Mathematical Models and Methods in Applied Sciences
Vol. 21, No. 8 (2011) 1601–1630

DOI: 10.1142/S0218202511005507

REDUCTION OF THE RESONANCE ERROR — PART 1: APPROXIMATION OF HOMOGENIZED COEFFICIENTS

ANTOINE GLORIA

Projet SIMPAF, INRIA Lille — Nord Europe, France
antoine.gloria@inria.fr

Received 16 February 2010
Revised 23 July 2010
Communicated by F. Brezzi

This paper is concerned with the approximation of effective coefficients in homogenization of linear elliptic equations. One common drawback among numerical homogenization methods is the presence of the so-called resonance error, which roughly speaking is a function of the ratio ε/η, where η is a typical macroscopic length scale and ε is the typical size of the heterogeneities. In the present work, we propose an alternative for the computation of homogenized coefficients (or more generally a modified cell-problem), which is a first brick in the design of effective numerical homogenization methods. We show that this approach drastically reduces the resonance error in some standard cases.

Keywords: Numerical homogenization; resonance error; effective coefficients.

AMS Subject Classification: 35B27, 60F99

1. Introduction

This paper deals with numerical homogenization. In essence, numerical homogenization techniques aim at computing solutions to partial differential equations whose coefficients depend on a small parameter denoted by ε, without resolving all the details at the scale ε (see for instance Refs. 16 and 17 for the multiscale finite element method, Refs. 6 and 5 for the heterogeneous multiscale method, and Ref. 8 for a general analysis). In the case of periodic coefficients, it amounts to approximating the solution to the homogenized problem.

To be more precise, let us consider the scalar linear elliptic equation

$$\begin{cases} -\nabla \cdot A_\varepsilon(x)\nabla u_\varepsilon(x) = f(x) & \text{in } D, \\ u_\varepsilon(x) = 0 & \text{on } \partial D \end{cases} \tag{1.1}$$

on a domain D with suitable assumptions on A_ε and f. Assume furthermore that A_ε is symmetric and can be homogenized in the following sense: There exists $A_{\rm hom}$ such

that the solution u_ε to (1.1) converges to the solution $u_{\rm hom}$ to

$$\begin{cases} -\nabla \cdot A_{\rm hom}(x)\nabla u_{\rm hom}(x) = f(x) & \text{in } D, \\ u_{\rm hom}(x) = 0 & \text{on } \partial D \end{cases} \tag{1.2}$$

for all suitable f (as well as the flux $A_\varepsilon \nabla u_\varepsilon$ to $A_{\rm hom}\nabla u_{\rm hom}$). Such a homogenization property typically arises when A_ε is the combination of a smooth function and an oscillating part at scale $\varepsilon > 0$. Unfortunately, $A_{\rm hom}$ is not explicit in general. The aim of numerical homogenization can now be rephrased as: Given A_ε (and not $A_{\rm hom}$), approximate $u_{\rm hom}$ without solving (1.1) at scale ε.

A very general approach is as follows. For all $\eta \geq \varepsilon > 0$, let $A_{\eta,\varepsilon}$ be defined by

$$\xi \cdot A_{\eta,\varepsilon}(x)\xi := \inf\left\{ \fint_{D\cap B(x,\eta)} (\xi + \nabla\phi)\cdot A_\varepsilon(y)(\xi+\nabla\phi)\,dy, \phi \in H^1_0(D\cap B(x,\eta)) \right\} \tag{1.3}$$

for all ξ, where $B(x,\eta)$ is the ball centered at x and of radius η. An approximation of $u_{\rm hom}$ is then given by $u_{\eta,\varepsilon}$, solution to

$$\begin{cases} -\nabla \cdot A_{\eta,\varepsilon}(x)\nabla u_{\eta,\varepsilon}(x) = f(x) & \text{in } D, \\ u_{\eta,\varepsilon}(x) = 0 & \text{on } \partial D. \end{cases} \tag{1.4}$$

In particular (see Ref. 8 for general results), we have

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} A_{\eta,\varepsilon} = A_{\rm hom} \tag{1.5}$$

in $L^p(D)$ for all $p < \infty$, and

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} \|u_{\eta,\varepsilon} - u_{\rm hom}\|_{H^1(D)} = 0. \tag{1.6}$$

The main difference between (1.4) and (1.1) is that $A_{\eta,\varepsilon}$ is expected not to oscillate at scale ε, which is a big advantage for the numerical practice.

In order to make a numerical analysis of the above method, and in particular to quantify the convergences (1.5) and (1.6), we turn to the simplest case possible: $A_\varepsilon(x) = A(x/\varepsilon)$, where A is periodic. Then, one can prove that

$$A_{\eta,\varepsilon} = A_{\rm hom} + O(\varepsilon/\eta) \tag{1.7}$$

and

$$\|u_{\eta,\varepsilon} - u_{\rm hom}\|_{H^1(D)} = O(\varepsilon/\eta). \tag{1.8}$$

The term $O(\varepsilon/\eta)$ is called the *resonance error*. In practice, that is when (1.3) and (1.4) are solved numerically (finite element, finite volume, finite difference, FFT methods etc.), the dominant term in the overall error can precisely be the resonance error $O(\varepsilon/\eta)$. Several refinements have then been introduced to reduce this error (such as oversampling in Ref. 18, and filtering in Ref. 23). However, as shown in Ref. 9, even with these refinements, although the *prefactor* may have been reduced, the error is still of order $O(\varepsilon/\eta)$.

The aim of this paper is to reduce the order of magnitude of the resonance error. Let us first perform a change of variables to make the length scale of the oscillations of A_ε be of order 1, and rewrite (1.3) as

$$\xi \cdot A_{\eta,\varepsilon}(x)\xi := \inf\left\{ \fint_{\frac{D-x}{\varepsilon}\cap B(0,\frac{\eta}{\varepsilon})} (\xi + \nabla\phi_{\eta,\varepsilon}) \cdot A(y)(\xi + \nabla\phi_{\eta,\varepsilon})\, dy, \right.$$
$$\left. \phi_{\eta,\varepsilon} \in H_0^1\left(\frac{D-x}{\varepsilon} \cap B\Big(0,\frac{\eta}{\varepsilon}\Big)\right)\right\}, \tag{1.9}$$

where $A(y) := A_\varepsilon(x + \varepsilon y)$. Essentially, (1.9) can be seen as an approximation of the averaged energy density

$$\mathcal{M}((\xi + \nabla\phi) \cdot A(y)(\xi + \nabla\phi)) := \lim_{R\to\infty} \frac{1}{|Q_R|} \int_{Q_R} (\xi + \nabla\phi) \cdot A(y)(\xi + \nabla\phi)\, dx, \tag{1.10}$$

where $Q_R = (-R, R)^d$, and ϕ is a solution (in a suitable sense) to

$$-\nabla \cdot A(\xi + \nabla\phi) = 0 \quad \text{in } \mathbb{R}^d. \tag{1.11}$$

This paper is devoted to the approximation of (1.10). Our method is inspired by the analysis by Otto and the author in Refs. 12, 13 and 10 and a recent article Ref. 3 by Blanc and Le Bris. We show that for some benchmark tests, the proposed method effectively reduces the resonance error. In particular, in the periodic case, estimate (1.7) is replaced by

$$A_{\eta,\varepsilon} = A_{\text{hom}} + O(\varepsilon/\eta)^p$$

for all $p < 4$. In the case of stochastic homogenization of discrete elliptic equations, the method also performs quite well, as recently proved by Otto and the author in Refs. 12, 13 and 10. The coupling with numerical homogenization methods will be discussed in a subsequent work.

The paper is organized as follows. In Sec. 2, we precisely define the resonance error when approximating homogenized coefficients and introduce our new method, that is an alternative formula to (1.9), see (2.5). The convergence of the method is analyzed in Sec. 3 for periodic coefficients (and results on the stochastic case are recalled). A numerical study completes the analysis in Sec. 4.

We make use of the following notation:

- $d \geq 1$ is the dimension;
- $\lesssim$ and $\gtrsim$ stand for $\leq$ and $\geq$ up to a multiplicative constant which only depends on the dimension d and the constants α, β (see Definition 3.2 below) if not otherwise stated;
- when both $\lesssim$ and $\gtrsim$ hold, we simply write $\sim$;
- we use $\gg$ instead of $\gtrsim$ when the multiplicative constant is (much) larger than 1.

2. Resonance Error and Proposed Strategy

In this section, we assume (as this is the case in periodic, quasi-periodic or stochastic homogenization) that the corrector problem is posed on $\mathbb{R}^d$, the corrector ϕ being solution to

$$-\nabla \cdot A(\xi + \nabla\phi) = 0 \quad \text{in } \mathbb{R}^d, \tag{2.1}$$

where A is a symmetric matrix. The homogenized coefficients are then given by

$$\xi \cdot A_{\text{hom}}\xi = \mathcal{M}\{(\xi + \nabla\phi) \cdot A(\xi + \nabla\phi)\}, \tag{2.2}$$

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

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

and $Q_R := (-R, R)^d$. Such quantities are well-defined in periodic, quasi-periodic and stochastic homogenization (see for instance the monograph Ref. 19).

To numerically compute A_{hom} via (2.2), one needs to approximate the corrector field ϕ and the average operator $\mathcal{M}\{\cdot\}$. These two approximations lead to the so-called resonance error.

2.1. *The resonance error*

The simplest way to approximate ϕ and $\mathcal{M}\{\cdot\}$ consists of solving (2.1) on a large domain $Q_R = (-R, R)^d$ (with suitable boundary conditions, say homogeneous Dirichlet), and taking the average of the energy density on Q_R. Doing so, and recalling Ref. 9, we make at least two errors:

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

More precisely, A_{hom} is approximated by

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

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

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

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

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

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

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

where μ_R is typically a smooth non-negative mask such that

$$\begin{aligned} \|\nabla \mu_R\|_{L^\infty} &\lesssim R^{-d-1}, \\ \int_{Q_R} \mu_R(x)\, dx &= 1, \\ \mu_{R|Q_R \setminus Q_{R/2}} &\equiv 0, \end{aligned}$$

we may hope to reduce both sources of the error. However, the overall error is still of order

$$|\tilde{A}_R - A_{\text{hom}}| \sim \frac{1}{R} \tag{2.3}$$

in any dimension, as already noticed by E and Yue in Ref. 23. Only the prefactor may have been reduced (see a related numerical test on Fig. 7).

