ESAIM: M2AN 46 (2012) 1–38
DOI: 10.1051/m2an/2011018

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

NUMERICAL APPROXIMATION OF EFFECTIVE COEFFICIENTS IN STOCHASTIC HOMOGENIZATION OF DISCRETE ELLIPTIC EQUATIONS

Antoine Gloria[1]

Abstract. We introduce and analyze a numerical strategy to approximate effective coefficients in stochastic homogenization of discrete elliptic equations. In particular, we consider the simplest case possible: An elliptic equation on the d-dimensional lattice $\mathbb{Z}^d$ with independent and identically distributed conductivities on the associated edges. Recent results by Otto and the author quantify the error made by approximating the homogenized coefficient by the averaged energy of a regularized corrector (with parameter T) on some box of finite size L. In this article, we replace the regularized corrector (which is the solution of a problem posed on $\mathbb{Z}^d$) by some practically computable proxy on some box of size $R \geq L$, and quantify the associated additional error. In order to improve the convergence, one may also consider N independent realizations of the computable proxy, and take the empirical average of the associated approximate homogenized coefficients. A natural optimization problem consists in properly choosing T, R, L and N in order to reduce the error at given computational complexity. Our analysis is sharp and sheds some light on this question. In particular, we propose and analyze a numerical algorithm to approximate the homogenized coefficients, taking advantage of the (nearly) optimal scalings of the errors we derive. The efficiency of the approach is illustrated by a numerical study in dimension 2.

Mathematics Subject Classification. 35B27, 39A70, 60H25, 65N99.

Received August 10, 2010.
Published online July 22, 2011

1. Introduction

In this article, we continue the analysis begun with Otto in [9,10] on stochastic homogenization of discrete elliptic equations. More precisely, we consider real functions u of the sites x in a d-dimensional Cartesian lattice $\mathbb{Z}^d$. Every edge e of the lattice is endowed with a "conductivity" $a(e) > 0$. This defines a discrete elliptic differential operator $-\nabla^* \cdot A\nabla$ *via*

$$-\nabla^* \cdot (A\nabla u)(x) := \sum_y a(e)(u(x) - u(y)),$$

where the sum is over the $2d$ sites y which are connected by an edge $e = [x, y] = [y, x]$ to the site x (the precise definitions of the discrete gradient and divergence are given in Sect. 2). We assume the conductivities a to be

Keywords and phrases. Stochastic homogenization, effective coefficients, difference operator, numerical method.

[1] Project-Team SIMPAF, INRIA Lille-Nord Europe, France and Laboratoire Paul Painlevé (UMR CNRS 8524), Université Lille 1, 59655 Villeneuve d'Ascq, France. antoine.gloria@inria.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2011

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

uniformly elliptic in the sense of

$$\alpha \leq a(e) \leq \beta \quad \text{for all edges } e$$

for some fixed constants $0 < \alpha \leq \beta < \infty$.

We are interested in random coefficients. To fix ideas, we consider the simplest situation possible:

$$\{a(e)\}_e \quad \text{are independent and identically distributed (i. i. d.).}$$

Hence the statistics are described by a distribution on the finite interval $[\alpha, \beta]$.

Classical results in stochastic homogenization of linear elliptic equations (see [15,21] for the continuous case, and [16,17] for the discrete case) state that there exist *homogeneous and deterministic* coefficients $A_{\rm hom}$ such that the solution operator of the continuous differential operator $-\nabla \cdot A_{\rm hom} \nabla$ describes the large scale behavior of the solution operator of the discrete differential operator $-\nabla^* \cdot A \nabla$. As a by product of this homogenization result, one obtains a characterization of the homogenized coefficients $A_{\rm hom}$: It is shown that for every direction $\xi \in \mathbb{R}^d$, there exists a unique scalar field ϕ such that $\nabla \phi$ is stationary (stationarity implies that the fields $\nabla \phi(\cdot)$ and $\nabla \phi(\cdot + z)$ have the same statistics for all shifts $z \in \mathbb{Z}^d$) and $\langle \nabla \phi \rangle = 0$, solving the equation

$$-\nabla^* \cdot (A(\xi + \nabla \phi)) = 0 \quad \text{in } \mathbb{Z}^d, \tag{1.1}$$

and normalized by $\phi(0) = 0$. As in periodic homogenization, the function $\mathbb{Z}^d \ni x \mapsto \xi \cdot x + \phi(x)$ can be seen as the A-harmonic function which macroscopically behaves as the affine function $\mathbb{Z}^d \ni x \mapsto \xi \cdot x$. With this "corrector" ϕ, the homogenized coefficients $A_{\rm hom}$ (which in general form a symmetric matrix and for our simple statistics in fact a multiple of the identity: $A_{\rm hom} = a_{\rm hom}{\rm Id}$) can be characterized as follows:

$$\xi \cdot A_{\rm hom} \xi = \langle (\xi + \nabla \phi) \cdot A(\xi + \nabla \phi) \rangle. \tag{1.2}$$

Since the scalar field $(\xi + \nabla \phi) \cdot A(\xi + \nabla \phi)$ is stationary, it does not matter (in terms of the distribution) at which site x it is evaluated in the formula (1.2), so that we suppress the argument x in our notation.

When one is interested in explicit values for $A_{\rm hom}$, one has to solve (1.1) and compute (1.2). Since this is not possible in practice, one has to make approximations. For a discussion of the literature on error estimates, in particular the pertinent work by Yurinskii [23] and Naddaf and Spencer [19], we refer to [9], Section 1.2. As recalled in [10], a standard approach used in practice consists in solving (1.1) in a box $Q_L = [-L, L)^d$ with periodic boundary conditions

$$-\nabla^* \cdot (A(\xi + \nabla \phi_{L,\#})) = 0 \quad \text{in } Q_L, \tag{1.3}$$

and replacing (1.2) by a space average

$$\xi \cdot A_{L,\#} \xi = \fint_{Q_L} (\xi + \nabla \phi_{L,\#}) \cdot A(\xi + \nabla \phi_{L,\#}). \tag{1.4}$$

Such an approach is consistent in the sense that

$$\lim_{L \to \infty} A_{L,\#} = A_{\rm hom}$$

almost surely, as proved in [20] for both the continuous and discrete cases (see also [3,4]). Numerical experiments tend to show that the use of periodic boundary conditions gives better results than other choices such as homogeneous Dirichlet boundary conditions, see [14,22]. As argued in [10], the error analysis for $\big\langle |A_{L,\#} - A_{\rm hom}|^2 \big\rangle^{1/2}$ is however not obvious *a priori* since $\nabla \phi$ and $\nabla \phi_{L,\#}$ are not jointly stationary. In [10], we have followed a somewhat different route by considering the standard regularization of (1.1) to prove existence of correctors. In particular, we have introduced a zero-order term in (1.1) and considered the unique stationary solution to

$$T^{-1} \phi_T - \nabla^* \cdot A(\xi + \nabla \phi_T) = 0 \qquad \text{in } \mathbb{Z}^d. \tag{1.5}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

The advantage of (1.5) for the analysis is that $\nabla\phi$ and $\nabla\phi_T$ are jointly stationary and solve an equation of the same type as (1.1) and (1.5):

$$-\nabla^* \cdot A(\nabla\phi - \nabla\phi_T) = T^{-1}\phi_T \qquad \text{in } \mathbb{Z}^d.$$

This has been of great help to estimate $|A_T - A_{\rm hom}|$ *via* $\langle|\nabla\phi_T - \nabla\phi|^2\rangle^{1/2}$ in [10], Theorem 1, where

$$\xi \cdot A_T\xi := \langle(\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\rangle. \tag{1.6}$$

Yet, the defining equation (1.5) for ϕ_T is still posed on the whole space $\mathbb{Z}^d$, which is a handicap for the numerical practice.

