs ont montré dans [19], qu'on dispose d'une caractérisation explicite de la matrice homogénéisée. Ainsi, on a

$$[A^\star]_{ij} = \frac{\mathbb{E}\left(\int_{\Phi(Q,\cdot)} (e_i + \nabla w_{e_i}(y,\cdot))^T A_{\mathrm{per}} \left(\Phi^{-1}(y,\cdot)\right) \left(e_j + \nabla w_{e_j}(y,\omega)\right) dy\right)}{\det\left(\mathbb{E}\left(\int_Q \nabla\Phi(y,\cdot)dy\right)\right)},$$

où, pour tout $p \in \mathbb{R}^d$ fixé, w_p est l'unique solution dans $H^1_{\mathrm{loc}}(\mathbb{R}^d, L^2(\Omega))$, définie à une constante (aléatoire) près, du système

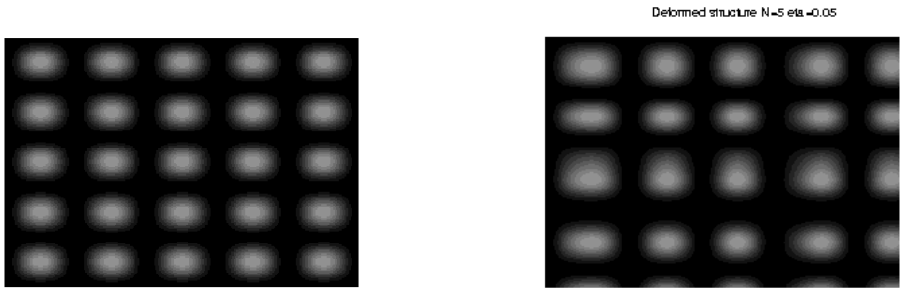


FIG. 1.1 – Un exemple de petite déformation aléatoire. A gauche, le matériau périodique de référence $A_{\rm per}$. A droite, le matériau déformé $A_{\rm per} \circ \Phi^{-1}$ avec Φ de la forme (1.25) et $\eta = 0.05$.

$$\left\{ \begin{array}{l} \mathrm{div}\left[A_{\rm per}\left(\Phi^{-1}(y,\omega)\right)(p+\nabla w_p)\right] = 0, \\ w_p(y,\omega) = \widetilde{w}_p\left(\Phi^{-1}(y,\omega),\omega\right), \quad \nabla\widetilde{w}_p \text{ est stationnaire au sens de (1.12)}, \\ \mathbb{E}\left(\displaystyle\int_{\Phi(Q,\cdot)} \nabla w_p(y,\cdot)dy\right) = 0. \end{array} \right. \tag{1.24}$$

Il convient de remarquer à nouveau que w_p est solution d'un problème posé sur $\mathbb{R}^d$ tout entier. Mais ce n'est pas ici la seule difficulté. En effet, sous cette forme, l'équation (1.24) ne fait pas intervenir de quantité stationnaire. Ceci a pour conséquence l'impossibilité d'utiliser en l'état la procédure de troncature issue de [23]. En d'autres termes, il faut proposer une nouvelle approximation $A^\star_N$ de $A^\star$ et redémontrer le cas échéant la convergence de $A^\star_N$ vers $A^\star$. Une telle approximation a été proposée dans [P1] (cf **Chapitre 6**).

1.4.2 Cadres perturbatifs

Nous avons donc vu ci-dessus deux cadres pour l'homogénéisation stochastique : le cas où $A^\varepsilon(x) = A(x/\varepsilon,\omega)$ avec A stationnaire, et le cas où $A^\varepsilon(x) = A_{\rm per}\left(\Phi^{-1}(x/\varepsilon,\omega)\right)$ où Φ est un difféomorphisme aéatoire. Ces deux cadres admettent une version faiblement stochastique que nous décrivons maintenant.

Dans le cas stochastique, la première occurence d'un modèle perturbatif pour l'homogénéisation se trouve à notre connaissance dans [19]. Rappelons que, dans ce contexte, la forme de A^ε est donnée par :

$$A^\varepsilon(x,\omega) = A_{\rm per}\left(\Phi^{-1}\left(\frac{x}{\varepsilon},\omega\right)\right),$$

où Φ est un difféomorphisme aléatoire. Si $\Phi = \mathbf{Id}$, alors on est revenu au cas périodique et $A^\star$ est simple à calculer. Un cas intéressant est alors celui où Φ est

une petite perturbation aléatoire de l'identité. Dans [20], les auteurs considèrent le cas où

$$\Phi_\eta(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2), \tag{1.25}$$

où le développement vaut dans un espace à préciser. Sous ces hypothèses, on montre que (cf [19]) la matrice homogénéisée $A^\star_\eta$ possède un développement limité en le petit paramètre η

$$A^\star_\eta = A^\star_{\rm per} + \eta A^\star_1 + O(\eta^2), \tag{1.26}$$

où $A^\star_{\rm per}$ est la matrice homogénéisée associée à la matrice $A_{\rm per}$, facile à calculer. Le point crucial est que la définition de la matrice $A^\star_1$ fait intervenir une fonction déterministe, solution d'un problème déterministe posé sur $Q = (0,1)^d$, et non plus sur $\mathbb{R}^d$ comme c'était le cas pour le correcteur exact. Plus précisément

$$\begin{aligned}
[A^\star_1]_{ij} &= -\int_Q \mathbb{E}(\mathrm{div}\Psi)\ \left[A^\star_{\rm per}\right]_{ij} + \int_Q (e_i + \nabla w^0_{e_i})^T A_{\rm per}\, e_j\, \mathbb{E}(\mathrm{div}\Psi) \\
&+ \int_Q \left(\nabla \overline{w}^1_{e_i} - \mathbb{E}(\nabla\Psi)\nabla w^0_{e_i}\right)^T A_{\rm per}\, e_j,
\end{aligned} \tag{1.27}$$

où $\overline{w}^1_p$ est la solution Q-périodique (unique à une constante près) du problème

$$\begin{aligned}
-\mathrm{div}\left(A_{\rm per}\,\nabla\overline{w}^1_p\right) &= \mathrm{div}\left(-A_{\rm per}\,\mathbb{E}(\nabla\Psi)\,\nabla w^0_p\right) \\
&- \mathrm{div}\left(\left(\mathbb{E}(\nabla\Psi^T) - \mathbb{E}(\mathrm{div}\Psi)\mathrm{Id}\right) A_{\rm per}\,(p + \nabla w^0_p)\right).
\end{aligned} \tag{1.28}$$

La résolution de (1.28) est incomparablement moins coûteuse que celle du problème des correcteurs stochastiques car ce problème est déterministe et posé sur le domaine borné Q. La caractérisation de la matrice homogénéisée devient effectivement explicite au premier ordre en η, au sens où elle ne fait intervenir que des quantités aisément calculables en pratique.

Dans le cadre classique de l'homogénéisation stochastique, la transcription des idées issues de [20] consiste à choisir

$$A^\varepsilon_\eta(x) = A_{\rm per}\left(\frac{x}{\varepsilon}\right) + \eta A_1\left(\frac{x}{\varepsilon},\omega\right) + O(\eta^2). \tag{1.29}$$

On trouvera dans [P3] une justification rigoureuse du fait, que dans ce cas la matrice homogénéisée possède également un développement en η

$$A^\star_\eta = A^\star_{\rm per} + \eta A^\star_1 + O(\eta^2),$$

avec cette fois

$$[A^\star_1]_{ij} = \int_Q e_i^T A_{\rm per}\overline{w}^1_{e_j} + \int_Q e_i^T \mathbb{E}(A_1)\left(\nabla w^0_{e_j} + e_j\right)$$

et

$$-\mathrm{div}\left(A\,\nabla\overline{w}^1_p\right) = \mathrm{div}\left(\mathbb{E}\left(A_1\right)(p + \nabla w^0_p)\right). \tag{1.30}$$

A nouveau le calcul au premier ordre revient à calculer les fonctions *déterministes* $\overline{w}_p^1$ et w_p^0, solutions de problèmes posés sur la cellule Q.

Un cadre voisin du précédent a été étudié dans une série d'articles par A. Anantharaman et C. Le Bris (cf les articles [7–9] et le Chapitre 2 pour une introduction). Il s'agit d'autoriser des déformations rares mais d'amplitudes potentiellement importantes du matériau de référence associé à $A_{\rm per}$. La manière précise de traduire cette intuition dépend de l'espace fonctionnel au sein duquel est supposé valoir le développement de A_η^ε. Pour certains choix d'espaces, on sait que la matrice homogénéisée exacte possède un développement limité en η, dont le premier terme *et les suivants* sont peu coûteux à calculer.

1.4.3 Les contributions de la Partie II

Dans le cadre de cette thèse, nous nous sommes concentrés sur les cadres perturbatifs associés au cadre standard de l'homogénéisation stochastique (expression (1.29)) ainsi qu'à sa variante introduite dans [19, 20] (expressions (1.23) et (1.25)) . Dans les deux cas, on sait que la matrice homogénéisée *exacte* admet un développement limité en η du type (1.26), dont les deux premiers termes sont simples à calculer, car ils font intervenir une équation aux dérivées partielles déterministe posée sur un domaine de petite taille Q. On a donc réalisé un gain important de coût calcul *si le développement à l'ordre 1 en η de $A_\eta^\star$ est suffisamment précis.*

Vérifier cette hypothèse en pratique implique l'introduction d'une méthode d'approximation pour le calcul de $\overline{w}_p^1$ qui permet d'accéder à une approximation $\overline{w}_p^{1,h}$ (type éléments finis) de $\overline{w}_p^1$. On calcule donc *in fine* un terme $A_1^{\star,h}$ qui n'est pas le premier terme du développement en η de la matrice $A_\eta^\star$ ($A_1^{\star,h}$ est seulement une approximation de $A_1^\star$ convergente quand $h \to 0$). Pour vérifier numériquement la pertinence du développement en η, on est donc amené à trouver une approximation numérique de $A_\eta^\star$ cohérente avec celle employée pour approximer $\overline{w}_p^1$ et $A_1^\star$. Or une telle approximation est, comme nous l'avons vu plus haut, nécessairement stochastique. La conséquence de cette discussion est la suivante : vérifier numériquement la validité du développement au premier ordre en η de la matrice homogénéisée *exacte* $A_\eta^\star$ conduit à analyser l'erreur au deuxième ordre d'une approximation aléatoire de cette matrice. Nous n'avons *a priori* aucune information sur cette quantité. En particulier nous ignorons le comportement de sa variance lorsque $\eta \to 0$.

Le **Chapitre 6** examine ce problème dans le cadre du modèle issu de [19, 20]. La matrice $A_\eta^\star$ est approximée via une méthode de discrétisation pour le

problème (1.24). On définit ainsi

$$\left[A_{h,N}^{\star,\eta}\right]_{ij}(\omega) = \frac{\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N} \det(\nabla\Phi_\eta)\left(e_i + (\nabla\Phi_\eta)^{-1}\nabla\widetilde{w}_{e_i}^{h,N}\right)^T A_{\rm per}\left(e_j + \nabla\widetilde{w}_{e_j}^{h,N}\right)}{\det\left(\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\nabla\Phi_\eta\right)}, \tag{1.31}$$

où $\widetilde{w}_{e_i}^{h,N}$ est solution de

$$\left\{\begin{array}{l} \text{Trouver } \widetilde{w}_p^{h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N) \text{ tel que, pour tout } \widetilde{v}_h \in V_h^{\rm per}(Q_N), \\ \int_{Q_N} \det\left(\nabla\Phi_\eta\right)\left(\nabla\widetilde{v}_h\right)^T\left(\nabla\Phi_\eta\right)^{-T} A_{\rm per}\left(p + (\nabla\Phi_\eta)^{-1}\nabla\widetilde{w}_p^{h,N}(\cdot,\omega)\right) = 0 \quad \text{p.s.} \end{array}\right. \tag{1.32}$$

et $V_h^{\rm per}(Q_N)$ est l'ensemble des fonctions définies sur $\mathbb{R}^d$, Q_N-périodiques dont la restriction à Q_N appartient à un espace d'éléments finis périodiques construit sur un maillage adéquat. Notre objectif est de juger de la qualité du développement limité. Pour ce faire, on est tenté de s'intéresser à la variance des composantes de la matrice

$$\eta^{-2}\left(A_{h,N}^{\star,\eta} - A_{\rm per}^\star - \eta A_{1,h}^\star\right).$$

Dans [P1] nous expliquons pourquoi il ne faut précisément pas considérer cette quantité. En effet, nous annonçons dans [P1] un résultat qui trouvera dans [P6] (voir **Chapitre 7**) sa forme définitive : la matrice $A_{h,N}^{\star,\eta}$, obtenue après troncature et discrétisation éléments finis, possède elle-même un développement limité en la variable η :

$$A_{h,N}^{\star,\eta}(\omega) = A_{{\rm per},h}^\star + \eta A_{1,h,N}^\star(\omega) + O(\eta^2).$$

Nous montrons *numériquement* sur un exemple relativement simple que la variance du terme en $O(\eta^2)$ dans le développement ci-dessus est la quantité à laquelle il faut s'intéresser et qu'en outre elle est bien bornée quand $\eta \to 0$. Ceci légitime l'usage de l'approximation au premier ordre de la matrice $A_\eta^\star$. La vérification *théorique* de ce résultat est l'objet du **Chapitre 7**. Celui-ci démontre comment sous certaines hypothèses relatives au développement de la matrice A_η^ε, le terme d'ordre 2 est effectivement borné *indépendamment des paramètres h et N* associés à la discrétisation numérique. Ceci confirme les résultats numériques du **Chapitre 6** et prouve que le cadre perturbatif est conservé dans le passage aux quantités discrétisées, seules accessibles à la pratique numérique.

1.5 Perspectives

Dans le cadre de cette thèse, nos contributions, détaillées dans les sections précédentes, se sont donc articulées autour d'un même objectif : *faciliter le calcul des propriétés effectives des matériaux aléatoires*. Dans cette optique, nous avons adopté deux types d'approches : une approche générale correspondant aux travaux de la **Partie I**, qui reposent sur l'utilisation de techniques de réduction

de variance dans le contexte de l'homogénéisation stochastique ; et une approche dite perturbative, associée aux travaux de la **Partie II**, qui a pour principe une modification de la modélisation. Nos apports dans chacune de ces directions ont été les suivants.

Dans la **Partie I**, nous avons démontré et illustré numériquement que la technique des variables antithétiques permet d'obtenir, pour une grande variété de quantités d'intérêt, des gains significatifs en terme de coût calcul ; et ce indépendamment des caractéristiques du champ hétérogène originel auquel s'applique la procédure d'homogénéisation. Ceci ouvre la voie à d'autres développements dans cette direction rassemblés dans la section 1.5.1 : application d'autres techniques de réduction de variance, combinaison de ces techniques avec des estimateurs alternatifs de la matrice homogénéisée, application de ces techniques à d'autres quantités. Nous allons détailler plus bas chacune de ces directions.

Dans la **Partie II**, nous avons démontré que les résultats théoriques relatifs au cadre perturbatif, décrits dans [20], sont également valables lorsqu'on s'intéresse à des quantités discrétisées. Ceci permet de construire des approximations numériques déterministes et peu coûteuses de la matrice homogénéisée, dont le bon comportement, lorsque l'amplitude de la perturbation du champ initial tend vers zéro, est garanti. Ainsi, nous avons montré que les travaux de [20] conduisent bien à des stratégies numériques fiables et efficaces. L'intérêt pratique de ce cadre perturbatif dans le cas d'opérateurs linéaires sous forme divergence permet d'envisager sa généralisation à d'autres types d'équations. Il s'agit de vérifier que des hypothèses perturbatives sur une donnée initiale du problème, du même type que celles effectuées dans la **Partie II**, induisent bien des développements des propriétés homogénéisées *exactes* et *approximées* ; développements dont les premiers termes sont simples à calculer. La section 1.5.2 s'intéresse à cette direction de travail, et montre, au travers d'un exemple précis, comment la démarche semble se généraliser.

Enfin, nous abordons en conclusion des directions de travail qui s'inscrivent moins directement dans la continuité des travaux de cette thèse, et dont l'exploration est *a priori* plus difficile.

1.5.1 Réduction de variance

1.5.1.1 D'autres techniques de réduction de variance

Dans le cadre de la **Partie I**, nous avons choisi d'étudier l'application d'une technique particulière de réduction de variance au cadre de l'homogénéisation stochastique : la technique des variables antithétiques. Nous avons montré que cette méthode permet de réduire systématiquement la variance de quantités qui

dépendent de $A_N^\star$. Le gain en terme de coût calcul est systématique, mais il n'est généralement pas spectaculaire (du moins pour les exemples les plus difficiles de [P2]). Ceci motive notre tentative d'appliquer à l'homogénéisation stochastique d'autres techniques de réduction de variance, *a priori* plus difficiles à mettre en oeuvre, mais qu'on espère plus efficaces.

Cette mise en oeuvre est l'objet d'un travail en cours qui s'intéresse plus spécialement à la technique des variables de contrôle. Nous expliquons ici le principe de la méthode, et montrons pourquoi le cas monodimensionnel laisse espérer, pour un choix habile de la variable de contrôle, des gains en terme de coût calcul bien supérieurs à ceux observés dans le cas des variables antithétiques. Notons enfin qu'il existe d'autres techniques de réduction de variance (échantillonnage préférentiel, stratification, ...) dont l'adaptation au cadre de l'homogénéisation stochastique mériterait que l'on s'y intéresse.

Variable de contrôle. Le principe de la technique de la variable de contrôle peut s'exprimer de la manière suivante. On veut évaluer l'espérance d'une variable aléatoire Y. On suppose que l'on connaît une variable aléatoire X corrélée à Y, dont l'espérance est calculable à moindre coût numérique. En pratique, on construira des estimateurs de l'espérance de la variable

$$\widetilde{Y}_b := Y + b\left(\mathbb{E}\left(X\right) - X\right), \tag{1.33}$$

et non de Y. Ainsi $\mathbb{E}(Y) = \mathbb{E}(\widetilde{Y}_b)$ sera approchée par $\frac{1}{M}\sum_{k=1}^{M} \widetilde{Y}_b^k$, où les $\left(\widetilde{Y}_b^k\right)_{1\leq k\leq M}$ sont des copies indépendantes et identiquement distribuées de la variable aléatoire $\widetilde{Y}_b$. Dans la définition (1.33), b désigne un paramètre déterministe dont la valeur est inconnue *a priori*. La variance de $\widetilde{Y}_b$ est donnée par

$$\mathbb{V}\mathrm{ar}\left(\widetilde{Y}_b\right) = \mathbb{V}\mathrm{ar}\left(Y\right) + b^2\mathbb{V}\mathrm{ar}\left(X\right) - 2b\mathbb{C}\mathrm{ov}\left(Y,X\right).$$

Elle est minimale pour la valeur

$$b_0 = \frac{\mathbb{C}\mathrm{ov}\left(Y,X\right)}{\mathbb{V}\mathrm{ar}\left(X\right)}.$$

La variance de $\widetilde{Y}_{b_0}$ correspondante est

$$\mathbb{V}\mathrm{ar}\left(\widetilde{Y}_{b_0}\right) = \mathbb{V}\mathrm{ar}\left(Y\right)\left(1-\rho_{X,Y}^2\right) \leq \mathbb{V}\mathrm{ar}\left(Y\right),$$

où

$$\rho_{X,Y} = \frac{\mathbb{C}\mathrm{ov}\left(Y,X\right)}{\sqrt{\mathbb{V}\mathrm{ar}\left(X\right)}\sqrt{\mathbb{V}\mathrm{ar}\left(Y\right)}}.$$

Cette formule s'étend aux estimateurs statistiques intervenant dans la méthode de Monte Carlo. L'estimateur classique de $\mathbb{E}(Y)$ est

$$\mu_M(Y) = \frac{1}{M}\sum_{k=1}^{M} Y_k, \tag{1.34}$$

et on propose maintenant d'utiliser

$$\mu_M\left(\widetilde{Y}_b\right) = \frac{1}{M}\sum_{k=1}^{M} Y_k + b\left(\mathbb{E}(X) - X_k\right). \tag{1.35}$$

On voit alors que

$$\mathbb{V}\mathrm{ar}\left(\mu_M\left(\widetilde{Y}_{b_0}\right)\right) = \mathbb{V}\mathrm{ar}\left(\mu_M(Y)\right)\left(1-\rho_{X,Y}^2\right) \leq \mathbb{V}\mathrm{ar}\left(\mu_M(Y)\right).$$

Ainsi l'utilisation de (1.35) plutôt que (1.34) est systématiquement intéressante, car la variance de cet estimateur est plus petite. En pratique, la valeur de b_0 est inconnue et doit être estimée. A cet effet, on utilise l'estimateur

$$\widehat{b}_0 = \frac{\displaystyle\sum_{k=1}^{M}\left(\left(X_k - \mu_M(X)\right)\left(Y_k - \mu_M(Y)\right)\right)}{\displaystyle\sum_{k=1}^{M} X_k - \mu_M(X)}.$$

Dans le cadre de l'homogénéisation stochastique, on cherche à estimer $\mathbb{E}\left(A_N^\star\right)$. Appliquer la technique de la variable de contrôle consiste donc à trouver une variable $B_N^\star \in \mathbb{R}^{d\times d}$ et des paramètres b_{ij} tels que, pour $1 \leq i,j \leq d$, la variance de

$$\left[\widetilde{A}_N^\star\right]_{ij} = [A_N^\star]_{ij} + b_{ij}\left(\mathbb{E}\left([B_N^\star]_{ij}\right) - [B_N^\star]_{ij}\right)$$

soit très inférieure à celle de $[A_N^\star]_{ij}$.

Afin de construire une telle variable $B_N^\star$, nous avons commencé par examiner le cas monodimensionnel. On suppose, comme dans [P2], que le champ initial s'écrit sous la forme

$$a(x,\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1]} a_k(\omega),$$

où les variables $(a_k)_{k\in\mathbb{Z}}$ sont supposées indépendantes et identiquement distribuées : pour tout $k \in \mathbb{Z}^d$ $a_k \sim \mathcal{U}([\alpha,\beta])$. Dans ce cadre monodimensionnel, le coefficient homogénéisé $a_N^\star$ est donné par la formule

$$a_N^\star(\omega) = \left(\frac{1}{N}\sum_{k=1}^{N}\frac{1}{a_k(\omega)}\right)^{-1}.$$

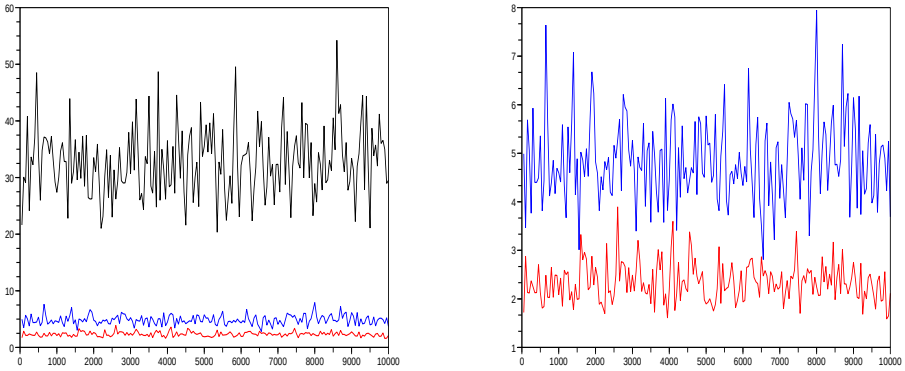


FIG. 1.2 – Ratios des variances des estimateurs de $a_N^\star$ et $\widetilde{a}_N^\star = a_N^\star + \widehat{b}_0 \left(\mathbb{E}\left(b_N^{\star,1}\right) - b_N^{\star,1} \right)$ en fonction de N (abscisse) pour différentes valeurs du contraste β/α. La figure de droite est un simple agrandissement de celle de gauche.

Une première idée consiste à utiliser la variable de contrôle

$$b_N^{\star,1}(\omega) = \frac{1}{N} \sum_{k=1}^{N} a_k(\omega),$$

dont l'espérance est explicite et vaut $(\alpha + \beta)/2$. La figure 1.2 représente pour différentes valeurs des paramètres α, β (différents contrastes) le ratio des *variances* des estimateurs $a_N^\star$ et $\widetilde{a}_N^\star$ (pour le choix optimal de b estimé). La courbe rouge est associée au plus fort contraste, la noire au plus faible, et la courbe bleue correspond au cas examiné dans [P2] : $\alpha = 3, \beta = 20$. On constate que les réductions de variance sont du même ordre de grandeur que celles obtenues dans le cas des variables antithétiques. On ne gagne donc rien à utiliser $b_N^{\star,1}(\omega)$.

Une seconde idée est d'utiliser la variable de contrôle

$$b_N^{\star,2}(\omega) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{a_k(\omega)}.$$

Dans ce cas également, $\mathbb{E}\left(b_N^{\star,2}\right)$ est connue. La figure 1.3 montre que l'emploi de cette variable de contrôle implique une réduction de variance très importante. Dans le cas $\beta/\alpha = 20/3$ (notre situation de référence dans [P2]), le ratio des variances est supérieur à 300 pour $N = 1000$. *Dans un cadre monodimensionnel, la méthode semble donc très efficace* même si, comme l'illustre la figure de droite, le ratio de variance reste sensible au contraste.

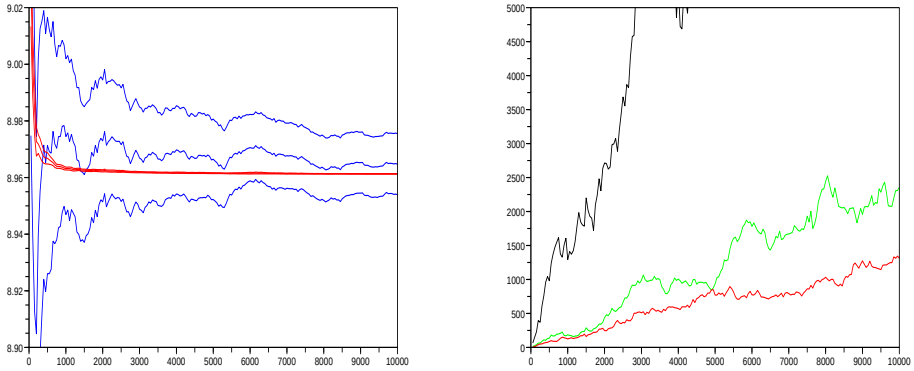


FIG. 1.3 – A gauche, estimations de la moyenne et intervalles de confiance en fonction de N pour $a_N^\star$ (courbe bleue) et $\widetilde{a}_N^\star = a_N^\star + \widehat{b}_0\left(\mathbb{E}\left(b_N^{\star,2}\right) - b_N^{\star,2}\right)$ (courbe rouge). A droite, ratios des variances de $a_N^\star$ et $\widetilde{a}_N^\star$ en fonction de N pour différentes valeurs du contraste β/α.

On est donc tenté de chercher en dimension supérieure une définition de $B_N^\star$ (la matrice homogénéisée de contrôle) qui généralise la définition de $b_N^{\star,2}$. L'extension la plus naturelle consiste à poser $B_N^\star = \left(A_N^\star\right)^{-1}$. Malheureusement, on ignore $\mathbb{E}\left(\left(A_N^\star\right)^{-1}\right)$ au sens où l'on ne dispose pas d'une méthode nous permettant d'estimer cette quantité de façon plus efficace que $\mathbb{E}(A_N^\star)$. Ainsi $\left(A_N^\star\right)^{-1}$ ne peut en tant que tel constituer une bonne variable de contrôle. On en est donc réduit à chercher une approximation de $\left(A_N^\star\right)^{-1}$ dont on sache estimer l'espérance à moindre coût.

Nous avons donc songé entre autres à la méthode suivante. Tout d'abord, on suppose que $N = cN_1$ avec $c, N_1 \in \mathbb{N}$. On calcule alors sur chacune des c^d sous-cellules $\mathcal{Q}_k := \mathcal{Q}_{(k_1,\cdots,k_d)} = [N_1 k_1, N_1(k_1+1)] \times \cdots \times [N_1 k_d, N_1(k_d+1)]$ où $k \in \mathbb{N}^d$ et $|k|_\infty \leq c$, une approximation locale de la matrice homogénéisée

$$\left[A_{N,k}^\star(\omega)\right]_{ij} = \frac{1}{|\mathcal{Q}_k|}\int_{\mathcal{Q}_k} (e_i + \nabla z_i^{N,k})^T A (e_j + \nabla z_j^{N,k}),$$

avec, pour tout $p \in \mathbb{R}^d$,

$$\begin{cases} -\mathrm{div}\left(A\left(p + \nabla z_p^{N,k}\right)\right) = 0 \ \text{ sur } \ \mathcal{Q}_k \text{ p.s.}, \\ \nabla z_p^{N,k} \ \ \mathcal{Q}_k\text{-périodique.} \end{cases}$$

Enfin on pose $B_N^\star(\omega) = \frac{1}{c^d}\sum_{k=1}^{c^d}\left(A_{N,k}^\star(\omega)\right)^{-1}$. A nouveau, on ignore $\mathbb{E}\left(B_N^\star\right)$. Comme $B_N^\star$ est la moyenne arithmétique de quantités calculées sur de petits domaines, on peut espérer estimer à moindre coût son espérance via l'application, en précalcul, de méthodes de Monte-Carlo indépendantes sur chaque petite cellule $\mathcal{Q}_k$.
Une autre approche pour chercher à mieux comprendre la manière de construire,

dans le cadre multidimensionnel, une variable de contrôle efficace est de commencer par étudier un opérateur plus simple que $-\mathrm{div}(A^\varepsilon \nabla)$.

1.5.1.2 Réduction de variance pour d'autres estimateurs de $A^\star$

Il serait intéressant de comprendre théoriquement et numériquement *comment utiliser simultanément les techniques de réduction de variance que nous avons développées avec les estimateurs de la matrice homogénéisée proposés par A. Gloria et F. Otto* dans le cas discret. Dans [49], les auteurs proposent des approximations de $A^\star$ du type

$$\forall\ 1 \leq i,j \leq d, \quad \left[A^\star_{N,T,L}\right]_{ij} = \frac{1}{|Q_N|}\int_{Q_N} \varphi_L(y)(e_i + \nabla w^{N,T}_{e_i})A(e_j + \nabla w^{N,T}_{e_j}), \quad (1.36)$$

où φ_L est une fonction filtre régulière, et où pour tout $p \in \mathbb{R}^d$, $w^{N,T}_p$ est la solution (définie à une constante près) de l'équation

$$\left\{ \begin{array}{l} -\mathrm{div}\left(A_N(\cdot,\omega)\left(p + \nabla w^{N,T}_p(\cdot,\omega)\right)\right) + T^{-1} w^{N,T}_p(\cdot,\omega) = 0 \quad \text{sur} \quad \mathbb{R}^d, \\ w^{N,T}_p \quad Q_N\text{-périodique.} \end{array} \right.$$

L'application de la méthode des variables antithétiques dans ce contexte ne présente *a priori* aucune difficulté sur le plan pratique. Si cette méthode s'avère efficace nous souhaiterions dans un second temps comprendre pourquoi. Car, sur le plan théorique, la formule (1.36) nous empêche d'avoir recours directement aux outils utilisés dans [P3] pour établir théoriquement l'efficacité de la méthode. La présence du filtre φ_L introduit en effet des difficultés pour utiliser les propriétés de monotonie de l'homogénéisation. En conséquence, et puisque de toute façon, comme nous l'avons déjà expliqué plus haut, une telle méthode de preuve échoue à quantifier la potentielle réduction de variance, il faudrait effectuer dans un premier temps un certain nombre de tests numériques.

1.5.1.3 Réduction de variance pour des estimateurs de u^ε

Dans le cadre de cette thèse nous nous sommes la plupart du temps concentrés sur la réduction de variance de quantités dérivant de la matrice homogénéisée $A^\star$ (à l'exception de la dernière section de [P2], où nous considérons également $u^\star$). Or, dans la pratique, on est également intéressé par $u^\varepsilon(\cdot,\omega)$, qui est une quantité *intrinsèquement* stochastique. Dans le cadre de l'homogénéisation stochastique, à partir de la matrice $A^\star_N(\omega)$ définie par (1.18), on introduit $u^\star_N(\omega)$ solution de

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A^\star_N(\omega)\nabla u^\star_N(\omega)\right) &=& f \quad \text{dans} \quad \mathcal{D}, \\ u^\star_N &=& 0 \quad \text{sur} \quad \partial\mathcal{D}. \end{array} \right.$$

Quand $\varepsilon \to 0$, une approximation de u^ε alternative à $u_N^\star$ est donnée par le développement à deux échelles au premier ordre en ε :

$$u_N^\varepsilon(x,\omega) := u_N^\star(x,\omega) + \sum_{i=1}^{d} w_{e_i}^N\left(\frac{x}{\varepsilon},\omega\right)\frac{\partial u_N^\star}{\partial x_i}(x,\omega),$$

où $w_{e_i}^N$ désigne la solution du problème des correcteurs tronqué (1.17) dans le cas $p = e_i$. On peut dès lors se demander dans quelle mesure u_N^ε approche $u^\star$ quand N est grand et ε petit et ensuite *si les techniques de réduction de variance exposées ici permettent de diminuer la variance de* u_N^ε.
On pourrait aussi s'interroger sur l'efficacité de ces techniques lorsqu'on approxime u^ε via une des méthodes numériques multi-échelles évoquées plus haut (MsFEM par exemple).

1.5.2 Cadre perturbatif

Dans la continuité des travaux de la **Partie II**, le premier point concerne *l'adaption du cadre perturbatif, faiblement stochastique, à d'autres équations* traitées par la théorie de l'homogénéisation stochastique. De ce point de vue, un premier pas pourrait consister en l'étude des équations d'Euler-Lagrange dérivant de problèmes de minimisation de la forme

$$\inf_{v\in H_0^1(\mathcal{D})}\int_{\mathcal{D}} W\left(\frac{x}{\varepsilon},\nabla v,\omega\right) - \int_{\mathcal{D}} fv, \tag{1.37}$$

où $W(y,F,\omega) : \mathbb{R}^d\times\mathbb{R}^d\times\Omega \to \mathbb{R}$ désigne une fonction aléatoire, régulière, strictement convexe en la variable F et stationnaire en les variables (y,ω). La fonctionnelle W et sa dérivée W_F en la variable F sont supposées vérifier des conditions adéquates. Sous ces conditions, lorsque $\varepsilon \to 0$, on sait (voir [31]) que l'*infimum* (1.37) tend vers

$$\inf_{v\in H_0^1(\mathcal{D})}\int_{\mathcal{D}} W^\star(\nabla v) - \int_{\mathcal{D}} fv, \tag{1.38}$$

où

$$W^\star(F) = \lim_{N\to+\infty}\mathbb{E}\left(\inf_{v\in H^1_{\mathrm{per}}(Q_N)}\frac{1}{|Q_N|}\int_{Q_N} W\left(y,F+\nabla v,\cdot\right)dy\right). \tag{1.39}$$

De plus le minimiseur u^ε de (1.37) tend vers le minimiseur $u^\star$ de (1.38). La définition (1.39) se récrit sous la forme

$$W^\star(F) = \lim_{N\to+\infty}\mathbb{E}\left(\frac{1}{|Q_N|}\int_{Q_N} W\left(y,F+\nabla w_F^N,\cdot\right)dy\right), \tag{1.40}$$

où la fonction w_F^N est le minimiseur du problème

$$\inf_{v\in H^1_{\rm per}(Q_N)} \frac{1}{|Q_N|}\int_{Q_N} W\left(y, F+\nabla v, \omega\right) dy,$$

solution de l'équation

$$\left\{\begin{array}{l} -\mathrm{div}\left(W_F^N\left(\cdot, F+\nabla w_F^N(\cdot,\omega),\omega\right)\right)=0 \quad \text{sur} \quad \mathbb{R}^d, \\ w_F^N(\cdot,\omega) \quad Q_N\text{-périodique}, \end{array}\right. \tag{1.41}$$

où W_F^N désigne la Q_N-périodisation de W_F.

Supposons maintenant que W n'est que faiblement stochastique, au sens où

$$W\left(y,F,\omega\right) = W_{\rm per}\left(y,F\right) + \eta W_1\left(y,F,\omega\right) + O(\eta^2), \tag{1.42}$$

où $W_{\rm per}(y,F) : \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ est Q-périodique en la variable y, et $W_1(y,F,\omega) : \mathbb{R}^d\times\mathbb{R}^d\times\Omega\to\mathbb{R}$ stationnaire en (y,ω). Admettons que la fonction w_F^N possède aussi un développement en η

$$w_F^N = w_{0,F}^N + \eta w_{1,F}^N + O(\eta^2). \tag{1.43}$$

On insère ce développement dans (1.41). Alors, toujours formellement, on constate que $w_{0,F}^N = w_{0,F}$ est solution de

$$\left\{\begin{array}{l} -\mathrm{div}\left(W_{{\rm per},F}\left(\cdot, F+\nabla w_{0,F}\right)\right)=0 \quad \text{sur} \quad \mathbb{R}^d, \\ w_{0,F} \quad Q\text{-périodique}, \end{array}\right.$$

et que $w_{1,F}^N$ est solution de

$$\left\{\begin{array}{l} -\mathrm{div}\left(W_{{\rm per},FF}\left(\cdot, F+\nabla w_{0,F}\right)\nabla w_{1,F}^N(\cdot,\omega)\right) = \\ \qquad \mathrm{div}\left(W_{1,F}^N\left(\cdot, F+\nabla w_{0,F},\omega\right)\right) \quad \text{sur} \quad \mathbb{R}^d \quad \text{p.s.}, \\ w_{1,F}(\cdot,\omega) \quad Q_N\text{-périodique}, \end{array}\right. \tag{1.44}$$

où W_{FF} désigne la dérivée d'ordre 2 en la variable F de la fonctionnelle d'énergie W et $W_{1,F}^N$ est la Q_N-périodisation de $W_{1,F}$. On insère alors (1.43) dans (1.40) et, en utilisant (1.42), on obtient

$$W^\star(F) = W^\star_{\rm per}(F) + \eta W_1^\star(F) + O(\eta^2),$$

où

$$\begin{aligned} W^\star_{\rm per}(F) &= \int_Q W_{\rm per}\left(y, F+\nabla w_{0,F}\right), \\ W_1^\star(F) &= \lim_{N\to+\infty}\frac{1}{|Q_N|}\int_{Q_N} W_{{\rm per},F}\left(y,F+\nabla w_{0,F}\right)\nabla\mathbb{E}\left(w_{1,F}^N\right) dy \\ &+ \mathbb{E}\left(\int_Q W_1\left(y,F+\nabla w_{0,F},\cdot\right) dy\right). \end{aligned}$$

On voit donc que $W^{\star}_{\text{per}}$ est facile à calculer. Enfin, on remarque en prenant l'espérance de (1.45) que $\overline{w}_{1,F} := \mathbb{E}\left(w^N_{1,F}\right)$ est solution d'un problème déterministe posé sur Q

$$\begin{cases} -\text{div}\left(W_{\text{per},FF}\left(\cdot, F+\nabla w_{0,F}\right)\nabla \overline{w}_{1,F}\right) = \\ \qquad \text{div}\left(\overline{W}_{1,F}\left(\cdot, F+\nabla w_{0,F}, \omega\right)\right) \quad \text{sur} \quad \mathbb{R}^d, \\ \overline{w}_{1,F} \quad Q\text{-périodique}, \end{cases}$$

où $\overline{W}_{1,F}(y,F) := \mathbb{E}\left(W^N_{1,F}(y,F,\cdot)\right) = \mathbb{E}\left(W_{1,F}(y,F,\cdot)\right)$ est une fonction $Q-$périodique, indépendante de N. Ainsi la définition du premier ordre en η de $W^{\star}_1(F)$ se récrit

$$\begin{aligned} W^{\star}_1(F) &= \int_Q W_{\text{per},F}\left(y, F+\nabla w_{0,F}, \cdot\right)\nabla \overline{w}_{1,F} dy \\ &+ \mathbb{E}\left(\int_Q W_1\left(y, F+\nabla w_{0,F}, \cdot\right) dy\right), \end{aligned}$$

et ne fait plus intervenir que des fonctions solutions de problèmes déterministes, posés sur Q. Il semble donc que l'on puisse généraliser le cadre perturbatif à ce type de situation. Les résultats décrits ici coïncident dans le cas où $W(y,F,\omega) = 1/2F^T A(y,\omega)F$ avec ceux de la **Partie II**. On note pour finir que ces résultats ont été dérivés de manière purement formelle. Il reste à démontrer tout cela de manière rigoureuse, puis à étudier l'influence de la discrétisation numérique employée pour résoudre en pratique le problème (1.41).
Notons enfin qu'il est également vraisemblable que le cadre pertubatif pour l'homogénéisation puisse s'étendre à des équations plus générales.

1.5.3 Perspectives plus générales

1.5.3.1 Systèmes linéaires aléatoires

Revenant au cas d'opérateurs de la forme $-\text{div}(A^{\varepsilon}\nabla)$, une autre direction de travail possible consiste à *s'intéresser directement aux systèmes algébriques linéaires aléatoires* qui résultent du problème (1.1) ou du problème des correcteurs. Ils s'écrivent sous la forme :

$$K^{\varepsilon}(\omega)U^{\varepsilon}(\omega) = B(\omega),$$

où la dimension de la matrice de rigidité aléatoire K^{ε} croît avec ε. Or ces matrices aléatoires possèdent une structure particulière, qui implique possiblement des propriétés limites exploitables dans le contexte de l'homogénéisation. Indépendamment du contexte de l'homogénéisation, certains travaux de mécanique choisissent de représenter l'incertitude directement sur la matrice de rigidité. En conséquence, ils choisissent une matrice K^{ε} appartenant à un ensemble de matrices dont les propriétés asymptotiques (quand la taille de la matrice tend vers l'infini) sont décrites par la théorie des matrices aléatoires. La matrice A^{ε} dérivant du problème (1.1)

n'appartient pas à un tel ensemble. Mais peut-être existe-t-il un moyen de s'y ramener et donc d'exploiter certaines propriétés limites de ces ensembles.

1.5.3.2 Modélisation

Tous les travaux de cette thèse reposent sur le fait que l'on se donne *a priori* un modèle pour la microstructure aléatoire associé à la matrice A^ε. Or, en pratique, dans bon nombre de cas, on ne dispose pas d'un tel modèle. Celui que l'on utilise le cas échéant dans le cadre de la simulation numérique est le produit d'une estimation statistique empirique du matériau considéré. Ainsi, si l'on suppose que le matériau réel est représenté par A^ε, nous n'avons accès en pratique qu'à une estimation de A^ε notée $\widehat{A}^\varepsilon$. La matrice $\widehat{A}^\varepsilon$ a pour matrice homogénéisée une matrice $\widehat{A}^\star$, sensée expliquer (à des approximations d'ordre ε près) le comportement macroscopique du matériau réel, lui même décrit par $A^\star$ résultant de l'homogénéisation de A^ε. Un exemple typique est le cas où l'on ne dispose que d'une information lacunaire sur le matériau (quelques coupes ou échantillons par exemple) qui ne suffisent pas à fournir un modèle fiable et précis.

Supposons dès lors que l'on dispose d'une information partielle sur $A^\star$. Si nous possédons un estimateur suffisamment précis de $\widehat{A}^\star$ (obtenu par exemple grace aux méthodes exposées dans les **Parties I** et **II**), alors nous sommes en mesure de tester la pertinence du modèle initial $\widehat{A}^\varepsilon$. L'idée générale est de permettre *un dialogue entre la modélisation et l'homogénéisation : la seconde constituant un moyen de tester et d'actualiser les hypothèses relatives au matériau auquel elle s'applique.* Il faut trouver un moyen *a posteriori* de contrôler les hypothèses sous lesquelles ont été calculées les propriétés homogénéisées.

Chapitre 2

Homogénéisation stochastique numérique : quelques contributions récentes

Ce chapitre est tiré de [P5]. Il reprend et détaille certains points du chapitre précédent : principe et éléments de théorie de l'homogénéisation stochastique, obtention de bornes sur la matrice homogénéisée, présentation de l'approche *Multiscale Finite Elements* (MsFEMs). L'objectif est de présenter de manière unifiée une série de travaux récents portant sur l'homogénéisation stochastique *numérique*. De plus, il annonce certains résultats obtenus dans le cadre des **Parties I** et **II**.

Introduction to numerical stochastic homogenization and the related computational challenges : some recent developments

Arnaud Anantharaman[a], Ronan Costaouec[a], Claude Le Bris[a], Frédéric Legoll[a] and Florian Thomines[a]

[a] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

{*anantharaman,costaour,lebris*}*@cermics.enpc.fr, legoll@lami.enpc.fr,florian.thomines@enpc.fr*

2.1 Introduction

The purpose of this set of Notes is to present some recent developments, both on the theoretical and on the computational sides, in stochastic homogenization. These developments, obtained in collaboration within a whole group of researchers, have already been presented in a scattered manner in the literature (see [P1, P2, P3, P4, 5, 7–9, 19, 20, 101]). We present them here from a unified perspective.

In short, the bottom line for the series of works presented is the wish of the authors to contribute *to make numerical random homogenization more practical*. Random homogenization is indeed always very, and sometimes prohibitively, costly. Although a now traditional topic for mathematical analysis, random homogenization is also rather poorly known from the numerical analysis viewpoint. There is thus a definite interest in putting efforts on such issues.

Because we cannot embrace all difficulties at once, the case under consideration here is homogenization of a *simple, linear, scalar second order elliptic partial differential equation in divergence form*. We focus on the highly oscillating coefficient present in the equation, the different manner this coefficient can be modelled, and how this affects the various computational approaches.

The Notes begin with, in Section 2.2, a brief, hopefully pedagogic introduction to periodic, general (in Subsection 2.2.1) and next random (in Subsection 2.2.2) homogenization theories. There is of course no novelty in such an introduction, the only purpose of which is the consistency of the contribution and the convenience of

the reader not familiar with the theory. We refer to, *e.g.*, the monographs [16,27,58] for more details on homogenization theory, to [2, Chapters 1 and 2] for a pedagogic presentation, and to [56] for a short non technical overview of related problems. A super elementary introduction is contained in [64].

The following Sections 2.3 and 2.4 present our recent series of works. They are respectively based upon more comprehensive texts published in [5, 7–9] (for Subsection 2.3.1), [P1, 19, 20] (for Subsection 2.3.2), and [P2, P3, P4] (for Section 2.4). Section 2.3 presents a set of studies that all aim at treating the random problems under consideration as "perturbations" of a periodic problem. This of course requires some appropriate assumptions on the coefficients in the equation, so that the case under consideration is close to periodic, or, otherwise stated, that the amount of randomness present in the system is, to some extent, small. We term such situations *weakly random* situations. The computational workload is of course expected to be lighter in such situations, and we design here numerical approaches so that it is indeed the case. Section 2.4 addresses an issue that, although different in nature, also strongly affects the computational workload in numerical random homogenization : *variance*. We present some elementary numerical strategies, along with the necessary theoretical ingredients, that reduce the variance in computational approaches. We also investigate related issues.

A quick overview of some of the issues and techniques considered here has appeared in [63].

2.2 Basics of stochastic homogenization

Stochastic homogenization is best understood in the light of the easiest context of homogenization : *periodic* homogenization. This is the reason why we begin with Section 2.2.1 laying some groundwork in the periodic (and next general) context, before turning to stochastic homogenization *per se* in Section 2.2.2.

2.2.1 General and Periodic homogenization

2.2.1.1 Periodic setting

Main result in the periodic setting To begin with, we recall some basic ingredients of elliptic homogenization theory in the periodic setting. As said above, we refer *e.g.* to the monographs [16, 27, 58] for more details on homogenization theory.

We consider, in a regular bounded domain $\mathcal{D}$ in $\mathbb{R}^d$, the problem

$$\begin{cases} -\text{div}\left[A_{per}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right] = f \quad \text{in} \quad \mathcal{D}, \\ \\ u^\varepsilon = 0 \quad \text{on} \quad \partial\mathcal{D}, \end{cases} \tag{2.1}$$

where the matrix A_{per} is symmetric definite positive and $\mathbb{Z}^d$-periodic. We manipulate for simplicity *symmetric* matrices, but the discussion carries over to non symmetric

matrices up to slight modifications.

The microscopic problem associated to (2.1), called the *corrector problem* in the terminology of homogenization theory, reads, for p fixed in $\mathbb{R}^d$,

$$\left\{ \begin{array}{l} -\mathrm{div}\left(A_{per}(y)\left(p+\nabla w_p(y)\right)\right)=0 \quad \text{in} \quad \mathbb{R}^d, \\ \\ w_p \text{ is } \mathbb{Z}^d\text{-periodic.} \end{array} \right. \tag{2.2}$$

It has a unique solution up to the addition of a constant. Then, the *homogenized* coefficients read

$$\begin{aligned} [A^\star]_{ij} &= \int_Q \left(e_i+\nabla w_{e_i}(y)\right)^T A_{per}(y)\left(e_j+\nabla w_{e_j}(y)\right) dy \\ &= \int_Q \left(e_i+\nabla w_{e_i}(y)\right)^T A_{per}(y) e_j dy, \end{aligned} \tag{2.3}$$

where Q is the unit cube, and where w_{e_i} denotes the solution to (2.2) for $p=e_i$, with e_i the canonical vectors of $\mathbb{R}^d$. The main result of periodic homogenization theory is that, as ε goes to zero, the solution u^ε to (2.1) converges to $u^\star$ solution to

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A^\star \nabla u^\star\right]=f \quad \text{in} \quad \mathcal{D}, \\ \\ u^\star=0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right. \tag{2.4}$$

The convergence holds in $L^2(\mathcal{D})$, and weakly in $H_0^1(\mathcal{D})$. The *correctors* w_{e_i} may then also be used to "correct" $u^\star$ in order to identify the behavior of u^ε in the strong topology of $H_0^1(\mathcal{D})$. We more precisely have

$$u^\varepsilon - u^\star - \varepsilon \sum_{i=1}^d w_{e_i}\left(\frac{\cdot}{\varepsilon}\right)\frac{\partial u^\star}{\partial x_i} \underset{\varepsilon\to 0}{\to} 0 \quad \text{in } H^1(\mathcal{D}). \tag{2.5}$$

Several other convergences on various products involving $A_{per}\left(\frac{x}{\varepsilon}\right)$ and u^ε also hold. All this is well documented.

The practical interest of the approach is evident. No small scale ε is present in the homogenized problem (2.4). At the price of only computing d periodic problems (2.2) (as many problems as dimensions in the ambient space) the solution to problem (2.1) can be efficiently approached for ε small. A direct attack of problem (2.1) would require taking a meshsize smaller than ε. The difficulty has been circumvented. Of course, many improvements and alternatives exist in the literature.

Some intuitive insight on the result It is not immediate, for the reader not familiar with the theory, to understand the nature of the above asymptotic result. So we briefly give some elementary instructive arguments. To begin with, we modify

(2.1) in a dramatic way : we consider a one-dimensional domain, we take A_{per} a scalar valued function (denoted a_{per} for simplicity), we assume it is bounded, positive, bounded away from zero, we delete the differential operators of (2.1), and we are left with the oversimplified equation $-a_{per}\left(\frac{x}{\varepsilon}\right) u^\varepsilon(x) = f(x)$, which has obvious solution $u^\varepsilon = -a_{per}^{-1}\left(\frac{\cdot}{\varepsilon}\right) f$. We see that finding an homogenized problem then amounts to identifying the behaviour, as ε vanishes, of the sequence of oscillating functions $a_{per}^{-1}\left(\frac{x}{\varepsilon}\right)$. The topology of *weak convergence*, more precisely here that of the weak-$\star$ convergence in L^∞, is the appropriate tool for this purpose. Since the weak limit of a periodic function is evidently its average, it is immediately seen that the weak-$\star$ limit of u^ε when ε vanishes is

$$u^\star = -\langle a_{per}^{-1}\rangle\, f\,. \tag{2.6}$$

Based on this, it is natural to guess that the homogenized equation arising from (2.1) will be related to some average of the periodic function a_{per}^{-1}. This is what is expressed, in the one-dimensional setting, in (2.9) below. The higher dimensional setting is even more intricate.

We remark that a naive approach would consist in guessing that the homogenized equation is obtained by simply replacing $A_{per}\left(\frac{x}{\varepsilon}\right)$ in (2.1) by its average $\langle A_{per}\rangle$. This is *not* correct, as will be seen already in the simple one-dimensional setting below. It is actually even clear on our "zero-dimensional" calculation, since we obtain (2.6) and *not* : $u^\star = -\left(\langle a_{per}\rangle\right)^{-1} f$. Otherwise stated, and somewhat in the language of the engineering sciences, knowing the statistics of A_{per} is not sufficient to know the statistics on the solution u, and this of course owes to the fact that the map $A_{per} \longrightarrow u$ that associates to A_{per} the solution u to the equation is *not* linear.

2.2.1.2 The one-dimensional setting

We now reinstall the differential operators, and consider the one-dimensional setting :

$$\left\{\begin{array}{rcl} -\dfrac{d}{dx}(a(\dfrac{\cdot}{\varepsilon})\dfrac{d}{dx}u^\varepsilon) &=& f \quad \text{in }]0,1[, \\ \\ u^\varepsilon(0) = u^\varepsilon(1) &=& 0, \end{array}\right. \tag{2.7}$$

where a is a 1-periodic, scalar valued function. We assume it is positive, bounded and bounded away from zero, that is, there exist $0 < c_1 \leq c_2 < +\infty$ such that

$$0 < c_1 \leq a(x) \leq c_2,\ \forall x \in]0,1[. \tag{2.8}$$

We additionally assume that $f \in L^2(]0,1[)$. It is easily seen that the unique solution $u^\varepsilon \in H_0^1(]0,1[)$ to problem (2.7) converges in $L^2(]0,1[)$, and weakly in $H_0^1(]0,1[)$, to the solution $u^\star \in H_0^1(]0,1[)$ of the *homogenized* equation

$$-\frac{d}{dx}(\langle a^{-1}\rangle^{-1}\frac{d}{dx}u^\star) = f. \tag{2.9}$$

The proof can be performed at least in three different manners. First, simple compactness arguments, which we do not reproduce here for the sake of brevity (we refer to the bibliography), allow to argue on the equation (2.7) itself and to easily conclude. Second, the one-dimensional context allows to explicitly find u^ε by quadrature and thus identify its limit. Indeed, it can be shown that

$$u^\varepsilon(x) = -c_\varepsilon \int_0^x a(\frac{\cdot}{\varepsilon})^{-1} - \int_0^x a(\frac{\cdot}{\varepsilon})^{-1} F \tag{2.10}$$

with the constant

$$c_\varepsilon = -\langle a(\frac{\cdot}{\varepsilon})^{-1} \rangle^{-1} \int_0^1 a(\frac{\cdot}{\varepsilon})^{-1} F, \tag{2.11}$$

where we have denoted by $F(x) = \displaystyle\int_0^x f(t)dt$ and $\langle \cdot \rangle$ the integral over the range $[0,1]$. On the other hand, we know the solution to (2.9) is

$$u^\star(x) = \langle a^{-1} \rangle \left(x \int_0^1 F - \int_0^x F \right), \tag{2.12}$$

the first derivative of which reads :

$$(u^\star)'(x) = \langle a^{-1} \rangle \left(\int_0^1 F - F(x) \right). \tag{2.13}$$

Consider then

$$(u^\varepsilon)' - (u^\star)' = \left(\langle a^{-1} \rangle - a(\frac{\cdot}{\varepsilon})^{-1} \right) \left(c_\varepsilon + F \right) - < a^{-1} > \left(c_\varepsilon + \int_0^1 F \right).$$

Again using that the weak limit of a periodic function is its average, we easily note that

$$c_\varepsilon + \int_0^1 F = O(\varepsilon), \tag{2.14}$$

and that the first term of the right hand side *weakly* converges to zero. That convergence is (generically) never strong, and we deduce that $(u^\varepsilon)' - (u^\star)'$ (only) *weakly* converges to zero in L^2.

Third, we may use the two-scale expansion technique that will be outlined below. It allows to claim that the limit $u^\star$ is the solution to

$$\left\{ \begin{array}{rcl} -\dfrac{d}{dx}\left(a^\star \dfrac{d}{dx} u^\star\right) & = & f \quad \text{in } [0,1], \\ u^\star(0) = u^\star(1) & = & 0, \end{array} \right.$$

where the homogenized coefficient $a^\star$ is

$$a^\star = \int_0^1 a(y)(1 + w'(y))^2 \, dy, \tag{2.15}$$

with w solution to

$$\left\{ \begin{array}{rl} -\dfrac{d}{dy}(a(y)(1+w'(y))) & = 0 \quad \text{in}\,\mathbb{R}, \\ w\ 1-\text{periodic}, & \end{array} \right.$$

which is the one-dimensional version of (2.2)-(2.4). It is straightforward to see that

$$w'(y) = -1 + \frac{1}{\langle a^{-1} \rangle}\, a^{-1}(y), \tag{2.16}$$

and thus that (up to an additive constant)

$$w(y) = -y + \frac{1}{\langle a^{-1} \rangle} \int_0^y a^{-1}. \tag{2.17}$$

Inserting (2.16) into (2.15), we find the explicit value

$$a^\star = \int_0^1 a(y)(1+w'(y))^2\, dy = \int_0^1 a(y)\, \frac{1}{\langle a^{-1} \rangle^2}\, a^{-2}(y)\, dy = \langle a^{-1} \rangle^{-1}$$

which agrees with the coefficient of (2.9) above.

The interest of the latter approach is that it allows, using the function w, to improve the approximation. After some simple computations, we obtain

$$\begin{array}{rcl} (u^\varepsilon)'(x) - \left(u^\star + \varepsilon (u^\star)' w(\dfrac{\cdot}{\varepsilon})\right)'(x) & = & -a(\dfrac{x}{\varepsilon})^{-1} \left(c_\varepsilon + \displaystyle\int_0^1 F \right) \\ & & + \varepsilon f(x) \displaystyle\int_0^{\frac{x}{\varepsilon}} \left[a(y)^{-1} - < a^{-1} > \right] dy, \end{array} \tag{2.18}$$

where we now have strong convergence of the right-hand side (in fact (2.18) is $O(\varepsilon)$ in $L^\infty([0,1])$). We therefore obtain

$$u^\varepsilon(x) - (u^\star(x) + \varepsilon (u^\star)'(x) w(\frac{x}{\varepsilon})) \text{ converges to } 0 \text{ in } H^1(\mathcal{D}). \tag{2.19}$$

2.2.1.3 Higher dimensions

Two-scale expansion and related tools In dimensions higher than one, the elementary arguments based on explicitness of the functions, or simple compactness theorems, do not carry over. Only a few cases are explicitly solvable (we refer to the bibliography for the homogenization of *laminated* and *checkerboard*-type materials). The proof of the asymptotic behaviour is established otherwise. And it requires much stronger technologies. One possible approach is the *energy method* (a.k.a. the *method*

of oscillating test functions) by Murat and Tartar (see [75, 94]). Another possible approach is to use the notion of *two-scale convergence* introduced by G. Nguetseng and developed by G. Allaire (see [1, 77]). We refer to the bibliography for more details on both approaches.

On a purely formal level (the manipulations are next justified using one of the above approaches), the result can be obtained using the technique of *two-scale expansion*. We *postulate* that the solution u^ε to (2.1) writes as the following expansion in powers of ε :

$$u^\varepsilon(x) = u_0(x, \frac{x}{\varepsilon}) + \varepsilon u_1(x, \frac{x}{\varepsilon}) + \varepsilon^2 u_2(x, \frac{x}{\varepsilon}) + ..., \tag{2.20}$$

where, at each order k, the function u_k is assumed to depend on two arguments : a macroscopic variable x and a microscopic variable $\frac{x}{\varepsilon}$. The dependence upon this latter variable is assumed periodic, that is

$$y \longrightarrow u_k(x, y) \text{ is periodic with periodic cell } Q = [0, 1]^d. \tag{2.21}$$

The approach next consists in inserting expansion (2.20) in (2.1) and equating powers in ε. The calculation is tedious but not difficult. All that needs to be borne in mind is the derivation rule :

$$\nabla \left(v(x, \frac{x}{\varepsilon}) \right) = (\nabla_x v)(x, y) + \frac{1}{\varepsilon}(\nabla_y v)(x, y), \text{ where } y = \frac{x}{\varepsilon}. \tag{2.22}$$

In the periodic setting, (2.2) and (2.3) are obtained. They define the homogenized problem (2.4).

The general theory The periodic setting we have considered so far is a simple setting. In fact, the convergence results established above can be extended to a more general setting. It usually comes at the price of losing the "explicitness", that is, we do not necessarily have the characterization of the limit matrix by formulae of the type (2.2)-(2.3). The general result reads :

Proposition 2.2.1 *Let $\mathcal{D}$ be a bounded domain in $\mathbb{R}^d$, and let A^ε be a sequence of invertible matrices with entries in $L^\infty(\mathbb{R}^d)$, satisfying $A^\varepsilon \geq c_1 Id$ and $(A^\varepsilon)^{-1} \geq c_2 Id$ (in the sense $\forall x \in \mathbb{R}^d$, $(A^\varepsilon x, x) \geq c_1 \|x\|^2$, and likewise for $(A^\varepsilon)^{-1}$) for two constants $c_i > 0$ independent from ε. Then, there exists a* homogenized *matrix $A^\star$ satisfying the same properties as A^ε and a subsequence $A^{\varepsilon'}$ of the original sequence A^ε so that, for all functions $f \in H^{-1}(\mathcal{D})$, the function u^ε solution in $H^1_0(\mathcal{D})$ to*

$$-\text{div}\,(A^\varepsilon \nabla u^\varepsilon) = f \tag{2.23}$$

converges in the following sense

$$u^{\varepsilon'} \rightharpoonup u^\star \,,\; A^{\varepsilon'} \nabla u^{\varepsilon'} \rightharpoonup A^\star \nabla u^\star \,,\; A^{\varepsilon'} \nabla u^{\varepsilon'} \cdot \nabla u^{\varepsilon'} \rightharpoonup A^\star \nabla u^\star \cdot \nabla u^\star \tag{2.24}$$

respectively in weak-$H^1_0(\mathcal{D})$, weak-$L^2(\mathcal{D})$ and in the distribution sense. In addition,

$$\int_{\mathcal{D}} A^{\varepsilon'}\nabla u^{\varepsilon'}\cdot\nabla u^{\varepsilon'}\,dx \longrightarrow \int_{\mathcal{D}} A^\star\nabla u^\star\cdot\nabla u^\star\,dx,$$

where $u^\star \in H^1_0(\mathcal{D})$ solves

$$-\mathrm{div}\,(A^\star\nabla u^\star) = f. \tag{2.25}$$

This proposition is part of the *theory of H-convergence*, by F. Murat and L. Tartar in the 1970s, a theory that generalizes the *theory of G-convergence*, by S. Spagnolo, which was restricted to the case of *symmetric* operators.

The main two points of the result are

- that the matrix $A^\star$ and the subsequence ε' do *not* depend on the right-hand side f : in the terminology of Mechanics, an homogenized "material" exists, the same whatever the load ;
- that the homogenized equation obtained is of the same form as the original equation (this is not evident, and many counter examples exist for other settings, we refer to the bibliography).

In contrast, the major weakness of the result is that *no explicit* expression of $A^\star$ is provided. The periodic context is one possible context to obtain explicit expressions. The stationary ergodic context presented in Section 2.2.2 is another one, although the explicitness of the result is more questionable *computationally*, as will be seen below.

2.2.2 Basic theory of stochastic homogenization

Mathematical setting The present section introduces the classical *stationary ergodic setting*. We choose to present the theory in a *discrete* stationary setting, which is more appropriate for our specific purpose in the next sections. Random homogenization is more often presented in the *continuous* stationary setting. This is only a matter of small modifications. We refer to the bibliography for the latter.

Throughout these notes, $(\Omega,\mathcal{F},\mathbb{P})$ denotes a probability space. For a random variable $X\in L^1(\Omega,d\mathbb{P})$, we denote by $\mathbb{E}(X)=\int_\Omega X(\omega)d\mathbb{P}(\omega)$ its expectation value. We assume that the group $(\mathbb{Z}^d,+)$ acts on Ω. We denote by $(\tau_k)_{k\in\mathbb{Z}^d}$ this action, and assume that it preserves the measure $\mathbb{P}$, i.e

$$\forall k\in\mathbb{Z}^d,\quad \forall A\in\mathcal{F},\quad \mathbb{P}(\tau_k A)=\mathbb{P}(A). \tag{2.26}$$

We assume that τ is *ergodic*, that is,

$$\forall A\in\mathcal{F},\quad \big(\forall k\in\mathbb{Z}^d,\quad \tau_k A = A\big)\Rightarrow(\mathbb{P}(A)=0\quad\text{or}\quad 1). \tag{2.27}$$

In addition, we define the following notion of stationarity : any $F\in L^1_{\mathrm{loc}}\left(\mathbb{R}^d,L^1(\Omega)\right)$ is said to be *stationary* if

$$\forall k\in\mathbb{Z}^d,\quad F(x+k,\omega)=F(x,\tau_k\omega)\text{ almost everywhere in }x,\text{ almost surely.} \tag{2.28}$$

In this setting, the ergodic theorem [61, 89] can be stated as follows :

Theorem 2.2.1 *Let $F \in L^\infty\left(\mathbb{R}^d, L^1(\Omega)\right)$ be a stationary random variable in the sense of (2.28). For $k = (k_1, k_2, \dots k_d) \in \mathbb{R}^d$, we set $|k|_\infty = \sup_{1\leq i\leq d} |k_i|$. Then*

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \leq N} F(x, \tau_k \omega) \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left(F(x,\cdot)\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.} \tag{2.29}$$

This implies that (denoting by Q the unit cube in $\mathbb{R}^d$)

$$F\left(\frac{x}{\varepsilon}, \omega\right) \overset{\star}{\underset{\varepsilon\to 0}{\rightharpoonup}} \mathbb{E}\left(\int_Q F(x,\cdot)dx\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.} \tag{2.30}$$

For illustration, it is useful to intuitively define stationarity and ergodicity in terms of materials modeling. Pick, at the microscale in the material, two points x and $y \neq x$, differing from x from an integer shift k. The particular local environment seen from x (that is, the microstructure present at x) is generically different from what is seen from y (that is, the microstructure present at y). However, the *average* local environment in x is identical to that in y (considering the various realizations of the random material). In mathematical terms, the *law* of microstructures is the same at all points. This is *stationarity*. On the other hand, *ergodicity* means that considering all the points in the material amounts to fixing a point x in this material and considering all the possible microstructures present there.

Main result We now fix $\mathcal{D}$ an open, smooth, bounded subset of $\mathbb{R}^d$, and a square matrix A of size d, which is assumed stationary in the sense defined above, and which is assumed to enjoy the classical assumptions of uniform ellipticity and boundedness. Then we consider the boundary value problem

$$\begin{cases} -\text{div}\left(A\left(\frac{\cdot}{\varepsilon}, \omega\right)\nabla u^\varepsilon\right) = f \quad \text{in} \quad \mathcal{D}, \\ \\ u^\varepsilon = 0 \quad \text{on} \quad \partial\mathcal{D}. \end{cases} \tag{2.31}$$

Observe that when A does not depend on ω, then, given (2.28), it is periodic and we recover the periodic problem (2.1) introduced above. Standard results of stochastic homogenization [16, 58] apply (see an outline of the proof below) and allow to find the homogenized problem for problem (2.31). These results generalize the periodic results recalled in Subsection 2.2.1. The solution u^ε to (2.31) converges to the solution to (2.4) where the homogenized matrix is now defined as :

$$[A^\star]_{ij} = \mathbb{E}\left(\int_Q \left(e_i + \nabla w_{e_i}(y,\cdot)\right)^T A\left(y,\cdot\right) e_j \, dy\right), \tag{2.32}$$

where for any $p \in \mathbb{R}^d$, w_p is the solution (unique up to the addition of a (random) constant) in $\left\{ w \in L^2_{\rm loc}(\mathbb{R}^d, L^2(\Omega)), \quad \nabla w \in L^2_{\rm unif}(\mathbb{R}^d, L^2(\Omega)) \right\}$ to

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A\left(y,\omega\right)\left(p+\nabla w_p(y,\omega)\right)\right] = 0, \quad \text{a.s. on} \quad \mathbb{R}^d \\ \\ \nabla w_p \quad \text{is stationary in the sense of (2.28),} \\ \\ \mathbb{E}\left(\displaystyle\int_Q \nabla w_p(y,\cdot)\, dy \right) = 0. \end{array} \right. \tag{2.33}$$

We have used above the notation $L^2_{\rm unif}$ for the *uniform* L^2 space, that is the space of functions for which, say, the L^2 norm on a ball of unit size is bounded above independently from the center of the ball.

A striking difference between the stochastic setting and the periodic setting can be observed comparing (2.2) and (2.33). In the periodic setting, the corrector problem is posed on a bounded domain, namely the periodic cell Q. In sharp contrast, the corrector problem (2.33) of the random setting is posed on the whole space $\mathbb{R}^d$, and cannot be reduced to a problem posed on a bounded domain. The reason is, condition $\mathbb{E}\left(\displaystyle\int_Q \nabla w_p(y,\cdot)\, dy \right) = 0$ in (2.33) is a *global* condition. It indeed equivalently reads, because of the ergodic theorem, a.s. $-\lim_{R\longrightarrow+\infty} \dfrac{1}{|B_R|} \displaystyle\int_{B_R} \nabla w_p(y,\cdot)\, dy = 0$ for any sequence of balls B_R of radii R. The fact that the random corrector problem is posed on the entire space has far reaching consequences for numerical practice. Truncations of problem (2.33) have to be considered, and the actual homogenized coefficients are only obtained in the asymptotic regime. This will be the main source of all the *practical* difficulties we will describe, and try to overcome, in the following sections.

Elementary analysis As in the periodic case, it is instructive to consider the "zero-dimensional" and one-dimensional settings. This allows one to highlight some important differences between the periodic and the stationary ergodic settings. Besides the little technicalities mentioned here, the reader should bear in mind that the outcome of the discussion is that the random context for homogenization is incredibly much richer than the periodic context. To some extent, random homogenization is *generic*.

Deleting all the differential operators in (2.31) immediately shows that the question is, again, to identify the weak limit, as ε vanishes, of the sequence $b(\frac{x}{\varepsilon},\omega)$ for b a stationary ergodic function. The ergodic theorem above states that the almost sure weak limit is the expectation value $\mathbb{E}(\int_0^1 b)$.

Moving on to the one-dimensional setting

$$\left\{ \begin{array}{rcll} -\dfrac{d}{dx}(a(\dfrac{\cdot}{\varepsilon},\omega)\dfrac{d}{dx}u^\varepsilon) & = & f & \text{in }]0,1[\\ \\ u^\varepsilon(0,\omega) = u^\varepsilon(1,\omega) & = & 0, & \end{array} \right. \tag{2.34}$$

we of course can exploit our computations in the periodic case. On the one hand, we know the exact solution u^ε by (2.10)-(2.11), the formula holding almost surely. On the other hand, we may solve the corrector problem

$$\left\{ \begin{array}{ll} -\dfrac{d}{dy}(a(y,\omega)(1+\dfrac{d}{dy}w(y,\omega))) = 0, & \text{in } [0,1], \\ \qquad\qquad w' \text{ stationary} & \mathbb{E}(\int_0^1 w') = 0 \end{array} \right.$$

and find, similarly to (2.17) that, up to an additive constant,

$$w(y,\omega) = -y + \left(\mathbb{E}(\int_0^1 a^{-1})\right)^{-1} \int_0^y a^{-1}(x,\omega)\,dx. \tag{2.35}$$

This of course yields the value $a^\star = \left(\mathbb{E}(\int_0^1 a^{-1})\right)^{-1}$ for the homogenized coefficient appearing in the homogenized equation, that is the one-dimensional version of (2.4). As in the periodic case, it is elementary to prove that u^ε converges to the solution $u^\star$ of the latter equation. The convergence holds weakly in H^1 and, this time, almost surely in ω. Including the corrector allows for a (almost sure in ω) strong convergence in H^1 of $u^\varepsilon(x,\omega) - \big(u^\star(x) + \varepsilon\,(u^\star)'(x)\,w(\frac{x}{\varepsilon},\omega)\big)$ to zero.