In the following section, we propose to treat separately the two sources of error. The geometric error is an error localized at the boundary, and a filtering method with a suitable mask is enough to significantly reduce it. The error we make on the boundary conditions has however nonlocal effects due to the poor decay of the Green function of the Laplace operator. To reduce this effect, it is natural to add a zeroth-order term to the equation, which makes the associated Green function decay exponentially fast. This allows one to drastically reduce the spurious effect of the boundary condition away from a boundary layer. Yet, this modifies the corrector equation and introduces a bias, which has to be quantified. The last task consists in suitably choosing the different parameters at stake.

2.2. *Proposed strategy*

As a proxy for the corrector field ϕ solution to (2.1), we consider $\phi_{T,R}$, solution to

$$\begin{cases} T^{-1}\phi_{T,R} - \nabla \cdot A(\xi + \nabla \phi_{T,R}) = 0 & \text{in } Q_R, \\ \phi_{T,R} = 0 & \text{on } \partial Q_R, \end{cases} \tag{2.4}$$

where $T > 0$ controls the importance of the zeroth-order term and $R > 0$ is the size of the finite domain Q_R. We then approximate the homogenized coefficient by taking the filtered average

$$\xi \cdot A_{T,R,L} \xi := \int_{Q_R} (\xi + \nabla \phi_{T,R}(x)) \cdot A(x)(\xi + \nabla \phi_{T,R}(x)) \mu_L(x)\, dx, \tag{2.5}$$

where μ_L is a smooth mask (whose properties will be fixed in Definition 3.1) supported in $Q_L = (-L, L)^d$, $L \leq R$. Note that we do not consider the full energy associated

with (2.4) and disregard the contribution of the zeroth-order term $T^{-1}\phi_{T,R}(x)^2$ in (2.5). The reason for this choice is made clear in the following discussion.

In order to choose the parameters T, R, L properly, we first make a coarse analysis of the error, that we split into three parts:

$$|A_{T,R,L} - A_{\rm hom}| \leq |A_T - A_{\rm hom}| + |A_{T,L} - A_T| + |A_{T,R,L} - A_{T,L}|,$$

where the different quantities are described below. We first define an approximate corrector ϕ_T, solution to

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

Such a function is well-defined in the periodic, quasi-periodic and stochastic cases. The matrix A_T is then characterized by

$$\xi \cdot A_T \xi := \mathcal{M}\{(\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\}.$$

The error term $|A_T - A_{\rm hom}|$ depends on T and on the structure properties of A (periodicity, quasi-periodicity, stochastic stationarity, etc.). As will be seen in the proof for the periodic case, we have the following universal lower bound:

$$|A_T - A_{\rm hom}| \gtrsim T^{-2}.$$

Note that if we had added the zeroth-order term $T^{-1}\mathcal{M}\{\phi_T^2\}$ in the definition of A_T, the error $|A_T - A_{\rm hom}|$ would have been at least of order T^{-1}. This motivates the definition of A_T, and more generally of (2.5). We refer to the recent work Ref. 11 by Mourrat and the author for related questions, and other pertinent approximations. The second error term $|A_T - A_{T,L}|$, where

$$\xi \cdot A_{T,L}\xi := \int_{Q_L} (\xi + \nabla\phi_T(x)) \cdot A(x)(\xi + \nabla\phi_T(x))\mu_L(x)dx,$$

is partly geometric. It can be reduced if a suitable mask μ_L is used. Note that, as opposed to $\phi_{T,R}$, ϕ_T has the same structure property as A (periodicity, quasi-periodicity, stationarity, etc.), which is a big advantage for the analysis (it is crucial for the analysis of the discrete stochastic case in Ref. 12). The last error term $|A_{T,R,L} - A_{T,L}|$ can be treated using standard elliptic estimates. It is essentially of infinite order in units of $(R-L)/\sqrt{T}$.

The combination of these three error terms allows us to make reasonable choices for L, R, T: $R \sim L$ and $L \gg \sqrt{T}$. The error is at least of order T^{-2}, hence R^{-4} (recall that R quantifies the cost to compute $\phi_{T,R}$). As will be seen in the following section, this strategy outperforms (both in terms of precision and/or computational cost) most of the other existing methods.

3. Analysis of Some Standard Cases

Before turning to the analysis proper, let us make precise the form of the masks we use.

Definition 3.1. A function $\mu : [-1,1] \to \mathbb{R}^+$ is said to be a filter of order $p \geq 0$ if

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

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

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

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

3.1. *The periodic case*

Let us first introduce the class of matrices we consider.

Definition 3.2. For all $d \geq 1$, $\beta \geq \alpha > 0$, we define $\mathcal{A}_{\alpha\beta}$ as the set of those symmetric matrices of order d such that for all $x \in \mathbb{R}^d$,

$$x \cdot Ax \geq \alpha|x|^2,$$
$$|Ax| \leq \beta|x|.$$

For all D open subset of $\mathbb{R}^d$, and all $A : D \to \mathcal{A}_{\alpha\beta}$, we use the shorthand notation $A \in \mathcal{A}_{\alpha\beta}$.

Theorem 3.1. *Let $d \geq 2$, $A \in \mathcal{A}_{\alpha\beta}$ be Q-periodic, μ be a filter of order $p \geq 0$, and A_{hom} and $A_{T,R,L}$ be the homogenized matrix and its approximation (2.5) respectively, where $R^2 \gtrsim T \gtrsim R$, $R \geq L \sim R \sim R - L$. Then, there exists $c > 0$ depending only on α, β and d such that we have*

$$|A_{T,R,L} - A_{\text{hom}}| \lesssim L^{-(p+1)} + T^{-2} + T^{1/4} \exp\left(-c\frac{R-L}{\sqrt{T}}\right). \tag{3.1}$$

We postpone the proof to Sec. 3.3, and directly turn to an application of Theorem 3.1. For $p \geq 3$, the rate in (3.1) is controlled by the last two terms. In particular, the last term requires T to be such that $L \gg \sqrt{T}$. A possible choice is then given by

- $T = L^2(\ln L)^{-4}$,
- $R = 3L/2$,

for which (3.1) reads:

$$|A_{T,R,L} - A_{\text{hom}}| \lesssim R^{-4} \ln^8 R. \tag{3.2}$$

Whereas estimate (2.3) in the basic approach of Sec. 2.1 is of order 1, the present approach yields an estimate (3.2) up to order 4^-.

Note that recently, Blanc and Le Bris have developed another strategy in Ref. 3, where essentially the mask is introduced in the very definition of the bilinear form associated with the equation, and not as a post-processing tool. Their formal analysis and numerical tests show a convergence of order 2 in the periodic case, which cannot be improved in general. The better result (3.2) of the present strategy is mainly due to the efficient treatment, by the zeroth-order term, of the spurious effects of the boundary conditions.

3.2. *A stochastic example*

In this section, we quickly review the results obtained by Otto and the author in Refs. 12, 13 and 10, since they complement the analysis of the periodic case quite well. In these papers, the elliptic equation is discrete, and A is a set of conductivities on the edges of $\mathbb{Z}^d$. The strategy remains the same, and we keep the notation of Sec. 2.1. The results are as follows: Let $d \geq 2$, $A \in \mathcal{A}_{\alpha\beta}$ be an independently and identically distributed conductivity function in the sense of Ref. 12, and $A_{\rm hom}$, A_T, $A_{T,L}$ and $A_{T,R,L}$ be the homogenized matrix and its approximations, where $T > 0$, $R \geq L \sim R \sim R - L$ and μ_L as in Definition 3.1 with $p \geq 1$.[a] Then, there exist $c > 0$ and $q > 0$ depending only on α, β and d such that we have

$$|A_T - A_{\rm hom}| \lesssim \begin{cases} d = 2 \ : \ T^{-1} \ln^q T \\ d = 3 \ : \ T^{-3/2} \\ d = 4 \ : \ T^{-2} \ln T \\ d > 4 \ : \ T^{-2} \end{cases}, \tag{3.3}$$

$$\langle |A_{T,L} - A_T|^2 \rangle^{1/2} \lesssim \begin{cases} d = 2 \ : \ L^{-1} \ln^q T \\ d > 2 \ : \ L^{-d/2} \end{cases}, \tag{3.4}$$

$$|A_{T,R,L} - A_{T,L}| \overset{\text{almost surely}}{\lesssim} T^{3/4} \exp\left(-c \frac{R-L}{\sqrt{T}}\right), \tag{3.5}$$

where $\langle \cdot \rangle$ denotes the ensemble average, or equivalently expectation in the underlying probability space. Let us comment on the above results. In the periodic case, the estimate $L^{-(p+1)}$ for the error term $|A_{T,L} - A_T|$ can be made of any order provided the use of an appropriate mask μ_L. In the stochastic case, the order of accuracy is naturally limited by the central limit theorem scaling, which we recognize in (3.4) (up to the logarithmic correction for $d = 2$). As for the periodic case, the estimate (3.5) for the error term $|A_{T,R,L} - A_{T,L}|$ is of infinite order in units of $(R-L)/\sqrt{T}$ (up to the multiplicative constant $T^{3/4}$). Finally, the estimate (3.3) of the error term $\langle |A_T - A_{\rm hom}|^2 \rangle^{1/2}$ also saturates at T^{-2} in high dimension, but further depends on the dimension for $d < 4$, which is a consequence of the stochastic structure. Note that

[a] The role of the mask is different in this case since the conductivities are i.i.d., so that there is *a priori* no "geometric error" involved.

the class of approximation formulas introduced by Mourrat and the author in Ref. 11 allow one to reach the convergence rate of the central limit theorem in any dimension.

In this stochastic example, a natural choice for T, L, R is as follows:

- $T = R$,
- $L = R(1 - \ln^2 R/\sqrt{R})$,

and the global error estimate reads

$$\langle |A_{T,R,L} - A_{\rm hom}|^2 \rangle^{1/2} \lesssim \begin{cases} d = 2 \ : \ R^{-1} \ln^q R, \\ d = 3 \ : \ R^{-3/2}, \\ d = 4 \ : \ R^{-2} \ \ln R, \\ d > 4 \ : \ R^{-2}, \end{cases} \tag{3.6}$$

which is sharp.

The extension of these results to the continuous setting is currently under investigation.

3.3. *Proof in the periodic case*

We divide the proof of Theorem 3.1 in three steps. Let $H^1_{\rm per}(Q)$ denote the subspace of periodic functions of $H^1(Q)$ with zero average, and $\xi \in \mathbb{R}^d$ with $|\xi| = 1$.