Turning back to the idea leading to (1.3), one may approximate the regularized corrector ϕ_T by solving (1.5) in a box $Q_R = [-R, R)^d$, $R \geq L$, with periodic boundary conditions

$$T^{-1}\phi_{T,R,\#} - \nabla^* \cdot A(\xi + \nabla\phi_{T,R,\#}) = 0 \qquad \text{in } Q_R, \tag{1.7}$$

and replace (1.6) by

$$\xi \cdot A_{T,R,L,\#}\xi := \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,R,\#}) \cdot A(\xi + \nabla\phi_{T,R,\#})\mu_L,$$

where μ_L is a suitable mask with support in Q_L (see Thm. 2.10). As opposed to the case without the zero-order term, estimating $|\nabla\phi_T - \nabla\phi_{T,R,\#}|$ in a box Q_L is made easy if $R - L \gg \sqrt{T}$ due to the exponential decay of the Green function associated with $T^{-1} - \nabla^* \cdot A\nabla$ (see Lem. 3.2). Therefore, $\langle|A_{T,R,L,\#} - A_{T,L}|^2\rangle^{1/2}$ is expected to be of infinite order in terms of $\frac{R-L}{\sqrt{T}}$, where

$$\xi \cdot A_{T,L}\xi := \int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\mu_L \tag{1.8}$$

(note that this definition slightly differs from the corresponding definition in [9], Thm. 2.1, since we do not consider the contribution of the zero-order term here). One crucial feature of the zero-order term is to make the dependence of $\nabla\phi_{T,L,\#}$ upon the boundary value be exponentially small in terms of the distance to the boundary measured in units of $\sqrt{T}$. Hence, although the zero-order term in (1.5) has been introduced for the "convenience" of the analysis, it turns out that such a term is also very pertinent from the numerical point of view, as further illustrated at the end of this article in our discrete stochastic case. Even in the much more studied continuous periodic case (for which the addition of a zero-order term is not needed for the analysis), such a term yields a striking improvement of the order of convergence for the approximation of the homogenized coefficients (see [8]). Let us now give the argument to conclude the numerical analysis. To pass from an estimate on $\langle|A_{T,R,L,\#} - A_{T,L}|^2\rangle^{1/2}$ to an estimate on $\langle|A_{T,R,L,\#} - A_{\rm hom}|^2\rangle^{1/2}$, we use the triangle inequality in the form

$$\langle|A_{T,R,L,\#} - A_{\rm hom}|^2\rangle^{1/2} \leq \langle|A_{T,R,L,\#} - A_{T,L}|^2\rangle^{1/2} + \langle|A_{T,L} - A_T|^2\rangle^{1/2} + \langle|A_T - A_{\rm hom}|^2\rangle^{1/2},$$

and then appeal to [9], Theorem 2.1 and Remark 2.1, to deal with the second term of the r. h. s. (which is the variance of $A_{T,L}$), and to [10], Theorem 1, for the last term (which is the systematic error due to the zero-order perturbation). Note that the natures of the three terms are different: the first and last terms are "deterministic errors" (or at least estimated by deterministic arguments) whereas the second term measures fluctuations. In particular other norms could be considered than the second moment, and one may wish to obtain large deviation estimates instead of a variance estimate, in the spirit of the work by Caputo and Ioffe in [4]. To do so, only the second term has to be further analyzed.

Since the zero-order term reduces the dependence of the solution upon the boundary conditions far from the boundary, the precise nature of the boundary conditions is somewhat irrelevant (in contrast to the numerical evidence in [22] without the zero-order term). Hence, one may safely replace the periodic boundary conditions of (1.7) by homogeneous Dirichlet boundary conditions, *i.e.* without changing the order of convergence of the

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

method. For the numerical practice the use of Dirichlet boundary conditions is an advantage (sparsity of the matrix, efficient preconditioner, and so on). In this article we will therefore focus on the following proxy for ϕ_T in the box Q_R: The unique solution $\phi_{T,R}$ to

$$\left\{ \begin{array}{rcll} T^{-1}\phi_{T,R} - \nabla^* \cdot A(\xi + \nabla\phi_{T,R}) & = & 0 & \text{in } Q_R, \\ \phi_{T,R} & = & 0 & \text{in } \mathbb{Z}^d \setminus Q_R, \end{array} \right.$$

and define an approximation $A_{T,R,L}$ of $A_{\rm hom}$ by

$$\xi \cdot A_{T,R,L}\xi \, = \, \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,R}) \cdot A(\xi + \nabla\phi_{T,R})\mu_L. \tag{1.9}$$

As we shall prove in Theorem 2.10, there exists $c > 0$ (depending only on d and the ellipticity constants α, β) such that for all $R \sim L \sim R - L \gtrsim \sqrt{T}$,

$$\left\langle |A_{T,R,L} - A_{T,L}|^2 \right\rangle^{1/2} \lesssim T^{3/4} \exp\left(-c \frac{R-L}{\sqrt{T}} \right). \tag{1.10}$$

In combination with [9], Theorem 2.1 and Remark 2.1 and [10], Theorem 1 (see the argument hereafter), this yields the following complete error estimate for the choice $R = 2L$ and $T = L$

$$\left\langle |A_{L,2L,L} - A_{\rm hom}|^2 \right\rangle^{1/2} \lesssim \left\{ \begin{array}{lcl} d = 2 & : & L^{-1} \ln^q L, \\ d = 3 & : & L^{-3/2}, \\ d = 4 & : & L^{-2} \ln L, \\ d > 4 & : & L^{-2}, \end{array} \right. \tag{1.11}$$

for some q depending only on α, β, where "$\lesssim$" stands for "$\leq$" up to a multiplicative constant depending only on α, β, and d. This estimate relies on the variance estimate and the estimate of the systematic error, which are both optimal in the sense that they coincide with the explicit rates obtained in the regime of vanishing ellipticity ratio $1 - \frac{\alpha}{\beta} \ll 1$ (see [9], Appendix). The error due to the boundary conditions is of higher order. Hence (1.11) is optimal (except for the exponent on the logarithmic correction for $d = 2$). This result is the first optimal estimate of the convergence rate in stochastic homogenization (of discrete elliptic equations) for $d > 1$ (estimates for $d > 2$ were obtained in [3,7] using Yurinskii's results in [23], they are however suboptimal in the case of stochastic coefficients with finite correlation-length, see [9], introduction). For the extension of this method to continuous elliptic equations, we refer the reader to the end of this introduction.

In the applied mechanics community, the periodization approach is usually combined with an empirical average: N independent realizations $\{A_{L,\#,k}\}_{k\in\{1,\ldots,N\}}$ of $A_{L,\#}$ are computed, and $A_{\rm hom}$ is approximated by the empirical average

$$A^N_{L,\#} \, = \, \frac{1}{N} \sum_{k=1}^N A_{L,\#,k}.$$

Some numerical experiments on such a method with partial conclusions are reported on in [14]. Proceeding the same way, we may consider N independent realizations of $\{A_k\}_{k\in\{1,\ldots,N\}}$ of $A_{|Q_R}$ and approximate $A_{\rm hom}$ by

$$A^N_{T,R,L} \, = \, \frac{1}{N} \sum_{k=1}^N A_{T,R,L,k},$$

where $A_{T,R,L,k}$ is the approximation (1.9) for the realization A_k of A, $k \in \{1, \ldots, N\}$.