Although the stationary ergodic setting shares many properties of the periodic setting, even the consideration of the one-dimensional simple situation allows to exhibit some striking differences. We have already seen that a major difference is that the corrector problem is posed on the entire space, and not on a bounded domain (the periodic cell). We now mention two additional differences : the nature of the corrector and the rate of convergence.

It is easily remarked on (2.35) that the corrector w itself is *not* a stationary function. Only its first derivative $w'(x,\omega) = -1 + \left(\mathbb{E}(\int_0^1 a^{-1})\right)^{-1} a^{-1}(x,\omega)$ is. In the periodic setting, both functions are periodic. This is in fact related to the following phenomenon, not restricted to the one-dimensional setting. If ∇w is a periodic function that has zero mean, then w is a periodic function (and, of course, conversely). In contrast, if ∇w is a stationary function with $\mathbb{E}(\int_Q \nabla w) = 0$, then we do not necessarily have w stationary, but we only can claim that w is *sublinear at infinity*, that is $(1+|x|)^{-1}\,w$ converges to zero as $|x| \longrightarrow +\infty$, almost surely in ω (this is a consequence of the ergodic theorem, and can be easily verified *e.g.* using (2.35)). Even though the homogenized matrix involves only the gradient of the corrector, and not the corrector itself, this observation complicates the situation.

In the one-dimensional setting, we have the explicit expressions of u^ε and $u^\star$, which are of course valid, only changing the notation (the average is replaced by the expectation value, and the dependence upon ω is possibly indicated), both in the periodic and the stationary ergodic settings. We are thus able to explicitly express the rate of the convergence. In the periodic setting, we have seen in (2.18) the explicit expression of the difference $(u^\varepsilon)' - \left(u^\star + \varepsilon\,(u^\star)'\,w(\frac{\cdot}{\varepsilon})\right)'$. We easily deduce that

$$u'_\varepsilon(x) - (1 + w'\left(\frac{x}{\varepsilon}\right))\,(u^\star)'(x) = -\,(c_\varepsilon - c_\star)\;a(\frac{x}{\varepsilon})^{-1}, \tag{2.36}$$

where c_ε is, we recall, defined in (2.11) and where we denote by $c_\star = -\displaystyle\int_0^1 F$. It follows that, in the periodic case, the difference scales, in L^∞-norm (and thus L^2-norm), as ε :

$$(u^\varepsilon)' - \left(1 + w'(\frac{\cdot}{\varepsilon})\right)\,(u^\star)' = O(\varepsilon). \tag{2.37}$$

The exact same expression for the stochastic case shows, using a central limit theorem, that

$$\sqrt{\varepsilon}^{-1}\left[(u^\varepsilon)' - \left(1 + w'(\frac{\cdot}{\varepsilon})\right)\,(u^\star)'\right] \text{ converges in law to a Gaussian,} \tag{2.38}$$

(up to the multiplication by the function $a(\frac{\cdot}{\varepsilon})^{-1}$ which, almost surely, weakly converges to $(a^\star)^{-1}$). We note that for a central limit theorem to hold, we need to assume more than ergodicity. The appropriate setting is to assume *mixing*, which is a condition ensuring that correlations become sufficiently small at large distance. In this simplified presentation, let us only mention that if the coefficients $a(x,\omega)$ and $a(y,\omega)$ are independent for x and y not in the same unit cell, then mixing holds. Likewise, considering now the functions and not their first derivatives, two straightforward calculations show that in the periodic setting

$$u^\varepsilon - u^\star = O_{L^2}(\varepsilon), \tag{2.39}$$

while in the stochastic setting

$$\sqrt{\varepsilon}^{-1}\,[u^\varepsilon - u^\star] \text{converges in law.} \tag{2.40}$$

A striking contrast between the periodic and the stationary setting is the above different scalings. We refer to [13, 22] for more details on this topic, and also to [66] for a similar analysis, on a variant of the classical theory of stochastic homogenization (this variant will be presented in Subsection 2.3.2 below).

Some elements of theory In dimensions higher than one, no explicit expression for the functions manipulated is available. As in the periodic case, there is almost no exactly solvable model either. The only famous exception is the so-called *random checkerboard* : the (scalar) coefficient appearing in (2.31) is defined piecewise

constant and randomly takes values α and β, each with probability 1/2. All random variables are independent. Then (see [58],p 236) the homogenized matrix is given by $A^\star = \sqrt{\alpha\beta}\ Id$.

In full generality, delicate arguments have to be conducted. In our simple linear elliptic setting, the arguments can however be kept elementary. We outline here the proof of the existence of the corrector, that is the solution to (2.33). As mentioned above, the difficulty is that we need to work on the entire space $\mathbb{R}^d$. The idea is to lift all the differential calculus on the abstract probability space, and use the variational theory (Lax-Milgram Lemma and related notions). But, even then, a remaining difficulty is the absence of any Poincaré inequality. To circumvent the difficulty, we need to *approximate* problem (2.33). This is typically performed introducing the following auxiliary problem : *search for a stationary function* $w_{p,\eta}$ *such that*

$$-\mathrm{div}\left[A\left(y,\omega\right)\left(p+\nabla w_{p,\eta}(y,\omega)\right)\right]+\eta\, w_{p,\eta}=0, \tag{2.41}$$

on the whole space , for $\eta>0$ *supposedly small*. Equation (2.33) is obtained passing to the limit of vanishing η (note that the approximating term $\eta\, w_{p,\eta}$ vanishes because it reads $\sqrt{\eta}$ times a bounded function). The condition $\mathbb{E}\left(\int_Q \nabla w_p(y,\cdot)\right)=0$ is also easily obtained from the weak convergence and $\mathbb{E}\left(\int_Q \nabla w_{p,\eta}(y,\cdot)\right)=0$ which holds because $w_{p,\eta}$ is itself stationary. We note that, as expected, no stationarity is known on w_p itself.

Other random problems such as nonlinear variational problems and, foremost, nonlinear non variational problems, are substantially more difficult. Many mathematical questions are still unsolved. We refer to [68, 91] for examples of important contributions on various issues.

2.2.3 Bounds and numerical approach

2.2.3.1 Direct approach

Practical approximations of the homogenized problem in random homogenization are not easily obtained, and can even be prohibitively expensive computationally. The major difficulty owes to the fact that the corrector problem (2.33) is set on the entire space, and not on a bounded unit cell as is the case in periodic homogenization. In practice, the matrix $A^\star$ is approximated by the matrix

$$\left[A_N^\star\right]_{ij}(\omega)=\frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^N(y,\omega)\right)^T A(y,\omega)\left(e_j+\nabla w_{e_j}^N(y,\omega)\right)\,dy \tag{2.42}$$

which is in turn obtained by solving the corrector problem on a *truncated* domain, say the cube $Q_N\subset\mathbb{R}^d$ of size $(2N+1)^d$ centered at the origin :

$$\left\{\begin{array}{l} -\mathrm{div}\,\left(A(\cdot,\omega)\left(p+\nabla w_p^N(\cdot,\omega)\right)\right)=0 \quad \text{on} \quad \mathbb{R}^d,\\ w_p^N(\cdot,\omega)\,\text{is}\,Q_N\text{-periodic.}\end{array}\right. \tag{2.43}$$

Although $A^\star$ itself is a deterministic object, its practical approximation $A_N^\star$ is random. It is only in the limit of infinitely large domains Q_N that the deterministic value is attained. Theoretically, convergence is obtained using an easy translation of Theorem 1 in [23] from the continuous stationary setting to our discrete stationary setting. At fixed N, a set of M independent realizations of the random coefficient A are therefore considered. The corresponding truncated problems (2.43) are solved, and an empirical mean of the truncated coefficients (2.42) is inferred. This empirical mean only agrees with the theoretical mean value of the truncated coefficient within a margin of error which is given by the central limit theorem (in terms of M). The variance of the coefficients therefore plays a role. For a sufficiently large truncation size N, this truncated value is admittedly the exact value of the coefficient. Of course, the overall computation described above is expensive, because each realization requires a new solution to the d-dimensional boundary value problem (2.43) of presumably large a size since N is taken large. There is therefore a huge interest in reducing the cost of the computation, or, otherwise stated, in reaching a better accuracy at a given computational cost. Since the variance of the truncated homogenized matrix is an important ingredient, reducing the variance becomes a challenging and sensitive issue. Reducing this variance is the problem we will consider in Section 2.4 below.

For now, let us be slightly more explicit regarding the computational approach. Let $(A^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq M}$ denote M independent and identically distributed random fields. We define a family $(A_N^{\star,\mathbf{m}})_{1\leq \mathbf{m}\leq M}$ of i.i.d. homogenized matrices by, for any $1\leq i,j\leq d$,

$$[A_N^{\star,\mathbf{m}}]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} \left(e_i+\nabla w_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T A^{\mathbf{m}}(\cdot,\omega)\ \left(e_j+\nabla w_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right),$$

where $w_{e_j}^{N,\mathbf{m}}$ is the solution of the corrector problem associated to $A^{\mathbf{m}}$. Then we define for each component of $A_N^\star$ the empirical mean and variance

$$\begin{aligned} \mu_M\left([A_N^\star]_{ij}\right) &= \frac{1}{M}\sum_{\mathbf{m}=1}^{M}[A_N^{\star,\mathbf{m}}]_{ij}\,, \\ \sigma_M\left([A_N^\star]_{ij}\right) &= \frac{1}{M-1}\sum_{\mathbf{m}=1}^{M}\left([A_N^{\star,\mathbf{m}}]_{ij}-\mu_M\left([A_N^\star]_{ij}\right)\right)^2. \end{aligned} \tag{2.44}$$

Since the matrices $A_N^{\star,\mathbf{m}}$ are i.i.d., the strong law of large numbers applies :

$$\mu_M\left([A_N^\star]_{ij}\right)(\omega) \underset{M\to+\infty}{\longrightarrow} \mathbb{E}\left([A_N^\star]_{ij}\right) \text{ almost surely.}$$

The central limit theorem then yields

$$\sqrt{M}\left(\mu_M\left([A_N^\star]_{ij}\right)-\mathbb{E}\left([A_N^\star]_{ij}\right)\right) \underset{M\to+\infty}{\xrightarrow{\mathcal{L}}} \sqrt{\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)}\,\mathcal{N}(0,1), \tag{2.45}$$

where the convergence holds in law, and $\mathcal{N}(0,1)$ denotes the standard gaussian law. Introducing its 95 percent quantile, it is standard to consider that the exact mean $\mathbb{E}\left([A_N^\star]_{ij}\right)$ is equal to $\mu_M\left([A_N^\star]_{ij}\right)$ within a margin of error $1.96\dfrac{\sqrt{\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}$. The exact variance $\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)$ being unknown in practice, it is customary to replace it by the empirical variance given in (2.44) above. It is therefore considered that the expectation $\mathbb{E}\left([A_N^\star]_{ij}\right)$ lies in the interval

$$\left[\mu_M\left([A_N^\star]_{ij}\right) - 1.96\frac{\sqrt{\sigma_M\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}\,,\ \mu_M\left([A_N^\star]_{ij}\right) + 1.96\frac{\sqrt{\sigma_M\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}\right]. \quad (2.46)$$

The value $\mu_M\left([A_N^\star]_{ij}\right)$ is thus, for both M and N sufficiently large, adopted as the approximation of the exact value $[A^\star]_{ij}$.

Remark 2.2.1 *In fact, the situation considered here is simple : it is the linear elliptic case. It is well known that the difficulties we mention for the random setting already arise in the periodic setting when the operator is, for instance, nonlinear. Then determining the periodic homogenized problem cannot always be reduced to a simple computation on one single periodic cell of the problem. Similarly to the random context, the question arises to speed up the convergence in these types of periodic problems that give rise to corrector problems set on the whole space. We refer to [18, 46] for more details on possible techniques.*

Given the above computational workload, practitioners, especially scientists from the applied communities (Mechanics, ...) sometimes choose to avoid computing actual homogenized equations and concentrate on *bounds* on the homogenized matrices $A^\star$. It is useful to briefly present the approach and its (strong, as will be seen below) limitations.

2.2.3.2 Bounds for homogenization

For the specific case of two-phase composite materials, some bounds on the homogenized coefficients may be established. In some situations these bounds may be useful for a first qualitative evaluation of the properties of the material. We briefly present in this section one example of such bounds (actually the most famous one). The case we consider is a scalar equation of the type (2.1) with a matrix coefficient A that needs not be periodic, nor stationary ergodic. Obtaining estimates on $A^\star$ without being in position to explicitly compute $A^\star$ at a reasonable computational price is the whole interest of the approach by "bounds".

The materials we consider are composite materials consisting of only two phases that we denote by $\mathcal{A}$ and $\mathcal{B}$. The problem is to find all possible homogenized materials, that is, mathematically, matrices $A^\star$, that can be attained homogenizing such materials $\mathcal{A}$ and $\mathcal{B}$. More precisely, let us denote the volume fraction of material $\mathcal{A}$ at point x by $\chi_\varepsilon(x)$, and thus the volume fraction of material $\mathcal{B}$ by $1-\chi_\varepsilon(x)$. The matrix coefficient in (2.1) therefore reads $A^\varepsilon(x) = \chi_\varepsilon(x)\, A + (1-\chi_\varepsilon(x))\, B$, where, with obvious notation, A is the matrix coefficient of material $\mathcal{A}$, and B that of $\mathcal{B}$. We assume that both A and B are symmetric matrices and belong to the set of suitable matrices

$$\mathcal{M}_{\alpha,\beta} = \left\{ M \in \mathbb{R}^{d\times d}, \text{ such that } \forall \xi \in \mathbb{R}^d,\ \alpha\,|\xi|^2 \leq M\xi\,.\xi \leq \beta\,|\xi|^2 \right\},$$

for some positive scalar constants α, β fixed. We define by $\mathcal{M}^s_{\alpha,\beta}$ the subspace of $\mathcal{M}_{\alpha,\beta}$ of symmetric matrices. We assume that χ_ε has a weak-$\star$ limit in $L^\infty(\mathcal{D},[0,1])$ and denote by $\theta(x)$ this weak limit. We next introduce the *G-closure* set, denoted by $\mathcal{G}_\theta$, which models all possible homogenized materials that have the appropriate *average* volume fraction θ. More precisely, $\mathcal{G}_\theta$ is defined as the set of all matrices $A^\star \in L^\infty(\mathcal{D}, \mathcal{M}^s_{\alpha,\beta})$ for which there exists some sequence χ_ε that weakly-$\star$ converges to the function $\theta \in L^\infty(\mathcal{D},[0,1])$ and for which A^ε yields $A^\star$ as homogenized matrix. The latter assertion is formalized by the notion of *H-convergence* (a notion actually already manipulated in Proposition 2.2.1 above) : for all right-hand side $f \in H^{-1}(\mathcal{D})$ inserted in (2.1), the solution u^ε weakly converges in $H^1_0(\mathcal{D})$ to $u^\star$, and $A^\varepsilon \nabla u^\varepsilon$ weakly converges in $L^2(\mathcal{D})^d$ to $A^\star \nabla u^\star$, where $u^\star$ solves (2.4).

A crucial property is that, in fact, $\mathcal{G}_\theta$ is essentially obtained using only periodic matrices. More precisely, let us denote by P_θ, for θ a scalar in $[0,1]$, the set of symmetric matrices obtained homogenizing, in a periodic way, A and B with fixed volume fraction θ, that is the set of matrices $A^\star$ writing, for $\xi \in \mathbb{R}^d$,

$$A^\star\xi.\xi = \mathrm{Inf}_{w\in H^1_\#(Q)} \int_Q \left(\chi(y)\, A + (1-\chi(y)\, B\right)(\xi + \nabla w(y)).(\xi + \nabla w(y)))\ dy$$

where the function χ is Q-periodic, with mean $\displaystyle\int_Q \chi = \theta$, and $H^1_\#(Q)$ denotes the set of H^1 Q-periodic functions. Then, we have that, for any function $\theta \in L^\infty(\mathcal{D},[0,1])$,

$$\mathcal{G}_\theta = \left\{ A^\star \in L^\infty(\mathcal{D}, \mathcal{M}^s_{\alpha,\beta}) \quad \text{s.t.} \quad A^\star(x) \in \overline{P_{\theta(x)}} \text{ a.e. in } \mathcal{D} \right\}.$$

The Hashin-Shtrikman bounds Based on the density of the matrices obtained by periodic homogenization in the set of matrices obtained by arbitrary homogenization, it is possible to derive bounds on the elements of $\mathcal{G}_\theta$. These bounds are the *Hashin-Shtrikman bounds*, and read as follows.

Let θ be a scalar in $[0,1]$ and assume $B \geq A$ in the sense of symmetric matrices. Any matrix $A^\star \in \overline{P}_\theta$ satisfies the upper bound :

$$\forall p \in \mathbb{R}^d,\ A^\star p\cdot p \leq Bp\cdot p + \theta \min_{\eta\in\mathbb{R}^d} \left[2p\cdot\eta + (B-A)^{-1}\eta\cdot\eta + (1-\theta)h(\eta)\right], \quad (2.47)$$

where $h(\eta)$ is defined by

$$h(\eta) = \min_{k\in\mathbb{Z}^d, k\neq 0} \frac{|\eta\cdot k|^2}{Bk\cdot k}.$$

Similarly, any matrix $A^\star \in \overline{P}_\theta$ satisfies the lower bound

$$A^\star p\cdot p \geq Ap\cdot p + (1-\theta)\max_{\eta\in\mathbb{R}^d}\left[2p\cdot\eta - (B-A)^{-1}\eta\cdot\eta - \theta g(\eta)\right], \tag{2.48}$$

where $g(\eta)$ is defined by

$$g(\eta) = \max_{k\in\mathbb{Z}^d, k\neq 0} \frac{|\eta\cdot k|^2}{Ak\cdot k}.$$

Furthermore, this upper bound can always be attained *i.e.* : for all $p\in\mathbb{R}^d$, there exists a function χ, Q-periodic and that generally depends on p, such that for the matrix $A_p^\star$ obtained by periodic homogenization of

$$A(\frac{x}{\varepsilon}) = \chi(\frac{x}{\varepsilon})A + (1-\chi(\frac{x}{\varepsilon}))B,$$

the inequality (2.47) becomes an equality (see *e.g.* [95]). Likewise, the lower bound (2.48) can always be attained.

Hashin-Shtrikman Variational Principle Let us briefly show how to obtain (2.48) for $A^\star \in \overline{P}_\theta$ obtained by periodic homogenization of the matrix $A(\frac{x}{\varepsilon}) = \chi(\frac{x}{\varepsilon})A+(1-\chi(\frac{x}{\varepsilon}))B$. The proof is a two step approach. First we show that solving the corrector problem is equivalent to solving the so-called Hashin-Shtrikman variational problem (or Hashin-Shtrikman variational principle). The second step consists in choosing an appropriate test function in the variational problem, so that analytical computations can be performed and yield the so-called Hashin-Shtrikman bounds.

We first recast the corrector problem (2.2) as

$$\left\{\begin{array}{ll} \text{Find } u \text{ such that :} & \\ -\mathrm{div}(A(y)\nabla u(y)) = 0 & \text{in } \mathbb{R}^d, \\ u(y) = p\cdot y + v_{per}(y) & \text{in } \mathbb{R}^d, \end{array}\right. \tag{2.49}$$

where v_{per} is Q-periodic, $v_{per}\in H^1(Q)$. We now introduce an arbitrary homogeneous, symmetric, positive matrix A_0 and rewrite (2.49) as

$$\left\{\begin{array}{ll} -\mathrm{div}(A_0\nabla u(y)) = \mathrm{div}((A(y)-A_0)\nabla u(y)) & \text{in } \mathbb{R}^d \\ u(y) = p\cdot y + v_{per}(y) & \text{in } \mathbb{R}^d, \\ v_{per} \text{ is } Q\text{-periodic.} & \end{array}\right. \tag{2.50}$$

Using the linearity of (2.50), its solution u satisfies

$$u = u_0 + u_{\underline{\tau}}, \tag{2.51}$$

where u_0 and $u_{\underline{\tau}}$ are the solutions (unique up to the addition of a constant) to

$$\left\{ \begin{array}{ll} -\mathrm{div}(A_0 \nabla u_0(y)) = 0 & \text{in } \mathbb{R}^d \\ u_0(y) = p \cdot y + v^0_{per}(y) & \text{in } \mathbb{R}^d,\ v^0_{per} \text{ is } Q\text{-periodic,} \end{array} \right.$$

and

$$\left\{ \begin{array}{l} -\mathrm{div}(A_0 \nabla u_{\underline{\tau}}(y)) = \mathrm{div}(\underline{\tau}(y)) \quad \text{in } \mathbb{R}^d \\ u_{\underline{\tau}} \text{ is } Q\text{-periodic,} \end{array} \right.$$

where

$$\underline{\tau}(y) := (A(y) - A_0)\nabla u(y) \tag{2.52}$$

is the so-called polarization field. We denote by Γ the linear operator such that

$$\nabla u_{\underline{\tau}}(y) = -\left(\Gamma \underline{\tau}\right)(y). \tag{2.53}$$

Taking the gradient of (2.51), we infer from (2.52) and (2.53) that

$$\nabla u_0(y) = \left(A(y) - A_0\right)^{-1} \underline{\tau}(y) + \left(\Gamma \underline{\tau}\right)(y). \tag{2.54}$$

As Γ is a self-adjoint operator, the equation (2.54) is the Euler-Lagrange equation of the variational problem

$$\inf_{\tau \in L^2(Q)^d} \{J(\tau)\}$$

where

$$\begin{aligned} J(\tau) &= \frac{1}{2}\left(\int_Q \left(\tau(y) \cdot (A(y) - A_0)^{-1} \tau(y) + \tau(y) \cdot (\Gamma\tau)(y) \right) dy \right) \\ &\quad - \int_Q \tau(y) \nabla u_0(y) dy. \end{aligned}$$

A tedious but straightforward computation yields

$$J(\underline{\tau}) = \frac{1}{2} A_0 p \cdot p - \frac{1}{2} A^\star p \cdot p. \tag{2.55}$$

Consider now the choice $A_0 = (1 - s)\, A$, with $s > 0$ an arbitrary small positive constant. The matrix $A(y) - A_0$ is then positive definite a.e. on $\mathbb{R}^d$. In that case, J turns out to be convex, thus

$$\forall \tau \in \left(L^2(Q)\right)^d \ Q\text{-periodic},\ J(\tau) \geq J(\underline{\tau}). \tag{2.56}$$

Choosing $\widetilde{\tau}(y) = (1 - \chi(y))\, \eta$, where $\eta \in \mathbb{R}^d$, the value $J(\widetilde{\tau})$ is shown to satisfy

$$J(\widetilde{\tau}) \leq (1 - \theta) \left(\frac{1}{2} (B - A + sA)^{-1} \eta \cdot \eta + \frac{1}{2} \theta \max_{k \in \mathbb{Z}^d, k \neq 0} \left(\frac{|\eta \cdot k|^2}{Ak \cdot k} \right) - \eta \cdot p \right). \tag{2.57}$$

Collecting (2.55), (2.56) and (2.57), and letting s go to 0, we obtain the bound (2.48).

Other estimates Besides the Hashin-Shtrikman bounds, many other estimates have been proposed, such as the dilute approximation, the self-consistent method [102] and the Mori Tanaka methods [73]. They are all based on the fact that the problem of a single inclusion in an infinite material (Eshelby problem) is analytically solvable [41]. Similarly to the Hashin-Shtrikman bounds, the spatial distribution of the phases is not taken into account in these other bounds. The accuracy of these estimates and bounds strongly depends on the contrast between A and B and the volume fraction θ as shown on Figure 2.1 below.

Numerical illustration We consider a two-phase composite with A and B. We denote by a the scalar conductivity of A (respectively b the conductivity of B) with $a < b$. We denote by d the dimension, and by θ the volume fraction of A.

We consider the case of the random checkerboard, for which the exact homogenized matrix is known : $A^\star = \sqrt{ab}\ \mathrm{Id}$. In this simple case, the different bounds and estimates have an analytical form : the homogenized coefficient $a^\star$ is bounded from below by the harmonic mean (often called the Reuss bound) and from above by the arithmetic mean (often called the Voigt bound)

$$\frac{1}{\theta/a + (1-\theta)/b} \leq a^\star \leq \theta a + (1-\theta)b.$$

These bounds are also called Wiener Bounds or Paul bounds. In this case, the Hashin-Shtrikman bounds detailed above read (see *e.g.* [58], p. 188)

$$a\left(1 + \frac{d(1-\theta)(b-a)}{da + \theta(b-a)}\right) \leq a^\star \leq b\left(1 - \frac{d\theta(b-a)}{db + (1-\theta)(a-b)}\right),$$

and the Self-Consistent model leads to an estimate λ_{eff} of the effective conductivity $a^\star$ implicitly defined (see [59]) by

$$\theta\frac{a - \lambda_{\text{eff}}}{a + 2\lambda_{\text{eff}}} + (1-\theta)\frac{b - \lambda_{\text{eff}}}{b + 2\lambda_{\text{eff}}} = 0.$$

On Figure 2.1 we plot these bounds and estimates for different values of the contrast, defined by $\frac{b}{a}$, for $a = 1$. Note that in this case, by construction, the volume fraction for any a and b is $\theta = \frac{1}{2}$.

In Tab.2.1 we collect the values of all these bounds and estimates, for the particular case $a = 1$ and $b = 10$.

$A^\star$	Harmonic	HS-	SC Model	HS+	Arithmetic
3.16	1.81	2.38	4.00	4.19	5.50

Tab. 2.1 – Values of bounds and estimate for a contrast of $\frac{b}{a} = 10$

We check that, for the critical volume fraction $\theta = 0.5$ and even for a contrast which is not extremely large ($a = 1$ and $b = 10$), the range of achievable homogenized

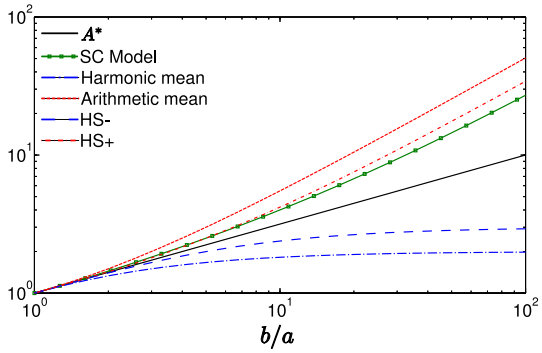


FIG. 2.1 – Different bounds for the checkerboard test case.

matrices, as given by the Hashin-Shtrikman bounds, is wide. In such a case, the spatial distribution of phases, which is not taken into account on the bounds, is certainly important.

2.3 Perturbation of periodic settings and related approaches

As we have just seen, random homogenization, as is, can be extremely, and even prohibitively, computationally expensive. The question arises to know whether the general random context is really relevant and practical, and whether adequate modifications of the general setting can lead to substantial improvements. The present section discusses some options, all extracted from two series of recent works [5, 7–9] and [P1, 19, 20], slightly different in nature. The following two sections 2.3.1 and 2.3.2 discuss these works, respectively. The common denominator is that the two series of works start from the same observation : many actual situations considered as random situations are in fact close, in a vague sense at least, to some periodic situation. The question is how to benefit from this and design some appropriate numerical approaches.

2.3.1 Perturbation theory

As announced above, our line of thought in this section is based on the following two-fold observation : classical random homogenization is costly but perhaps, in a number of situations, not necessary. A careful examination of the prototypical example material shown on Figure 2.2 indeed shows that albeit not periodic, the material is not totally random. It is probably fair to consider it as a perturbation of a periodic material. The homogenized behaviour should expectedly be close to that of the underlying periodic material, up to an error depending on the amount of randomness present.

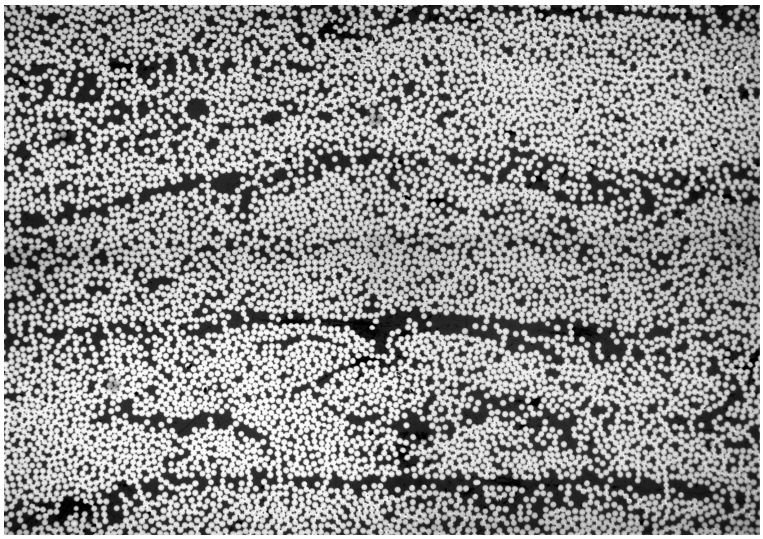


FIG. 2.2 – Two-dimensional cut of a composite material used in the aeronautics industry, extracted from [100] and reproduced with permission of the author. It is clear that this material is not periodic, yet there is some type of underlying periodic arrangement of the fibers.

We introduce and study a specific model for a randomly perturbed periodic material that we term a *weakly random material*. More precisely, we are interested in the homogenization of the following elliptic problem

$$\begin{cases} -\operatorname{div}\left(A_\eta(\dfrac{x}{\varepsilon},\omega)\,\nabla u_\varepsilon\right) = f(x) \text{ in } \mathcal{D}\subset\mathbb{R}^d, \\ u_\varepsilon = 0 \text{ on } \partial\mathcal{D}, \end{cases} \tag{2.58}$$

that is (2.31) where, specifically here,

$$A_\eta(\cdot,\omega) = A_{per} + b_\eta(\cdot,\omega)\,C_{per}. \tag{2.59}$$

Here the matrix A_{per} models a reference $\mathbb{Z}^d$-periodic material which is randomly perturbed by the $\mathbb{Z}^d$-periodic matrix C_{per}, the stochastic perturbation being encoded in the stationary ergodic scalar field b_η. Our purpose is to derive an expansion

$$A_\eta^\star = A_{per}^\star + \eta\tilde{A}_1^\star + \eta^2\tilde{A}_2^\star + o(\eta^2)$$

where $A_\eta^\star$ and $A_{per}^\star$ are the homogenized matrices associated with A_η and A_{per} respectively. The first-order and second-order corrections $\tilde{A}_1^\star$ and $\tilde{A}_2^\star$ will be obtained as the limit when $N\to\infty$ of a sequence of matrices computed on the supercell $[-N/2,N/2]^d$, using, this is exactly the point, only purely deterministic computations.

The above setting is of course *one* possible setting where the theory may be developed. Other forms of random perturbations of periodic problems could also be addressed. Also, we have again deliberately considered the simplest possible equation (a scalar, linear second order elliptic equation in divergence form) to avoid any unnecessary technicalities and fundamental difficulties. Other equations could be considered, although it is not currently clear (to us, at least) how general our theory is in this respect.

The ideas developed here (and originally introduced and further developed in [5, 7–9]) are reminiscent of many previous contributions that have considered perturbative approaches in homogenization. In [96] and [4], a deterministic setting in which an asymptotic expansion is assumed on the properties of the material (the latter being not necessarily periodic) is studied under the name "small amplitude homogenization". In [88], the case of a Gaussian perturbation with a small variance is addressed from a mechanical point of view.

To encode that the perturbation is "small", we assume that the *random* field $b_\eta(x,\omega)$ in (2.59) satisfies

$$\|b_\eta\|_{L^\infty(Q;L^p(\Omega))} \underset{\eta\to 0}{\longrightarrow} 0, \tag{2.60}$$

for some $1 \leq p < \infty$. Condition (2.60) states that the perturbation is small *on average*. However, it does not prevent the perturbation from being large, once in a while, because we only have $p < \infty$. The intuitive image behind the present setting is *'perturb the periodic material only rarely, but then possibly in a way that is* not *small'*. We shall now first expose the necessary mathematical setting (in Section 2.3.1.1) and next proceed to some illustrative numerical tests, in Section 2.3.1.2.

2.3.1.1 Mathematical setting

When the exponent p in (2.60) is strictly larger than one, a classical "expansion" theory can be developed. Assuming that $m_\eta := \|b_\eta\|_{L^\infty(Q;L^p(\Omega))} \longrightarrow 0$ as η vanishes, it may be proved, using an expansion in powers of m_η, and up to the extraction of a subsequence, that the homogenized matrix $A^\star_\eta$ admits a first order expansion in terms of the small 'coefficient' m_η. The coefficients can be expressed using periodic corrector problems built from the matrices A_{per} and C_{per}. The remainder in the expansion can indeed be shown to be $o(m_\eta)$ in a certain sense and under appropriate assumptions. Higher order terms can be obtained likewise. All this will be made precise below (for the prototypical case $p = 2$).

The difficulty is that it is not always the case that the first order term obtained is not trivial (we may typically have $\dfrac{b_\eta}{m_\eta}$ that vanishes in the weak limit), and also that the above argument does not carry over to the case $p = 1$. There is some interest in considering a drastically different situation. As an instance of such a situation,

we will consider the case where b_η follows a Bernoulli law. Let us actually begin by focusing our attention on this case.

A defect-type approach Consider the case where b_η is uniform in each cell of $\mathbb{Z}^d$ and writes

$$b_\eta(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) B_\eta^k(\omega), \tag{2.61}$$

where the B_η^k are independent identically distributed random variables. Their common law is assumed to be a Bernoulli law of parameter η (They all take value 1 with probability η, and take value 0 with probability $1-\eta$). This setting satisfies condition (2.60) for all $p \geq 1$. The difficulty with a possible expansion in 'powers' of b_η is intuitively that, a Bernoulli variable B, being valued in $\{0,1\}$, is such that $B^p = B$ for all p. So all terms in the expansion are potentially of the same order. A different strategy is needed. We now explain an alternative, *formal* approach, for which we only have *partial* rigorous foundations to date. Although definite conclusions on the validity of the approach have yet to be obtained, the numerical tests we performed show its practical correctness and efficiency.

Heuristically, on the cube $Q_N = [-N/2, N/2]^d$ and at order 1 in η, the probability to get the perfect periodic material (entirely modeled by the matrix A_{per}) is $(1-\eta)^{N^d} \approx 1 - N^d\eta + O(\eta^2)$, while the probability to obtain the unperturbed material on all cells except one (where the material has matrix $A_{per}+C_{per}$) is $N^d\,(1-\eta)^{N^d-1}\eta \approx N^d\eta + O(\eta^2)$. All other configurations, with more than two cells perturbed, yield contributions of orders higher than or equal to η^2. This gives the intuition that the first order correction indeed comes from the difference between the material perfectly periodic except on one cell and the perfect material itself. We may proceed likewise for higher order terms. We now make all this more precise.

For well-posedness let us assume that there exist $0 < \alpha \leq \beta$ such that for all $\xi \in \mathbb{R}^d$ and almost all $x \in \mathbb{R}^d$,

$$\alpha|\xi|^2 \leq A_{per}(x)\xi\cdot\xi, \qquad \alpha|\xi|^2 \leq \left(A_{per}+C_{per}\right)(x)\xi\cdot\xi,$$

$$|A_{per}(x)\xi| \leq \beta|\xi|, \qquad |\left(A_{per}+C_{per}\right)(x)\xi| \leq \beta|\xi|.$$

We can therefore use for every $0 \leq \eta \leq 1$ the stochastic homogenization results recalled in Section 2.2.2. The cell problems associated to (2.58) read

$$\begin{cases} -\operatorname{div}\left(A_\eta(\nabla w_i^\eta + e_i)\right) = 0 \quad \text{in } \mathbb{R}^d, \\ \nabla w_i^\eta \ \text{ stationary}, \ \ \mathbb{E}\left(\displaystyle\int_Q \nabla w_i^\eta\right) = 0, \end{cases} \tag{2.62}$$

and the homogenized matrix $A_\eta^\star$ is given by

$$A_\eta^\star e_i = \mathbb{E}\left(\int_Q A_\eta(\nabla w_i^\eta + e_i)\right). \tag{2.63}$$

We will denote by w_i^0 the solution to the i-th *periodic* cell problem (2.2) associated with A_{per}.

Because of the specific form of A_η, and more precisely because A_η converges strongly to A_{per} in $L^2(Q \times \Omega)$ as $\eta \to 0$, it is easy to see that, when $\eta \to 0$, $A_\eta^\star \to A_{per}^\star$. Our purpose is to make this convergence more precise in terms of η.

We consider a specific realization $\tilde{\omega} \in \Omega$ of the matrix A_η in the truncated domain $Q_N = [-N/2, N/2]^d$ with N an even integer, and solve the following "supercell" problem :

$$\begin{cases} -\operatorname{div}\left(A_\eta(x,\tilde{\omega})(\nabla w_i^{\eta,N,\tilde{\omega}} + e_i)\right) = 0 \quad \text{in } Q_N, \\ w_i^{\eta,N,\tilde{\omega}} \ (N\mathbb{Z})^d - \text{periodic}, \end{cases} \tag{2.64}$$

as in (2.43). Then $\dfrac{1}{N^d}\displaystyle\int_{Q_N} A_\eta(x,\tilde{\omega})(\nabla w_i^{\eta,N,\tilde{\omega}}(x) + e_i)dx$ converges to $A_\eta^\star e_i$ almost surely. We deduce the convergence in expectation :

$$A_\eta^\star e_i = \lim_{N\to+\infty} \frac{1}{N^d}\mathbb{E}\left(\int_{Q_N} A_\eta(x,\omega)(\nabla w_i^{\eta,N,\omega}(x) + e_i)dx\right). \tag{2.65}$$

Using now the fact that b_η has a Bernoulli distribution in each cell of $\mathbb{Z}^d$, it is a simple matter to count the events (in the manner we briefly formally did above) and to make (2.65) more precise. We first define the set

$$\mathcal{T}_N = \{k \in \mathbb{Z}^d, Q + k \subset Q_N\}. \tag{2.66}$$

The cardinal of $\mathcal{T}_N$ is of course N^d, and

$$\bigcup_{k\in\mathcal{T}_N} \{Q + k\} = Q_N.$$

We then have the following possible values for A_η :

– $A_\eta(x,\tilde{\omega}) = A_{per}$ with probability $(1-\eta)^{N^d}$. In this case $w_i^{\eta,N,\tilde{\omega}} = w_i^0$ solves the usual periodic cell problem :

$$\begin{cases} -\operatorname{div}\left(A_{per}(\nabla w_i^0 + e_i)\right) = 0 \quad \text{in } Q, \\ w_i^0 \ \mathbb{Z}^d - \text{periodic}. \end{cases}$$

– $A_\eta(x,\tilde{\omega}) = A_{per} + \mathbf{1}_{Q+k}C_{per}$ for $k \in \mathcal{T}_N$, with probability $\eta(1-\eta)^{N^d-1}$. In this case $w_i^{\eta,N,\tilde{\omega}} = w_i^{1,k,N}$ solves the following problem, which we call here a "one defect" supercell problem :

$$\begin{cases} -\operatorname{div}\left(\left(A_{per} + \mathbf{1}_{Q+k}C_{per}\right)(\nabla w_i^{1,k,N} + e_i)\right) = 0 \quad \text{in } Q_N, \\ w_i^{1,k,N} \ (N\mathbb{Z})^d - \text{periodic}. \end{cases} \tag{2.67}$$

– $A_\eta(x,\tilde{\omega}) = A_{per} + \mathbf{1}_{Q+l\cup Q+m} C_{per}$ for $(l,m) \in \mathcal{T}_N$, $l \neq m$, with probability $\eta^2(1-\eta)^{N^d-2}$. In this case $w_i^{\eta,N,\tilde{\omega}} = w_i^{2,l,m,N}$ solves the following problem, which we call here a "two defects" supercell problem :

$$\begin{cases} -\operatorname{div}\left((A_{per} + \mathbf{1}_{Q+l\cup Q+m} C_{per}) \, (\nabla w_i^{2,l,m,N} + e_i) \right) = 0 \quad \text{in } Q_N, \\ w_i^{2,l,m,N} \ (N\mathbb{Z})^d - \text{periodic.} \end{cases} \tag{2.68}$$

– the other possible values for A_η, which we will not use henceforth, can be obtained through similar computations.

An instance of a setting with zero, one and two defects is shown in Figure 2.3 in the case of a material A_{per} consisting of a lattice of inclusions and for a supercell of size $N = 10$.

Defining $A_1^k = A_{per} + \mathbf{1}_{Q+k} C_{per}$ and $A_2^{l,m} = A_{per} + \mathbf{1}_{Q+l\cup Q+m} C_{per}$, we then easily obtain

$$\begin{aligned} A_\eta^\star &= \lim_{N\to\infty} \sum_{k=0}^{N^d} \eta^k A_k^{\star,N} \\ &= \lim_{N\to\infty} \left(A_0^{\star,N} + \eta A_1^{\star,N} + \eta^2 A_2^{\star,N} + o_N(\eta^2) \right), \end{aligned} \tag{2.69}$$

where the remainder $o_N(\eta^2)$ depends on N, and where

$$A_0^{\star,N} = A_{per}^\star,$$

$$A_1^{\star,N} e_i = \int_{Q_N} A_1^0 (\nabla w_i^{1,0,N} + e_i) - \int_{Q_N} A_{per} (\nabla w_i^0 + e_i),$$

$$\begin{aligned} A_2^{\star,N} e_i \ = \ & \frac{1}{2} \sum_{l\in\mathcal{T}_N, l\neq 0} \left(\int_{Q_N} A_2^{0,l} (\nabla w_i^{2,0,l,N} + e_i) \right. \\ & \left. -2 \int_{Q_N} A_1^0 (\nabla w_i^{1,0,N} + e_i) + \int_{Q_N} A_{per} (\nabla w_i^0 + e_i) \right). \end{aligned}$$

It is proved in [7], formally exchanging the limits $N \to \infty$ and $\eta \to 0$ in (2.69), that $A_1^{\star,N}$ and $A_2^{\star,N}$ are convergent sequences. We are not able to prove though that $A_\eta^\star - \lim_{N\to\infty} (A_{per}^\star + \eta A_1^{\star,N} + \eta^2 A_2^{\star,N}) = o(\eta^2)$ with $o(\eta^2)$ independent of N. Of course, in the one-dimensional setting, everything can be rigorously justified (see [7]).

Remark 2.3.1 *The expression of $A_1^{\star,N}$ (and likewise $A_2^{\star,N}$) is reminiscent of standard expressions in solid-state theory : each of the two integrals in the definition of $A_1^{\star,N}$ scales as the volume N^d of the domain Q_N, and a priori needs to be renormalized in order to give a finite limit. The difference however has a finite limit without renormalization. In solid-state physics, it is common to subtract a jellium, that is a uniform background, to proceed similarly.*

A theory in L^2 As mentioned above, we now consider a setting in which we are able to *prove* that our expansion indeed holds. We assume here that the random field b_η satisfies :

$$\exists M > 0, \forall \eta > 0, \|b_\eta\|_{L^\infty(Q\times\Omega)} \leq M, \tag{2.70}$$

$$\|b_\eta\|_{L^\infty(Q;L^2(\Omega))} \underset{\eta\to 0^+}{\to} 0 \ . \tag{2.71}$$

The asymptotic expansion for A_η with respect to η is given by the following theorem announced as Theorem 1 in [7] and proved in [8] :

Theorem 2.3.1 *Assume that b_η satisfies (2.70) and (2.71), and denote by $m_\eta = \|b_\eta\|_{L^\infty(Q;L^2(\Omega))}$. There exists a subsequence of η, still denoted η for the sake of simplicity, such that $\frac{b_\eta}{m_\eta}$ converges weakly-* in $L^\infty(Q;L^2(\Omega))$ to a limit field denoted by $\bar{b}_0$. Then*

– *for all $i \in \{1,d\}$, the following expansion*

$$\nabla w_i^\eta = \nabla w_i^0 + m_\eta \nabla v_i^0 + o(m_\eta)$$

holds weakly in $L^2(Q;L^2(\Omega))$ where w_i^0 is the solution to the i-th periodic cell problem and v_i^0 is solution in $\mathbb{R}^d$ to

$$\begin{cases} -\operatorname{div}(A_{per}\nabla v_i^0) = \operatorname{div}\left(\bar{b}_0 C_{per}(\nabla w_i^0 + e_i)\right), \\ \nabla v_i^0 \ \text{stationary}, \ \mathbb{E}\left(\displaystyle\int_Q \nabla v_i^0\right) = 0. \end{cases} \tag{2.72}$$

– *$A_\eta^\star$ can be expanded up to first order as*

$$A_\eta^\star = A_{per}^\star + m_\eta \tilde{A}_1^\star + o(m_\eta), \tag{2.73}$$

where, for all $i \in \{1,d\}$,

$$\tilde{A}_1^\star e_i = \int_Q \mathbb{E}(\bar{b}_0) C_{per}(\nabla w_i^0 + e_i) + \int_Q A_{per}\nabla\mathbb{E}(v_i^0). \tag{2.74}$$

Notice that taking the expectation of both sides of (2.72), $\mathbb{E}(v_i^0)$ is actually the $\mathbb{Z}^d$-periodic function that is the unique solution (up to an additive constant) to

$$\begin{cases} -\operatorname{div}\left(A_{per}\nabla\mathbb{E}(v_i^0)\right) = \operatorname{div}\left(\mathbb{E}(\bar{b}_0) C_{per}\left(\nabla w_i^0 + e_i\right)\right) \quad \text{in} \quad Q, \\ \mathbb{E}(v_i^0) \ \ \mathbb{Z}^d - \text{periodic}. \end{cases} \tag{2.75}$$

The computation of $A_\eta^\star$ up to the first order in m_η only requires solving $2d$ deterministic problems, namely (2.2) and (2.75), in the unit cell Q.

In fact, the situation is even more advantageous when A_{per} is a symmetric matrix (see Remark 2 in [8]). Up to some mathematical manipulations based upon the consideration of an *adjoint problem*, the computation of $A_\eta^\star$ up to the first order in m_η only requires solving the cell problems (2.2).

Pushing expansion (2.73) to second order requires more information on b_η :

Theorem 2.3.2 *Assume in addition to (2.70) and (2.71) that*

$$b_\eta = \eta\bar{b}_0 + \eta^2\bar{r}_0 + o(\eta^2) \quad \text{weakly} - \star \text{ in } L^\infty(Q; L^2(\Omega)).$$

Then

– *for all $i \in \{1, d\}$, the following expansion*

$$\nabla w_i^\eta = \nabla w_i^0 + \eta\nabla v_i^0 + \eta^2\nabla z_i^0 + o(\eta^2)$$

holds weakly in $L^2(Q; L^2(\Omega))$ where z_i^0 is solution in $\mathbb{R}^d$ to

$$\begin{cases} -\operatorname{div}(A_{per}\nabla z_i^0) = \operatorname{div}\left(\bar{r}_0 C_{per}(\nabla w_i^0 + e_i)\right) + \operatorname{div}\left(\bar{b}_0 C_{per}\nabla v_i^0\right), \\ \nabla z_i^0 \text{ stationary, } \mathbb{E}\left(\displaystyle\int_Q \nabla z_i^0\right) = 0. \end{cases}$$

– *$A_\eta^\star$ can be expanded up to second order as*

$$A_\eta^\star = A_{per}^\star + \eta\tilde{A}_1^\star + \eta^2\tilde{A}_2^\star + o(\eta^2),$$

where $\tilde{A}_1^\star$ is defined by (2.74) and for all $i \in \{1, d\}$,

$$\begin{aligned} \tilde{A}_2^\star e_i \;=\; & \int_Q \mathbb{E}(\bar{r}_0) C_{per}(\nabla w_i^0 + e_i) + \int_Q C_{per}\mathbb{E}(\bar{b}_0\nabla v_i^0) \\ & + \int_Q A_{per}\nabla\mathbb{E}(z_i^0). \end{aligned}$$

The computation of $A_\eta^\star$ up to the order η^2 is much more intricate than that up to the order η, for it requires determining $\mathbb{E}(\bar{b}_0\nabla v_i^0)$. Computing the periodic deterministic function $\mathbb{E}(v_i^0)$ solution to the simpler problem (2.75) is not sufficient in general. We have to determine the stationary random field v_i^0 solution to (2.72) in $\mathbb{R}^d$. A particular, practically relevant setting where the solution of the random problem (2.72) can be avoided, is discussed in Corollary 2 in [8].

Theorem 2.3.1, and the related results, are only of interest if $\mathbb{E}(\bar{b}_0) \neq 0$. Indeed, if $\mathbb{E}(\bar{b}_0) = 0$ it only states that $A_\eta^\star = A_{per}^\star + o(m_\eta)$.

The prototypical case where Theorem 2.3.1 does not provide valuable information is the case studied in the previous section, namely (2.61) where the B_η^k are independent identically distributed variables that have Bernoulli law with parameter η. Then, using the notations of Theorem 2.3.1, $b_\eta^2 = b_\eta$, $m_\eta = \sqrt{\eta}$ and $\bar{b}_0 = 0$, and we only get $A_\eta^\star = A_{per}^\star + o(\sqrt{\eta})$ (while it may be shown (see [7]) that there exists a matrix $\bar{A}_1^\star$ such that $A_\eta^\star = A_{per}^\star + \eta\bar{A}_1^\star + o(\eta)$ at least in dimension 1). Omitting the dependence on the space variables since b_η is uniform in each cell of $\mathbb{Z}^d$ in this particular setting, a suitable functional space F to obtain a non trivial weak limit of $\frac{b_\eta}{\|b_\eta\|_F}$ would be $L^1(\Omega)$ since the norm of each B_η^k in $L^1(\Omega)$ is equal to η. The

Dunford-Pettis weak compactness criterion in that space is however not satisfied by $\frac{b_\eta}{\|b_\eta\|_{L^1(\Omega)}}$. The reason is of course that $\frac{b_\eta}{\|b_\eta\|_{L^1(\Omega)}}$ converges in the set of bounded measures to a Dirac mass. The techniques used in the proof of Theorem 2.3.1 do not work in this setting.

The above considerations somehow suggest that an alternative viewpoint might be useful. Because of (2.71), the image measure dP_η^x of $b_\eta(x,\cdot)$ converges to a Dirac mass in the sense of distributions. We discuss in the next paragraph an alternate approach, which, as in the previous paragraph, consists in working out an expansion of the image measure (or of the law), rather than an expansion of the random variable.

Toward a general theory We now assume that the distribution of our random field b_η, denoted by dP_η, satisfies

$$dP_\eta = \delta_0 + \eta d\bar{P}_1 + \eta^2 d\bar{P}_2 + o(\eta^2) \quad \text{in } \mathcal{E}'(]-M,M[), \tag{2.76}$$

where $\mathcal{E}'(]-M,M[)$ denotes the set of distributions with compact support in the bounded interval $]-M,M[$. This is equivalent to

$$\mathbb{E}(\varphi(b_\eta)) = \langle dP_\eta, \varphi\rangle = \varphi(0) + \eta\langle d\bar{P}_1,\varphi\rangle + \eta^2\langle d\bar{P}_2,\varphi\rangle + o(\eta^2)$$

for all $\varphi \in \mathcal{C}^\infty(]-M,M[)$.

Considering a specific realization $\tilde{\omega}\in\Omega$ of A_η in $Q_N = [-N/2,N/2]^d$, we solve as above the supercell problem (2.64) and obtain the homogenized matrix (2.65). For convenience, we label the unit cells of Q_N from 1 to N^d. The k-th cell is denoted by Q_k, for $1\le k\le N^d$. A given realization $A_\eta(x,\tilde{\omega})$ can then be rewritten

$$A_\eta(x,\tilde{\omega}) = A_{per}(x) + \sum_{k=1}^{N^d} \mathbf{1}_{Q_k}(x) s_k C_{per}(x)$$

with $s_k = B_\eta^k(\tilde{\omega})$ for all $k\in\{1,N^d\}$. The $B_\eta^k(\tilde{\omega})$ being independent random variables, the joint probability of the N^d-tuple $(s_1,\cdots,s_{N^d})$ is simply the product $\displaystyle\prod_{k=1}^{N^d} dP_\eta(s_k)$.

We now define $A^{s_1,\cdots,s_{N^d}} = A_{per} + \displaystyle\sum_{k=1}^{N^d} \mathbf{1}_{Q_k} s_k C_{per}$ for $(s_1,\cdots,s_{N^d})\in[-M,M]^{N^d}$. We denote by $w_i^{s_1,\cdots,s_{N^d}}$ the solution of the i-th cell problem for the periodic homogenization of $A^{s_1,\cdots,s_{N^d}}$ on Q_N, that is

$$\begin{cases} -\operatorname{div}\left(A^{s_1,\cdots,s_{N^d}}(\nabla w_i^{s_1,\cdots,s_{N^d}} + e_i)\right) = 0 \quad \text{in } Q_N, \\ w_i^{s_1,\cdots,s_{N^d}}\ (N\mathbb{Z})^d - \text{periodic.} \end{cases}$$

Then, defining

$$A_\eta^{\star,N} = \frac{1}{N^d}\mathbb{E}\left(\int_{Q_N} A_\eta(x,\omega)(\nabla w_i^{\eta,N,\omega}(x) + e_i)dx\right),$$

we have

$$A_\eta^{\star,N} = \int \dots \int \frac{1}{N^d} \left(\int_{Q_N} A^{s_1,\cdots,s_{N^d}} \left(\nabla w_i^{s_1,\cdots,s_{N^d}} + e_i \right) \right) \prod_{k=1}^{N^d} dP_\eta(s_k). \tag{2.77}$$

Inserting the expansion

$$\begin{aligned} \prod_{k=1}^{N^d} dP_\eta(s_k) = & \prod_{k=1}^{N^d} \delta_0(s_k) + \eta \sum_{l=1}^{N^d} d\bar{P}_1(s_l) \prod_{k=1,k\neq l}^{N^d} \delta_0(s_k) \\ & + \frac{\eta^2}{2} \sum_{l=1}^{N^d} \sum_{m\neq l=1}^{N^d} d\bar{P}_1(s_l) d\bar{P}_1(s_m) \prod_{k=1,k\neq\{l,m\}}^{N^d} \delta_0(s_k) \\ & + \eta^2 \sum_{l=1}^{N^d} d\bar{P}_2(s_l) \prod_{k=1,k\neq l}^{N^d} \delta_0(s_k) + o_N(\eta^2) \end{aligned}$$

(with $o_N \in \mathcal{E}'(]-M, M[^{N^d})$) in (2.77), we obtain the following second-order expansion

$$A_\eta^{\star,N} = \tilde{A}_0^{\star,N} + \eta \tilde{A}_1^{\star,N} + \eta^2 \tilde{A}_2^{\star,N} + o_N(\eta^2).$$

The terms of the expansion can be explicitly determined. The zero order term is

$$\tilde{A}_0^{\star,N} = A_{per}^\star.$$

The first order term can be shown to read

$$\tilde{A}_1^{\star,N} e_i = \left\langle d\bar{P}_1(s), \int_{Q_N} A_1^{s,0} \left(\nabla w_i^{1,s,0,N} + e_i \right) \right\rangle, \tag{2.78}$$

where we denote by, for $s \in [-M, M]$,

$$A_1^{s,0} = A^{s,0\cdots,0} = A_{per} + s \mathbf{1}_Q C_{per},$$

and $w_i^{1,s,0,N} = w_i^{s,0,\cdots,0}$ solution to

$$\begin{cases} -\operatorname{div} \left(A_1^{s,0} (\nabla w_i^{1,s,0,N} + e_i) \right) = 0 \quad \text{in } Q_N, \\ w_i^{1,s,0,N} \ (N\mathbb{Z})^d - \text{periodic.} \end{cases}$$

The matrix $A_1^{s,0}$ corresponds to the periodic material with a defect of amplitude s located in Q (i.e at a position $0 \in \mathbb{Z}^d$ in Q_N), and $w_i^{1,s,0,N}$ is the i-th cell solution

for the periodic homogenization of $A_1^{s,0}$ in Q_N.

For the second-order term, we recall our definition (2.66) of $\mathcal{T}_N$. For $(s,t) \in [-M,M]^2$ and $l \in \mathcal{T}_N$, we define

$$A_2^{s,t,0,l} = A_{per} + s\mathbf{1}_Q C_{per} + t\mathbf{1}_{Q+l} C_{per},$$

and $w_i^{2,s,t,0,l,N}$ solution to

$$\begin{cases} -\operatorname{div}\left(A_2^{s,t,0,l}(\nabla w_i^{2,s,t,0,l,N} + e_i)\right) = 0 \quad \text{in } Q_N, \\ w_i^{2,s,t,0,l,N} \ (N\mathbb{Z})^d - \text{periodic.} \end{cases}$$

The matrix $A_2^{s,t,0,l}$ corresponds to the periodic material with two defects of amplitude s and t located in Q and $Q+l$ (i.e at positions $0 \in \mathbb{Z}^d$ and $l \in \mathbb{Z}^d$ in Q_N) respectively. The function $w_i^{2,s,t,0,l,N}$ is the i-th cell solution for the periodic homogenization of $A_2^{s,t,0,l}$ in Q_N.

Then computations similar to that for the first order yield

$$\begin{aligned} \tilde{A}_2^{\star,N} e_i = \frac{1}{2} \sum_{l \in \mathcal{T}_N, l \neq 0} & \left\langle d\bar{P}_1(s) d\bar{P}_1(t), \int_{Q_N} A_2^{s,t,0,l}(\nabla w_i^{2,s,t,0,l,N} + e_i) \right\rangle \\ & + \left\langle d\bar{P}_2(s), \int_{Q_N} A_1^{s,0}\left(\nabla w_i^{1,s,0,N} + e_i\right) \right\rangle. \end{aligned} \tag{2.79}$$

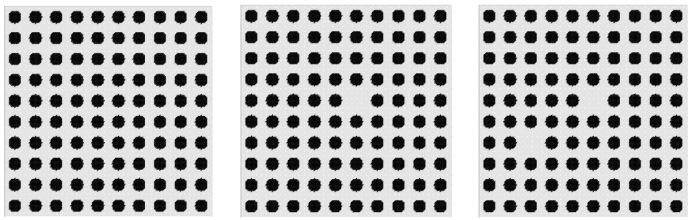


FIG. 2.3 – From left to right : zero defect, one defect and two defects.

Remark 2.3.2 *It is illustrative to consider the particular case studied in a previous paragraph where the random variable B_η has a Bernoulli law. Then, expansion (2.76) holds exactly with $d\bar{P}_1 = \delta_1 - \delta_0$. The distribution $d\bar{P}_2$ and all other terms of higher order identically vanish. The expressions (2.78) and (2.79) then coincide with the expressions previously established.*

It is proved in [8] that $A_1^{\star,N}$ and $A_2^{\star,N}$ converge to a finite limit when $N \to \infty$.

2.3.1.2 Numerical methodology and results

Our purpose in this section is to assess the approximation of $A_\eta^\star$ by the second-order expansion $A_{per}^\star + \eta A_1^{\star,N} + \eta^2 A_2^{\star,N}$. The limited computational facilities we have access to impose that we restrict ourselves to the two-dimensional case.

We consider a commonly used two-phase composite material as periodic reference material A_{per}. The material consists of a constant background reinforced by a periodic lattice of circular inclusions, that is

$$A_{per}(x_1, x_2) = 20\, Id + 100 \sum_{k \in \mathbb{Z}^2} \mathbf{1}_{B(k,1)}(x_1, x_2)\, Id,$$

where $B(k,1)$ is the ball of center k and radius 1. The role of the perturbation is, loosely speaking, to randomly eliminate some fibers :

$$C_{per}(x_1, x_2) = -100 \sum_{k \in \mathbb{Z}^2} \mathbf{1}_{B(k,1)}(x_1, x_2)\, Id.$$

We have chosen the coefficients 20 and 100 in order to have a high contrast between A_{per} and $A_{per} + C_{per}$, and thus for the perturbation to be significant. The material is shown on Figure 2.3, left.

Our goal is to compare $A_\eta^\star$ with its approximation $A_{per}^\star + \eta \tilde{A}_1^{\star,N} + \eta^2 \tilde{A}_2^{\star,N}$. A major computational difficulty is the computation of the "exact" matrix $A_\eta^\star$ given by formula (2.63). It ideally requires to solve the stochastic cell problems (2.62) on $\mathbb{R}^d$. To this end, we actually compute, for a given realization ω and a truncated domain Q_N the approximation (2.42)-(2.43). In a second step, we take averages over the realizations ω.

For each ω, we use the finite element software `FreeFem++` [54] to solve the boundary value problems (2.64) (or, otherwise stated, (2.43)) and compute the integrals (2.42). We work with standard P1 finite elements on a triangular mesh such that there are 10 degrees of freedom on each edge of the unit cell Q. We define an approximate value $A_\eta^{\star,N}$ as the average of $A_\eta^{\star,N}(\omega)$ over 40 realizations ω. Our numerical experiments indeed show that the number 40 is high enough for the convergence of this kind of Monte-Carlo computation to be reached. We then let N grow from 5 to 80 by steps of 5. We observe that $A_\eta^{\star,N}$ stabilizes at a fixed value around $N = 80$ and thus take $A_\eta^{*,80}$ as the reference value for $A_\eta^\star$ in our subsequent tests. The next step is to compute the zero-order term $A_{per}^\star$, and the first-order and second-order deterministic corrections $A_1^{\star,N}$ and $A_2^{\star,N}$. Using the same mesh and finite elements as for our reference computation above, we compute $A_{per}^\star$ using the periodic cell problem, and for each N we compute $A_1^{\star,N}$ and $A_2^{\star,N}$ using the formulae established in the previous section. We again let N grow from 5 to 80 by steps of 5 for $A_1^{\star,N}$. The computation of $A_2^{\star,N}$ being significantly more expensive

(there is not only an integral over Q_N but also a sum over the N^2 cells) we have to limit ourselves to $N = 25$ and approximate the value for N larger than 25 by the value obtained for $N = 25$.

Note that there are three distinct sources of error in our computations :
- the finite elements discretization error ;
- the truncation error due to the replacement of $\mathbb{R}^d$ with Q_N, in the computation of the stochastic cell problems (2.62) that are replaced with (2.64), as well as in the computation of the integrals ;
- the stochastic error arising from the approximation of the expectation value (2.65) by an empirical mean.

As regards the discretization error, we only have in practice access to the finite element approximations of all the functions manipulated here (such as w_i^0, $w_i^{\eta,N,\omega}$,...). Although we have not proved it in the specific context of our work, it is shown in a similar weakly random setting (see [P1, P6]) that all the convergences stated and used here still hold true for the finite-dimensional approximations of the objects. This is likely to be the case in our framework, and since our numerical results additionally confirm it we will admit it. Our practical approach to this issue consists in adopting the same finite element space for all approximations of the cell and supercell problems, independently of N.

The truncation error is a different issue. For the "exact" computation of $A_\eta^\star$ (we mean not using the second-order expansion), we only know from [23], for a continuous notion of stationarity instead of the discrete one we use, and under mixing conditions which are satisfied in our setting, that the convergence of the truncated approximation to the ideal value holds at a rate $N^{-\kappa}$ with $\kappa > 0$ an undetermined function of the dimension, the mixing and the properties of the material. Concerning the second-order expansion, the zero-order term $A_{per}^\star$ is of course free of any truncation error. We have only partial knowledge regarding the first-order term $\tilde{A}_1^{\star,N}$, we have no insight on the truncation error for the second-order correction $\tilde{A}_2^{\star,N}$ and we also wish to study its convergence from a numerical point of view.

Finally, we have a practical approach to the stochastic error : besides the empirical mean, we provide for each N the minimum and the maximum of $A_\eta^{\star,N}(\omega)$ achieved over the 40 computations.

We only hope here to demonstrate, and we indeed do so, that the second-order expansion is an approximation to $A_\eta^\star$ sufficiently good for all practical purposes, and in particular for η not too small ! We will observe that, when $N \longrightarrow +\infty$, $A_2^{\star,N}$ converges to a limit $\bar{A}_2^\star$, and that both $A_1^{\star,N}$ and $A_2^{\star,N}$ converge to their respective limits faster than $A_\eta^{\star,N}$ to $A_\eta^\star$ (which is expected since the former quantities are deterministic and contain less information). We will also observe that $A_{per}^\star + \eta A_1^{\star,N}$

is closer to $A^\star_\eta$ than $A^\star_{per}$. The inclusion of the second order improves the situation : $A^\star_{per} + \eta A_1^{\star,N} + \eta^2 A_2^{\star,N}$ is even closer.

We will use the following legend in the graphs :
- *periodic* : gives the value of the periodic homogenized matrix $A^\star_{per}$;
- *first-order* : gives the value of $A^\star_{per} + \eta A_1^{\star,N}$;
- *second-order* : gives the value of $A^\star_{per} + \eta A_1^{\star,N} + \eta^2 A_2^{\star,N}$;
- *stochastic mean, minima and maxima* : respectively give the values of $A_\eta^{\star,N}$ and the extrema obtained in the computation of the empirical mean.

The first test we present corresponds to the Bernoulli setting of formula (2.61). We choose $\eta = 0.4$, so that the perturbation is important. The results are shown in Figure 2.4.

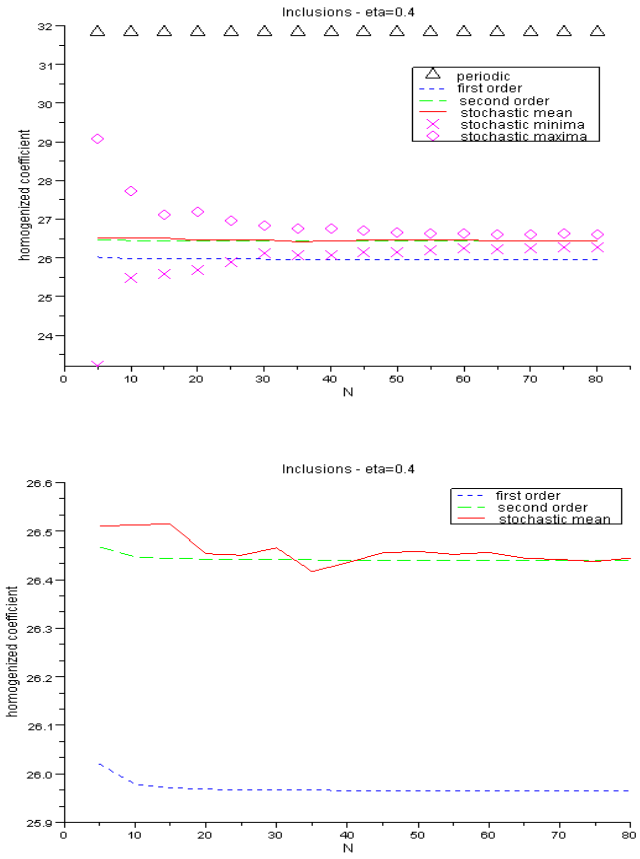


FIG. 2.4 – Results for a Bernoulli perturbation and $\eta = 0.4$. Above : complete results. Below : close-up on $A_\eta^{*,N}$ and the first and second-order corrections.

The second test is an illustration of our general theory based upon an expansion of the image measure of the perturbation, as given by formula (2.76).

In this test, we consider R_η a random variable having Bernoulli law with parameter η, and U a uniform variable on $[0,1]$ independent of R_η. We define $B_\eta = R_\eta - \eta U$. Then a simple computation (see [8]) shows that the following expansion holds :

$$\begin{aligned} dP_\eta =& \delta_0 + \eta\left(-\mathbb{E}(U)\delta_0' + \delta_1 - \delta_0\right) \\ &+ \eta^2\left(-\mathbb{E}(U)(\delta_1' - \delta_0') + \frac{1}{2}\mathbb{E}(U^2)\delta''(0)\right) + o(\eta^2) \quad \text{in } \mathcal{E}'(\mathbb{R}). \end{aligned} \tag{2.80}$$

We insert (2.80) in (2.78) and (2.79) for $\eta = 0.1$. The results are shown in Figure 2.5.

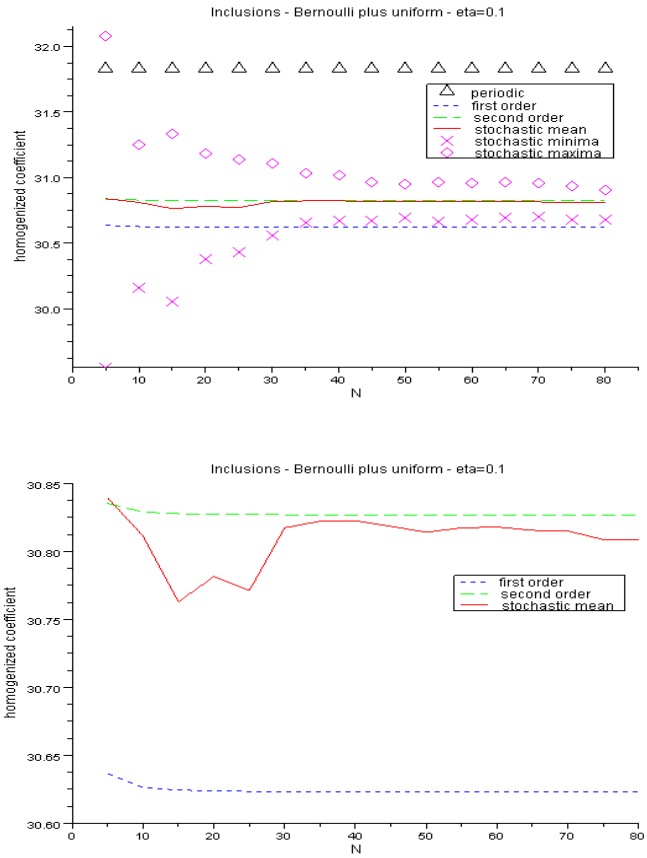


FIG. 2.5 – Results for perturbation (2.80) and $\eta = 0.1$. Above : complete results. Below : close-up on $A_\eta^{*,N}$ and the first and second-order corrections.

Notice on both Figures 2.4 and 2.5 (it is especially clear on the close-ups) that the first and second-order corrections $A_1^{*,N}$ and $A_2^{*,N}$ converge very fast in function of N, and in particular, as expected, much faster than the stochastic computation.

Convergence of these deterministic computations is actually typically reached for $N = 10$.

Then, it is clear in these two instances that the first-order correction enables to get substantially closer to A^*_η. The interest of the second-order term is also obvious for it provides excellent accuracy.

We stress that our perturbative approach has limitations and deteriorates, like any asymptotic approach, for large values of η, the threshold being case dependent (note that the Bernoulli setting with η as large as 0.4 still works extremely well for the material of Figure 2.3, left). However, the comprehensive numerical experiments that we have performed in [7] and [8], which provide results qualitatively similar to those above, lead us to conclude that this approach is generically robust.