Step 1. Proof of

$$|A_T - A_{\rm hom}| \lesssim T^{-2}. \tag{3.7}$$

We recall that $A_{\rm hom}$ is given by

$$\xi \cdot A_{\rm hom} \xi = \fint_Q (\xi + \nabla \phi) \cdot A(\xi + \nabla \phi),$$

where ϕ is the unique weak solution in $H^1_{\rm per}(Q)$ to

$$-\nabla \cdot A(\xi + \nabla \phi) = 0. \tag{3.8}$$

We have, using Eq. (3.8) and the symmetry of A,

$$\begin{aligned} \xi \cdot (A_T - A_{\rm hom}) \xi &= \int_Q (\xi + \nabla \phi_T) \cdot A(\xi + \nabla \phi_T) - (\xi + \nabla \phi) \cdot A(\xi + \nabla \phi) \\ &= \int_Q (\xi + \nabla \phi_T) \cdot A \nabla (\phi_T - \phi) + \nabla (\phi_T - \phi) \cdot A(\xi + \nabla \phi) \\ &\overset{(3.8)}{=} \int_Q (\xi + \nabla \phi_T) \cdot A \nabla (\phi_T - \phi) - \nabla (\phi_T - \phi) \cdot A(\xi + \nabla \phi) \\ &= \int_Q \nabla (\phi_T - \phi) \cdot A \nabla (\phi_T - \phi). \end{aligned}$$

Introducing ψ_T defined by

$$\psi_T = -T(\phi_T - \phi),$$

this identity turns into

$$\xi \cdot (A_T - A_{\rm hom})\xi = T^{-2} \int_Q \nabla\psi_T \cdot A\nabla\psi_T. \tag{3.9}$$

Note that ψ_T is the unique solution in $H^1_{\rm per}(Q)$ to

$$T^{-1}\psi_T - \nabla \cdot A\nabla\psi_T = \phi. \tag{3.10}$$

We then appeal to Eq. (3.10) in the form of the *a priori* estimate

$$T^{-1}\int_Q \psi_T^2 + \int_Q \nabla\psi_T \cdot A\nabla\psi_T = \int_Q \phi\psi_T,$$

which turns into

$$\int_Q |\nabla\psi_T|^2 \lesssim \int_Q \phi^2,$$

by uniform ellipticity of A and Poincaré's inequality in $H^1_{\rm per}(Q)$. Combined with (3.9), this yields

$$\begin{aligned} |\xi \cdot (A_T - A_{\rm hom})|\xi &\lesssim T^{-2}\int_Q \phi^2 \\ &\lesssim T^{-2}\int_Q |\nabla\phi|^2 \\ &\lesssim T^{-2}|\xi|^2 = T^{-2}, \end{aligned}$$

using Poincaré's inequality in $H^1_{\rm per}(Q)$, and an *a priori* estimate on $\nabla\phi$. This proves (3.7).

Step 2. Proof of

$$|A_T - A_{T,L}| \lesssim L^{-(p+1)}. \tag{3.11}$$

Estimate (3.11) would be a consequence of the following lemma if the energy density

$$\mathcal{E}_T : x \mapsto (\xi + \nabla\phi_T(x)) \cdot A(x)(\xi + \nabla\phi_T(x)),$$

which is a Q-periodic function, were square integrable uniformly in T.

Lemma 3.1. *Let μ be a filter of order $p \geq 0$ according to Definition 3.1. Then, for all $\phi \in L^2_{\rm per}(Q)$ and all $L > 0$, we have*

$$\left| \int_{Q_L} \phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi) \right| \lesssim L^{-(p+1)}\|\phi\|_{L^2(Q)}, \tag{3.12}$$

where $Q_L = (-L, L)^d$ and $\mu_L(x) := L^{-d}\prod_{i=1}^d \mu(L^{-1}x_i)$, $x = (x_1, \ldots, x_d)$. The constant in (3.12) only depends on p and μ.

Note that Lemma 3.1 holds for general periodic functions (not necessarily Q-periodic).

Yet, $\mathcal{E}_T$ is only in $L^1(Q)$. As proved in Appendix A, the higher integrability $\mathcal{E}_T \in L^q(Q)$ for some $q > 1$ is enough to conclude. This higher integrability of the energy density is a standard consequence of Meyers' estimate (see the original paper by Meyers Ref. 20 and its use for homogenization problems by Murat and Tartar in Ref. 21), noting that q depends on α, β and d, but not on T (see for instance the argument in Ref. 12 in the discrete case).

Step 3. Proof of

$$|A_{T,R,L} - A_{T,L}| \lesssim T^{1/4} \exp\left(-c\frac{R-L}{\sqrt{T}}\right). \tag{3.13}$$

Our proof of (3.13), which is self-contained, slightly departs from the approach by Bourgeat and Piatnitski in Ref. 4. The argument is based on the exponential decay of the Green function, which is the object of the following auxiliary lemma.

Lemma 3.2. *Let G_T be the Green function associated with the operator $(T^{-1} - \nabla \cdot A\nabla)$ on $\mathbb{R}^d$. Then there exists $c > 0$ depending only on α, β and d such that the following pointwise estimate holds for all $|x - y| \gtrsim \sqrt{T}$:*

$$G_T(x,y) \lesssim \frac{1}{|x-y|^{d-2}} \exp\left(-c\frac{|x-y|}{\sqrt{T}}\right). \tag{3.14}$$

Although this result is common knowledge, we did not find any precise reference for it. A proof is given in Appendix B.

By definition of ϕ_T and $\phi_{T,R}$, we have

$$\begin{cases} T^{-1}(\phi_T - \phi_{T,R}) - \nabla \cdot A(\nabla\phi_T - \nabla\phi_{T,R}) = 0 & \text{in } Q_R, \\ \phi_T - \phi_{T,R} = \phi_T & \text{on } \partial Q_R. \end{cases}$$

We then consider a lifting ϕ_1 of $\phi_{T|\partial Q_R}$ on Q_R such that $\phi_{1|Q_{R-1}} \equiv 0$ and $\|\phi_1\|^2_{H^1(Q_R)} \lesssim R^{d-1}\|\phi_T\|^2_{H^1(Q)}$ (recall that ϕ_T is Q-periodic). The function $\phi_2 := \phi_T - \phi_{T,R} - \phi_1$ then satisfies the equation

$$\begin{cases} T^{-1}\phi_2 - \nabla \cdot A\nabla\phi_2 = -T^{-1}\phi_1 + \nabla \cdot A\nabla\phi_1 & \text{in } Q_R, \\ \phi_2 = 0 & \text{on } \partial Q_R. \end{cases} \tag{3.15}$$

Let $G_{T,R} : Q_R \times Q_R \to \mathbb{R}^+$ be the Green function associated with the operator $(T^{-1} - \nabla \cdot A\nabla)$ on Q_R with homogeneous Dirichlet boundary conditions. The function ϕ_2 can be written as

$$\phi_2(x) = -\int_{Q_R} (T^{-1}\phi_1(y) G_{T,R}(x,y) + \nabla G_{T,R}(x,y) \cdot A(y)\nabla\phi_1(y))\, dy.$$

By Cauchy−Schwarz's inequality, this becomes

$$|\phi_2(x)| \leq \|\phi_1\|_{H^1(Q_R)} \left(T^{-1} \left(\int_{Q_R \setminus Q_{R-1}} G_{T,R}(x,y)^2 dy \right)^{1/2} + \left(\int_{Q_R \setminus Q_{R-1}} |\nabla G_{T,R}(x,y)|^2 dy \right)^{1/2} \right). \tag{3.16}$$

To control the first term on the right-hand side of (3.16), we use Lemma 3.2, which, combined with the maximum principle, yields an estimate for $0 \leq G_{T,R} \leq G_T$. For the second term of the right-hand side of (3.16), we use Cacciopoli's inequality. To this aim, let $\eta : Q_R \to \mathbb{R}^+$ be a function of class C^1. We multiply the defining equation for $G_{T,R}$ by $\eta^2 G_{T,R}$, integrate on Q_R, and obtain after integration by parts

$$\begin{aligned} 0 = {} & T^{-1} \int_{Q_R} \eta^2(y) G_{T,R}(x,y)^2 dy \\ & + \int_{Q_R} \nabla G_{T,R}(x,y) \cdot A(y) \nabla (\eta(y)^2 G_{T,R}(x,y)) dy \\ = {} & T^{-1} \int_{Q_R} \eta^2(y) G_{T,R}(x,y)^2 dy \\ & + \int_{Q_R} \nabla (\eta(y) G_{T,R}(x,y)) \cdot A(y) \nabla (\eta(y) G_{T,R}(x,y)) dy \\ & - \int_{Q_R} G_{T,R}(x,y)^2 \nabla \eta(y) \cdot A(y) \nabla \eta(y) dy, \end{aligned}$$

provided $\eta_{|\mathcal{N}(x)} \equiv 0$ on an open neighborhood $\mathcal{N}(x)$ of x. Hence, by the uniform bounds on A,

$$\int_{Q_R} |\nabla (\eta(y) G_{T,R}(x,y))|^2 dy \lesssim \int_{Q_R} G_{T,R}(x,y)^2 |\nabla \eta(y)|^2 dy. \tag{3.17}$$

Taking now $\sqrt{T} \leq \rho \leq R/2$, and $\eta : Q_R \to [0,1]$ such that

$$\begin{aligned} \text{for } y \in Q_{R-\rho/2} &: \eta(y) = 0, \\ \text{for } y \in Q_R \setminus Q_{R-1} &: \eta(y) = 1, \\ \text{for } y \in Q_R &: |\nabla \eta(y)| \lesssim \rho^{-1}, \end{aligned}$$