An important question for practical purposes is to quantify the dependence of the error $\left\langle |A^N_{T,R,L} - A_{\rm hom}|^2 \right\rangle^{1/2}$ in terms of T, R, L and N. Relying only on the results of [9,10], we can already give some pieces of answer to

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

this question. In particular, in view of (1.10), one needs $R-L \gg \sqrt{T}$, which we replace at first order for this discussion by $R=L$ and $T \leq L^2$. The following coarse complexity analysis gives a hint on the relative cost of the method in terms of L and N. Yet, as we shall discuss in the core of this paper, the careful analysis of the effect of boundary conditions will significantly modify this picture (see Sect. 4). In the rest of this introduction, we focus on the error we make by approximating $A_{\rm hom}$ by the matrix $A^N_{T,L}$ defined in (1.8). As shown in [9, Introduction] when $N=1$ (the argument does not depend on N), the error is made of two contributions, a "systematic error" and a "random error":

$$\begin{aligned}\left\langle |A^N_{T,L}-A_{\rm hom}|^2 \right\rangle &= \underbrace{\left\langle (\nabla\phi_T-\nabla\phi)\cdot A(\nabla\phi_T-\nabla\phi)\right\rangle^2}_{=:\ \mathrm{Error}_{\rm sys}(T)^2} \\ &\quad + \underbrace{\left\langle \left(\left\langle (\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\right\rangle - \frac{1}{N}\sum_{k=1}^N \int_{\mathbb{Z}^d} (\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\,\mathrm{d}x \right)^2 \right\rangle}_{=:\ \mathrm{Error}_{\rm rand}(T,L,N)^2}. \end{aligned} \tag{1.12}$$

The systematic error has been estimated in [10], Theorem 1

$$\mathrm{Error}_{\rm sys}(T) \lesssim \left\{ \begin{array}{lcl} d=2 & : & T^{-1}\ln^q T, \\ d=3 & : & T^{-3/2}, \\ d=4 & : & T^{-2}\ln T, \\ d>4 & : & T^{-2}. \end{array} \right. \tag{1.13}$$

It vanishes when $T\uparrow\infty$. Using the independence of the $\phi_{T,k}$, we may rewrite the random error as

$$\begin{aligned}\mathrm{Error}_{\rm rand}(T,L,N)^2 &= \left\langle \left(\frac{1}{N}\sum_{k=1}^N \left(\left\langle \int_{\mathbb{Z}^d} (\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\mu_L\,\mathrm{d}x \right\rangle \right.\right.\right. \\ &\qquad \left.\left.\left. - \int_{\mathbb{Z}^d} (\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\mu_L\,\mathrm{d}x \right) \right)^2 \right\rangle \\ &= \frac{1}{N}\mathrm{var}\left[\int_{\mathbb{Z}^d} (\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\mu_L\,\mathrm{d}x \right]. \end{aligned}$$

It measures the fluctuations of the energy density. This error vanishes as $L\uparrow\infty$, but also when the number of realizations $N\uparrow\infty$. In [9], Theorem 2.1 and Remark 2.1, we have proved that

$$\mathrm{var}\left[\int_{\mathbb{Z}^d} (\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\mu_L\,\mathrm{d}x \right]^{1/2} \lesssim \left\{ \begin{array}{lcl} d=2 & : & (L^{-1}+T^{-1})\ln^q T, \\ d>2 & : & L^{-d/2}(1+T^{-1}L). \end{array} \right. \tag{1.14}$$

Hence, if we further assume that $T\geq L$, this yields

$$\mathrm{Error}_{\rm rand}(T,L,N) \lesssim \left\{ \begin{array}{lcl} d=2 & : & (N^{1/2}L)^{-1}\ln^q T, \\ d>2 & : & (N^{1/d}L)^{-d/2} \end{array} \right.$$

where q only depends on the coercivity constants. The estimate of the random error singles out a quantity which plays an important role: the product $N(2L)^d$, which we will denote by M, and call the *effective number of sites* for the triplet (T,L,N) (this is the number of sites at which the energy density of the proxy for the regularized corrector is considered in the definition of $A^N_{T,L}$). In particular, since the error estimates are optimal (at least in the regime of vanishing ellipticity ratio), due to (1.12), the global error $\left\langle |A^N_{T,L}-A_{\rm hom}|^2\right\rangle$ scales at least as

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

M^{-1}. This error is intrinsic – note that this scaling coincides with the central limit theorem scaling associated with M independent realizations of one single random variable. The optimization problem we shall address is the following: *find a triplet (T, L, N) which yields the best error possible M^{-1} at the lowest computational cost.* This optimization problem is completed by the following constraints:

- fixed effective number of sites: $N(2L)^d = M$;
- effect of boundary conditions: $T \leq L^2$;
- optimal form of the variance estimate: $T \geq L$.

We focus on dimensions $d = 2, 3, 4$. The combination of the estimates of the random error and systematic error yields the following global error estimate completed by the bounds on T:

$$\left\langle |A^N_{T,L} - A_{\rm hom}|^2 \right\rangle \lesssim \left\{ \begin{array}{lcll} d=2 & : & (M^{-1}+T^{-2})\ln^q T, & (M/N)^{1/2} \leq T \leq (M/N), \\ d=3 & : & M^{-1}+T^{-3}, & (M/N)^{1/3} \leq T \leq (M/N)^{2/3}, \\ d=4 & : & M^{-1}+T^{-4}\ln^2 T, & (M/N)^{1/4} \leq T \leq (M/N)^{1/2}. \end{array} \right.$$

In order for the systematic error to be of higher order than the random error, one needs:

$$T^{-d} \lesssim M^{-1},$$

which is only possible if $N \leq \sqrt{M}$ for $d = 2, 3, 4$ in view of the bounds on T. Hence, among the triplets (T, L, N) with $N(2L)^d = M$, only those with $N \leq \sqrt{M}$ may yield the optimal scaling M^{-1} (with the logarithmic correction in dimensions $d = 2$ and $d = 4$). In addition, to minimize further the error, T should be chosen as large as possible, that is $T = (M/N)^{2/d}$. Since the cost of solving a linear system is a convex function (superlinear) of the number of unknows, it is more favourable to solve $\sqrt{M}$ systems of $\sqrt{M}$ unknowns than N systems of $N^{-1}M$ unknowns for all $N \in \{1, \ldots, \sqrt{M}\}$. *In particular, for $d \leq 4$ it seems best to evenly split a given number M of effective sites into the number N of realizations and the number L^d of sites per realization, i.e. $N = (2L)^d = \sqrt{M}$.*

Within the first order version $T \leq L^2$ of the fact that the regularized corrector equation has to be solved on a finite box, the discussion above gives a clear answer to the optimization problem: The larger N, the better, provided N remains bounded by $\sqrt{M}$. Yet, in practice, the effect of solving the regularized corrector equation on a large box Q_R cannot be reduced to the inequality $T \leq L^2$, and the parameter R has to be considered explicitly in the optimization process. The presence of the "buffer region" $Q_R \setminus Q_L$ makes the effective number of sites $N(2L)^d$ different from the total number of unknowns $N(2R)^d$ of the problem, so that the above discussion has to be refined. In particular, the difference $R-L$ should be large with respect to $\sqrt{T}$. As a consequence, at fixed effective number of sites M, the larger N, the smaller L, and therefore the larger the ratio $(R-L)/L \gg \sqrt{T}/L$. Hence they are two competing phenomena: the total number of unknowns increases with N (the ratio of the buffer regions increases with the number N of boxes Q_R) whereas at fixed number of unknowns the cost of solving linear systems decreases with N. One aim of this article is to make use of the sharp analysis of the error of Theorem 2.10 to further study this nontrivial interplay.