2.3.2 A variant of the general setting

We now move on to a slightly different "perturbative" setting. We will consider a random problem that still is, in some vague sense, close to a periodic problem. However, in contrast to the previous section, the setting is not, originally at least, obtained as an *expansion* over a periodic setting. It is a *genuinely* random setting, on which an expansion can be superimposed. This specific stochastic setting has been introduced and studied in [19, 20]. It is *not* a particular case of the classical stationary setting defined in Section 2.2.2. As briefly mentioned in the introduction, it is motivated by the consideration of random geometries (we mean, materials) that have some relation to the periodic setting. Here, the periodic setting is taken as a *reference* configuration, somewhat similarly to the classical mathematical formalization of continuum mechanics where a reference configuration is used to define the state of the material under study. The real situation is seen *via* a *mapping* from the reference configuration to the actual configuration. Otherwise stated, the setting is a *periodic setting seen through random glasses*...

2.3.2.1 Mathematical setting and approximation

We fix some $\mathbb{Z}^d$-periodic, square matrix A_{per} of size d, assumed to satisfy

$$\exists \gamma > 0 \quad / \quad \forall \xi \in \mathbb{R}^d, \quad \xi^T A_{per}(y)\xi \geq \gamma |\xi|^2, \quad \text{a.e. in } \mathbb{R}^d, \tag{2.81}$$

$$\forall i, j \in \{1, 2, \ldots, d\}, \quad [A_{per}]_{ij} \in L^\infty\left(\mathbb{R}^d\right). \tag{2.82}$$

We consider the following specific form of problem (2.31)

$$\begin{cases} -\text{div}\left(A_{per}\left(\Phi^{-1}\left(\frac{\cdot}{\varepsilon}, \omega\right)\right)\nabla u^\varepsilon\right) = f \quad \text{in} \quad \mathcal{D}, \\ \\ u^\varepsilon = 0 \quad \text{on} \quad \partial\mathcal{D}, \end{cases} \tag{2.83}$$

where the function $\Phi(\cdot,\omega)$ is assumed to be a diffeomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$ for $\mathbb{P}$-almost every ω. The diffeomorphism is assumed to additionally satisfy

$$\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\text{EssInf}}\, [\det(\nabla\Phi(x,\omega))] = \nu > 0, \tag{2.84}$$

$$\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\text{EssSup}}\, (|\nabla\Phi(x,\omega)|) = M < \infty, \tag{2.85}$$

$$\nabla\Phi \quad \text{is stationary in the sense of (2.28).} \tag{2.86}$$

Such a Φ is called a *random stationary diffeomorphism.*

The following result is proved in [19, 20] :

Theorem 2.3.3 *Let $\mathcal{D}$ be a bounded smooth open subset of $\mathbb{R}^d$, and let $f \in H^{-1}(\mathcal{D})$. Let $A_{\rm per}$ be a square matrix which is $\mathbb{Z}^d$-periodic and satisfies (2.81)-(2.82). Let Φ be a random stationary diffeomorphism satisfying hypotheses (2.84)-(2.85)-(2.86). Then the solution $u^\varepsilon(x,\omega)$ to (2.83) satisfies the following properties :*
(i) $u^\varepsilon(x,\omega)$ converges to some $u_0(x)$ strongly in $L^2(\mathcal{D})$ and weakly in $H^1(\mathcal{D})$, almost surely ;
(ii) the function u_0 is the solution to the homogenized problem :

$$\begin{cases} -\text{div}\,(A^\star\nabla u_0) = f \quad in \quad \mathcal{D}, \\ \\ u_0 = 0 \quad on \quad \partial\mathcal{D}. \end{cases} \tag{2.87}$$

In (2.87), the homogenized matrix $A^\star$ is defined by :

$$\begin{aligned} [A^\star]_{ij} &= \det\left(\mathbb{E}\left(\int_Q \nabla\Phi(z,\cdot)dz\right)\right)^{-1} \dots \\ &\times \quad \mathbb{E}\left(\int_{\Phi(Q,\cdot)} (e_i + \nabla w_{e_i}(y,\cdot))^T A_{per}\left(\Phi^{-1}(y,\cdot)\right) e_j\, dy\right), \end{aligned} \tag{2.88}$$

where for any $p \in \mathbb{R}^d$, w_p is the solution (unique up to the addition of a (random) constant) in $\left\{w \in L^2_{\rm loc}(\mathbb{R}^d, L^2(\Omega)), \quad \nabla w \in L^2_{\rm unif}(\mathbb{R}^d, L^2(\Omega))\right\}$ to

$$\begin{cases} -\text{div}\left[A_{per}\left(\Phi^{-1}(y,\omega)\right)(p + \nabla w_p)\right] = 0, \\ \\ w_p(y,\omega) = \tilde{w}_p\left(\Phi^{-1}(y,\omega),\omega\right), \quad \nabla\tilde{w}_p \quad \textit{is stationary as in (2.28)}, \\ \\ \mathbb{E}\left(\displaystyle\int_{\Phi(Q,\cdot)} \nabla w_p(y,\cdot)dy\right) = 0. \end{cases} \tag{2.89}$$

At the discrete level, we proceed as follows. The corrector problems (2.89) are solved numerically by truncation and approximation. For a meshsize h fixed, we give ourselves a periodic triangulation $\mathcal{T}_h^{(Q)}$ (in two dimensions, to fix the ideas) of the

unit cell Q. By periodicity, we next obtain the triangulation $\mathcal{T}_h = \cup_{k\in\mathbb{Z}^d}\left(k + \mathcal{T}_h^{(Q)}\right)$ of $\mathbb{R}^d$. For all $N \in \mathbb{N}^*$, we denote by $Q_N = [0, N]^d$ and $\mathcal{T}_h^N = \mathcal{T}_h \cap Q_N$ the associated triangulation. Mimicking the approach used in [23] (and recalled above) to approximate standard corrector problems in classical random homogenization, we consider

$$\left\{\begin{array}{l} \text{Find } \widetilde{w}_p^{h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N) \ \text{ such that, for all } \widetilde{v}_h \in V_h^{\rm per}(Q_N), \\ \\ \displaystyle\int_{Q_N} \det\left(\nabla\Phi\right)\left(\nabla\widetilde{v}_h\right)^T\left(\nabla\Phi\right)^{-T} A_{per}\left(p + \left(\nabla\Phi\right)^{-1}\nabla\widetilde{w}_p^{h,N}(\cdot,\omega)\right) = 0 \ \text{ a.s.}, \end{array}\right. \tag{2.90}$$

where $V_h^{\rm per}(Q_N)$ is the set of Q_N-periodic functions defined on $\mathbb{R}^d$, with restriction on Q_N that belongs to a periodic finite element space constructed from the mesh $\mathcal{T}_h^N$. Notice it is more convenient to work on the function $\widetilde{w}_p$ rather than on the function w_p because the stationarity assumption holds on the former. We next define the following approximation of the entries of the matrix $A^\star$:

$$\begin{array}{l} \left(A^{\star,h,N}\right)_{ij}(\omega) = \\ \det\left(\displaystyle\frac{1}{|Q_N|}\int_{Q_N}\nabla\Phi\right)^{-1}\frac{1}{|Q_N|}\int_{Q_N}\det(\nabla\Phi)\left(e_i + \left(\nabla\Phi\right)^{-1}\nabla\widetilde{w}_{e_i}^{h,N}\right)^T A_{per}\, e_j. \end{array} \tag{2.91}$$

For a numerical illustration, we work in two dimensions : $x = (x_1, x_2)$, and we give ourselves two families $(X_k)_{k\in\mathbb{Z}}$ and $(Y_k)_{k\in\mathbb{Z}}$ of i.i.d. random variables, all sharing the uniform law $\mathcal{U}([a, b])$. We consider the diffeomorphism : $\Phi(x,\omega) = 6\,x + \Psi(x,\omega)$, with $\Psi(x,\omega) = (\psi_X(x_1,\omega), \psi_Y(x_2,\omega))$, where

$$\psi_X(x_1,\omega) = \sum_{k\in\mathbb{Z}} 1_{[k,k+1[}(x_1)\left(\sum_{q=0}^{k-1} X_q(\omega) + 2X_k(\omega)\int_k^{x_1}\sin^2(2\pi t)\,dt\right),$$

and likewise for ψ_Y. The periodic matrix A_{per} is defined, for all $x \in Q$, by

$$A_{per}(x) = a_{\rm per}(x)\,{\rm Id}_2, \quad a_{\rm per}(x_1,x_2) = \beta + (\alpha-\beta)\sin^2(\pi x_1)\sin^2(\pi x_2).$$

Note that ψ_X is not stationary but its gradient is. As announced above, the case under consideration is not a particular case of classical stationary random homogenization.

The parameters used in the computation are made precise in [P1]. All our numerical results are obtained using `FreeFem++` [54]. Figure 2.6 shows the convergence obtained for (2.91).

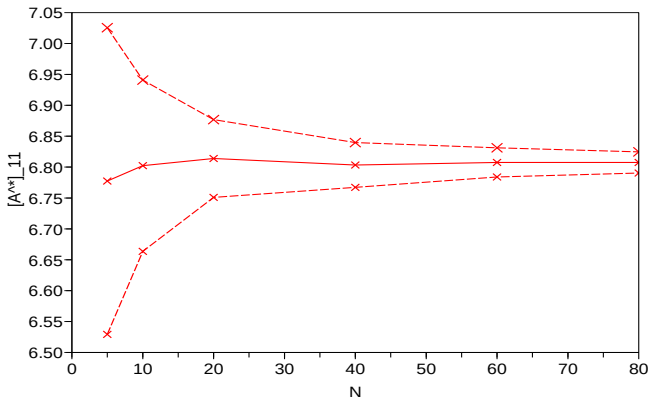


FIG. 2.6 – Results for (2.91) : convergence of the estimator as $N \longrightarrow +\infty$ and associated interval of confidence.

2.3.2.2 Perturbations of Identity

Theory We now superimpose to the formalism presented above the idea of considering *small* perturbations, in the spirit of the previous sections. When the diffeomorphism Φ is a perturbation of Identity, in the sense

$$\Phi(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2), \tag{2.92}$$

the solution of the corrector problem (2.89) can be expanded in powers of η. More precisely, $\widetilde{w}_p(x,\omega) = w_p^0(x) + \eta w_p^1(x,\omega) + O(\eta^2)$, where w_p^0 satisfies

$$-\text{div}\left[A_{per}\left(p+\nabla w_p^0\right)\right] = 0, \quad w_p^0 \text{ is } Q\text{-periodic}, \tag{2.93}$$

while w_p^1 is defined by

$$\begin{cases} -\text{div}\left[A_{per}\,\nabla w_p^1\right] = \text{div}\left[-A_{per}\,\nabla\Psi\,\nabla w_p^0 \right. \\ \qquad\qquad \left. -(\nabla\Psi^T - (\text{div}\,\Psi)\text{Id})\,A_{per}\left(p+\nabla w_p^0\right)\right], \\ \nabla w_p^1 \text{ is stationary and } \mathbb{E}\left(\int_Q \nabla w_p^1\right) = 0. \end{cases} \tag{2.94}$$

The problem defining w_p^1 is stochastic, thus computationally expensive. However, as remarked in [20], the average $\overline{w}_p^1 = \mathbb{E}(w_p^1)$ is Q-periodic and satisfies the closed *deterministic* equation

$$\begin{aligned} -\text{div}\left[A_{per}\,\nabla\overline{w}_p^1\right] \;=\; & \text{div}\left[-A_{per}\,\mathbb{E}(\nabla\Psi)\,\nabla w_p^0 \right. \\ & \left. -(\mathbb{E}(\nabla\Psi^T) - \mathbb{E}(\text{div}\,\Psi)\text{Id})\,A_{per}\left(p+\nabla w_p^0\right)\right]. \end{aligned} \tag{2.95}$$

This problem is much easier to solve than (2.94). In addition, the knowledge of only w_p^0 and $\overline{w}_p^1$ is sufficient to obtain an approximation, at the first order in η, of the homogenized matrix $A^\star$. It can be proven (see [20]) that we indeed have

$$A^\star = A^0 + \eta A^1 + O(\eta^2), \tag{2.96}$$

with $A^0_{ij} = \displaystyle\int_Q \left(e_i + \nabla w^0_{e_i}\right)^T A_{per}\, e_j$ and

$$\begin{aligned} A^1_{ij} &= -\int_Q \mathbb{E}(\operatorname{div}\Psi)\, A^0_{ij} + \int_Q (e_i + \nabla w^0_{e_i})^T A_{per}\, e_j\, \mathbb{E}(\operatorname{div}\Psi) \\ &\quad + \int_Q \left(\nabla \overline{w}^1_{e_i} - \mathbb{E}(\nabla\Psi)\nabla w^0_{e_i}\right)^T A_{per}\, e_j. \end{aligned}$$

At the first order in η, we can thus eliminate all the randomness of the problem.

Discretization We assume (2.92). Inserting the formal expansion $\widetilde{w}^{h,N}_p = w^{0,h,N}_p + \eta w^{1,h,N}_p + O(\eta^2)$ in the *discretized* formulation (2.90), we obtain that $w^{0,h,N}_p$ is independent of N (we henceforth denote it by $w^{0,h}_p$), and that $w^{0,h}_p \in V^{\rm per}_h(Q)$ and $w^{1,h,N}_p \in V^{\rm per}_h(Q_N)$ respectively solve, for all $v_h \in V^{\rm per}_h(Q)$,

$$\int_Q (\nabla v_h)^T\, A_{per}\, (p + \nabla w^{0,h}_p) = 0, \tag{2.97}$$

and, for all $v_h \in V^{\rm per}_h(Q_N)$,

$$\begin{aligned} \int_{Q_N} (\nabla v_h)^T\, A_{per}\, \nabla w^{1,h,N}_p &= \int_{Q_N} (\nabla v_h)^T \left[A_{per}\, \nabla\Psi\, \nabla w^{0,h}_p \right. \\ &\quad \left. + \left(\nabla\Psi^T - (\operatorname{div}\Psi)\,\mathrm{Id}\right) A_{per}\left(p + \nabla w^{0,h}_p\right)\right]. \end{aligned} \tag{2.98}$$

Expressions (2.97) and (2.98) are discretized formulations of (2.93) and (2.94), respectively. The following proposition extends to the discretized formulation ($h > 0$ and $N < \infty$) the expansion performed in the previous paragraph at the continuous level. The result has been announced in [P1]. Its proof follows the lines of that of Proposition 3.1 in [20], and is made precise in [P6].

Proposition 2.3.1 *Assume that $\Phi(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2)$ as $\eta \to 0$, in $C^1(\mathbb{R}^d, L^2(\Omega))$, where $\nabla\Psi$ is stationary. Then there exists a constant $C(h,N,\omega)$ such that, for all η sufficiently small*

$$\eta^{-2} \left\|\nabla\widetilde{w}^{h,N}_p(\cdot,\omega) - \nabla w^{0,h}_p - \eta\nabla w^{1,h,N}_p(\cdot,\omega)\right\|_{L^2(Q_N)} \leq \sqrt{|Q_N|}\; C(h,N,\omega),$$

where $\widetilde{w}^{h,N}_p$, $w^{0,h}_p$ and $w^{1,h,N}_p$ respectively solve (2.90), (2.97) and (2.98), and such that

$$\eta^{-2} \left|A^{\star,h,N}(\omega) - A^{0,h} - \eta A^{1,h,N}(\omega)\right| \leq C(h,N,\omega), \tag{2.99}$$

where $A^{\star,h,N}$ is the matrix defined by (2.91), $(A^{0,h})_{ij} = \displaystyle\int_Q \left(e_i + \nabla w^{0,h}_{e_i}\right)^T A_{per}\, e_j$ and

$$\begin{aligned} (A^{1,h,N})_{ij} &= -(A^{0,h})_{ij}\, \frac{1}{|Q_N|}\int_{Q_N} \operatorname{div}\Psi + \frac{1}{|Q_N|}\int_{Q_N} (e_i + \nabla w^{0,h}_{e_i})^T A_{per}\, e_j \operatorname{div}\Psi \\ &\quad + \frac{1}{|Q_N|}\int_{Q_N} \left(\nabla w^{1,h,N}_{e_i} - \nabla\Psi\, \nabla w^{0,h}_{e_i}\right)^T A_{per}\, e_j. \end{aligned}$$

Let w_p^0 and $w_p^{0,h}$ respectively solve (2.93) and (2.97). Under classical assumptions, $\nabla w_p^{0,h}$ converges to ∇w_p^0 in $L^2(Q)$, whence $\lim_{h\to 0} A^{0,h} = A^0$. The convergence as $N \to \infty$ of $A^{\star,h,N}(\omega)$ and of $\lim_{h\to 0} A^{\star,h,N}(\omega)$ are questions different in nature. In the one-dimensional setting, it can be shown that $\lim_{N\to\infty} \lim_{h\to 0} A^{\star,h,N}(\omega) = A^\star$ almost surely. The study of the convergence in N of $\lim_{h\to 0} A^{\star,h,N}(\omega)$ in higher dimensions is left for future work. Our numerical results consistently show that $\lim_{h\to 0} \lim_{N\to\infty} A^{\star,h,N}(\omega) = A^\star$. The situation is different for $A^{1,h,N}(\omega)$ since (2.94) and (2.98) are not classical corrector problems. Similar remarks however apply.

At the continuous level, it has been established (and briefly recalled above) that one needs not determine w_p^1 to compute A^1. Only the average $\overline{w}_p^1$, solution to a simpler, deterministic problem is required. The observation carries over to the discretized setting ($h > 0$ and $N < \infty$). Introduce indeed $\overline{w}_p^{1,h,N} = \mathbb{E}(w_p^{1,h,N})$, which satisfies, for all $v_h \in V_h^{\rm per}(Q_N)$,

$$\begin{aligned} \int_{Q_N} (\nabla v_h)^T \, A_{per} \, \nabla \overline{w}_p^{1,h,N} &= \int_{Q_N} (\nabla v_h)^T \Big[A_{per} \, \mathbb{E}(\nabla \Psi) \, \nabla w_p^{0,h} \\ &+ \Big(\mathbb{E}\left(\nabla \Psi\right)^T - \mathbb{E}\left(\operatorname{div} \Psi\right) \mathrm{Id} \Big) \, A_{per} \left(p + \nabla w_p^{0,h} \right) \Big] . \end{aligned} \tag{2.100}$$

Since A_{per} and $\mathbb{E}(\nabla\Psi)$ are Q-periodic, we obtain that the solution to (2.100) with $N = 1$, which is Q-periodic, is solution to (2.100) for all $N \in \mathbb{N}^*$. Since the solution to (2.100) is unique, up to the addition of a random constant, we obtain that $\overline{w}_p^{1,h,N}$ is actually independent from N (we denote it by $\overline{w}_p^{1,h} = \mathbb{E}(w_p^{1,h,N})$), Q-periodic and solution to (2.100) with $N = 1$. The latter is a discretization of (2.95) converging as $h \to 0$. As a function of $\nabla \overline{w}_p^{1,h}$, the matrix $A^{1,h} = \mathbb{E}(A^{1,h,N})$ is also independent of N, and can be straightforwardly computed from $\nabla \overline{w}_p^{1,h}$.

We now have at hand two approximations of A^1, namely $A^{1,h,N}(\omega)$ and $A^{1,h} = \mathbb{E}(A^{1,h,N})$. For the analysis of the approach, we use $A^{1,h,N}$ rather than $A^{1,h}$. We introduce two error estimators

$$\begin{aligned} e^{h,N}(\omega) &:= \eta^{-2} \left(A^{\star,h,N}(\omega) - A^{0,h} - \eta A^{1,h,N}(\omega) \right) \\ \overline{e}^{h,N}(\omega) &:= \eta^{-2} \left(A^{\star,h,N}(\omega) - A^{0,h} - \eta A^{1,h} \right) \\ &= e^{h,N}(\omega) + \eta^{-1} \left(A^{1,h,N}(\omega) - A^{1,h} \right) . \end{aligned} \tag{2.101}$$

Given that this is true in the one-dimensional setting, we assume that the constant $C(h, N, \omega)$ in (2.99) has finite expectation value and variance. Since $A^{1,h} = \mathbb{E}(A^{1,h,N})$, we deduce that the expectation of $\overline{e}^{h,N}(\omega)$ is bounded when $\eta \to 0$. Expression (2.101) however implies that

$$\eta^{-2} \, \mathbb{V}\mathrm{ar} \left(A^{1,h,N}(\omega) - A^{1,h} \right) \leq 2 \mathbb{V}\mathrm{ar} \left(\overline{e}^{h,N} \right) + 2 \mathbb{V}\mathrm{ar} \left(e^{h,N} \right) .$$

We hence observe that the variance of $\overline{e}^{h,N}$ blows up when $\eta \to 0$. As a consequence, for the specific purpose of the *analysis* of the approach, we rather use $e^{h,N}$. This allows us to verify that $A^{0,h}+\eta A^{1,h,N}$ is an accurate estimator for $A^{\star,h,N}$, at the second order in η. *In practice*, we of course approach $A^\star$ itself and we use $A^{0,h}+\eta A^{1,h}$, much easier to compute, for this purpose.

Proposition 2.3.1 shows that $e^{h,N}(\omega)$ is bounded when $\eta \to 0$, by a constant $C(h,N,\omega)$ that does *a priori* depend on h, N and ω. It can be easily seen in the one-dimensional setting that this constant is actually uniformly bounded in h, N and ω, assuming that $\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\text{EssSup}}\,(|\nabla\Psi(x,\omega)|) = \overline{M} < \infty$. The proof can be extended to the multi-dimensional setting if we assume the expansion $\Phi(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2)$ holds in $C^1(\mathbb{R}^d, L^\infty(\Omega))$ and not only in $C^1(\mathbb{R}^d, L^2(\Omega))$. We refer to [P6] for these results. Our numerical results in fact show that this is generically true, although we are not in position to prove this to date.

Example of numerical results We again work in dimension two, in the same context as for our numerical test above in Section 2.3.2.1. The only modification concerns the definition of the diffeomorphism, which now writes $\Phi(x,\omega) = x + \eta\Psi(x,\omega)$ for the same function $\Psi(x,\omega)$ as above, but where the coefficient η is taken small. Figure 2.7 shows $A_{per}\circ\Phi^{-1}(x,\omega)$ for a particular realization. Our results on $\left(A^{\star,h,N}\right)_{11}$ and $\left(e^{h,N}\right)_{11}$ are summarized on Table 2.2. The parameters used in the computations are made precise in [P1].

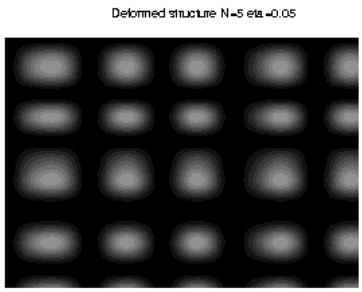


FIG. 2.7 – We show a particular realization of $A_{per}\circ\Phi^{-1}(x,\omega)$ on the domain $Q_{N=5}$, for $\eta = 0.05$.

When $\eta \to 0$, the estimator $e^{h,N}$ is bounded both in expectation and in variance, by a small constant, thereby showing the efficiency of the perturbative approach.

η	$\left(A^{\star,h,N}\right)_{11}$	$\left(e^{h,N}\right)_{11}$
0.1	3.073 $\pm$ 0.00928	- 4.233 $\pm$ 0.216
0.01	2.839 $\pm$ 0.00111	- 5.009 $\pm$ 0.254
0.001	2.812 $\pm$ 0.000113	- 5.104 $\pm$ 0.259
0.0001	2.809 $\pm$ 0.0000113	- 5.114 $\pm$ 0.259

TAB. 2.2 – Values of $\left(A^{\star,h,N}\right)_{11}$ and $\left(e^{h,N}\right)_{11}$ in function of η, for $h = 1/3$ and $N = 20$. Qualitatively similar results are obtained for the other entries of the matrix.

2.3.3 MsFEM-type approach

We now present the Multiscale Finite Element Method (MsFEM), which is designed to directly address the original elliptic problem, keeping ε at its fixed value. For clarity, we begin by presenting the approach in a deterministic setting. We will next turn to the stochastic setting. We refer to [65] for more details.

2.3.3.1 Deterministic setting

We consider problem (2.1), which we reproduce here for convenience,

$$\begin{cases} -\mathrm{div}(A^\varepsilon(x)\nabla u^\varepsilon(x)) = f(x) & \text{in } \mathcal{D}, \\ u^\varepsilon = 0 & \text{on } \partial\mathcal{D}, \end{cases} \tag{2.102}$$

where $A^\varepsilon \in (L^\infty(\mathcal{D}))^d$ is a symmetric matrix satisfying the standard coercivity and boundedness conditions. Note that the MsFEM approach we describe is not restricted to the periodic setting, so we do not assume that $A^\varepsilon(x) = A_{per}(\frac{x}{\varepsilon})$ for a periodic matrix A_{per}.

Rather than letting ε go to 0 (and approximate u^ε by the solution of the homogenized problem, see Section 2.2.1) we wish here to keep ε fixed at a (small) given value. The MsFEM approach consists in performing a variational approximation of (2.102) where the basis functions are defined *numerically* and *encode the fast oscillations* present in (2.102). In the sequel we will argue on the variational formulation of (2.102)

$$\text{Find } u^\varepsilon \in H^1_0(\mathcal{D}) \text{ such that : } \quad a_\varepsilon(u^\varepsilon, v) = b(v) \quad \forall v \in H^1_0(\mathcal{D}), \tag{2.103}$$

where

$$a_\varepsilon(u,v) = \sum_{i,j} \int_{\mathcal{D}} A^\varepsilon_{ij} \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} dx \text{ and } b(v) = \int_{\mathcal{D}} f\, v\, dx.$$

The MsFEM consists of the following three steps :

1. we introduce a standard discretization of the domain $\mathcal{D}$ using a coarse mesh,
2. for each element $\mathbf{K}$ of the coarse grid we compute the basis function $\phi_i^{\varepsilon,\mathbf{K}}$ as the solution of an elliptic equation posed in $\mathbf{K}$,
3. we solve a Galerkin approximation of (2.103), where the basis functions are those defined at step 2.

Definition of the coarse grid We consider a classical P1 discretization of the domain $\mathcal{D}$, denoted by $\mathcal{T}_h$, with N nodes. Let $\phi_i^0, i = 1, \cdots, N$, be the basis functions. We define the restriction of these functions in the element $\mathbf{K}$ by

$$\phi_i^{0,\mathbf{K}} := \phi_i^0|_{\mathbf{K}}.$$

Definition of the MsFEM basis Several definitions of the MsFEM basis functions have been proposed in the literature (see *e.g.* [3, 36, 37, 55]). They give rise to different methods. We present in the following one of these methods. We consider the problem

$$\begin{array}{l} \text{Find } \phi_i^{\varepsilon,\mathbf{K}} \in H^1(\mathbf{K}) \text{ such that :} \\ \left\{ \begin{array}{ll} -\text{div}(A^\varepsilon(x)\nabla\phi_i^{\varepsilon,\mathbf{K}}) = 0 & \text{in } \mathbf{K}, \\ \phi_i^{\varepsilon,\mathbf{K}} = \phi_i^{0,\mathbf{K}} & \text{on } \partial\mathbf{K}, \end{array} \right. \end{array} \tag{2.104}$$

which, in practice, is numerically solved *e.g.* using a finite element method with a mesh size adapted to the small scale ε. Note the similarity between (2.104) and the corrector problem (see equation (2.2)). As there exists a function $\phi_i^0 \in H^1(\mathcal{D})$ such that $\phi_i^{0,\mathbf{K}} = \phi_i^0|_{\mathbf{K}}$ for all $\mathbf{K}$, there also exists a function $\phi_i^\varepsilon \in H^1(\mathcal{D})$ such that $\phi_i^{\varepsilon,\mathbf{K}} = \phi_i^\varepsilon|_{\mathbf{K}}$ for all $\mathbf{K}$.

Note that the problems (2.104), indexed by $\mathbf{K}$, are all independent from one another. They can hence be solved in parallel.

Remark 2.3.3 *It has been observed that the boundary conditions in (2.104) give rise to an error called* resonance error*, which can be avoided using an oversampling technique. We refer to the literature (for instance [55]) for more details.*

Macro scale problem We now introduce the finite dimensional space

$$\mathcal{W}_h := span\,\{\phi_i^\varepsilon\} \subset H_0^1(\mathcal{D}),$$

and proceed with a standard Galerkin approximation of (2.103) using $\mathcal{W}_h$

$$\text{Find } u_h^\varepsilon \in \mathcal{W}_h \text{ such that : } \quad a_\varepsilon(u_h^\varepsilon, v) = b(v) \quad \forall v \in \mathcal{W}_h. \tag{2.105}$$

Observe that, due to the choice of boundary conditions in (2.104), $\phi_i^\varepsilon \in H^1(\mathcal{D})$, hence (2.105) is a conforming Galerkin approximation. The dimension of $\mathcal{W}_h$ is equal to N : the formulation (2.105) hence requires solving a linear system with only a limited number of degrees of freedom.

Numerical illustration In order to illustrate the MsFEM approach, we solve (2.102) in a one dimensional setting with

$$A^\varepsilon(x) = 5 + 50\sin^2\left(\frac{\pi x}{\varepsilon}\right),$$

on the domain $\mathcal{D} = [0, 1]$, with $\varepsilon = 0.025$ and $f = 1000$. We subdivide the interval $[0, 1]$ in $N = 10$ elements (i.e. $h = 0.1$). On Figure 2.8, we plot the MsFEM basis functions in a reference element and, on Figure 2.9, the solution u_h^ε in the domain $[0, 1]$. We can see 4 periods on Figure 2.8, which is consistent with the fact that $\frac{h}{\varepsilon} = 4$.

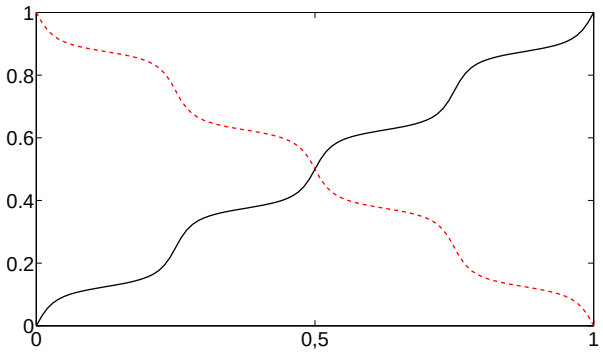


FIG. 2.8 – Multiscale basis functions $\phi^{\varepsilon,\mathbf{K}}$.

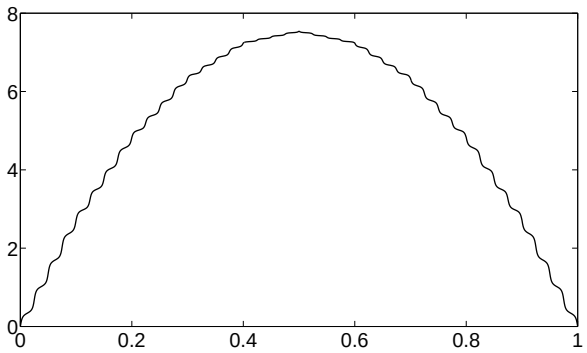


FIG. 2.9 – MsFEM solution u_h^ε.

2.3.3.2 Stochastic setting

We now consider problem (2.31)

$$\begin{array}{l}\text{Find } u^\varepsilon(\cdot,\omega) \in H^1(\mathcal{D}) \text{ such that :} \\ \left\{\begin{array}{ll} -\text{div}(A^\varepsilon(x,\omega)\nabla u^\varepsilon(x,\omega)) = f(x) & \text{in } \mathcal{D} \\ u^\varepsilon = 0 & \text{on } \partial\mathcal{D}. \end{array}\right.\end{array} \tag{2.106}$$

We are interested here in the statistical properties of the solution $u^\varepsilon(x,\omega)$ of (2.106), such as its mean $\mathbb{E}(u^\varepsilon(x,\cdot))$. A direct approach consists in building an estimate of the mean with a Monte-Carlo simulation method, *i.e.*, for each random realization m, first constructing a MsFEM basis and next solving the macro scale problem. We

denote $u_M^m(\cdot,\omega)$ the solution obtained. This approach requires a significantly large number of computations.

In the case when the matrix A^ε in (2.106) is a *perturbation of a deterministic* matrix, an alternative method can be proposed (see [65]). Consider

$$A^\varepsilon(.,\omega) = A^\varepsilon_\eta(.,\omega) = A^\varepsilon_{det}(.) + \eta\, C^\varepsilon(.,\omega)$$

(with obvious notation and properties) where $\eta \in \mathbb{R}$ is a small parameter. The matrix A^ε_η is hence a perturbation of A^ε_{det}. The principle of the proposed approach is to compute the MsFEM basis functions with the deterministic part of the matrix A^ε_η and next to perform Monte-Carlo realizations only for the macro scale problems. As in Section 2.3.3.1, we hence first solve (2.104), with $A^\varepsilon = A^\varepsilon_{det}$, and construct the finite dimensional space

$$\mathcal{W}_h := span\,\{\phi_i^\varepsilon\} \subset H_0^1(\mathcal{D}).$$

We next proceed with a standard Galerkin approximation of (2.106) using $\mathcal{W}_h$: for each $m \in \{1,\cdots,M\}$ we consider a realization $A^\varepsilon_{\eta,m}(\cdot,\omega)$ and compute $u_S^m(\cdot,\omega) \in \mathcal{W}_h$ such that

$$\sum_{i,j} \int_{\mathcal{D}} \left[A^\varepsilon_{\eta,m}\right]_{ij}(x,\omega) \frac{\partial u_S^m}{\partial x_i}(x,\omega) \frac{\partial v}{\partial x_j}(x)\, dx = \int_{\mathcal{D}} f(x)\, v(x)\, dx \quad \forall v \in \mathcal{W}_h. \tag{2.107}$$

Since the MsFEM basis functions are only computed once, a large computational gain is expected, and this is indeed the case.

Numerical studies We now estimate the performance of the approach. To this aim we compare the solution u_S of the above approach with the solution u_M of the direct approach and, for reference, the solution u_{ref} obtained using a finite element method with a mesh size adapted to the small scale ε. Our estimators are built as follows

$$e(u_1,u_2) = \mathbb{E}\left(\frac{||u_1-u_2||_N}{||u_2||_N}\right), \tag{2.108}$$

where N is the norm employed, u_1 and u_2 are the solutions obtained with any two different methods. The expectation is in turn computed using a Monte-Carlo method. Considering M realizations $\{X_m(\omega)\}_{1\le m\le M}$ of a random variable (here $X(\omega) = \dfrac{||u_1(\cdot,\omega)-u_2(\cdot,\omega)||_N}{||u_2(\cdot,\omega)||_N}$), we compute the empirical mean μ_M and the empirical standard deviation σ_M as

$$\mu_M(X) = \frac{1}{M}\sum_{m=1}^M X_m(\omega), \tag{2.109}$$

$$\sigma_M^2(X) = \frac{1}{M-1}\sum_{m=1}^M \left(X_m(\omega)-\mu_M(X)\right)^2. \tag{2.110}$$

As a classical consequence of the Central Limit Theorem, it is commonly admitted that $\mathbb{E}(X)$ satisfies

$$|\mathbb{E}(X) - \mu_M(X)| \leq 1,96\frac{\sigma_M}{\sqrt{M}}. \tag{2.111}$$

We consider the following numerical example. Let $(X_{k,l})_{(k,l)\in\mathbb{Z}^2}$ denote a sequence of independent, identically distributed scalar random variables uniformly distributed over the interval $[0,1]$. We construct the random conductivity matrix as follows

$$\begin{aligned} A_\eta^\varepsilon(x,y,\omega) \;=\; & \sum_{(k,l)\in\mathbb{Z}^2} \mathbf{1}_{]k,k+1]}(\frac{x}{\varepsilon})\mathbf{1}_{]l,l+1]}(\frac{y}{\varepsilon}) \left(\frac{2+P\sin(2\pi x/\varepsilon)}{2+P\sin(2\pi y/\varepsilon)}\right. \\ & \left. +\frac{2+\sin(2\pi y/\varepsilon)}{2+P\sin(2\pi x/\varepsilon)}\right)(1+\eta X_{k,l}(\omega))\,\mathbf{I}d_2, \end{aligned} \tag{2.112}$$

with parameters $P = 1.8$ and $\varepsilon = 0.025$. Then we compute u_{ref} solution to

$$\begin{cases} \operatorname{div}\left(A_\eta^\varepsilon \nabla u_{ref}\right) = -1 & \text{in } \mathcal{D}, \\ u_{ref} = 0 & \text{on } \partial\mathcal{D}, \end{cases} \tag{2.113}$$

on the domain $\mathcal{D} = [0,1]^2$. Let u_M and u_S be its approximation by the two MsFEM approaches described above. The numerical parameters for the computation are determined using an empirical study of convergence. We used for the reference solution a fine mesh of size $h_f = \varepsilon/40$. The MsFEM basis functions are computed in each element $\mathbf{K}$ using a mesh of size $h_M = \varepsilon/80$. The coarse mesh size is $H = 1/30$. We consider $M = 30$ realizations.

We report in Tables 2.3 and 2.4 the estimator (2.108), along with its confidence interval, for the norms $H^1(\mathcal{D})$ and $L^2(\mathcal{D})$, respectively.

η	$e(u_M, u_{ref})$	$e(u_S, u_{ref})$	$e(u_S, u_M)$
1	8.12±0.19	17.37±0.78	15.51±0.87
0.1	7.17±0.02	7.62±0.07	2.56±0.10
0.01	7.15±0.002	7.28±0.007	1.39±0.002

TAB. 2.3 – $H^1(\mathcal{D})$ error (in %)

η	$e(u_M, u_{ref})$	$e(u_S, u_{ref})$	$e(u_S, u_M)$
1	0.56±0.08	1.69±0.49	1.47±0.50
0.1	0.54±0.01	0.57±0.06	0.20±0.07
0.01	0.53±0.001	0.62±0.005	0.11±0.003

TAB. 2.4 – $L^2(\mathcal{D})$ error (in %)

We observe that, when η is small, the alternative approach provides a function u_S that is an approximation of u_{ref} as accurate as u_M. Recall that, since the MsFEM

basis has only been computed once, the cost for obtaining an empirical approximation of $\mathbb{E}(u_S)$ is much smaller than that for obtaining the corresponding empirical estimator of $\mathbb{E}(u_M)$. This demonstrates the efficiency of the approach.

2.4 Variance issues

Suppose now that no satisfactory approximation can be performed using periodic problems in the manner described in Section 2.3. The genuine random problem (2.31) needs to be addressed. As we briefly mentioned in Section 2.2.2, variance issues are now central to the practical approximation. We examine here the situation and show a possible appropriate technique that, for finite N, allows to compute the approximation $A^\star_N$ more effectively, that is, with a smaller variance. The present section summarizes arguments and computations more exhaustively developed in [P2, P3, P4].

A quick overview of the literature The issue of variance in stochastic homogenization is not new, although the concern has long been a theoretical concern. It is not our purpose here to review in details the important contributions existing in the literature. We however wish to cite some results particularly relevant to our own study :

- the original contribution [103] by Yurinski, where the convergence of some truncated approximation of $A^\star$ is established, along with an estimate of the rate of convergence (in short, problem (2.33) is regularized and then $A^\star$ is approximated on a bounded domain),
- a similar study [23] by Bourgeat and Piatnitsky for a specific approximation more relevant to actual numerical practice (in short, both problem (2.33) and the integral in (2.32) are truncated as in (2.42)–(2.43)),
- the work [76] by Naddaf and Spencer on a discrete ("lattice-type") approximation of the differential operator present in the original problem (2.31),
- and the enterprise by Gloria and Otto (see [49] for homogenization problems set on random lattices and publications announced in preparation for some problems for differential operators) to establish sharp estimates of the convergence of the numerical approximation in terms of size of the truncation domain and other discretization parameters.

In all the above works, the convergence and the rate of convergence are studied. We take here the problem from a slightly different perspective : we are interested in *basic practical* issues. Can we improve the prefactor in the convergence of $A^\star_N$ to $A^\star$ as $N \to +\infty$ (loosely stated, the variance in a Central Limit Theorem type result) ? or, even more practically, can we reduce the interval of confidence for empirical means approximating $\mathbb{E}(A^\star_N)$? and similar issues.

We present here a first attempt to reduce the variance in stochastic homogenization. For this purpose, we consider a simple situation, and a simple variance

reduction technique. The equation under consideration is the same simple elliptic equation in divergence form considered above, with a scalar coefficient. The coefficient is assumed to consist of independent, identically distributed random variables set on a simple mesh. The technique used for variance reduction is that of antithetic variables. Our setting is academic in nature, somewhat far from physically relevant cases, and elementary. Many more difficult situations could be addressed : other types of stationary ergodic coefficients, other types of equations, other techniques for variance reduction, ...

Variance reduction using antithetic variables As we know, the direct approach for numerical random homogenization consists in using empirical means approximating $\mathbb{E}\left(A_N^\star\right)$.

In generality, fix $M = 2\mathcal{M}$. Suppose that we give ourselves $\mathcal{M}$ i.i.d. copies $\left(A^{\mathbf{m}}(x,\omega)\right)_{1\leq \mathbf{m}\leq \mathcal{M}}$ of $A(x,\omega)$. Construct next $\mathcal{M}$ i.i.d. *antithetic* random fields

$$B^{\mathbf{m}}(x,\omega) = T\left(A^{\mathbf{m}}(x,\omega)\right), \quad 1 \leq \mathbf{m} \leq \mathcal{M},$$

from the $\left(A^{\mathbf{m}}(x,\omega)\right)_{1\leq \mathbf{m}\leq \mathcal{M}}$. The map T transforms the random field $A^{\mathbf{m}}$ into another, so-called *antithetic*, field $B^{\mathbf{m}}$. The transformation is performed in such a way that, for each $\mathbf{m}$, $B^{\mathbf{m}}$ has the same law as $A^{\mathbf{m}}$, namely the law of the matrix A. Somewhat vaguely stated, if A was obtained in a coin tossing game (using a fair coin), then $B^{\mathbf{m}}$ would be head each time $A^{\mathbf{m}}$ is tail and vice versa. We refer the reader to (2.119)-(2.121) below for an explicit example of such a construction. Then, for each $1 \leq \mathbf{m} \leq \mathcal{M}$, we solve two corrector problems. One is associated to the original $A^{\mathbf{m}}$, the other one is associated to the antithetic field $B^{\mathbf{m}}$. Using its solution $v_p^{N,\mathbf{m}}$, we define the *antithetic homogenized matrix* $B_N^{\star,\mathbf{m}}$, whose elements read, for $1 \leq i,j \leq d$,

$$\left[B_N^{\star,\mathbf{m}}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}\left(e_i + \nabla v_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T B^{\mathbf{m}}(\cdot,\omega)\,\left(e_j + \nabla v_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right).$$

And we finally set, for any $1 \leq \mathbf{m} \leq \mathcal{M}$,

$$\widetilde{A}_N^{\star,\mathbf{m}}(\omega) := \frac{1}{2}\left(A_N^{\star,\mathbf{m}}(\omega) + B_N^{\star,\mathbf{m}}(\omega)\right). \tag{2.114}$$

Since $A^{\mathbf{m}}$ and $B^{\mathbf{m}}$ are identically distributed, so are $A_N^{\star,\mathbf{m}}$ and $B_N^{\star,\mathbf{m}}$. Thus, $\widetilde{A}_N^{\star,\mathbf{m}}$ is unbiased (that is, $\mathbb{E}\left(\widetilde{A}_N^{\star,\mathbf{m}}\right) = \mathbb{E}\left(A_N^{\star,\mathbf{m}}\right)$). In addition, it satisfies :

$$\widetilde{A}_N^{\star,\mathbf{m}} \underset{N\to+\infty}{\longrightarrow} A^\star \text{ almost surely,}$$

because $B^{\mathbf{m}}$ is ergodic.

Let us define new estimators

$$\begin{aligned}\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) &= \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\left[\widetilde{A}_N^{\star,\mathbf{m}}\right]_{ij},\\ \sigma_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) &= \frac{1}{\mathcal{M}-1}\sum_{\mathbf{m}=1}^{\mathcal{M}}\left(\left[\widetilde{A}_N^{\star,\mathbf{m}}\right]_{ij}-\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)\right)^2,\end{aligned} \tag{2.115}$$

which require $2\mathcal{M}$ resolutions of corrector problems, i.e. as many as the classical estimators (2.44), since we choose $M = 2\mathcal{M}$. In addition, note that we have built a new random variable whose variance is

$$\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) = \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\left[A_N^\star\right]_{ij}\right) + \frac{1}{2}\mathbb{C}\mathrm{ov}\left(\left[A_N^\star\right]_{ij},\left[B_N^\star\right]_{ij}\right). \tag{2.116}$$

Applying the central limit theorem to $\widetilde{A}_N^\star$, we obtain

$$\sqrt{\mathcal{M}}\left(\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)-\mathbb{E}\left(\left[A_N^\star\right]_{ij}\right)\right)\xrightarrow[\mathcal{M}\to+\infty]{\mathcal{L}}\sqrt{\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)}\,\mathcal{N}(0,1). \tag{2.117}$$

Similarly to (2.46), we deduce a confidence interval from this convergence. The exact mean $\mathbb{E}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)$ is equal to $\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)$ within a margin of error $1.96\dfrac{\sqrt{\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)}}{\sqrt{\mathcal{M}}}$. It results from (2.116) that, if

$$\mathbb{C}\mathrm{ov}\left(\left[A_N^\star\right]_{ij},\left[B_N^\star\right]_{ij}\right)\leq 0, \tag{2.118}$$

then the width of this interval has been reduced by the new approach, and, correspondingly, the quality of approximation at given computational cost has increased.

2.4.1 One-dimensional situation

To understand slightly more in details at the theoretical level why the approach is likely to perform well, we again consider the one-dimensional setting for which we recall we have all necessary explicit expressions. Suppose for illustration that

$$a(x,\omega) = \sum_{k\in\mathbb{Z}} a_k(\omega)\mathbf{1}_{[k,k+1[}(x), \tag{2.119}$$

where $a_k(\omega)$ is a family of independent random variables, identically distributed according to a Bernoulli law :

$$\mathbb{P}(a_0=\alpha)=1/2 \quad\text{and}\quad \mathbb{P}(a_0=\beta)=1/2,$$

for some $0 < \alpha < \beta$. Defining the antithetic variable

$$b_k(\omega) = \alpha + \beta - a_k(\omega) \tag{2.120}$$

and next the antithetic field

$$b(x,\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1[}(x)\, b_k(\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1[}(x)\, (\alpha+\beta-a_k(\omega))\,, \tag{2.121}$$

it is immediately seen that

$$\frac{1}{2}\left(\frac{1}{a_N^\star(\omega)} + \frac{1}{b_N^\star(\omega)}\right) = \mathbb{E}\left(\frac{1}{a_0}\right) = \frac{1}{a^\star}.$$

The variance of the inverse of the truncated coefficient has vanished. This example might seem oversimplified because we are indeed making use of two peculiarities of the problem : the set $\{\alpha,\beta\}$ of values taken by the coefficient a has cardinality two, and we may explicitly manipulate the inverse of the homogenized coefficient. The situation, although oversimplified, is yet a first good indicator of the interest of the approach. We can be slightly more general, by considering for instance that the random coefficient a is now uniformly distributed over a given interval, say $a_0 \sim \mathcal{U}([\alpha,\beta])$. Then,

$$\frac{1}{2}\left(\frac{1}{a_N^\star(\omega)} + \frac{1}{b_N^\star(\omega)}\right) = \frac{1}{2N}\sum_{k=-N}^{N-1} \frac{1}{2}\left(\frac{1}{a_k(\omega)} + \frac{1}{b_k(\omega)}\right), \tag{2.122}$$

where the antithetic variable $b_k(\omega)$ is again defined by (2.120). It is a simple matter (see [P3] for details) to show that, because the function $x \mapsto 1/x$ is decreasing, we have

$$\mathbb{C}\mathrm{ov}\left(\frac{1}{a_0}, \frac{1}{b_0}\right) \leq 0.$$

Since

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\frac{1}{a_N^\star} + \frac{1}{b_N^\star}\right)\right) = \frac{1}{4N}\mathbb{V}\mathrm{ar}\left(\frac{1}{a_0}\right) + \frac{1}{4N}\mathbb{C}\mathrm{ov}\left(\frac{1}{a_0}, \frac{1}{b_0}\right),$$

we conclude that

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\frac{1}{a_N^\star} + \frac{1}{b_N^\star}\right)\right) \leq \mathbb{V}\mathrm{ar}\left(\frac{1}{a_{2N}^\star}\right).$$

Therefore, $\mathbb{E}(1/a_0)$ can be approximated either by (2.122) or by $1/a_{2N}^\star$, with an equal cost (i.e. an equal number of random variables in both sums), but the former has a smaller variance than the latter. It is hence of better quality.

As mentioned above, the practice in dimensions higher than one is to generate a set of identically distributed coefficients for each truncated corrector problem, and

to use (2.114). The appropriate analogous one-dimensional approach is to consider $\mathcal{M} = \dfrac{M}{2}$ independent copies of $a(x,\omega)$ and set

$$\begin{aligned}\widetilde{a}_N^{\star,\mathbf{m}}(\omega) &:= \frac{1}{2}\left(a_N^{\star,\mathbf{m}}(\omega) + b_N^{\star,\mathbf{m}}(\omega)\right) \\ &= \frac{1}{2}\left(\frac{1}{2N}\sum_{k=-N}^{N-1}\frac{1}{a_k^{\mathbf{m}}(\omega)}\right)^{-1} + \frac{1}{2}\left(\frac{1}{2N}\sum_{k=-N}^{N-1}\frac{1}{b_k^{\mathbf{m}}(\omega)}\right)^{-1}\end{aligned}$$

with empirical mean

$$\mu_{\mathcal{M}}\left(\widetilde{a}_N^{\star}\right)(\omega) = \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\widetilde{a}_N^{\star,\mathbf{m}}(\omega).$$

We approach more generality since

$$\mu_{\mathcal{M}}\left(\widetilde{a}_N^{\star}\right)(\omega) \underset{\mathcal{M}\to+\infty}{\longrightarrow} \mathbb{E}\left(\widetilde{a}_N^{\star}\right) = \mathbb{E}\left(a_N^{\star}\right) \quad \text{almost surely,}$$

but $\mathbb{E}\left(a_N^{\star}\right) \neq a^{\star}$. It can again be remarked that $a_N^{\star}(\omega)$ is an increasing function of the uniform variables $(a_k(\omega))_{k\in\mathbb{Z}}$. From this observation, it is possible to show that $\mathbb{C}\mathrm{ov}\left(a_N^{\star}, b_N^{\star}\right) \leq 0$, and to conclude that the variance of $\mu_{\mathcal{M}}\left(\widetilde{a}_N^{\star}\right)$ is smaller than that of $\mu_{2\mathcal{M}}\left(a_N^{\star}\right)$.

2.4.2 Elements of proof for two simple situations

Since considering one-dimensional settings is obviously not sufficient to have a good theoretical insight of the variance reduction approach using antithetic variables, we now continue our investigation by considering two simple cases in dimension higher than one. The first case is a "genuinely" random setting. The second case is a "weakly random" case, in the spirit of a case already considered previously in Section 2.3.

It is important to note that some of the results mentioned here are limited to the technique of antithetic variables. Some others are not. The latter can therefore be useful for other variance reduction techniques.

2.4.2.1 A strongly random case

We consider a random matrix A constructed with independent, identically distributed random variables on the cells of our periodic lattice (although A needs not be constant on each cell and equal to these random variables; see *e.g.* example (2.126) below). We consider the truncations (2.42)–(2.43) on a finite domain $Q_N = [-N, N]^d$, solved for a set of realizations of the random matrix. Empirical means of the truncated homogenized matrix $A_N^{\star}(\omega)$ are obtained, along with an (approximation of an) interval of confidence involving the variance. The consideration of

antithetic variables allows to improve the approximation. This is numerically observed, and theoretically investigated, in [P2, P4, P3]. We establish there theoretically that the variance of the homogenized objects is indeed diminished by the technique. The present section summarizes our main results in this setting.

We make here the following two assumptions on the matrix A of (2.31). First, we assume that, for any N, there exists an integer n (possibly $n = |Q_N|$, but not necessarily) and a function $\mathcal{A}$, defined on $Q_N \times \mathbb{R}^n$, such that the matrix $A(x,\omega)$ writes

$$\forall x \in Q_N, \quad A(x,\omega) = \mathcal{A}(x, X_1(\omega), \ldots, X_n(\omega)) \quad \text{a.s.}, \tag{2.123}$$

where $\{X_k(\omega)\}_{1\le k\le n}$ are independent scalar random variables, which are all distributed according to the uniform law $\mathcal{U}[0,1]$. In general, the function $\mathcal{A}$, as well as the number n of independent, identically distributed variables involved in (2.123), depend on N, the size of Q_N, although this dependency is not made explicit in (2.123).

Second, we assume that the function $\mathcal{A}$ in (2.123) is such that, for all $x \in Q_N$, and any vector $\xi \in \mathbb{R}^d$, the map

$$(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \xi^T \mathcal{A}(x, x_1, \ldots, x_n)\xi \tag{2.124}$$

is non-decreasing with respect to each of its arguments.

A typical example of matrix A that satisfies the above assumptions is the case we considered several times so far in these notes. Consider a family $(a_k(\omega))_{k\in\mathbb{Z}^d}$ of independent, identically distributed random variables, and set

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega)\,\mathrm{Id}, \tag{2.125}$$

where Q is the unit cube of $\mathbb{R}^d$, centered at the origin, and $Q+k$ is the cube Q translated by the vector $k \in \mathbb{Z}^d$. We assume that there exist $\alpha > 0$ and $\beta < \infty$ such that, for all k, $0 < \alpha \le a_k \le \beta < +\infty$ almost surely. Consequently, A is uniformly coercive and bounded. Example (2.125) corresponds to a spherical matrix $A(x,\omega)$ that is constant in each cube $Q+k$, with independent, identically distributed values $a_k(\omega)$. We refer to [P3] for showing that Example (2.125) indeed satisfies (2.123)–(2.124). Of course, Example (2.125) can be readily extended to the case of non-spherical matrices.

An example using variables that, on each cell, are not equal to independent, identically distributed variables is as follows. For this purpose, define positive coefficients κ_p for $|p|_\infty \le 1$, and consider a non-decreasing function f. We then set

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \left[\mathbf{1}_{Q+k}(x) \sum_{j\in\mathbb{Z}^d,\ |j-k|_\infty\le 1} \kappa_{j-k}\ f(X_j(\omega))\ \mathrm{Id} \right], \tag{2.126}$$

which clearly satisfies (2.123)–(2.124). In (2.126), it is immediately seen that the value of $A(x,\omega)$ in the cell $Q+k$ is a local average of the values $f(X_j(\omega))\,\mathrm{Id}$, for $|j-k|_\infty \le 1$.

Our main results (from [P3]) are :

Proposition 2.4.1 *We assume* (2.123)–(2.124). *Let* $A_N^\star(\omega)$ *be the approximated homogenized matrix, obtained by solving the corrector problem* (2.43) *on the truncated domain* $Q_N \subset \mathbb{R}^d$. *We define on* Q_N *the field*

$$B(x,\omega) := \mathcal{A}(x, 1 - X_1(\omega), \ldots, 1 - X_n(\omega)),$$

antithetic to $A(\cdot,\omega)$ *defined by* (2.123). *We associate to this field the corrector problem* (2.43) *(replacing* A *by* B*), the solution of which is denoted by* v_p^N, *and the matrix* $B_N^\star(\omega)$, *defined from* v_p^N *using* (2.42). *Recall that* $\widetilde{A}_N^\star(\omega)$ *is defined in* (2.114). *Then*

$$\mathbb{E}\left(\widetilde{A}_N^\star\right) = \mathbb{E}\left(A_N^\star\right),$$

and, for any $\xi \in \mathbb{R}^d$,

$$\mathbb{V}ar\left(\xi^T \widetilde{A}_N^\star \xi\right) \leq \frac{1}{2}\mathbb{V}ar\left(\xi^T A_N^\star \xi\right). \tag{2.127}$$

Otherwise stated, $\widetilde{A}_N^\star$ *is an unbiased estimator of* $\mathbb{E}\left(A_N^\star\right)$, *and its variance is smaller than half of that of* $A_N^\star$, *in the sense of* (2.127).

Taking $\xi = e_i$ in (2.127) implies that the variance of the diagonal coefficient $\left[\widetilde{A}_N^\star\right]_{ii}$ is reduced by a factor two. Similar results can be obtained, first on the eigenvalues of the matrix $A_N^\star$ (as well as on its trace and its determinant, see Remark 2 in [P3])), and second on the eigenvalues of the (approximate) homogenized elliptic operator $L_A = -\text{div}\,[A_N^\star(\omega)\nabla\,\cdot]$. This is the purpose of the following two corollaries.

Corollary 2.4.1 *We assume* (2.123)–(2.124). *Denote by* $\{\lambda_k(A,\omega)\}_{1\leq k\leq d}$ *and* $\{\lambda_k(B,\omega)\}_{1\leq k\leq d}$ *the eigenvalues of* $A_N^\star(\omega)$ *and* $B_N^\star(\omega)$ *respectively, sorted in non-decreasing order, where* $A_N^\star(\omega)$ *and* $B_N^\star(\omega)$ *are defined as in Proposition 2.4.1. Define*

$$\widetilde{\lambda}_k(\omega) := \frac{1}{2}\big[\lambda_k(A,\omega) + \lambda_k(B,\omega)\big].$$

Then, for all $1 \leq k \leq d$,

$$\mathbb{E}\left(\widetilde{\lambda}_k\right) = \mathbb{E}\left(\lambda_k(A,\cdot)\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\lambda}_k\right) \leq \frac{1}{2}\mathbb{V}ar\left(\lambda_k(A,\cdot)\right).$$

We likewise have

Corollary 2.4.2 *We assume* (2.123)–(2.124). *Let* $(\lambda_k(L_A,\omega), u_k(L_A,\omega))_{k\in\mathbb{N}}$ *be the eigenelements of the operator* $L_A = -\text{div}\,[A_N^\star(\omega)\nabla\cdot]$ *with homogeneous Dirichlet boundary conditions, i.e.*

$$-\text{div}\,[A_N^\star(\omega)\nabla u_k(L_A,\omega)] = \lambda_k(L_A,\omega)\, u_k(L_A,\omega)$$

with $u_k(L_A,\omega) \in H^1_0(\mathcal{D})$ *and* $\|u_k(L_A,\omega)\|_{L^2(\mathcal{D})} = 1$. *We similarly consider the eigenelements of* $L_B = -\text{div}\,[B^\star_N(\omega)\nabla\cdot]$ *:*

$$-\text{div}\,[B^\star_N(\omega)\nabla u_k(L_B,\omega)] = \lambda_k(L_B,\omega)\,u_k(L_B,\omega).$$

We assume that, almost surely, $\lambda_k(L_A,\omega)$ *and* $\lambda_k(L_B,\omega)$ *are sorted in non-decreasing order. Define*

$$\widetilde{\lambda}_k(L,\omega) := \frac{1}{2}\left(\lambda_k(L_A,\omega) + \lambda_k(L_B,\omega)\right).$$

Then, for all $k \in \mathbb{N}$,

$$\mathbb{E}\left(\widetilde{\lambda}_k(L,\cdot)\right) = \mathbb{E}\left(\lambda_k(L_A,\cdot)\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\lambda}_k(L,\cdot)\right) \leq \frac{1}{2}\mathbb{V}ar\left(\lambda_k(L_A,\cdot)\right).$$

The proofs of the above results are given in [P3]. They are obtained combining some classical results on variance reduction using antithetic variables [69] and some monotonicity results from the theory of homogenization.

The proof goes as follows. First, we recall that the technique of antithetic variables reduces variance for the computation of $\mathbb{E}(f(X_1,\ldots,X_n))$, when f is a real-valued function, that is *non-decreasing* of each of its arguments, and $X = (X_1,\ldots,X_n)$ is a vector of independent random variables. Second, we assume that the matrix field $A(x,\omega)$ of (2.31) writes as a non-decreasing function (in the sense of symmetric positive matrices) of independent random variables $X_k(\omega)$ (these are assumptions (2.123) and (2.124)). Then, we use that the homogenization process *preserves the order of symmetric matrices* to claim that $A^\star_N(\omega)$ is likewise a non-decreasing function of the random variables $X_k(\omega)$. Consequently, we obtain variance reduction for $A^\star_N$. In the same vein, since the map that associates to a symmetric matrix its eigenvalues is increasing, we obtain variance reduction for the eigenvalues of $A^\star_N$.

2.4.2.2 A weakly random case

Our second setting is a "weakly" random setting. We mean that the random matrix A is a small perturbation of a deterministic, periodic matrix. Consequently, the solution of the problem is only seeked *at the first order in the size of the perturbation.* A similar setting has been introduced in Section 2.3 above. Its practical interest is that the computation comes down to a set of *fully deterministic* computations. So in practice, no variance issue is relevant. We however consider this case *pretending* not to exploit the simplification : we treat the problem stochastically and prove that the technique of variance reduction still works. This is useful to get theoretical insight on the technique when applied to more general cases.

We now assume, somewhat in the vein of what we did in Section 2.3 above, that the matrix A in (2.31) is a *perturbation* of a periodic matrix :

$$A(x,\omega) = A_{\text{per}}(x) + \eta A_1(x,\omega), \tag{2.128}$$

where $A_{\rm per}$ is Q-periodic, A_1 is stationary in the sense (2.28), and η is a deterministic, supposedly small, parameter. We assume that A, $A_{\rm per}$ and A_1 are all symmetric, uniformly coercive and bounded matrices. We also assume the following special structure for A_1 :

$$A_1(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x)\, X_k(\omega)\ \mathrm{Id}, \tag{2.129}$$

where Id denotes the identity matrix and $(X_k)_{k\in\mathbb{Z}^d}$ is a sequence of (scalar) independent, identically distributed random variables, satisfying, for some $\alpha > 0$ and $\beta < \infty$, the bound $0 < \alpha \leq X_k \leq \beta < +\infty$ almost surely and for all k.

Expansion for small η Since, in view of (2.128), A is a perturbation of $A_{\rm per}$, we may expand the homogenized matrix $A^\star$ in powers of η, as shown in Section 2.3. The setting is slightly different from that of Sections 2.3.1 and 2.3.2, but the techniques of proof are however similar, in particular to those of Section 2.3.2. For convenience, we recall that, first, the corrector ∇w_p, solution to (2.33), has an expansion in powers of η :

$$\nabla w_p = \nabla w_p^0 + \eta \nabla w_p^1 + \dots, \tag{2.130}$$

where w_p^0 is the unique (up to the addition of a constant) solution to

$$\left\{ \begin{array}{l} -\mathrm{div}\,\left[A_{\rm per}(\nabla w_p^0 + p)\right] = 0, \\ w_p^0 \text{ is } Q\text{-periodic}, \end{array} \right. \tag{2.131}$$

and w_p^1 is the unique (up to the addition of a random constant) solution to

$$\left\{ \begin{array}{l} -\mathrm{div}\,\left[A_{\rm per}\nabla w_p^1\right] = \mathrm{div}\,\left[A_1\left(\nabla w_p^0 + p\right)\right] \quad \text{a.s. on } \mathbb{R}^d, \\ \nabla w_p^1 \text{ is stationary}, \\ \displaystyle\int_Q \mathbb{E}(\nabla w_p^1) = 0. \end{array} \right. \tag{2.132}$$

Likewise, the homogenized matrix reads

$$A^\star = A^\star_{\rm per} + \eta A_1^\star + O(\eta^2), \tag{2.133}$$

where

$$\left[A^\star_{\rm per}\right]_{ij} = \int_Q (\nabla w^0_{e_i} + e_i)^T A_{\rm per}(\nabla w^0_{e_j} + e_j),$$

and

$$\begin{aligned} [A_1^\star]_{ij} \;=\; & \int_Q \mathbb{E}(\nabla w^1_{e_i})^T A_{\rm per}(\nabla w^0_{e_j} + e_j) + \int_Q (\nabla w^0_{e_i} + e_i)^T A_{\rm per}\mathbb{E}(\nabla w^1_{e_j}) \\ & + \int_Q (\nabla w^0_{e_i} + e_i)^T \mathbb{E}(A_1)(\nabla w^0_{e_j} + e_j). \end{aligned} \tag{2.134}$$

In essence, this specific setting does not give rise to any variance concerns, for two reasons at least. First, as observed in [20], and as evident on (2.134), the knowledge of $\overline{w}_p^1 := \mathbb{E}(w_p^1)$ is actually sufficient to compute $A_1^\star$. Taking expectations in (2.134), we see that $\overline{w}_p^1$ is solution to the closed, cell problem

$$\left\{ \begin{array}{l} -\text{div } \left[A_{\text{per}} \nabla \overline{w}_p^1\right] = \text{div } \left[\mathbb{E}(A_1)\left(\nabla w_p^0 + p\right)\right] \quad \text{on } \mathbb{R}^d, \\ \overline{w}_p^1 \text{ is } Q\text{-periodic}, \end{array} \right. \tag{2.135}$$

which is much easier to solve than the corrector problem (2.33) or its truncated version (2.43), since it is a *deterministic* problem set on a *single* cell. Hence, as pointed out above, the determination of $A^\star$ comes down, in practice, to solving the deterministic problems (2.131) and (2.135).

Second, in practice, the *exact* corrector w_p, which solves (2.33), is approximated by the solution w_p^N to the truncated problem (2.43), which also has an expansion in powers of η, as shown in [P1, P6] : similarly to (2.130), we have

$$\nabla w_p^N = \nabla w_p^0 + \eta \nabla w_p^{1,N} + \dots,$$

where $w_p^{1,N}$ solves the *truncated* problem

$$\left\{ \begin{array}{l} -\text{div } \left[A_{\text{per}} \nabla w_p^{1,N}\right] = \text{div } \left[A_1\left(\nabla w_p^0 + p\right)\right], \\ w_p^{1,N} \text{ is } Q_N - \text{periodic}. \end{array} \right. \tag{2.136}$$

In turn, the approximate homogenized matrix (2.42) satisfies

$$A_N^\star(\omega) = A_{\text{per}}^\star + \eta A_{1,N}^\star(\omega) + O(\eta^2),$$

where

$$\begin{aligned} \left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{1}{|Q_N|} \Bigg[& \int_{Q_N} \nabla w_{e_i}^{1,N}(x,\omega)^T A_{\text{per}}(x)(\nabla w_{e_j}^0(x) + e_j) dx \\ & + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x,\omega)(\nabla w_{e_j}^0(x) + e_j) dx \\ & + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_{\text{per}}(x) \nabla w_{e_j}^{1,N}(x,\omega) dx \Bigg]. \end{aligned} \tag{2.137}$$

Observe now that the first term of (2.137) reads

$$\begin{aligned} \int_{Q_N} \left(\nabla w_{e_i}^{1,N}\right)^T A_{\text{per}} \left(\nabla w_{e_j}^0 + e_j\right) &= -\int_{Q_N} w_{e_i}^{1,N} \text{div } \left[A_{\text{per}}(\nabla w_{e_j}^0 + e_j)\right] \\ &+ \int_{\partial Q_N} w_{e_i}^{1,N} \left[A_{\text{per}}(\nabla w_{e_j}^0 + e_j)\right] \cdot n. \end{aligned}$$

Using (2.131), we see that the first term of the right hand side vanishes. Since $w_{e_i}^{1,N}$, $A_{\rm per}$ and $w_{e_j}^0$ are Q_N-periodic, the second term vanishes as well. Hence (2.137) reads

$$\left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}(\nabla w_{e_i}^0(x)+e_i)^T A_1(x,\omega)(\nabla w_{e_j}^0(x)+e_j)dx. \qquad (2.138)$$

In the specific case (2.129), we thus have

$$\left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{\gamma_{ij}}{|Q_N|}\sum_{|k|\le N} X_k(\omega),$$

with $\gamma_{ij} = \displaystyle\int_Q \left(\nabla w_{e_i}^0 + e_i\right)^T \left(\nabla w_{e_j}^0 + e_j\right)$. The variables X_k being independent and identically distributed, we obtain that $\left[A_{1,N}^\star\right]_{ij}$ converges almost surely to $\gamma_{ij}\mathbb{E}(X_0)$, the rate of convergence being given by the central limit theorem. Of course, using (2.138), this argument is not restricted to the form (2.129) of A_1, and can be extended to more general cases.

Observe yet that the fact that the first and third terms of (2.137) vanish strongly relies on the specific equation (and boundary conditions) used to build a numerical approximation of the corrector, and may not hold true in a more general setting.

Estimates on the variance As announced above, our viewpoint in this section is *different* from that of Section 2.3. With a view to use (2.128)-(2.129) to test and further understand our variance reduction approach, we *pretend* not to exploit the various simplifications, and we thus treat the problem entirely stochastically.

Our aim is to show that, for any fixed N, applying the variance reduction strategy, we obtain a better estimate of the approximate homogenized matrix $\mathbb{E}(A_N^\star)$ *using empirical means* in the spirit of (2.44). When the number of independent realizations M increases to $+\infty$, the rates at which the empirical means converge to the expectation are identical, but the prefactor is better in the latter case. In addition, using the specificities of this setting, we are also able to analyze the convergence of the approximation procedure when N goes to $+\infty$.

Our main result (from [P3]) is the following proposition.

Proposition 2.4.2 *Let A be defined by* (2.128)*, where $A_{\rm per}$ is periodic, and A_1 satisfies* (2.129)*, with $(X_k)_{k\in\mathbb{Z}^d}$ a sequence of independent, identically distributed scalar random variables. Assume in addition that A, $A_{\rm per}$ and A_1 are symmetric, uniformly coercive and bounded matrices. For any $N\in\mathbb{N}$ and $1\le i,j\le d$, define*

$$\begin{aligned}\left[A_{1,N}^{\star,\rm exact}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\Bigg[&\int_{Q_N}\nabla w_{e_i}^1(x,\omega)^T A_{\rm per}(x)(\nabla w_{e_j}^0(x)+e_j)dx\\ &+\int_{Q_N}(\nabla w_{e_i}^0(x)+e_i)^T A_1(x,\omega)(\nabla w_{e_j}^0(x)+e_j)dx\\ &+\int_{Q_N}(\nabla w_{e_i}^0(x)+e_i)^T A_{\rm per}(x)\nabla w_{e_j}^1(x,\omega)dx\Bigg], \qquad (2.139)\end{aligned}$$

where w^0 and w^1 are solution to (2.131) *and* (2.134) *respectively, and $Q_N = \cup_{|k|_\infty \leq N}(Q+k)$ is the cube of size $(2N+1)^d$ centered at the origin. Then, there exist d^2 coefficients $C_N^{ij} > 0$, independent of $(X_k)_{k\in\mathbb{Z}^d}$, such that*

$$\mathbb{V}ar\left(\left[A_{1,N}^{\star,\text{exact}}\right]_{ij}\right) = C_N^{ij}\ \mathbb{V}ar(X_0). \tag{2.140}$$

In addition, we have

$$C_N^{ij} \leq \frac{C}{|Q_N|},$$

where C does not depend on i, j, and N, and only depends on A_{per}.

The estimate (2.140) above shows that reducing the variance of X_0 (for instance using the technique described in Proposition 2.4.1) reduces the variance on $A_{1,N}^{\star,\text{exact}}$. This also gives a quantitative estimate of the variance reduction.

Note that, in the above Proposition, we have used the first order term w_p^1 in the expansion of the *exact* corrector. If this term is replaced by the first order term $w_p^{1,N}$ of the *approximate* corrector (solution to a truncated problem), then we recover $A_{1,N}^\star$ defined by (2.137).