(3.17) turns into

$$\int_{Q_R \setminus Q_{R-1}} |\nabla G_{T,R}(x,y)|^2 dy \lesssim \rho^{-2} \int_{Q_R \setminus Q_{R-\rho/2}} G_{T,R}(x,y)^2 dy \tag{3.18}$$

for all $x \in Q_{R-\rho}$. Inserting now (3.18) into (3.16) and using Lemma 3.2 to control $G_{T,R}$, we obtain for all $x \in Q_{R-\rho}$

$$\begin{aligned}
|\phi_2(x)| &\lesssim \|\phi_1\|_{H^1(Q_R)} \left(T^{-1} \left(\int_{Q_R \setminus Q_{R-1}} G_{T,R}(x,y)^2 \, dy \right)^{1/2} \right. \\
&\quad \left. + \rho^{-1} \left(\int_{Q_R \setminus Q_{R-\rho/2}} G_{T,R}(x,y)^2 \, dy \right)^{1/2} \right) \\
&\lesssim R^{(d-1)/2} \left(T^{-1} R^{(d-1)/2} \rho^{2-d} \exp\left(-c \frac{\rho}{\sqrt{T}} \right) \right. \\
&\quad \left. + R^{d/2} \rho^{-1} \rho^{2-d} \exp\left(-c \frac{\rho}{2\sqrt{T}} \right) \right) \\
&\leq \left(\frac{R^{d-1}}{\rho^{d-2} T} + \frac{R^{d-1/2}}{\rho^{d-1}} \right) \exp\left(-c \frac{\rho}{2\sqrt{T}} \right).
\end{aligned}$$

Combined with the assumption $R^2 \gtrsim T \gtrsim R$ and taking $\rho = R/4$, this inequality can be further simplified to

$$|\phi_2(x)| \lesssim T^{1/4} \exp\left(-c \frac{\rho}{2\sqrt{T}} \right), \tag{3.19}$$

for some slightly smaller $c > 0$, noting that

$$\rho^{1/2} \exp\left(-c \frac{\rho}{2\sqrt{T}} \right) = T^{1/4} \left(\frac{\rho}{\sqrt{T}} \right)^{1/2} \exp\left(-c \frac{\rho}{2\sqrt{T}} \right).$$

Hence,

$$\int_{Q_{R-\rho}} \phi_2(x)^2 \, dx \lesssim (R-\rho)^d \sqrt{T} \exp\left(-c \frac{\rho}{\sqrt{T}} \right).$$

Another use of Cacciopoli's inequality, this time for ϕ_2 (recall that the right-hand side of (3.15) vanishes identically in Q_{R-1}), then yields

$$\int_{Q_{R-\rho}} |\nabla \phi_2(x)|^2 \, dx \lesssim (R-\rho)^d \sqrt{T} \exp\left(-c \frac{\rho}{\sqrt{T}} \right) \tag{3.20}$$

as well.

We are now in a position to conclude the proof of (3.13):

$$\begin{aligned}
&|\xi \cdot (A_{T,L,R} - A_{T,L}) \xi| \\
&\quad = \left| \fint_{Q_L} \big((\xi + \nabla \phi_{T,R}) \cdot A(\xi + \nabla \phi_{T,R}) - (\xi + \nabla \phi_T) \cdot A(\xi + \nabla \phi_T) \big) \mu_L(x) \, dx \right| \\
&\quad = \left| \fint_{Q_L} \big((\nabla \phi_{T,R} - \nabla \phi_T) \cdot A(\xi + \nabla \phi_{T,R}) - (\xi + \nabla \phi_T) \cdot A(\nabla \phi_T - \nabla \phi_{T,R}) \big) \mu_L(x) \, dx \right|
\end{aligned}$$

$$\lesssim \left(\fint_{Q_L} |\nabla\phi_{T,R} - \nabla\phi_T|^2 dx\right)^{1/2} \left(\left(\fint_{Q_L} |\xi + \nabla\phi_T|^2 dx\right)^{1/2} + \left(\fint_{Q_L} |\xi + \nabla\phi_{T,R}|^2 dx\right)^{1/2}\right)$$
$$= \left(\fint_{Q_L} |\nabla\phi_2|^2 dx\right)^{1/2} \left(\left(\fint_{Q_L} |\xi + \nabla\phi_T|^2 dx\right)^{1/2} + \left(\fint_{Q_L} |\xi + \nabla\phi_{T,R}|^2 dx\right)^{1/2}\right).$$

Combined with the *a priori* estimates

$$\|\nabla\phi_T\|_{H^1(Q)} \lesssim 1,$$
$$\|\nabla\phi_{T,R}\|^2_{H^1(Q_L)} \leq \|\nabla\phi_{T,R}\|^2_{H^1(Q_R)} \lesssim R^d,$$

and (3.20) with $R - \rho = L$, this proves the claim of Step 3, and concludes the proof of Theorem 3.1.

4. Numerical Study

In this section, we present numerical tests which show that Theorem 3.1 is sharp. Note however that Theorem 3.1 only gives an asymptotic rate of convergence, whereas in practice one is interested in values for the number $2R$ of periodic cells per dimension of the order $1 \leq 2R \leq 50$. In this case, it is not clear how the method behaves. We therefore present numerical tests in the two different regimes: $2R \gg 50$ and $2R \leq 50$. For the asymptotic regime, we need to reach large values of R. To this aim, we have preferred to treat the case of a discrete elliptic equation, for which the numerical simulations are exact (there is no further approximation in terms of finite element method, and the simulations are much cheaper in terms of computational cost). For the regime $2R \leq 50$, we have considered a continuous equation and numerically solved the problems by a finite element method, since this is the interesting case in practice. Two cases have been considered: Periodic and quasi-periodic coefficients.

4.1. *Asymptotic regime*

The discrete corrector equation is

$$-\nabla^* \cdot A(\xi + \nabla\phi) = 0 \quad \text{in } \mathbb{Z}^2, \tag{4.1}$$

where for all $u : \mathbb{Z}^2 \to \mathbb{R}$,

$$\nabla u(x) := \begin{bmatrix} u(x+\mathbf{e}_1) - u(x) \\ u(x+\mathbf{e}_2) - u(x) \end{bmatrix}, \quad \nabla^* u(x) := \begin{bmatrix} u(x) - u(x-\mathbf{e}_1) \\ u(x) - u(x-\mathbf{e}_2) \end{bmatrix}$$

and

$$A(x) := \operatorname{diag}[a(x, x+\mathbf{e}_1), a(x, x+\mathbf{e}_2)].$$

The matrix A is $[0,4)^2$-periodic, and sketched on a periodic cell on Fig. 1. In the example considered, $a(x, x+\mathbf{e}_1)$ and $a(x, x+\mathbf{e}_2)$ represent the conductivities 1 or

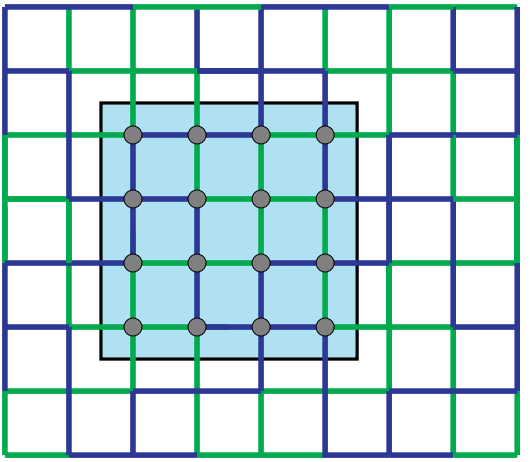

Fig. 1. (Color online) Periodic cell in the discrete case.

100 of the horizontal edge $[x, x + \mathbf{e}_1]$ and the vertical edge $[x, x + \mathbf{e}_2]$ respectively, according to the colors on Fig. 1. The homogenization theory for such discrete elliptic operators is similar to the continuous case (see for instance Ref. 22 in two dimensions, and Ref. 2 in the general case). By symmetry arguments, the homogenized matrix associated with A is a multiple of the identity. It can be evaluated numerically (note that we do not make any other error than the machine precision). Its numerical value is $A_{\text{hom}} = 26.240099009901\ldots$.

In order to illustrate the discrete counterpart to Theorem 3.1 (which is similar, both in terms of estimates and proof, cf. Refs. 12, 13 and 10 for related arguments), we have conducted the following series of tests. We have considered:

(a) Five values for the zeroth-order term: $T = \infty$ (no zeroth-order term), $T \sim R$, $T \sim R^{3/2}$, $T \sim R^{7/4}$, and $T \sim R^2(\ln R)^{-4}$;
(b) Two different filters: Orders $p = 0$ (no filter) and $p = \infty$;
(c) $L = R/3$.

The linear problem has been solved by a gradient conjugate method preconditioned by an incomplete Cholesky factorization. The predictions of Theorem 3.1 in terms of convergence rate of $A_{T,R,L}$ to A_{hom} in function of R are gathered and compared to the results of numerical tests in Table 1. More details are also given in Figs. 2–6, where

Table 1. Order of convergence: Predictions and numerical results.

	$T = \infty$		$T \sim R$		$T \sim R^{3/2}$		$T \sim R^{7/4}$		$T \sim R^2(\ln R)^{-8}$	
	Pred.	Test	Pred.	Test	Pred.	Test	Pred.	Test	Pred.	Test
$p = 0$	1	1	1	1	1	1	1	1	1	1
$p = \infty$	1	1	2	2	3	3.1	3.5	3.4	4^-	$\simeq 3$

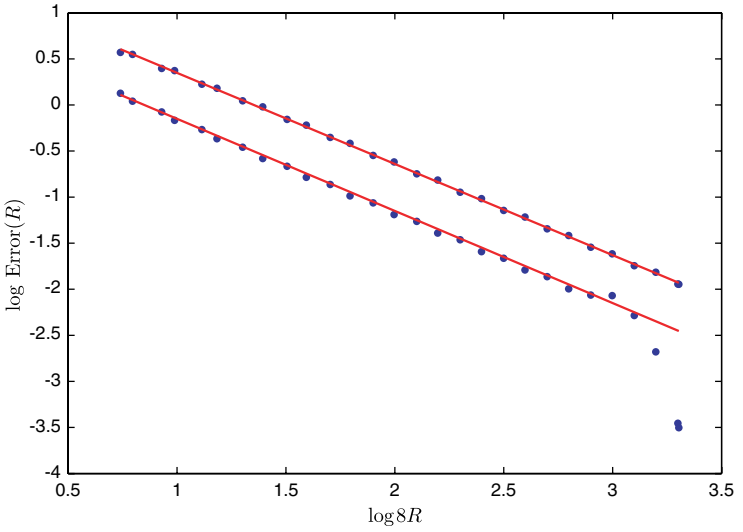

Fig. 2. Absolute error in log scale without zeroth-order term, no filter (slope −1), infinite order filter (slope −1, better prefactor).