The article is organized as follows: in Section 2, we introduce the general framework and state the main result of this paper, *i.e.* an optimal estimate of $\left\langle |A^N_{T,R,L} - A_{\rm hom}|^2 \right\rangle^{1/2}$ (with respect to the case of vanishing ellipticity ratio), whose proof is the object of Section 3. In Section 4, we take advantage of this error analysis to address the optimization of the number N of subproblems given a fixed effective number of sites M. This allows us in particular to illustrate the sharpness of our result on a two-dimensional example.

To conclude this introduction, let us mention that we'd like to see the discrete stochastic elliptic operator under investigation here as a good model problem for continuous elliptic operators with random coefficients of correlation length unity. As will be clear in Section 3, the results of this paper rely on two types of results: The estimates of [9,10] (which heavily use the discreteness of the conductivity function on $\mathbb{Z}^d$) and deterministic estimates on elliptic equations. The deterministic estimates derived in this paper do not exploit the specific structure of the random coefficients (that is, i. i. d.) and Proposition 2.8 actually holds for any coefficients

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

A (satisfying the ellipticity conditions). In addition, the proof of Proposition 2.8 only uses one feature of the discreteness: The fact that the gradient $\nabla u(x)$ of a discrete function $u : \mathbb{Z}^d \to \mathbb{R}$ is controlled by $\sum_{i=1}^d (|u(x)| + |u(x+\mathbf{e}_i)|)$. This convenient estimate is not essential for our argument and may be replaced in the continuous case by Cacciopoli's inequality in the form of

$$\int_{B_1(x)} |\nabla u(y)|^2 \mathrm{d}y \lesssim \int_{B_2(x)} u(y)^2 \mathrm{d}y$$

for A-harmonic functions. In particular, Proposition 2.8 holds as well in the continuous case (see [8] for similar results). Hence, provided one extends the results of [9,10] to the continuous case – see in particular [11] –, the results of the present paper (that is essentially Thm. 2.10) will hold as well.

Throughout the paper, we make use of the following notation:

- $d \geq 2$ is the dimension;
- $\int_{\mathbb{Z}^d} \mathrm{d}x$ denotes the sum over $x \in \mathbb{Z}^d$, and $\int_D \mathrm{d}x$ denotes the sum over $x \in \mathbb{Z}^d$ such that $x \in D$, D of $\mathbb{R}^d$;
- $\langle \cdot \rangle$ is the ensemble average, or equivalently the expectation in the underlying probability space;
- $\mathrm{var}\,[\cdot]$ is the variance associated with the ensemble average;
- $\lesssim$ 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 Def. 2.1 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;
- $(\mathbf{e}_1, \ldots, \mathbf{e}_d)$ denotes the canonical basis of $\mathbb{Z}^d$.

2. Main result

2.1. General framework

Definition 2.1. We say that a is a conductivity function if there exist $0 < \alpha \leq \beta < \infty$ such that for every edge e of the square lattice generated by $\mathbb{Z}^d$, one has $a(e) \in [\alpha, \beta]$. We denote by $\mathcal{A}_{\alpha\beta}$ the set of such conductivity functions.

Definition 2.2. The elliptic operator L associated with a conductivity function $a \in \mathcal{A}_{\alpha\beta}$ is defined for all $u : \mathbb{Z}^d \to \mathbb{R}$ and $x \in \mathbb{Z}^d$ by

$$(Lu)(x) = -\nabla^* \cdot A(x) \nabla u(x) \tag{2.1}$$

where

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

the divergence of some vector V is given by the "formal" scalar product between ∇^* and V, that is $\nabla^* \cdot V(x) = \sum_{i=1}^d \big(V_i(x+\mathbf{e}_i) - V_i(x)\big)$, and

$$A(x) := \mathrm{diag}\,[a(e_1), \ldots, a(e_d)]\,,$$

$e_1 = [x, x+\mathbf{e}_1], \ldots, e_d = [x, x+\mathbf{e}_d]$.

We now turn to the definition of the statistics of the conductivity function.

Definition 2.3. A conductivity function is said to be independent and identically distributed (i. i. d.) if the coefficients $a(e)$ are i. i. d. random variables.

Definition 2.4. The conductivity matrix A is obviously stationary in the sense that for all $k \in \mathbb{N}$, all $x_1, \ldots, x_k \in \mathbb{Z}^d$, and $z \in \mathbb{Z}^d$, the random "vectors" $(A(x_1+z), \ldots, A(x_k+z))$ and $(A(x_1), \ldots, A(x_k))$ have the same statistics. Therefore, any translation invariant function of A, such as the regularized corrector ϕ_T (see Lem. 2.6), is jointly stationary with A. In particular, not only are ϕ_T and its gradient $\nabla\phi_T$ stationary, but also

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

any function of A, ϕ_T and $\nabla\phi_T$. A useful such example is the energy density $(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)$, which is stationary by joint stationarity of A and $\nabla\phi_T$.

Lemma 2.5 (corrector)**.** ([17], *Thm. 3*). *Let $a\in\mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function, then for all $\xi\in\mathbb{R}^d$, there exists a unique random function $\phi:\mathbb{Z}^d\to\mathbb{R}$ which satisfies the corrector equation*

$$-\nabla^*\cdot A(x)\left(\xi+\nabla\phi(x)\right)=0\qquad \text{in } \mathbb{Z}^d, \tag{2.2}$$

and such that $\phi(0)=0$, $\nabla\phi$ is stationary and $\langle\nabla\phi\rangle=0$. In addition, $\langle|\nabla\phi|^2\rangle\lesssim|\xi|^2$.

We also define a regularization of the corrector as follows:

Lemma 2.6 (regularized corrector)**.** ([17], *Proof of Theorem 3*). *Let $a\in\mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function, then for all $T>0$ and $\xi\in\mathbb{R}^d$, there exists a unique stationary random function $\phi_T:\mathbb{Z}^d\to\mathbb{R}$ which satisfies the regularized corrector equation*

$$T^{-1}\phi_T(x)-\nabla^*\cdot A(x)\left(\xi+\nabla\phi_T(x)\right)=0\qquad \text{in } \mathbb{Z}^d. \tag{2.3}$$

In addition, $\langle\phi_T\rangle=0$, and $T^{-1}\langle\phi_T^2\rangle+\left\langle|\nabla\phi_T|^2\right\rangle\lesssim|\xi|^2$.

Definition 2.7 (homogenized coefficients)**.** Let $a\in\mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function and let $\xi\in\mathbb{R}^d$ and ϕ be as in Lemma 2.5. We define the homogenized $d\times d$-matrix $A_{\rm hom}$ as

$$\xi\cdot A_{\rm hom}\xi\;=\;\langle(\xi+\nabla\phi)\cdot A(\xi+\nabla\phi)\rangle\,. \tag{2.4}$$

Note that (2.4) fully characterizes $A_{\rm hom}$ since $A_{\rm hom}$ is a symmetric matrix (it is in particular of the form $a_{\rm hom}{\rm Id}$ for an i. i. d. conductivity function).