Corollary 2.4.3 *Under the assumptions of Proposition 2.4.2, we have that*

$$\lim_{N\to\infty} A_{1,N}^{\star,\text{exact}} = A_1^\star \quad a.s.,$$

where $A_1^\star$ is defined by (2.134)*. Assume in addition that*

$$\lim_{N\to\infty} |Q_N|^2\ \mathbb{V}ar\left(\left[A_{1,N}^{\star,\text{exact}}\right]_{ij}\right) = +\infty.$$

Then, $\dfrac{\left[A_{1,N}^{\star,\text{exact}}\right]_{ij} - [A_1^\star]_{ij}}{\sqrt{\mathbb{V}ar\left(\left[A_{1,N}^{\star,\text{exact}}\right]_{ij}\right)}}$ is a random variable that converges in law to a normal Gaussian random variable.

The proof of Proposition 2.4.2 (and of Corollary 2.4.3) is performed in [P3]. We use there estimates on the asymptotic behaviour of Green functions of the operator $-\text{div}\,[A_{\text{per}}\nabla\cdot]$, namely the solution to

$$-\text{div}_x\ (A_{\text{per}}(x)\nabla_x G(x,y)) = \delta_y(x) \quad \text{in } \mathbb{R}^d,$$

which is a topic of independent interest. See [6] for a proof of these estimates.

2.4.3 Numerical tests

We now provide some representative numerical test cases that show that we can actually achieve variance reduction using the technique of antithetic variable. These test cases are extracted from [P2, P4] where more comprehensive numerical experiments are also presented.

Our tests concern three different "input" random fields $A(\cdot,\omega)$ in (2.31) :

– **i.i.d. isotropic fields :**

$$A(x,\omega)=\sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)a_k(\omega)\mathrm{Id}=\sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)F\left(X_k(\omega)\right)\mathrm{Id} \tag{2.141}$$

where $(X_k(\omega))_{k\in\mathbb{Z}^d}$ is a family of independent uniform random variables (note we manipulated this case theoretically in (2.125) above)

– **correlated isotropic fields :**

$$\begin{aligned} A(x,\omega) &= \sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)a_k(\omega)\mathrm{Id} \\ &= \sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)F\left(\left\{X_{k+j}\right\}_{|j|_\infty\leq p}(\omega)\right)\mathrm{Id}, \end{aligned} \tag{2.142}$$

where p is some fixed nonnegative integer, $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ is a family of independent real-valued random variables and F is defined on $\mathbb{R}^{2p+1}$ and real valued (note we manipulated this case theoretically in (2.126) above) ;

– **i.i.d. anisotropic fields :**

$$A(x,\omega)=\sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)A_k(\omega)=\sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)F(X_k(\omega)), \tag{2.143}$$

where $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ is a family of independent $\mathbb{R}^{N_{rv}}$-valued random vectors, the components of which are independent and identically distributed (we choose the uniform law) ; F, defined on $\mathbb{R}^{N_{rv}}$ is valued in the set of symmetric matrices.

We equip these three settings with appropriate assumptions so that the elliptic problem (2.31) is well posed. In these settings, we will investigate variance reduction on a typical diagonal $[A_N^\star]_{11}$, or off-diagonal $[A_N^\star]_{12}$ entry of the approximate homogenized matrix $A_N^\star$, as well as on the eigenvalue of the matrix, and the eigenvalues of the associated differential operator $L_A=-\mathrm{div}\left[A_N^\star(\omega)\nabla\cdot\right]$ (supplied with homogeneous Dirichlet boundary conditions). We do so with different functions F (either monotone or non monotone with respect to their arguments) in (2.141)-(2.142)-(2.143), and different correlation lengths : $p=0$ in (2.141) and (2.143), p variable in (2.142).

We consider four test cases.

Example 1 : We first investigate the case of an i.i.d. isotropic field for the choice $F(x)=\alpha+(\beta-\alpha)x$ in (2.141) (that is, the value of the field A on each periodic cell is uniformly distributed between $\alpha=3$ and $\beta=20$). We focus on the output $[A_N^\star]_{11}$.

Figure 2.10 shows that variance is indeed reduced. A quantitative estimation of the gain is given by the ratio

$$R\left([A_N^\star]_{11}\right) = \frac{\sigma_{100}\left([A_N^\star]_{11}\right)}{2\sigma_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)},$$

the definition of which can be easily extended to the other outputs considered in the sequel. Basically, R corresponds to the ratio of the (square of the) widths of the intervals of confidence. It may also be seen as the gain in computational time at fixed accuracy. Table 2.5 shows the values of this ratio for three different i.i.d isotropic fields : $a_0 \sim \mathcal{B}(1/2)$, $\mathcal{B}(1/3)$ and $\mathcal{U}([\alpha,\beta])$. The latter field corresponds to Figure 2.10.

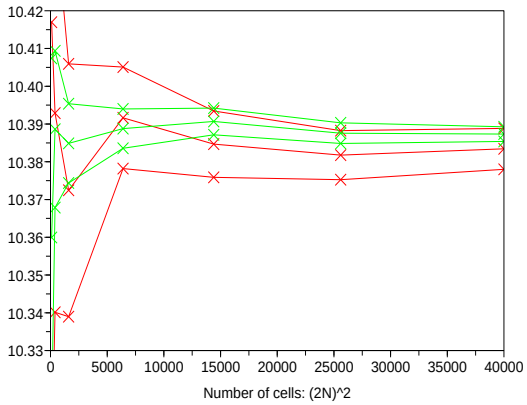


FIG. 2.10 – Example 1 : estimated means (with confidence intervals) for $[A_N^\star]_{11}$ (red) and $\left[\widetilde{A}_N^\star\right]_{11}$ (green) in the case $a_0 \sim \mathcal{U}([\alpha,\beta])$.

N	$a_0 \sim \mathcal{B}(1/2)$	$a_0 \sim \mathcal{B}(1/3)$	$a_0 \sim \mathcal{U}([\alpha,\beta])$
5	5.34	2.06	6.31
10	3.91	1.56	6.46
20	5.41	2.92	10.2
40	3.07	2.31	6.67
60	4.41	2.47	6.16
80	4.49	1.95	5.68
100	4.28	2.99	7.89

TAB. 2.5 – Example 1 : Representative effectivity ratios $R\left([A_N^\star]_{11}\right)$ for several i.i.d isotropic test cases

Example 2 : Second, we consider an i.i.d anisotropic field of the form (2.143) for

the specific choice

$$F(x,y) = P \begin{pmatrix} \alpha + (\beta - \alpha)x & 0 \\ 0 & \delta + (\gamma - \delta)y \end{pmatrix} P^T,$$

with $\alpha = 5$, $\beta = 20$, $\delta = 25$ and $\gamma = 40$. Such a field is monotone, in the sense mentioned above, with respect to each of its variables. The analysis of variance reduction for diagonal terms leads to the same conclusion as in the former case. Since the corresponding homogenized matrix is not diagonal, we concentrate on new outputs : λ_1 the first eigenvalue of $A_N^\star$, $\lambda_1(L_A, \cdot)$ that of the related operator $-\mathrm{div}\,(A_N^\star \nabla \cdot)$ and the off-diagonal term $[A_N^\star]_{12}$. Note the latter output is a non monotone function of the input fields. The corresponding values of ratio R are reported in Tables 2.6, 2.7 and 2.8.

N	$\mu_{100}(\lambda_1)$	$\sqrt{\sigma_{100}}(\lambda_1)$	$\mu_{50}\left(\widetilde{\lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_1\right)$	$R(\lambda_1)$
40	11.9703	0.0572	11.9702	0.0075	29.19
60	11.9650	0.0385	11.9696	0.0043	39.36
80	11.9670	0.0267	11.9696	0.0035	28.63
100	11.9672	0.0233	11.9692	0.0033	24.00

TAB. 2.6 – Example 2 : mean and standard deviation of the first eigenvalue λ_1 of the homogenized matrix.

	$\lambda_1(L_A, \cdot)$		$\widetilde{\lambda}_1(L, \cdot)$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
40	842.9851	1.5576	842.9140	0.2084	27.94
60	842.7855	1.0364	842.8998	0.1284	32.60
80	842.8560	0.6882	842.9044	0.0900	29.21
100	842.8422	0.6977	842.9013	0.0891	30.67

TAB. 2.7 – Example 2 : mean and standard deviation of $\lambda_1(L_A, \cdot)$.

	$[A_N^\star]_{12}$		$\left[\widetilde{A}_N^\star\right]_{12}$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
40	- 10.0897	0.0389	- 10.0877	0.0043	41.32
60	- 10.0902	0.0266	- 10.0880	0.0024	63.93
80	- 10.0899	0.0215	- 10.0882	0.0022	48.82
100	- 10.0892	0.0169	- 10.0886	0.0019	39.10

TAB. 2.8 – Example 2 : mean and standard deviation of the off-diagonal term $[A_N^\star]_{12}$.

The results are self explanatory : the technique is efficient whatever the output considered. It has to be stressed that, however, no theoretical argument guarantees the efficiency of the method for the off-diagonal term.

Example 3 : We consider another i.i.d field of form (2.143) associated to $F : \mathbb{R}^3 \to \mathbb{R}^{2\times 2}$ defined by

$$F(x,y,z) = \begin{pmatrix} \alpha + (\beta-\alpha)x & \iota + (\kappa-\iota)z \\ \iota + (\kappa-\iota)z & \delta + (\gamma-\delta)y \end{pmatrix}$$

where parameters are fixed to $\alpha = 15, \beta = 30, \delta = 20, \gamma = 40, \iota = 5, \kappa = 15$. This situation is very general but is not covered by our theoretical study. The results reported in Tables 2.9 and 2.10 however show that the method is efficient for both monotone and non-monotone outputs respectively.

	$[A_N^\star]_{11}$		$\left[\widetilde{A}_N^\star\right]_{11}$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
20	22.0121	0.1239	22.0116	0.0125	49.44
40	22.0105	0.0571	22.0086	0.0073	30.54
60	22.0050	0.0387	22.0086	0.0046	34.39
80	22.0073	0.0282	22.0079	0.0037	28.55

TAB. 2.9 – Example 3 : mean and standard deviation of the diagonal term $[A_N^\star]_{11}$.

	$[A_N^\star]_{12}$		$\left[\widetilde{A}_N^\star\right]_{12}$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
20	2.5031	0.0369	2.5024	0.0046	31.61
40	2.5017	0.0193	2.5021	0.0025	30.42
60	2.5012	0.0121	2.5028	0.0016	27.25
80	2.5012	0.0092	2.5012	0.0092	37.47

TAB. 2.10 – Example 3 : mean and standard deviation of the off-diagonal term $[A_N^\star]_{12}$.

Example 4 : The approach may also be successfully applied to correlated fields. As an example, we consider the case when F is defined by

$$F\left(\{x_j\}_{|j|\le p}\right) = \frac{1}{(2p+1)^2}\sum_{|j|\le p} \alpha + (\beta-\alpha)x_j,$$

with $\alpha = 3$ and $\beta = 20$. The values of $R\left([A_N^\star]_{11}\right)$ corresponding to correlation lengths $p = 1$ and 2 are given in tables 2.11 and 2.12 respectively.

We conclude that the technique may be applied to a large variety of situations, and that variance is systematically reduced.

	$[A_N^\star]_{11}$		$\left[\widetilde{A}_N^\star\right]_{11}$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
40	11.3791	0.0637	11.3824	0.0054	69.25
60	11.3794	0.0438	11.3818	0.0028	121.07
80	11.3765	0.0310	11.3821	0.0029	58.42
100	11.3794	0.0259	11.3818	0.0018	97.21

TAB. 2.11 – Example 4, $p = 1$: mean and standard deviation of $[A_N^\star]_{11}$.

	$[A_N^\star]_{11}$		$\left[\widetilde{A}_N^\star\right]_{11}$		R
N	μ_{100}	$\sqrt{\sigma_{100}}$	μ_{50}	$\sqrt{\sigma_{50}}$	
40	11.4563	0.05697	11.4574	0.0033	145.75
60	11.4556	0.03641	11.4580	0.0023	121.15
80	11.4528	0.03053	11.4579	0.0017	169.72
100	11.4554	0.02641	11.4579	0.0013	214.58

TAB. 2.12 – Example 4, $p = 2$: mean and standard deviation of $[A_N^\star]_{11}$.

2.5 Trends

As was repeatedly mentioned, the many topics and techniques we have described above can certainly be adapted to a lot of other situations, although we have not done so yet. We have addressed here the very specific case of a *linear*, *elliptic* second order equation *in divergence form*. Other types of equations, of settings, of random variables, could be addressed, provided the theoretical foundations are sound, and the theoretical knowledge is sufficient to seriously consider a numerical approach. Also, when variance is examined, we have only employed *one* possible variance reduction approach, and tested it on *some* outputs of the computation. Alternate options can be considered. More generally, we have essentially presented our developments on the traditional direct numerical approach : first compute the solution to the cell problems, next compute the homogenized matrix, and finally solve the homogenized problem. A number of important contributions in the past two decades have shown that alternate routes can be followed for the numerical approximation, in particular approaches that entangle theory and numerics in order to achieve better efficiency. Multiscale finite element methods are one example among many, outlined in Section 2.3.3. We believe that some of the tools and techniques developed throughout this set of notes can be adapted to these many contexts. It is on our agenda to investigate such issues in the near future.

Acknowledgments

The authors thank Xavier Blanc for a critical reading of the manuscript. The present set of notes is related to a lecture given by CLB at IMS of National University of Singapore in August 2009, within the program on "Mathematical Theory and Numerical Methods for Computational Materials Simulation and Design" - (1 Jul - 31 Aug 2009). Special thanks are due to the IMS for providing such a nice opportunity. The support of ONR and EOARD is gratefully acknowledged.

Première partie

Réduction de variance pour l'homogénéisation stochastique

Chapitre 3

Etude d'un cas simple

Ce chapitre correspond à l'article [P2] dans lequel nous nous intéressons à l'application des techniques classiques de réduction de variance dans le contexte de l'homogénéisation stochastique. Nous avons choisi d'étudier la plus générique et *a priori* la simple d'entre elles : la technique dite des variables antithétiques. Nous nous restreignons au cas de matériaux simples (en particulier isotropes $A^\varepsilon = a^\varepsilon \mathbb{I}_d$). Dans cet article, nous montrons qu'on peut effectivement réduire la variance à l'aide de cette technique dans le contexte de l'homogénéisation stochastique, et donnons une preuve du fonctionnement de la méthode dans un cadre monodimensionnel. Ce travail a ensuite été complété par deux autres articles [P3, P4] qui font l'objet des chapitres 4 et 5, où des tests numériques plus complets sont décrits et des éléments de preuve dans des cas plus généraux sont donnés.

Variance reduction in stochastic homogenization : proof of concept, using antithetic variables

Ronan Costaouec[a], Claude Le Bris[a] and Frédéric Legoll[a]

[a] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

{*costaour,lebris*}*@cermics.enpc.fr,*
legoll@lami.enpc.fr

We show that we can reduce the variance in a simple problem of stochastic homogenization using the classical technique of antithetic variables. The setting, and the presentation, are deliberately kept elementary. We point out the main issues, show some illustrative results, and demonstrate, both theoretically and numerically, the efficiency of the approach on simple cases.

3.1 Introduction

Several settings in homogenization require the solution of corrector problems posed on the entire space $\mathbb{R}^d$. In practice, truncations of these problems over bounded domains are considered and the homogenized coefficients are obtained in the limit of large domains. The question arises as to how such computations can be accelerated. In the deterministic case, acceleration techniques reminiscent of signal filtering have been introduced in [18]. The work has since then been significantly improved by A. Gloria in [46]. In [18], it was shown that acceleration techniques efficient for deterministic problems do not necessarily perform well in the stochastic framework. In the latter case, the main difficulty is related to the intrinsic noise present in the simulation. The challenge is consequently not that much to improve the rate of convergence, which is intrinsically that of the central limit theorem, but rather to reduce the variance, thereby improving the prefactor of the convergence given by the central limit theorem. Although very well investigated in other application fields such as financial mathematics, variance reduction techniques seem to have not been applied to the context of stochastic homogenization. The purpose of the present contribution is to present a first attempt to reduce the variance in stochastic homogenization. For this purpose, we consider a simple situation, and a simple variance reduction technique. The probability theoretic arguments we will make use of are elementary. The equation under consideration is a simple elliptic equation in

divergence form, with a scalar coefficient. The coefficient is assumed to consist of independent, identically distributed random variables set on a simple mesh (see (3.2) below). The technique used for variance reduction is that of antithetic variables. Our setting is academic in nature, somewhat far from physically relevant cases, and elementary. Many more difficult situations could be addressed : other types of stationary ergodic coefficients, matrix rather than scalar coefficients, other types of equations, other techniques for variance reduction, ... The present contribution is a proof of concept : variance reduction *can* be achieved in stochastic homogenization. Future works [P3, P4] will provide more details on the numerics and the theory, and also address some of the many possible extensions mentioned above. We also mention the related work [49] on stochastic homogenization of *discrete* elliptic equations.

3.2 Stochastic homogenization theory

Although we wish to keep the mathematical formalism as limited as possible in our exposition, we need to introduce the basic setting of stochastic homogenization (see [63] for a similar presentation and related issues). Throughout this article, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and we denote by $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ the expectation value of any random variable $X \in L^1(\Omega, d\mathbb{P})$. We next fix $d \in \mathbb{N}^*$ (the ambient physical dimension), and assume that the group $(\mathbb{Z}^d, +)$ acts on Ω. We denote by $(\tau_k)_{k\in\mathbb{Z}^d}$ this action, and assume that it preserves the measure $\mathbb{P}$, that is, for all $k \in \mathbb{Z}^d$ and all $A \in \mathcal{F}$, $\mathbb{P}(\tau_k A) = \mathbb{P}(A)$. We assume that the action τ is *ergodic*, that is, if $A \in \mathcal{F}$ is such that $\tau_k A = A$ for any $k \in \mathbb{Z}^d$, then $\mathbb{P}(A) = 0$ or 1. In addition, we define the following notion of stationarity (see [20]) : any $F \in L^1_{\rm loc}\left(\mathbb{R}^d, L^1(\Omega)\right)$ is said to be *stationary* if, for all $k \in \mathbb{Z}^d$,

$$F(x + k, \omega) = F(x, \tau_k \omega), \tag{3.1}$$

almost everywhere in x and almost surely. In this setting, the ergodic theorem [61,89] can be stated as follows : *Let $F \in L^\infty\left(\mathbb{R}^d, L^1(\Omega)\right)$ be a stationary random variable in the above sense. For $k = (k_1, k_2, \dots k_d) \in \mathbb{Z}^d$, we set $|k|_\infty = \sup_{1\le i\le d} |k_i|$. Then*

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \le N} F(x, \tau_k \omega) \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left(F(x, \cdot)\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

This implies that (denoting by Q the unit cube in $\mathbb{R}^d$)

$$F\left(\frac{x}{\varepsilon}, \omega\right) \overset{*}{\underset{\varepsilon\to 0}{\rightharpoonup}} \mathbb{E}\left(\int_Q F(x, \cdot) dx\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

Besides technicalities, the purpose of the above setting is simply to formalize that, even though realizations may vary, the function F at point $x \in \mathbb{R}^d$ and the function F at point $x + k$, $k \in \mathbb{Z}^d$, share the same law. In the homogenization context

we now turn to, this means that the local, microscopic environment (encoded in the coefficient a, see (3.3) below) is everywhere the same *on average*. From this, homogenized, macroscopic properties will follow.

We now fix an open, regular, bounded subset $\mathcal{D}$ of $\mathbb{R}^d$, a L^2 function f on $\mathcal{D}$, and a random function a assumed stationary in the sense (3.1) defined above. We also assume a is bounded, positive and almost surely bounded away from zero. For simplicity, we take a random piecewise constant function of the form :

$$a(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega), \tag{3.2}$$

where Q is the unit cube of $\mathbb{R}^d$ and $(a_k(\omega))_{k\in\mathbb{Z}^d}$ denotes a family of i.i.d. random variables. The standard results of stochastic homogenization [16, 58] apply to the boundary value problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(a\left(\dfrac{x}{\varepsilon},\omega\right)\nabla u^\varepsilon\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u^\varepsilon &=& 0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right. \tag{3.3}$$

These results state that, in the limit $\varepsilon \longrightarrow 0$, the homogenized problem obtained from (3.3) reads :

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A^\star \nabla u^\star\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u^\star &=& 0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right. \tag{3.4}$$

The homogenized matrix $A^\star$ is defined as

$$[A^\star]_{ij} = \mathbb{E}\left(\int_Q \left(e_i + \nabla w_{e_i}(y,\cdot)\right)^T a\left(y,\cdot\right)\left(e_j + \nabla w_{e_j}(y,\cdot)\right) dy \right), \tag{3.5}$$

where, for any $p \in \mathbb{R}^d$, w_p is the solution (unique up to the addition of a (random) constant) in $\left\{ w \in L^2_{\mathrm{loc}}(\mathbb{R}^d, L^2(\Omega)), \quad \nabla w \in L^2_{\mathrm{unif}}(\mathbb{R}^d, L^2(\Omega)) \right\}$ to

$$\left\{ \begin{array}{l} -\mathrm{div}\left[a\left(y,\omega\right)\left(p + \nabla w_p(y,\omega)\right)\right] = 0 \quad \text{a.s. on } \mathbb{R}^d, \\ \nabla w_p \quad \text{is stationary in the sense of (3.1)}, \\ \mathbb{E}\left(\displaystyle\int_Q \nabla w_p(y,\cdot)\, dy \right) = 0, \end{array} \right. \tag{3.6}$$

where we have used the notation L^2_{unif} for the *uniform* L^2 space, that is the space of functions for which, say, the L^2 norm on a ball of unit size is bounded above independently from the center of the ball.

The solution u^ε to (3.3) is known to converge to the solution $u^\star$ to (3.4) in various appropriate senses. The tensor and function $A^\star$ and $u^\star$ are *deterministic quantities*, although they originate from a series of random problems. This is a consequence of the *ergodic* setting described above, which allows random microscopic

quantities to average out in deterministic macroscopic quantities. Note however that the computation of $A^\star$ requires the computation of the so-called corrector functions w_p, which are random.

The above result generalizes that of the classical periodic setting (see *e.g.* [16,27]) where, instead of being stationary ergodic, the function a in (3.3) is periodic. Then, although the homogenized problem can be expressed similarly, the crucial difference is that (at least in this simple linear case) the corrector problem can, in the periodic case, be reduced to the equation $-\text{div}\left[a(y)\left(p+\nabla w_p(y)\right)\right]=0$ set on the periodic cell $Q=[0,1]^d$, and not on the entire space $\mathbb{R}^d$ as in (3.6). Correspondingly, the terms of the homogenized tensor in (3.5) are simple deterministic integrals on Q. In the random case, the corrector problem (3.6) is intrinsically set on the entire space and the numerical approximation of its solution w_p is the main computational challenge. Problem (3.6) is in practice truncated on a bounded domain $Q_N=[-N,N]^d$ and usually supplied with periodic boundary conditions :

$$\left\{\begin{array}{l} -\text{div}\left(a(\cdot,\omega)\left(p+\nabla w_p^N(\cdot,\omega)\right)\right)=0 \quad \text{on} \quad Q_N, \\ w_p^N \text{ is } Q_N\text{-periodic.} \end{array}\right. \tag{3.7}$$

Correspondingly, we set :

$$\left[A_N^\star\right]_{ij}(\omega)=\frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^N(y,\omega)\right)^T a(y,\omega)\left(e_j+\nabla w_{e_j}^N(y,\omega)\right)dy. \tag{3.8}$$

In the limit of large domains Q_N, the homogenized tensor (3.5) is recovered. In addition, the rate of convergence with which the truncated values approach the exact homogenized value $A^\star$ can be assessed theoretically. We refer to [23, 103] for the proof of all the above statements. As will be seen below, the variance of the random variables involved plays a role in the approximation procedure. Reducing this variance is the problem we now consider.

3.3 Variance reduction

3.3.1 Classical Monte Carlo method

As mentioned above, the large size (large N) limit of the coefficient (3.8) obtained using the solution of the truncated corrector problem (3.7) gives the value of the homogenized coefficient (3.5). Formally, this is a convergence of the type $A_N^\star(\omega)\longrightarrow A^\star$ as $N\longrightarrow+\infty$ almost surely. The practical approach to this problem is the Monte Carlo approach. We now briefly investigate the role of the variance in the problem.

To start with, we consider the one-dimensional setting. Although this setting is very particular (and sometimes misleading because oversimplified), it also allows to already understand the basic features of the problem and the bottom line of the approach, with the economy of many unnecessary technicalities.

In the one-dimensional setting, the definition (3.2) reads

$$a(x,\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1[}(x) a_k(\omega) \tag{3.9}$$

with $(a_k(\omega))_{k\in\mathbb{Z}}$ a family of i.i.d. random variables. It is easily seen that the truncated corrector problem (3.7) can be explicitly solved and leads to the value

$$a^\star_N(\omega) = \left(\frac{1}{2N} \sum_{k=-N}^{N-1} \frac{1}{a_k(\omega)} \right)^{-1} \tag{3.10}$$

of the approximation for the homogenized tensor (here, a scalar coefficient of course). In the limit of large N, it almost surely converges to the value of the *exact* homogenized coefficient

$$a^\star = \mathbb{E}\left(\frac{1}{a_0} \right)^{-1}. \tag{3.11}$$

This exact value is readily obtained explicitly solving (3.5)-(3.6). The simplest possible argument consists now in considering $(a^\star_N(\omega))^{-1} = \dfrac{1}{2N} \displaystyle\sum_{k=-N}^{N-1} \frac{1}{a_k(\omega)}$ and remark that the rate of convergence of this quantity to $(a^\star)^{-1}$ is evidently given by the central limit theorem, where the variance of the random variable $(a_k(\omega))^{-1}$ plays a crucial role. Although correct, this argument exploits too much the very peculiar nature of the one-dimensional setting (we have taken the inverse of the coefficient and recasted it as a sum, a fact that is not possible otherwise than in one dimension). An argument with slightly more generality consists in considering $a^\star_N(\omega)$ itself – and not its inverse–, and, using elementary calculus, showing that it also converges to $a^\star$ with a rate of convergence where the variance of $a_0(\omega)$ again plays the crucial role. Indeed, one may for instance remark that $\mathbb{E}\left(\left| \left(\frac{1}{2N} \sum_{k=-N}^{N-1} \frac{1}{a_k} \right)^{-1} - \mathbb{E}\left(\frac{1}{a_0} \right)^{-1} \right|^2 \right)$ can be bounded from above (using a simple almost sure upper bound of $a_k(\omega)$) by $\mathbb{E}\left(\left| \left(\frac{1}{2N} \sum_{k=-N}^{N-1} \frac{1}{a_k} \right) - \mathbb{E}\left(\frac{1}{a_0} \right) \right|^2 \right)$ up to an irrelevant multiplicative constant and that the latter quantity, once easily computed, is of the form $\dfrac{1}{2N}\, \mathbb{V}\mathrm{ar}\left(\dfrac{1}{a_0} \right)$. Again, the variance of the random coefficient plays a role.

In dimensions higher than one, the situation is considerably more intricate and the rate of convergence with which the coefficient arising from the truncated computation converges to its limit is not so simple to evaluate. This is the purpose, under appropriate conditions (called *mixing conditions* and which are indeed met in our present setting), of the work [23].

The numerical practice is as follows. A set of M independent realizations of the random coefficient a are considered. The corresponding truncated problems (3.7) are

solved, and an empirical mean of the truncated coefficients (3.8) is inferred. This empirical mean only agrees with the theoretical mean value of the truncated coefficient within a margin of error which is given by the central limit theorem (in terms of M). The variance of the coefficients therefore again plays a role, as a prefactor. For a sufficiently large truncation size N, this truncated value is admitted to be the exact value of the coefficient. The error made is controlled by the estimations of the theoretical work [23]. Of course, the overall computation described above is expensive, because each realization requires a new solution to the d-dimensional boundary value problem (3.7) of presumably large a size since N is taken large. There is therefore a huge interest in reducing the cost of the computation, or, otherwise stated, in reaching a better accuracy at a given computational cost. Since the variance of the truncated homogenized tensor is an important ingredient, reducing the variance becomes a challenging and sensitive issue.

More explicitly, let $(a^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq M}$ denote M independent and identically distributed underlying random fields. We define a family $(A_N^{\star,\mathbf{m}})_{1\leq \mathbf{m}\leq M}$ of i.i.d. homogenized matrices by, for any $1 \leq i,j \leq d$,

$$[A_N^{\star,\mathbf{m}}]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} \left(e_i + \nabla w_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T a^{\mathbf{m}}(\cdot,\omega)\ \left(e_j + \nabla w_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right),$$

where $w_{e_j}^{N,\mathbf{m}}$ is the solution of the corrector problem associated to $a^{\mathbf{m}}$. Then we define for each component of $A_N^\star$ the empirical mean and variance

$$\begin{aligned} \mu_M\left([A_N^\star]_{ij}\right) &= \frac{1}{M}\sum_{\mathbf{m}=1}^{M} [A_N^{\star,\mathbf{m}}]_{ij}, \\ \sigma_M\left([A_N^\star]_{ij}\right) &= \frac{1}{M-1}\sum_{\mathbf{m}=1}^{M} \left([A_N^{\star,\mathbf{m}}]_{ij} - \mu_M\left([A_N^\star]_{ij}\right)\right)^2. \end{aligned} \tag{3.12}$$

Since the matrices $A_N^{\star,\mathbf{m}}$ are i.i.d., the strong law of large numbers applies :

$$\mu_M\left([A_N^\star]_{ij}\right)(\omega) \underset{M\to+\infty}{\longrightarrow} \mathbb{E}\left([A_N^\star]_{ij}\right) \text{ almost surely.}$$

The central limit theorem then yields

$$\sqrt{M}\left(\mu_M\left([A_N^\star]_{ij}\right) - \mathbb{E}\left([A_N^\star]_{ij}\right)\right) \underset{M\to+\infty}{\xrightarrow{\mathcal{L}}} \sqrt{\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)}\,\mathcal{N}(0,1), \tag{3.13}$$

where the convergence holds in law, and $\mathcal{N}(0,1)$ denotes the standard gaussian law. Introducing its 95 percent quantile, it is standard to consider that the exact mean $\mathbb{E}\left([A_N^\star]_{ij}\right)$ is equal to $\mu_M\left([A_N^\star]_{ij}\right)$ within a margin of error $1.96\dfrac{\sqrt{\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}$. The exact variance $\mathbb{V}\mathrm{ar}\left([A_N^\star]_{ij}\right)$ being unknown in practice, it is customary to replace

it by the empirical variance given in (3.12) above. It is therefore considered that the expectation $\mathbb{E}\left([A_N^\star]_{ij}\right)$ lies in the interval

$$\left[\mu_M\left([A_N^\star]_{ij}\right) - 1.96\frac{\sqrt{\sigma_M\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}\,,\ \mu_M\left([A_N^\star]_{ij}\right) + 1.96\frac{\sqrt{\sigma_M\left([A_N^\star]_{ij}\right)}}{\sqrt{M}}\right]. \quad (3.14)$$

The value $\mu_M\left([A_N^\star]_{ij}\right)$ is thus, for both M and N sufficiently large, adopted as the approximation of the exact value $[A^\star]_{ij}$.

Of course, a tensorial argument could be applied here, not considering separately each entry of the matrix but treating the matrix as a whole. The approach developed above, component by component, is sufficient for the simple cases considered in the present work.

3.3.2 Antithetic variable for stochastic homogenization

We know from the previous section that constructing empirical means approximating $\mathbb{E}\left(A_N^\star\right)$ with a smaller variance at the same computational cost is of high interest. We now describe a possible approach to achieve this goal.

In generality, fix $M = 2\mathcal{M}$. Suppose that we give ourselves $\mathcal{M}$ i.i.d. copies $(a^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq \mathcal{M}}$ of $a(x,\omega)$. Construct next $\mathcal{M}$ i.i.d. *antithetic* random fields

$$b^{\mathbf{m}}(x,\omega) = T\left(a^{\mathbf{m}}(x,\omega)\right), \quad 1 \leq \mathbf{m} \leq \mathcal{M},$$

from the $(a^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq \mathcal{M}}$. The map T transforms the random field $a^{\mathbf{m}}$ into another, so-called *antithetic*, field $b^{\mathbf{m}}$. Explicit examples of such T are given in the sequel (see (3.20) and Section 3.4 below). The transformation is performed in such a way that, for each $\mathbf{m}$, $b^{\mathbf{m}}$ should have the same law as $a^{\mathbf{m}}$, namely the law of the coefficient a. Somewhat vaguely stated, if the coefficient a was obtained in a coin tossing game (using a fair coin), then $b^{\mathbf{m}}$ would be head each time $a^{\mathbf{m}}$ is tail and vice versa. We refer the reader to Figure 3.1 below for explicit illustrative examples of such a construction. Then, for each $1 \leq \mathbf{m} \leq \mathcal{M}$, we solve two corrector problems. One is associated to the original $a^{\mathbf{m}}$, the other one is associated to the antithetic field $b^{\mathbf{m}}$. Using its solution $v_p^{N,\mathbf{m}}$, we define the *antithetic homogenized matrix* $B_N^{\star,\mathbf{m}}$, whose elements read, for $1 \leq i,j \leq d$,

$$[B_N^{\star,\mathbf{m}}]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}\left(e_i + \nabla v_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T b^{\mathbf{m}}(\cdot,\omega)\ \left(e_j + \nabla v_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right).$$

And finally we set, for any $1 \leq \mathbf{m} \leq \mathcal{M}$,

$$\widetilde{A}_N^{\star,\mathbf{m}}(\omega) := \frac{1}{2}\left(A_N^{\star,\mathbf{m}}(\omega) + B_N^{\star,\mathbf{m}}(\omega)\right). \quad (3.15)$$

Since $a^{\mathbf{m}}$ and $b^{\mathbf{m}}$ are identically distributed, so are $A_N^{\star,\mathbf{m}}$ and $B_N^{\star,\mathbf{m}}$. Thus, $\widetilde{A}_N^{\star,\mathbf{m}}$ is unbiased (that is, $\mathbb{E}\left(\widetilde{A}_N^{\star,\mathbf{m}}\right) = \mathbb{E}\left(A_N^{\star,\mathbf{m}}\right)$). In addition, it satisfies :

$$\widetilde{A}_N^{\star,\mathbf{m}} \underset{N\to+\infty}{\longrightarrow} A^\star \text{ almost surely,}$$

because b is ergodic.

Let us define new estimators

$$\begin{aligned} \mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) &= \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\left[\widetilde{A}_N^{\star,\mathbf{m}}\right]_{ij}, \\ \sigma_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) &= \frac{1}{\mathcal{M}-1}\sum_{\mathbf{m}=1}^{\mathcal{M}}\left(\left[\widetilde{A}_N^{\star,\mathbf{m}}\right]_{ij} - \mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)\right)^2, \end{aligned} \tag{3.16}$$

which require $2\mathcal{M}$ resolutions of corrector problems, i.e. as many as the classical estimators (3.12), since we choose $M = 2\mathcal{M}$. In addition, note that we have built a new random variable whose variance is

$$\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) = \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\left[A_N^\star\right]_{ij}\right) + \frac{1}{2}\mathbb{C}\mathrm{ov}\left(\left[A_N^\star\right]_{ij}, \left[B_N^\star\right]_{ij}\right). \tag{3.17}$$

Applying the central limit theorem to $\widetilde{A}_N^\star$, we obtain

$$\sqrt{\mathcal{M}}\left(\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right) - \mathbb{E}\left(\left[A_N^\star\right]_{ij}\right)\right) \underset{\mathcal{M}\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \sqrt{\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)}\,\mathcal{N}(0,1). \tag{3.18}$$

Similarly to (3.14), we deduce a confidence interval from this convergence. The exact mean $\mathbb{E}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)$ is equal to $\mu_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)$ within a margin of error $1.96\dfrac{\sqrt{\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}_N^\star\right]_{ij}\right)}}{\sqrt{\mathcal{M}}}$. It results from (3.17) that, if

$$\mathbb{C}\mathrm{ov}\left(\left[A_N^\star\right]_{ij}, \left[B_N^\star\right]_{ij}\right) \leq 0, \tag{3.19}$$

then the width of this interval has been diminished by the new approach, and, correspondingly, the quality of approximation at given computational cost has increased.

To understand slightly more in details at the theoretical level why the approach is likely to perform well, we again consider the one-dimensional setting (3.9) for which we recall the explicit expressions (3.10) and (3.11) for the truncated and the exact homogenized coefficients, respectively.

Suppose as a first illustration that a_0 is a Bernoulli distributed random variable $a_0 \sim \mathcal{B}(1/2)$:

$$\mathbb{P}(a_0 = \alpha) = 1/2 \quad \text{and} \quad \mathbb{P}(a_0 = \beta) = 1/2,$$

for some $0 < \alpha < \beta$. Defining the antithetic variable

$$b_k(\omega) = \alpha + \beta - a_k(\omega)$$

and next the antithetic field

$$b(x,\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1[}(x)\, b_k(\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1[}(x)\, (\alpha+\beta-a_k(\omega))\,, \tag{3.20}$$

it is immediately seen that

$$\frac{1}{2}\left(\frac{1}{a_N^\star(\omega)} + \frac{1}{b_N^\star(\omega)}\right) = \mathbb{E}\left(\frac{1}{a_0}\right).$$

The variance of the inverse of the truncated coefficient has vanished. This example might seem oversimplified because we are indeed making use of two peculiarities of the problem : the set $\{\alpha,\beta\}$ of values taken by the coefficient a has cardinality two, and the explicit expression (3.10) allows us to explicitly manipulate the inverse of the homogenized coefficient. The situation, although oversimplified, is yet a first good indicator of the interest of the approach. As in the previous section, we can be slightly more general, by considering for instance that the random coefficient a is now uniformly distributed over a given interval, say $a_0 \sim \mathcal{U}([\alpha,\beta])$. Then,

$$\frac{1}{2}\left(\frac{1}{a_N^\star(\omega)} + \frac{1}{b_N^\star(\omega)}\right) = \frac{1}{2N}\sum_{k=-N}^{N-1} \frac{1}{2}\left(\frac{1}{a_k(\omega)} + \frac{1}{b_k(\omega)}\right). \tag{3.21}$$

It is a simple matter to show that, because the function $x \mapsto 1/x$ is decreasing, we have

$$\mathbb{C}\mathrm{ov}\left(\frac{1}{a_0}, \frac{1}{b_0}\right) \leq 0. \tag{3.22}$$

Consider indeed a decreasing function f, and X and Y two independent random variables, identically distributed according to $\mathcal{U}([\alpha,\beta])$. Since $x \mapsto f(\alpha+\beta-x)$ is increasing, we observe that

$$(f(X) - f(Y))\ (f(\alpha+\beta-X) - f(\alpha+\beta-Y)) \leq 0,$$

hence

$$\mathbb{E}[f(X)\ f(\alpha+\beta-X)] \leq \mathbb{E}[f(X)]\ \mathbb{E}[f(\alpha+\beta-X)],$$

which reads $\mathbb{C}\mathrm{ov}[f(X), f(\alpha+\beta-X)] \leq 0$. Choosing $f(x) = 1/x$ yields (3.22).

Since

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\frac{1}{a_N^\star} + \frac{1}{b_N^\star}\right)\right) = \frac{1}{4N}\mathbb{V}\mathrm{ar}\left(\frac{1}{a_0}\right) + \frac{1}{4N}\mathbb{C}\mathrm{ov}\left(\frac{1}{a_0}, \frac{1}{b_0}\right),$$

we conclude that

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\frac{1}{a_N^\star} + \frac{1}{b_N^\star}\right)\right) \leq \mathbb{V}\mathrm{ar}\left(\frac{1}{a_{2N}^\star}\right).$$

Therefore, $\mathbb{E}(1/a_0)$ can be approximated either by (3.21) or by $1/a^\star_{2N}$, with an equal cost (i.e. an equal number of random variables in both sums), but the former has a smaller variance than the latter. It is hence of better quality.

As mentioned above, the practice in dimensions higher than one is to generate a set of identically distributed coefficients for each truncated corrector problem, and to use (3.15). The appropriate analogous one-dimensional approach is to consider $\mathcal{M} = \dfrac{M}{2}$ independent copies of $a(x,\omega)$ and set

$$\begin{aligned}\widetilde{a}^{\star,\mathbf{m}}_N(\omega) &:= \frac{1}{2}\left(a^{\star,\mathbf{m}}_N(\omega) + b^{\star,\mathbf{m}}_N(\omega)\right)\\ &= \frac{1}{2}\left(\frac{1}{2N}\sum_{k=-N}^{N-1}\frac{1}{a^{\mathbf{m}}_k(\omega)}\right)^{-1} + \frac{1}{2}\left(\frac{1}{2N}\sum_{k=-N}^{N-1}\frac{1}{b^{\mathbf{m}}_k(\omega)}\right)^{-1}\end{aligned}$$

with empirical mean

$$\mu_{\mathcal{M}}\left(\widetilde{a}^\star_N\right)(\omega) = \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\widetilde{a}^{\star,\mathbf{m}}_N(\omega).$$

We approach more generality since

$$\mu_{\mathcal{M}}\left(\widetilde{a}^\star_N\right)(\omega) \underset{\mathcal{M}\to+\infty}{\longrightarrow} \mathbb{E}\left(\widetilde{a}^\star_N\right) = \mathbb{E}\left(a^\star_N\right) \quad \text{almost surely,}$$

but $\mathbb{E}\left(a^\star_N\right) \neq a^\star$. It can again be remarked that $a^\star_N(\omega)$ is an increasing function of the uniform variables $(a_k(\omega))_{k\in\mathbb{Z}}$. From this observation, it is possible to show that $\mathbb{C}\mathrm{ov}\left(a^\star_N, b^\star_N\right) \leq 0$, and to conclude that the variance of $\mu_{\mathcal{M}}\left(\widetilde{a}^\star_N\right)$ is smaller than that of $\mu_{2\mathcal{M}}\left(a^\star_N\right)$. For this proof on a model by analogy, as well as for proofs that variance reduction is indeed achieved for some *actual* settings in dimensions higher than one (such as for instance those from [P1, 7, 19]), we refer to [P3]. The above simplified arguments were only meant to have pedagogic value.

3.4 Numerical experiments

The previous section provides some elementary ingredients for a theoretical analysis of the efficiency of the approach. The one-dimensional setting is however too particular. More convincing theoretical arguments have to be developed. As announced, this will be the purpose of future publications. Meanwhile, it is possible to test the approach on actual two-dimensional cases, and it is the purpose of this section to report on such tests. As above, we only consider random coefficients that are piecewise constant and of the form (3.2). The test cases we choose correspond to three different laws for a_0 :

– case (i) : a Bernoulli law of parameter 1/2, namely $a_0 \sim \mathcal{B}(1/2)$, $\mathbb{P}\left(a_0 = \alpha\right) = 1/2$ and $\mathbb{P}\left(a_0 = \beta\right) = 1/2$;

– case (ii) : a Bernoulli law of parameter 1/3, namely $a_0 \sim \mathcal{B}(1/3)$, $\mathbb{P}(a_0 = \alpha) = 1/3$ and $\mathbb{P}(a_0 = \beta) = 2/3$;
– case (iii) : a uniform law, namely $a_0 \sim \mathcal{U}([\alpha, \beta])$.

We take the specific values $\alpha = 3$ and $\beta = 20$, just to fix the ideas. Similar qualitative conclusions would be reached with other generic values. Figure 3.1 shows a realization of a and its antithetic field b in cases (i) and (iii).

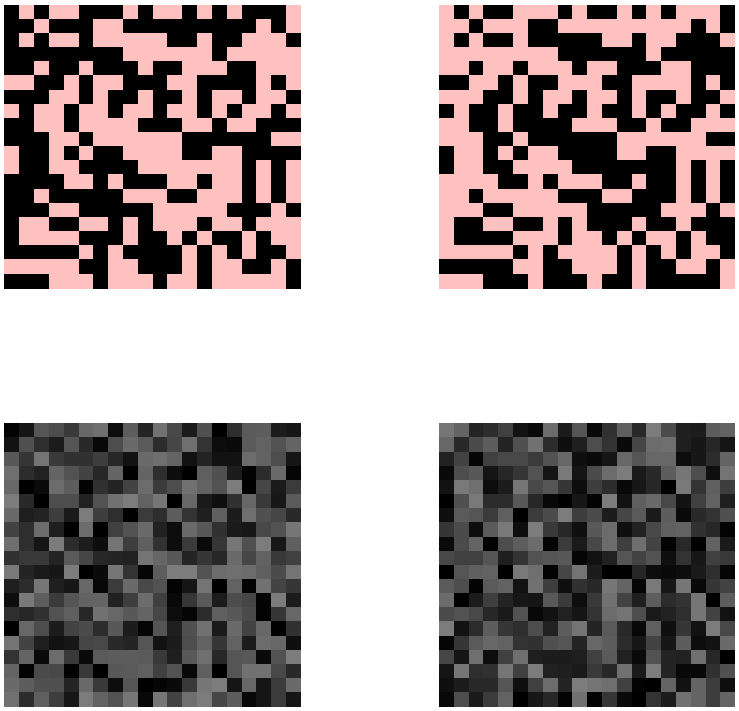


FIG. 3.1 – Realization of $a(x, \omega)$ given by (3.2) (left) and the associated antithetic field $b(x, \omega)$ (right). Top figures : $a_0 \sim \mathcal{B}(1/2)$; bottom figures : $a_0 \sim \mathcal{U}([\alpha, \beta])$.

Our numerical tests have been performed using the finite elements software FreeFem++ developed by F. Hecht (Paris VI, see [54]). The discretization of the corrector problem is performed using $\mathbb{P}1$ Lagrange finite elements, and a regular Q-periodic mesh of Q_N. The discretization meshsize is fixed and has value $h = 0.2$.

It is worth mentioning how we practically proceed to generate an antithetic variable. This may indeed be delicate. We have taken random coefficients that can all originally be expressed in terms of a uniformly distributed random variable (with a view, notably, to be consistent with the way a random variable is practically generated on a computer). We then build the antithetic variable precisely using the 'mother' uniform random variable. The technique is best explained on case (ii).

Write the variable $a_0 \sim \mathcal{B}(1/3)$ as $a_0 \sim \alpha + (\beta - \alpha)\mathbf{1}_{\{1/3 \leq U_0 \leq 1\}}$ where $U_0 \sim \mathcal{U}([0,1])$ denotes a random variable that has uniform law on the interval $[0,1]$. The antithetic variable is then taken as $b_0 \sim \alpha + (\beta - \alpha)\mathbf{1}_{\{0 \leq U_0 \leq 2/3\}}$ and the correspondence is made realization by realization using the actual realization of U_0.

In cases (i) and (ii), in dimension 2, the *exact* homogenized tensor is known to be *isotropic*, $A^\star = a^\star \mathbb{I}_2$ (see [58, Chap. 7, pp. 234-237] for a proof). Of course, for N finite, $A_N^\star$ is a generic matrix, but our numerical experiments consistently show that, for N sufficiently large, the off-diagonal terms are very small on average compared to the diagonal terms, in the three cases we have considered. Table 3.1 summarizes, in case (iii), the estimated means and standard deviations of the components of $A_N^\star$ for different values of N. It confirms that the main sources of variance are the diagonal terms. The same conclusion holds in cases (i) and (ii).

N	$[A_N^\star]_{11}$	$[A_N^\star]_{22}$	$[A_N^\star]_{12}$
5	10.42 (0.608)	10.39 (0.620)	0.00391 (0.074)
10	10.39 (0.269)	10.39 (0.273)	0.00369 (0.033)
20	10.37 (0.171)	10.37 (0.162)	0.00089 (0.017)
40	10.39 (0.069)	10.39 (0.070)	-0.00219 (0.0095)
60	10.38 (0.045)	10.38 (0.045)	0.00059 (0.0069)
80	10.38 (0.033)	10.38 (0.034)	0.00013 (0.0047)
100	10.38 (0.028)	10.38 (0.028)	0.00010 (0.0033)

TAB. 3.1 – For each entry of $A_N^\star$, empirical mean $\mu_{100}\left([A_N^\star]_{ij}\right)$ (and empirical standard deviation $\sigma_{100}^{1/2}\left([A_N^\star]_{ij}\right)$, in brackets), in the case (iii).

In our three test cases, we have compared for different values of N the estimated variance of $\left[\widetilde{A}_N^\star\right]_{11}$ with that of $[A_N^\star]_{11}$. In order to quantitatively assess the efficiency of the antithetic variables method, we introduce the effectivity ratio

$$R\left([A_N^\star]_{11}\right) = \frac{\sigma_{100}\left([A_N^\star]_{11}\right)}{2\sigma_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)}.$$

The factor 2 at the denominator accounts for the number of realizations associated to the classical and antithetic Monte Carlo methods, given that we wish to work at fixed computational cost. Indeed, after solving $M = 2\mathcal{M}$ corrector problems (3.7), one can either build a confidence interval of size $1.96\sqrt{\sigma_M\left([A_N^\star]_{11}\right)/M}$ following (3.13) and (3.14), or a confidence interval of size $1.96\sqrt{\sigma_{\mathcal{M}}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)/\mathcal{M}}$ following (3.18).

Our next table, Table 3.2, contains the values of this representative ratio for each test case. We have also plotted on Figure 3.2 the curves of estimated means (3.12) and (3.16), with their confidence intervals, for the three cases under study here.

If we admit that the theory developed in the previous section applies to the two-dimensional case, another manner to check variance reduction is to compute the empirical covariance between $[A_N^\star]_{11}$ and $[B_N^\star]_{11}$ (recall (3.19)). This is the reason why we have also plotted on Figure 3.2 the normalized empirical value of this covariance,

$$\frac{\mathbb{C}\mathrm{ov}\left([A_N^\star]_{11}\,,[B_N^\star]_{11}\right)}{\sqrt{\mathbb{V}\mathrm{ar}\left([A_N^\star]_{11}\right)\ \mathbb{V}\mathrm{ar}\left([B_N^\star]_{11}\right)}}, \tag{3.23}$$

for test case (iii) (similar results have been obtained for the two other test cases).

N	$a_0 \sim \mathcal{B}(1/2)$	$a_0 \sim \mathcal{B}(1/3)$	$a_0 \sim \mathcal{U}([\alpha,\beta])$
5	5.34	2.06	6.31
10	3.91	1.56	6.46
20	5.41	2.92	10.2
40	3.07	2.31	6.67
60	4.41	2.47	6.16
80	4.49	1.95	5.68
100	4.28	2.99	7.89

TAB. 3.2 – Representative effectivity ratios $R\left([A_N^\star]_{11}\right)$ for test cases (i), (ii) and (iii). The number shown gives the gain in computational time or, equivalently, at given computational cost, the square of the gain in the width of the confidence interval.

The results are self-explanatory : the variance is reduced. The reduction is not spectacular, but it is definite, and, equally importantly, systematic. Considering that the approach induces no additional computational cost at all, this is very good. Other more adapted, but also more delicate to design and implement, variance reduction approaches will be tested in the future [P4], and one may expect even more significant reductions.

3.5 Variance reduction for the solution $u^\star$

We conclude this article examining the problem of variance reduction from a slightly different perspective. We have so far investigated the question of variance reduction for the homogenized tensor $A^\star$. This is the question typically relevant in Mechanics, where for instance determining the homogenized tensor is an important issue because it allows to define, say, the Young modulus or the Poisson ratio of the homogenized material. In some contexts however, the focus is more on the solution of the homogenized problem, rather than on the coefficients of the homogenized equation. For a given right-hand side f in (3.3) (or for a *set* of such right-hand sides), one wishes to know the behaviour of the solution u^ε for small ε. Now, reducing the variance on the solution u is not exactly the same question as reducing the variance

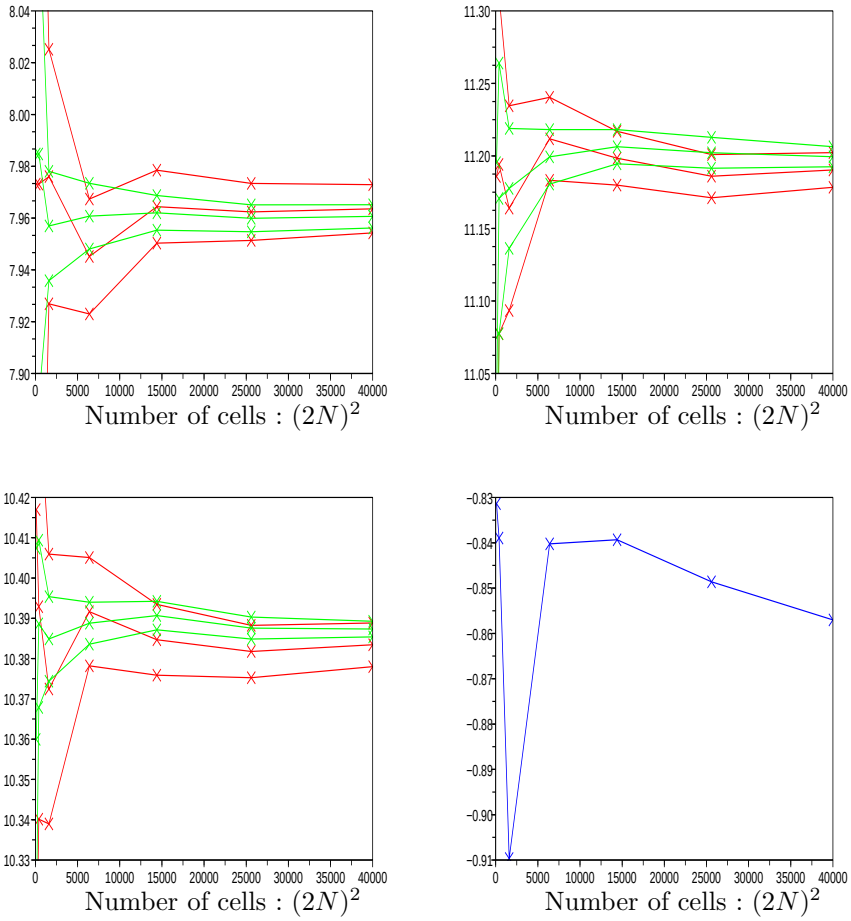


FIG. 3.2 – Estimated means (with confidence intervals) for $[A_N^\star]_{11}$ (red) and $\left[\widetilde{A}_N^\star\right]_{11}$ (green), in the cases $a_0 \sim \mathcal{B}(1/2)$ (top left), $a_0 \sim \mathcal{B}(1/3)$ (top right) and $a_0 \sim \mathcal{U}([\alpha, \beta])$ (bottom left). In the latter case, we also plot the estimator (3.23) of the normalized covariance between $[A_N^\star]_{11}$ and $[B_N^\star]_{11}$ (bottom right).

on the coefficients of the equation (because the map that associates the solution to the coefficients of the equation is a highly nonlinear nonlocal map). Note also that a systematic way to investigate the question would of course be to study the variance of the homogenized operator itself (or of its eigenelements) and it is indeed on our agenda to do so in a more extensive article [P4]. But for the time being, we briefly mention here a possible variance reduction approach on the solution $u^\star$, for a given representative right-hand side f.

In principle, one may think of several possible ways for computing the solution $u^\star$ to the homogenized problem (3.4). A first approach, which we denote by **(M1)**, consists in the following schematic sequence of computations

$$(a^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq M} \overset{\text{corrector pb}}{\longrightarrow} (A_N^{\star,\mathbf{m}}(\omega))_{1\leq \mathbf{m}\leq M} \overset{\frac{1}{M}\sum}{\longrightarrow} \mu_M\left(A_N^\star\right) \overset{(3.24)}{\longrightarrow} u^\star_{N,M},$$

where $u^\star_{N,M}$ solves the boundary value problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(\mu_M\left(A_N^\star\right)(\omega)\,\nabla u^\star_{N,M}(x,\omega)\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u^\star_{N,M}(x,\omega) &=& 0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right. \tag{3.24}$$

In short, **(M1)** consists in *first* approximating $A^\star$ using the Monte Carlo approach and its outcome $\mu_M\left(A_N^\star\right)$, and *next* to solve for $u^\star_{N,M}$.

A second approach, **(M2)**, consists in the sequence

$$(a^{\mathbf{m}}(x,\omega))_{1\leq \mathbf{m}\leq M} \overset{\text{corrector pb}}{\longrightarrow} (A_N^{\star,\mathbf{m}}(\omega))_{1\leq \mathbf{m}\leq M} \overset{(3.25)}{\longrightarrow} (u_N^{\star,\mathbf{m}}(\cdot,\omega))_{1\leq \mathbf{m}\leq M}.$$

Otherwise stated, for each $1\leq \mathbf{m}\leq M$, the problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A_N^{\star,\mathbf{m}}\nabla u_N^{\star,\mathbf{m}}\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u_N^{\star,\mathbf{m}} &=& 0 \quad \text{on} \quad \partial\mathcal{D}, \end{array} \right. \tag{3.25}$$

is first solved, and the empirical mean and variance of the corresponding solutions are next constructed :

$$\begin{array}{rcl} \mu_M\left(u_N^\star\right)(x,\omega) &=& \dfrac{1}{M}\displaystyle\sum_{\mathbf{m}=1}^{M} u_N^{\star,\mathbf{m}}(x,\omega), \\ \sigma_M\left(u_N^\star\right)(x,\omega) &=& \dfrac{1}{M-1}\displaystyle\sum_{\mathbf{m}=1}^{M} \left(u_N^{\star,\mathbf{m}}(x,\omega)-\mu_M\left(u_N^\star\right)(x,\omega)\right)^2. \end{array} \tag{3.26}$$

The empirical mean is then taken as the approximation of our seeked solution $u^\star$.

Of course, it is immediately seen that a set of approaches, intermediate between **(M1)** and **(M2)**, can be designed. This is the set of approaches **(M3)**. For each $1\leq \mathbf{m}\leq M$, we first solve the corrector problem, and thus obtain $A_N^{\star,\mathbf{m}}(\omega)$. We next set $M=PR$, and define, for each $1\leq \mathbf{r}\leq R$,

$$\mu_P^{\mathbf{r}}\left(A_N^\star\right)(\omega) = \frac{1}{P}\sum_{\mathbf{p}=1}^{P} A_N^{\star,\mathbf{p}+(\mathbf{r}-1)P}(\omega),$$

which is an empirical mean computed with P realizations among the M available realizations. For each $1 \leq \mathbf{r} \leq R$, we next solve the boundary value problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(\mu_P^{\mathbf{r}}\left(A_N^\star\right)\nabla u_N^{\star,\mathbf{r}}\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u_N^{\star,\mathbf{r}} &=& 0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right.$$

The estimators for $u^\star$ then are

$$\begin{aligned} \mu_{R,P}\left(u_N^\star\right)(x,\omega) &= \frac{1}{R}\sum_{\mathbf{r}=1}^{R} u_N^{\star,\mathbf{r}}(x,\omega), \\ \sigma_{R,P}\left(u_N^\star\right)(x,\omega) &= \frac{1}{R-1}\sum_{\mathbf{r}=1}^{R}\left(u_N^{\star,\mathbf{r}}(x,\omega) - \mu_{R,P}\left(u_N^\star\right)(x,\omega)\right)^2. \end{aligned}$$

We observe that, in dimension one, the solution of (3.25) satisfies

$$\left(u_N^{\star,\mathbf{m}}\right)'(x,\omega) = -\frac{1}{a_N^{\star,\mathbf{m}}(\omega)}\left(F(x) - \frac{1}{|\mathcal{D}|}\int_{\mathcal{D}} F\right),$$

where $F(x)$ is such that $F'(x) = f(x)$. Hence, in view of (3.10) and (3.11), we have

$$\mathbb{E}\left[\left(u_N^{\star,\mathbf{m}}\right)'\right] = -\frac{1}{a^\star}\left(F(x) - \frac{1}{|\mathcal{D}|}\int_{\mathcal{D}} F\right) = \mathbb{E}\left[\left(u^\star\right)'\right].$$

As a consequence, the empirical mean built following approach **(M2)**, namely $\mu_M\left(u_N^\star\right)(x,\omega)$ defined by (3.26), is an unbiased estimator of $u^\star(x)$, for any finite N and M, in the one-dimensional case. The estimators built following approaches **(M1)** and **(M3)** do not share this property.

In the present work, we only consider approach **(M2)**, leaving the study of the other approaches for future works. We apply the exact same technique as above, considering antithetic variables to reduce the variance. The variance under consideration is however now that of the approximation of $u^\star$.

We consider the test case (iii) defined in the previous section. We choose the right-hand side $f(x,y) = (x-0.5)^2 + (y-0.5)^2$ on the domain $\mathcal{D} = (0,1)^2$ (similar results have been obtained with other right-hand sides). The efficiency of the antithetic variable technique is assessed using the following ratio

$$R\left(u_N^\star\right) = \operatorname*{Ess\,Inf}_{x\in\mathcal{D}} \frac{\sigma_{100}\left(u_N^\star\right)}{2\sigma_{50}\left(\widetilde{u}_N^\star\right)}. \tag{3.27}$$

We have also checked that the technique does not introduce any bias by monitoring the estimator

$$\operatorname*{Ess\,Sup}_{x\in\mathcal{D}} \left|\frac{\mu_{100}\left(u_N^\star\right) - \mu_{50}\left(\widetilde{u}_N^\star\right)}{\mu_{100}\left(u_N^\star\right)}\right|. \tag{3.28}$$

Numerical results are gathered in Table 3.3. We observe that the technique does not introduce any bias, and that, again, a significant variance reduction, at fixed computational cost, is obtained.

N	Estimator (3.28)	Estimator (3.27)
5	4.20×10^{-4}	10.1
10	3.80×10^{-4}	10.9
20	1.56×10^{-3}	14.6
40	4.05×10^{-4}	11.8
80	5.21×10^{-4}	9.10
100	3.24×10^{-4}	9.02

TAB. 3.3 – Estimator (3.28) of the bias, and estimator (3.27) of the variance reduction, in the case $a_0 \sim \mathcal{U}([\alpha, \beta])$ (the equation (3.25) has been solved on a mesh of size $h = 0.1$).

Acknowledgments

The authors thank Xavier Blanc for several stimulating discussions. Part of this work was initiated while the second author was visiting the Institute for Mathematics and its Applications and the Department of Mathematics of the University of Minnesota. The hospitality of these institutions is gratefully acknowledged. The work of the authors is partially supported by ONR under contract Grant 00014-09-1-0470.

Chapitre 4

Tests numériques

Le contenu de ce chapitre correspond au travail publié dans [P4]. Nous avons montré dans le chapitre 3, à l'aide de tests numériques dans des cas simples, que la technique des variables antithétiques réduit la variance de manière systématique. Cependant l'analyse théorique ne permet pas de quantifier cette réduction de variance constatée en pratique. L'objectif de ce chapitre est précisément de quantifier la réduction de variance résultant de l'usage des variables antithétiques pour des modèles de matériaux plus représentatifs des matériaux réels. On cherche à vérifier la *robustesse* de notre approche vis à vis du champ hétérogène d'entrée et de la quantité de sortie. Ainsi on s'intéresse à des champs hétérogènes plus variés : des champs isotropes *corrélés* correspondant à des matrices A de forme

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega)\mathrm{Id} = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F\left(\{X_{k+j}\}_{|j|_\infty\leq p}(\omega)\right)\mathrm{Id},$$

où $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ désigne une suite de variables aléatoires indépendantes et identiquement distribuées (voir (4.18)). On traite aussi le cas de champs non isotropes

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) A_k(\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F(X_k(\omega)),$$

où $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ désigne une famille de vecteurs aléatoires (voir (4.19)).
Nous nous intéressons également à la réduction de variance pour des quantités d'intérêt non étudiées dans le cadre de [P2] : termes hors-diagonaux $A^\star_{12}$ de la matrice homogénéisée, valeurs propres de la matrice $A^\star$ et valeurs propres de l'opérateur $-\mathrm{div}\,(A^\star\nabla)$.

Nous constatons ainsi que la corrélation et l'anisotropie n'affectent pas l'efficacité de la méthode à variance relative du champ d'entrée constante. La méthode est robuste. Nous constatons en outre que la variance de certaines quantités qui n'entrent pas dans le cadre théorique décrit dans [P3] est aussi réduite.

Variance reduction in stochastic homogenization : the technique of antithetic variables

Xavier Blanc[a], Ronan Costaouec[b], Claude Le Bris[b] and Frédéric Legoll[b]

[a] *CEA, DAM, DIF, 91297 Arpajon, France*
[b] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

{*costaour,lebris*}*@cermics.enpc.fr,*
legoll@lami.enpc.fr,blanc@ann.jussieu.fr

4.1 Introduction

Several settings in homogenization require the solution of corrector problems posed on the entire space $\mathbb{R}^d$. In practice, truncations of these problems over bounded domains are considered and the homogenized coefficients are obtained in the limit of large domains. The question arises as to accelerate such computations. In the random case, the main difficulty is related to the intrinsic noise present in the simulation. Although very well investigated in other application fields such as financial mathematics, variance reduction techniques seem to have not been applied to the context of stochastic homogenization. In a previous article (see [P2]), we have presented a first attempt to reduce the variance in stochastic homogenization using antithetic random variables. For this purpose, we have considered a simple situation. In particular, the equation under consideration was an elliptic equation in divergence form, with a *scalar* coefficient. In addition, the coefficient was assumed to consist of independent, identically distributed random variables set on a simple mesh (see (4.16) below). Though a bit restrictive, this situation pointed out that using antithetic variables results practically in diminishing the variance for the diagonal terms of the approximated homogenized matrix. We thus obtained an effective gain in computational time at fixed accuracy. Beyond this practical validation, we have also demonstrated, on a theoretical level and for some sufficiently simple situations, that the technique does reduce variance. The theoretical arguments of [P3] not only apply to the examples of scalar random fields that we previously considered in [P2] but they extend to a wider range of random fields. We present here some numerical tests that show that the technique still efficiently reduces variance in the presence

of correlations and for matrices more general than those considered in our previous contributions. We also investigate variance reduction for eigenproblems. For convenience of the reader and consistency, we devote the rest of the present section to a brief introductory exposition of random homogenization, the related numerical challenges, and the technique of antithetic variables. We turn in Section 4.2 to homogenization problems for materials that have correlations or that are anisotropic. Section 4.3 discusses variance reduction for eigenproblems.

Other variance reduction techniques, such as techniques based on control variates, will be the subject of future investigations and will be reported on elsewhere.

4.1.1 Homogenization theoretical setting

To begin with, we introduce the basic setting of stochastic homogenization we will employ. We refer to [40] for a general, numerically oriented presentation, and to [16,27,58] for classical textbooks. We also refer to [19,20] or [63] for a presentation of our particular setting. Throughout this article, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and we denote by $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ the expectation value of any random variable $X \in L^1(\Omega, d\mathbb{P})$. We next fix $d \in \mathbb{N}^*$ (the ambient physical dimension), and assume that the group $(\mathbb{Z}^d, +)$ acts on Ω. We denote by $(\tau_k)_{k\in\mathbb{Z}^d}$ this action, and assume that it preserves the measure $\mathbb{P}$, that is, for all $k \in \mathbb{Z}^d$ and all $A \in \mathcal{F}$, $\mathbb{P}(\tau_k A) = \mathbb{P}(A)$. We assume that the action τ is *ergodic*, that is, if $A \in \mathcal{F}$ is such that $\tau_k A = A$ for any $k \in \mathbb{Z}^d$, then $\mathbb{P}(A) = 0$ or 1. In addition, we define the following notion of stationarity (see [20]) : any $F \in L^1_{\rm loc}\left(\mathbb{R}^d, L^1(\Omega)\right)$ is said to be *stationary* if, for all $k \in \mathbb{Z}^d$,

$$F(x + k, \omega) = F(x, \tau_k \omega), \tag{4.1}$$

almost everywhere in x and almost surely. In this setting, the ergodic theorem [61,89] can be stated as follows : *Let $F \in L^\infty\left(\mathbb{R}^d, L^1(\Omega)\right)$ be a stationary random variable in the above sense. For $k = (k_1, k_2, \ldots, k_d) \in \mathbb{Z}^d$, we set $|k|_\infty = \sup_{1\le i\le d} |k_i|$. Then*

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \le N} F(x, \tau_k \omega) \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left(F(x, \cdot)\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

This implies that (denoting by Q the unit cube in $\mathbb{R}^d$)

$$F\left(\frac{x}{\varepsilon}, \omega\right) \overset{*}{\underset{\varepsilon\to 0}{\rightharpoonup}} \mathbb{E}\left(\int_Q F(x, \cdot) dx\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

Besides technicalities, the purpose of the above setting is simply to formalize that, even though realizations may vary, the function F at point $x \in \mathbb{R}^d$ and the function F at point $x + k$, $k \in \mathbb{Z}^d$, share the same law. In the homogenization context we now turn to, this means that the local, microscopic environment (encoded in the matrix A) is everywhere the same *on average*. From this, homogenized, macroscopic properties will follow. In addition, and this is evident reading the above setting,

the microscopic environment considered has a relation to an underlying *periodic* structure (thus the integer shifts k in (4.1)).

We now consider the elliptic boundary value problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A\left(\dfrac{x}{\varepsilon},\omega\right)\nabla u^\varepsilon\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u^\varepsilon &=& 0 \quad \text{on} \quad \partial\mathcal{D}, \end{array} \right. \tag{4.2}$$

set on a domain $\mathcal{D}$ that is an open, regular, bounded subset of $\mathbb{R}^d$. The right-hand side is an L^2 function f on $\mathcal{D}$. The random symmetric matrix A is assumed stationary in the sense (4.1) defined above. We also assume that A is bounded and that, in the sense of quadratic forms, A is positive and almost surely bounded away from zero. The mathematical theory of homogenization states that when ε goes to zero, up to the extraction of a subsequence, u^ε converges to a *deterministic* function $u^\star$ that is solution of the so-called homogenized problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A^\star\nabla u^\star\right) &=& f \quad \text{in} \quad \mathcal{D}. \\ u^\star &=& 0 \quad \text{on} \quad \partial\mathcal{D}. \end{array} \right. \tag{4.3}$$

In contrast to problem (4.2), problem (4.3) is *deterministic* and *does not involve the small scale* ε. It is hence easier to solve. Yet, the practical computation of the homogenized matrix $A^\star$, necessary for solving (4.3), is challenging. In our specific setting, this matrix reads

$$A^\star_{ij} = \int_Q \mathbb{E}\left[\left(\nabla w_{e_j}(x,\cdot)+e_j\right)^T A(x,\cdot)\left(\nabla w_{e_i}(x,\cdot)+e_i\right)\right] dx, \tag{4.4}$$

where, for any vector $p \in \mathbb{R}^d$, the *corrector* w_p is the solution (unique up to the addition of a random constant) in $\left\{w \in L^2_{\mathrm{loc}}(\mathbb{R}^d, L^2(\Omega)),\ \nabla w \in L^2_{\mathrm{unif}}(\mathbb{R}^d, L^2(\Omega))\right\}$ to

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A(\nabla w_p + p)\right] = 0 \quad \text{on } \mathbb{R}^d \ \text{ a.s.}, \\ \nabla w_p \text{ is stationary in the sense of (4.1)}, \\ \displaystyle\int_Q \mathbb{E}(\nabla w_p) = 0. \end{array} \right. \tag{4.5}$$

The major practical difficulty of random homogenization lies in the fact that the above problem (4.5), necessary for determining the homogenized matrix $A^\star$, is posed on the entire space $\mathbb{R}^d$.