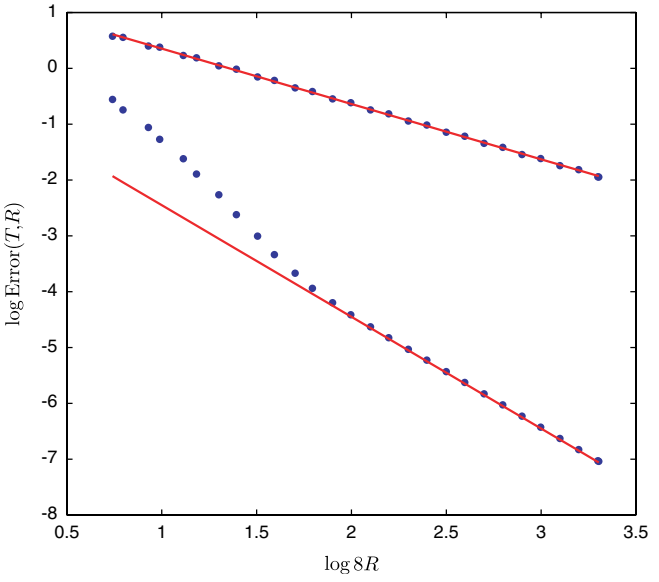

Fig. 3. Absolute error in log scale for $T = 2R/25$, no filter (slope −1), infinite order filter (slope −2).

the overall error

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

is plotted in log scale in function of R, according to (a), (b) and (c). Let us quickly comment on the values of T in Figs. 2–6. For the five dependences of T upon R in (a),

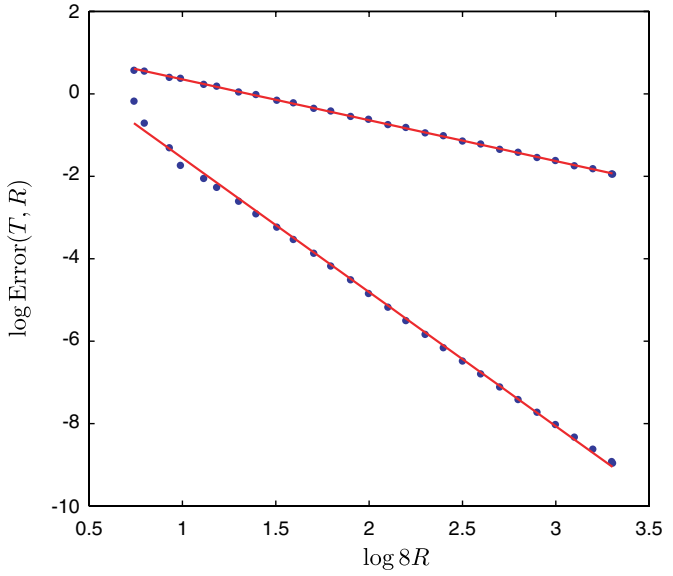

Fig. 4. Absolute error in log scale for $T = (8R)^{3/2}/1000$, no filter (slope -1), infinite order filter (slope -3.1).

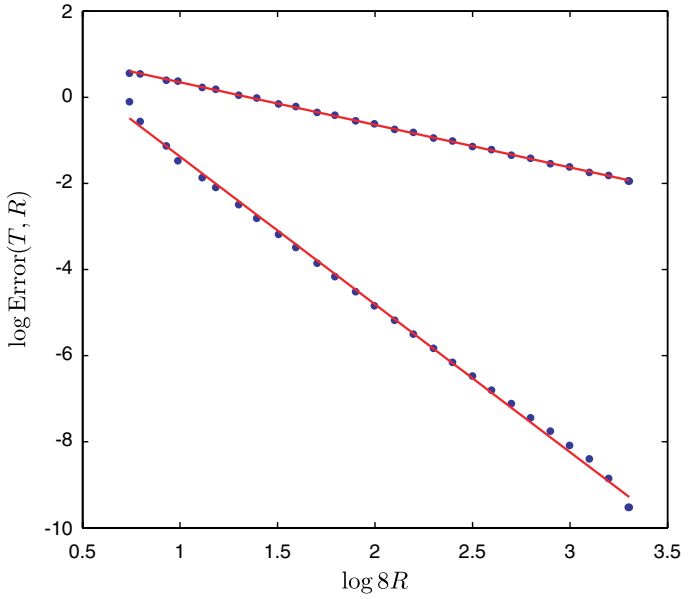

Fig. 5. Absolute error in log scale for $T = (8R)^{7/4}/5000$, no filter (slope -1), infinite order filter (slope -3.4).

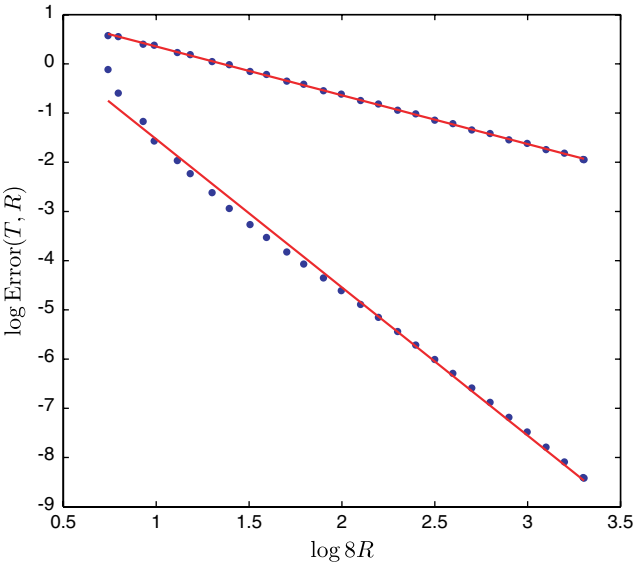

Fig. 6. Absolute error in log scale for $T = (8R)^2/(25\ln^4(8R))$, no filter (slope -1), infinite order filter (slope -3).

we have chosen the prefactors so that their values roughly coincide for $2R = 25$ (that is for 25 periodic cells per dimension):

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

The numerical results widely confirm the analysis, and perfectly illustrate the specific influences of the two parameters p and T. The convergence rate for $T \sim R^2(\ln R)^{-4}$ does not seem to meet the theoretical prediction. Indeed, the effect of the logarithm to the power 4 is still not negligible for $8R = 10^4$: The asymptotic regime is not yet captured by the tests.

For similar numerical tests in the stochastic case, we refer the reader to Ref. 10, where the model and the results are described in detail. Note that, there again, the role of the zeroth-order term T is very important *in practice*.

4.2. *Regime* $2R \leq 50$

As can be seen on the previous numerical tests, the asymptotic regime is not met for small values of R. For instance, in the case $T = 2R/25$, the apparent rate of convergence on Fig. 3 is closer to 4 than to 2 up to $2R \simeq 25$, whereas it is clearly 2 asymptotically. In this section, we focus on continuous differential operators in the regime $2R \leq 50$ (that is a number of periodic cells per dimension less than 50). Since

the multiplicative constants and coefficients in estimate (3.1) only depend on the dimension, and on the ellipticity and continuity constants α and β of A, there exists a choice of the parameter T in function of R which is efficient for a wide class of coefficients A.

In what follows, the partial differential equations are numerically solved using $P2$-finite elements (the diffusion coefficients are chosen regular) on a mesh with 20 points per periodic cell per dimension (that is 400 degrees of freedom per periodic cell). When the coefficients are periodic, the reference homogenized matrix A_{hom} is approximated using the same procedure (same meshsize, same finite elements) on one single cell with periodic boundary conditions.

We consider the following matrix A:

$$A(x) = \left(\frac{2 + 1.8\sin(2\pi x_1)}{2 + 1.8\cos(2\pi x_2)} + \frac{2 + \sin(2\pi x_2)}{2 + 1.8\cos(2\pi x_1)}\right)\mathrm{Id}, \tag{4.2}$$

used as benchmark tests in Refs. 18 and 3, and for which $\alpha \simeq 0.35$, $\beta \simeq 20.5$, and $A_{\text{hom}} \simeq 2.75\,\mathrm{Id}$. We take $L = R/3$, $T = R/5$ and a filter of order 2. The global error $|A_{T,R,L} - A_{\text{hom}}|$ and the error without zeroth-order term and without filtering are plotted on Figs. 7 and 8. Without zeroth-order term, the convergence rate is R^{-1} as expected, and the use of a filtering method reduces the prefactor but does not change the rate. With the zeroth-order term and the filtering method, the apparent convergence rate is R^{-3} (note that the asymptotic theoretical rate R^{-2} is not attained yet), which coincides with the convergence rate associated with filters of order 2

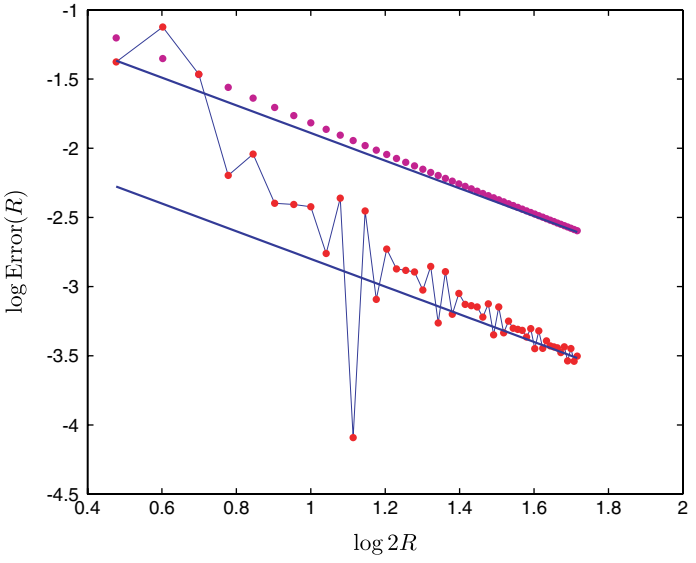

Fig. 7. Error in log−log for (4.2) in function of the number of cells per dimension $2R \in [3, 52]$ without zeroth-order term, with and without filtering: Slope -1 in both cases.