2.2. Statement of the main result

We replace ϕ_T by the computable function $\phi_{T,R}$, which is an approximation of ϕ_T on a bounded domain of size $2R$. Let $a\in\mathcal{A}_{\alpha\beta}$, $T>0$, $R\gg1$ and $\xi\in\mathbb{R}^d$. We set $Q_R:=[-R,R)^d\cap\mathbb{Z}^d$ and we let $\phi_{T,R}$ be the solution in $L^2(\mathbb{Z}^d)$ to

$$\left\{\begin{array}{rcll} T^{-1}\phi_{T,R}-\nabla^*\cdot A(\xi+\nabla\phi_{T,R}) &=& 0 & \text{in } Q_R,\\ \phi_{T,R} &=& 0 & \text{on } \mathbb{Z}^d\setminus Q_R.\end{array}\right. \tag{2.5}$$

We then quantify the error we make by replacing ϕ_T by the computable $\phi_{T,R}$. It is given by the following:

Proposition 2.8. *Let $a\in\mathcal{A}_{\alpha\beta}$ be a conductivity function, $T>0$, $R\gg1$ and $\xi\in\mathbb{R}^d$, $|\xi|=1$. Let ϕ_T denote the regularized corrector, and $\phi_{T,R}$ be the solution of (2.5). Then there exists $c>0$ depending only on α,β, and d, such that for all $R-L\geq\sqrt{T}$, we have almost surely*

$$\int_{Q_L}|\nabla\phi_{T,R}(x)-\nabla\phi_T(x)|^2{\rm d}x\;\lesssim\;L^d\left(\left(\frac{R}{R-L}\right)^{d-1/2}T^{3/4}\exp\left(-c\frac{R-L}{\sqrt{T}}\right)\right)^2. \tag{2.6}$$

Remark 2.9. In Proposition 2.8, we only assume that $a\in\mathcal{A}_{\alpha\beta}$. In particular, if a is i. i. d. or more generally stationary ergodic, then ϕ_T is the usual stationary regularized corrector. However, for a general conductivity function a (not necessarily stationary) equation (2.3) cannot be interpreted in some probability space, so that the arguments of [17] do not apply (recall that the r. h. s. $\nabla^*\cdot A\xi$ is not in $L^2(\mathbb{Z}^d)$, which prevents from using standard variational arguments). Instead, we define ϕ_T punctually using the Green representation formula

$$\phi_T(x)=\int_{\mathbb{Z}^d}G_T(x,y)\nabla^*\cdot A(y)\xi{\rm d}y,$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

where the Green's function G_T is defined in Definition 3.1 for any $a \in \mathcal{A}_{\alpha\beta}$ by Riesz' representation theorem. This formula makes sense since $G_T(x,\cdot)$ is in $L^1(\mathbb{Z}^d)$. If a is stationary ergodic, both definitions of ϕ_T are equivalent.

From this proposition, we deduce the main result of this paper.

Theorem 2.10. *Let $a \in \mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function, $T > 0$, $R \gg 1$ and $\xi \in \mathbb{R}^d$, $|\xi| = 1$. Let $\{\phi_{T,R,k}\}_{k=1,\dots,N}$ be the solutions of (2.5) with $N \geq 1$ independent realizations A_k of A, and $A_{\rm hom}$ be the homogenized matrix. For all L such that $R - L \geq \sqrt{T}$ and $L \lesssim T$, we denote by $\mu_L : \mathbb{Z}^d \to [0,1]$ a mask such that $\int_{\mathbb{Z}^d} \mu_L(x)\mathrm{d}x = 1$, $|\nabla \mu_L(x)| \lesssim L^{-d-1}$ and $\mathrm{supp}\,(\mu_L) \subset Q_L$. For all $k \in \{1,\dots,N\}$ we define*

$$\xi \cdot A_{T,R,L,k}\xi \;=\; \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,R,k}(x)) \cdot A_k(x)(\xi + \nabla\phi_{T,R,k}(x))\mu_L(x)\mathrm{d}x,$$

and set

$$\xi \cdot A^N_{T,R,L}\xi \;:=\; \xi \cdot \left(\frac{1}{N}\sum_{k=1}^{N} A_{T,R,L,k}\right)\xi.$$

Then, we have the following error estimate

$$\begin{aligned}\left\langle \left(\xi \cdot A^N_{T,R,L}\xi - \xi \cdot A_{\rm hom}\xi\right)^2 \right\rangle^{1/2} \;\lesssim\; & \left(\frac{R}{L}\right)^{d/2} \left(\frac{R}{R-L}\right)^{d-1/2} T^{3/4} \exp\left(-c\frac{R-L}{\sqrt{T}}\right) \\ & + \begin{cases} d=2 & : \quad (T^{-1} + (NL^2)^{-1/2})(\ln T)^q \\ d=3 & : \quad T^{-3/2} + (NL^3)^{-1/2} \\ d=4 & : \quad T^{-2}\ln T + (NL^4)^{-1/2} \\ d>4 & : \quad T^{-2} + (NL^d)^{-1/2} \end{cases}\end{aligned} \tag{2.7}$$

for some q depending only on α, β, and some c depending further on d.

Let us apply Theorem 2.10 to the strategies described in the introduction. In particular, we have seen that at first approximation, for $d \leq 4$ it seems best to evenly split a given number M of effective sites into the number N of realizations and the number L^d of sites per realization, *i.e.* $N = (2L)^d = \sqrt{M}$. For the reasoning, we consider a buffer zone of size $\sqrt{M}^{1/d}\ln^2 M$ so that the effect due to the boundary conditions is of infinite order. In this case, we split the number of effective sites as

$$\begin{cases} N \;=\; \sqrt{M}\ln^{-2d} M, \\ 2L \;=\; \sqrt{M}^{1/d}\ln^2 M, \\ T \;=\; M^{1/d}, \\ 2R \;=\; 3\sqrt{M}^{1/d}\ln^2 M. \end{cases} \tag{2.8}$$

This choice of parameters amounts to solving $\sqrt{M}\ln^{-2d} M$ equations with $3^d\sqrt{M}\ln^{2d} M$ unknowns each, which gives a total number of $3^d M$ unknowns. Theorem 2.10 then provides the (nearly) optimal error estimate in dimensions 2, 3 and 4:

$$\left\langle \left(\xi \cdot A^N_{T,R,L}\xi - \xi \cdot A_{\rm hom}\xi\right)^2 \right\rangle^{1/2} \;\lesssim\; \begin{cases} d=2 & : \quad M^{-1/2}(\ln M)^q, \\ d=3 & : \quad M^{-1/2}, \\ d=4 & : \quad M^{-1/2}\ln M. \end{cases} \tag{2.9}$$

since for all $\gamma > 0$

$$\exp\left(-c\ln^2 M\right) \;=\; o(M^{-\gamma}).$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

This scaling indeed coincides with the explicit scaling obtained in the case of vanishing ellipticity ratio $1-\frac{\beta}{\alpha} \ll 1$.

Compared to the informal statement of introduction, the effect of the boundary conditions makes this strategy "slightly more expensive" than expected (3^d times as many unknowns). The comparison to the strategy which splits the effective sites as

$$\left\{\begin{array}{l} N = 1, \\ 2L = M^{1/d}, \\ T = M^{1/d}, \\ 2R = 2\left(1+\dfrac{\ln^2(2L)}{\sqrt{2L}}\right)L, \end{array}\right. \tag{2.10}$$

is therefore much less clear. The optimization of N (and R, L) in (2.7) at fixed complexity and fixed rate of convergence is thus nontrivial. It is the object of Section 4.