4.1.2 Numerical approach

In practice, the corrector problem (4.5), posed on the whole space $\mathbb{R}^d$, is approximated by the *truncated* corrector problem :

$$\left\{ \begin{array}{l} -\mathrm{div}\left(A(\cdot,\omega)\left(p+\nabla w_p^N(\cdot,\omega)\right)\right) = 0 \quad \text{on} \quad \mathbb{R}^d, \\ w_p^N(\cdot,\omega) \text{ is } Q_N\text{-periodic}, \end{array} \right. \tag{4.6}$$

posed on the cube Q_N, of size $(2N+1)^d$, centered at the origin. Correspondingly, the matrix $A^\star$ is then approximated by the *random* matrix

$$[A_N^\star]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} \left(e_i + \nabla w_{e_i}^N(y,\omega)\right)^T A(y,\omega)\left(e_j + \nabla w_{e_j}^N(y,\omega)\right) dy. \quad (4.7)$$

Although $A^\star$ itself is a deterministic object, its practical approximation $A_N^\star$ is random. It is only in the limit of infinitely large domains Q_N that the deterministic value is attained [23].

Besides the homogenized matrix $A^\star$ itself, other related quantities, such as the eigenelements of the matrix $A^\star$, the solution $u^\star$ of the homogenized problem (4.3), and the eigenelements of the operator $L_{A^\star} = -\mathrm{div}\,(A^\star\nabla\cdot)$, are of major interest. They all reflect some property of the homogenized material. As is the case for $A^\star$, only *random* approximations of those quantities are accessible. We formalize this saying that all these quantities, denoted by $\mathcal{F}(A^\star)$, are approximated by the corresponding random variables $\mathcal{F}(A_N^\star(\omega))$ obtained by truncation and approximation (using a Monte Carlo method). For simplicity, we will suppose from now on that $\mathcal{F} : \mathbb{S}_d(\mathbb{R}) \to \mathbb{R}$. Let $(A^{\mathbf{m}}(x,\omega))_{1\le\mathbf{m}\le M}$ denote M independent and identically distributed underlying random fields. We define a family $(A_N^{\star,\mathbf{m}})_{1\le\mathbf{m}\le M}$ of i.i.d. homogenized matrices by, for any $1 \le i,j \le d$,

$$[A_N^{\star,\mathbf{m}}]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} \left(e_i + \nabla w_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T A^{\mathbf{m}}(\cdot,\omega)\ \left(e_j + \nabla w_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right),$$

where $w_{e_j}^{N,\mathbf{m}}$ is the solution of the corrector problem associated to $A^{\mathbf{m}}$. Then we define for each quantity $\mathcal{F}(A_N^\star)$ the empirical mean and variance

$$\begin{aligned} \mu_M(\mathcal{F}(A_N^\star)) &= \frac{1}{M}\sum_{\mathbf{m}=1}^{M} \mathcal{F}(A_N^{\star,\mathbf{m}}), \\ \sigma_M(\mathcal{F}(A_N^\star)) &= \frac{1}{M-1}\sum_{\mathbf{m}=1}^{M} \left(\mathcal{F}(A_N^{\star,\mathbf{m}}) - \mu_M(\mathcal{F}(A_N^\star))\right)^2. \end{aligned} \quad (4.8)$$

Since the matrices $A_N^{\star,\mathbf{m}}$ are i.i.d. the strong law of large numbers applies :

$$\mu_M(\mathcal{F}(A_N^\star))(\omega) \underset{M\to+\infty}{\longrightarrow} \mathbb{E}(\mathcal{F}(A_N^\star)) \text{ almost surely.}$$

The central limit theorem then yields

$$\sqrt{M}\left(\mu_M(\mathcal{F}(A_N^\star)) - \mathbb{E}(\mathcal{F}(A_N^\star))\right) \underset{M\to+\infty}{\xrightarrow{\mathcal{L}}} \sqrt{\mathbb{V}\mathrm{ar}(\mathcal{F}(A_N^\star))}\,\mathcal{N}(0,1), \quad (4.9)$$

where the convergence holds in law, and $\mathcal{N}(0,1)$ denotes the standard Gaussian law. Introducing its 95 percent quantile, it is standard to consider that the exact mean $\mathbb{E}(\mathcal{F}(A_N^\star))$ lies in the interval

$$\left[\mu_M(\mathcal{F}(A_N^\star)) - 1.96\frac{\sqrt{\sigma_M(\mathcal{F}(A_N^\star))}}{\sqrt{M}},\ \mu_M(\mathcal{F}(A_N^\star)) + 1.96\frac{\sqrt{\sigma_M(\mathcal{F}(A_N^\star))}}{\sqrt{M}}\right]. \quad (4.10)$$

The value $\mu_M\left(\mathcal{F}\left(A_N^\star\right)\right)$ is thus, for both M and N sufficiently large, adopted as the approximation of the exact value $\mathcal{F}\left(A^\star\right)$.

Our aim is to design a numerical technique that, for finite N, allows to compute a better approximation of $\mathbb{E}\left(\mathcal{F}\left(A_N^\star\right)\right)$, *e.g.* an approximation with smaller variance.

4.1.3 The technique of antithetic variables

The application of variance reduction using the antithetic variable technique, a classical variance reduction technique ubiquitous in many applied fields, to the specific framework of stochastic homogenization was first performed in [P2]. For the sake of completeness we outline here the basic steps of the approach in our specific context. For an elementary introduction to the technique, we refer to [69].

Fix $M = 2\mathcal{M}$ and suppose that we give ourselves $\mathcal{M}$ i.i.d. copies $\left(A^{\mathbf{m}}(x,\omega)\right)_{1\leq \mathbf{m}\leq \mathcal{M}}$ of $A(x,\omega)$. Construct next $\mathcal{M}$ i.i.d. *antithetic* random fields

$$B^{\mathbf{m}}(x,\omega) = T\left(A^{\mathbf{m}}(x,\omega)\right), \quad 1 \leq \mathbf{m} \leq \mathcal{M},$$

from the $\left(A^{\mathbf{m}}(x,\omega)\right)_{1\leq \mathbf{m}\leq \mathcal{M}}$. The map T transforms the random field $A^{\mathbf{m}}$ into another, so-called *antithetic*, field $B^{\mathbf{m}}$. The transformation is performed in such a way that, for each $\mathbf{m}$, $B^{\mathbf{m}}$ has the same law as $A^{\mathbf{m}}$. Somewhat vaguely stated, if the coefficient was obtained in a coin tossing game (using a fair coin), then the antithetic coefficient would be head each time the original coefficient is tail and vice versa. More specifically in our context, see an example in (4.30)-(4.31) below. Then, for each $1 \leq \mathbf{m} \leq \mathcal{M}$, we solve two corrector problems. One is associated to the original $A^{\mathbf{m}}$, the other one is associated to the antithetic field $B^{\mathbf{m}}$. Using its solution $v_p^{N,\mathbf{m}}$, we define the *antithetic homogenized matrix* $B_N^{\star,\mathbf{m}}$, the entries of which read, for $1 \leq i,j \leq d$,

$$\left[B_N^{\star,\mathbf{m}}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}\left(e_i + \nabla v_{e_i}^{N,\mathbf{m}}(\cdot,\omega)\right)^T B^{\mathbf{m}}(\cdot,\omega)\,\left(e_j + \nabla v_{e_j}^{N,\mathbf{m}}(\cdot,\omega)\right).$$

And finally we set, for any $1 \leq \mathbf{m} \leq \mathcal{M}$,

$$\widetilde{A}_N^{\star,\mathbf{m}}(\omega) := \frac{1}{2}\left(A_N^{\star,\mathbf{m}}(\omega) + B_N^{\star,\mathbf{m}}(\omega)\right). \tag{4.11}$$

Since $A^{\mathbf{m}}$ and $B^{\mathbf{m}}$ are identically distributed, so are $A_N^{\star,\mathbf{m}}$ and $B_N^{\star,\mathbf{m}}$. Thus, $\widetilde{A}_N^{\star,\mathbf{m}}$ is unbiased (that is, $\mathbb{E}\left(\widetilde{A}_N^{\star,\mathbf{m}}\right) = \mathbb{E}\left(A_N^{\star,\mathbf{m}}\right)$). In addition, it satisfies :

$$\widetilde{A}_N^{\star,\mathbf{m}} \underset{N\to+\infty}{\longrightarrow} A^\star \text{ almost surely,}$$

because B is ergodic. The matrix $\widetilde{A}_N^\star$ is thus an alternative random variable that converges almost surely to $A^\star$ when $N \to \infty$. In addition, for any N, the mean of $\widetilde{A}_N^\star$ is equal to that of $A_N^\star$. Consequently, $\widetilde{A}_N^\star$ can be used to define new estimators.

In order to compute an approximation of $\mathbb{E}\left(\mathcal{F}(A_N^\star)\right)$, we use the antithetic variable defined above, and define

$$\begin{aligned}\mu_{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right) &= \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\widetilde{\mathcal{F}}\left(A_N^{\star,\mathbf{m}}\right) = \frac{1}{\mathcal{M}}\sum_{\mathbf{m}=1}^{\mathcal{M}}\frac{\mathcal{F}\left(A_N^{\star,\mathbf{m}}\right)+\mathcal{F}\left(B_N^{\star,\mathbf{m}}\right)}{2},\\ \sigma_{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right) &= \frac{1}{\mathcal{M}-1}\sum_{\mathbf{m}=1}^{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^{\star,\mathbf{m}}\right)-\mu_{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)\right)^2,\end{aligned} \tag{4.12}$$

which require $2\mathcal{M}$ resolutions of corrector problems, i.e. as many as the classical estimators (4.8). Our new random variable has variance

$$\mathbb{V}\mathrm{ar}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right) = \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\mathcal{F}\left(A_N^\star\right)\right)+\frac{1}{2}\mathbb{C}\mathrm{ov}\left(\mathcal{F}\left(A_N^\star\right),\mathcal{F}\left(B_N^\star\right)\right). \tag{4.13}$$

Applying the central limit theorem to $\widetilde{\mathcal{F}}\left(A_N^\star\right)$, we obtain

$$\sqrt{\mathcal{M}}\left(\mu_{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)-\mathbb{E}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)\right)\underset{\mathcal{M}\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}}\sqrt{\mathbb{V}\mathrm{ar}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)}\,\mathcal{N}(0,1). \tag{4.14}$$

Similarly to (4.10), we deduce a confidence interval from this convergence. The exact mean $\mathbb{E}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)$ is equal to $\mu_{\mathcal{M}}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)$ within a typical margin of error

$$1.96\frac{\sqrt{\mathbb{V}\mathrm{ar}\left(\widetilde{\mathcal{F}}\left(A_N^\star\right)\right)}}{\sqrt{\mathcal{M}}}.$$

We see on (4.13) that, when

$$\mathbb{C}\mathrm{ov}\left(\mathcal{F}\left(A_N^\star\right),\mathcal{F}\left(B_N^\star\right)\right)\leq 0, \tag{4.15}$$

the width of the interval of confidence has been reduced by the approach, and, consequently, the quality of approximation at given computational cost has improved.

4.1.4 A brief summary of our former results

We have considered in [P2] the case of an isotropic random field

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)a_k(\omega)\mathrm{Id} = \sum_{k\in\mathbb{Z}^d}\mathbf{1}_{Q+k}(x)f\left(X_k(\omega)\right)\mathrm{Id} \tag{4.16}$$

where $(X_k(\omega))_{k\in\mathbb{Z}^d}$ is a family of independent uniform random variables and f is a monotone function. For well-posedness, we assume that there exist $\alpha>0$ and $\beta<\infty$ such that, for all k, $0<\alpha\leq a_k\leq\beta<+\infty$ almost surely. Consequently, A is uniformly coercive and bounded. The quantity we mainly considered in [P2] is

$\mathcal{F}(A_N^\star) = [A_N^\star]_{ii}$, an approximation of a diagonal entry of the matrix $A^\star$. We have demonstrated numerically the efficiency of the approach.

Another purpose of [P2] was to investigate the approach theoretically. The one-dimensional setting was addressed. The study has been complemented by a study in higher dimensions in [P3]. A particularly useful ingredient, theoretically, is, somewhat vaguely stated, the monotonicity of the homogenized objects in function of the original random field. More precisely, we proved in [P3] that variance is indeed reduced as long as the output $\mathcal{F}(A_N^\star)$ we consider is monotone with respect to each of the uniform random variables. The arguments given in [P3] apply under the following *structure hypothesis* on A : for any N, there exists an integer n (possibly $n = |Q_N|$, but not necessarily) and a function $\mathcal{A}$, defined on $Q_N \times \mathbb{R}^n$, such that the tensor $A(x,\omega)$ writes

$$\forall x \in Q_N, \quad A(x,\omega) = \mathcal{A}(x, X_1(\omega), \ldots, X_n(\omega)) \quad \text{a.s.}, \tag{4.17}$$

where $\{X_k(\omega)\}_{1\leq k\leq n}$ are independent scalar random variables, which are all distributed according to the uniform law $\mathcal{U}[0,1]$. Then the global monotonicity of $\mathcal{F}(A_N^\star)$ is related to the following composition scheme

$$\{X_k(\omega)\}_{1\leq k\leq n} \xrightarrow{\mathcal{A}} A(x,\omega) \xrightarrow{\mathcal{H}} A_N^\star(\omega) \xrightarrow{\mathcal{F}} \mathcal{F}(A_N^\star),$$

where $\mathcal{H}$ denotes the application associated to periodic homogenization. Since $\mathcal{H}$ is increasing in the sense of symmetric matrices, the global monotonicity only depends on our way to model randomness $\mathcal{A}$ and the output $\mathcal{F}$ we are interested in. In [P3], we proved that variance is indeed reduced by the approach described in Section 4.1.3 when $\mathcal{A}$ is non-decreasing with respect to each of its argument, and $\mathcal{F}$ is monotone.

4.2 Variance reduction for problems involving correlations or anisotropy

Our theoretical results encourage us to apply the technique to more general cases than the simple cases addressed in [P2]. We will subsequently consider in this section two specific situations :

– **correlated isotropic fields**, that is matrices A in (4.2) of the form

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega) \mathrm{Id} = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F\left(\{X_{k+j}\}_{|j|_\infty\leq p}(\omega)\right) \mathrm{Id}, \tag{4.18}$$

where p is some fixed non-negative integer, $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ is a family of independent real-valued random variables and F is defined on $\mathbb{R}^{2p+1}$ and real valued ;

– **i.i.d. anisotropic fields**, that is matrices A in (4.2) of the form

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) A_k(\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F(X_k(\omega)), \tag{4.19}$$

where $\{X_k(\omega)\}_{k\in\mathbb{Z}^d}$ is a family of independent $\mathbb{R}^{N_{rv}}$-valued random vectors, the components of which are independent and identically distributed (we choose the uniform law). The function F, defined on $\mathbb{R}^{N_{rv}}$, is valued in the set of symmetric matrices.

Of course, we could combine the structure assumptions (4.18) and (4.19) to form correlated anisotropic random fields, but we will not proceed in this direction here.

In the case of correlated fields, in line with the theoretical observations of [P3] recalled in the previous section, we assume that the function F is non decreasing with respect to each of its arguments. In the case of anisotropic fields, we will first consider functions F that are non-decreasing. To check the robustness of the approach, we will second consider functions F that are non monotone.

We will specifically investigate four questions.

First, considering the correlated isotropic case, we will try to understand how correlation affects the efficiency of our variance reduction technique (see Section 4.2.1). To this end, we consider variance reduction of the diagonal entries of the matrix $A^\star$, first on the correlated case (4.18), second on an uncorrelated case, as we previously did in [P2]. Comparing the two cases will outline the influence of correlation. In this context, the monotonicity assumptions are satisfied and we are thus proceeding on a sound theoretical ground.

Second, we will use anisotropic fields generated using *monotone* functions $\mathcal{A}$ in (4.17) (that is, monotone functions F in (4.19)), and that have homogenized matrices with non trivial off-diagonal terms (see Section 4.2.2, Example 1). We will double-check that variance is reduced on diagonal terms as was the case in our previous study. As for off-diagonal terms, which are *not* monotone functions of the random fields, we cannot rely on any theoretical guideline. As our experiments will show, we still reduce variance, though.

Third, we will consider anisotropic fields that do not correspond to monotone functions $\mathcal{A}$ (they are of the form (4.19) with a non-monotone F). Absent any theoretical analysis, we investigate numerically variance reduction on both diagonal and off-diagonal terms (see Section 4.2.2, Examples 2 and 3).

Fourth, again using anisotropic fields, we will consider variance reduction of eigenelements (see Section 4.3).

4.2.1 Correlated cases

We consider a two dimensional situation and proceed computationally as explained in Section 4.1. We restrict ourselves to considering the first diagonal entry

$[A_N^\star]_{11}$. In order to investigate the role of correlation, we consider random fields of form (4.18)

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F\left(\{a_{k+j}\}_{|j|_\infty\leq p}(\omega)\right) \mathrm{Id}, \tag{4.20}$$

with correlation length p. We begin with the case $p = 1$ and next consider some larger values of p. In order to focus on the effect of correlation, we will not only monitor the variance reduction for the homogenized matrix $A^\star$ associated to the above matrix A. We also consider a similar matrix, where the correlation is set to zero, and apply the variance reduction technique for its homogenization. More precisely, we introduce

$$C(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F\left(\{c_{k,j}\}_{|j|_\infty\leq 1}(\omega)\right) \mathrm{Id}, \tag{4.21}$$

where $\left\{(c_{k,j})_{|j|_\infty\leq 1}\right\}_{k\in\mathbb{Z}^d}$ denotes a family of i.i.d. random *vectors*, the components of which are *independent from one another* and share the exact same law as the a_k (which we take here as the uniform law). The local behaviour (meaning, the behaviour on a single unit cell) of the field C is similar to that of the field A. However, when it comes to the global fields seen as functions on the entire space, the behaviours differ, because correlation is turned off in the case of C. In the very peculiar one-dimensional situation (where homogenization is a local, pointwise, process), the homogenized matrices $A^\star$ and $C^\star$ respectively obtained from A and C are identical. The variance of the approximate matrices $[A_N^\star]$ and $[C_N^\star]$ can be different, though. Some elementary arguments allow to prove that in both cases we reduce variance using the technique of antithetic variable. In dimensions higher than or equal to 2, $A^\star \neq C^\star$. The matrix $C^\star$ serves as a useful reference to evaluate how correlation affects the efficiency of our variance reduction technique.

In the numerical examples below, the random variables $\{a_k\}_{k\in\mathbb{Z}^d}$ and $\{c_{k,j}\}_{k\in\mathbb{Z}^d,|j|_\infty\leq 1}$ are all uniformly distributed between $\alpha = 3$ and $\beta = 20$.

4.2.1.1 Influence of correlation : identical local behaviour

We define the function F in (4.20) as

$$F\left(\{a_{k+q}\}_{|q|\leq 1}\right) = \frac{1}{9}\sum_{|q|\leq 1} a_{k+q}. \tag{4.22}$$

For comparison purposes, the field C of (4.21) is, as announced above, defined by

$$F\left(\{c_{k,l}\}_{|l|\leq 1}\right) = \frac{1}{9}\sum_{|l|\leq 1} c_{k,l}. \tag{4.23}$$

Our results for the variance reduction on $[A_N^\star]_{11}$ and $[C_N^\star]_{11}$ are reported in Tables 4.1 and 4.2, respectively. As mentioned in [P2], because of isotropy and invariance by

rotation of angle $\pi/2$, the corresponding approximated homogenized matrix $A_N^\star$ and $C_N^\star$ are, like the exact homogenized matrices $A^\star$ and $C^\star$, isotropic. In each table, the ratio of variance

$$R\left([A_N^\star]_{11}\right) = \frac{\sigma_{100}\left([A_N^\star]_{11}\right)}{2\sigma_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)} \tag{4.24}$$

measures the reduction of uncertainty on estimations of $\mathbb{E}\left(A_N^\star\right)$ at fixed computational cost, that is, the efficiency of the variance reduction technique. It corresponds to the ratio of the square of the widths of intervals of confidence. We will use a similar ratio (with obvious definition and notation) for all the tables presented throughout this article.

From the consideration of Tables 4.1 and 4.2 we conclude that correlation does not affect the efficiency of the technique.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	11.3791	0.0637	11.3824	0.0054	69.25
60	11.3794	0.0438	11.3818	0.0028	121.07
80	11.3765	0.0310	11.3821	0.0029	58.42
100	11.3794	0.0259	11.3818	0.0018	97.21

TAB. 4.1 – Correlated equidistributed case (4.20)-(4.22), $p = 1$: mean and standard deviation of $[A_N^\star]_{11}$.

N	$\mu_{100}\left([C_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([C_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{C}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{C}_N^\star\right]_{11}\right)$	$R\left([C_N^\star]_{11}\right)$
40	11.3843	0.0226	11.3859	0.0021	55.64
60	11.3850	0.0153	11.3858	0.0017	40.95
80	11.3863	0.0111	11.3858	0.0012	40.53
100	11.3863	0.0091	11.3860	0.0009	51.00

TAB. 4.2 – Uncorrelated equidistributed case (4.21)-(4.23), $p = 1$: mean and standard deviation of $[C_N^\star]_{11}$.

Note that we observe here a ratio R of the order of 40, better than in [P2]. It owes to the fact that we deliberately considered in [P2] more challenging test cases in order to prove that variance can be reduced in generic situations, even demanding ones in terms of normalized variance. Here *we are focusing on the effect of correlation only*, and our purpose, different in nature from that of [P2], is to compare the correlated and the uncorrelated situations. Indeed, denoting by $A_k(\omega) = F\left(\{a_{k+q}\}_{|q|\leq 1}\right)$ and $C_k(\omega) = F\left(\{c_{k,l}\}_{|l|\leq 1}\right)$ the local values of the correlated field (4.20)-(4.22) and of

the uncorrelated field (4.21)-(4.23) respectively, we see here that the corresponding normalized variance reads

$$\frac{\mathbb{V}\text{ar } A_k}{\left(\mathbb{E}\ A_k\right)^2} = \frac{\mathbb{V}\text{ar } C_k}{\left(\mathbb{E}\ C_k\right)^2} = \frac{\mathbb{V}\text{ar } c_{0,0}}{9\left(\mathbb{E}\ c_{0,0}\right)^2} = \frac{(\beta-\alpha)^2}{27(\beta+\alpha)^2} \approx 0.0202.$$

In contrast, in the test case (iii) of [P2], the random field is

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega)\,\mathrm{Id}, \tag{4.25}$$

where $\{a_k\}_{k\in\mathbb{Z}^d}$ is a family of i.i.d. variables uniformly distributed between $\alpha_0 = 3$ and $\beta_0 = 20$. The normalized variance of the local value of $A(x,\omega)$ hence reads

$$\frac{\mathbb{V}\text{ar } a_k}{\left(\mathbb{E}\ a_k\right)^2} = \frac{(\beta_0-\alpha_0)^2}{3(\beta_0+\alpha_0)^2} \approx 0.182, \tag{4.26}$$

and is indeed 9 times as large as the normalized local variance considered here. Our formal considerations above are confirmed by the numerical results shown in Table 4.3, where we consider the test case (4.25), this time with $\alpha_0 = 3$ and $\beta_0 = 5$, so that the normalized local variance (which is now equal to 0.0208, in view of (4.26)) is close to the one of the fields (4.20)-(4.22) and (4.21)-(4.23). We again obtain an efficiency ratio close to 50. So, the normalized local variance of the fields (4.20)-(4.22), (4.21)-(4.23) (with $\alpha = 3$ and $\beta = 20$) and (4.25) (with $\alpha_0 = 3$ and $\beta_0 = 5$) are of the same order, and we indeed observe an efficiency ratio R of the same order.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	3.9597	0.0071	3.9595	0.00069	52.87
60	3.9591	0.0049	3.9595	0.00045	59.18
80	3.9589	0.0037	3.9594	0.00035	55.66
100	3.9590	0.0030	3.9594	0.00025	71.53

TAB. 4.3 – Uncorrelated case (4.25), where $a_k \sim \mathcal{U}[\alpha_0,\beta_0]$, with $\alpha_0 = 3$ and $\beta_0 = 5$: mean and standard deviation of $[A_N^\star]_{11}$.

Remark 4.2.1 *Consider, in the one-dimensional setting, the case*

$$A(x,\omega) = \sum_{k\in\mathbb{Z}} \mathbf{1}_{[k,k+1)}(x) a_k(\omega),$$

where $\{a_k\}_{k\in\mathbb{Z}}$ is a family of i.i.d. variables uniformly distributed between α_0 and β_0. Then the efficiency ratio (4.24)*, which we write here as*

$$R_N = \frac{\mathbb{V}ar\,(A_N^\star)}{2\mathbb{V}ar\left(\widetilde{A}_N^\star\right)},$$

is analytically computable. After tedious but straightforward computations, we obtain

$$\lim_{N\to\infty} R_N = R_\infty = \left[1 - \frac{g(x)\ \ln(x)}{(x-1)[1/x - (\ln(x)/(x-1))^2]}\right]^{-1} \tag{4.27}$$

where $x = \beta_0/\alpha_0 > 1$ *and* $g(x) = \ln(x)/(x-1) - 2/(1+x)$*. On Fig. 4.1, we plot* R_∞ *as a function of* x*. For any* x*, we observe that* $R_\infty > 1$*, that is the variance reduction technique is indeed efficient, and provides a more accurate estimation of* $\mathbb{E}\,(A_N^\star)$ *for an equal computational cost. We also observe that* R_∞ *is a decreasing function of* x*, which tends to* 1 *as* x *tends to infinity. This one-dimensional study also confirms our considerations above : the technique always allows to reduce the variance, but is all the more efficient as the original normalized variance (here intuitively measured by the quotient* x*, and above measured by the ratio* (4.26)*) is small.*

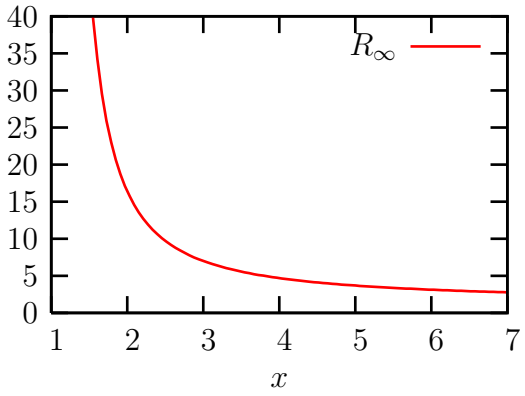


FIG. 4.1 – Variance reduction efficiency R_∞ defined by (4.27), as a function of $x = \beta_0/\alpha_0$.

4.2.1.2 Centered vs equidistributed correlation structure

We now compare two different correlation structures sharing the same correlation length $p = 1$. The first structure is the equidistributed case (4.22). As for the second structure, we consider

$$F\left(\{a_{k+q}\}_{|q|\le 1}\right) = \frac{1}{2}a_k + \frac{1}{16}\sum_{|q|\le 1; q\neq 0} a_{k+q}, \tag{4.28}$$

where, as in (4.22), $\{a_k\}_{k\in\mathbb{Z}^d}$ refers to a family of i.i.d. random variables uniformly distributed between $\alpha = 3$ and $\beta = 20$. From Tables 4.1 and 4.4, we see that the correlation structure affects the efficiency of the method, but the reduction remains significant.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	11.1531	0.0655	11.1563	0.0083	31.17
60	11.1527	0.0448	11.1551	0.0050	39.89
80	11.1494	0.0319	11.1552	0.0045	25.43
100	11.1523	0.0267	11.1548	0.0028	45.52

TAB. 4.4 – Correlated centered case (4.20)-(4.28), $p = 1$: mean and standard deviation of $[A_N^\star]_{11}$.

4.2.1.3 Longer correlation lengths

We now let our parameter p modelling the correlation length increase, and consider (4.20), with F defined by

$$F\left(\{a_{k+q}\}_{|q|\leq p}\right) = \frac{1}{(2p+1)^2}\sum_{|q|\leq p} a_{k+q}, \tag{4.29}$$

with values $p = 2$, $p = 5$ and $p = 10$ (the case $p = 1$ corresponds to (4.22)). Results are reported in Tables 4.5 through 4.7. Comparing also with Table 4.1, we see that increasing the correlation length indeed affects, in fact advantageously, the efficiency of the variance reduction.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	11.4563	0.05697	11.4574	0.0033	145.75
60	11.4556	0.03641	11.4580	0.0023	121.15
80	11.4528	0.03053	11.4579	0.0017	169.72
100	11.4554	0.02641	11.4579	0.0013	214.58

TAB. 4.5 – Correlated equidistributed case (4.20)-(4.29), $p = 2$: mean and standard deviation of $[A_N^\star]_{11}$.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	11.4871	0.0592	11.4912	0.0014	886.83
60	11.4853	0.0413	11.4914	0.0010	873.11
80	11.4882	0.0313	11.4915	0.0007	994.81
100	11.4888	0.0246	11.4914	0.0005	1015.36

TAB. 4.6 – Correlated equidistributed case (4.20)-(4.29), $p = 5$: mean and standard deviation of $[A_N^\star]_{11}$.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\tilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	11.4970	0.0596	11.4978	0.0007	3360.50
60	11.4940	0.0392	11.4977	0.0006	2473.20
80	11.4963	0.0301	11.4977	0.0003	3719.40
100	11.4937	0.0251	11.4977	0.0002	4640.11

TAB. 4.7 – Correlated equidistributed case (4.20)-(4.29), $p = 10$: mean and standard deviation of $[A_N^\star]_{11}$.

4.2.2 Anisotropic cases

4.2.2.1 Test cases

To begin with, we introduce three test cases we will focus on in the remainder of this section. They correspond to different *deterministic* functions $\mathcal{A}$, that is, different ways of constructing in (4.17) the field $A(x,\omega)$ from the random variables.

Example 1 We consider a random matrix

$$A_1(x,\omega) = P\left(\sum_{k\in\mathbb{Z}^2} \mathbf{1}_{Q+k}(x)\begin{pmatrix}\lambda_k^1(\omega) & 0\\ 0 & \lambda_k^2(\omega)\end{pmatrix}\right)P^T \quad \text{with} \quad P = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}, \tag{4.30}$$

where $\{\lambda_k^1\}_{k\in\mathbb{Z}^2}$ and $\{\lambda_k^2\}_{k\in\mathbb{Z}^2}$ are two independent families of independent random variables uniformly distributed on $[\alpha,\beta]$ and $[\delta,\gamma]$ respectively. We assume that

$$\rho_1 = \min(\alpha,\delta) > 0,$$

so that, for all $k\in\mathbb{Z}^2$, $\lambda_k^1(\omega)\geq\rho_1>0$ and $\lambda_k^2(\omega)\geq\rho_1>0$ almost surely.

This case corresponds to the *deterministic* function

$$\mathcal{A}_1\left(x,\{(y_k,z_k)\}_{k\in\mathbb{Z}^2}\right) = P\left(\sum_{k\in\mathbb{Z}^2}\mathbf{1}_{Q+k}(x)\begin{pmatrix}\alpha+(\beta-\alpha)y_k & 0\\ 0 & \delta+(\gamma-\delta)z_k\end{pmatrix}\right)P^T,$$

and to the choice

$$A_1(x,\omega) = \mathcal{A}_1\left(x,\{(Y_k(\omega),Z_k(\omega))\}_{k\in\mathbb{Z}^2}\right),$$

where Y_k and Z_k are i.i.d. random variables with uniform law on $[0,1]$. Note that $\mathcal{A}_1$ is indeed non-decreasing with respect to any y_k and z_k. The associated antithetic field is

$$B_1(x,\omega) = P\left(\sum_{k\in\mathbb{Z}^2}\mathbf{1}_{Q+k}(x)\begin{pmatrix}\alpha+\beta-\lambda_k^1(\omega) & 0\\ 0 & \delta+\gamma-\lambda_k^2(\omega)\end{pmatrix}\right)P^T. \tag{4.31}$$

Parameters values are fixed at $\alpha=5$, $\beta=20$, $\delta=25$ and $\gamma=40$.

Example 2 We choose

$$A_2(x,\omega) = \sum_{k\in\mathbb{Z}^2} \mathbf{1}_{Q+k}(x) A_k(\omega) \text{ with } A_k(\omega) = \begin{pmatrix} a_k(\omega) & b_k(\omega) \\ b_k(\omega) & a_k(\omega) \end{pmatrix}, \tag{4.32}$$

where $\{a_k\}_{k\in\mathbb{Z}^2}$ and $\{b_k\}_{k\in\mathbb{Z}^2}$ are two independent families of i.i.d. random variables uniformly distributed in $[\alpha,\beta]$ and $[\delta,\gamma]$ respectively, with

$$\alpha > 0.$$

The eigenvalues of A_k are $\lambda_k^1(\omega) = a_k(\omega) - b_k(\omega)$ and $\lambda_k^2(\omega) = a_k(\omega) + b_k(\omega)$. We thus assume that there exists $\rho_2 > 0$ such that

$$\text{for all } k \in \mathbb{Z}^2, \quad a_k(\omega) - |b_k(\omega)| \geq \rho_2 \quad \text{almost surely,}$$

so that A_2 is uniformly coercive. Note that the deterministic function $\mathcal{A}_2$ associated to A_2, which reads

$$\mathcal{A}_2\left(x, \{(y_k,z_k)\}_{k\in\mathbb{Z}^2}\right) = \sum_{k\in\mathbb{Z}^2} \mathbf{1}_{Q+k}(x) \begin{pmatrix} \alpha+(\beta-\alpha)y_k & \delta+(\gamma-\delta)z_k \\ \delta+(\gamma-\delta)z_k & \alpha+(\beta-\alpha)y_k \end{pmatrix},$$

is not monotone with respect to z_k. This case does not fall in the framework of [P3]. The antithetic field we will consider is

$$B_2(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) \begin{pmatrix} \alpha+\beta-a_k(\omega) & \gamma+\delta-b_k(\omega) \\ \gamma+\delta-b_k(\omega) & \alpha+\beta-a_k(\omega) \end{pmatrix}.$$

The numerical tests have been performed with the following parameters : $\alpha = 25$, $\beta = 40$, $\delta = 5$ and $\gamma = 20$.

Example 3 We define the random matrix

$$A_3(x,\omega) = \sum_{k\in\mathbb{Z}^2} \mathbf{1}_{Q+k}(x) A_k(\omega) \text{ with } A_k(\omega) = \begin{pmatrix} a_k(\omega) & c_k(\omega) \\ c_k(\omega) & b_k(\omega) \end{pmatrix} \tag{4.33}$$

where $\{a_k\}_{k\in\mathbb{Z}^2}$, $\{b_k\}_{k\in\mathbb{Z}^2}$ and $\{c_k\}_{k\in\mathbb{Z}^2}$ are three independent families of independent uniform random variables, uniformly distributed in $[\alpha,\beta]$, $[\delta,\gamma]$ and $[\iota,\kappa]$ respectively, with

$$\alpha > 0, \quad \delta > 0 \quad \text{and} \quad \iota > 0. \tag{4.34}$$

Uniform coercivity holds if and only if the two eigenvalues of $A_k(\omega)$ are positive and uniformly bounded away from 0. A necessary condition is that the trace and the determinant of $A_k(\omega)$ are positive and uniformly bounded away from 0, which is guaranteed under the assumptions (4.34) and the existence of $\rho_3 > 0$ such that

$$\alpha\delta - \kappa^2 \geq \rho_3 > 0. \tag{4.35}$$

The lowest eigenvalue $\lambda_k^1(\omega)$ of $A_k(\omega)$ then reads

$$\lambda_k^1(\omega) = \frac{2 \det\, A_k(\omega)}{\mathrm{Tr}\, A_k(\omega) + \sqrt{(\mathrm{Tr}\, A_k(\omega))^2 - 4 \det\, A_k(\omega)}},$$

which is bounded from below as $\det\, A_k$ is bounded from below by $\rho_3 > 0$ and $\mathrm{Tr}\, A_k$ is bounded from above by $\beta + \gamma$.

The corresponding antithetic field reads

$$B_3(x,\omega) = \sum_{k \in \mathbb{Z}^d} \mathbf{1}_{Q+k}(x) \begin{pmatrix} \alpha + \beta - a_k(\omega) & \iota + \kappa - c_k(\omega) \\ \iota + \kappa - c_k(\omega) & \delta + \gamma - b_k(\omega) \end{pmatrix}. \tag{4.36}$$

The numerical tests have been performed with the following parameters : $\alpha = 15$, $\beta = 30$, $\delta = 20$, $\gamma = 40$, $\iota = 5$ and $\kappa = 15$.

4.2.2.2 Numerical results

We consider both a diagonal term, namely $[A_N^\star]_{11}$ and an off-diagonal term, namely $[A_N^\star]_{12}$. For the former, in the case when monotonicity holds, we expect the results to be qualitatively good, since we have a theoretical result ensuring variance reduction. Our purpose is to evaluate the reduction *quantitatively*. When monotonicity does not hold, because of the particular structure considered, then we also test the reduction itself. We mention that the other diagonal entry $[A_N^\star]_{22}$ of the matrix would yield results qualitatively similar to those for $[A_N^\star]_{11}$.

Table 4.8 confirms that variance of the diagonal terms is reduced in our Example 1. The gain is rather significant. We also observe on Table 4.9 the same computational gain for the off-diagonal term, although our theoretical arguments in [P3] do not cover this case. The other Tables (Table 4.10 through Table 4.13) show that variance reduction is also obtained for our Examples 2 and 3, although no theoretical argument holds in these non-monotone settings.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	22.0595	0.0362	22.0577	0.0066	14.88
60	22.0550	0.0240	22.0575	0.0036	21.80
80	22.0570	0.0161	22.0578	0.0020	32.96
100	22.0565	0.0166	22.0578	0.0025	21.74

Tab. 4.8 – Example 1 : mean and standard deviation of the diagonal term $[A_N^\star]_{11}$.

4.3 Variance reduction for eigenproblems

As announced in the introduction, we now turn to the issue of variance reduction for eigenproblems.

N	$\mu_{100}\left([A_N^\star]_{12}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{12}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$R\left([A_N^\star]_{12}\right)$
40	- 10.0897	0.0389	- 10.0877	0.0043	41.32
60	- 10.0902	0.0266	- 10.0880	0.0024	63.93
80	- 10.0899	0.0215	- 10.0882	0.0022	48.82
100	- 10.0892	0.0169	- 10.0886	0.0019	39.10

TAB. 4.9 – Example 1 : mean and standard deviation of the off-diagonal term $[A_N^\star]_{12}$.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
40	31.8828	0.0498	31.8783	0.0098	12.85
60	31.8768	0.0337	31.8765	0.0049	23.48
80	31.8761	0.0284	31.8779	0.0036	31.50
100	31.8776	0.0242	31.8781	0.0032	29.30

TAB. 4.10 – Example 2 : mean and standard deviation of the diagonal term $[A_N^\star]_{11}$.

N	$\mu_{100}\left([A_N^\star]_{12}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{12}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$R\left([A_N^\star]_{12}\right)$
40	12.6126	0.0561	12.6118	0.0085	21.69
60	12.6083	0.0342	12.6125	0.0051	16.94
80	12.6071	0.0270	12.6127	0.0042	20.86
100	12.6106	0.0226	12.6123	0.0038	18.18

TAB. 4.11 – Example 2 : mean and standard deviation of the off-diagonal term $[A_N^\star]_{12}$.

N	$\mu_{100}\left([A_N^\star]_{11}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{11}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{11}\right)$	$R\left([A_N^\star]_{11}\right)$
20	22.0121	0.1239	22.0116	0.0125	49.44
40	22.0105	0.0571	22.0086	0.0073	30.54
60	22.0050	0.0387	22.0086	0.0046	34.39
80	22.0073	0.0282	22.0079	0.0037	28.55

TAB. 4.12 – Example 3 : mean and standard deviation of the diagonal term $[A_N^\star]_{11}$.

N	$\mu_{100}\left([A_N^\star]_{12}\right)$	$\sqrt{\sigma_{100}}\left([A_N^\star]_{12}\right)$	$\mu_{50}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$\sqrt{\sigma_{50}}\left(\left[\widetilde{A}_N^\star\right]_{12}\right)$	$R\left([A_N^\star]_{12}\right)$
20	2.5031	0.0369	2.5024	0.0046	31.61
40	2.5017	0.0193	2.5021	0.0025	30.42
60	2.5012	0.0121	2.5028	0.0016	27.25
80	2.5012	0.0092	2.5012	0.0092	37.47

TAB. 4.13 – Example 3 : mean and standard deviation of the off-diagonal term $[A_N^\star]_{12}$.

We respectively denote by $\left\{\lambda_k^A(\omega)\right\}_{1\leq k\leq d}$ and $\left\{\lambda_k^B(\omega)\right\}_{1\leq k\leq d}$ the eigenvalues of the (approximate) homogenized matrix $A_N^\star(\omega)$ and the (approximate) homogenized matrix $B_N^\star(\omega)$ obtained using the antithetic field $B(x,\omega)$. We sort these eigenvalues in non-decreasing order.

Likewise, we denote by $\left(\Lambda_k^A(\omega), u_k^A(\omega)\right)_{k\in\mathbb{N}}$ the eigenelements of the operator $L_A = -\mathrm{div}\left[A_N^\star(\omega)\nabla\cdot\right]$ with homogeneous Dirichlet boundary conditions, i.e.

$$-\mathrm{div}\left[A_N^\star(\omega)\nabla u_k^A(\omega)\right] = \Lambda_k^A(\omega)\, u_k^A(\omega)$$

with $u_k^A(\omega) \in H_0^1(\mathcal{D})$ and $\|u_k^A(\omega)\|_{L^2(\mathcal{D})} = 1$. We proceed similarly for the matrix obtained using the antithetic field $B(x,\omega)$ and consider the eigenelements of $L_B = -\mathrm{div}\left[B_N^\star(\omega)\nabla\cdot\right]$:

$$-\mathrm{div}\left[B_N^\star(\omega)\nabla u_k^B(\omega)\right] = \Lambda_k^B(\omega)\, u_k^B(\omega).$$

We also assume that, almost surely, $\Lambda_k^A(\omega)$ and $\Lambda_k^B(\omega)$ are sorted in non-decreasing order.

Our purpose here is to reduce the variance on $\mathcal{F}\left(A_N^\star\right) = \lambda_k^A$ or Λ_k^A for some $k \in \mathbb{N}$. Note that this is a monotone function of the random field (see [P3]). In the case when $\mathcal{A}$ is also monotone, the following result from [P3] applies.

Proposition 4.3.1 *Define* $\widetilde{\lambda}_k(\omega) := \frac{1}{2}\left[\lambda_k^A(\omega) + \lambda_k^B(\omega)\right]$. *Then, for all* $1 \leq k \leq d$,

$$\mathbb{E}\left(\widetilde{\lambda}_k\right) = \mathbb{E}\left(\lambda_k^A\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\lambda}_k\right) \leq \frac{1}{2}\mathbb{V}ar\left(\lambda_k^A\right).$$

Define $\widetilde{\Lambda}_k(\omega) := \frac{1}{2}\left(\Lambda_k^A(\omega) + \Lambda_k^B(\omega)\right)$. *Then, for all* $k \in \mathbb{N}$,

$$\mathbb{E}\left(\widetilde{\Lambda}_k\right) = \mathbb{E}\left(\Lambda_k^A\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\Lambda}_k\right) \leq \frac{1}{2}\mathbb{V}ar\left(\Lambda_k^A\right).$$

This guarantees that the technique indeed reduces variance. We however need a *quantitative* evaluation of the efficiency of variance reduction.

To begin with, we mention that in the one-dimensional setting, or in the case of diagonal homogenized matrices, the question of variance reduction for eigenelements reduces to elementary questions already addressed. Indeed, in the one-dimensional setting, the approximate homogenized operator reads $-a_N^\star(\omega)\dfrac{d^2}{dx^2}$ and thus its eigenfunctions are the deterministic eigenfunctions of the one-dimensional Laplacian, and its eigenvalues are likewise the deterministic eigenvalues of the one-dimensional Laplacian *multiplied* by the random quantity $a_N^\star(\omega)$. The variance reduction of the eigenelements comes down to that of $a_N^\star(\omega)$. Similarly, in the two-dimensional setting when the approximate homogenized matrix is diagonal, namely $A_N^\star(\omega) = \begin{pmatrix} a_N^\star(\omega) & 0 \\ 0 & b_N^\star(\omega) \end{pmatrix}$, the eigenfunctions and eigenvalues may again be explicitly expressed in terms of those (deterministic) of the Laplacian. All is a matter of scaling, and again the question of variance reduction is elementary and already covered by that of reducing the variance on $A_N^\star(\omega)$.

Besides these oversimplified cases, additional numerical experiments are in order. We consider the three examples defined in Section 4.2.2.1. For each of them, and for the eigenvalues of the matrix $A_N^\star$ and the eigenvalues of the operator $L_{A_N^\star}$, we study an effectivity ratio R similar to that defined in (4.24).

Tables 4.14, 4.16 and 4.18 illustrate the efficiency of the technique for the computation of the first eigenvalue for any structure of the random fields. Our results for the second eigenvalue are displayed on Tables 4.15, 4.17 and 4.19. These results show the good efficiency of the approach, for all the test cases considered.

Tables 4.20 through 4.25 illustrate the variance reduction for the first two eigenvalues of $L_{A_N^\star}$. Again, the approach performs very well. We omit to present here our results for higher eigenvalues of $L_{A_N^\star}$. They lead to similar qualitative conclusions on the good efficiency of the approach.

Acknowledgments The work of RC, CLB and FL is partially supported by ONR under Grant 00014-09-1-0470.

N	$\mu_{100}\left(\lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_1\right)$	$R\left(\lambda_1^A\right)$
40	11.9703	0.0572	11.9702	0.0075	29.19
60	11.9650	0.0385	11.9696	0.0043	39.36
80	11.9670	0.0267	11.9696	0.0035	28.63
100	11.9672	0.0233	11.9692	0.0033	24.00

TAB. 4.14 – Example 1 : mean and standard deviation of the first eigenvalue λ_1^A of the homogenized matrix.

N	$\mu_{100}\left(\lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_2\right)$	$R\left(\lambda_2^A\right)$
40	32.1496	0.0480	32.1456	0.0049	47.62
60	32.1454	0.0327	32.1455	0.0031	56.60
80	32.1467	0.0270	32.1459	0.0022	78.10
100	32.1456	0.0241	32.1463	0.0019	79.77

TAB. 4.15 – Example 1 : mean and standard deviation of the second eigenvalue λ_2^A of the homogenized matrix.

N	$\mu_{100}\left(\lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_1\right)$	$R\left(\lambda_1^A\right)$
40	19.2698	0.07870	19.2670	0.01353	16.91
60	19.2688	0.05414	19.2650	0.00799	22.96
80	19.2690	0.04196	19.2652	0.00593	25.02
100	19.2668	0.03411	19.2659	0.00557	18.70

TAB. 4.16 – Example 2 : mean and standard deviation of the first eigenvalue λ_1^A of the homogenized matrix.

N	$\mu_{100}\left(\lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_2\right)$	$R\left(\lambda_2^A\right)$
40	44.4951	0.0704	44.4905	0.0079	39.31
60	44.4854	0.0471	44.4900	0.0052	41.47
80	44.4832	0.0365	44.4905	0.0040	41.93
100	44.4880	0.0326	44.4905	0.0035	42.98

TAB. 4.17 – Example 2 : mean and standard deviation of the second eigenvalue λ_2^A of the homogenized matrix.

N	$\mu_{100}\left(\lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_1\right)$	$R\left(\lambda_1^A\right)$
20	16.2994	0.0465	16.2998	0.0051	41.59
40	16.3016	0.0235	16.2995	0.0029	32.82
60	16.3005	0.0148	16.2991	0.0016	44.61
80	16.3001	0.0115	16.2989	0.0011	50.42

TAB. 4.18 – Example 3 : mean and standard deviation of the first eigenvalue λ_1^A of the homogenized matrix.

N	$\mu_{100}\left(\lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\lambda}_2\right)$	$R\left(\lambda_2^A\right)$
20	23.1095	0.1053	23.1079	0.0099	56.70
40	23.1069	0.0472	23.1052	0.0060	31.18
60	23.1018	0.0337	23.1057	0.0043	31.24
80	23.1035	0.0249	23.1051	0.0032	29.77

TAB. 4.19 – Example 3 : mean and standard deviation of the second eigenvalue λ_2^A of the homogenized matrix.

N	$\mu_{100}\left(\Lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_1\right)$	$R\left(\Lambda_1^A\right)$
40	842.9851	1.5576	842.9140	0.2084	27.94
60	842.7855	1.0364	842.8998	0.1284	32.60
80	842.8560	0.6882	842.9044	0.0900	29.21
100	842.8422	0.6977	842.9013	0.0891	30.67

TAB. 4.20 – Example 1 : mean and standard deviation of Λ_1^A.

N	$\mu_{100}\left(\Lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_2\right)$	$R\left(\Lambda_2^A\right)$
40	1847.7161	4.4299	1847.5915	0.5976	27.47
60	1847.1983	2.9559	1847.5500	0.3593	33.86
80	1847.3861	1.9570	1847.5559	0.2678	29.45
100	1847.3705	1.8871	1847.5370	0.2614	26.05

TAB. 4.21 – Example 1 : mean and standard deviation of Λ_2^A.

N	$\mu_{100}\left(\Lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_1\right)$	$R\left(\Lambda_1^A\right)$
40	611.2645	1.0533	611.1924	0.1631	27.87
60	611.1149	0.7222	611.1646	0.0985	26.86
80	611.1549	0.6078	611.1718	0.0657	42.82
100	611.1685	0.5140	611.1796	0.0639	32.33

TAB. 4.22 – Example 2 : mean and standard deviation of Λ_1^A.

N	$\mu_{100}\left(\Lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_2\right)$	$R\left(\Lambda_2^A\right)$
40	1351.9939	2.9920	1351.8258	0.4960	19.95
60	1351.7016	2.0189	1351.7458	0.2958	23.29
80	1351.7995	1.6841	1351.7609	0.2063	33.31
100	1351.7827	1.3957	1351.7865	0.1986	24.69

TAB. 4.23 – Example 2 : mean and standard deviation of Λ_2^A.

N	$\mu_{100}\left(\Lambda_1^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_1^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_1\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_1\right)$	$R\left(\Lambda_1^A\right)$
20	389.7542	1.2979	389.7434	0.1258	53.26
40	389.7524	0.5981	389.7139	0.0774	29.86
60	389.6902	0.4072	389.7138	0.0467	38.06
80	389.7038	0.2967	389.7064	0.0375	31.37

TAB. 4.24 – Example 3 : mean and standard deviation of Λ_1^A.

N	$\mu_{100}\left(\Lambda_2^A\right)$	$\sqrt{\sigma_{100}}\left(\Lambda_2^A\right)$	$\mu_{50}\left(\widetilde{\Lambda}_2\right)$	$\sqrt{\sigma_{50}}\left(\widetilde{\Lambda}_2\right)$	$R\left(\Lambda_2^A\right)$
20	901.8832	2.1182	901.8619	0.1934	60.01
40	901.9316	0.9934	901.8325	0.1221	33.09
60	901.8420	0.6740	901.8300	0.0650	53.95
80	901.8335	0.5028	901.8193	0.0514	53.95

TAB. 4.25 – Example 3 : mean and standard deviation of Λ_2^A.

Chapitre 5

Eléments d'analyse

Le contenu de ce chapitre correspond à la publication [P3]. Son objectif est de préciser sous quelles hypothèses il est possible d'appliquer la technique des variables antithétiques pour l'homogénéisation stochastique introduite dans [P2]. La perspective est différente de celle du chapitre précédent. On établit ici pourquoi selon la théorie on doit s'attendre à ce que l'usage de la méthode conduise à une réduction de variance pour certaines quantités (éléments diagonaux de $A_N^\star$, valeurs propres,...). Il s'agit d'un résultat qualitatif. Dans un second temps, nous traitons d'un exemple particulier : un cas perturbatif linéaire pour lequel la perturbation aléatoire d'un champs périodique sous-jacent est constante sur chaque cellule de périodicité $Q + k$ avec $k \in \mathbb{Z}^d$. Dans ce contexte on dispose de formules fermées pour les premiers ordres des développements du correcteur et de la matrice homogénéisée en la variable η qui correspond à l'amplitude de la perturbation. Ces formules font intervenir uniquement les variables aléatoires qui définissent le champ initial. Ces structures illustrent donc l'intuition selon laquelle réduire la variance du champ initial signifie également réduire la variance du champ homogénéisé.

Variance reduction in stochastic homogenization using antithetic variables

X. Blanc[a], Ronan Costaouec[b], Claude Le Bris[b] and Frédéric Legoll[b]

[a] *CEA, DAM, DIF, 91297 Arpajon, France*
[b] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

blanc@ann.jussieu.fr,Xavier.Blanc@cea.fr,
{costaour,lebris}@cermics.enpc.fr,
legoll@lami.enpc.fr.

5.1 Introduction

The present article examines some theoretical questions related to variance reduction techniques that can be successfully applied to some stochastic homogenization problems. It is a follow-up to an introductory article [P2] where some one-dimensional settings are considered theoretically and some two-dimensional numerical experiments are presented. In particular because of space limitation, the present contribution is supplemented by another publication [P4] presenting numerical results on a broad set of test cases.

The stochastic homogenization problem we consider here writes as follows. We consider the elliptic boundary value problem

$$\left\{ \begin{array}{rcl} -\mathrm{div}\left(A\left(\dfrac{x}{\varepsilon},\omega\right)\nabla u^\varepsilon\right) &=& f \quad \text{in} \quad \mathcal{D}, \\ u^\varepsilon &=& 0 \quad \text{on} \quad \partial\mathcal{D}, \end{array} \right. \tag{5.1}$$

set on a domain $\mathcal{D}$ in $\mathbb{R}^d$. Here, ε denotes a supposedly small, positive constant that models the smallest possible scale present in the problem. The matrix A is assumed random and stationary in a sense that will be made precise below. Somewhat loosely stated, A typically models a material that has a periodic pattern (with a basic unit cell Q) and for which, in each cell, some stationary random structure is present. For ε small, it is almost impossible, practically, to directly attack (5.1) with a numerical

discretization. A useful practical approach is to *first* transform (5.1) in the associated homogenized problem :

$$\left\{ \begin{array}{rcl} -\mathrm{div}\,(A^\star \nabla u^\star) & = & f \quad \text{in} \quad \mathcal{D}, \\ u^\star & = & 0 \quad \text{on} \quad \partial\mathcal{D}, \end{array} \right. \tag{5.2}$$

and *next* numerically solve the latter problem. The two-fold advantage of (5.2) as compared to (5.1) is that *it is deterministic* and *it does not involve the small scale* ε. This simplification comes at a price. The homogenized matrix $A^\star$ in (5.2) is given by an average of an integral involving the corrector function (a solution to an (random) auxiliary problem, reminiscent of (5.1), and set at the scale of the fine structure of the material). All this will be made precise below.

Now, practically computing the corrector function and the homogenized matrix $A^\star$ requires to generate several realizations of the material over a finite, supposedly large volume at the microscale, and approach the matrix by some empirical means. Although the *theoretical* value of $A^\star$ is deterministic (and this is the whole point and the definite success of homogenization theory to obtain this), it is because of the numerical approximation process itself that randomness again comes into the picture. Generating different configurations of the material and then efficiently averaging over these realizations require to understand how variance affects the result. This is the purpose of the present article to investigate some theoretical questions in this direction. Before proceeding and for the sake of consistency, we now present in more details some basic elements of stochastic homogenization, and situate the questions under consideration in a more general existing literature.

5.1.1 Homogenization theoretical setting

To begin with, we introduce the basic setting of stochastic homogenization we will employ. We refer to [40] for a general, numerically oriented presentation, and to [16, 27,58] for classical textbooks. We also refer to [19,20], [63] and the lecture notes [P5] for a presentation of our particular setting. Throughout this article, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and we denote by $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ the expectation value of any random variable $X \in L^1(\Omega, d\mathbb{P})$. We next fix $d \in \mathbb{N}^*$ (the ambient physical dimension), and assume that the group $(\mathbb{Z}^d, +)$ acts on Ω. We denote by $(\tau_k)_{k\in\mathbb{Z}^d}$ this action, and assume that it preserves the measure $\mathbb{P}$, that is, for all $k \in \mathbb{Z}^d$ and all $A \in \mathcal{F}$, $\mathbb{P}(\tau_k A) = \mathbb{P}(A)$. We assume that the action τ is *ergodic*, that is, if $A \in \mathcal{F}$ is such that $\tau_k A = A$ for any $k \in \mathbb{Z}^d$, then $\mathbb{P}(A) = 0$ or 1. In addition, we define the following notion of stationarity (see [19,20]) : any $F \in L^1_{\mathrm{loc}}\left(\mathbb{R}^d, L^1(\Omega)\right)$ is said to be *stationary* if, for all $k \in \mathbb{Z}^d$,

$$F(x + k, \omega) = F(x, \tau_k \omega), \tag{5.3}$$

almost everywhere in x and almost surely. In this setting, the ergodic theorem [61,89] can be stated as follows : *Let* $F \in L^\infty\left(\mathbb{R}^d, L^1(\Omega)\right)$ *be a stationary random variable*

in the above sense. For $k = (k_1, k_2, \ldots, k_d) \in \mathbb{Z}^d$, *we set* $|k|_\infty = \sup_{1 \le i \le d} |k_i|$. *Then*

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \le N} F(x, \tau_k \omega) \underset{N \to \infty}{\longrightarrow} \mathbb{E}\left(F(x, \cdot)\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

This implies that (denoting by Q *the unit cube in* $\mathbb{R}^d$*)*

$$F\left(\frac{x}{\varepsilon}, \omega\right) \overset{*}{\underset{\varepsilon \to 0}{\rightharpoonup}} \mathbb{E}\left(\int_Q F(x, \cdot) dx\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

Besides technicalities, the purpose of the above setting is simply to formalize that, even though realizations may vary, the function F at point $x \in \mathbb{R}^d$ and the function F at point $x+k$, $k \in \mathbb{Z}^d$, share the same law. In the homogenization context we now turn to, this means that the local, microscopic environment (encoded in the matrix A) is everywhere the same *on average*. From this, homogenized, macroscopic properties will follow. In addition, and this is evident reading the above setting, the microscopic environment considered has a relation to an underlying *periodic* structure (thus the integer shifts k in (5.3)).

As briefly introduced above, we now wish to solve the *multiscale random* elliptic problem (5.1). Let us formalize this. The domain $\mathcal{D}$ is an open, regular, bounded subset of $\mathbb{R}^d$. The right-hand side is an L^2 function f on $\mathcal{D}$. The random symmetric matrix A is assumed stationary in the sense (5.3). We also assume that A is bounded and that, in the sense of quadratic forms, A is positive and almost surely bounded away from zero.

In this specific setting, the homogenized matrix $A^\star$, that appears in the homogenized problem (5.2) obtained in the limit of small ε, reads

$$A^\star_{ij} = \int_Q \mathbb{E}\left[(\nabla w_{e_j}(x, \cdot) + e_j)^T \, A(x, \cdot) \, (\nabla w_{e_i}(x, \cdot) + e_i) \right] dx, \tag{5.4}$$

where, for any vector $p \in \mathbb{R}^d$, the *corrector* w_p is the solution (unique up to the addition of a random constant) in $\left\{ w \in L^2_{\text{loc}}(\mathbb{R}^d, L^2(\Omega)),\ \nabla w \in L^2_{\text{unif}}(\mathbb{R}^d, L^2(\Omega)) \right\}$ to

$$\begin{cases} -\text{div}\left[A(\nabla w_p + p)\right] = 0 \quad \text{on } \mathbb{R}^d \ \text{ a.s.}, \\ \nabla w_p \text{ is stationary in the sense of (5.3)}, \\ \displaystyle\int_Q \mathbb{E}(\nabla w_p) = 0. \end{cases} \tag{5.5}$$

We have used the notation L^2_{unif} for the *uniform* L^2 space, that is the space of functions for which, say, the L^2 norm on a ball of unit size is bounded above independently from the center of the ball.

5.1.2 The questions we consider

Now that the theory has been briefly presented, we turn to practice. The homogenized matrix $A^\star$ needs to be computed, so that in a second step the homogenized solution $u^\star$ may be approximated. By classical results in homogenization theory, we know $u^\star$ is a good approximation of u^ε, in a sense made precise in the literature (see *e.g.* [27]). In practice, the matrix $A^\star$ is approximated by the matrix

$$[A_N^\star]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} \left(e_i + \nabla w_{e_i}^N(y,\omega)\right)^T A(y,\omega)\left(e_j + \nabla w_{e_j}^N(y,\omega)\right) dy, \tag{5.6}$$

which is in turn obtained by solving the corrector problem on a *truncated* domain, say the cube $Q_N \subset \mathbb{R}^d$ of size $(2N+1)^d$ centered at the origin :

$$\begin{cases} -\mathrm{div}\left(A(\cdot,\omega)\left(p + \nabla w_p^N(\cdot,\omega)\right)\right) = 0, \\ w_p^N(\cdot,\omega) \text{ is } Q_N\text{-periodic.} \end{cases} \tag{5.7}$$

As briefly explained above, although $A^\star$ itself is a deterministic object, its practical approximation $A_N^\star$ is random. It is only in the limit of infinitely large domains Q_N that the deterministic value is attained. Our aim is to design a numerical technique that, for finite N, allows to compute $A_N^\star$ more effectively, that is, with a smaller variance.

The issue of variance in stochastic homogenization is not new. It has seen lately a definite revival, mainly motivated by numerical concerns. It is not our purpose here to review in details the important contributions existing in the literature. We however wish to cite some results particularly relevant to our own study :

- the original contribution [103] by Yurinski, where the convergence of some truncated approximation of $A^\star$ is established, along with an estimate of the rate of convergence (in short, problem (5.5) is regularized and then $A^\star$ is approximated on a bounded domain),
- a similar study [23] by Bourgeat and Piatnitsky for a specific approximation more relevant to actual numerical practice (in short, both problem (5.5) and the integral in (5.4) are truncated as in (5.6)–(5.7)),
- the work [76] by Naddaf and Spencer on a discrete ("lattice-type") approximation of the differential operator present in the original problem (5.1),
- and the enterprise by Gloria and Otto (see [49] for homogenization problems set on random lattices and publications announced in preparation for some problems for differential operators) to establish sharp estimates of the convergence of the numerical approximation in terms of the size of the truncation domain and other discretization parameters.

In all the above works, the convergence and the rate of convergence are studied. We take here the problem from a slightly different perspective : we are interested in *basic practical* issues. Can we improve the prefactor in the convergence of $A_N^\star$ to $A^\star$ as $N \to +\infty$ (loosely stated, the variance in a Central Limit Theorem type result) ?

or, even more practically, can we reduce the confidence interval for empirical means approximating $\mathbb{E}(A_N^\star)$? and similar issues.

To better understand the issue of reducing variance in stochastic homogenization, we consider a specific, well known variance reduction technique, the technique of *antithetic variables* [69, page 27]. In the sequel of this article, we consider two specific cases.

Our first setting (in Section 5.2) is a "genuinely" random setting (this terminology will be clear when we introduce our second setting shortly below). We consider a random matrix A constructed with independent, identically distributed random variables on the cells of our periodic lattice (although A does not need to be constant on each cell and equal to these random variables ; see *e.g.* example (5.13) below). Since solving problem (5.5) and directly computing $A^\star$ is out of reach practically, our numerical approach considers the truncations (5.6)–(5.7) on a finite domain Q_N, solved for a set of realizations of the random matrix. Empirical means of the truncated homogenized matrix $A_N^\star(\omega)$ are obtained, along with a (approximation of a) confidence interval involving the variance. The consideration of antithetic variables allows to improve the approximation. This is experimentally observed, and documented in [P2, P4]. We establish here theoretically that the variance of the homogenized objects is indeed diminished by the technique (we are, unfortunately, unable to explicitly estimate the gain). In our study, the matrix $A_N^\star$, its eigenvalues, its trace and determinant, and the eigenvalues of the elliptic operator associated to $A_N^\star$ are considered, but other objects could be studied : the eigenvectors, the differential operator itself, the approximation of the homogenized solution $u^\star$, the residual $u^\varepsilon - u^\star$ (somewhat as a follow-up to the studies [13, 23]), etc.

Our second setting is a "weakly" random setting. By this we mean that the random matrix A is a small perturbation of a deterministic, periodic matrix. Consequently, the solution of the problem is only seeked *at the first order in the size of the perturbation*. The setting has been introduced in [20] (and is recalled in Section 5.3.1 below). Its practical interest is that the computation comes down to a set of *fully deterministic* computations. So in practice, no variance issue is relevant. We however consider this case *pretending* not to exploit the simplification : we treat the problem stochastically and prove that the technique of variance reduction still works. As we can compute everything deterministically "in the backroom", the setting, although clearly particular and not general, is an appropriate test-bed to get some insight on some of the generically relevant issues.

It is important to note that some of the results we establish are limited to the technique of antithetic variables. Some others are not. They can therefore be useful for other variance reduction techniques. This is the case for our estimates of variance of the output of the computations in terms of the variance of the original parameters (see *e.g.* estimate (5.46)).

Let us conclude this introduction mentioning that of course there exists many other settings where similar questions can be asked. We treat here the very specific

case of a *linear*, *elliptic* second order equation *in divergence form*. The coefficient is assumed to be constructed with independent, identically distributed random variables set on a simple underlying periodic structure. The technique used for variance reduction is that of antithetic variables. Many more difficult situations could be addressed : other types of stationary ergodic coefficients, other types of equations, other techniques of truncations and regularizations of the original problems, other techniques for variance reduction, other numerical approaches, ...

Some of these issues (but clearly not all!) will be addressed in [P4].

5.2 A "fully" stochastic case

We consider in this section a "genuinely" random setting, in contrast to the setting of Section 5.3, where randomness will come as a *small perturbation* of a deterministic periodic setting.

5.2.1 Mathematical setting and statement of our main result

In this section, we make the following two assumptions on the matrix A of (5.1). First, we assume that, for any N, there exists an integer n (possibly $n = |Q_N|$, but not necessarily) and a function $\mathcal{A}$, defined on $Q_N \times \mathbb{R}^n$, such that the tensor $A(x,\omega)$ writes

$$\forall x \in Q_N, \quad A(x,\omega) = \mathcal{A}(x, X_1(\omega), \ldots, X_n(\omega)) \quad \text{a.s.}, \tag{5.8}$$

where $\{X_k(\omega)\}_{1\leq k\leq n}$ are independent scalar random variables, which are all distributed according to the uniform law $\mathcal{U}[0,1]$. In general, the function $\mathcal{A}$, as well as the number n of independent, identically distributed variables involved in (5.8), depend on N, the size of Q_N, although this dependency is not made explicit in (5.8).

Second, we assume that the function $\mathcal{A}$ in (5.8) is such that, for all $x \in Q_N$, and any vector $\xi \in \mathbb{R}^d$, the map

$$(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \xi^T \mathcal{A}(x, x_1, \ldots, x_n)\xi \tag{5.9}$$

is non-decreasing with respect to each of its arguments.

Before proceeding, we give a set of specific examples of matrices A that satisfy the above assumptions. Consider a family $(a_k(\omega))_{k\in\mathbb{Z}^d}$ of independent, identically distributed random variables, and set

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega)\,\mathrm{Id}, \tag{5.10}$$

where Q is the unit cube of $\mathbb{R}^d$, centered at the origin, and $Q+k$ is the cube Q translated by the vector $k \in \mathbb{Z}^d$. We assume that there exist $\alpha > 0$ and $\beta < \infty$ such

that, for all k, $0 < \alpha \leq a_k \leq \beta < +\infty$ almost surely. Consequently, A is uniformly coercive and bounded. Example (5.10) corresponds to a spherical matrix $A(x,\omega)$ that is constant in each cube $Q+k$, with independent, identically distributed values $a_k(\omega)$.

Introduce now the cumulative distribution function $P(x) = \nu(-\infty, x)$, where ν is the common probability measure of all the a_k, and next the non-decreasing function $f(x) = \inf\{y; P(x) \geq y\}$. Then, for any random variable $X(\omega)$ uniformy distributed in $[0,1]$, the random variable $f(X(\omega))$ is distributed according to the measure ν. As a consequence, we can recast (5.10) in the form

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) f(X_k(\omega))\,\mathrm{Id}, \tag{5.11}$$

where $(X_k(\omega))_{k\in\mathbb{Z}^d}$ is a family of independent random variables that are all uniformly distributed in $[0,1]$, and f is non-decreasing. This yields an example falling in our framework (5.8)–(5.9).

Remark 5.2.1 *Consider the example when, in* (5.10)*, the variables a_k are all distributed according to a Bernoulli law of parameter $r \in (0,1)$, that is, $a_0 \sim \mathcal{B}(r)$, $\mathbb{P}(a_0 = \alpha) = r$ and $\mathbb{P}(a_0 = \beta) = 1 - r$, for some $0 < \alpha < \beta$. In that case, the function f mentioned above reads*

$$f(x) = \alpha + (\beta - \alpha)\mathbf{1}_{\{r\leq x\leq 1\}},$$

and we may write $a_0(\omega) = f(X_0(\omega))$ with $X_0 \sim \mathcal{U}([0,1])$.

Of course, Example (5.10) can be readily extended to the case of non-spherical matrices. Consider a function F, defined on $[0,1]$, such that, for each $x \in [0,1]$, $F(x)$ is a symmetric matrix. We also assume that $F(x)$ is uniformly coercive and bounded, and that, for any $\xi \in \mathbb{R}^d$, the function $x \in [0,1] \mapsto \xi^T F(x)\xi$ is non-decreasing. Then

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) F(X_k(\omega)) \tag{5.12}$$

also satisfies our assumptions.

We eventually give an example of a matrix A satisfying our assumptions and that, on each cell, is not equal to independent, identically distributed variables. For this purpose, define positive coefficients κ_p for $|p|_\infty \leq 1$, and consider a non-decreasing function f. We then set

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \left[\mathbf{1}_{Q+k}(x) \sum_{j\in\mathbb{Z}^d,\ |j-k|_\infty\leq 1} \kappa_{j-k}\ f(X_j(\omega))\ \mathrm{Id} \right], \tag{5.13}$$

which clearly satisfies (5.8)–(5.9). In (5.13), it is immediately seen that the value of $A(x,\omega)$ in the cell $Q+k$ is a local average of the values $f(X_j(\omega))\,\mathrm{Id}$, for $|j-k|_\infty \leq 1$.

The main result of this section is the following :

Proposition 5.2.1 *We assume* (5.8)–(5.9). *Let* $A_N^\star(\omega)$ *be the approximated homogenized matrix, obtained by solving the corrector problem* (5.7) *on the truncated domain* $Q_N \subset \mathbb{R}^d$. *We define on* Q_N *the field*

$$B(x,\omega) := \mathcal{A}(x, 1 - X_1(\omega), \ldots, 1 - X_n(\omega)), \tag{5.14}$$

antithetic to $A(\cdot,\omega)$ *defined by* (5.8). *We associate to this field the corrector problem* (5.7) *(replacing* A *by* B*), the solution of which is denoted by* v_p^N, *and the matrix* $B_N^\star(\omega)$, *defined from* v_p^N *using* (5.6). *Set*

$$\widetilde{A}_N^\star(\omega) := \frac{1}{2}\left(A_N^\star(\omega) + B_N^\star(\omega)\right). \tag{5.15}$$

Then

$$\mathbb{E}\left(\widetilde{A}_N^\star\right) = \mathbb{E}\left(A_N^\star\right), \tag{5.16}$$

and, for any $\xi \in \mathbb{R}^d$,

$$\mathbb{V}ar\left(\xi^T \widetilde{A}_N^\star \xi\right) \leq \frac{1}{2}\mathbb{V}ar\left(\xi^T A_N^\star \xi\right). \tag{5.17}$$

Otherwise stated, $\widetilde{A}_N^\star$ *is a random variable which has the same expectation as* $A_N^\star$, *and its variance is smaller than half of that of* $A_N^\star$, *in the sense of* (5.17).