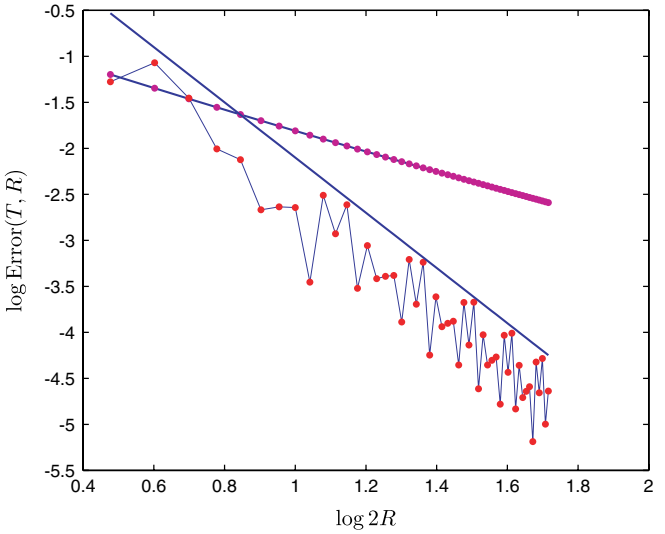

Fig. 8. Error in log–log for (4.2) in function of the number of cells per dimension $2R \in [3, 52]$ with a zeroth-order term $T = R/5$, with and without filtering: Slopes -1 and -3.

(cf. Lemma 3.1). This is in agreement with the tests in the discrete case, and confirms the analysis.

Let us now consider the following other matrix:

$$A(x) = (1 + 30(2 + \sin(2\pi x_1)\sin(2\pi x_2)))\,\mathrm{Id}, \tag{4.3}$$

for which $\alpha = 31$, $\beta = 91$ and $A_{\mathrm{hom}} \simeq 59.1\,\mathrm{Id}$. This example is much easier to deal with than the previous one (the homogenized coefficient is close to the arithmetic mean 61). Hence, Dirichlet boundary conditions are expected to perform well, even without zeroth-order term. This is confirmed by the numerical tests (see Fig. 9). Interestingly, the result seems to be better without filtering for $T = \infty$ in that case (such a behavior has also been reported by E and Yue in Ref. 23). Adding now the zeroth-order term with $T = R/100, 2R/300, 2R/500$, the results are better provided the use of a filtering method (order 2), as can be seen in Figs. 10–12. Note that the homogenized coefficient is 22 times larger than in the previous case, so that we could expect T to be 22 times smaller than in the case (4.2).

These two series of tests clearly show that the numerical method performs quite well in this continuous periodic setting, even with a limited number of periodic cells.

Besides the periodic and stochastic settings, another standard benchmark case is the (academic) quasi-periodic setting. The last series of tests is dedicated to this case, and we consider the following quasi-periodic coefficients used in Ref. 3:

$$A(x) = \begin{pmatrix} 4 + \cos(2\pi(x_1 + x_2)) \\ \quad + \cos(2\pi\sqrt{2}(x_1 + x_2)) & 0 \\ 0 & 6 + \sin^2(2\pi x_1) + \sin^2(2\pi\sqrt{2}x_1) \end{pmatrix}. \tag{4.4}$$

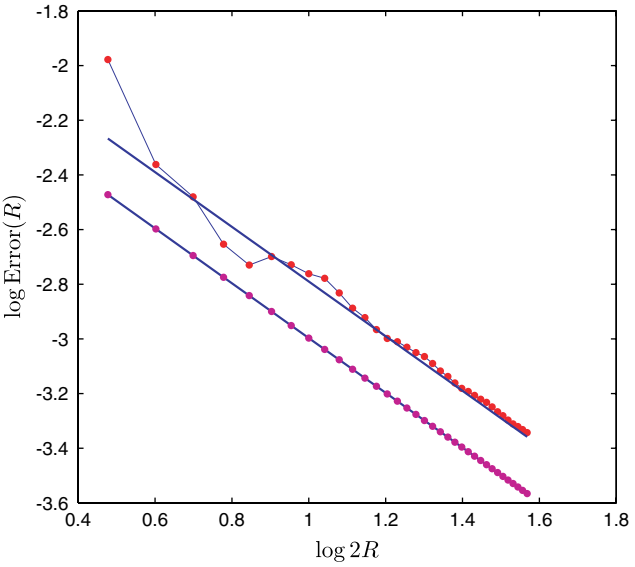

Fig. 9. Error in log−log for (4.3) in function of the number of cells per dimension $2R \in [3, 35]$ without zeroth-order term, with and without filtering: Slope -1 in both cases.

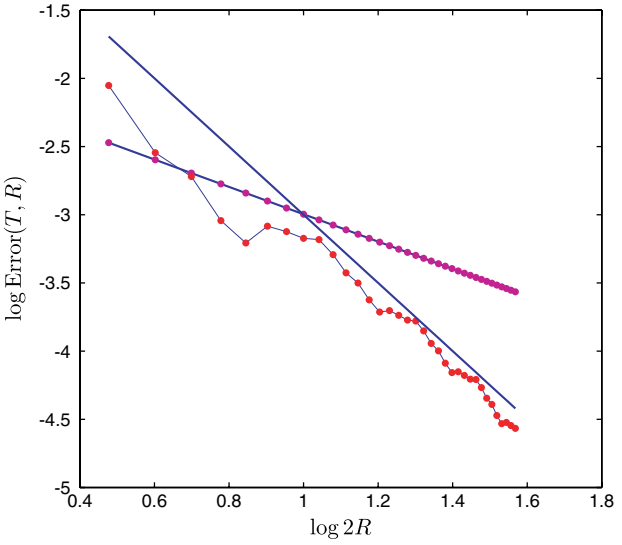

Fig. 10. Error in log−log for (4.3) in function of the number of cells per dimension $2R \in [3, 38]$ with a zeroth-order term $T = R/100$, with and without filtering: Slopes -1 and -2.5.

In this case, the homogenized coefficients are not easy to compute. They can only be extrapolated. We have taken for the approximation of the homogenized coefficients (that we call coefficient of reference) the output of the computation with $T = R/50$ and $2R = 52$. Although this may introduce a bias in favor of the proposed

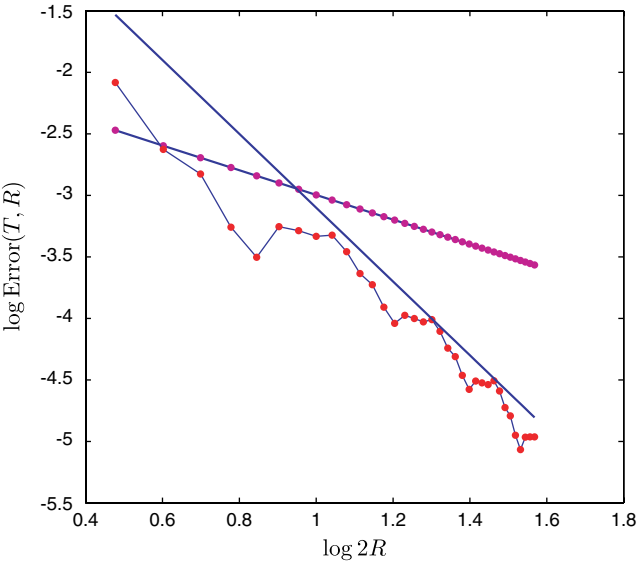

Fig. 11. Error in log−log for (4.3) in function of the number of cells per dimension $2R \in [3, 38]$ with a zeroth-order term $T = 2R/300$, with and without filtering: Slopes -1 and -3.

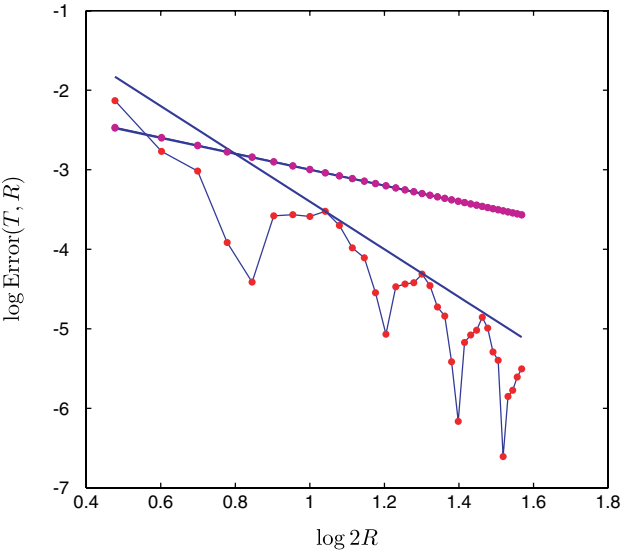

Fig. 12. Error in log−log for (4.3) in function of the number of cells per dimension $2R \in [3, 38]$ with a zeroth-order term $T = 2R/500$, with and without filtering: Slopes -1 and -3.

strategy, it can be checked *a posteriori*: The method without zeroth-order term and without filtering is expected to converge at a rate R^{-1}. This is effectively what we observe in Fig. 13 using this coefficient of reference. Instead, if we use as a reference the output of the computation for $2R = 52$ without zeroth-order term nor filtering,

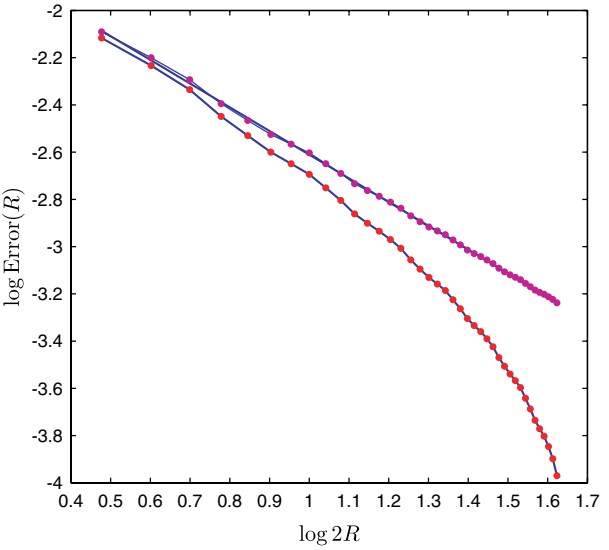

Fig. 13. Error in log−log for (4.4) in function of the number of cells per dimension $2R \in [3, 42]$ without zeroth-order term and without filtering, for the two different coefficients of reference: Slope -1 and artificial super-linear convergence.