3. Proofs of the results

We define discrete Green's functions as follows:

Definition 3.1 (discrete Green's function)**.** Let $d \geq 2$. For all $T > 0$, the Green's function $G_T : \mathcal{A}_{\alpha\beta} \times \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{Z}^d, (a, x, y) \mapsto G_T(x, y; a)$ associated with the conductivity function a is defined for all $y \in \mathbb{Z}^d$ and $a \in \mathcal{A}_{\alpha\beta}$ as the unique solution in $L^2_x(\mathbb{Z}^d)$ to

$$\int_{\mathbb{Z}^d} T^{-1} G_T(x, y; a) v(x)\,\mathrm{d}x + \int_{\mathbb{Z}^d} \nabla v(x) \cdot A(x) \nabla_x G_T(x, y; a)\,\mathrm{d}x = v(y), \qquad \forall v \in L^2(\mathbb{Z}^d), \tag{3.1}$$

where A is as in (2.1).

Note that the existence and uniqueness of G_T follows in the discrete case from Riesz' representation theorem. Throughout this paper, we use the shorthand notation $G_T(x, y)$ for $G_T(x, y; a)$.

3.1. Proof of Proposition 2.8

This proof is inspired by the analysis by Bourgeat and Piatnitski in [3], that we adapt here to the discrete setting. In order to prove Proposition 2.8, we need to estimate the pointwise decay of the Green's function G_T and to prove a uniform bound on the approximate corrector field ϕ_T. These are given by the following two auxiliary lemmas.

Lemma 3.2 (pointwise decay estimates)**.** *There exists $c > 0$ depending only on α, β, and d, such that for all $a \in \mathcal{A}_{\alpha\beta}$ and $T > 0$, the Green's function G_T satisfies the pointwise estimates: For all $x, y \in \mathbb{Z}^d$,*

$$\text{for } d = 2: \quad G_T(x, y) \lesssim \ln\Big(\frac{\sqrt{T}}{1+|x-y|}\Big) \exp\left(-c\frac{|x-y|}{\sqrt{T}}\right), \tag{3.2}$$

$$\text{for } d > 2: \quad G_T(x, y) \lesssim (1+|x-y|)^{2-d} \exp\left(-c\frac{|x-y|}{\sqrt{T}}\right). \tag{3.3}$$

Lemma 3.3. *Let $a \in \mathcal{A}_{\alpha\beta}$, $T \gg 1$, and $\xi \in \mathbb{R}^d$, $|\xi| = 1$. The approximate corrector ϕ_T satisfies the following uniform bound*

$$\sup_{\mathbb{Z}^d} |\phi_T| \lesssim \sqrt{T}. \tag{3.4}$$

Note that this bound is sharper than the one used for the continuous case in [3], Formula (25). We first prove Proposition 2.8, and then turn to Lemma 3.3. The proof of Lemma 3.2 is postponed to Appendix B.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

Proof of Proposition 2.8. We divide the proof into two steps.

Step 1. Proof of the estimate

$$|\phi_T(x) - \phi_{T,R}(x)| \lesssim \left(\frac{R}{\rho}\right)^{d-1/2} T^{3/4} \exp\left(-c\frac{\rho}{\sqrt{T}}\right) \tag{3.5}$$

for all $x \in Q_{R-\rho}$, with $\rho \gtrsim \sqrt{T}$, and some c depending only on d and the ellipticity constants α, β.

The function $\phi_T - \phi_{T,R}$ is solution to

$$\left\{ \begin{array}{rcll} T^{-1}(\phi_T - \phi_{T,R}) - \nabla^* \cdot A(\nabla(\phi_T - \phi_{T,R})) &=& 0 & \text{in } Q_R, \\ \phi_T - \phi_{T,R} &=& \phi_T & \text{on } \mathbb{Z}^d \setminus Q_R. \end{array} \right. \tag{3.6}$$

Let φ_0 denote the trivial lifting of $\phi_{T|\mathbb{Z}^d \setminus Q_R}$ in Q_R:

$$\left\{ \begin{array}{rcll} \varphi_0(x) &=& 0 & \text{in } Q_R, \\ \varphi_0(x) &=& \phi_T(x) & \text{on } \mathbb{Z}^d \setminus Q_R. \end{array} \right.$$

We may then write $\phi_T - \phi_{T,R} = \varphi_0 + \varphi_1$, where φ_1 is the solution to

$$\left\{ \begin{array}{rcll} T^{-1}\varphi_1 - \nabla^* \cdot A\nabla\varphi_1 &=& \underbrace{-T^{-1}\varphi_0}_{=0} + \underbrace{\nabla^* \cdot A\nabla\varphi_0}_{=:\tilde{\varphi}_0} & \text{in } Q_R, \\ \varphi_1 &=& 0 & \text{on } \mathbb{Z}^d \setminus Q_R. \end{array} \right. \tag{3.7}$$

Note that

$$\left\{ \begin{array}{rcll} \nabla\varphi_0(x) &=& 0 & \text{for } x \in Q_R,\ d(x, \mathbb{Z}^d \setminus Q_R) \geq 2, \\ \sup_{x \in Q_R} |\varphi_0(x)| &\lesssim& \sup |\phi_T| \overset{(3.4)}{\lesssim} \sqrt{T}. & \end{array} \right. \tag{3.8}$$

We next use the Green representation formula. To this aim, we define the Green's function $G_{T,R}(\cdot, y) : \mathbb{Z}^d \to \mathbb{R}$ for all $y \in Q_R$ as the unique solution to

$$\left\{ \begin{array}{rcll} T^{-1}G_{T,R}(x,y) - \nabla_x^* \cdot A(\nabla_x G_{T,R}(x,y))) &=& \delta(y-x) & \text{in } Q_R, \\ G_{T,R}(x,y) &=& 0 & \text{on } \mathbb{Z}^d \setminus Q_R. \end{array} \right. \tag{3.9}$$

By the maximum principle, for all $x \in Q_R$ we have

$$0 \leq G_{T,R}(x,y) \leq G_T(x,y). \tag{3.10}$$

For all $x \in Q_R$, we then rewrite $\varphi_1(x)$ as

$$\begin{aligned} \varphi_1(x) &= \int_{Q_R} G_{T,R}(x,y)\tilde{\varphi}_0(y)\,\mathrm{d}y \\ &= \int_{Q_R} G_{T,R}(x,y)\nabla^* \cdot A(y)\nabla\varphi_0(y)\,\mathrm{d}y. \end{aligned}$$

By integration by parts (recall that $G_{T,R}(x,y) = 0$ for $y \in \mathbb{Z}^d \setminus Q_R$) and using the fact that $\nabla\varphi_0$ is supported on the boundary of Q_R, this yields

$$\varphi_1(x) = -\int_{Q_R \setminus Q_{R-1}} \nabla_y G_{T,R}(x,y) \cdot A(y)\nabla\varphi_0(y)\,\mathrm{d}y. \tag{3.11}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

We use Cauchy-Schwarz' inequality, the boundedness of A, and (3.8) in the form of the uniform bound $\sup|\nabla\varphi_0| \lesssim \sup|\varphi_0| \lesssim \sqrt{T}$ to obtain

$$\begin{aligned}\left|\int_{Q_R\setminus Q_{R-1}} \nabla_y G_{T,R}(x,y)\cdot A(y)\nabla\varphi_0(y)\,\mathrm{d}y\right| \\ \lesssim\ R^{(d-1)/2}\sqrt{T}\left(\int_{Q_R\setminus Q_{R-1}}|\nabla_y G_{T,R}(x,y)|^2\mathrm{d}y\right)^{1/2} \\ =\ R^{(d-1)/2}\sqrt{T}\left(\int_{Q_R\setminus Q_{R-1}}|\nabla_y G_{T,R}(y,x)|^2\mathrm{d}y\right)^{1/2}.\end{aligned} \tag{3.12}$$

In the last line we've used that the symmetry property $G_{T,R}(x,y) = G_{T,R}(y,x)$ of the Green's function (see [9], Proof of Corollary 2.3, Step 1) yields the identity $\nabla_y G_{T,R}(x,y) = \nabla_y G_{T,R}(y,x)$.