The practical consequences of this result for variance reduction techniques are discussed below.

Taking $\xi = e_i$ in (5.17) implies that the variance of the diagonal coefficient $\left[\widetilde{A}_N^\star\right]_{ii}$ is reduced by a factor two. Similar results can be obtained, first on the eigenvalues of the matrix $A_N^\star$, and second on the eigenvalues of the (approximate) homogenized elliptic operator $L_A = -\mathrm{div}\left[A_N^\star(\omega)\nabla\,\cdot\right]$. This is the purpose of the following two corollaries.

Corollary 5.2.1 *We assume* (5.8)–(5.9). *Denote by* $\{\lambda_k(A,\omega)\}_{1\leq k\leq d}$ *and* $\{\lambda_k(B,\omega)\}_{1\leq k\leq d}$ *the eigenvalues of* $A_N^\star(\omega)$ *and* $B_N^\star(\omega)$ *respectively, sorted in non-decreasing order, where* $A_N^\star(\omega)$ *and* $B_N^\star(\omega)$ *are defined as in Proposition 5.2.1. Define*

$$\widetilde{\lambda}_k(\omega) := \frac{1}{2}\big[\lambda_k(A,\omega) + \lambda_k(B,\omega)\big].$$

Then, for all $1 \leq k \leq d$,

$$\mathbb{E}\left(\widetilde{\lambda}_k\right) = \mathbb{E}\left(\lambda_k(A,\cdot)\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\lambda}_k\right) \leq \frac{1}{2}\mathbb{V}ar\left(\lambda_k(A,\cdot)\right). \tag{5.18}$$

Remark 5.2.2 *The above corollary shows variance reduction for each eigenvalue of the homogenized matrix. It is easily seen that the proof of this result carries over to the case when the quantity of interest is a function* $\mathcal{F}\left(\lambda_1(A,\omega), \ldots, \lambda_d(A,\omega)\right)$ *of*

these eigenvalues, provided $\mathcal{F}$ is a real-valued function that is non-decreasing with respect to each of its arguments. We indeed have

$$\mathbb{E}\left(\widetilde{Z}\right) = \mathbb{E}\left(Z\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{Z}\right) \leq \frac{1}{2}\mathbb{V}ar(Z)\,,$$

where $Z(\omega) = \mathcal{F}\left(\lambda_1(A,\omega),\ldots,\lambda_d(A,\omega)\right)$ *and*

$$\widetilde{Z}(\omega) = \frac{1}{2}\big[\mathcal{F}\left(\lambda_1(A,\omega),\ldots,\lambda_d(A,\omega)\right) + \mathcal{F}\left(\lambda_1(B,\omega),\ldots,\lambda_d(B,\omega)\right)\big].$$

Typical examples for such functions are

$$\begin{aligned}\mathcal{F}\left(\lambda_1(A,\omega),\ldots,\lambda_d(A,\omega)\right) &= \sum_{k=1}^{d}\lambda_k(A,\omega) = \operatorname{Tr} A_N^\star(\omega),\\ \mathcal{F}\left(\lambda_1(A,\omega),\ldots,\lambda_d(A,\omega)\right) &= \prod_{k=1}^{d}\lambda_k(A,\omega) = \det A_N^\star(\omega).\end{aligned}$$

Using the technique of antithetic variables, we hence achieve variance reduction for any diagonal coefficient of the matrix $A_N^\star(\omega)$, as well as for its trace and its determinant, and for any quantity of the form $\xi^T A_N^\star \xi$, for any given vector $\xi \in \mathbb{R}^d$. Note however that our argument does not cover the case of the off-diagonal coefficients of $A_N^\star(\omega)$, although numerical results on several test cases indicate that the variance of these coefficients is also reduced by the present method (see [P4]).

Corollary 5.2.2 *We assume* (5.8)–(5.9). *Let* $(\lambda_k(L_A,\omega), u_k(L_A,\omega))_{k\in\mathbb{N}}$ *be the eigenelements of the operator* $L_A = -\operatorname{div}\left[A_N^\star(\omega)\nabla\cdot\right]$ *with homogeneous Dirichlet boundary conditions, i.e.*

$$-\operatorname{div}\left[A_N^\star(\omega)\nabla u_k(L_A,\omega)\right] = \lambda_k(L_A,\omega)\,u_k(L_A,\omega)$$

with $u_k(L_A,\omega) \in H_0^1(\mathcal{D})$ *and* $\|u_k(L_A,\omega)\|_{L^2(\mathcal{D})} = 1$. *We similarly consider the eigenelements of* $L_B = -\operatorname{div}\left[B_N^\star(\omega)\nabla\cdot\right]$ *:*

$$-\operatorname{div}\left[B_N^\star(\omega)\nabla u_k(L_B,\omega)\right] = \lambda_k(L_B,\omega)\,u_k(L_B,\omega).$$

We assume that, almost surely, $\lambda_k(L_A,\omega)$ *and* $\lambda_k(L_B,\omega)$ *are sorted in non-decreasing order. Define*

$$\widetilde{\lambda}_k(L,\omega) := \frac{1}{2}\left(\lambda_k(L_A,\omega) + \lambda_k(L_B,\omega)\right).$$

Then, for all $k \in \mathbb{N}$,

$$\mathbb{E}\left(\widetilde{\lambda}_k(L,\cdot)\right) = \mathbb{E}\left(\lambda_k(L_A,\cdot)\right) \quad \textit{and} \quad \mathbb{V}ar\left(\widetilde{\lambda}_k(L,\cdot)\right) \leq \frac{1}{2}\mathbb{V}ar\left(\lambda_k(L_A,\cdot)\right). \tag{5.19}$$

The proofs of the above results (Proposition 5.2.1 and Corollaries 5.2.1 and 5.2.2) are given in Section 5.2.4. They are obtained combining some classical results on variance reduction using antithetic variables [69, page 27] and some monotonicity results from the theory of homogenization. For consistency, these results are recalled in Sections 5.2.2 and 5.2.3, respectively.

The proof goes as follows. First, we recall that the technique of antithetic variables reduces variance for the computation of $\mathbb{E}(f(X_1,\dots,X_n))$, when f is a real-valued function, that is *non-decreasing* of each of its argument, and $X = (X_1,\dots,X_n)$ is a vector of independent random variables. This is made precise in Section 5.2.2. Second, we assume that the tensor field $A(x,\omega)$ of (5.1) writes as a non-decreasing function (in the sense of symmetric positive matrices) of independent random variables $X_k(\omega)$ (these are assumptions (5.8) and (5.9)). Then, as recalled in Section 5.2.3, we use that the homogenization process *preserves the order of symmetric matrices* to claim that $A_N^\star(\omega)$ is likewise a non-decreasing function of the random variables $X_k(\omega)$. Consequently, we obtain variance reduction for $A_N^\star$. In the same vein, since the map that associates to a symmetric matrix its eigenvalues is non-decreasing, we obtain variance reduction for the eigenvalues of $A_N^\star$. This argument is formalized in Sections 5.2.2, 5.2.3 and 5.2.4.

Before proceeding, we briefly explain the usefulness of the above results for variance reduction techniques. Assume we want to compute the expectation of $\xi^T A_N^\star(\omega)\xi$, for some fixed vector $\xi \in \mathbb{R}^d$ (similar arguments hold for the computation of the expectation of any quantity considered above : the eigenvalues of the matrix $A_N^\star(\omega)$, its trace, its determinant, or the eigenvalues of the associated elliptic operator). Following a classical Monte-Carlo method, we estimate $\mathbb{E}\left(\xi^T A_N^\star \xi\right)$ by its empirical mean. To this end, we consider $2M$ independent, identically distributed copies $\{A_m(x,\omega)\}_{1\le m\le 2M}$ of the random field $A(x,\omega)$ on Q_N. To each copy A_m, we associate an approximate homogenized matrix $A_N^{\star,m}(\omega)$, obtained using the solution to the corrector problem (5.7) in the average (5.6). We next introduce the empirical mean

$$\mu_{2M}\left(\xi^T A_N^\star \xi\right)(\omega) = \frac{1}{2M}\sum_{m=1}^{2M} \xi^T A_N^{\star,m}(\omega)\xi, \tag{5.20}$$

and consider that, in practice, the mean $\mathbb{E}\left(\xi^T A_N^\star \xi\right)$ is equal to its estimator $\mu_{2M}\left(\xi^T A_N^\star \xi\right)$ within an approximate margin of error $1.96\dfrac{\sqrt{\mathbb{V}\mathrm{ar}\left(\xi^T A_N^\star \xi\right)}}{\sqrt{2M}}$.

Alternate to considering (5.20), we may consider

$$\mu_M\left(\xi^T \widetilde{A}_N^\star \xi\right)(\omega) = \frac{1}{M}\sum_{m=1}^{M} \xi^T \widetilde{A}_N^{\star,m}(\omega)\xi, \tag{5.21}$$

where $\widetilde{A}_N^{\star,m}$ is defined by (5.15). Again, in practice, the mean $\mathbb{E}\left(\xi^T A_N^\star \xi\right) = \mathbb{E}\left(\xi^T \widetilde{A}_N^\star \xi\right)$ is equal to $\mu_M\left(\xi^T \widetilde{A}_N^\star \xi\right)$ within an approximate margin of error

$1.96\frac{\sqrt{\mathbb{V}\mathrm{ar}\left(\xi^T \widetilde{A}^\star_N \xi\right)}}{\sqrt{M}}$. Observe now that both estimators (5.20) and (5.21) are of equal cost, since they require the same number $2M$ of corrector problems to be solved. The accuracy of the latter is better if and only if $\mathbb{V}\mathrm{ar}\left(\xi^T \widetilde{A}^\star_N \xi\right) \leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\xi^T A^\star_N \xi\right)$, which is exactly the bound (5.17) of Proposition 5.2.1.

We conclude this discussion by describing for illustration a typical numerical result (see [P2, P4] for more comprehensive numerical experiments). We consider the case (5.10), with $(a_k(\omega))_{k\in\mathbb{Z}^d}$ a family of independent random variables that are all distributed according to a Bernoulli law of parameter $r \in (0,1) : \mathbb{P}(a_0 = \alpha) = r$ and $\mathbb{P}(a_0 = \beta) = 1 - r$, for some $0 < \alpha < \beta$. As noted above, this case falls in our framework (5.8)–(5.9). Following (5.14), we introduce the antithetic field

$$B(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) b_k(\omega)\,\mathrm{Id},$$

where the variable b_k, antithetic to a_k, is defined as follows. We recall from Remark 5.2.1 that $a_k(\omega) = \alpha + (\beta-\alpha)\mathbf{1}_{\{r\leq X_k(\omega)\leq 1\}}$, with $X_k \sim \mathcal{U}([0,1])$. We then set $b_k(\omega) = \alpha + (\beta-\alpha)\mathbf{1}_{\{0\leq X_k(\omega)\leq 1-r\}}$.

Focusing, to fix the ideas, on the computation of $\mathbb{E}\left([A^\star_N]_{11}\right)$, we introduce the effectivity ratio

$$R = \frac{\mathbb{V}\mathrm{ar}\left([A^\star_N]_{11}\right)}{2\mathbb{V}\mathrm{ar}\left(\left[\widetilde{A}^\star_N\right]_{11}\right)},$$

where in practice the above variances are replaced by empirical variances. This ratio quantifies the gain in computational time (at fixed accuracy). Results are reported in Table 5.1, for Bernoulli variables of parameter $\alpha = 3$, $\beta = 20$ and $r = 1/2$ (numerical tests have been performed using the finite elements software FreeFem++, see `http ://www.freefem.org`). On Figure 5.1, we plot the estimated means (5.20) and (5.21) along with their confidence intervals. These results indeed show the efficiency of the approach.

N	5	10	20	40	60	80	100
R	5.34	3.91	5.41	3.07	4.41	4.49	4.28

TAB. 5.1 – Effectivity ratio R, in the case of a Bernoulli variable of parameter $\alpha = 3$, $\beta = 20$ and $r = 1/2$ (results extracted from [P2]).

Note that the above test case is a challenging one, as the ratio β/α is large. When the normalized variance of the field is smaller, even larger effectivity ratios R are obtained. Consider again the case (5.10), namely

$$A(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) a_k(\omega)\,\mathrm{Id},$$

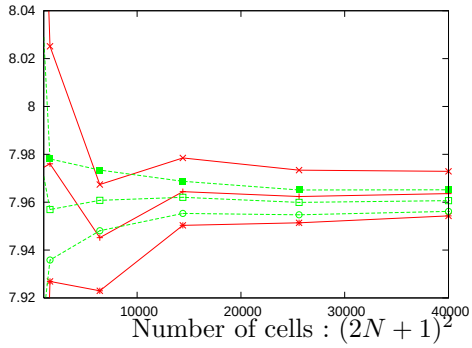


FIG. 5.1 – Estimated means (with confidence intervals) for $[A_N^\star]_{11}$ (red) and $\left[\widetilde{A}_N^\star\right]_{11}$ (green), in the case of a Bernoulli variable of parameter $\alpha = 3$, $\beta = 20$ and $r = 1/2$. Results are extracted from [P2] (in each case, $2M = 100$ corrector problems have been solved).

where now $(a_k(\omega))_{k\in\mathbb{Z}^d}$ is a family of independent random variables that are all uniformly distributed between $\alpha_0 = 3$ and $\beta_0 = 5$. In that case, we obtain effectivity ratios R of the order of 50, as shown in Table 5.2.

N	40	60	80	100
R	52.9	59.2	55.7	71.5

TAB. 5.2 – Effectivity ratio R, in the case of a random variable uniformly distributed between $\alpha_0 = 3$ and $\beta_0 = 5$ (results extracted from [P4]).

5.2.2 Classical results on antithetic variables

We first recall the following lemma, and provide its proof for consistency. This result is crucial for our proof of variance reduction using the technique of antithetic variables.

Lemma 5.2.1 ([69], page 27) *Let f and g be two real-valued functions defined on $\mathbb{R}^n$, which are non-decreasing with respect to each of their arguments. Consider $X = (X_1, \ldots, X_n)$ a vector of random variables, which are all independent from one another. Then*

$$\mathbb{C}ov(f(X), g(X)) \geq 0. \tag{5.22}$$

Proof : We prove the lemma by induction. Consider X and Y two independent scalar random variables, identically distributed. Both functions f and g are non-decreasing, so

$$(f(X) - f(Y))\ (g(X) - g(Y)) \geq 0.$$

We now take the expectation of the above inequality :

$$\mathbb{E}(f(X)\ g(X)) + \mathbb{E}(f(Y)\ g(Y)) \geq \mathbb{E}(f(Y)\ g(X)) + \mathbb{E}(f(X)\ g(Y)).$$

As X and Y share the same law, and are independent, this yields

$$\mathbb{E}(f(X)\ g(X)) \geq \mathbb{E}(f(X))\ \mathbb{E}(g(X)),$$

and (5.22) follows for $n = 1$.

Assume now that, for some N, we have proved the result for any random vector X of any dimension $n \leq N-1$. Let us now prove the result for a vector X of dimension N. For any fixed x_N, the functions $(x_1, \ldots, x_{N-1}) \mapsto f(x_1, \ldots, x_{N-1}, x_N)$ and $(x_1, \ldots, x_{N-1}) \mapsto g(x_1, \ldots, x_{N-1}, x_N)$ are non-decreasing with respect to each of their arguments. It follows from the induction assumption that, for any x_N,

$$\mathbb{E}(f(X_1, \ldots, X_{N-1}, x_N)\ g(X_1, \ldots, X_{N-1}, x_N)) \geq \mathbb{E}(f(X_1, \ldots, X_{N-1}, x_N))\ \mathbb{E}(g(X_1, \ldots, X_{N-1}, x_N)). \quad (5.23)$$

Introducing the conditional expectations

$$\begin{aligned} F(x) &= \mathbb{E}(f(X_1, \ldots, X_{N-1}, X_N)|X_N = x), \\ G(x) &= \mathbb{E}(g(X_1, \ldots, X_{N-1}, X_N)|X_N = x), \end{aligned}$$

inequality (5.23) reads

$$\mathbb{E}(f(X_1, \ldots, X_{N-1}, X_N)\ g(X_1, \ldots, X_{N-1}, X_N)|X_N = x_N) \geq F(x_N)\ G(x_N).$$

Integrating in x_N with respect to the law of X_N, we thus obtain

$$\mathbb{E}(f(X)\ g(X)) \geq \mathbb{E}(F(X_N)\ G(X_N)). \quad (5.24)$$

By construction, F and G are non-decreasing functions. The result for $n = 1$ applies, and yields

$$\mathbb{E}(F(X_N)\ G(X_N)) \geq \mathbb{E}(F(X_N))\ \mathbb{E}(G(X_N)) = \mathbb{E}(f(X))\ \mathbb{E}(g(X)). \quad (5.25)$$

Collecting (5.24) and (5.25), we conclude the proof. □

Remark 5.2.3 *The proof clearly shows that the result also holds if, for each variable, f and g are either both non-decreasing or both non-increasing.*

The following result is a simple consequence of the above lemma.

Corollary 5.2.3 *Let f be a function defined on $\mathbb{R}^n$, which is non-decreasing with respect to each of its arguments. Consider $X = (X_1, \ldots, X_n)$ a vector of random variables, which are all independent from one another, and distributed according to the uniform law $\mathcal{U}[0,1]$. Then*

$$\mathbb{V}ar\left(\frac{1}{2}\left(f(X) + f(1-X)\right)\right) \leq \frac{1}{2}\mathbb{V}ar\left(f(X)\right),$$

where we denote $1 - X = (1 - X_1, \ldots, 1 - X_n) \in \mathbb{R}^n$.

Proof : Choosing $g(x_1, \ldots, x_n) = -f(1-x_1, \ldots, 1-x_n)$ in Lemma 5.2.1, we obtain that

$$\mathbb{C}\mathrm{ov}(f(X), f(1-X)) = \mathbb{C}\mathrm{ov}(f(X_1, \ldots, X_n), f(1-X_1, \ldots, 1-X_n)) \leq 0.$$

We next observe that

$$\begin{aligned} \mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(f(X) + f(1-X)\right)\right) &= \frac{1}{2}\mathbb{V}\mathrm{ar}(f(X)) + \frac{1}{2}\mathbb{C}\mathrm{ov}\left(f(X), f(1-X)\right) \\ &\leq \frac{1}{2}\mathbb{V}\mathrm{ar}(f(X)), \end{aligned}$$

where we have used that $\mathbb{V}\mathrm{ar}(f(X)) = \mathbb{V}\mathrm{ar}(f(1-X))$. □

Remark 5.2.4 *In the proofs of Lemma 5.2.1 and Corollary 5.2.3, we have not used the fact that the random variables X_i are independent from one another. Lemma 5.2.1 and Corollary 5.2.3 actually hold without this assumption. However, this assumption will be needed in the sequel.*

5.2.3 Monotonicity in periodic homogenization

We first recall a classical result of periodic homogenization, which is useful in the sequel. We provide its proof for consistency.

Lemma 5.2.2 ([99], page 12) *Consider, for all x in $\mathbb{R}^d$, a symmetric matrix $A(x) \in \mathbb{R}^{d\times d}$. Assume that A is $\overline{Q}$-periodic, uniformly coercive and bounded, for $\overline{Q}$ a cube of $\mathbb{R}^d$. We denote by $A^\star$ the matrix obtained from A by homogenization of the operator $L_\varepsilon = -\mathrm{div}\left[A\left(\frac{x}{\varepsilon}\right)\nabla\cdot\right]$.*

We now consider a matrix $B(x)$ enjoying the same properties as $A(x)$, and such that

$$\forall \xi \in \mathbb{R}^d, \quad \xi^T B(x)\xi \geq \xi^T A(x)\xi \quad \text{a.e. on } \overline{Q}.$$

We correspondingly denote by $B^\star$ the matrix obtained from B by homogenization of L_ε, replacing A by B. Then

$$\forall \xi \in \mathbb{R}^d, \quad \xi^T B^\star \xi \geq \xi^T A^\star \xi.$$

Proof : For any $p \in \mathbb{R}^d$, let w_p be the corrector function associated to the matrix $A(x)$, i.e. the function defined on $\mathbb{R}^d$, $\overline{Q}$-periodic, and solution to

$$-\mathrm{div}\left(A\left(p + \nabla w_p\right)\right) = 0 \quad \text{on} \quad \mathbb{R}^d.$$

By definition, $w_p \in H^1_\square(\overline{Q})$ satisfies

$$\forall \theta \in H^1_\square(\overline{Q}), \quad \int_{\overline{Q}} \nabla\theta^T A(x)(p + \nabla w_p)\, dx = 0, \tag{5.26}$$

where $H^1_\square(\overline{Q})$ denotes the set of functions that belong to $H^1(\overline{Q})$ and are $\overline{Q}$-periodic. On the other hand, by definition of the homogenized matrix $A^\star$,

$$p^T A^\star p = \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla w_p)^T A(x)(p + \nabla w_p).$$

Similar assertions hold for the matrix B. Denoting by v_p the corrector function associated to that matrix, we have

$$\begin{aligned}
p^T B^\star p &= \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla v_p)^T B(x)(p + \nabla v_p) \\
&\geq \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla v_p)^T A(x)(p + \nabla v_p) \\
&= \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla w_p + \nabla(v_p - w_p))^T A(x)(p + \nabla w_p + \nabla(v_p - w_p)) \\
&= \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla w_p)^T A(x)(p + \nabla w_p) \\
&\quad + \frac{2}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla w_p)^T A(x) \nabla(v_p - w_p) \\
&\quad + \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (\nabla(v_p - w_p))^T A(x) \nabla(v_p - w_p).
\end{aligned}$$

Since $A(x)$ is coercive, the third term is non-negative. Using (5.26), we see that the second term vanishes. We are left with

$$p^T B^\star p \geq \frac{1}{|\overline{Q}|} \int_{\overline{Q}} (p + \nabla w_p)^T A(x)(p + \nabla w_p) = p^T A^\star p.$$

This concludes the proof. □

We now recall the following elementary result on the monotonicity of eigenvalues of symmetric matrices. Consider two symmetric matrices A and B of size $d \times d$, such that $\xi^T B \xi \geq \xi^T A \xi$ for any $\xi \in \mathbb{R}^d$. Then, for any $1 \leq k \leq d$, $\lambda_k(B) \geq \lambda_k(A)$, where $\lambda_k(A)$ and $\lambda_k(B)$ are the eigenvalues of A and B respectively, sorted in non-decreasing order.

This result can be readily extended to the eigenvalues of the corresponding elliptic operators, as stated in the following lemma :

Lemma 5.2.3 *Consider two symmetric matrices $A(x)$ and $B(x)$, of size $d \times d$, defined on a bounded domain $\mathcal{D} \subset \mathbb{R}^d$. We assume that these two matrices are uniformly coercive and bounded, and that*

$$\forall \xi \in \mathbb{R}^d, \quad \xi^T B(x) \xi \geq \xi^T A(x) \xi \quad \textit{a.e. on } \mathcal{D}. \tag{5.27}$$

Consider the eigenelements $(\lambda_k(L_A), u_k(L_A))_{k\in\mathbb{N}}$ *of the operator* $L_A = -\mathrm{div}\,[A(x)\nabla\cdot]$ *with homogeneous Dirichlet boundary conditions, i.e.*

$$-\mathrm{div}\,[A(x)\nabla u_k(L_A)] = \lambda_k(L_A)u_k(L_A),$$

with $u_k(L_A) \in H^1_0(\mathcal{D})$ *and* $\|u_k(L_A)\|_{L^2(\mathcal{D})} = 1$. *We assume that* $\lambda_k(L_A)$ *are sorted in non-decreasing order. Let* $(\lambda_k(L_B), u_k(L_B))_{k\in\mathbb{N}}$ *be the eigenelements of* L_B. *Then*

$$\forall k \in \mathbb{N},\ \lambda_k(L_B) \geq \lambda_k(L_A). \tag{5.28}$$

This result is an immediate consequence of the Courant-Fisher characterization of the eigenvalues.

5.2.4 Proof of Proposition 5.2.1 and Corollaries 5.2.1 and 5.2.2

Now that we have collected all the necessary ingredients, we may turn here to the proof of our main results.

Proof of Proposition 5.2.1 : As $1 - X_k(\omega)$ and $X_k(\omega)$ share the same law, so do the fields $A(x,\omega)$ and $B(x,\omega)$, on Q_N. Hence, the homogenized matrices $A^\star_N(\omega)$ and $B^\star_N(\omega)$ have the same law, and we obtain (5.16).

We now prove (5.17). Let $\mathcal{P}_N$ be the operator that associates to a given Q_N-periodic tensor field A the effective tensor $\mathcal{P}_N(A)$ obtained by periodic homogenization. Then, by construction, the approximation $A^\star_N(\omega)$ defined by (5.6)–(5.7) is the effective matrix obtained by periodic homogenization of $A_{|Q_N}(\cdot,\omega)$:

$$A^\star_N(\omega) = \mathcal{P}_N\left[A_{|Q_N}(\cdot,\omega)\right] \quad \text{almost surely.}$$

In the framework of our Assumption (5.8), we have

$$\forall x \in Q_N,\ A(x,\omega) = \mathcal{A}(x, X_1(\omega), \ldots, X_n(\omega)),$$

where $\{X_k(\omega)\}_{1\leq k\leq n}$ are independent scalar random variables that are all uniformly distributed in $[0,1]$. Setting $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, we have

$$A^\star_N(\omega) = \mathcal{P}_N\left[\mathcal{A}(\cdot, X(\omega))\right] \in \mathbb{R}^{d\times d}. \tag{5.29}$$

We now choose a vector $\xi \in \mathbb{R}^d$, and introduce the function

$$\begin{array}{rcl} f : \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \xi^T \mathcal{P}_N\left[\mathcal{A}(\cdot, x)\right]\xi. \end{array}$$

By construction, we have

$$f(X(\omega)) = \xi^T A^\star_N(\omega)\xi \tag{5.30}$$

and, using the definition (5.15) of $\widetilde{A}^\star_N(\omega)$, we have

$$\frac{1}{2}\left(f(X(\omega)) + f(1-X(\omega))\right) = \frac{1}{2}\xi^T\left(A^\star_N(\omega) + B^\star_N(\omega)\right)\xi = \xi^T\widetilde{A}^\star_N(\omega)\xi. \tag{5.31}$$

We now show that f is non-decreasing with respect to each of its arguments. We infer from Assumption (5.9) that, for any $\eta \in \mathbb{R}^d$ and any $y \in Q_N$, the function $x \in \mathbb{R}^n \mapsto \eta^T\mathcal{A}(y,x)\eta$ is non-decreasing with respect to each of its arguments. In view of Lemma 5.2.2, we thus obtain that, for any $\eta \in \mathbb{R}^d$, the function $x \in \mathbb{R}^n \mapsto \eta^T\mathcal{P}_N\left(\mathcal{A}(\cdot,x)\right)\eta$ is non-decreasing with respect to each of its arguments. As a consequence, f is non-decreasing.

We are now in position to use Corollary 5.2.3, which yields

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(f(X) + f(1-X)\right)\right) \leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(f(X)\right).$$

Using (5.31) and (5.30), we obtain

$$\mathbb{V}\mathrm{ar}\left(\xi^T\widetilde{A}^\star_N\xi\right) = \mathbb{V}\mathrm{ar}\left[\frac{1}{2}\left(f(X) + f(1-X)\right)\right] \leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(f(X)\right) = \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\xi^T A^\star_N\xi\right),$$

which concludes the proof of (5.17). □

Proof of Corollary 5.2.1 : As shown in the proof of Proposition 5.2.1, $A^\star_N(\omega)$ and $B^\star_N(\omega)$ share the same law, so their eigenvalues also share the same law, and we thus obtain the first assertion in (5.18).

We now prove the second assertion. As in the proof of Proposition 5.2.1, we make use of the operator $\mathcal{P}_N$ that associates to any Q_N-periodic tensor field the effective tensor obtained by periodic homogenization. Expression (5.29) holds.

We next fix some k, $1 \leq k \leq d$, and consider the function

$$\begin{array}{rcl} \Lambda : \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & k\text{-th eigenvalue of the matrix } \mathcal{P}_N\left[\mathcal{A}(\cdot,x)\right]. \end{array}$$

We clearly have

$$\lambda_k(A,\omega) = \Lambda(X(\omega)) \quad \text{and} \quad \lambda_k(B,\omega) = \Lambda(1-X(\omega)), \tag{5.32}$$

where $\lambda_k(A,\omega)$ and $\lambda_k(B,\omega)$ denote the k-th eigenvalue (sorted increasingly) of $A^\star_N(\omega)$ and $B^\star_N(\omega)$, respectively.

The function Λ is non-decreasing with respect to each of its arguments. It is an immediate consequence of the proof of Proposition 5.2.1 and of the monotonicity of eigenvalues of symmetric matrices recalled in Section 5.2.3.

Corollary 5.2.3 then yields

$$\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\Lambda(X)+\Lambda(1-X)\right)\right) \leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\Lambda(X)\right).$$

By definition of $\widetilde{\lambda}_k$, and using (5.32), we obtain

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(\widetilde{\lambda}_k\right) &= \mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\lambda_k(A,\cdot)+\lambda_k(B,\cdot)\right)\right) \\
&= \mathbb{V}\mathrm{ar}\left(\frac{1}{2}\left(\Lambda(X)+\Lambda(1-X)\right)\right) \\
&\leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\Lambda(X)\right) \\
&\leq \frac{1}{2}\mathbb{V}\mathrm{ar}\left(\lambda_k(A,\cdot)\right),
\end{aligned}$$

which concludes the proof of (5.18). □

The proof of Corollary 5.2.2 follows the same pattern as that of Corollary 5.2.1.

5.3 A "weakly" stochastic case

In this section, we assume that the matrix A in (5.1) is a *perturbation* of a periodic matrix :

$$A(x,\omega) = A_{\mathrm{per}}(x) + \eta A_1(x,\omega), \tag{5.33}$$

where A_{per} is Q-periodic and δ-Hölder continuous (for some $0 < \delta < 1$), A_1 is stationary in the sense (5.3), and η is a deterministic, supposedly small, parameter. We assume that A, A_{per} and A_1 are all symmetric and bounded matrices, and that A and A_{per} are uniformly coercive. We also assume the following special structure for A_1 :

$$A_1(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x)\, X_k(\omega)\ \mathrm{Id}, \tag{5.34}$$

where Id denotes the identity matrix and $(X_k)_{k\in\mathbb{Z}^d}$ is a sequence of (scalar) independent, identically distributed random variables, satisfying, for some $0 < \alpha < \infty$, the bound $|X_k| \leq \alpha$ almost surely and for all k. Some extensions of the above setting will be considered in Section 5.3.4 below.

Since, in view of (5.33), A is a perturbation of A_{per}, we may expand the homogenized matrix $A^\star$ in powers of η, as shown in [20]. Let us first recall this expansion, at first order.

5.3.1 Main result

In [20], an expansion of the homogenized matrix (and of all the relevant quantities, such as the corrector function) in terms of a series in powers of η is shown to exist. The setting in [20] is slightly different from the present setting. It is however straightforward to check that our arguments carry over to the present case.

First, the corrector ∇w_p, solution to (5.5), is easily proved to have an expansion in powers of η :

$$\nabla w_p = \nabla w_p^0 + \eta \nabla w_p^1 + \dots, \tag{5.35}$$

where w_p^0 is the unique (up to the addition of a constant) solution to

$$\begin{cases} -\mathrm{div}\left[A_{\mathrm{per}}(\nabla w_p^0 + p)\right] = 0, \\ w_p^0 \text{ is } Q\text{-periodic}, \end{cases} \tag{5.36}$$

and w_p^1 is the unique (up to the addition of a random constant) solution to

$$\begin{cases} -\mathrm{div}\left[A_{\mathrm{per}} \nabla w_p^1\right] = \mathrm{div}\left[A_1\left(\nabla w_p^0 + p\right)\right] \quad \text{a.s. on } \mathbb{R}^d, \\ \nabla w_p^1 \text{ is stationary}, \\ \displaystyle\int_Q \mathbb{E}(\nabla w_p^1) = 0. \end{cases} \tag{5.37}$$

Second, the homogenized matrix (5.4) satisfies

$$A^\star = A^\star_{\mathrm{per}} + \eta A_1^\star + O(\eta^2), \tag{5.38}$$

where

$$\left[A^\star_{\mathrm{per}}\right]_{ij} = \int_Q (\nabla w^0_{e_i} + e_i)^T A_{\mathrm{per}} (\nabla w^0_{e_j} + e_j), \tag{5.39}$$

and

$$\begin{aligned} [A_1^\star]_{ij} \;=\; & \int_Q \mathbb{E}(\nabla w^1_{e_i})^T A_{\mathrm{per}} (\nabla w^0_{e_j} + e_j) + \int_Q (\nabla w^0_{e_i} + e_i)^T A_{\mathrm{per}} \mathbb{E}(\nabla w^1_{e_j}) \\ & + \int_Q (\nabla w^0_{e_i} + e_i)^T \mathbb{E}(A_1)(\nabla w^0_{e_j} + e_j). \end{aligned} \tag{5.40}$$

In essence, this specific setting does not give rise to any variance concerns, for two reasons at least. First, as observed in [20], and as evident on (5.40), the knowledge of $\overline{w}_p^1 := \mathbb{E}(\nabla w_p^1)$ is actually sufficient to compute $A_1^\star$. Taking expectations in (5.37), we indeed see that $\overline{w}_p^1$ is solution to the cell problem

$$\begin{cases} -\mathrm{div}\left[A_{\mathrm{per}} \nabla \overline{w}_p^1\right] = \mathrm{div}\left[\mathbb{E}(A_1)\left(\nabla w_p^0 + p\right)\right] \quad \text{on } \mathbb{R}^d, \\ \overline{w}_p^1 \text{ is } Q\text{-periodic}, \end{cases} \tag{5.41}$$

which is much easier to solve than the corrector problem (5.5) or its truncated version (5.7), since it is a *deterministic* problem set on a *single* cell. Hence, as pointed out in the introduction, the determination of $A^\star$ comes down, in practice, to solving the deterministic problems (5.36) and (5.41).

Second, in practice, the *exact* corrector w_p, which solves (5.5), is approximated by the solution w_p^N to the truncated problem (5.7), which also has an expansion in powers of η, as shown in [P1, P6] : similarly to (5.35), we have

$$\nabla w_p^N = \nabla w_p^0 + \eta \nabla w_p^{1,N} + \dots,$$

where $w_p^{1,N}$ solves the *truncated* problem

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A_{\mathrm{per}} \nabla w_p^{1,N}\right] = \mathrm{div}\left[A_1\left(\nabla w_p^0 + p\right)\right], \\ w_p^{1,N} \text{ is } Q_N - \text{periodic.} \end{array} \right. \tag{5.42}$$

In turn, the approximate homogenized matrix (5.6) satisfies

$$A_N^\star(\omega) = A_{\mathrm{per}}^\star + \eta A_{1,N}^\star(\omega) + O(\eta^2),$$

where

$$\begin{aligned} \left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{1}{|Q_N|}\Bigg[& \int_{Q_N} \nabla w_{e_i}^{1,N}(x,\omega)^T A_{\mathrm{per}}(x)(\nabla w_{e_j}^0(x) + e_j)dx \\ & + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x,\omega)(\nabla w_{e_j}^0(x) + e_j)dx \\ & + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_{\mathrm{per}}(x) \nabla w_{e_j}^{1,N}(x,\omega)dx \Bigg]. \end{aligned} \tag{5.43}$$

Observe now that the first term of (5.43) reads

$$\begin{aligned} \int_{Q_N} \left(\nabla w_{e_i}^{1,N}\right)^T A_{\mathrm{per}} \left(\nabla w_{e_j}^0 + e_j\right) &= -\int_{Q_N} w_{e_i}^{1,N} \mathrm{div}\left[A_{\mathrm{per}}(\nabla w_{e_j}^0 + e_j)\right] \\ &+ \int_{\partial Q_N} w_{e_i}^{1,N}\left[A_{\mathrm{per}}(\nabla w_{e_j}^0 + e_j)\right] \cdot n. \end{aligned}$$

Using (5.36), we see that the first term of the right hand side vanishes. Since $w_{e_i}^{1,N}$, A_{per} and $w_{e_j}^0$ are Q_N-periodic, the second term vanishes as well. Hence (5.43) reads

$$\left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x,\omega)(\nabla w_{e_j}^0(x) + e_j)dx. \tag{5.44}$$

In the specific case (5.34), we thus have

$$\left[A_{1,N}^\star\right]_{ij}(\omega) = \frac{\gamma_{ij}}{|Q_N|}\sum_{|k|_\infty \le N} X_k(\omega),$$

with $\gamma_{ij} = \int_Q \left(\nabla w^0_{e_i} + e_i\right)^T \left(\nabla w^0_{e_j} + e_j\right)$. The variables X_k being independent and identically distributed, we obtain that $\left[A^\star_{1,N}\right]_{ij}$ converges almost surely to $\gamma_{ij}\mathbb{E}(X_0)$, the rate of convergence being given by the central limit theorem. Of course, using (5.44), this argument is not restricted to the form (5.34) of A_1, and can be extended to more general cases.

Observe yet that the fact that the first and third terms of (5.43) vanish strongly relies on the specific equation (and boundary conditions) used to build a numerical approximation of the corrector, and may not be expected in a more general setting.

As announced in the introduction, with a view to use (5.33)-(5.34) to test and further understand our variance reduction approach, we *pretend* in this section not to exploit the various simplifications, and we thus treat the problem entirely stochastically. We make use of this opportunity to derive a series of technical lemmas (see Lemmas 5.3.1 and 5.3.2) that we believe might be useful for a similar study in a more general setting.

Our aim is to show that, for any fixed N, applying the variance reduction strategy described in Section 5.2.1, we obtain a better estimate of the approximate homogenized matrix $\mathbb{E}(A^\star_N)$ *using empirical means*, in the spirit of (5.20) and (5.21). When the number of independent realizations M increases to $+\infty$, the rates at which the empirical means (5.20) and (5.21) converge to the expectation are identical, but the prefactor is better in the latter case. In addition, using the specificities of this setting, we are also able to analyze the convergence of the approximation procedure when N goes to $+\infty$.

The main result of this section is the following proposition.

Proposition 5.3.1 *Let A be defined by* (5.33)*, where $A_{\rm per}$ is periodic, Hölder continuous, and A_1 satisfies* (5.34)*, with $(X_k)_{k\in\mathbb{Z}^d}$ a sequence of independent, identically distributed scalar random variables. Assume in addition that A, $A_{\rm per}$ and A_1 are symmetric and bounded matrices, and that A and $A_{\rm per}$ are uniformly coercive. For any $N \in \mathbb{N}$ and $1 \le i, j \le d$, define*

$$\begin{aligned}\left[A^{\star,\rm exact}_{1,N}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\Bigg[&\int_{Q_N} \nabla w^1_{e_i}(x,\omega)^T A_{\rm per}(x)(\nabla w^0_{e_j}(x) + e_j)dx \\ &+ \int_{Q_N} (\nabla w^0_{e_i}(x) + e_i)^T A_1(x,\omega)(\nabla w^0_{e_j}(x) + e_j)dx \\ &+ \int_{Q_N} (\nabla w^0_{e_i}(x) + e_i)^T A_{\rm per}(x)\nabla w^1_{e_j}(x,\omega)dx\Bigg], \quad (5.45)\end{aligned}$$

where w^0 and w^1 are solution to (5.36) *and* (5.37) *respectively, and $Q_N = \cup_{|k|_\infty \le N}(Q + k)$ is the cube of size $(2N+1)^d$ centered at the origin. Then, there exist d^2 coefficients $C^{ij}_N > 0$, independent of $(X_k)_{k\in\mathbb{Z}^d}$, such that*

$$\mathbb{V}ar\left(\left[A^{\star,\rm exact}_{1,N}\right]_{ij}\right) = C^{ij}_N\ \mathbb{V}ar(X_0). \quad (5.46)$$

In addition, we have

$$C_N^{ij} \leq \frac{C}{|Q_N|}, \tag{5.47}$$

where C does not depend on i, j, and N, and only depends on $A_{\rm per}$.

The estimate (5.46) above shows that reducing the variance of X_0 (for instance using the technique described in Proposition 5.2.1) reduces the variance on $A_{1,N}^{\star,\rm exact}$. This also gives a quantitative estimate of the variance reduction.

Note that, in the above Proposition, we have used the first order term w_p^1 in the expansion of the *exact* corrector. If this term is replaced by the first order term $w_p^{1,N}$ of the *approximate* corrector (solution to a truncated problem), then we recover $A_{1,N}^\star$ defined by (5.43).

Corollary 5.3.1 *Under the assumptions of Proposition 5.3.1, we have that*

$$\lim_{N\to\infty} A_{1,N}^{\star,\rm exact} = A_1^\star \quad a.s.,$$

where $A_1^\star$ is defined by (5.40)*. Assume in addition that*

$$\lim_{N\to\infty} |Q_N|^2 \, \mathbb{V}ar\left(\left[A_{1,N}^{\star,\rm exact}\right]_{ij}\right) = +\infty. \tag{5.48}$$

Then, $\dfrac{\left[A_{1,N}^{\star,\rm exact}\right]_{ij} - [A_1^\star]_{ij}}{\sqrt{\mathbb{V}ar\left(\left[A_{1,N}^{\star,\rm exact}\right]_{ij}\right)}}$ is a random variable that converges in law to a normal Gaussian random variable.

The proof of Proposition 5.3.1 (and of Corollary 5.3.1) is the purpose of Section 5.3.3 below. It makes use of some preliminary results established in Section 5.3.2. Section 5.3.4 presents some extensions.

5.3.2 Decomposition of ∇w_p^1

In this section, we study the first order term in the expansion (5.35) of the corrector function. Our purpose is to prove the following two lemmas.

Lemma 5.3.1 *Let $p \in \mathbb{R}^d$. The problem*

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A_{\rm per}\nabla\phi_p\right] = \mathrm{div}\left[\mathbf{1}_Q\left(\nabla w_p^0 + p\right)\right], \\ \phi_p \in L^2_{\rm loc}(\mathbb{R}^d), \quad \nabla\phi_p \in \left(L^2(\mathbb{R}^d)\right)^d, \end{array} \right. \tag{5.49}$$

has a solution, which is unique up to the addition of a constant. Moreover, there exists a constant $C > 0$ such that

$$\forall x \in \mathbb{R}^d \text{ with } |x| \geq 1, \quad |\nabla\phi_p(x)| \leq \frac{C}{|x|^d}, \tag{5.50}$$

and

$$\forall x \in \mathbb{R}^d, \quad |\phi_p(x)| \leq \frac{C}{1+|x|^{d-1}}. \tag{5.51}$$

Lemma 5.3.2 *Let $p \in \mathbb{R}^d$, and let w_p^1 be defined by (5.37). Let $\overline{w}_p^1$ be the unique solution (up to the addition of a constant) to*

$$\begin{cases} -\mathrm{div}\left[A_{\mathrm{per}}\nabla\overline{w}_p^1\right] = \mathrm{div}\left[\mathbb{E}(A_1)\left(\nabla w_p^0 + p\right)\right], \\ \overline{w}_p^1 \text{ is } Q-\text{periodic.} \end{cases} \tag{5.52}$$

Then, we have

$$\nabla w_p^1(x,\omega) = \nabla\overline{w}_p^1(x) + \sum_{k\in\mathbb{Z}^d} \nabla\phi_p(x-k)\left(X_k(\omega) - \mathbb{E}(X_k)\right), \tag{5.53}$$

and the sum in (5.53) is a convergent series in $L^2(Q\times\Omega)$.

Proof of Lemma 5.3.1 : We first prove the existence of a solution to (5.49). The argument is standard and we provide it here for consistency. To begin with, we define a regularized version of the equation : given $\delta > 0$, we consider the problem

$$\begin{cases} -\mathrm{div}\left[A_{\mathrm{per}}\nabla\phi_{p,\delta}\right] + \delta\phi_{p,\delta} = \mathrm{div}\left[\mathbf{1}_Q\left(\nabla w_p^0 + p\right)\right], \\ \phi_{p,\delta} \in H^1(\mathbb{R}^d). \end{cases} \tag{5.54}$$

Applying the Lax-Milgram lemma, it is clear that (5.54) has a unique solution in $H^1(\mathbb{R}^d)$. Next, multiplying the equation by $\phi_{p,\delta}$ and integrating over $\mathbb{R}^d$, we find

$$\int_{\mathbb{R}^d} \nabla\phi_{p,\delta}^T A_{\mathrm{per}}\nabla\phi_{p,\delta} + \delta\int_{\mathbb{R}^d}\phi_{p,\delta}^2 = -\int_Q \nabla\phi_{p,\delta}^T\left(\nabla w_p^0 + p\right),$$

hence, using Cauchy-Schwarz inequality and elementary calculus,

$$\int_{\mathbb{R}^d}|\nabla\phi_{p,\delta}|^2 \leq C \quad \text{and} \quad \int_{\mathbb{R}^d}\phi_{p,\delta}^2 \leq \frac{C}{\delta} \tag{5.55}$$

for some constant $C > 0$ independent of δ. One can thus define a sequence $\delta_n \to 0$ such that

$$\nabla\phi_{p,\delta_n} \rightharpoonup T \in \left(L^2(\mathbb{R}^d)\right)^d$$

as $n \to \infty$, weakly in $\left(L^2(\mathbb{R}^d)\right)^d$. Hence, the equality $\partial_i\partial_j\phi_{p,\delta_n} = \partial_j\partial_i\phi_{p,\delta_n}$ passes to the limit in the sense of distributions and implies $\partial_i T_j = \partial_j T_i$. This implies that $T = \nabla\phi_p$ for some $\phi_p \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$, with $\nabla\phi_p \in \left(L^2(\mathbb{R}^d)\right)^d$. Moreover, for any $\xi \in \mathcal{D}(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d}\nabla\phi_{p,\delta_n}^T A_{\mathrm{per}}\nabla\xi + \delta_n\int_{\mathbb{R}^d}\phi_{p,\delta_n}\xi = -\int_Q \nabla\xi^T\left(\nabla w_p^0 + p\right).$$

Passing to the limit $n \to \infty$ in this equation, we find that ϕ_p is a solution to (5.49), in the appropriate functional spaces. This concludes the proof of existence.

Proving uniqueness (up to the addition of a constant) of the solution to (5.49) amounts to proving that

$$\begin{cases} -\mathrm{div}\,[A_{\mathrm{per}}\nabla\phi] = 0, \\ \phi \in L^2_{\mathrm{loc}}(\mathbb{R}^d), \quad \nabla\phi \in \left(L^2(\mathbb{R}^d)\right)^d, \end{cases} \tag{5.56}$$

implies $\nabla\phi \equiv 0$. For this purpose, we consider $\chi \in C^\infty(\mathbb{R}^d)$ such that $\chi = 1$ in B_R, $\chi = 0$ in B^c_{R+1}, $\chi \geq 0$ on $\mathbb{R}^d$ and $|\nabla\chi| \leq 1$. Multiplying (5.56) by $\chi\phi$, we find that

$$\int_{\mathbb{R}^d} \left(\nabla\phi^T A_{\mathrm{per}}\nabla\phi\right)\chi + \int_{\mathbb{R}^d} \left(\nabla\chi^T A_{\mathrm{per}}\nabla\phi\right)\phi = 0,$$

hence

$$m_{\mathrm{per}} \int_{B_R} |\nabla\phi|^2 \leq M_{\mathrm{per}} \int_{B_{R+1}\setminus B_R} |\nabla\phi|\ |\phi|,$$

where $M_{\mathrm{per}} = \|A_{\mathrm{per}}\|_{L^\infty}$ and $m_{\mathrm{per}} > 0$ is the coercivity constant of A_{per}. Next, we point out that the above computations are also valid if we replace ϕ by $\phi - \phi_R$, where ϕ_R is the constant

$$\phi_R = \frac{1}{|B_{R+1} \setminus B_R|} \int_{B_{R+1}\setminus B_R} \phi.$$

Hence, using the Cauchy-Schwarz inequality,

$$\int_{B_R} |\nabla\phi|^2 \leq \frac{M_{\mathrm{per}}}{m_{\mathrm{per}}} \left(\int_{B_{R+1}\setminus B_R} |\nabla\phi|^2\right)^{1/2} \left(\int_{B_{R+1}\setminus B_R} (\phi - \phi_R)^2\right)^{1/2}.$$

We next make use of the Poincaré-Wirtinger inequality [45, page 164] on the bounded domain $B_{R+1} \setminus B_R$:

$$\|\phi - \phi_R\|_{L^2(B_{R+1}\setminus B_R)} \leq C\|\nabla\phi\|_{L^2(B_{R+1}\setminus B_R)},$$

for some constant C independent of R and ϕ. We thus obtain

$$\int_{B_R} |\nabla\phi|^2 \leq C\,\frac{M_{\mathrm{per}}}{m_{\mathrm{per}}} \int_{B_{R+1}\setminus B_R} |\nabla\phi|^2.$$

As $\nabla\phi \in \left(L^2(\mathbb{R}^d)\right)^d$, we have that $\lim\limits_{R\to\infty} \int_{B_{R+1}\setminus B_R} |\nabla\phi|^2 = 0$. The above bound thus yields $\lim\limits_{R\to\infty} \int_{B_R} |\nabla\phi|^2 = 0$. As a consequence, $\int_{\mathbb{R}^d} |\nabla\phi|^2 = 0$, which implies $\nabla\phi \equiv 0$. This concludes the proof of uniqueness.

It now remains to prove (5.50) and (5.51). Let G be the Green function associated with the operator $-\mathrm{div}\,[A_{\mathrm{per}}\nabla\cdot]$, that is, the solution of

$$-\mathrm{div}_x\left(A_{\mathrm{per}}(x)\nabla_x G(x,y)\right)=\delta_y(x), \tag{5.57}$$

with $G(\cdot,y)\in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p<d/(d-1)$ and $\nabla_x G(\cdot,y)\in L^p(\mathbb{R}^d\setminus B_r(y))$, for all $p>d/(d-1)$ and $r>0$. Such a solution exists and is unique according to [52, Theorem 1.1] (actually, the results of [52] are cited only for $d\geq 3$, but the existence proof there carries through to the case $d=2$. The uniqueness may be proved by standard arguments, as pointed out in [6]).

We then have

$$\begin{aligned}\phi_p(x) &= \int_{\mathbb{R}^d} G(x,y)\mathrm{div}_y\left[\mathbf{1}_Q(y)\left(\nabla w_p^0(y)+p\right)\right]dy\\ &= -\int_Q \nabla_y G(x,y)\left(\nabla w_p^0(y)+p\right)dy.\end{aligned} \tag{5.58}$$

We now recall that, as A_{per} is Hölder continuous, the solution w_p^0 to (5.36) satisfies $w_p^0\in W^{1,\infty}(Q)$ (see [52, Lemma 3.1]). We also recall that, if $d\geq 3$, according to [52, Theorem 3.3] (see also [15, Sec. 2] and [10, Theorem 13]), G satisfies the estimate :

$$\forall |x-y|\geq 1,\quad |G(x,y)|\leq C|x-y|^{2-d}.$$

In addition (see [10, Theorem 13]), we have, in the case $d=2$,

$$\forall |x-y|\geq 1,\quad |G(x,y)|\leq C\left(1+\log|x-y|\right).$$

We also recall (see [6]) that, for any $d\geq 2$,

$$\forall |x-y|\geq 1,\quad |\nabla_x G(x,y)|+|\nabla_y G(x,y)|\leq C|x-y|^{1-d} \tag{5.59}$$

and

$$\forall |x-y|\geq 1,\quad |\nabla_x\nabla_y G(x,y)|\leq C|x-y|^{-d}, \tag{5.60}$$

for some constant $C>0$.

Collecting (5.58) and (5.59), and as $\nabla w_p^0\in L^\infty(Q)$, we obtain

$$\forall x\in\mathbb{R}^d,\quad |\phi_p(x)|\leq C\int_Q |\nabla_y G(x,y)|\,dy.$$

Recall now that, as $G(\cdot,y)$, the function $G(x,\cdot)$ belongs to $W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p<d/(d-1)$. Thus $\nabla_y G(x,\cdot)\in L^1(Q)$ and ϕ_p is bounded on $\mathbb{R}^d$. In addition, we infer from (5.59) that, when $|x|\geq 1$, $|\phi_p(x)|\leq C/(1+|x|^{d-1})$ for some constant C. We thus obtain (5.51).

We next infer from (5.58) that

$$\nabla\phi_p(x)=-\int_Q \nabla_x\nabla_y G(x,y)\left(\nabla w_p^0(y)+p\right)\,dy. \tag{5.61}$$

Consider x with $|x| \geq 1$ and $y \in Q$. As Q is the unit cube centered at the origin, $|x-y|$ is isolated from 0, and we can thus use (5.60). As $\nabla w_p^0 \in L^\infty(Q)$, we deduce from (5.61) that, for all $|x| \geq 1$,

$$|\nabla \phi_p(x)| \leq \overline{C} \int_Q |x-y|^{-d} dy \leq C|x|^{-d}$$

for two constants C and $\overline{C}$. This gives (5.50), and concludes the proof of Lemma 5.3.1. □

Proof of Lemma 5.3.2 : We first point out that (5.52) admits a solution in $H^1_\square(Q)$. It is a simple consequence of the Lax-Milgram lemma. Next, we prove that the sum in (5.53) is a convergent series in $L^2(Q \times \Omega)$. For this purpose, we compute the norm of the remainder of the series, using the independence of the X_k :

$$\left\| \sum_{|k| \geq N+1} \nabla \phi_p(\cdot - k) \left(X_k - \mathbb{E}(X_k)\right) \right\|^2_{L^2(Q\times\Omega)} = \mathbb{V}\mathrm{ar}(X_0) \sum_{|k|\geq N+1} \int_{Q+k} |\nabla \phi_p|^2,$$

which converges to 0 as $N \to \infty$ since $\nabla \phi_p \in \left(L^2(\mathbb{R}^d)\right)^d$. Hence, the right-hand side of (5.53) defines a function $T \in \left(L^2(Q\times\Omega)\right)^d$. As $\partial_i T_j = \partial_j T_i$, there exists a function $\widetilde{w}_p^1$ such that

$$\nabla \widetilde{w}_p^1 = T = \nabla \overline{w}_p^1 + \sum_{k \in \mathbb{Z}^d} \nabla \phi_p(x-k) \left(X_k(\omega) - \mathbb{E}(X_k)\right).$$

As $\overline{w}_p^1$ is Q-periodic, we infer from the above equality that

$$\nabla \widetilde{w}_p^1 \text{ is stationary and } \int_Q \mathbb{E}(\nabla \widetilde{w}_p^1) = 0. \tag{5.62}$$

Next, we compute

$$A_{\mathrm{per}} \nabla \widetilde{w}_p^1 = \sum_{k \in \mathbb{Z}^d} A_{\mathrm{per}} \nabla \phi_p(x-k) \left(X_k - \mathbb{E}(X_k)\right) + A_{\mathrm{per}} \nabla \overline{w}_p^1.$$

Taking the divergence of this equation, we thus find that

$$\begin{aligned} -\mathrm{div}\left[A_{\mathrm{per}} \nabla \widetilde{w}_p^1\right] &= \sum_{k\in\mathbb{Z}^d} -\mathrm{div}\left[A_{\mathrm{per}} \nabla \phi_p(x-k)\right] \left(X_k - \mathbb{E}(X_k)\right) \\ &\quad -\mathrm{div}\left[A_{\mathrm{per}} \nabla \overline{w}_p^1\right] \\ &= \sum_{k\in\mathbb{Z}^d} \mathrm{div}\left[\mathbf{1}_{Q+k}\left(\nabla w_p^0 + p\right)\right] \left(X_k - \mathbb{E}(X_k)\right) \\ &\quad +\mathrm{div}\left[\nabla w_p^0 + p\right] \mathbb{E}(X_k) \\ &= \mathrm{div}\left[A_1\left(\nabla w_p^0 + p\right)\right]. \end{aligned} \tag{5.63}$$

Collecting (5.62) and (5.63), we see that $\widetilde{w}_p^1$ solves (5.37). As the solution to this equation is unique up to the addition of a (possibly random) constant $C(\omega)$, we obtain that $\widetilde{w}_p^1 = w_p^1 + C(\omega)$, hence proving (5.53). □

5.3.3 Variance of $A_{1,N}^{\star,\mathrm{exact}}$

We are now in position to prove Proposition 5.3.1. Let us briefly explain the structure of the argument. Exploiting the fact that the randomness in $A(x,\omega)$ only appears as a small perturbation, it turns out that $A_{1,N}^{\star,\mathrm{exact}}$, which is the *first order* correction of the homogenized matrix, depends *linearly* on the random variables $X_k(\omega)$ involved in the matrix $A(x,\omega)$. This can be clearly seen on (5.45), (5.34) and (5.53). As a consequence, it is possible to explicitly, analytically, write how the random matrix $A_{1,N}^{\star,\mathrm{exact}}$ depends on the random variables X_k. This explicit expression yields the relation between the variance of $A_{1,N}^{\star,\mathrm{exact}}$ and that of X_k, namely (5.46) above.

Proof of Proposition 5.3.1 : Using (5.45) and the expression (5.53) of ∇w_p^1 provided by Lemma 5.3.2, we have

$$\begin{aligned}
\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij} &= \frac{1}{|Q_N|}\sum_{|k|\le N}\Bigg[\int_Q \left(\nabla w_{e_i}^0 + e_i\right)^T \left(\nabla w_{e_j}^0 + e_j\right) X_k \\
&+ \sum_{\ell\in\mathbb{Z}^d}\int_{Q+k} \nabla\phi_{e_i}(x-\ell)^T A_{\mathrm{per}}(x)\left(\nabla w_{e_j}^0(x)+e_j\right)dx\,(X_\ell - \mathbb{E}(X_\ell)) \\
&+ \sum_{\ell\in\mathbb{Z}^d}\int_{Q+k} \left(\nabla w_{e_i}^0(x)+e_i\right)^T A_{\mathrm{per}}(x)\nabla\phi_{e_j}(x-\ell)dx\,(X_\ell - \mathbb{E}(X_\ell))\Bigg] \\
&+ \frac{1}{|Q_N|}\int_{Q_N}\left(\nabla\overline{w}_{e_i}^1\right)^T A_{\mathrm{per}}\left(\nabla w_{e_j}^0+e_j\right) \\
&+ \frac{1}{|Q_N|}\int_{Q_N}\left(\nabla w_{e_i}^0+e_i\right)^T A_{\mathrm{per}}\nabla\overline{w}_{e_j}^1.
\end{aligned}$$

Setting, for any $1\le i,j\le d$ and $m\in\mathbb{Z}^d$,

$$\gamma_{ij} = \int_Q \left(\nabla w_{e_i}^0+e_i\right)^T\left(\nabla w_{e_j}^0+e_j\right), \tag{5.64}$$

$$\alpha_{ij}(m) = \int_Q \nabla\phi_{e_i}(x-m)^T A_{\mathrm{per}}(x)\left(\nabla w_{e_j}^0(x)+e_j\right)dx, \tag{5.65}$$

and using that, for any $p\in\mathbb{R}^d$, the functions w_p^0, $\overline{w}_p^1$ and A_{per} are Q-periodic, we obtain

$$\begin{aligned}
\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij} &= \frac{1}{|Q_N|}\sum_{|k|\le N}\left[\gamma_{ij}X_k + \sum_{\ell\in\mathbb{Z}^d}\left(\alpha_{ij}(\ell)+\alpha_{ji}(\ell)\right)\left(X_{\ell+k}-\mathbb{E}(X_{\ell+k})\right)\right] \\
&+\frac{1}{|Q|}\int_Q\left(\nabla\overline{w}_{e_i}^1\right)^T A_{\mathrm{per}}\left(\nabla w_{e_j}^0+e_j\right) \\
&+\frac{1}{|Q|}\int_Q\left(\nabla w_{e_i}^0+e_i\right)^T A_{\mathrm{per}}\nabla\overline{w}_{e_j}^1.
\end{aligned}$$

Introducing

$$\beta_{ij}(\ell) = \alpha_{ij}(\ell) + \alpha_{ji}(\ell), \quad Y_k = X_k - \mathbb{E}(X_k), \tag{5.66}$$

we next obtain, using (5.40), that

$$\left[A_{1,N}^{\star,\text{exact}}\right]_{ij} = [A_1^\star]_{ij} + \frac{1}{|Q_N|}\sum_{|k|\leq N}\left[\gamma_{ij}Y_k + \sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)Y_{\ell+k}\right]. \tag{5.67}$$

Note that

$$\begin{aligned} |\beta_{ij}(\ell)|^2 &\leq 2\left(\alpha_{ij}^2(\ell)+\alpha_{ji}^2(\ell)\right) \\ &\leq 2\|A_{\text{per}}\|_{L^\infty}^2\ C(w^0)\int_{Q+\ell}\left(|\nabla\phi_{e_i}|^2+|\nabla\phi_{e_j}|^2\right), \end{aligned} \tag{5.68}$$

where $C(w^0) = \max_j \|\nabla w^0_{e_j} + e_j\|^2_{L^2(Q)}$. As $\nabla\phi_p$ belongs to $\left(L^2(\mathbb{R}^d)\right)^d$ for any $p\in\mathbb{R}^d$, we first deduce that, for any $1\leq i,j\leq d$,

$$\sum_{\ell\in\mathbb{Z}^d}|\beta_{ij}(\ell)|^2 < \infty. \tag{5.69}$$

Furthermore, we have

$$\mathbb{E}\left(|\beta_{ij}(\ell)Y_{k+\ell}|^2\right) = \mathbb{V}\text{ar}\left(X_0\right)\ |\beta_{ij}(\ell)|^2.$$

Since Y_k are independent, identically distributed random variables of zero mean, we infer from (5.69) that the sum $\sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)Y_{\ell+k}$ in (5.67) is convergent in $L^2(\Omega)$, for any $k\in\mathbb{Z}^d$, and any $1\leq i,j\leq d$.

As $\mathbb{E}\left(Y_k\right)=0$, we infer from (5.67) that $\mathbb{E}\left(A_{1,N}^{\star,\text{exact}}\right) = A_1^\star$. Since $A_1^\star$, γ_{ij} and $\beta_{ij}(\ell)$ are deterministic and the variables $(Y_k)_{k\in\mathbb{Z}^d}$ are independent, identically distributed, and of mean zero, we deduce from (5.67) that

$$\begin{aligned} \mathbb{V}\text{ar}\left(\left[A_{1,N}^{\star,\text{exact}}\right]_{ij}\right) &= \mathbb{V}\text{ar}(X_0)\frac{\gamma_{ij}^2}{|Q_N|} \\ &+ \frac{1}{|Q_N|^2}\sum_{|k|\leq N}\sum_{|k'|\leq N}\mathbb{C}\text{ov}\left(\sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)Y_{\ell+k},\sum_{\ell'\in\mathbb{Z}^d}\beta_{ij}(\ell')Y_{\ell'+k'}\right) \\ &+ \frac{2}{|Q_N|^2}\sum_{|k|\leq N}\sum_{|k'|\leq N}\mathbb{C}\text{ov}\left(\gamma_{ij}Y_{k'},\sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)Y_{\ell+k}\right), \\ &= \mathbb{V}\text{ar}(X_0)\frac{\gamma_{ij}^2}{|Q_N|} \\ &+ \frac{1}{|Q_N|^2}\sum_{|k|\leq N}\sum_{|k'|\leq N}\sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)\beta_{ij}(\ell+k-k')\mathbb{V}\text{ar}(X_0) \\ &+ \frac{2\gamma_{ij}}{|Q_N|^2}\sum_{|k|\leq N}\sum_{|k'|\leq N}\beta_{ij}(k'-k)\mathbb{V}\text{ar}(X_0) \\ &= C_N^{ij}\ \mathbb{V}\text{ar}(X_0), \end{aligned}$$

where

$$C_N^{ij} = \frac{\gamma_{ij}^2}{|Q_N|} + \frac{1}{|Q_N|^2} \sum_{|k|\le N} \sum_{|k'|\le N} \sum_{\ell\in\mathbb{Z}^d} \beta_{ij}(\ell)\beta_{ij}(\ell+k-k') + \frac{2\gamma_{ij}}{|Q_N|^2} \sum_{|k|\le N} \sum_{|k'|\le N} \beta_{ij}(k'-k). \quad (5.70)$$

Note that the sum $\sum_{\ell\in\mathbb{Z}^d} \beta_{ij}(\ell)\beta_{ij}(\ell+k-k')$ is finite. This is a simple consequence of the Cauchy-Schwarz inequality and (5.69). Hence C_N^{ij} is finite and we thus have proved (5.46).

We next prove (5.47). For this purpose, we examine each term of the right-hand side of (5.70) separately. It is clear that the first one satisfies (5.47). The second one writes

$$\frac{1}{|Q_N|^2} \sum_{\ell\in\mathbb{Z}^d} \sum_{|k|\le N} \sum_{|k'|\le N} \beta_{ij}(\ell+k')\beta_{ij}(\ell+k) = \sum_{\ell\in\mathbb{Z}^d} \left[\frac{1}{|Q_N|} \sum_{|k|\le N} \beta_{ij}(k+\ell) \right]^2. \quad (5.71)$$

We first consider the large values of ℓ, namely $|\ell| \ge 2N$. As $|k| \le N$, we have $|\ell+k| \ge N$. Using (5.50) and (5.68), we deduce $|\beta_{ij}(k+\ell)| \le C|k+\ell|^{-d}$ for some constant C independent of ℓ and k. Hence, we have

$$\begin{aligned} \sum_{|\ell|\ge 2N} \left(\frac{1}{|Q_N|} \sum_{|k|\le N} \beta_{ij}(k+\ell) \right)^2 &\le \sum_{|\ell|\ge 2N} \left(\frac{1}{|Q_N|} \sum_{|k|\le N} \frac{C}{|k+\ell|^d} \right)^2 \\ &\le \sum_{|\ell|\ge 2N} \frac{C^2}{(|\ell|-N)^{2d}} \\ &\le \sum_{|\ell|\ge 2N} \frac{2^{2d}C^2}{|\ell|^{2d}} \\ &\le \frac{C'}{N^d}, \end{aligned} \quad (5.72)$$

for some constant C' independent of N. Here, we have used the fact that if $|\ell| \ge 2N$, then $|\ell| - N \ge |\ell|/2$.

Next, we consider the small values of ℓ, namely $|\ell| \le 2N$. For this purpose, we

note that

$$\begin{aligned}\sum_{|k|\leq N}\beta_{ij}(k+\ell) &= \sum_{|k|\leq N}\Big[\int_{Q+\ell+k}\left(\nabla w^0_{e_i}(x)+e_i\right)^T A_{\rm per}(x)\nabla\phi_{e_j}(x)dx \\ &\quad + \int_{Q+\ell+k}\nabla\phi_{e_i}(x)^T A_{\rm per}(x)\left(\nabla w^0_{e_j}(x)+e_j\right)dx\Big] \\ &= \int_{Q_N+\ell}\left(\nabla w^0_{e_i}(x)+e_i\right)^T A_{\rm per}(x)\nabla\phi_{e_j}(x)dx \\ &\quad + \int_{Q_N+\ell}\nabla\phi_{e_i}(x)^T A_{\rm per}(x)\left(\nabla w^0_{e_j}(x)+e_j\right)dx.\end{aligned}$$

We use an integration by parts and (5.36) to write this as

$$\sum_{|k|\leq N}\beta_{ij}(k+\ell) = \int_{\partial(Q_N+\ell)}\phi_{e_j}n^T A_{\rm per}\left(\nabla w^0_{e_i}+e_i\right) + \int_{\partial(Q_N+\ell)}\phi_{e_i}\left(\nabla w^0_{e_j}+e_j\right)^T A_{\rm per}n.$$

Using (5.51) and the fact that the periodic function w^0_p satisfies $w^0_p\in W^{1,\infty}(Q)$, we obtain

$$\left|\sum_{|k|\leq N}\beta_{ij}(k+\ell)\right| \leq C\int_{\partial(Q_N+\ell)}\frac{dx}{1+|x|^{d-1}},$$

where C is a constant which does not depend on N and ℓ, and dx is the Lebesgue measure on the $(d-1)$-dimension surface $\partial(Q_N+\ell)$. In order to estimate the right-hand side of this inequality, we change variables in the integral, getting

$$\begin{aligned}\left|\sum_{|k|\leq N}\beta_{ij}(k+\ell)\right| &\leq C\int_{\partial(Q+\frac{\ell}{N})}\frac{N^{d-1}dy}{1+N^{d-1}|y|^{d-1}}, \\ &= C\int_{\partial Q}\frac{dy}{\frac{1}{N^{d-1}}+\left|y-\frac{\ell}{N}\right|^{d-1}}. \qquad (5.73)\end{aligned}$$

We use a Riemann sum to write

$$\begin{aligned}\frac{1}{|Q_N|}\sum_{|\ell|\leq 2N}\left(\sum_{|k|\leq N}\beta_{ij}(k+\ell)\right)^2 &\leq \frac{C}{|Q_N|}\sum_{|\ell|\leq 2N}\left(\int_{\partial Q}\frac{dy}{\left|y-\frac{\ell}{N}\right|^{d-1}}\right)^2 \\ &\to 2^dC\int_{2Q}\left(\int_{\partial Q}\frac{dy}{|y-z|^{d-1}}\right)^2 dz. \qquad (5.74)\end{aligned}$$

Note that the integrals in (5.74) converge since ∂Q is a $(d-1)$-dimensional manifold, hence

$$\int_{\partial Q}\frac{dy}{|y-z|^{d-1}} \leq C\log\left(\mathrm{dist}(z,\partial Q)\right),$$

which, as a function of z, belongs to $L^2(2Q)$. Collecting (5.72) and (5.74), we find

$$\sum_{\ell\in\mathbb{Z}^d}\left(\frac{1}{|Q_N|}\sum_{|k|\le N}\beta_{ij}(k+\ell)\right)^2 \le \frac{C}{|Q_N|},$$

thereby proving that the second term of the right-hand side of (5.70) satisfies (5.47).

For the last term of the right-hand side of (5.70), we return to (5.73), and use it exactly as in (5.74) to prove

$$\begin{aligned}\frac{1}{|Q_N|}\sum_{|k'|\le N}\left|\sum_{|k|\le N}\beta_{ij}(k+k')\right| &\le \frac{C}{|Q_N|}\sum_{|k'|\le N}\left|\int_{\partial Q}\frac{dy}{\left|y-\frac{k'}{N}\right|^{d-1}}\right| \\ &\longrightarrow C\int_Q\left|\int_{\partial Q}\frac{dy}{|y-z|^{d-1}}\right|dz.\end{aligned}$$

Collecting all the above estimates concludes the proof of (5.47). □

Proof of Corollary 5.3.1 : By construction (see (5.45)),

$$\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}F_{ij}(x,\omega)\,dx,$$

where F_{ij} is a stationary function. The ergodic theorem thus applies :

$$\lim_{N\to\infty}\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij} = \mathbb{E}\left(\int_Q F_{ij}(x,\cdot)\,dx\right) = [A_1^\star]_{ij} \quad \text{a.s.},$$

which proves the first assertion.