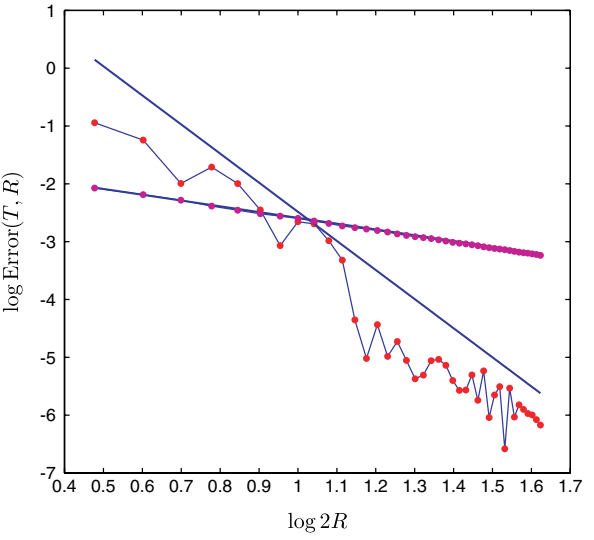

Fig. 14. Error in log−log for (4.4) in function of the number of cells per dimension $2R \in [3, 42]$ with a zeroth-order term $T = R/100$, with and without filtering: Slopes -1 and -5.

then we observe a super-linear convergence which is artificial (see Fig. 13). With the proposed method, as can be seen in Fig. 14, the rate of convergence seems to be much better (the slope of the straight line is -5). Hence the method also performs quite well for this quasi-periodic example.

To conclude, the numerical tests performed clearly demonstrate that the proposed strategy effectively reduces the resonance error, in the periodic, quasi-periodic and stochastic cases, both asymptotically and in the small number of periods regime. In addition, the analysis is sharp and both the roles of the zeroth-order term and of the filter are crucial.

Appendix A. Proof of Lemma 3.1

We first prove Lemma 3.1 for $d = 1$, and expand ϕ in Fourier series:

$$\phi(x) = \sum_{k\in\mathbb{Z}} c_k \exp(i2\pi kx). \tag{A.1}$$

Step 1. Proof of

$$\left|\int_{-L}^{L} \exp(i2\pi kx)\mu_L(x)\,dx\right| \lesssim |k|^{-p} L^{-(p+1)}, \tag{A.2}$$

for $k \neq 0$ and $p > 0$. After p integrations by parts, and using Definition 3.1(iii), we have

$$\begin{aligned}&\int_{-L}^{L} \exp(i2\pi kx)\mu_L(x)\,dx\\ &\quad= (-1)^p \int_{-L}^{L} \frac{1}{(2i\pi k)^p} \exp(i2\pi kx) L^{-p} L^{-1} \mu^{(p)}(L^{-1}x)\,dx.\end{aligned} \tag{A.3}$$

Extending μ by zero on $\mathbb{R}\backslash(-1,1)$, we may introduce

$$\begin{aligned}\overline{\mu}_L : (-1,1) &\to \mathbb{R}\\ x &\mapsto L^{-1} \sum_{k=-[L/2]-1}^{[L/2]+1} \mu^{(p)}(L^{-1}(x+2k)),\end{aligned} \tag{A.4}$$

where $[\cdot]$ denotes the integer part. Since $\mu^{(p)}$ is Lipschitz (with a constant C_μ), for all $x \in (-1,1)$ one has

$$|\overline{\mu}_L(x) - m_\mu| \leq C_\mu L^{-1}|x|, \tag{A.5}$$

where $m_\mu = \int_{-1}^{1} \overline{\mu}_L(x)\,dx$. Using $\overline{\mu}_L$ and the periodicity of $x \mapsto \exp(i2\pi kx)$, we rewrite (A.3) as

$$\begin{aligned}\left|\int_{-L}^{L} \exp(i2\pi kx)\mu_L(x)\,dx\right| &\overset{(A.4)}{=} \left|\frac{L^{-p}}{(-2i\pi k)^p} \int_{-1}^{1} \exp(i2\pi kx)\overline{\mu}_L(x)\,dx\right|\\ &\overset{(A.5)}{\leq} \frac{C_\mu L^{-(p+1)}}{(2\pi k)^p},\end{aligned}$$

which is (A.2).

Step 2. Proof of (3.12) for $d = 1$. For $p = 0$, (3.12) is trivial, and we only consider $p > 0$. We first note that (A.1) and Definition 3.1(ii) imply

$$\int_{-L}^{L} \phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi) = \sum_{k\in\mathbb{Z}\setminus\{0\}} c_k \int_{-L}^{L} \exp(i2\pi kx)\mu_L(x)\,dx.$$

Combined with (A.2), this becomes

$$\left|\int_{-L}^{L} \phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi)\right| \lesssim L^{-(p+1)} \sum_{k\in\mathbb{Z}\setminus\{0\}} c_k k^{-p}.$$

Using then Cauchy−Schwarz's inequality and Parseval's identity, one concludes

$$\begin{aligned}\left|\int_{-L}^{L} \phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi)\right| &\lesssim L^{-(p+1)} \left(\sum_{k\in\mathbb{Z}\setminus\{0\}} c_k^2\right)^{1/2} \left(\sum_{k\in\mathbb{Z}\setminus\{0\}} k^{-2p}\right)^{1/2} \\ &\lesssim L^{-(p+1)} \|\phi\|_{L^2(0,1)},\end{aligned}$$

since $p \geq 1$.

Step 3. Extension to dimension $d > 1$. We use a Fourier expansion as above:

$$\phi(x) = \sum_{k_1,\ldots,k_d\in\mathbb{Z}} c_{k_1,\ldots,k_d} \exp(i2\pi k \cdot x), \tag{A.6}$$

where $k = (k_1, \ldots, k_d)$ and $x = (x_1, \ldots, x_d)$. In this case, (A.2) is replaced by

$$\left|\int_{Q_L} \exp(i2\pi k \cdot x)\mu_L(x)\,dx\right| \lesssim L^{-(p+1)} \prod_{l=1}^{d} (1 + |k_l|)^{-p}$$

for all $k \neq 0$. As in Step 3, this leads to

$$\begin{aligned}\left|\int_{Q_L} \phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi)\right| &\lesssim L^{-(p+1)} \left(\sum_{k\in\mathbb{Z}^d} c_k^2\right) \left(\sum_{k\in\mathbb{Z}^d} \prod_{l=1}^{d} (1 + |k_l|)^{-2p}\right) \\ &\lesssim L^{-(p+1)} \|\phi\|_{L^2(Q)}.\end{aligned}$$

Step 4. The case $\phi \in L^q(Q)$, $1 < q < 2$. In this case, one cannot use Parseval's identity any longer. For $p \geq 2$, it is enough to use $|c_k| \leq \|\phi\|_{L^1(Q)}$ and the summability of $\prod_{l=1}^{d} (1 + |k_l|)^{-p}$ on $\mathbb{Z}^d$ to conclude. For $p = 1$, we appeal to Hardy−Littlewood's inequality (see for instance Ref. 7) for functions in $L^q(Q)$, $1 < q < 2$:

$$\left(\sum_{k\in\mathbb{Z}^d} \prod_{l=1}^{d} (1 + |k_l|)^{q-2} c_k^q\right)^{1/q} \lesssim \|\phi\|_{L^q(Q)},$$

and conclude by Hölder's inequality with exponents $(q/(q-1), q)$

$$\begin{aligned}
&\left|\int_{-L}^{L}\phi(x)\mu_L(x)\,dx - \mathcal{M}(\phi)\right| \\
&\quad\lesssim L^{-(p+1)}\sum_{k\in\mathbb{Z}^d}\prod_{l=1}^{d}(1+|k_l|)^{-1}c_k \\
&\quad= L^{-(p+1)}\sum_{k\in\mathbb{Z}^d}\prod_{l=1}^{d}(1+|k_l|)^{-2(q-1)/q}(1+|k_l|)^{(q-2)/q}c_k \\
&\quad\le L^{-(p+1)}\left(\sum_{k\in\mathbb{Z}^d}\prod_{l=1}^{d}(1+|k_l|)^{-2}\right)^{(q-1)/q}\left(\sum_{k\in\mathbb{Z}^d}\prod_{l=1}^{d}(1+|k_l|)^{q-2}c_k^q\right)^{1/q} \\
&\quad\lesssim L^{-(p+1)}\|\phi\|_{L^q(Q)}.
\end{aligned}$$

Appendix B. Proof of Lemma 3.2

The following proof is standard and relies on three arguments:

- Harnack's inequality,
- pointwise estimates for the Green function of second-order elliptic equations,
- the operator positivity method due to Agmon (see Ref. 1).

W.l.o.g. we assume $y = 0$, and use the shorthand notation $G_T(x)$ for $G_T(x, y)$.

Step 1. Operator positivity method. Let $b > 0$, and for all $j, k \in \mathbb{N}$ let $\chi_{T,j} : \mathbb{R}^d \to \mathbb{R}^+$ and $g_{T,j,k} : \mathbb{R}^d \to \mathbb{R}^+$ be given by

$$\chi_{T,j}(x) = \begin{cases} \text{for } |x| \le 2^j\sqrt{T} & : 0, \\ \text{for } 2^j\sqrt{T} \le |x| \le 2^{j+1}\sqrt{T} & : (2^j\sqrt{T})^{-1}(|x| - 2^j\sqrt{T}), \\ \text{for } |x| \ge 2^{j+1}\sqrt{T} & : 1 \end{cases}$$

and

$$g_{T,j,k}(x) = \begin{cases} \text{for } |x| \le 2^{j+1}\sqrt{T} & : 0, \\ \text{for } 2^{j+1}\sqrt{T} \le |x| \le 2^{j+1}\sqrt{T} + 2^k & : b(|x| - 2^{j+1}\sqrt{T}), \\ \text{for } 2^{j+1}\sqrt{T} + 2^k \le |x| \le 2^{j+1}\sqrt{T} + 2^{k+1} & : 2^{k+1}b + b(2^{j+1}\sqrt{T} - |x|), \\ \text{for } 2^{j+1}\sqrt{T} + 2^{k+1} \le |x| & : 0. \end{cases}$$

These functions are plotted for convenience in Fig. 15 for $d = 1$.