We then appeal to Cacciopoli's inequality. To this aim, we recall that $\rho \gtrsim \sqrt{T}$, and we let $\eta_\rho : Q_R \to [0,1]$ be a cut-off function such that

$$\begin{aligned}\text{for } y\in Q_{R-\rho/2} &: \eta_\rho(y) = 0,\\ \text{for } y\in Q_R\setminus Q_{R-1} &: \eta_\rho(y) = 1,\\ \text{for } y\in Q_R &: |\nabla\eta_\rho(y)| \lesssim \rho^{-1}.\end{aligned} \tag{3.13}$$

Since $G_{T,R} = 0$ on $\mathbb{Z}^d\setminus Q_R$, multiplying the defining equation (3.9) for $G_{T,R}$ by $\eta_\rho^2 G_{T,R}$, and integrating by parts on Q_R yield the following discrete Cacciopoli estimate (see [9], Proof of Lemma 2.8, Step 1, for details)

$$\int_{Q_R}\eta_\rho^2(y)|\nabla_y G_{T,R}(y,x)|^2\mathrm{d}y \lesssim \int_{Q_R} G_{T,R}^2(y,x)|\nabla\eta_\rho(y)|^2\mathrm{d}y,$$

provided that $x\in Q_{R-\rho}$. By the properties (3.13) of η_ρ, this implies

$$\int_{Q_R\setminus Q_{R-1}}|\nabla G_{T,R}(y,x)|^2\mathrm{d}y \lesssim \rho^{-2}\int_{Q_R\setminus Q_{R-\rho/2}} G_{T,R}^2(y,x)\mathrm{d}y. \tag{3.14}$$

We are now in position to estimate (3.12) for all $x\in Q_{R-\rho}$. By the Cacciopoli estimate (3.14), the maximum principle (3.10), and the pointwise estimates on G_T from Lemma 3.2, (3.12) turns into

$$\begin{aligned}\left|\int_{Q_R\setminus Q_{R-1}} \nabla_y G_{T,R}(x,y)\cdot A(y)\nabla\varphi_0(y)\,\mathrm{d}y\right| \\ \lesssim\ R^{(d-1)/2}\left(\rho^{-2}R^d\rho^{2(2-d)}\exp\left(-2c\frac{\rho/2}{\sqrt{T}}\right)\right)^{1/2}\sqrt{T} \\ =\ \left(\frac{R}{\rho}\right)^{d-1/2}T^{3/4}\left(\frac{\rho}{\sqrt{T}}\right)^{1/2}\exp\left(-c\frac{\rho}{2\sqrt{T}}\right)\end{aligned} \tag{3.15}$$

for all $x\in Q_{R-\rho}$.

The combination of (3.11), (3.15), and the definition of φ_1 shows (3.5) for some constant $c>0$ depending only on d, and α,β.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

Step 2. *Proof of (2.6).*
We first bound $|\nabla\phi_T(x)|$ by $\sum_{i=1}^d |\phi_T(x)| + |\phi_T(x+\mathbf{e}_1)|$ and integrate inequality (3.5) over Q_L for $L = R-\rho$, $\rho \gtrsim \sqrt{T}$, obtaining

$$\int_{Q_L} |\nabla\phi_{T,R}(x) - \nabla\phi_T(x)|^2 \mathrm{d}x \lesssim L^d \left(\left(\frac{R}{R-L} \right)^{d-1/2} T^{3/4} \exp\left(-c\frac{R-L}{\sqrt{T}} \right) \right)^2,$$

as desired. □

Proof of Lemma 3.3. We start with the Green representation formula, and perform an integration by parts using that G_T is in $L^1(\mathbb{Z}^d)$ by [9], Corollary 2.2:

$$\begin{aligned} \phi_T(x) &= \int_{\mathbb{Z}^d} G_T(x,y)\nabla^* \cdot A(y)\xi \,\mathrm{d}y \\ &= -\int_{\mathbb{Z}^d} \nabla_y G_T(x,y) \cdot A(y)\xi \,\mathrm{d}y. \end{aligned}$$

This yields

$$|\phi_T(x)| \lesssim \int_{|x-y|\le\sqrt{T}} |\nabla_y G_T(x,y)| \,\mathrm{d}y + \int_{|x-y|>\sqrt{T}} |\nabla_y G_T(x,y)| \,\mathrm{d}y. \tag{3.16}$$

To proceed with the estimate, we reproduce [9], Lemma 2.9, for the reader's convenience.

Lemma 3.4. *Let $a \in \mathcal{A}_{\alpha\beta}$ be a conductivity function, and G_T be its associated Green's function. Then, for $d \ge 2$, for all $T > 0$, $k > 0$, $R \gg 1$, and $x \in \mathbb{Z}^d$*

$$\int_{R\le|x-y|\le 2R} |\nabla_y G_T(x,y)|^2 \mathrm{d}y \lesssim R^d (R^{1-d})^2 \min\{1, \sqrt{T}R^{-1}\}^k.$$

We begin with the second term of the r. h. s. of (3.16). We divide the integration on $\{y : |x-y| > \sqrt{T}\}$ as the integration on annuli of the form $\{y : 2^i\sqrt{T} < |x-y| \le 2^{i+1}\sqrt{T}\}$ for $i \in \mathbb{N}$, and appeal to the decay of ∇G_T on such annuli from Lemma 3.4 for $k = 4$. This yields by Cauchy-Schwarz' inequality

$$\begin{aligned} &\int_{|x-y|>\sqrt{T}} |\nabla G_T(x,y)| \,\mathrm{d}y \\ &\quad\le \sum_{i\in\mathbb{N}} \left(\int_{2^i\sqrt{T}<|x-y|\le 2^{i+1}\sqrt{T}} |\nabla G_T(x,y)|^2 \mathrm{d}y \right)^{1/2} (2^i\sqrt{T})^{d/2} \\ &\overset{\text{Lemma 3.4}}{\lesssim} \sum_{i\in\mathbb{N}} \left((2^i\sqrt{T})^{d+2(1-d)} (2^i)^{-4} \right)^{1/2} (2^i\sqrt{T})^{d/2} \\ &\quad= \sqrt{T} \sum_{i\in\mathbb{N}} (2^i)^{-1} \lesssim \sqrt{T}. \end{aligned} \tag{3.17}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