We now prove the second assertion. In view of (5.67), we have

$$S_N^{ij} := \left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij} - [A_1^\star]_{ij} = \frac{1}{|Q_N|}\sum_{|k|\le N}\theta_k^{ij} \tag{5.75}$$

with

$$\theta_k^{ij} = \gamma_{ij}Y_{-k} + \sum_{\ell\in\mathbb{Z}^d}\beta_{ij}(\ell)Y_{\ell-k} = \sum_{\ell\in\mathbb{Z}^d}\overline{\beta}_{ij}(\ell+k)Y_\ell, \tag{5.76}$$

where

$$\overline{\beta}_{ij}(\ell) = \beta_{ij}(\ell) \quad \text{if } \ell\neq 0, \quad \overline{\beta}_{ij}(0) = \beta_{ij}(0)+\gamma_{ij},$$

with Y_k and β_{ij} defined by (5.65)-(5.66), and γ_{ij} by (5.64). Thus, Y_k are independent, identically distributed random variables with $\mathbb{E}(Y_0)=0$ and $\mathbb{E}(Y_0^2)<\infty$. We infer from (5.69) that $\sum_{\ell\in\mathbb{Z}^d}\left(\overline{\beta}_{ij}(\ell)\right)^2 < \infty$. Hence, equations (5.75) and (5.76) exactly correspond to the setting considered in [86].

Observe now that

$$\begin{aligned}
|Q_N|^2 \, \mathbb{V}\mathrm{ar}\left(\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij}\right) &= \mathbb{E}\left[\left(\sum_{|k|\leq N} \theta_k^{ij}\right)^2\right] \\
&= \sum_{|k|\leq N}\sum_{|k'|\leq N} \mathbb{E}\left[\theta_k^{ij}\ \theta_{k'}^{ij}\right] \\
&= \sum_{|k|\leq N}\sum_{|k'|\leq N}\sum_{\ell\in\mathbb{Z}^d} \overline{\beta}_{ij}(\ell+k)\overline{\beta}_{ij}(\ell+k')\mathbb{E}(Y_0^2) \\
&= \mathbb{E}(Y_0^2)\sum_{\ell\in\mathbb{Z}^d}\left(\sum_{|k|\leq N}\overline{\beta}_{ij}(\ell+k)\right)^2 .
\end{aligned}$$

Our assumption (5.48) hence reads $\lim_{N\to\infty}\sum_{\ell\in\mathbb{Z}^d}\left(\sum_{|k|\leq N}\overline{\beta}_{ij}(\ell+k)\right)^2 = \infty$. We are thus in position to apply [86, Theorem 1], which yields the second assertion. □

5.3.4 Extensions

We briefly mention here some possible extensions of Proposition 5.3.1.

Our proof, performed in the setting (5.34), can be extended to the case

$$A_1(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x)\, X_k(\omega)\ \varphi_{\mathrm{per}}(x)\,\mathrm{Id},$$

where $(X_k)_{k\in\mathbb{Z}^d}$ is again a sequence of independent, identically distributed scalar random variables, and φ_{per} is a scalar-valued function, Q-periodic, and in $L^\infty(Q)$. Our proof can also be extended to the case

$$A_1(x,\omega) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{Q+k}(x) B_k(\omega),$$

where $(B_k)_{k\in\mathbb{Z}^d}$ is a sequence of independent, identically distributed symmetric bounded matrices. A result similar to (5.46)–(5.47) is then valid, in the following sense. Define the matrix $\Sigma\in\mathbb{R}^{d^2\times d^2}$ as the covariance matrix of A_1, that is

$$\Sigma_{mn} = \mathbb{C}\mathrm{ov}\left(\left[A_1\right]_{ij}, \left[A_1\right]_{k\ell}\right), \quad \text{where} \quad \left\{\begin{array}{l} m=(i-1)d+j, \\ n=(k-1)d+\ell. \end{array}\right.$$

Similarly, define $\Sigma_N\in\mathbb{R}^{d^2\times d^2}$ as the covariance matrix of $A_{1,N}^{\star,\mathrm{exact}}$. Then we have

$$\Sigma_N = C_N^T\ \Sigma\ C_N, \tag{5.77}$$

where the matrix $C_N \in \mathbb{R}^{d^2 \times d^2}$ does not depend on the random matrices B_k, and satisfies the bound

$$|C_N| \leq \frac{C}{\sqrt{|Q_N|}}, \tag{5.78}$$

for C a constant independent of N. Note that in order to prove the above property, one needs to carry out computations similar to those in Sections 5.3.2 and 5.3.3, except that (5.49) is now a system of d^2 partial differential equations indexed by $1 \leq n, m \leq d$:

$$\left\{ \begin{array}{l} -\mathrm{div}\left[A_{\mathrm{per}} \nabla \phi_p^{m,n}\right] = \mathrm{div}\left[E_{mn} \mathbf{1}_Q \left(\nabla w_p^0 + p\right)\right], \\ \phi_p^{m,n} \in L^2_{\mathrm{loc}}(\mathbb{R}^d), \quad \nabla \phi_p^{m,n} \in \left(L^2(\mathbb{R}^d)\right)^d, \end{array} \right. \tag{5.79}$$

where E_{mn} is the canonical basis of the space of matrices of size $d \times d$. This problem is easily shown to have a unique solution, by the same method as in the proof of Lemma 5.3.1, with the estimate (5.50). Then, Lemma 5.3.2 reads

$$\nabla w_p^1(x,\omega) = \nabla \overline{w}_p^1(x) + \sum_{k \in \mathbb{Z}^d} \sum_{1 \leq m,n \leq d} \nabla \phi_p^{m,n}(x-k) \left(B_k^{mn}(\omega) - \mathbb{E}(B_k^{mn})\right),$$

with $\overline{w}_p^1$ defined by (5.52), and where $B_k^{mn}(\omega)$ is the coefficient mn of the matrix $B_k(\omega)$. This decomposition may then be used to prove (5.77) and (5.78) above, by the same method as in Section 5.3.3.

Another observation is the following. Return to the case (5.34), but assume now that X_k is a stationary sequence (we do not assume independence anymore). Then a similar result holds. In such a case, one needs to assume that

$$\mathcal{E} = \sum_{k \in \mathbb{Z}^d} |\mathbb{C}\mathrm{ov}(X_0, X_k)| < +\infty, \tag{5.80}$$

and equality (5.46) is then replaced by

$$\mathbb{V}\mathrm{ar}\left(\left[A_{1,N}^{\star,\mathrm{exact}}\right]_{ij}\right) \leq \mathcal{E}\, C_N^{ij}, \tag{5.81}$$

where C_N^{ij} is independent from the random variables X_k and satisfies (5.47).

Acknowledgments

The work of RC, CLB and FL is partially supported by ONR under Grant 00014-09-1-0470. The authors thank Eric Bonnetier and Arnaud Debussche for stimulating discussions and for pointing out useful references. We also thank Raphaël Roux for his suggestions on a previous version of this article.

Deuxième partie

Problèmes faiblement stochastiques

Chapitre 6

Approximation numérique d'une classe de problèmes en homogénéisation stochastique

Ce chapitre a donné lieu à la publication [P1]. L'objectif est d'étudier dans le cadre du modèle perturbatif introduit par X. Blanc, P.L. Lions et C. Le Bris dans [19,20] le comportement du terme d'ordre 2 dans le développement de la matrice homogénéisée $A^\star_\eta = A^0 + \eta A^1 + O(\eta^2)$ où η est un paramètre réel associé à l'amplitude de la perturbation aléatoire. Si ce développement de la matrice homogénéisée est en théorie déterministe, le recours à des méthodes numériques pour approximer $A^\star_\eta$ nous oblige à considérer des quantités aléaoires. C'est en particulier le cas pour l'approximation du terme en $O(\eta^2)$. On montre ici numériquement que sa variance est bornée lorsque $\eta \to 0$.

Approximation numérique d'une classe de problémes en homogénéisation stochastique

Ronan Costaouec[a], Claude Le Bris[a] and Frédéric Legoll[a]

[a] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

{*costaour,lebris*}*@cermics.enpc.fr,*
legoll@lami.enpc.fr

6.1 Introduction (in english)

We study in this Note a numerical strategy for the stochastic homogenization of an elliptic boundary problem, in the case when the random coefficients are a small perturbation of deterministic periodic coefficients. More precisely, we follow [19, 20] and consider the equation (6.1) below, for an elliptic bounded periodic matrix A, and for a diffeomorphism Φ whose gradient is stationary in the sense of (6.2) below. The stochastic formalism is identical to that of [19, 20], and briefly recalled in the French version.

In the case when Φ is a perturbation of the identity, that is $\Phi(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2)$, it is possible to efficiently compute an approximation of the homogenized matrix $A_\star$ (see [20]) : $A_\star = A^0 + \eta A^1 + O(\eta^2)$, where A^0 and A^1 are known from the solution of two *deterministic periodic* corrector problems, whereas $A_\star$ depends on the solution of a stochastic equation posed on $\mathbb{R}^d$, which is therefore delicate and expensive to compute.

In this Note, we show that such an approximation result also holds in the case when the corrector problems are discretized, that is they are solved on the truncated domain $Q_N = [0,N]^d \subset \mathbb{R}^d$, using a Finite Element method based on a mesh of size h. In that discrete setting, the deterministic matrix $A_\star$ is approximated by a stochastic matrix $A_\star^{h,N}(\omega)$. Estimating the error introduced by the truncation in the series in power of η brings up questions related to variance, that we analyze and discuss. Indeed, it turns out that two natural quantities can be considered to approximate $A_\star^{h,N}(\omega)$ when $\eta \to 0$, namely a stochastic matrix $A_{\rm app}^{h,N}(\omega)$ and a deterministic matrix $\overline{A}^h_{\rm app}$. We hence introduce the estimators (see (6.16) and (6.17) below)

$$e^{h,N}(\omega) = \eta^{-2}\left(A_\star^{h,N}(\omega) - A_{\rm app}^{h,N}(\omega)\right) \quad \text{and} \quad \overline{e}^{h,N}(\omega) = \eta^{-2}\left(A_\star^{h,N}(\omega) - \overline{A}^h_{\rm app}\right),$$

which are two different numerical approximations of $\eta^{-2}\left(A_\star - A^0 - \eta A^1\right)$.

The approximation $A^{h,N}_{\rm app}(\omega)$ is useful for the *analysis* of the method, since the associated estimator $e^{h,N}(\omega)$ has a bounded expectation and variance when $\eta \to 0$, which is not the case with the estimator $\overline{e}^{h,N}(\omega)$. On the other hand, the objective of the numerical *practice* is different, and leads to rather working with $\overline{A}^h_{\rm app}$, which is cheaper to evaluate.

Numerical simulations performed on a representative example illustrate the efficiency of the approach : we observe that, for any fixed $h > 0$ and $N < \infty$, the quantity $e^{h,N}(\omega)$ is bounded by a small constant when $\eta \to 0$.

6.2 Introduction

Nous étudions dans cette Note une stratégie numérique pour l'homogénéisation d'un problème aux limites elliptique dont les coefficients sont stochastiques. On s'intéresse ici au cas où l'équation considérée est une *petite perturbation* aléatoire d'un problème déterministe périodique. Plus précisément, on se place dans le cadre étudié dans [19, 20], et on considère l'équation

$$-\mathrm{div}\left(A\left(\Phi^{-1}\left(\frac{x}{\varepsilon},\omega\right)\right)\nabla u_\varepsilon\right) = f \quad \text{dans} \quad \mathcal{D}, \quad u_\varepsilon = 0 \quad \text{sur} \quad \partial\mathcal{D}, \tag{6.1}$$

où la matrice A est périodique et où, presque sûrement, Φ est un difféomorphisme de $\mathbb{R}^d$ dans $\mathbb{R}^d$, dont le gradient est stationnaire au sens de (6.2) ci-dessous. On s'intéresse au problème homogénéisé obtenu dans la limite $\varepsilon \to 0$, dans le cas où Φ est une perturbation de l'identité.

On rappelle brièvement le cadre de travail adopté dans [19, 20]. Dans toute la suite, $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé. Pour toute variable aléatoire $X \in L^1(\Omega, d\mathbb{P})$, on note $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ son espérance. On suppose que le groupe $(\mathbb{Z}^d, +)$ agit sur Ω, par une action notée $(\tau_k)_{k\in\mathbb{Z}^d}$, qui préserve la mesure $\mathbb{P}$ et est ergodique. On dira que $F \in L^1_{\rm loc}\left(\mathbb{R}^d, L^1(\Omega)\right)$ est *stationnaire* si elle vérifie :

$$\forall k \in \mathbb{Z}^d, \quad F(x+k,\omega) = F(x,\tau_k\omega) \quad \text{presque partout en } x, \text{ presque sûrement.} \tag{6.2}$$

Pour toute la suite, on se donne une matrice $A(y) = [a_{ij}(y)] \in \left(L^\infty(\mathbb{R}^d)\right)^{d\times d}$, fonction Q-périodique ($Q = [0,1]^d$) et telle que, pour un certain $\gamma > 0$, on ait $\xi^T A(y)\xi \geq \gamma|\xi|^2$ pour tout $\xi \in \mathbb{R}^d$ et presque partout en $y \in \mathbb{R}^d$. On supposera pour simplifier que cette matrice est symétrique. On se donne aussi Φ tel que, pour presque tout ω, l'application $\Phi(\cdot,\omega)$ est un difféomorphisme de $\mathbb{R}^d$ dans $\mathbb{R}^d$, vérifiant d'une part que $\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\mathrm{EssInf}} \left[\det(\nabla\Phi(x,\omega))\right] = \nu > 0$, d'autre part que $\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\mathrm{EssSup}} \left(|\nabla\Phi(x,\omega)|\right) = M < \infty$, et enfin que $\nabla\Phi(x,\omega)$ est stationnaire au sens de (6.2).

Dans [19], les auteurs montrent le résultat suivant.

Théorème 6.2.1 *Pour tout $p \in \mathbb{R}^d$ fixé, le système*

$$\begin{cases} \operatorname{div}\left[A\left(\Phi^{-1}(y,\omega)\right)(p+\nabla w_p)\right]=0, \qquad \mathbb{E}\left(\int_{\Phi(Q,\cdot)} \nabla w_p(y,\cdot)dy\right)=0, \\ w_p(y,\omega)=\widetilde{w}_p\left(\Phi^{-1}(y,\omega),\omega\right), \quad \nabla\widetilde{w}_p \text{ est stationnaire au sens de (6.2)}, \end{cases} \tag{6.3}$$

admet dans $H^1_{\mathrm{loc}}(\mathbb{R}^d, L^2(\Omega))$ une solution, qui est unique à l'ajout d'une constante (aléatoire) près.

De plus, la matrice homogénéisée $A_\star$ associée au problème (6.1) est définie par

$$(A_\star)_{ij} = \frac{\mathbb{E}\left(\int_{\Phi(Q,\cdot)} \left(e_i+\nabla w_{e_i}(y,\cdot)\right)^T A\left(\Phi^{-1}(y,\cdot)\right) e_j\, dy\right)}{\det\left(\mathbb{E}\left(\int_Q \nabla\Phi(z,\cdot)dz\right)\right)}$$

Lorsque Φ est une perturbation de l'identité, soit

$$\Phi(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2), \tag{6.4}$$

on peut développer la solution du problème du correcteur (6.3), comme montré dans [20] (notons qu'un autre cadre de travail, lui aussi perturbatif quoique différent, est étudié dans [7]). Plus précisément, on a $\widetilde{w}_p(x,\omega) = w^0_p(x) + \eta w^1_p(x,\omega) + O(\eta^2)$, où w^0_p vérifie

$$-\operatorname{div}\left[A\left(p+\nabla w^0_p\right)\right]=0, \quad w^0_p \text{ est } Q\text{-périodique}, \tag{6.5}$$

tandis que w^1_p est donné par

$$\begin{cases} -\operatorname{div}\left[A\nabla w^1_p\right] = \operatorname{div}\left[-A\nabla\Psi\nabla w^0_p - \left(\nabla\Psi^T - (\operatorname{div}\Psi)\mathrm{Id}\right)A\left(p+\nabla w^0_p\right)\right], \\ \nabla w^1_p \text{ est stationnaire et } \mathbb{E}\left(\int_Q \nabla w^1_p\right)=0. \end{cases} \tag{6.6}$$

Le problème en w^1_p est stochastique, donc coûteux et délicat à résoudre. Cependant, comme remarqué dans [20], la fonction $\overline{w}^1_p = \mathbb{E}(w^1_p)$ est Q-périodique et vérifie le problème déterministe

$$-\operatorname{div}\left[A\nabla\overline{w}^1_p\right] = \operatorname{div}\left[-A\,\mathbb{E}(\nabla\Psi)\nabla w^0_p - \left(\mathbb{E}(\nabla\Psi^T) - \mathbb{E}(\operatorname{div}\Psi)\mathrm{Id}\right)A\left(p+\nabla w^0_p\right)\right], \tag{6.7}$$

qui est bien plus facile à résoudre que (6.6). De plus, la connaissance de w^0_p et $\overline{w}^1_p$ suffit pour obtenir une approximation de la matrice homogénéisée $A_\star$ au premier ordre en η. On a en effet

$$A_\star = A^0 + \eta A^1 + O(\eta^2), \tag{6.8}$$

avec $A^0_{ij} = \displaystyle\int_Q \left(e_i+\nabla w^0_{e_i}\right)^T A\, e_j$ et

$$\begin{aligned} A^1_{ij} &= -\int_Q \mathbb{E}(\operatorname{div}\Psi)\, A^0_{ij} + \int_Q (e_i+\nabla w^0_{e_i})^T A\, e_j\, \mathbb{E}(\operatorname{div}\Psi) \\ &+ \int_Q \left(\nabla\overline{w}^1_{e_i} - \mathbb{E}(\nabla\Psi)\nabla w^0_{e_i}\right)^T A\, e_j. \end{aligned}$$

6.3 Approximation numérique

Dans cette section, nous établissons un résultat analogue à (6.8), dans le cas où les problèmes des correcteurs sont résolus *numériquement*. Pour h fixé, on se donne une triangulation périodique $\mathcal{T}_h^{(Q)}$ de Q. Par périodicité, on obtient ainsi la triangulation $\mathcal{T}_h = \cup_{k\in\mathbb{Z}^d}\left(k+\mathcal{T}_h^{(Q)}\right)$ de $\mathbb{R}^d$. Pour tout $N\in\mathbb{N}^*$, on note $Q_N = [0,N]^d$, et $\mathcal{T}_h^N = \mathcal{T}_h \cap Q_N$ la triangulation associée. En s'inspirant de [23, 87] pour discrétiser (6.3), on considère le problème

$$\left\{\begin{array}{l} \text{Trouver } \widetilde{w}_p^{h,N}(\cdot,\omega)\in V_h^{\rm per}(Q_N) \text{ tel que, pour tout } \widetilde{v}_h\in V_h^{\rm per}(Q_N),\\ \int_{Q_N}\det\left(\nabla\Phi\right)\left(\nabla\widetilde{v}_h\right)^T\left(\nabla\Phi\right)^{-T}A\left(p+\left(\nabla\Phi\right)^{-1}\nabla\widetilde{w}_p^{h,N}(\cdot,\omega)\right)=0 \text{ presque sûrement,}\end{array}\right. \tag{6.9}$$

où $V_h^{\rm per}(Q_N)$ est l'ensemble des fonctions définies sur $\mathbb{R}^d$, Q_N-périodiques, et dont la restriction à Q_N appartient à un espace d'éléments finis périodiques construit à partir de $\mathcal{T}_h^N$. On a choisi de travailler avec $\widetilde{w}_p$ plutôt qu'avec w_p car les propriétés de stationnarité de (6.3) portent sur $\widetilde{w}_p$. On introduit ensuite la matrice

$$\left(A_\star^{h,N}\right)_{ij}(\omega) = \frac{\displaystyle\frac{1}{|Q_N|}\int_{Q_N}\det(\nabla\Phi)\left(e_i+\left(\nabla\Phi\right)^{-1}\nabla\widetilde{w}_{e_i}^{h,N}\right)^T A\,e_j}{\displaystyle\det\left(\frac{1}{|Q_N|}\int_{Q_N}\nabla\Phi\right)}. \tag{6.10}$$

On fait maintenant l'hypothèse perturbative (6.4). En insérant le développement limité formel $\widetilde{w}_p^{h,N} = w_p^{0,h,N}+\eta w_p^{1,h,N}+O(\eta^2)$ dans (6.9), on obtient que $w_p^{0,h,N}$ est indépendant de N (on le note désormais $w_p^{0,h}$), et que $w_p^{0,h}$ et $w_p^{1,h,N}$ sont solutions des problèmes

$$\left\{\begin{array}{l}\text{Trouver } w_p^{0,h}\in V_h^{\rm per}(Q) \text{ tel que, pour tout } v_h\in V_h^{\rm per}(Q),\\ \int_Q\left(\nabla v_h\right)^T A\,\left(p+\nabla w_p^{0,h}\right)=0,\end{array}\right. \tag{6.11}$$

et

$$\left\{\begin{array}{l}\text{Trouver } w_p^{1,h,N}(\cdot,\omega)\in V_h^{\rm per}(Q_N) \text{ tel que, pour tout } v_h\in V_h^{\rm per}(Q_N), \text{ et p.s,}\\ \int_{Q_N}\left(\nabla v_h\right)^T A\,\nabla w_p^{1,h,N} =\\ \int_{Q_N}\left(\nabla v_h\right)^T\left[A\,\nabla\Psi\,\nabla w_p^{0,h}+\left(\nabla\Psi^T-(\mathrm{div}\Psi)\,\mathrm{Id}\right)A\left(p+\nabla w_p^{0,h}\right)\right].\end{array}\right. \tag{6.12}$$

On reconnaît dans (6.11) et (6.12) une formulation discrétisée de (6.5) et (6.6), respectivement. Le résultat principal de cette section, qui est la proposition ci-dessous, fournit dans le *cadre discrétisé* ($h>0$ et $N<\infty$) l'analogue du résultat rappelé dans la section précédente, et qui a été obtenu dans le *cadre continu*.

Proposition 6.3.1 *On suppose que $\Phi(x,\omega) = x+\eta\Psi(x,\omega)+O(\eta^2)$ quand $\eta\to 0$, dans $C^1(\mathbb{R}^d, L^2(\Omega))$, où $\nabla\Psi$ est stationnaire. Il existe une constante $C(h,N,\omega)$ telle que, pour η suffisamment petit,*

$$\eta^{-2}\left\|\nabla\widetilde{w}_p^{h,N}(\cdot,\omega)-\nabla w_p^{0,h}-\eta\nabla w_p^{1,h,N}(\cdot,\omega)\right\|_{L^2(Q_N)}\leq\sqrt{|Q_N|}\;C(h,N,\omega), \tag{6.13}$$

où $\widetilde{w}_p^{h,N}$, $w_p^{0,h}$ *et* $w_p^{1,h,N}$ *sont solutions de (6.9), (6.11) et (6.12), respectivement, et telle que*

$$\eta^{-2}\left|A_\star^{h,N}(\omega) - A^{0,h} - \eta A^{1,h,N}(\omega)\right| \leq C(h,N,\omega), \tag{6.14}$$

où $A_\star^{h,N}$ *est la matrice définie par (2.91),* $(A^{0,h})_{ij} = \displaystyle\int_Q \left(e_i + \nabla w_{e_i}^{0,h}\right)^T A\, e_j$ *et*

$$\begin{aligned}(A^{1,h,N})_{ij} &= -(A^{0,h})_{ij}\,\frac{1}{|Q_N|}\int_{Q_N}\mathrm{div}\Psi + \frac{1}{|Q_N|}\int_{Q_N}(e_i+\nabla w_{e_i}^{0,h})^T A\, e_j\,\mathrm{div}\Psi \\ &+ \frac{1}{|Q_N|}\int_{Q_N}\left(\nabla w_{e_i}^{1,h,N} - \nabla\Psi\,\nabla w_{e_i}^{0,h}\right)^T A\, e_j.\end{aligned}$$

La preuve de cette proposition suit les mêmes arguments que celle de la Proposition 3.1 de [20], et sera détaillée dans [P6].

Remarque 6.3.1 *Soient* w_p^0 *et* $w_p^{0,h}$ *respectivement solutions de (6.5) et (6.11). Sous des hypothèses classiques,* $\nabla w_p^{0,h}$ *converge vers* ∇w_p^0 *dans* $L^2(Q)$*, et donc* $\lim_{h\to 0} A^{0,h} = A^0$*. La convergence lorsque* $N\to\infty$ *de* $A_\star^{h,N}(\omega)$ *et de* $\lim_{h\to 0} A_\star^{h,N}(\omega)$ *sont des questions différentes. Dans un cadre monodimensionnel, on montre que* $\lim_{N\to\infty}\lim_{h\to 0} A_\star^{h,N}(\omega) = A_\star$ *presque sûrement. L'étude, dans un cadre multidimensionnel, de la convergence avec* N *de* $\lim_{h\to 0} A_\star^{h,N}(\omega)$ *fera l'objet d'une publication future (cf. [23] pour une étude similaire). D'autre part, les résultats numériques de la section suivante sont consistants avec le fait que* $\lim_{h\to 0}\lim_{N\to\infty} A_\star^{h,N}(\omega) = A_\star$*. Le cas de* $A^{1,h,N}(\omega)$ *est différent, car les problèmes (6.6) et (6.12) ne sont pas des problèmes de correcteur standards. Cependant, les mêmes remarques s'appliquent.*

En l'absence de discrétisation, on a vu à la section précédente qu'il n'est pas nécessaire de connaître w_p^1 pour calculer A^1, et que son espérance $\overline{w}_p^1$, solution d'un problème plus simple, suffit.

Dans le cas discrétisé ($h>0$ et $N<\infty$), cette observation subsiste. En effet, introduisons $\overline{w}_p^{1,h,N} = \mathbb{E}(w_p^{1,h,N})$, qui vérifie, pour tout $v_h \in V_h^{\mathrm{per}}(Q_N)$,

$$\begin{aligned}\int_{Q_N}(\nabla v_h)^T A\,\nabla\overline{w}_p^{1,h,N} &= \int_{Q_N}(\nabla v_h)^T A\,\mathbb{E}(\nabla\Psi)\,\nabla w_p^{0,h} \\ &+ \int_{Q_N}(\nabla v_h)^T\left(\mathbb{E}\left(\nabla\Psi\right)^T - \mathbb{E}\left(\mathrm{div}\Psi\right)\mathrm{Id}\right)A\left(p+\nabla w_p^{0,h}\right).\end{aligned} \tag{6.15}$$

Comme A et $\mathbb{E}(\nabla\Psi)$ sont Q-périodiques, on obtient que la solution de (6.15) avec $N=1$, qui est Q-périodique, est solution de (6.15) pour tout $N\in\mathbb{N}^*$. Comme la solution de (6.15) est unique à l'ajout d'une constante aléatoire près, on obtient donc que $\overline{w}_p^{1,h,N}$ est en fait indépendant de N (on la note $\overline{w}_p^{1,h} = \mathbb{E}(w_p^{1,h,N})$), Q-périodique,

et solution de (6.15) avec $N = 1$, ce qui est une discrétisation du problème (6.7) convergente quand $h \to 0$. Comme fonction de $\nabla \overline{w}_p^{1,h}$, la matrice $A^{1,h} = \mathbb{E}(A^{1,h,N})$ est elle aussi indépendante de N, et se calcule facilement à partir de $\nabla \overline{w}_p^{1,h}$.

Cependant, pour la stricte question de l'*analyse* de la méthode, il va nous falloir utiliser $A^{1,h,N}$ plutôt que $A^{1,h}$. Introduisons en effet les deux estimateurs

$$e^{h,N}(\omega) := \eta^{-2}\left(A_\star^{h,N}(\omega) - A^{0,h} - \eta A^{1,h,N}(\omega)\right) \tag{6.16}$$

$$\overline{e}^{h,N}(\omega) := \eta^{-2}\left(A_\star^{h,N}(\omega) - A^{0,h} - \eta A^{1,h}\right) = e^{h,N}(\omega) + \eta^{-1}\left(A^{1,h,N}(\omega) - A^{1,h}\right) \tag{6.17}$$

En s'appuyant sur le fait que c'est le cas dans un cadre monodimensionnel, on suppose que la constante $C(h, N, \omega)$ qui apparaît dans (6.14) a une espérance et une variance finies. Comme $A^{1,h} = \mathbb{E}(A^{1,h,N})$, on obtient donc que l'espérance de $\overline{e}^{h,N}(\omega)$ est bornée quand $\eta \to 0$. Cependant, la relation (6.17) implique que

$$\eta^{-2}\ \mathbb{V}\mathrm{ar}\left(A^{1,h,N}(\omega) - A^{1,h}\right) \leq 2\mathbb{V}\mathrm{ar}\ \left(\overline{e}^{h,N}\right) + 2\mathbb{V}\mathrm{ar}\left(e^{h,N}\right).$$

La variance de $\overline{e}^{h,N}$ diverge donc lorsque $\eta \to 0$. Il s'ensuit donc que, dans un but d'*analyse*, on travaille ci-dessous avec l'estimateur $e^{h,N}$. Ceci va permettre de vérifier que $A^{0,h} + \eta A^{1,h,N}$ est une bonne estimation de $A_\star^{h,N}$, à l'ordre 2 en η. La *pratique* numérique, elle, est différente : elle consistera en effet à approcher $A_\star$, et on utilisera plutôt la quantité $A^{0,h} + \eta A^{1,h}$, beaucoup plus facile à calculer.

Remarque 6.3.2 *La Proposition 6.3.1 montre que $e^{h,N}(\omega)$ est borné, quand $\eta \to 0$, par une constante $C(h, N, \omega)$ qui dépend* a priori *de h, N et ω. Dans un cadre monodimensionnel, on montre que cette constante est bornée uniformément en h, N et ω, sous l'hypothèse que $\underset{\omega\in\Omega,\, x\in\mathbb{R}^d}{\mathrm{EssSup}}\ (|\nabla\Psi(x,\omega)|) = \overline{M} < \infty$. Dans un cadre multidimensionnel, les résultats numériques de la section suivante sont consistants avec une telle borne uniforme sur $C(h, N, \omega)$.*

6.4 Résultats numériques

On se place en deux dimensions : $x = (x_1, x_2)$, et on se donne deux familles $(X_k)_{k\in\mathbb{Z}}$ et $(Y_k)_{k\in\mathbb{Z}}$ de variables aléatoires indépendantes et identiquement distribuées, de même loi uniforme $\mathcal{U}([a, b])$. On considère le difféomorphisme $\Phi(x) = x + \eta\Psi(x, \omega)$, avec $\Psi(x, \omega) = (\psi_X(x_1, \omega), \psi_Y(x_2, \omega))$, où ψ_X est défini par

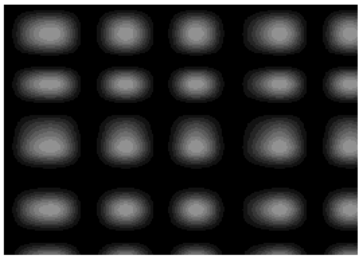
$$\psi_X(x_1, \omega) = \sum_{k\in\mathbb{Z}} 1_{[k,k+1[}(x_1)\left(\sum_{q=0}^{k-1} X_q(\omega) + 2X_k(\omega)\int_k^{x_1} \sin^2(2\pi t)\, dt\right),$$

et de même pour ψ_Y. La matrice périodique A est définie par

$$\forall x \in Q,\ A(x) = a_{\mathrm{per}}(x)\,\mathrm{Id}_2, \quad a_{\mathrm{per}}(x_1, x_2) = \beta + (\alpha - \beta)\sin^2(\pi x_1)\sin^2(\pi x_2).$$

Ceci revient donc à considérer un matériau $\mathbb{Z}^2$-périodique, où la conductivité varie de façon régulière de α à $\beta \leq \alpha$. Dans la cellule Q, la conductivité est maximale au centre, et minimale sur le bord de Q. L'application ψ_X n'est pas stationnaire, mais sa dérivée l'est. Cet exemple est un cas typique de [20]. Il ne relève pas de la théorie classique de l'homogénéisation stochastique : $A \circ \Phi^{-1}$ n'est pas stationnaire.

On travaille avec $a = -2.25$, $b = 5.75$, $\alpha = 10$, et $\beta = 1$. La Figure 6.1 montre la valeur de $A \circ \Phi^{-1}(x, \omega)$, pour une réalisation particulière de l'aléa. Tous les résultats numériques ont été obtenus à l'aide du logiciel d'éléments finis FreeFem++. Avec les paramètres de discrétisation $h = 1/3$ et $N = 20$, on calcule $\left(A^{0,h}\right)_{11} \approx 2.809$ et $\left(A^{1,h,N}\right)_{11} \approx 3.065 \pm 0.113$ (cette valeur est la moyenne sur 10 réalisations indépendantes, avec un intervalle de confiance à 95 %). Les résultats pour $\left(A^{h,N}_\star\right)_{11}$ et $\left(e^{h,N}\right)_{11}$ sont rassemblés dans la Figure 6.1. On vérifie bien que, dans la limite $\eta \to 0$, l'estimateur $e^{h,N}$ reste borné en espérance et en variance, par une petite constante, d'où l'efficacité avérée de la méthode. Par ailleurs, pour $\eta = 10^{-4}$, on constate que $e^{h,N}$ est bien borné indépendamment de h et de N : pour $h = 0.05$ et $N = 20$, on a $\left(e^{h,N}\right)_{11} = -2.254 \pm 0.149$, tandis que, pour $h = 1/3$ et $N = 40$, on a $\left(e^{h,N}\right)_{11} = -4.411 \pm 0.207$ (La variation de $e^{h,N=20}$ avec h montre que la valeur $h = 1/3$ est une valeur grossière, comparée par exemple à la valeur $h = 0.05$. Nous avons fait le choix $h = 1/3$ afin que les calculs numériques aient un coût raisonnable. De plus, la convergence des résultats avec h est une question classique, qui n'est pas l'objet de la présente Note).



η	$\left(A^{h,N}_\star\right)_{11}$	$\left(e^{h,N}\right)_{11}$
0.1	3.073 $\pm$ 0.00928	- 4.233 $\pm$ 0.216
0.01	2.839 $\pm$ 0.00111	- 5.009 $\pm$ 0.254
0.001	2.812 $\pm$ 0.000113	- 5.104 $\pm$ 0.259
0.0001	2.809 $\pm$ 0.0000113	- 5.114 $\pm$ 0.259

FIG. 6.1 – A gauche : valeur de $A \circ \Phi^{-1}(x, \omega)$, pour une réalisation de l'aléa, sur le domaine $Q_{N=5}$ ($\eta = 0.05$). A droite : valeurs de $\left(A^{h,N}_\star\right)_{11}$ et $\left(e^{h,N}\right)_{11}$, en fonction de η, pour $h = 1/3$ et $N = 20$. Les résultats pour les autres composantes des matrices sont similaires.

Chapitre 7

Estimation de l'erreur dans l'approximation numérique de deux modèles faiblement stochastiques

Cet article fera l'objet d'une prochaine publication dans AMRX (cf [P6]). Il a pour objet l'étude rigoureuse du développement limité de la matrice homogénéisée dans deux cadres perturbatifs, ou faiblement stochastiques. On s'intéresse dans un premier temps au cas d'une perturbation linéaire $A_\eta^\varepsilon(x,\omega) = A_{\text{per}}(x/\varepsilon) + \eta A_1(x/\varepsilon,\omega) + O(\eta^2)$, où A_1 est supposée stationnaire. Enfin, on traite, comme annoncé dans [P1], le cas où $A_\eta^\varepsilon(x,\omega) = A_{\text{per}}\left(\Phi_\eta^{-1}(x/\varepsilon,\omega)\right)$ où $\Phi_\eta(x,\omega) = x+\eta\Psi(x,\omega)+O(\eta^2)$ est une petite perturbation de l'identité.

Asymptotic expansion of the homogenized matrix in two weakly stochastic homogenization settings

Ronan Costaouec[a]

[a] *École Nationale des Ponts et Chaussées, 6 & 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2 and INRIA Rocquencourt, MICMAC team-project, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*

costaour@cermics.enpc.fr

7.1 Introduction

The context of this work is the homogenization of stochastic linear elliptic partial differential equations in divergence form

$$\left\{ \begin{array}{r} -\mathrm{div}\left(A_\eta\left(\frac{\cdot}{\varepsilon},\omega\right)\nabla u_\eta^\varepsilon(\cdot,\omega)\right) = f \text{ on } \mathcal{D}, \\ u_\eta^\varepsilon(\cdot,\omega) = 0 \text{ on } \partial\mathcal{D}, \end{array} \right. \tag{7.1}$$

where $\mathcal{D}$ denotes a bounded domain of $\mathbb{R}^d$, $d \in \mathbb{N}^*$ being the ambient dimension, and $f \in L^2(\mathcal{D})$. The matrix A_η is random, symmetric, uniformly bounded and coercive, that is :

$$\left\{ \begin{array}{l} \exists \gamma > 0,\ \forall \xi \in \mathbb{R}^d,\ \xi^T A_\eta(x,\omega)\xi \geq \gamma|\xi|^2 \text{ almost everywhere and almost surely,} \\ \exists M > 0 \text{ such that } \|A_\eta\|_{\left(L^\infty(\mathbb{R}^d\times\Omega)\right)^{d\times d}} \leq M, \end{array} \right. \tag{7.2}$$

where γ and M do not depend on η. Under additional stationarity hypotheses (the sense of which will be made precise below), it is classical that, when $\varepsilon \to 0$, the *random* solutions u_η^ε of (7.1) converge in some appropriate sense to a *deterministic* function $u_\eta^\star$ solution to the following problem

$$\left\{ \begin{array}{r} -\mathrm{div}\left(A_\eta^\star \nabla u_\eta^\star\right) = f \text{ on } \mathcal{D}, \\ u_\eta^\star = 0 \text{ on } \partial\mathcal{D}, \end{array} \right. \tag{7.3}$$

where $A_\eta^\star$ is the constant homogenized matrix. To compute $A_\eta^\star$, one needs first to compute some random functions, the so-called correctors w_p^η (where $p \in \mathbb{R}^d$), that

are solutions to random auxiliary problems, namely the *corrector problems* posed on the whole space $\mathbb{R}^d$. As a consequence, solving these corrector problems, and hence computing $A_\eta^\star$, is in general computationally challenging. Appropriate truncations have to be considered, and the standard numerical strategies lead to extremely expensive computations. In this article, we do not deal with the very general framework of stochastic homogenization. We rather focus on the specific case when the matrix A_η is a *small* perturbation of an underlying periodic matrix denoted by $A_{\rm per}$. The parameter η somehow encodes the degree of randomness of A_η : the smaller η, the less random A_η is. When $\eta = 0$, the matrix A_η is deterministic. There are several ways to formalize the notion of a small perturbation in the context of stochastic homogenization (see [P5, P3, P1, 7, 8, 19, 20]). In this article, we will consider the following two models (for which all the details will be provided in Section 7.1 below) :

- **Model 1** [Standard discrete stationary setting for stochastic homogenization] : the matrix $A_\eta(x,\omega)$ in (7.1) is stationary, and the small perturbation is linear (see [P3]) :
$$A_\eta(x,\omega) = A_{\rm per}(x) + \eta A_1(x,\omega) + O(\eta^2),$$
where the expansion holds in some appropriate functional space.
- **Model 2** [Setting introduced in [19]] : the matrix reads
$$A_\eta(x,\omega) = A_{\rm per}\left(\Phi_\eta^{-1}(x,\omega)\right),$$
where Φ_η is referred to as a stochastic diffeomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$ and $A_{\rm per}$ is a periodic matrix. An important point is that in this case A_η is not necessarily stationary. The corresponding perturbative approach has been developed in [20]. The stochastic diffeomorphism Φ_η is supposed to be a linear perturbation of the identity, namely
$$\Phi_\eta(x,\omega) = x + \eta\Psi(x,\omega) + O(\eta^2),$$
in some appropriate functional space.

In both settings, it has been shown that the gradients of the correctors ∇w_p^η and the homogenized matrix $A_\eta^\star$ possess an expansion in powers of η : $\nabla w_p^\eta = \nabla w_p^0 + \eta\nabla w_p^1 + O(\eta^2)$ and $A_\eta^\star = A_{\rm per}^\star + \eta A_1^\star + O(\eta^2)$. Approximating $A_\eta^\star$ by $A_{\rm per}^\star + \eta A_1^\star$, one thus makes an error of the order of η^2. The functions w_p^0 and w_p^1 are solutions to a deterministic and a random problem respectively. The dominant term $A_{\rm per}^\star$ in the expansion of $A_\eta^\star$ is simply the homogenized matrix associated to $A_{\rm per}$. The definition of $A_1^\star$ involves only ∇w_p^0 and $\nabla\mathbb{E}\left(w_p^1\right)$. Since $\mathbb{E}\left(w_p^1\right)$ can be shown to solve a deterministic problem, the computation of the first order approximation $A_{\rm per}^\star + \eta A_1^\star$ only requires the solution of two *deterministic* partial differential equations. This is far less demanding that solving a single *stochastic* partial differential equation. This is the major advantage of these perturbative approaches.

Of course, in practice, one cannot exactly solve the corrector problems that are posed on an unbounded domain. As will be seen below, the classical discretization approach consists in two steps : truncation of the corrector problem on a bounded computational domain of size N and finite elements approximation with mesh size h. This introduces two types of error related to the parameters N and h, respectively. Because of the truncation, the approximated homogenized matrix $A_\eta^{\star,h,N}$ is random. We will see that it also possesses an expansion in the variable η, almost surely. The main difference between this expansion and the exact one is that, at the discrete level, the difference $A_\eta^{\star,h,N} - A_{\rm per}^{\star,h,N} - \eta A_1^{\star,h,N}$ is random in nature, and so is the second order error estimate. Therefore, in order to ensure the relevance of the first order approximation of $A_\eta^{\star,h,N}$, one needs to bound the random error $\eta^{-2}(A_\eta^{\star,h,N} - A_{\rm per}^{\star,h,N} - \eta A_1^{\star,h,N})$ of second order, in some appropriate probability space, namely $(L^\infty(\Omega))^{d\times d}$ in the sequel. In the setting of Model 2, this question has already been addressed from a numerical point of view in [P1]. The aim of the present work is to theoretically derive rigorous bounds on the second order error for both models and to understand how these bounds depend on the parameters of the discretization procedure. To this end, we first have to make precise the appropriate functional spaces in which the original expansions hold. We next deduce for each model in which sense the second order error terms are bounded. Our presentation elaborates on the previous works [19, 20]. We only provide the formalism for self-consistency and we refer to [20] for more details.

7.1.1 Probabilistic setting

Throughout this article, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and we denote by $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ the expectation value of any random variable $X \in L^1(\Omega, d\mathbb{P})$. We assume that the group $(\mathbb{Z}^d, +)$ acts on Ω. We denote by $(\tau_k)_{k\in\mathbb{Z}^d}$ this action, and assume that it preserves the measure $\mathbb{P}$, that is, for all $k \in \mathbb{Z}^d$ and all $A \in \mathcal{F}$, $\mathbb{P}(\tau_k A) = \mathbb{P}(A)$. We assume that the action τ is *ergodic*, that is, if $A \in \mathcal{F}$ is such that $\tau_k A = A$ for any $k \in \mathbb{Z}^d$, then $\mathbb{P}(A) = 0$ or 1. In addition, we define the following discrete notion of stationarity (see [19]) : any $F \in L^1_{\rm loc}\left(\mathbb{R}^d, L^1(\Omega)\right)$ is said to be *stationary* if, for all $k \in \mathbb{Z}^d$,

$$F(x + k, \omega) = F(x, \tau_k \omega), \text{ almost everywhere and almost surely.} \tag{7.4}$$

Note that this setting is a straightforward classical variant of the more commonly used (continuous) stationary setting for random homogenization, for which the shift τ is indexed by elements of the group $\mathbb{R}^d$ and (7.4) holds for all $k \in \mathbb{R}^d$ instead of $\mathbb{Z}^d$. In our discrete setting, the ergodic theorem [61, 89] can be stated as follows :

Theorem 7.1.1 *Let $F \in L^\infty\left(\mathbb{R}^d, L^1(\Omega)\right)$ be a stationary random variable in the*

above sense. For $k = (k_1, k_2, \dots k_d) \in \mathbb{Z}^d$*, we set* $|k|_\infty = \sup_{1 \le i \le d} |k_i|$*. Then*

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \le N} F(x, \tau_k \omega) \underset{N \to \infty}{\longrightarrow} \mathbb{E}\left(F(x, \cdot)\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

This implies that (denoting by Q *the unit cube in* $\mathbb{R}^d$*)*

$$F\left(\frac{x}{\varepsilon}, \omega\right) \overset{*}{\underset{\varepsilon \to 0}{\rightharpoonup}} \mathbb{E}\left(\int_Q F(x, \cdot) dx\right) \quad \text{in } L^\infty(\mathbb{R}^d), \text{ almost surely.}$$

7.1.2 Stochastic homogenization results

Standard results of stochastic homogenization (see [16, 58]) apply to problem (7.1) when the matrix A_η is stationary. For any fixed value of η, they provide the following result.

Theorem 7.1.2 *We consider Model 1. Suppose that* A_η *is a symmetric matrix, stationary in the sense of* (7.4)*, uniformly bounded and coercive in the sense of* (7.2)*. Then the homogenized matrix* $A_\eta^\star$ *appearing in* (7.3) *is defined by*

$$\left[A_\eta^\star\right]_{ij} = \mathbb{E}\left[\int_Q e_i^T A_\eta \left(e_j + \nabla w_{e_j}^\eta\right)\right], \tag{7.5}$$

where, for any $p \in \mathbb{R}^d$*,* w_p^η *denotes the unique solution (up to the addition of a random constant) of the corrector problem*

$$\begin{cases} -\mathrm{div}\left(A_\eta\left(p + \nabla w_p^\eta\right)\right) = 0 \text{ on } \mathbb{R}^d, \text{ almost surely,} \\ \nabla w_p^\eta \text{ is stationary in the sense of (7.4)}, \\ \mathbb{E}\left[\displaystyle\int_Q \nabla w_p^\eta\right] = 0. \end{cases} \tag{7.6}$$

In the sequel we will always assume that $|p| = 1$.

In the case $A_\eta(x, \omega) = A_{\mathrm{per}}\left(\Phi_\eta^{-1}(x, \omega)\right)$, there exists an analogous result due to Blanc, Le Bris and Lions (see [19]). Its statement requires to make precise the notion of *stochastic diffeomorphism* mentioned above. The map Φ_η is said to be a stochastic diffeomorphism if it satisfies

$$\Phi_\eta(\cdot, \omega) \text{ is a diffeomorphism almost surely,} \tag{7.7}$$

$$\operatorname*{Ess\,inf}_{x \in \mathbb{R}^d, \omega \in \Omega} \left(\det\left(\nabla \Phi_\eta(x, \omega)\right)\right) = \nu > 0, \tag{7.8}$$

$$\operatorname*{Ess\,sup}_{x \in \mathbb{R}^d, \omega \in \Omega} \left(|\nabla \Phi_\eta(x, \omega)|\right) = M' < \infty, \tag{7.9}$$

$$\nabla \Phi_\eta \text{ is stationary in the sense of (7.4).} \tag{7.10}$$

Under these hypotheses on Φ_η, the following theorem gives the homogenized problem associated to (7.1).

Theorem 7.1.3 *We consider Model 2. Suppose that* $A_\eta = A_{\rm per} \circ \Phi_\eta^{-1}$ *where* $A_{\rm per}$ *is a periodic symmetric matrix, uniformly bounded and coercive, and* Φ_η *satisfies* (7.7), (7.8), (7.9) *and* (7.10). *Then the homogenized matrix* $A_\eta^\star$ *appearing in* (7.3) *is defined by*

$$\left[A_\eta^\star\right]_{ij} = \left(\mathbb{E}\left[\int_Q \det\left(\nabla\Phi_\eta\right)\right]\right)^{-1} \mathbb{E}\left[\int_Q \det\left(\nabla\Phi_\eta\right) e_i^T A_{\rm per}\left(e_j + \left(\nabla\Phi_\eta\right)^{-1}\nabla w_{e_j}^\eta\right)\right], \tag{7.11}$$

where, for any $p \in \mathbb{R}^d$, w_p^η *denotes the unique solution (up to the addition of a random constant) of the corrector problem*

$$\begin{cases} -\mathrm{div}\left(\det\left(\nabla\Phi_\eta(\cdot,\omega)\right)\left(\nabla\Phi_\eta(\cdot,\omega)\right)^{-T} A_{\rm per}\left(p + \left(\nabla\Phi_\eta(\cdot,\omega)\right)^{-1}\nabla w_p^\eta(\cdot,\omega)\right)\right) = 0, \\ \text{on} \quad \mathbb{R}^d, \ \text{almost surely}, \\ \nabla w_p^\eta \ \text{is stationary in the sense of (7.4)}, \\ \mathbb{E}\left[\int_Q \nabla w_p^\eta\right] = 0. \end{cases} \tag{7.12}$$

7.1.3 Standard numerical approximation

In practice, problems (7.6) and (7.12) are solved numerically. The first step is to introduce a truncation. Following [23], we approximate (7.6) by

$$\begin{cases} -\mathrm{div}\left(A_\eta(\cdot,\omega)\left(p + \nabla w_p^{\eta,N}(\cdot,\omega)\right)\right) = 0 \text{ on } Q_N \text{ almost surely}, \\ w_p^{\eta,N}(\cdot,\omega) \ \text{ is } Q_N - \text{periodic}, \end{cases} \tag{7.13}$$

where $Q_N = \left[-N - \frac{1}{2}, N + \frac{1}{2}\right]^d$. We consider an analogous truncated problem for (7.12). A classical finite element method is then used to approximate the solutions of (7.13). We consider a periodic triangulation $\mathcal{T}_h^{(Q)}$ of the unit cell $Q = \left[-\frac{1}{2}, \frac{1}{2}\right]^d$. Replicating it, we obtain a triangulation

$$\mathcal{T}_h = \cup_{k\in\mathbb{Z}^d}\left(k + \mathcal{T}_h^{(Q)}\right)$$

of $\mathbb{R}^d$. We denote by $V_h^{\rm per}(Q_N)$ the space of functions φ_h defined on $\mathbb{R}^d$, Q_N-periodic, whose restriction to Q_N is in a $\mathbb{P}_1$-Lagrange finite element space built from $\mathcal{T}_h^N = \mathcal{T}_h \cap Q_N$, and which satisfy

$$\int_{Q_N} \varphi_h = 0.$$

Let $\{\phi_k\}_{1\le k\le N_v}$ be a basis of $V_h^{\rm per}\left(Q_N\right)$:

$$V_h^{\rm per}\left(Q_N\right) = \mathrm{span}\left(\{\phi_k\}_{1\le k\le N_v}\right),$$

where $N_v = N_v(N)$ is the number of degrees of freedom considered.

In the standard case (Model 1), we define the *approximated corrector* $w_p^{\eta,h,N}$ as the solution to the variational formulation

$$\begin{cases} \text{Find } w_p^{\eta,h,N}(\cdot,\omega) \in V_h^{\text{per}}(Q_N) \text{ such that,} \\ \forall \varphi_h \in V_h^{\text{per}}(Q_N), \quad \displaystyle\int_{Q_N} A_\eta(\cdot,\omega)\left(p + \nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h = 0 \text{ almost surely,} \end{cases} \tag{7.14}$$

and the *approximated homogenized matrix* by

$$\forall 1 \leq i,j \leq d, \quad \left[A_\eta^{\star,h,N}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_\eta(\cdot,\omega)\left(e_j + \nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)\right). \tag{7.15}$$

In the second setting (Model 2), we similarly define the approximated corrector as the solution to

$$\begin{cases} \text{Find } w_p^{\eta,h,N}(\cdot,\omega) \in V_h^{\text{per}}(Q_N) \text{ such that, for any } \varphi_h \in V_h^{\text{per}}(Q_N), \\ \displaystyle\int_{Q_N} \det(\nabla\Phi_\eta(\cdot,\omega))(\nabla\Phi_\eta(\cdot,\omega))^{-T} A_{\text{per}}(\nabla\Phi_\eta(\cdot,\omega))^{-1}\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h \\ \displaystyle + \int_{Q_N} \det(\nabla\Phi_\eta(\cdot,\omega))(\nabla\Phi_\eta(\cdot,\omega))^{-T} A_{\text{per}} p\cdot\nabla\varphi_h = 0 \text{ almost surely,} \end{cases} \tag{7.16}$$

and the homogenized matrix is defined by

$$\left[A_\eta^{*,h,N}\right]_{ij}(\omega) = \frac{\displaystyle\int_{Q_N} \det\left(\nabla\Phi_\eta(\cdot,\omega)\right) e_i^T A_{\text{per}}\left(e_j + \left(\nabla\Phi_\eta(\cdot,\omega)\right)^{-1}\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)\right)}{\det\left(\displaystyle\int_{Q_N}\nabla\Phi_\eta(\cdot,\omega)\right)}. \tag{7.17}$$

Note that for both (7.14) and (7.16) the solution $w_p^{\eta,h,N}$ is a *random* field. It follows that both (7.15) and (7.17) are *random* matrices. It is only in the limit $N = \infty$ that these objects become *deterministic*. In the numerical practice, one then commonly considers that the best approximation of $A_\eta^\star$ is given by $\mathbb{E}\left(A_\eta^{*,h,N}\right)$ which is in turn estimated using an empirical mean, computed with standard Monte-Carlo methods.

7.2 Expansion of the random matrix A_η (Model 1)

In this section, we consider the weakly stochastic setting of Model 1. We will turn our attention to Model 2 in Section 3. Our main result for Model 1 is Proposition 7.2.1. As announced above, it makes precise the behaviour of the *random* second order error in the expansion of the approximated homogenized matrix $A_\eta^{\star,h,N}$, under appropriate hypotheses. Then, passing to the limit $h \to 0$, we prove that this result extends to the approximated homogenized matrix when only truncation is taken into account. Finally, letting N go to infinity, we recover the expansion of the

exact homogenized matrix $A_\eta^\star$ derived in [P3]. The functional setting described below is simple and the arguments we use in the proof of Proposition 2.1 are standard. Our aim is to illustrate in a simple and relevant framework that the randomness of the second order error in the expansions of $\nabla w_p^{\eta,h,N}$ and $A_\eta^{\star,h,N}$ does not affect the validity of the approximation $A_{\rm per}^{\star,h}+\eta\mathbb{E}(A_1^{\star,h,N})$. In particular, this amounts to prove that the quantity $\eta^{-2}(A_\eta^{\star,h,N}-A_{\rm per}^{\star,h}-\eta A_1^{\star,h,N})$ is bounded independently of h, N and η in some probability space. This is precisely what ensures Proposition 2.1 below.

7.2.1 Assumptions

We suppose that the matrix A_η admits in $\left(L^\infty(Q\times\Omega)\right)^{d\times d}$ the expansion

$$A_\eta(x,\omega)=A_{\rm per}(x)+\eta A_1(x,\omega)+R_\eta(x,\omega), \tag{7.18}$$

where $A_{\rm per}\in\left(L^\infty(Q)\right)^{d\times d}$, $A_1\in\left(L^\infty(Q\times\Omega)\right)^{d\times d}$ is stationary and the stationary matrix R_η satisfies $R_\eta=O(\eta^2)$ in $\left(L^\infty(Q\times\Omega)\right)^{d\times d}$. This means that

$$\lim_{\eta\to 0}\left\|A_\eta-A_{\rm per}\right\|_{\left(L^\infty(Q\times\Omega)\right)^{d\times d}} = 0, \tag{7.19}$$

$$\lim_{\eta\to 0}\left\|\eta^{-1}\left(A_\eta-A_{\rm per}\right)-A_1\right\|_{\left(L^\infty(Q\times\Omega)\right)^{d\times d}} = 0, \tag{7.20}$$

and there exists a *deterministic* constant C_R independent of η such that, when $|\eta|\leq 1$,

$$\eta^{-2}\left\|R_\eta\right\|_{\left(L^\infty(Q\times\Omega)\right)^{d\times d}}\leq C_R, \tag{7.21}$$

where we recall that

$$\|v\|_{\left(L^\infty(Q\times\Omega)\right)^{d\times d}}=\operatorname*{Ess\,sup}_{(x,\omega)\in Q\times\Omega}\left|v(x,\omega)\right|.$$

From (7.2), we know that A_η is bounded and uniformly coercive. Using (7.19) and (7.2), we see that

$$\left\|A_{\rm per}\right\|_{\left(L^\infty(Q)\right)^{d\times d}}=\lim_{\eta\to 0}\left\|A_\eta\right\|_{\left(L^\infty(Q\times\Omega)\right)^{d\times d}}\leq M, \tag{7.22}$$

and, for any $\xi\in\mathbb{R}^d$, we have

$$\xi^T A_{\rm per}(x)\xi=\lim_{\eta\to 0}\xi^T A_\eta(x,\omega)\xi\geq\gamma\left|\xi\right|^2 \text{ almost everywhere.} \tag{7.23}$$

The matrix $A_{\rm per}$ is thus also uniformly bounded and coercive.

Assuming that the expansion of A_η holds in $\left(L^\infty\left(Q\times\Omega\right)\right)^{d\times d}$ is relevant from the point of view of modeling, and it somehow simplifies the proof of Proposition 2.1. It is nevertheless to be mentioned that our proof does not directly extend to the case when the expansion of A_η holds in spaces of the form $\left(L^\infty\left(Q;L^p(\Omega)\right)\right)^{d\times d}$ with $0<p<+\infty$, which models the idea of possibly large but rare local perturbations (see [P5, 7, 8] for a detailed presentation of this latter model).

7.2.2 Formal expansion

Following the method introduced in [20], we first postulate the formal expansion of the solution to (7.14) :

$$w_p^{\eta,h,N} = w_p^{0,h,N} + \eta w_p^{1,h,N} + r_p^{\eta,h,N}, \tag{7.24}$$

where $\nabla r_p^{\eta,h,N} = O(\eta^2)$ in some appropriate space. We will sucessively identify $w_p^{0,h,N}$ and $w_p^{1,h,N}$ and prove the validity of the expansion. Formally inserting this expression in (7.14), we obtain that the function $w_p^{0,h,N}$ is independent of N (we denote it by $w_p^{0,h}$ in the sequel) and solves

$$\left\{ \begin{array}{l} \text{Find } w_p^{0,h} \in V_h^{\text{per}}(Q) \text{ such that,} \\ \forall \varphi_h \in V_h^{\text{per}}(Q), \quad \displaystyle\int_Q A_{\text{per}} \left(p + \nabla w_p^{0,h}\right) \cdot \nabla \varphi_h = 0, \end{array} \right. \tag{7.25}$$

and that the function $w_p^{1,h,N}$ is solution to

$$\left\{ \begin{array}{l} \text{Find } w_p^{1,h,N}(\cdot,\omega) \in V_h^{\text{per}}(Q_N) \text{ such that, } \forall \varphi_h \in V_h^{\text{per}}(Q_N), \\ \displaystyle\int_{Q_N} A_{\text{per}} \nabla w_p^{1,h,N}(\cdot,\omega) \cdot \nabla \varphi_h + \int_{Q_N} A_1(\cdot,\omega) \left(p + \nabla w_p^{0,h}\right) \cdot \nabla \varphi_h = 0 \quad \text{a. s.} \end{array} \right. \tag{7.26}$$

In addition, substituting the expansions of A_η and $w_p^{\eta,h,N}$ into (7.15), we formally obtain

$$A_\eta^{\star,h,N}(\omega) = A_{\text{per}}^{\star,h} + \eta A_1^{\star,h,N}(\omega) + O(\eta^2), \tag{7.27}$$

where the terms of order zero and one are respectively defined by

$$\left[A_{\text{per}}^{\star,h}\right]_{ij} = \int_Q e_i^T A_{\text{per}} \left(e_j + \nabla w_{e_j}^{0,h}\right), \tag{7.28}$$

$$\begin{aligned} \left[A_1^{\star,h,N}\right]_{ij}(\omega) &= \frac{1}{|Q_N|} \int_{Q_N} e_i^T A_{\text{per}} \nabla w_{e_j}^{1,h,N}(\cdot,\omega) \\ &+ \frac{1}{|Q_N|} \int_{Q_N} e_i^T A_1(\cdot,\omega) \left(\nabla w_{e_j}^{0,h} + e_j\right). \end{aligned} \tag{7.29}$$

In the sequel, we will make precise and rigorously justify the expansions (7.24) and (7.27).

7.2.3 Main result

Our main result in this section is the following.

Proposition 7.2.1 *Suppose that A_η is a symmetric matrix that satisfies (7.2) and is stationary in the sense of* (7.4)*. Suppose, in addition, that it satisfies (7.18),*

(7.19), (7.20). We assume that (7.21) holds, namely the second order error is $O(\eta^2)$ in $(L^\infty(Q\times\Omega))^{d\times d}$. Then there exists a constant C independent of η, ω, N and h, such that, for $|\eta|\leq 1$,

$$\eta^{-2}\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)-\nabla w_p^{0,h}-\eta\nabla w_p^{1,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\leq C\sqrt{|Q_N|}\quad \text{almost surely,} \tag{7.30}$$

where $w_p^{\eta,h,N}$, $w_p^{0,h}$ and $w_p^{1,h,N}$ are solutions to (7.14), (7.25) and (7.26), respectively, and such that

$$\eta^{-2}\left|A_\eta^{\star,h,N}(\omega)-A_{\rm per}^{\star,h}-\eta A_1^{\star,h,N}(\omega)\right|\leq C\quad \text{almost surely,} \tag{7.31}$$

where $A_\eta^{\star,h,N}$, $A_{\rm per}^{\star,h}$ and $A_1^{\star,h,N}$ are defined by (7.15), (7.28) *and* (7.29).

Note that, as the constant C in (7.30) and (7.31) is independent of ω, the expansions of $\nabla w_p^{\eta,h,N}$ and $A_\eta^{\star,h,N}$ hold in $\left(L^\infty\left(\Omega;L^2(Q_N)\right)\right)^d$ and $(L^\infty(\Omega))^{d\times d}$, respectively.

Proof: Our goal is to prove estimates (7.30) and (7.31). To do so, following the methodology of [20], we begin with justifying the expansion (7.24) in our discrete framework. This is the purpose of Steps 1, 2 and 3. First we will check that $\nabla w_p^{\eta,h,N}$ is bounded independently of η,ω,N and h and that it converges to the gradient $\nabla w_p^{0,h}$ of a *deterministic* function (Step 1). We will then verify that the first order error term of $\nabla w_p^{\eta,h,N}$, namely $\nabla v_p^{\eta,h,N}$ defined below, is bounded independently of η,ω,N and h and that it converges to the gradient $\nabla w_p^{1,h,N}$ of a *random* function (Step 2). In Step 3, we prove that the second order error of $\nabla w_p^{\eta,h,N}$, namely $\nabla z_p^{\eta,h,N}$ defined below, is bounded independently of η,ω,N and h. Using bounds derived at Steps 1 and 2, this will prove (7.30). In Step 4, remarking that the second order error term of the homogenized matrix $A_\eta^{\star,h,N}$ depends on $\nabla w_p^{\eta,h,N}$, $\nabla v_p^{\eta,h,N}$ and $\nabla z_p^{\eta,h,N}$, we use the bounds from Steps 1, 2 and 3 to prove (7.31).

Step 1 : Our goal is first to prove that the gradient $\nabla w_p^{\eta,h,N}$ is bounded in $\left(L^2(Q_N)\right)^d$, almost surely and independently of η,ω,N and h, and next to show that $w_p^{\eta,h,N}$ converges to a *deterministic* function $w_p^{0,h}$ independent of N. Choosing $\varphi_h=w_p^{\eta,h,N}$ as test function in (7.14) and using (7.2), we obtain that

$$\gamma\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|^2_{(L^2(Q_N))^d}\leq\left(\int_{Q_N}|A_\eta(\cdot,\omega)p|^2\right)^{\frac12}\left(\int_{Q_N}|\nabla w_p^{\eta,h,N}|^2\right)^{\frac12},$$

which implies that (recall that $|p|=1$)

$$\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\leq\frac{\|A_\eta\|_{(L^\infty(Q\times\Omega))^{d\times d}}}{\gamma}|Q_N|^{\frac12}.$$

Using (7.2), we deduce that

$$\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\leq C_w|Q_N|^{\frac12}, \tag{7.32}$$

where

$$C_w = \frac{M}{\gamma}$$

is a constant independent of η, ω, N and h. Thus $\nabla w_p^{\eta,h,N}(\cdot,\omega)$ is bounded uniformly in $(L^2(Q_N))^d$ independently from η and almost surely. We then remark that

$$\nabla w_p^{\eta,h,N}(\cdot,\omega) = \sum_{i=1}^{N_v} \alpha_{p,i}^{\eta}(\omega) \nabla \phi_i,$$

where $\alpha_p^\eta = (\alpha_{p,i}^\eta)_{1\leq i\leq N_v}$ denotes the coordinates of $w_p^{\eta,h,N}(\cdot,\omega)$ in the basis of $V_h^{\rm per}(Q_N)$. In the finite dimensional space $V_h^{\rm per}(Q_N)$, all norms are equivalent and we deduce from (7.32) that α_p^η is bounded independently of η in $\mathbb{R}^{N_v}$. Thus, up to extracting a subsequence, α_p^η converges to a vector α_p^0 in $\mathbb{R}^{N_v}$ almost surely. Consequently, we see that

$$\nabla w_p^{\eta,h,N}(\cdot,\omega) \underset{\eta\to 0}{\longrightarrow} \sum_{i=1}^{N_v} \alpha_{p,i}^0(\omega)\nabla\phi_i = \nabla\left(\sum_{i=1}^{N_v} \alpha_{p,i}^0(\omega)\phi_i\right) = \nabla w_p^{0,h,N}(\cdot,\omega) \qquad (7.33)$$

in $(L^2(Q_N))^d$ almost surely, where $w_p^{0,h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N)$ is defined by

$$w_p^{0,h,N}(\cdot,\omega) := \sum_{i=1}^{N_v} \alpha_{p,i}^0(\omega)\phi_i.$$

We now return to (7.14) which we decompose using the expansion (7.18) of A_η : for any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$\begin{aligned}&\int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h + \eta\int_{Q_N} A_1(\cdot,\omega)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\nabla\varphi_h \\ &+ \int_{Q_N} R_\eta(\cdot,\omega)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h = 0. \qquad (7.34)\end{aligned}$$

We are going to pass to the limit $\eta\to 0$ in (7.34). By definition of $\nabla w_p^{0,h,N}(\cdot,\omega)$,

$$\int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \underset{\eta\to 0}{\longrightarrow} \int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h$$

almost surely, and

$$\int_{Q_N} A_1(\cdot,\omega)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \underset{\eta\to 0}{\longrightarrow} \int_{Q_N} A_1(\cdot,\omega)\left(p+\nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h$$

almost surely. Using (7.21), we see that, when $|\eta|\leq 1$,

$$\begin{aligned}&\left|\int_{Q_N} R_\eta(\cdot,\omega)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h\right| \\ &\leq \|R_\eta\|_{(L^\infty(Q\times\Omega))^{d\times d}}\left\|p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\|\nabla\varphi_h\|_{(L^2(Q_N))^d} \\ &\leq C_R\eta^2\left\|p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\|\nabla\varphi_h\|_{(L^2(Q_N))^d} \\ &\leq C_R\eta^2\left(\|p\|_{(L^2(Q_N))^d} + C_w|Q_N|^{\frac12}\right)\|\nabla\varphi_h\|_{(L^2(Q_N))^d},\end{aligned}$$

where we recall that C_w is independent of η. We thus obtain that

$$\lim_{\eta\to 0}\int_{Q_N} R_\eta(\cdot,\omega)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h = 0.$$

Passing to the limit $\eta\to 0$ in (7.34), we obtain that $w_p^{0,h,N}(\cdot,\omega)$ satisfies

$$\forall\varphi_h\in V_h^{\rm per}(Q_N),\quad \int_{Q_N} A_{\rm per}\nabla w_p^{0,h,N}(\cdot,\omega)\cdot\nabla\varphi_h + \int_{Q_N} A_{\rm per}p\cdot\nabla\varphi_h = 0. \qquad (7.35)$$

Note that this equation has a unique solution in $V_h^{\rm per}(Q_N)$. We now show that $w_p^{0,h,N} = w_p^{0,h}$, where $w_p^{0,h}$ is the solution of (7.25). Indeed, for any $\varphi_h\in V_h^{\rm per}(Q_N)$, we see that

$$\begin{aligned}\int_{Q_N} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\varphi_h &= \sum_{|k|_\infty\le N}\int_{Q+k} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\varphi_h\\ &= \sum_{|k|_\infty\le N}\int_{Q} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\varphi_h(\cdot+k)\\ &= \int_{Q} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\left(\sum_{|k|_\infty\le N}\varphi_h(\cdot+k)\right).\end{aligned}$$

Let

$$\theta_h := \sum_{|k|_\infty\le N}\varphi_h(\cdot+k).$$

Observe that θ_h is Q-periodic, and that its restriction to Q is in the $\mathbb{P}_1$-Lagrange finite elements space built from $\mathcal{T}_h^{(Q)}$. In addition, we have

$$\int_{Q_N}\theta_h = 0.$$

Thus $\theta_h\in V_h^{\rm per}(Q)$. Hence by definition of $w_p^{0,h}$, we obtain, for any $\varphi_h\in V_h^{\rm per}(Q_N)$,

$$\int_{Q_N} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\varphi_h = \int_{Q} A_{\rm per}\left(\nabla w_p^{0,h}+p\right)\cdot\nabla\theta_h = 0.$$

In addition, we have $w_p^{0,h}\in V_h^{\rm per}(Q)\subset V_h^{\rm per}(Q_N)$. As a consequence, $w_p^{0,h}$ is solution to (7.35). We conclude that $w_p^{0,h,N}(\cdot,\omega) = w_p^{0,h}$.

Step 2 : We introduce the function

$$v_p^{\eta,h,N} = \frac{w_p^{\eta,h,N} - w_p^{0,h}}{\eta}.$$

We want to prove that $\nabla v_p^{\eta,h,N}$ is bounded in $(L^2(Q_N))^d$ almost surely and independently from η, h, N and ω, and that it converges to the gradient of a random function, namely $w_p^{1,h,N}$ solution to (7.26). Substracting equation (7.35) from (7.34), we obtain that $v_p^{\eta,h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N)$ is such that, for any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$\begin{aligned}&\int_{Q_N} A_{\rm per}\nabla v_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h + \\ &\int_{Q_N}\left(A_1(\cdot,\omega)+\eta^{-1}R_\eta(\cdot,\omega)\right)\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h = 0. \qquad (7.36)\end{aligned}$$

Choosing $\varphi_h = v_p^{\eta,h,N}(\cdot,\omega)$ as test function in (7.36) and using (7.23), we have

$$\begin{aligned}\gamma\left\|\nabla v_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} &\le \left(\int_{Q_N}\left|\left(A_1(\cdot,\omega)+\eta^{-1}R_\eta(\cdot,\omega)\right)p\right|^2\right)^{\frac12} \\ &\quad+\left(\int_{Q_N}\left|\left(A_1(\cdot,\omega)+\eta^{-1}R_\eta(\cdot,\omega)\right)\nabla w_p^{\eta,h,N}(\cdot,\omega)\right|^2\right)^{\frac12}, \\ &\le \left(\|A_1\|_{(L^\infty(Q\times\Omega))^{d\times d}}+\left\|\eta^{-1}R_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}}\right) \\ &\quad\times\left(\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}+|Q_N|^{\frac12}\right). \qquad (7.37)\end{aligned}$$

Using (7.21), we have, when $|\eta|\le 1$,

$$\left\|\eta^{-1}R_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}} \le \left\|\eta^{-1}R_\eta\right\|_{(L^\infty(Q\times\Omega))^{d\times d}} \le C_R|\eta| \le C_R. \qquad (7.38)$$

We then deduce from (7.37), (7.32) and (7.38) that, for any $|\eta|\le 1$,

$$\begin{aligned}\left\|\nabla v_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} &\le \left(\|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}}+C_R\right)\left(\frac{C_w}{\gamma}+\frac{1}{\gamma}\right)|Q_N|^{\frac12} \\ &\le C_v|Q_N|^{\frac12}, \qquad (7.39)\end{aligned}$$

where

$$C_v = \left(\|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}}+C_R\right)\left(\frac{M}{\gamma^2}+\frac{1}{\gamma}\right)$$

is a constant independent of η, h, N and ω. Thus $\nabla v_p^{\eta,h,N}(\cdot,\omega)$ is also bounded independently of η in $(L^2(Q_N))^d$ almost surely.