We multiply the defining equation for G_T by the test function

$$x \mapsto \chi_{T,j}(x)^2 \exp(2g_{T,j,k}(x))\, G_T(x)$$

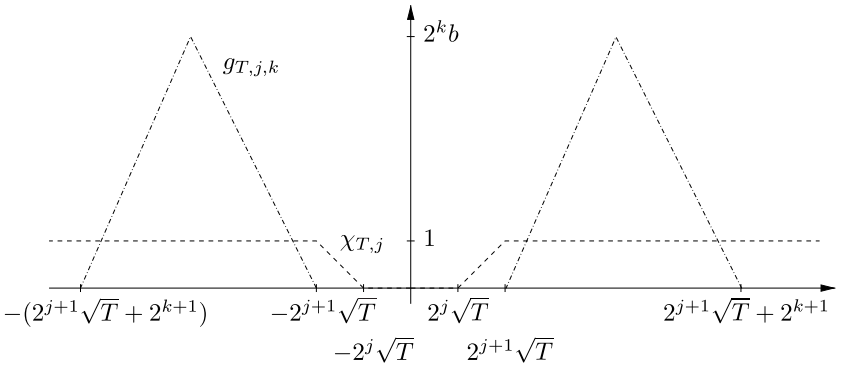

Fig. 15. Functions $g_{T,j,k}$ and $\chi_{T,j}$ for $d = 1$.

and integrate on $\mathbb{R}^d$, obtaining

$$T^{-1}\int_{\mathbb{R}^d}(\chi_{T,j}(x)\exp(g_{T,j,k}(x))G_T(x))^2dx + \int_{\mathbb{R}^d}\nabla(\chi_{T,j}(x)^2\exp(2g_{T,j,k}(x))G_T(x))\cdot A(x)\nabla G_T(x)dx = 0. \quad \text{(B.1)}$$

We focus on the second term of the equation and use Leibniz' rule. For the sake of clarity, we drop the subscripts and variables in the following calculation.

$$\begin{aligned}&\nabla(\chi^2\exp(2g)G)\cdot A\nabla G\\ &\quad=\nabla(\chi\exp(g)G)\cdot A\chi\exp(g)\nabla G+\chi\exp(g)G\nabla(\chi\exp(g))\cdot A\nabla G\\ &\quad=\nabla(\chi\exp(g)G)\cdot A\nabla(\chi\exp(g)G)\\ &\quad\underbrace{-\nabla(\chi\exp(g)G)\cdot A\nabla(\chi\exp(g))G}_{\clubsuit}+\chi\exp(g)G\nabla(\chi\exp(g))\cdot A\nabla G.\end{aligned}$$

We rewrite the second term of the right-hand side as follows:

$$\begin{aligned}\clubsuit&=-\chi\exp(g)\nabla G\cdot A\nabla(\chi\exp(g))G-G^2\nabla(\chi\exp(g))\cdot A\nabla(\chi\exp(g))\\ &=-G^2\nabla(\chi\exp(g))\cdot A\nabla(\chi\exp(g))-\chi\exp(g)G\nabla(\chi\exp(g))\cdot A\nabla G,\end{aligned}$$

by symmetry of A. The combination of these two identities yields

$$\begin{aligned}&\nabla(\chi^2\exp(2g)G)\cdot A\nabla G\\ &\quad=\nabla(\chi\exp(g)G)\cdot A\nabla(\chi\exp(g)G)-G^2\nabla(\chi\exp(g))\cdot A\nabla(\chi\exp(g))\\ &\quad\geq-\beta G^2|\nabla(\chi\exp(g))|^2\end{aligned}$$

by the uniform bound on A. Setting $\psi_{T,j,k}: x\mapsto\exp(g_{T,j,k}(x))G_T(x)$, we insert the latter inequality into (B.1) to get

$$\int_{\mathbb{R}^d}(T^{-1}\chi_{T,j}(x)^2\psi_{T,j,k}(x)^2-\beta G_T(x)^2|\nabla(\chi_{T,j}(x)\exp(g_{T,j,k}(x)))|^2)dx\leq 0.$$

Using the properties of $\chi_{T,j}$ and $g_{T,j,k}$, this becomes

$$\int_{|x|\geq 2^{j+1}\sqrt{T}} (T^{-1}\psi_{T,j,k}(x)^2 - \beta G_T(x)^2 |\nabla \exp(g_{T,j,k}(x))|^2)\,dx$$
$$\leq \int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} \beta |\nabla \chi_{T,j}(x)|^2 G_T(x)^2\,dx,$$

and finally

$$\int_{|x|\geq 2^{j+1}\sqrt{T}} (T^{-1} - \beta |\nabla g_{T,j,k}(x)|^2)\psi_{T,j,k}(x)^2\,dx$$
$$\leq \beta T^{-1} 2^{-2j} \int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} G_T(x)^2\,dx.$$

Choosing $b = (2\beta T)^{-1/2}$ then yields

$$\int_{|x|\geq 2^{j+1}\sqrt{T}} \psi_{T,j,k}(x)^2\,dx \lesssim 2^{-2j} \int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} G_T(x)^2\,dx.$$

We then pass to the limit $k \to \infty$ by the monotone convergence theorem to obtain

$$\int_{|x|\geq 2^{j+1}\sqrt{T}} \exp(2b(|x| - 2^{j+1}\sqrt{T}))\, G_T(x)^2\,dx \lesssim 2^{-2j} \int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} G_T(x)^2\,dx,$$

and therefore

$$\int_{|x|\geq 2^{j+2}\sqrt{T}} G_T(x)^2\,dx \lesssim 2^{-2j} \exp(-b2^{j+2}\sqrt{T}) \int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} G_T(x)^2\,dx. \tag{B.2}$$

Step 2. Decay estimates. We now appeal to standard decay estimates derived via the De Giorgi−Nash−Moser theory. We refer to Ref. 14 for $d > 2$. For $d = 2$, we refer to Ref. 12, whose proof is actually first presented in the continuous case (and later on adapted to the discrete case). In particular, for all $j \in \mathbb{N}$ and $d \geq 2$, we have

$$\int_{2^j\sqrt{T}\leq |x|<2^{j+1}\sqrt{T}} G_T(x)^2\,dx \lesssim (2^j\sqrt{T})^d ((2^j\sqrt{T})^{2-d})^2 = (2^j\sqrt{T})^{4-d}. \tag{B.3}$$

Note that the pointwise estimates in Ref. 14 for G_T are uniform in T for $d > 2$, so that (B.3) indeed holds on any annulus of the form $\{2^j R \leq |x| < 2^{j+1}R\}$, $R \geq 1$. This is not the case for $d = 2$ (the Green function of the Laplace operator diverges logarithmically).

Step 3. Harnack's inequality. We rewrite the defining equation for G_T as

$$-\nabla \cdot A\nabla G_T(x) = -T^{-1} G_T(x) \leq 0, \quad \text{for } |x| \geq 1,$$

so that we may use Harnack's inequality for non-negative subsolutions (see for instance Ref. 15):

$$\sup_{2^j\sqrt{T}\le|x|\le 2^{j+1}\sqrt{T}} G_T(x) \lesssim (2^j\sqrt{T})^{-d/2}\|G_T\|_{L^2(\{2^{j-1}\sqrt{T}\le|x|\le 2^{j+2}\sqrt{T}\})}. \tag{B.4}$$

The combination of (B.2)−(B.4) concludes the proof of the lemma.

Acknowledgment

The author gratefully acknowledges the help of Michel Hua for the numerical tests.

References

1. S. Agmon, *Lectures on Exponential Decay of Solutions of Second-Order Elliptic Equations*: *Bounds on Eigenfunctions of N-body Schrödinger Operators*, Mathematical Notes, Vol. 29 (Princeton Univ. Press, 1982).
2. R. Alicandro and M. Cicalese, A general integral representation result for the continuum limits of discrete energies with superlinear growth, *SIAM J. Math. Anal.* **36** (2004) 1−37.
3. X. Blanc and C. Le Bris, Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings, *Netw. Heterog. Media* **5** (2010) 1−29.
4. A. Bourgeat and A. Piatnitski, Approximations of effective coefficients in stochastic homogenization, *Ann. I. H. Poincaré* **40** (2005) 153−165.
5. W. E, B. Engquist, X. Li, W. Ren and E. Vanden-Eijnden, Heterogeneous multiscale methods: A review, *Commun. Comput. Phys.* **2** (2007) 367−450.
6. W. E, P. B. Ming and P. W. Zhang, Analysis of the heterogeneous multiscale method for elliptic homogenization problems, *J. Amer. Math. Soc.* **18** (2005) 121−156.
7. R. E. Edwards, *Fourier Series*, 2nd edn., Graduate Texts in Mathematics, Vol. 85 (Springer-Verlag, 1982).
8. A. Gloria, An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies, *Multiscale Model. Simul.* **5** (2006) 996−1043.
9. A. Gloria, An analytical framework for numerical homogenization — Part II: Windowing and oversampling, *Multiscale Model. Simul.* **7** (2008) 275−293.
10. A. Gloria, Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations, preprint, http://hal.archives-ouvertes.fr/inria-00510514/en/.
11. A. Gloria and J.-C. Mourrat, Spectral measure and approximation of homogenized coefficients, preprint, http://hal.archives-ouvertes.fr/inria-00510513/en/.
12. A. Gloria and F. Otto, An optimal variance estimate in stochastic homogenization of discrete elliptic equations, *Ann. Probab.* **39** (2011) 779−856.
13. A. Gloria and F. Otto, An optimal error estimate in stochastic homogenization of discrete elliptic equations, preprint, http://hal.archives-ouvertes.fr/inria-00457020/en/.
14. M. Grüter and K.-O. Widman, The Green function for uniformly elliptic equations, *Manuscripta Math.* **37** (1982) 303−342.
15. Q. Han and F. Lin, *Elliptic Partial Differential Equations* (Courant Institute of Mathematical Sciences, 1997).
16. T. Y. Hou and X. H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, *J. Comput. Phys.* **134** (1997) 169−189.
17. T. Y. Hou, X. H. Wu and Z. Q. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, *Math. Comput.* **68** (1999) 913−943.

18. T. Y. Hou, X. H. Wu and Y. Zhang, Removing the cell resonance error in the multiscale finite element method via a Petrov−Galerkin formulation, *Commun. Math. Sci.* **2** (2004) 185−205.
19. V. V. Jikov, S. M. Kozlov and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals* (Springer-Verlag, 1994).
20. N. Meyers, An L^p-estimate for the gradient of solutions of second order elliptic divergence equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)* **17** (1963) 189−206.
21. F. Murat and L. Tartar, H-convergence, in *Topics in the Mathematical Modelling of Composites Materials*, Progress in Nonlinear Differential Equations and Their Applications, Vol. 31, eds. A. V. Cherkaev and R. V. Kohn (Birkhäuser, 1997), pp. 21−44.
22. M. Vogelius, A homogenization result for planar, polygonal networks, *RAIRO Modél. Math. Anal. Numér.* **25** (1991) 483−514.
23. X. Yue and W. E, The local microscale problem in the multiscale modeling of strongly heterogeneous media: Effects of boundary conditions and cell size, *J. Comput. Phys.* **222** (2007) 556−572.