As above, we deduce that there almost surely exists $w_p^{1,h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N)$ such that, up to extracting a subsequence,

$$\nabla v_p^{\eta,h,N}(\cdot,\omega) \underset{\eta\to 0}{\longrightarrow} \nabla w_p^{1,h,N}(\cdot,\omega) \text{ in } \left(L^2(Q_N)\right)^d.$$

Using the same arguments as in Step 1, we can pass to the limit $\eta \to 0$ in (7.36). We obtain that $w_p^{1,h,N}(\cdot,\omega)$ satisfies

$$\forall \varphi \in V_h^{\rm per}(Q_N), \quad \int_{Q_N} A_{\rm per} \nabla w_p^{1,h,N}(\cdot,\omega) \cdot \nabla \varphi_h + \int_{Q_N} A_1(\cdot,\omega) \left(\nabla w_p^{0,h} + p\right) \cdot \nabla \varphi_h = 0,$$

almost surely. We thus recover (7.26).

Step 3 : Our purpose is now to obtain a second order approximation of $w_p^{\eta,h,N}$. To this end, we define

$$z_p^{\eta,h,N} = \frac{w_p^{\eta,h,N} - w_p^{0,h} - \eta w_p^{1,h,N}}{\eta^2} \in V_h^{\rm per}(Q_N).$$

Our goal is to prove that $\nabla z_p^{\eta,h,N}$ is bounded in $\left(L^2(Q_N)\right)^d$ almost surely and independently from η, N, h and ω. To do so, we will need the previous estimates (7.32) and (7.39) from Steps 1 and 2 respectively. Using (7.34), (7.35) and (7.26) we see that, for any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$\begin{aligned} &\int_{Q_N} A_{\rm per} \nabla z_p^{\eta,h,N}(\cdot,\omega) \cdot \nabla \varphi_h + \int_{Q_N} A_1(\cdot,\omega) \nabla v_p^{\eta,h,N}(\cdot,\omega) \cdot \nabla \varphi_h \\ &+ \int_{Q_N} \eta^{-2} R_\eta(\cdot,\omega) \left(p + \nabla w_p^{\eta,h,N}(\cdot,\omega)\right) \cdot \nabla \varphi_h = 0. \end{aligned} \tag{7.40}$$

We choose $\varphi_h = z_p^{\eta,h,N}(\cdot,\omega)$ as test function in (7.40). Using (7.23), we obtain

$$\begin{aligned} \gamma \left\|\nabla z_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} &\leq \left(\int_{Q_N} \left|A_1 \nabla v_p^{\eta,h,N}(\cdot,\omega)\right|^2\right)^{\frac{1}{2}} \\ &\quad + \left(\int_{Q_N} \left|\eta^{-2} R_\eta \left(p + \nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\right|^2\right)^{\frac{1}{2}}, \\ &\leq \|A_1\|_{(L^\infty(Q\times\Omega))^{d\times d}} \left\|\nabla v_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \\ &\quad + \left\|\eta^{-2} R_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}} \left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \\ &\quad + \left\|\eta^{-2} R_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}} |Q_N|^{\frac{1}{2}}. \end{aligned} \tag{7.41}$$

In view of (7.21) we have that, for any $\eta \leq 1$,

$$\left\|\eta^{-2} R_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}} \leq \left\|\eta^{-2} R_\eta\right\|_{(L^\infty(Q\times\Omega))^{d\times d}} \leq C_R.$$

We deduce from (7.41), (7.32) and (7.39) that

$$\begin{aligned} \left\|\nabla z_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} &\leq \frac{\|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}}}{\gamma} C_v |Q_N|^{\frac{1}{2}} + \frac{C_R}{\gamma}(C_w + 1)\,|Q_N|^{\frac{1}{2}} \\ &\leq C_z |Q_N|^{\frac{1}{2}}, \end{aligned} \tag{7.42}$$

with (see (7.32) and (7.39))

$$C_z = \frac{\|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}}}{\gamma}\left(\|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}} + C_R\right)\left(\frac{M}{\gamma^2}+\frac{1}{\gamma}\right) + \frac{C_R}{\gamma}\left(\frac{M}{\gamma}+1\right).$$

We observe that C_z is independent of N, h, η and ω. This concludes the proof of the first assertion of the proposition, namely the bound (7.30).

Step 4 : We now prove the second assertion of Proposition 7.2.1, namely the expansion of the approximated homogenized matrix. Using (7.15) and (7.18), we have for $1 \le i, j \le d$

$$\begin{aligned}\left[A_\eta^{\star,h,N}\right]_{ij}(\omega) &= \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_{\rm per}\left(\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)+e_j\right)\\ &+ \eta\frac{1}{|Q_N|}\int_{Q_N} e_i^T A_1(\cdot,\omega)\left(\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)+e_j\right)\\ &+ \frac{1}{|Q_N|}\int_{Q_N} e_i^T R_\eta(\cdot,\omega)\left(\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)+e_j\right).\end{aligned}$$

Using definitions (7.28), (7.29), (7.36) and (7.40), we write

$$\begin{aligned}&\eta^{-2}\left(\left[A_\eta^{\star,h,N}\right]_{ij}(\omega) - \left[A_{\rm per}^{\star,h}\right]_{ij} - \eta\left[A_1^{\star,h,N}\right]_{ij}(\omega)\right)\\ &= \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_{\rm per}\nabla z_{e_j}^{\eta,h,N}(\cdot,\omega)\\ &+\frac{1}{|Q_N|}\int_{Q_N} e_i^T A_1(\cdot,\omega)\nabla v_{e_j}^{\eta,h,N}(\cdot,\omega)\\ &+\frac{1}{|Q_N|}\int_{Q_N} e_i^T\eta^{-2}R_\eta(\cdot,\omega)\left(\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)+e_j\right).\end{aligned}$$

Applying Cauchy Schwarz inequality, and the bound (7.21), we obtain that, for any $|\eta| \le 1$,

$$\begin{aligned}&\eta^{-2}\left|\left[A_\eta^{\star,h,N}\right]_{ij}(\omega) - \left[A_{\rm per}^{\star,h}\right]_{ij} - \eta\left[A_1^{\star,h,N}\right]_{ij}(\omega)\right|\\ &\le \frac{1}{|Q_N|}\|A_{\rm per}\|_{(L^\infty(Q))^{d\times d}}|Q_N|^{\frac12}\left\|\nabla z_{e_j}^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\\ &+\frac{1}{|Q_N|}\|A_1\|_{(L^\infty(Q\times\Omega))^{d\times d}}|Q_N|^{\frac12}\left\|\nabla v_{e_j}^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\\ &+\frac{C_R}{|Q_N|}\left(|Q_N|^{\frac12}\left\|\nabla w_{e_j}^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}+|Q_N|\right).\end{aligned}$$

Thus using (7.32), (7.39) and (7.42), we deduce that

$$\begin{aligned}\eta^{-2}\left|\left[A_\eta^{\star,h,N}\right]_{ij}(\omega) - \left[A_{\rm per}^{\star,h}\right]_{ij} - \eta\left[A_1^{\star,h,N}\right]_{ij}(\omega)\right| &\le MC_z + \|A_1\|_{(L^\infty(\Omega\times Q))^{d\times d}}C_v\\ &\quad + C_R(1+C_w).\end{aligned}$$

Recall that C_w, C_v and C_z are all independent from η, N, h and ω. The estimate (7.31) is thus proved. This concludes the proof of Proposition 7.2.1. □

7.2.4 Convergence with respect to h and N

In Proposition 7.2.1, we have obtained bounds for quantities defined at the discrete level, namely after truncation and finite elements discretization. We now study the limit of (7.30) and (7.31) when $h \to 0$, and next $N \to +\infty$.

7.2.4.1 Convergence as $h \to 0$

First, let us define $w_p^{\eta,N}$, w_p^0 and $w_p^{1,N}$ solutions to the problems

$$\left\{ \begin{array}{l} \text{Find } w_p^{\eta,N}(\cdot,\omega) \in H^1_{\rm per}(Q_N) \text{ such that,} \\ \forall \varphi \in H^1_{\rm per}(Q_N), \quad \displaystyle\int_{Q_N} A_\eta(\cdot,\omega)\left(p + \nabla w_p^{\eta,N}(\cdot,\omega)\right)\cdot\nabla\varphi = 0 \text{ a. s.,} \end{array} \right. \tag{7.43}$$

$$\left\{ \begin{array}{l} \text{Find } w_p^{0} \in H^1_{\rm per}(Q) \text{ such that,} \\ \forall \varphi \in H^1_{\rm per}(Q), \quad \displaystyle\int_{Q} A_{\rm per}\left(p + \nabla w_p^{0}\right)\cdot\nabla\varphi = 0, \end{array} \right. \tag{7.44}$$

$$\left\{ \begin{array}{l} \text{Find } w_p^{1,N}(\cdot,\omega) \in H^1_{\rm per}(Q_N) \text{ such that,} \\ \forall \varphi \in H^1_{\rm per}(Q_N), \quad \displaystyle\int_{Q_N} A_{\rm per}\nabla w_p^{1,N}(\cdot,\omega)\cdot\nabla\varphi \\ \displaystyle + \int_{Q_N} A_1(\cdot,\omega)\left(p + \nabla w_p^{0}\right)\cdot\nabla\varphi = 0 \text{ a. s.,} \end{array} \right. \tag{7.45}$$

respectively, where $H^1_{\rm per}(Q)$ denotes the closure of $C^\infty_{\rm per}(Q)$, the space of infinitly derivable periodic functions, with respect to the H^1-norm. We also define the matrices $A_\eta^{\star,N}$, $A^\star_{\rm per}$ and $A_1^{\star,N}$ by

$$\forall 1 \le i,j \le d, \qquad \left[A_\eta^{\star,N}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_\eta(\cdot,\omega)\left(e_j + \nabla w_{e_j}^{\eta,N}(\cdot,\omega)\right), \tag{7.46}$$

$$\left[A^\star_{\rm per}\right]_{ij} = \int_Q e_i^T A_{\rm per}\left(e_j + \nabla w^0_{e_j}\right), \tag{7.47}$$

$$\begin{aligned} \left[A_1^{\star,N}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_{\rm per}\nabla w_{e_j}^{1,N}(\cdot,\omega) \\ + \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_1(\cdot,\omega)\left(\nabla w^0_{e_j} + e_j\right). \end{aligned} \tag{7.48}$$

We now prove the following result, which is a direct consequence of Proposition 7.2.1.

Corollary 7.2.1 *Suppose that A_η is a symmetric matrix that satisfies (7.2) and is stationary in the sense of* (7.4)*. Suppose, in addition, that it satisfies (7.18), (7.19), (7.20). We assume that (7.21) holds, namely the second order error is $O(\eta^2)$*

in $(L^\infty(Q\times\Omega))^{d\times d}$. *Then there exists a constant* C, *independent of* N, η *and* ω *such that, for* $|\eta|\leq 1$,

$$\eta^{-2}\left\|\nabla w_p^{\eta,N}(\cdot,\omega)-\nabla w_p^0-\eta\nabla w_p^{1,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\leq C\sqrt{|Q_N|}\quad \textit{almost surely,} \tag{7.49}$$

where $w_p^{\eta,N}$, w_p^0 *and* $w_p^{1,N}$ *are solutions to (7.43), (7.44) and (7.45), respectively, and such that*

$$\eta^{-2}\left|A_\eta^{\star,N}(\omega)-A_{\rm per}^\star-\eta A_1^{\star,N}(\omega)\right|\leq C\quad \textit{almost surely,} \tag{7.50}$$

where $A_\eta^{\star,N}$, $A_{\rm per}^\star$ *and* $A_1^{\star,N}$ *are defined by* (7.46), (7.47) *and* (7.48) *respectively.*

Proof: Using (7.30), remark that

$$\begin{aligned}
&\eta^{-2}\left\|\nabla w_p^{\eta,N}(\cdot,\omega)-\nabla w_p^0-\eta\nabla w_p^{1,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\\
&\leq \eta^{-2}\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)-\nabla w_p^{\eta,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}+\eta^{-2}\left\|\nabla w_p^{0,h}-\nabla w_p^0\right\|_{(L^2(Q_N))^d}\\
&\quad+\eta^{-1}\left\|\nabla w_p^{1,h,N}(\cdot,\omega)-\nabla w_p^{1,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}+C\sqrt{|Q_N|}
\end{aligned}\tag{7.51}$$

where C is independent of h, N, ω and η. Using standard properties of finite element approximations, we have that

- $\lim_{h\to 0}\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)-\nabla w_p^{\eta,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}=0$ almost surely,
- $\lim_{h\to 0}\left\|\nabla w_p^{1,h,N}(\cdot,\omega)-\nabla w_p^{1,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}=0$ almost surely,
- $\lim_{h\to 0}\left\|\nabla w_p^{0,h}-\nabla w_p^0\right\|_{(L^2(Q_N))^d}=0$.

Passing to the limit $h\to 0$ in (7.51), we obtain (7.49). The proof of (7.50) follows the same lines. □

Remark 7.2.1 *It is worth mentioning that Corollary 7.2.1 can also be proved directly, using essentially the same arguments as those of the proof of Theorem 7.2.1. Indeed, from* (7.43), *we deduce that* $\nabla w_p^{\eta,N}$ *is bounded in* $(L^2(Q_N))^d$. *Thus, it weakly converges to the gradient* ∇w_p^0 *of a function* $w_p^0\in H^1(Q_N)$ *that is independent from* N *(note that we invoke here weak convergence arguments, in constrast to the proof of Theorem 7.2.1, where we have used strong convergence arguments, which were easier to handle). Then, following the proof of Theorem 7.2.1, we introduce the function* $v_p^{\eta,N}=(w_p^{\eta,h,N}-w_p^{0,h})/\eta$, *the gradient of which is bounded in* $(L^2(Q_N))^d$ *(the inequalities of Theorem 7.2.1 still hold). Thus, there exists* $w_p^{1,N}\in H^1(Q_N)$ *such that* $\nabla v_p^{\eta,N}\longrightarrow\nabla w_p^{1,N}$ *in* $(L^2(Q_N))^d$. *The proof proceeds in the same manner.*

7.2.4.2 Convergence as $N\to\infty$

We now study the limit of (7.50) as $N\to+\infty$. From [23] we already know that

$$\lim_{N\to+\infty}A_\eta^{\star,N}(\omega)=A_\eta^\star\ \text{almost surely,} \tag{7.52}$$

where the exact homogenized matrix $A_\eta^\star$ is defined by (7.5). We now turn to $A_1^{\star,N}$ whose limit when $N\to+\infty$ is given by the following lemma.

Lemma 7.2.1 *Suppose that the matrices A_η, A_1 and $A_{\rm per}$ are symmetric, then the matrix $A_1^{\star,N}(\omega)$ defined by (7.48) satisfies*

$$\lim_{N\to+\infty} A_1^{\star,N}(\omega) = A_1^\star \text{ almost surely,} \tag{7.53}$$

where the deterministic matrix $A_1^\star$ is given by

$$\left[A_1^\star\right]_{ij}(\omega) = \int_Q \left(e_i + \nabla w_{e_i}^0\right)^T \mathbb{E}\left(A_1\right)\left(e_j + \nabla w_{e_j}^0\right),$$

where w_p^0 is the unique solution (up to an additive constant) to (7.44).

Proof: We observe that

$$\begin{aligned}\left[A_1^{\star,N}\right]_{ij}(\omega) &= \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_{\rm per}\nabla w_{e_j}^{1,N}(\cdot,\omega) + \frac{1}{|Q_N|}\int_{Q_N} e_i^T A_1(\cdot,\omega)\left(\nabla w_{e_j}^0 + e_j\right)\\ &= \frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^0\right)^T\left(A_{\rm per}\nabla w_{e_j}^{1,N}(\cdot,\omega) + A_1(\cdot,\omega)\left(\nabla w_{e_j}^0 + e_j\right)\right)\end{aligned} \tag{7.54}$$

because $w_{e_i}^0$ is Q_N-periodic and can therefore be chosen as a test function in (7.45). Since $A_{\rm per}$ is symmetric, we have, using (7.44), that

$$\int_{Q_N}\left(e_i+\nabla w_{e_i}^0\right)^T A_{\rm per}\nabla w_{e_j}^{1,N}(\cdot,\omega) = \int_{Q_N}\left(\nabla w_{e_j}^{1,N}(\cdot,\omega)\right)^T A_{\rm per}\left(e_i+\nabla w_{e_i}^0\right) = 0.$$

Thus

$$\left[A_1^{\star,N}\right]_{ij}(\omega) = \frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^0\right)^T A_1(\cdot,\omega)\left(\nabla w_{e_j}^0 + e_j\right),$$

and we conclude using the ergodic Theorem 7.1.1, as A_1 is stationary and ∇w_p^0 is periodic. □

Using (7.52) and (7.53), we can pass to the limit $N \to +\infty$ in (7.50) and we obtain that

$$\eta^{-2}\left|A_\eta^\star - A_{\rm per}^\star - \eta A_1^\star\right| \le C, \tag{7.55}$$

where C is independent of η and ω (note that $A_\eta^\star$, $A_{\rm per}^\star$ and $A_1^\star$ are all deterministic matrices). We thus recover the expansion of the exact *deterministic* homogenized matrix $A_\eta^\star$ as given in [P3].

7.3 Expansion of stochastic diffeomorphism (Model 2)

We now focus on Model 2. The goal of this section is the same as that of Section 2. We prove that the random second order error in the expansions of the corrector gradient $\nabla w_p^{\eta,h,N}$ and the homogenized matrix $A_\eta^{\star,h,N}$ is bounded independently of h, N and η in some appropriate space. The functional space in which the expansion of the original diffeomorphism Φ_η holds is simple and somehow corresponds to the one considered for Model 1 in Section 2.

7.3.1 Hypotheses

In this section, we consider the Model 2 mentioned above, where the random matrix A_η in (7.1) writes

$$A_\eta(x,\omega) = A_{\rm per}\left(\Phi_\eta^{-1}(x,\omega)\right), \tag{7.56}$$

where Φ_η is a *stochastic diffeomorphism*, that satisfies conditions (7.7), (7.8), (7.9) and (7.10). The periodic matrix $A_{\rm per}$ is supposed uniformly bounded and coercive :

$$\left\{\begin{array}{l} \exists \gamma > 0,\ \forall \xi \in \mathbb{R}^d,\ \xi^T A_{\rm per}(x)\xi \geq \gamma|\xi|^2 \text{ almost everywhere,} \\ \exists M > 0 \text{ such that } \|A_{\rm per}\|_{(L^\infty(Q))^{d\times d}} \leq M. \end{array}\right. \tag{7.57}$$

This model introduced by Blanc, Le Bris and Lions in [19] is not a particular case of the standard homogenization setting. Following [20] we now consider a weakly stochastic case, where the diffeomorphism Φ_η is close to the identity. More precisely, we suppose in the sequel that the following expansion holds in $\left(C^1\left(\mathbb{R}^d, L^\infty(\Omega)\right)\right)^d$:

$$\Phi_\eta(x,\omega) = x + \eta\Psi(x,\omega) + \Theta_\eta(x,\omega) \ \text{ with } \ \Theta_\eta = O(\eta^2), \tag{7.58}$$

where Ψ satisfies (7.9) and (7.10). This means that

$$\lim_{\eta\to 0} \|\Phi_\eta - {\rm Id}\|_{C^1\left(\mathbb{R}^d, L^\infty(\Omega)\right)^d} = 0, \tag{7.59}$$

$$\lim_{\eta\to 0} \left\|\eta^{-1}\left(\Phi_\eta - {\rm Id}\right) - \Psi\right\|_{C^1\left(\mathbb{R}^d, L^\infty(\Omega)\right)^d} = 0, \tag{7.60}$$

where Id denotes the identity mapping. In addition, $\Theta_\eta = O(\eta^2)$ means that there exists a *deterministic* constant C_Θ independent of η such that, when $|\eta| \leq 1$,

$$\left\|\eta^{-2}\Theta_\eta\right\|_{C^1\left(\mathbb{R}^d, L^\infty(\Omega)\right)^d} \leq C_\Theta. \tag{7.61}$$

In [20], it has been shown that, under these hypotheses, the exact homogenized matrix (7.11) possesses an expansion $A_\eta^\star = A_{\rm per}^\star + \eta A_1^\star + O(\eta^2)$, where $A_{\rm per}^\star$ and $A_1^\star$ are deterministic matrices that only involve solutions to deterministic problems posed on a bounded domain, in contrast to $A_\eta^\star$ (see (7.12) and (7.11)). Here, as above, we will prove that the *approximate* homogenized matrix, obtained after truncation and discretization, also possesses an expansion in powers of η and that the error at second order is bounded independently of the parameters of the discretization procedure.

7.3.2 Formal expansion

As in Section 7.2, we first present formally our main result. We assume the following formal expansion on the approximated corrector, solution to (7.16) :

$$w_p^{\eta,h,N} = w_p^{0,h,N} + \eta w_p^{1,h,N} + r_p^{\eta,h,N}, \tag{7.62}$$

where $\nabla r_p^{\eta,h,N} = O(\eta^2)$ in some appropriate space. Inserting this expansion in (7.16), we obtain that the function $w_p^{0,h,N}$ is in fact equal to $w_p^{0,h}$ solution to (7.25). The function $w_p^{1,h,N}(\cdot,\omega) \in V_{\rm per}^h(Q_N)$ is the unique function such that, for any $\varphi_h \in V_{\rm per}^h(Q_N)$, we have

$$\int_{Q_N} A_{\rm per}\nabla w_p^{1,h,N}(\cdot,\omega)\cdot\nabla\varphi_h + \int_{Q_N} A_{\rm per} + \nabla\Psi(\cdot,\omega)\nabla w_p^{0,h,N}(\cdot,\omega)\cdot\nabla\varphi_h$$
$$\int_{Q_N} \left(\mathrm{div}\Psi(\cdot,\omega)\mathbb{I}_d - \nabla\Psi(\cdot,\omega)^T\right) A_{\rm per}\left(p + \nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h = 0 \quad \text{a. s.} \tag{7.63}$$

In addition, substituting Φ_η and $w_p^{\eta,h,N}$ by their expansions into (7.17), we formally obtain that

$$A_\eta^{\star,h,N}(\omega) = A_{\rm per}^{\star,h} + \eta A_1^{\star,h,N}(\omega) + O(\eta^2), \tag{7.64}$$

where the term of order zero is defined by (7.25) and (7.28), as in the first model. The term of order one is defined by

$$\begin{aligned}\left[A_1^{\star,h,N}\right]_{ij}(\omega) &= -\left[A_{\rm per}^{\star,h}\right]_{ij}\frac{1}{|Q_N|}\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega)\\ &+ \frac{1}{|Q_N|}\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega)\left(e_i+\nabla w_{e_i}^{0,h}\right)^T A_{\rm per}e_j\\ &+ \frac{1}{|Q_N|}\int_{Q_N}\left(\nabla w_{e_i}^{1,h,N}(\cdot,\omega) - \nabla\Psi(\cdot,\omega)\nabla w_{e_i}^{0,h}\right)^T A_{\rm per}e_j. \end{aligned}\tag{7.65}$$

As in Section 7.2, we now make precise and rigorously justify the expansions (7.62) and (7.64).

7.3.3 Main result

The main result of this section is the following proposition, which is analogous to Proposition 7.2.1.

Proposition 7.3.1 *Suppose that Φ_η satisfies (7.7), (7.8), (7.9) and (7.10), and that it is a perturbation of the identity in the sense of (7.58), (7.59), (7.60) and (7.61). We suppose in addition that the symmetric periodic matrix $A_{\rm per}$ satisfies* (7.57)*. Then there exists a constant C independent of η, ω, N and h, such that, for sufficiently small values of η,*

$$\eta^{-2}\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega) - \nabla w_p^{0,h} - \eta\nabla w_p^{1,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \le C\sqrt{|Q_N|} \quad \textit{almost surely} \tag{7.66}$$

where $w_p^{\eta,h,N}$, $w_p^{0,h}$ and $w_p^{1,h,N}$ are solutions to (7.16), (7.25) and (7.63), respectively, and such that

$$\eta^{-2}\left|A_\eta^{\star,h,N}(\omega) - A_{\rm per}^{\star,h} - \eta A_1^{\star,h,N}(\omega)\right| \le C \quad \textit{almost surely,} \tag{7.67}$$

where $A_\eta^{\star,h,N}$, $A_{\rm per}^{\star,h}$ and $A_1^{\star,h,N}$ are given by (7.17), (7.28) *and* (7.65) *respectively.*

The proof, which is detailed below, follows the same lines as that of Proposition 7.2.1. It makes use of the following lemma.

Lemma 7.3.1 *Suppose that Φ_η satisfies (7.58), (7.59), (7.60) and (7.61). Then, there exist*

$$\Gamma_\eta \in \left(C^0\left(\mathbb{R}^d, L^\infty(\Omega)\right)\right)^{d\times d} \text{ and } \sigma_\eta \in C^0\left(\mathbb{R}^d, L^\infty(\Omega)\right),$$

such that

$$\left(\nabla\Phi_\eta\right)^{-1}(\cdot,\omega) = \mathbb{I}_d - \eta\nabla\Psi(\cdot,\omega) + \Gamma_\eta(\cdot,\omega) \text{ almost surely,} \tag{7.68}$$

$$\det(\nabla\Phi_\eta)(\cdot,\omega) = 1 + \eta\mathrm{div}\Psi(\cdot,\omega) + \sigma_\eta(\cdot,\omega) \text{ almost surely,} \tag{7.69}$$

with, when η is sufficiently small,

$$\|\Gamma_\eta\|_{C^0\left(\mathbb{R}^d, L^\infty(\Omega)\right)^{d\times d}} \leq C_\Gamma\eta^2 \quad \text{and} \quad \|\sigma_\eta\|_{C^0\left(\mathbb{R}^d, L^\infty(\Omega)\right)} \leq C_\sigma\eta^2, \tag{7.70}$$

for two deterministic constants C_Γ and C_σ independent of η. In addition, there exists η_0 such that, for all $|\eta| \leq \eta_0$, we have

$$\forall\xi\in\mathbb{R}^d,\ \forall x\in\mathbb{R}^d,\quad \frac{1}{2}|\xi|^2 \leq \xi^T\left(\nabla\Phi_\eta(x,\omega)\right)^{-T}\cdot\left(\nabla\Phi_\eta(x,\omega)\right)^{-1}\xi \leq \frac{3}{2}|\xi|^2 \quad \text{a. s. .} \tag{7.71}$$

Proof of Lemma 7.3.1. The proofs of (7.68), (7.69) and (7.70) are straightforward. We now prove (7.71). First, let us denote by Λ the functional that associates to any symmetric matrix $M \in \mathbb{S}_d(\mathbb{R})$ its ordered eigenvalues $\Lambda(M) = (\lambda_1, \cdots, \lambda_d)$. This application is continuous on $\mathbb{S}_d(\mathbb{R})$. As a consequence, there exists a value δ_0 such that

$$\forall\delta\leq\delta_0,\ |M-\mathbb{I}_d| \leq \delta \Rightarrow |\Lambda(M) - (1,\cdots,1)| \leq \frac{1}{2}. \tag{7.72}$$

We infer from (7.68) that the following expansion holds in $\left(C^0\left(\mathbb{R}^d, L^\infty(\Omega)\right)\right)^{d\times d}$:

$$\left(\nabla\Phi_\eta\right)^{-T}\left(\nabla\Phi_\eta\right)^{-1} = \mathbb{I}_d - \eta\left(\left(\nabla\Psi\right)^T + \nabla\Psi\right) + O(\eta^2).$$

As a consequence, there exists a *deterministic* value η_0 such that

$$\forall\eta\leq\eta_0,\ \forall x\in\mathbb{R}^d,\quad \left|\left(\nabla\Phi_\eta(x,\omega)\right)^{-T}\left(\nabla\Phi_\eta(x,\omega)\right)^{-1} - \mathbb{I}_d\right| \leq \delta_0, \quad \text{a. s..} \tag{7.73}$$

Collecting (7.72) and (7.73), we deduce that

$$\forall|\eta|\leq\eta_0,\ \forall x\in\mathbb{R}^d,\ \left|\Lambda\left(\left(\nabla\Phi_\eta(x,\omega)\right)^{-T}\left(\nabla\Phi_\eta(x,\omega)\right)^{-1}\right) - (1,\cdots,1)\right| \leq \frac{1}{2},\ \text{ a. s. .}$$

Hence all the eigenvalues of $\left(\nabla\Phi_\eta\right)^{-T}\left(\nabla\Phi_\eta\right)^{-1}$ belong to the interval $[1/2, 3/2]$. It implies (7.71).

Proof of Proposition 7.3.1. As announced above, the structure of the proof is the same as that of Proposition 2.1. First, we rigorously justify (7.62). Doing so, we obtain bounds on $\nabla w_p^{\eta,h,N}$, and on the error terms $\nabla v_p^{\eta,h,N}$ and $\nabla z_p^{\eta,h,N}$ that are defined below (Steps 1, 2 and 3 respectively). These bounds are independent of h, N, η and ω. We use them at Step 4 to control the second order error of the expansion of the approximated homogenized matrix $A_\eta^{\star,h,N}$.

Step 1 : Our goal is first to prove that $\nabla w_p^{\eta,h,N}(\cdot,\omega)$ is bounded in $(L^2(Q_N))^d$, almost surely. We next show that $w_p^{0,h,N}$, the limit of $w_p^{\eta,h,N}$, is in fact the unique solution to (7.25). Choosing $\varphi_h = w_p^{\eta,h,N}(\cdot,\omega)$ as test function in (7.16) and using (7.57), we obtain that

$$\gamma \left\| \sqrt{\det(\nabla\Phi_\eta)(\cdot,\omega)}\, (\nabla\Phi_\eta(\cdot,\omega))^{-1}\, \nabla w_p^{\eta,h,N}(\cdot,\omega) \right\|_{(L^2(Q_N))^d} \leq \left(\int_{Q_N} \det(\nabla\Phi_\eta)(\cdot,\omega) |A_{\rm per} p|^2 \right)^{\frac{1}{2}}.$$

Using Lemma 7.3.1 and (7.8) we obtain

$$\gamma \left\| \sqrt{\det(\nabla\Phi_\eta)(\cdot,\omega)}\, (\nabla\Phi_\eta(\cdot,\omega))^{-1}\, \nabla w_p^{\eta,h,N}(\cdot,\omega) \right\|_{(L^2(Q_N))^d} \geq \gamma \sqrt{\frac{\nu}{2}} \left\| \nabla w_p^{\eta,h,N}(\cdot,\omega) \right\|_{(L^2(Q_N))^d}. \tag{7.74}$$

From (7.9), we know that $|\det(\nabla\Phi_\eta)(\cdot,\omega)| \leq d!(M')^d$. We thus have (recalling that $|p| = 1$)

$$\left(\int_{Q_N} \det(\nabla\Phi_\eta)(\cdot,\omega) |A_{\rm per} p|^2 \right)^{\frac{1}{2}} \leq \left(d!(M')^d \right)^{\frac{1}{2}} M |Q_N|^{\frac{1}{2}}. \tag{7.75}$$

Collecting (7.74) and (7.75), we obtain that

$$\left\| \nabla w_p^{\eta,h,N}(\cdot,\omega) \right\|_{(L^2(Q_N))^d} \leq C_w |Q_N|^{\frac{1}{2}}, \tag{7.76}$$

where

$$C_w = \sqrt{\frac{2}{\nu}} \frac{\left(d!(M')^d \right)^{\frac{1}{2}} M}{\gamma}$$

is independent of η, h, N and ω. Hence, $\nabla w_p^{\eta,h,N}$ is bounded in $L^2(Q_N)^d$, independently of η. As in the proof of Proposition 7.2.1 we deduce that there exists a function $w_p^{0,h,N}(\cdot,\omega) \in V_h^{\rm per}(Q_N)$ such that

$$\nabla w_p^{\eta,h,N}(\cdot,\omega) \underset{\eta\to 0}{\longrightarrow} \nabla w_p^{0,h,N} \text{ in } \left(L^2(Q_N) \right)^d \text{ almost surely.} \tag{7.77}$$

We now insert in equation (7.16) the expansions (7.68) and (7.69). For any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$\begin{aligned} 0 &= \int_{Q_N} \det(\nabla\Phi_\eta(\cdot,\omega))(\nabla\Phi_\eta(\cdot,\omega))^{-T} A_{\rm per} \left(p + (\nabla\Phi_\eta(\cdot,\omega))^{-1} \nabla w_p^{\eta,h,N}(\cdot,\omega) \right) \cdot \nabla\varphi_h \\ &= E_0^\eta(\omega) + \eta E_1^\eta(\omega) + \eta^2 E_2^\eta(\omega) + E_3^\eta(\omega) + F_\eta(\omega), \end{aligned} \tag{7.78}$$

with

$$
\begin{aligned}
E_0^\eta(\omega) &= \int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\
E_1^\eta(\omega) &= \int_{Q_N} {\rm div}\Psi(\cdot,\omega) A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\
&- \int_{Q_N} \nabla\Psi(\cdot,\omega)^T A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\
&- \int_{Q_N} A_{\rm per}\nabla\Psi(\cdot,\omega)\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h, \qquad (7.79) \\
E_2^\eta(\omega) &= \int_{Q_N} \nabla\Psi(\cdot,\omega)^T A_{\rm per}\nabla\Psi(\cdot,\omega)\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h \\
&- \int_{Q_N} {\rm div}\Psi(\cdot,\omega)\nabla\Psi(\cdot,\omega)^T A_{\rm per}\left(\nabla w_p^{\eta,h,N}(\cdot,\omega)+p\right)\cdot\nabla\varphi_h \\
&- \int_{Q_N} {\rm div}\Psi(\cdot,\omega) A_{\rm per}\nabla\Psi(\cdot,\omega)\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h, \qquad (7.80) \\
E_3^\eta(\omega) &= \int_{Q_N} \sigma_\eta(\cdot,\omega) A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\
&+ \int_{Q_N} \Gamma_\eta(\cdot,\omega)^T A_{\rm per}\left(\nabla w_p^{\eta,h,N}(\cdot,\omega)+p\right)\cdot\nabla\varphi_h \\
&+ \int_{Q_N} A_{\rm per}\Gamma_\eta(\cdot,\omega)\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h, \qquad (7.81)
\end{aligned}
$$

and $F_\eta(\omega)$ is the remainder term. Formally, E_0^η is of order η^0, ηE_1^η is of order η, $\eta^2 E_2^\eta$ is of order η^2 (because it includes products of two quantities of the order η), E_3^η is of order η^2 as well (it includes products of a quantity of order η^2 with a quantity of order η^0) and F_η is of order η^3. We now prove rigorously boundedness and convergence results for each of the terms : $E_0^\eta, E_1^\eta, E_2^\eta, E_3^\eta$ and F_η. These estimates are very useful not only at Step 1 but also at Steps 2 and 3.

Convergences of E_0^η, E_1^η and E_2^η, as $\eta\to 0$. We deduce from (7.77) that

$$
E_0^\eta(\omega) = \int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \underset{\eta\to 0}{\longrightarrow} \int_{Q_N} A_{\rm per}\left(p+\nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \qquad (7.82)
$$

almost surely. The same argument gives

$$\begin{aligned} E_1^\eta(\omega) \underset{\eta\to 0}{\longrightarrow} & \int_{Q_N} \mathrm{div}\Psi(\cdot,\omega) A_{\mathrm{per}}\left(p+\nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\ & - \int_{Q_N} \nabla\Psi(\cdot,\omega)^T A_{\mathrm{per}}\left(p+\nabla w_p^{0,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h \\ & - \int_{Q_N} A_{\mathrm{per}}\nabla\Psi(\cdot,\omega)\nabla w_p^{0,h,N}(\cdot,\omega)\cdot\nabla\varphi_h, \quad \text{a. s.} \end{aligned} \tag{7.83}$$

$$\begin{aligned} E_2^\eta(\omega) \underset{\eta\to 0}{\longrightarrow} & \int_{Q_N} \nabla\Psi(\cdot,\omega)^T A_{\mathrm{per}}\nabla\Psi(\cdot,\omega)\nabla w_p^{0,h,N}(\cdot,\omega)\cdot\nabla\varphi_h \\ & - \int_{Q_N} \mathrm{div}\Psi(\cdot,\omega)\nabla\Psi(\cdot,\omega)^T A_{\mathrm{per}}\left(\nabla w_p^{0,h,N}(\cdot,\omega)+p\right)\cdot\nabla\varphi_h \\ & - \int_{Q_N} \mathrm{div}\Psi(\cdot,\omega) A_{\mathrm{per}}\nabla\Psi(\cdot,\omega)\nabla w_p^{0,h,N}(\cdot,\omega)\cdot\nabla\varphi_h, \quad \text{a. s.}. \end{aligned} \tag{7.84}$$

Bounds on E_1^η, E_2^η, E_3^η and F_η for sufficiently small values of η. First we remark that

$$\begin{aligned} |E_1^\eta(\omega)| \leq & \|A_{\mathrm{per}}\|_{(L^\infty(Q_N))^{d\times d}} \|\nabla\varphi_h\|_{(L^2(Q_N))^d} \\ & \times \left(|Q_N|^{\frac12} + \left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\right) \\ & \times \left(\|\mathrm{div}\Psi\|_{C^0(\mathbb{R}^d,L^\infty(\Omega))} + \|\nabla\Psi\|_{\left(C^0(\mathbb{R}^d,L^\infty(\Omega))\right)^{d\times d}}\right) \\ & + \|A_{\mathrm{per}}\|_{(L^\infty(Q_N))^{d\times d}} \|\nabla\Psi\|_{\left(C^0(\mathbb{R}^d,L^\infty(\Omega))\right)^{d\times d}} \\ & \times \|\nabla\varphi_h\|_{(L^2(Q_N))^d} \left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}, \\ |E_2^\eta(\omega)| \leq & \|A_{\mathrm{per}}\|_{(L^\infty(Q_N))^{d\times d}} \|\nabla\varphi_h\|_{(L^2(Q_N))^d} \\ & \times \left(|Q_N|^{\frac12} + \left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\right) \\ & \times \|\mathrm{div}\Psi\|_{C^0(\mathbb{R}^d,L^\infty(\Omega))} \|\nabla\Psi\|_{C^0(\mathbb{R}^d,L^\infty(\Omega))} \\ & + \|A_{\mathrm{per}}\|_{(L^\infty(Q_N))^{d\times d}} \|\nabla\varphi_h\|_{(L^2(Q_N))^d} \\ & \times \left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \\ & \times \left(\|\nabla\Psi\|^2_{C^0(\mathbb{R}^d,L^\infty(\Omega))} + \|\nabla\Psi\|_{C^0(\mathbb{R}^d,L^\infty(\Omega))} \|\mathrm{div}\Psi\|_{C^0(\mathbb{R}^d,L^\infty(\Omega))}\right), \end{aligned}$$

almost surely. Using (7.76), we deduce that there exist two constants C_{E_1} and C_{E_2} independent of η, h, N and ω such that,

$$|E_1^\eta(\omega)| \leq C_{E_1} \|\nabla\varphi_h\|_{(L^2(Q_N))^d} |Q_N|^{\frac12} \quad \text{and} \quad |E_2^\eta(\omega)| \leq C_{E_2} \|\nabla\varphi_h\|_{(L^2(Q_N))^d} |Q_N|^{\frac12}, \tag{7.85}$$

almost surely. Second, using (7.70), we observe that each term of $E_3^\eta(\omega)$ is bounded. As an example, the first term of $E_3^\eta(\omega)$ satisfies

$$\begin{aligned}&\left|\int_{Q_N} \sigma_\eta(\cdot,\omega) A_{\rm per}\left(p+\nabla w_p^{\eta,h,N}(\cdot,\omega)\right)\cdot\nabla\varphi_h\right| \\ &\leq C_\sigma \eta^2 \left\|\nabla\varphi_h\right\|_{(L^2(Q_N))^d} \left\|A_{\rm per}\right\|_{(L^\infty(Q_N))^{d\times d}} \left(\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} + |Q_N|^{\frac{1}{2}}\right).\end{aligned}$$

The other terms of $E_3^\eta(\omega)$ can be bounded likewise. Using (7.76), we obtain that there exists a constant C_{E_3} independent of η, h, N and ω such that, for sufficiently small values of η, we have

$$|E_3^\eta(\omega)| \leq C_{E_3} \left\|\nabla\varphi_h\right\|_{(L^2(Q_N))^d} |Q_N|^{\frac{1}{2}} \eta^2 \quad \text{almost surely.} \tag{7.86}$$

We conclude that

$$E_3^\eta(\omega) \underset{\eta\to 0}{\longrightarrow} 0 \quad \text{almost surely.} \tag{7.87}$$

Bounding from above $F_\eta(\omega)$ is technically more tedious, since it contains many terms. Nevertheless, each of them can be written under one of these two forms

$$\int_{Q_N} T_\eta(\cdot,\omega)\nabla w_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h \quad \text{or} \quad \int_{Q_N} T_\eta(\cdot,\omega)\left(\nabla w_p^{\eta,h,N}(\cdot,\omega)+p\right)\cdot\nabla\varphi_h,$$

where $T_\eta(\cdot,\omega) \to 0$ strongly in $(L^\infty(Q_N))^{d\times d}$ almost surely. As an exemple, let us consider a term corresponding to the choice $T_\eta(\cdot,\omega) = \eta {\rm div}\Psi(\cdot,\omega) A_{\rm per}\Gamma_\eta(\cdot,\omega)$. Using (7.70) and (7.76), we have

$$\begin{aligned}&\left|\int_{Q_N} \eta {\rm div}\Psi(\cdot,\omega) A_{\rm per}\Gamma_\eta(\cdot,\omega)\nabla w_p^{\eta,h,N}\cdot\nabla\varphi_h\right| \\ &\leq \eta \left\|{\rm div}\Psi(\cdot,\omega) A_{\rm per}\Gamma_\eta(\cdot,\omega)\right\|_{(L^\infty(Q_N))^{d\times d}} C_w |Q_N|^{\frac{1}{2}} \left\|\nabla\varphi_h\right\|_{(L^2(Q_N))^d} \\ &\leq \left\|A_{\rm per}\right\|_{(L^\infty(Q_N))^{d\times d}} \left\|{\rm div}\Psi(\cdot,\omega)\right\|_{L^\infty(Q_N)} C_\Gamma \eta^3 C_w |Q_N|^{\frac{1}{2}} \left\|\nabla\varphi_h\right\|_{(L^2(Q_N))^d}.\end{aligned}$$

This term tends to 0 as $\eta \to 0$. Carrying over the same analysis for each term of $F_\eta(\omega)$, we conclude that there exists a *deterministic* constant C_F independent of h, N and η such that, for sufficiently small values of η, we have

$$|F_\eta(\omega)| \leq C_F \eta^3 \left\|\nabla\varphi_h\right\|_{(L^2(Q_N))^d} |Q_N|^{\frac{1}{2}} \quad \text{almost surely.} \tag{7.88}$$

Of course, this implies

$$F_\eta(\omega) \underset{\eta\to 0}{\longrightarrow} 0 \quad \text{almost surely.} \tag{7.89}$$

Zero order equation. We now return to equation (7.78) and pass to the limit $\eta \to 0$. Using (7.85), it is obvious that

$$\eta E_1^\eta(\omega) \underset{\eta\to 0}{\longrightarrow} 0, \qquad \eta^2 E_2^\eta(\omega) \underset{\eta\to 0}{\longrightarrow} 0 \ \text{ almost surely.} \tag{7.90}$$

Collecting (7.82), (7.87), (7.89) and (7.90), we deduce that $w_p^{0,h,N}(\cdot,\omega)$ is a solution to (7.35). We have proved (see the first step of the proof of Proposition 7.2.1) that (7.35) has a unique solution which is independent of N and is equal to $w_p^{0,h}$ the unique solution $w_p^{0,h}$ to (7.25).

Step 2 : We introduce the function

$$v_p^{\eta,h,N}(\cdot,\omega) = \frac{w_p^{\eta,h,N}(\cdot,\omega) - w_p^{0,h}}{\eta}.$$

Our aim is to prove that $\nabla v_p^{\eta,h,N}$ is bounded in $(L^2(Q_N))^d$, almost surely. We next show that $w_p^{1,h,N}$, the limit of $v_p^{\eta,h,N}$, is in fact the unique solution to (7.63). Substracting equation (7.35) from (7.78), we obtain that $v_p^{\eta,h,N} \in V_h^{\rm per}(Q_N)$ is such that, for any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$0 \ = \ \int_{Q_N} A_{\rm per}\nabla v_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h + E_1^\eta(\omega) + \eta E_2^\eta(\omega) + \frac{E_3^\eta(\omega)}{\eta} + \frac{F_\eta(\omega)}{\eta}. \tag{7.91}$$

We are now going to use the bounds on E_1^η, E_2^η, E_3^η, F^η obtained below. Indeed, choosing $\varphi_h = v_p^{\eta,h,N}(\cdot,\omega)$ in (7.91), (7.85), (7.86), and (7.88), we obtain that, for sufficiently small values of η,

$$\gamma \left\|\nabla v_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \le \left(C_{E_1} + \eta C_{E_2} + \eta C_{E_3} + \eta^2 C_F\right)|Q_N|^{\frac{1}{2}},$$

where C_{E_1}, C_{E_2}, C_{E_3} and C_F are all independent of η, h, N and ω. As a consequence, $v_p^{\eta,h,N}(\cdot,\omega)$ is bounded in $(L^2(Q_N))^d$ almost surely. Thus, there exists a deterministic constant C_v independent of h, N and η, such that, for sufficiently small values of η,

$$\left\|\nabla v_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \ \le \ C_v|Q_N|^{\frac{1}{2}} \ \text{ almost surely.} \tag{7.92}$$

Thus, there exists a function $w_p^{1,h,N}(\cdot,\omega) \in V_{\rm per}^h(Q_N)$ such that, up to extracting a subsequence,

$$\nabla v_p^{\eta,h,N}(\cdot,\omega) \underset{\eta\to 0}{\longrightarrow} \nabla w_p^{1,h,N}(\cdot,\omega) \text{ in } \left(L^2(Q_N)\right)^d \text{ almost surely.} \tag{7.93}$$

Using (7.83), (7.85), (7.86), (7.88) and (7.93), we can pass to the limit $\eta \to 0$ in (7.91). We conclude that $w_p^{1,h,N}$ is the unique function such that, for any $\varphi_h \in$

$V_h^{\rm per}(Q_N)$

$$\int_{Q_N} A_{\rm per}\nabla w_p^{1,h,N}(\cdot,\omega)\cdot\nabla\varphi_h - \int_{Q_N} A_{\rm per}\nabla\Psi(\cdot,\omega)\nabla w_p^{0,h}\cdot\nabla\varphi_h$$
$$+ \int_{Q_N}\left(\mathrm{div}\Psi(\cdot,\omega)\mathbb{I}_d - \nabla\Psi(\cdot,\omega)^T\right) A_{\rm per}\left(p+\nabla w_p^{0,h}\right)\cdot\nabla\varphi_h = 0.$$

We recover the problem (7.63), the solution of which is unique.

Step 3 : We define

$$z_p^{\eta,h,N}(\cdot,\omega) = \frac{w_p^{\eta,h,N}(\cdot,\omega) - w_p^{0,h} - \eta w_p^{1,h,N}(\cdot,\omega)}{\eta^2}.$$

Using (7.35), (7.63), (7.78) and (7.93) we obtain that $z_p^{\eta,h,N}$ is such that, for any $\varphi_h \in V_h^{\rm per}(Q_N)$,

$$0 = \int_{Q_N} A_{\rm per}\nabla z_p^{\eta,h,N}(\cdot,\omega)\cdot\nabla\varphi_h + E_2^\eta(\omega) + \frac{E_3^\eta(\omega)}{\eta^2} + \frac{F_\eta(\omega)}{\eta^2}. \qquad (7.94)$$

Choosing $\varphi_h = z_p^{\eta,h,N}(\cdot,\omega)$ in (7.94), and using (7.85), (7.86), and (7.88), we obtain that, for sufficiently small values of η,

$$\gamma\left\|\nabla z_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \le \left(C_{E_2} + C_{E_3} + \eta C_F\right)|Q_N|^{\frac{1}{2}}.$$

We thus have

$$\left\|\nabla z_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d} \le C_z|Q_N|^{\frac{1}{2}}, \qquad (7.95)$$

where C_z is independent of h, N, η, and ω. Thus, inequality (7.66) is proved.

Step 4 : Our last step aims at bounding the second order error of the expansion of $A_\eta^{\star,h,N}$. We insert the expansions (7.68) and (7.69) into (7.17). We obtain that

$$\left[A_\eta^{\star,h,N}\right]_{ij}(\omega) =$$
$$\frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j$$
$$+\eta\left(\begin{array}{rl} & -\left(\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\left(e_i+\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j\right)\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega) \\ + & \dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega)\left(e_i+\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j \\ - & \dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\left(\nabla\Psi(\cdot,\omega)\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j \end{array}\right)$$
$$+\delta_\eta^{h,N}(\omega),$$

where $\delta_\eta^{h,N}(\omega)$ is a sum of terms of the forms

$$s_\eta(\omega)\frac{1}{|Q_N|}\int_{Q_N} T_\eta(\cdot,\omega)\left(\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}\cdot e_j$$
$$\text{and}\quad s_\eta(\omega)\frac{1}{|Q_N|}\int_{Q_N} T_\eta(\cdot,\omega)\left(\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)+e_i\right)^T A_{\rm per}\cdot e_j,$$

with $s_\eta(\omega)\in\mathbb{R}$ and $T_\eta\in C^0\left(\mathbb{R}^d,L^\infty(\Omega)\right)^{d\times d}$. For each of these (s_η,T_η), there exists a constant $C_{s,T}$ independent of h, N, η, and ω such that, for sufficiently small values of η,

$$\|s_\eta\|_{L^\infty(\Omega)}\,\|T_\eta\|_{(C^0(Q_N,L^\infty(\Omega)))^{d\times d}}\le C_{s,T}\eta^2.$$

Therefore, for each term of $\delta_\eta^{h,N}(\omega)$ we can write

$$\begin{aligned}
&\left|s_\eta(\omega)\frac{1}{|Q_N|}\int_{Q_N} T_\eta(\cdot,\omega)\left(\nabla w_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}\cdot e_j\right|\\
&\le \frac{1}{|Q_N|}\,\|s_\eta\|_{L^\infty(\Omega)}\,\|T_\eta\|_{\left(C^0\left(\mathbb{R}^d,L^\infty(Q_N)\right)\right)^{d\times d}}\\
&\quad\times\left\|\nabla w_p^{\eta,h,N}(\cdot,\omega)\right\|_{(L^2(Q_N))^d}\|A_{\rm per}\|_{(L^\infty(Q_N))^{d\times d}}\,|Q_N|^{\frac12}\\
&\le C_{s,T}C_w\,\|A_{\rm per}\|_{(L^\infty(Q_N))^{d\times d}}\,\eta^2.
\end{aligned}$$

As a consequence, there exists a constant C_δ independent of h, N, η and ω, such that, for sufficiently small values of η,

$$|\delta_\eta^{h,N}(\omega)|\le C_\delta\eta^2\ \text{ almost surely.}\tag{7.96}$$

Next, we compute

$$\begin{aligned}
&\eta^{-2}\left(A_\eta^{\star,h,N}(\omega)-A_{\rm per}^{\star,h}-\eta A_1^{\star,h,N}(\omega)\right)=\\
&\quad+\frac{1}{|Q_N|}\int_{Q_N}\left(e_i+\nabla z_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j\\
&\quad+\left(\begin{array}{rl}
- & \left(\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\left(\nabla v_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j\right)\dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega)\\
+ & \dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\mathrm{div}\Psi(\cdot,\omega)\left(\nabla v_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j\\
- & \dfrac{1}{|Q_N|}\displaystyle\int_{Q_N}\left(\nabla\Psi(\cdot,\omega)\nabla v_{e_i}^{\eta,h,N}(\cdot,\omega)\right)^T A_{\rm per}e_j
\end{array}\right)\\
&\quad+\frac{\delta_\eta^{h,N}(\omega)}{\eta^2}.
\end{aligned}$$

Using bounds (7.92), (7.95) and (7.96), we obtain (7.67). This concludes the proof.

Remark 7.3.1 *As in Section 7.2, we can pass to the limit* $h \to 0$ *in the bounds of Proposition 3.1 and prove that* (7.67) *extends to the case when only truncation is taken into account. Letting* N *go to infinity is more difficult. Indeed, we need to show that* $\lim_{N\to\infty}\left(\lim_{h\to 0} A_\eta^{\star,h,N}(\omega)\right) = A_\eta^\star$, *almost surely. Adapting the arguments of [23] to Model 2 is not straightforward. Indeed, the proof in [23] is based on a change of variable* $\widetilde{w}_p^N(x,\omega) = \frac{1}{N} w_p^N(Nx,\omega)$ *that allows to invoke classical properties of G-convergence of the operator* $-\mathrm{div}\left(A(Nx,\omega)\nabla\cdot\right)$. *For Model 2, due to the diffeomorphism, the situation is different. In [19], a G-convergence result have been shown for the operator* $T_N = -\mathrm{div}\left(A_{\mathrm{per}}\left(\Phi_\eta^{-1}(x,\omega)\right)\nabla\cdot\right)$. *If we rescale the corrector, solution to* (7.16), *in the same spirit as for Model 1, then we are left with showing a G-convergence result for an operator that is* not T_N.

AcknowledgmentsThe author thanks Frédéric Legoll for helpful discussions. This work is partially suppported by ONR under contract Grant 00014-09-1-0470.

Bibliographie

[1] G. Allaire, Homogenization and two-scale convergence, *SIAM J. Math. Anal.* 23 (6), pp 1482-1518, 1992.

[2] G. Allaire, Shape optimization by the homogenization method, **Applied Mathematical Sciences**, 146, Springer, 2002.

[3] G. Allaire and R. Brizzi, A multiscale finite element method for numerical homogenization, *Multiscale Modeling and Simulation*, 4(3) :790-812, 2005.

[4] G. Allaire and S. Gutierrez, Optimal design in small amplitude homogenization, *Math. Model. Numer. Anal. (M2AN)*, 41(3) :543-574, 2007.

[5] A. Anantharaman, Thèse de l' Université Paris Est, Ecole des Ponts, 2010. `http ://hal.archives-ouvertes.fr/tel-00558618`

[6] A. Anantharaman, X. Blanc, and F. Legoll, Asymptotic behaviour of Green functions of divergence form operators with periodic coefficients, in preparation.

[7] A. Anantharaman and C. Le Bris, Homogenization of a weakly randomly perturbed periodic material, *C. R. Acad. Sci. Sér. I*, vol. 348 (2010) : 529-534.

[8] A. Anantharaman and C. Le Bris, A numerical approach related to defect-type theories for some weakly random problems in homogenization, *Multiscale Model. Simul.*, vol. 9 (2011) : 513-544.

[9] A. Anantharaman and C. Le Bris, Elements of mathematical foundations for a numerical approach for weakly random homogenization problems, *Communications in Computational Physics*, to appear, preprint at `http ://arxiv.org/abs/1005.3922`

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

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

[12] I. Babuska, B. Andersson, P. J. Smith, and K. Levin, Damage analysis of fiber composites. Part I : Statistical analysis on fiber scale, *Comp. Meth. in Appl. Mech. Eng.*, 172 :27-77, 1999.

[13] G. Bal, J. Garnier, S. Motsch and V. Perrier, Random integrals and correctors in homogenization, *Asymptotic Analysis*, 59(1-2) :1-26, 2008.

[14] G. Bal, Central limits and homogenization in random media, *Multiscale Model. Simul.*, 7(2), pp. 677-702, 2008.

[15] M. F. Ben Hassen and E. Bonnetier, An asymptotic formula for the voltage potential in a perturbed ε-periodic composite medium containing misplaced inclusions of size ε, *Proc. R. Soc. Edinb., Sect. A*, 136 :669-700, 2006.

[16] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, **Asymptotic analysis for periodic structures**, Studies in Mathematics and its Applications, 5. North-Holland Publishing Co., Amsterdam-New York, 1978.

[17] L. Berlyand and H. Owhadi, Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast, *Arch. Rat. Mech. Anal.* , vol. 198 (2), pp. 677-721 (2010).

[18] 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.

[19] X. Blanc, C. Le Bris and P.-L. Lions, Une variante de la théorie de l'homogénéisation stochastique des opérateurs elliptiques [A variant of stochastic homogenization theory for elliptic operators], *C. R. Acad. Sci. Série I*, 343(11-12) :717-724, 2006.

[20] X. Blanc, C. Le Bris and P.-L. Lions, Stochastic homogenization and random lattices, *J. Math. Pures Appl.*, 88(1) :34-63, 2007.

[21] A. Bourgeat, A. Mikelic and S. Wright : Stochastic two-scale convergence in the mean and applications, *J. Reine Angew. Math.* 456, pp. 19-51, 1994.

[22] A. Bourgeat and A. Piatnitski, Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator, *Asymptotic Analysis*, 21 :303-315, 1999.

[23] A. Bourgeat and A. Piatnitski, Approximation of effective coefficients in stochastic homogenization, *Ann I. H. Poincaré - PR*, 40(2) :153-165, 2004.

[24] A. Braides, **Γ-convergence for beginners**, Oxford University Press, 2002.

[25] M. Briane, D. Manceau and G.W. Milton, Homogenization of the two-dimensional Hall effect,*J. Math. Ana. App.*, 339 :1468-1484, 2008.

[26] L. Caffarelli, P. Souganidis, L. Wang, Homogenization of Fully Nonlinear, Uniformly Elliptic and Parabolic Partial Differential Equations in Stationary Ergodic Media, *Comm. in Pure Applied Math.*, to appear.

[27] D. Cioranescu and P. Donato, **An introduction to homogenization**, Oxford Lecture Series in Mathematics and its Applications, 17. Oxford University Press, New York, 1999.

[28] A.-L. Dalibard, Homogenization of linear transport equations in a stationary ergodic setting, *Communications in partial differential equations*, 33(5) : 881 :921, 2008.

[29] A.-L. Dalibard, Homogenization of nonlinear scalar conservation laws, *Archive for Rational Mechanics and Analysis* 192 : 117-164, 2009.

[30] G. Dal Maso, **An Introduction to Γ-convergence**, Progress in Nonlinear Differential Equations and Their Applications, Volume 8, Birkhaüser, 1993.

[31] G. Dal Maso, L. Modica, Nonlinear Stochastic Homogenization, **Annali di Matematica Pura ed Applicata**, Vol 144 (1) :347-389, 1986.

[32] E. De Giorgi and S. Spagnolo, Sulla convergenza degli integrali dell'energia per operatori elltici del 2 ordine, Boll. Un. Mat. Ital (4) 8(1973), 391-411.

[33] W. E and B. Engquist, The heterogeneous multiscale methods, *Comm. Math. Sci.*, vol. 1, no. 1, pp. 87-132, 2003.

[34] W. E and B. Engquist, The Heterogeneous Multiscale Method for Homogenization Problems, *Multiscale Methods in Science and Engeneering*, pp. 89-110, Lect. Notes in Comput. Sci. Eng., vol. 44, Springer, Berlin, 2005.

[35] W. E, B. Engquist, X. Li, W. Ren and E. Vanden-Eijnden, Heterogeneous multiscale methods : A review, *Comm. Comput. Phys.*, vol. 2, no. 3, pp. 367-450, 2007.

[36] Y.R. Efendiev, T.Y. Hou and X.H. Wu, Convergence of a nonconforming multiscale finite element method, *SIAM Journal on Numerical Analysis*, 37(3) :888-910, 2000.

[37] Y.R. Efendiev, T.Y. Hou and V. Ginting, Multiscale finite element methods for nonlinear problems and their applications, Communications in Mathematical Sciences, 2(4) :553-589, 2004.

[38] Y. Efendiev and T.Y. Hou, **Multiscale Finite Element method**, Theory and applications. Surveys and Tutorials in the Applied Mathematical Sciences, 4. Springer, New York, 2009.

[39] B. Engquist, Wavelet based numerical homogenization, *Documenta Mathematica.*, Extra Volume ICM III, pp. 503-512, 1998.

[40] B. Engquist and P. E. Souganidis, Asymptotic and numerical homogenization, *Acta Numerica*, 17 :147–190, 2008.

[41] J. D. Eshelby, The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 241(1226) :376–396, 1957.

[42] J.-P. Fouque, J. Garnier, G. Papanicolaou and K. Solna, **Wave propagation and time reversal in randomly layered media**, Springer, 2007.

[43] G. Francfort, F. Murat and L.Tartar, Homogenization of monotone operators in divergence form with x-dependent multivalued graphs, *Ann. Mat. Pura Appl.*, 188(4) :631-652, 2009.

[44] M. Jardak and R.G. Ghanem, Spectral stochastic homogenization of divergence-type PDEs, Comp. Meth. in Appl. Mech. and Eng. Vol. 193, Issues 6-8,pp 429-447, 2004

[45] D. Gilbarg and N. S. Trudinger, **Elliptic partial differential equations of second order**, reprint of the 1998 ed., Classics in Mathematics, Springer, 2001.

[46] A. Gloria, Reduction of the resonance error. Part 1 : Approximation of homogenized coefficients, *Math. Models Methods Appl. Sci. (M3AS)* vol. 21, Issue : 8 pp. 1601-1630, 2011.

[47] A. Gloria, Analytical framework for numerical homogenization. Part 2 : Windowing and oversampling, *Multiscale Model. Simul.* 7(1), pp 274-293, 2008.

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

[49] A. Gloria and F. Otto, An optimal error estimate in stochastic homogenization of discrete elliptic equations, *Ann. Appl. Probab.* vol. 39(3), pp. 779-856, 2011.

[50] A. Gloria and F. Otto, An optimal variance estimate in stochastic homogenization of discrete elliptic equations, *Ann. of Probab.*, 39 (2011), No 3, pp 779-856.

[51] F. Golse, Particle transport in non-homogeneous media, *Mathematical aspects of fluid and plasma dynamics*, Lecture Notes in Maths. vol 1460, pp 152-170, Springer, 1991.

[52] M. Grüter and K. O. Widman, The Green function for uniformly elliptic equations, *Manuscripta Math.*, 37 :303–342, 1982.

[53] P. Hall and C. C. Heyde, **Martingale limit theory and its application**, Probability and Mathematical Statistics, Academic Press, 1980.

[54] `FreeFEM, http ://www.freefem.org`

[55] T. Y. Hou and X.-H. Wu, A multiscale Finite Element Method for elliptic problems in composite materials and porous media, *Journal of Computational Physics*, 134(1) : 169-189 (1997)

[56] U. Hornung, **Homogenization and Porous Media**, Interdisciplinary Applied Mathematics, Springer.

[57] V. V. Jikov, S. M. Kozlov, O. A. Oleinik, H. T. Ngoan, Averaging and G-convergence of differential operators, *Uspekhi Mat. Nauk*, 34 :5(209) (1979), 65-133.

[58] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, **Homogenization of differential operators and integral functionals**, Springer-Verlag, 1994.

[59] T. Kanit, S. Forest, I. Galliet, V. Mounoury, and D. Jeulin, *Determination of the size of the representative volume element for random composites : statistical and numerical approach*, International Journal of Solids and Structures, 40(13-14) :3647 – 3679, 2003.

[60] S. M. Kozlov, Averaging of random operators, *Mat. Sb. (N.S.)*, 109(151) :2(6) (1979), pp. 188 ?202.

[61] U. Krengel, **Ergodic theorems**, de Gruyter Studies in Mathematics, vol. 6, de Gruyter, 1985.

[62] C. Le Bris, **Systèmes multi-échelles : modélisation et simulation**, Mathématiques et Applications, volume 47, Springer, 2005.

[63] C. Le Bris, Some numerical approaches for "weakly" random homogenization, *Proceedings of ENUMATH 2009*, G. Kreiss, P. Lötstedt, A. Malqvist and M. Neytcheva Eds., pp. 29-45, Springer, 2010 .

[64] C. Le Bris, *A selection of mathematical topics in multiscale sciences*, **Numerical Simulation in Physics and Engineering, Proceedings of the XIV Spanish-French Jacques-Louis Lions School**, Universidade da Coruña (Spain), 6-10 September 2010, ISBN 978-84-9749-420-5, p 153-233, 2010.

[65] C. Le Bris, F. Legoll, F. Thomines, Multiscale FEM for weakly random problems and related issues, en préparation.

[66] F. Legoll, F. Thomines, *Convergence of the residual process of a variant of stochastic homogenization in dimension one*, en préparation.

[67] P.-L. Lions, P. Souganidis, Correctors for the homogenization theory of Hamilton-Jacobi equations, *Comm Pure Applied Math.* LVI (2003) 1501-1524.

[68] P.-L. Lions, P. Souganidis, Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media, *Comm. in Partial Dif. Equations*, to appear.

[69] J. S. Liu, **Monte-Carlo strategies in scientific computing**, Springer Series in Statistics, 2001.

[70] D. Manceau, Small amplitude homogenization applied to models of non-periodic fibered materials", *M2AN Math. Model. Numer. Anal.* 41 (2007), no. 6, pp. 1061-1087.

[71] D. Manceau, Quelques problèmes d'homogénéisation à faible et fort contraste, thèse de doctorat de l'université Rennes 1, 2007.

[72] F. Merlevède, M. Peligrad and S. Utev, Recent advances in invariance principles for stationary sequences, *Probability surveys*, 3 :1–36, 2006.

[73] T. Mori and K. Tanaka, *Average stress in matrix and average elastic energy of materials with misfitting inclusions*, Acta Metallurgica, 21(5) :571 – 574, 1973.

[74] F. Murat, H-convergence, Séminaire d'analyse fonctionnelle et numérique d'Alger (1977-78).

[75] F. Murat, Compacité par compensation, Ann. Scuola Norm. Sup. Pisa. Cl. Sci. 5 (4) (1978), pp. 485-507.

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

[77] G. Nguetseng, A general convergence result for a functional related to the theory of homogenization, *SIAM J. Math. Anal.* 20 (3), pp 608-623, 1989.

[78] G. Nguetseng, Homogenization structures and applications I, *Z. Anal. Anwendungen*, 22 (1), pp 73-107, 2003.

[79] G. Nguetseng, Homogenization Structures and Applications II, *Z. Anal. Anwendungen*, 23, No. 3, pp 483-508, 2004.

[80] G. Nguetseng, Homogenization of perforated domains beyond the periodic setting, *J. Math. Anal. Appl.*, 289, pp 608-628, 2004.

[81] J. Nolen, G. Papanicolaou, O. Pironneau, A Framework for Adaptive Multiscale Methods for Elliptic Problems, *SIAM Multiscale Modeling and Simulation*, vol. 7 (2008), pp. 171-196.

[82] H. Owhadi, L. Zhang, Localized bases for finite dimensional homogenization approximations with non-separated scales and high-contrast, accepted for publication in *SIAM Multiscale Modeling and Simulation*, 2011.

[83] G. Papanicolaou and S. R. S. Varadhan, Boundary value problems with rapidly oscillating random coefficients. In Proc. Colloq. on Random Fields : Rigorous Results in Statistical Mechanics and Quantum Field Theory (J. Fritz, J. L. Lebaritz and D. Szasz, eds), Vol. 10 of *Colloquia Mathematica Societ. Janos Bolyai*, pp. 835-873.

[84] G. Papanicolaou and S. R. S. Varadhan, Diffusions with random coefficients, Statistics and probability : essays in honor of C. R. Rao, pp. 547-552, North-Holland, Amsterdam, 1982.

[85] G. Pavliotis, A.M. Stuart, **Multiscale Methods : Averaging and Homogenization**, Texts in Applied Mathematics, Vol. 53, Springer, 2008.

[86] M. Peligrad and S. Utev, Central limit theorem for stationary linear processes, *The Annals of Probability*, 34(4) :1608–1622, 2006.

[87] K. Sab, On the homogenization and the simulation of random materials, *Eur. J. Mech.* A/Solids 11 (1992) 585–607.

[88] S. Sakata, F. Ashida, T. Kojima, and M. Zako, Three-dimensional stochastic analysis using a perturbation-based homogenization method for elastic properties of composite material considering microscopic uncertainty, *Int. J. Solids and Structures*, 45 :894-907, 2008.

[89] A. N. Shiryaev, **Probability**, Graduate Texts in Mathematics, vol. 95, Springer, 1984.

[90] P. Souganidis, Stochastic homogenization for Hamilton-Jacobi equations and applications, *Asympt. Anal.* 20 (1999) 1-11.

[91] P. Souganidis, Recent developments in the theory of homogenization for fully nonlinear first- and second-order PDE in random environments, *Bol. Soc. Esp. Mat. Apl. SEMA*, 34, 2006, 81-90.

[92] S. Spagnolo, Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche, *Annali della Scuola Normale Superiore di Pisa*, Classe di Scienze 3e série, tome 22, no 4 (1968), p. 571-597.

[93] S. Spagnolo, Convergence in energy for elliptic operators, Proc. Third Symp. Numer. Solut. Partial. Dif. Equat., Academic Press, New York (1976), pp. 469-498

[94] L. Tartar, Compensated compactness and applications to partial differential equations, *Nonlinear analysis and mechanics : Heriot-Watt Symp.*, Vol. 4, Edinburgh 1979, Res. Notes Math. 39, 136-212, 1979.

[95] L. Tartar, *Estimations fines des coefficients homogénéisés*, in *Ennio De Giorgi colloquium (Paris, 1983)*, volume 125 of *Res. Notes in Math.*, pages 168–187. Pitman, Boston, MA, 1985.

[96] L. Tartar, H-measures and small amplitude homogenization, in *Random media and composites*, Proc. SIAM Workshop, Leesburg/VA (USA) 1988, 89–99, 1989.

[97] L. Tartar, H-measures, a new approach for studying homogenization, oscillations and concentration effects in partial differential equations, *Proc. R. Soc. Edinb., Sect. A*, 115(3-4) :193-230, 1990.

[98] L. Tartar, Homogénéisation et H-mesures, *ESAIM : Proc.*, Vol. 6, pp. 111-131, 1999.

[99] L. Tartar, Estimations of homogenized coefficients, in *Topics in the mathematical modelling of composite materials*, A. Cherkaev and R. Kohn eds., Progress in nonlinear differential equations and their applications, vol. 31, Birkhäuser, 1987.

[100] M. Thomas, *Propriétés thermiques de matériaux composites : caractérisation expérimentale et approche microstructurale*, [Thermal properties of composite materials : experimental characterization and microstructural approach], Thèse de l' Université de Nantes, Laboratoire de Thermocinétique, CNRS - UMR 6607, 2008.

[101] F. Thomines, Thèse de l'Université Paris-Est, en préparation.

[102] J. R. Willis, *Bounds and self-consistent estimates for the overall properties of anisotropic composites*, Journal of the Mechanics and Physics of Solids, 25(3) :185–202, 1977.

[103] V. V. Yurinskii, Averaging of symmetric diffusion in random medium, *Sibirskii Mat. Zh.*, 27(4) :167–180, 1986.