For the first term of the r. h. s. of (3.16), we also make use of a dyadic decomposition of space. Let $R = 2^{-I}\sqrt{T} \sim 1$, $I \in \mathbb{N}$, be such that Lemma 3.4 applies on annuli of the form $\{y : 2^{-i-1}\sqrt{T} < |x-y| \le 2^{-i}\sqrt{T}\}$ for $i \le I-1$ (R has to be large enough although of order unity). We then split the integration on $\{y : |x-y| \le \sqrt{T}\}$ as the integration on the ball of radius $R \sim 1$, and the integration over annuli of the form $\{y : 2^{-i-1}\sqrt{T} < |x-y| \le 2^{-i}\sqrt{T}\}$ for $i \in \{0,\dots,I-1\}$. For the integral on the ball, we appeal to the uniform estimate $|\nabla G_T| \lesssim 1$ from [9], Corollary 2.3, and for the integrals on the annuli, we appeal once more to the decay of Lemma 3.4. By Cauchy-Schwarz' inequality, this yields

$$\begin{aligned}
&\int_{|x-y|\le\sqrt{T}} |\nabla G_T(x,y)|\,\mathrm{d}y \\
&\lesssim \int_{|x-y|\le R} |\nabla G_T(x,y)|\,\mathrm{d}y \\
&\quad + \sum_{i=0}^{I-1} \left(\int_{2^{-i-1}\sqrt{T}<|x-y|\le 2^{-i}\sqrt{T}} |\nabla G_T(x,y)|^2\,\mathrm{d}y \right)^{1/2} (2^{-i}\sqrt{T})^{d/2} \qquad (3.18)\\
&\lesssim 1 + \sum_{i=0}^{I-1} \Big((2^{-i}\sqrt{T})^{d+2(1-d)} \Big)^{1/2} (2^{-i}\sqrt{T})^{d/2} \\
&= 1 + \sqrt{T} \sum_{i=0}^{I-1} 2^{-i} \lesssim \sqrt{T}. \qquad (3.19)
\end{aligned}$$

The claim of the lemma now follows from the combination of (3.16), (3.17), and (3.19). □

3.2. **Proof of Theorem 2.10**

To prove Theorem 2.10, we combine the variance estimate of [9], Theorem 2.1, and Remark 2.1, and the estimate of the systematic error in [10], Theorem 1, with Proposition 2.8.

We start with the triangle inequality

$$\begin{aligned}
&\left\langle \left(\frac{1}{N} \sum_{k=1}^{N} \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,R,k}) \cdot A_k (\xi + \nabla\phi_{T,R,k}) \mu_L - \xi \cdot A_{\mathrm{hom}} \xi \right)^2 \right\rangle^{1/2} \\
&\quad \le \left\langle \left(\frac{1}{N} \sum_{k=1}^{N} \int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k}) \cdot A_k (\xi + \nabla\phi_{T,R,k}) \right. \right. \\
&\qquad \left. \left. - (\xi + \nabla\phi_{T,k}) \cdot A_k (\xi + \nabla\phi_{T,k}) \Big) \mu_L \right)^2 \right\rangle^{1/2} \\
&\qquad + \left\langle \left(\frac{1}{N} \sum_{k=1}^{N} \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,k}) \cdot A_k (\xi + \nabla\phi_{T,k}) \mu_L - \xi \cdot A_{\mathrm{hom}} \xi \right)^2 \right\rangle^{1/2}. \qquad (3.20)
\end{aligned}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

We first deal with the first term of the r. h. s. of (3.20). To this aim, we expand the square, which yields N^2 terms. By Cauchy-Schwarz' inequality in probability, each term is bounded by the same single term: For all $k, k' \in \{1, \dots, N\}$,

$$\begin{aligned}
&\Big\langle \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k}) \cdot A_k(\xi + \nabla\phi_{T,R,k}) - (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k}) \Big) \mu_L \Big) \\
&\qquad \times \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k'}) \cdot A_{k'}(\xi + \nabla\phi_{T,R,k'}) - (\xi + \nabla\phi_{T,k'}) \cdot A_{k'}(\xi + \nabla\phi_{T,k'}) \Big) \mu_L \Big) \Big\rangle \\
&\qquad \le \Big\langle \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k}) \cdot A_k(\xi + \nabla\phi_{T,R,k}) - (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k}) \Big) \mu_L \Big)^2 \Big\rangle^{1/2} \\
&\qquad \times \Big\langle \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k'}) \cdot A_{k'}(\xi + \nabla\phi_{T,R,k'}) - (\xi + \nabla\phi_{T,k'}) \cdot A_{k'}(\xi + \nabla\phi_{T,k'}) \Big) \mu_L \Big)^2 \Big\rangle^{1/2} \\
&\qquad = \Big\langle \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R}) \cdot A(\xi + \nabla\phi_{T,R}) - (\xi + \nabla\phi_{T}) \cdot A(\xi + \nabla\phi_{T}) \Big) \mu_L \Big)^2 \Big\rangle,
\end{aligned}$$

where we have dropped the subscripts k and k' since the A_k have the same law. From this and the symmetry of A, we deduce

$$\begin{aligned}
&\Big\langle \Big(\frac{1}{N} \sum_{k=1}^{N} \int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R,k}) \cdot A_k(\xi + \nabla\phi_{T,R,k}) \\
&\qquad - (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k}) \Big) \mu_L \Big)^2 \Big\rangle^{1/2} \\
&\qquad \le \Big\langle \Big(\int_{\mathbb{Z}^d} \Big((\xi + \nabla\phi_{T,R}) \cdot A(\xi + \nabla\phi_{T,R}) - (\xi + \nabla\phi_{T}) \cdot A(\xi + \nabla\phi_{T}) \Big) \mu_L \Big)^2 \Big\rangle^{1/2} \\
&\qquad = \Big\langle \Big(\int_{\mathbb{Z}^d} (\nabla\phi_{T,R} - \nabla\phi_T) \cdot A(2\xi + \nabla\phi_{T,R} + \nabla\phi_T) \mu_L \Big)^2 \Big\rangle^{1/2}.
\end{aligned}$$

To bound this term, we make use of the *a priori* estimate

$$\int_{\mathbb{Z}^d} |\nabla\phi_{T,R}(x)|^2 \mathrm{d}x \lesssim R^d, \tag{3.21}$$

that we obtain by integration by parts after testing (2.5) with $\phi_{T,R}$ itself. We then use Cauchy-Schwarz' inequality in $\mathbb{Z}^d$, Proposition 2.8, the properties of μ_L, and the *a priori* estimates on $\nabla\phi_T$ and $\nabla\phi_{T,R}$ to bound

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

the r. h. s.

$$\begin{aligned}
&\left\langle \left(\int_{\mathbb{Z}^d} (\nabla\phi_{T,R}(x) - \nabla\phi_T(x)) \cdot A(x)(2\xi + \nabla\phi_{T,R}(x) + \nabla\phi_T(x))\mu_L(x)\mathrm{d}x \right)^2 \right\rangle \\
&\quad \lesssim \quad \left\langle \frac{1}{L^d} \int_{Q_L} |\nabla\phi_{T,R}(x') - \nabla\phi_T(x')|^2 \mathrm{d}x' \int_{\mathbb{Z}^d} (1 + |\nabla\phi_{T,R}(x)|^2 + |\nabla\phi_T(x)|^2)\mu_L(x)\mathrm{d}x \right\rangle \\
&\quad \overset{(2.6)}{\lesssim} \quad \left(\left(\frac{R}{R-L} \right)^{d-1/2} T^{3/4} \exp\left(-c\frac{R-L}{\sqrt{T}} \right) \right)^2 \\
&\quad \times \quad \Big(1 + \underbrace{\left\langle \int_{\mathbb{Z}^d} |\nabla\phi_{T,R}(x)|^2 \mu_L(x)\mathrm{d}x \right\rangle}_{\overset{(3.21)}{\lesssim} \left(\frac{R}{L}\right)^d} + \int_{\mathbb{Z}^d} \underbrace{\langle |\nabla\phi_T|^2 \rangle}_{\underset{\lesssim}{\text{Lemma 2.6}} 1} \mu_L(x)\mathrm{d}x \Big) \\
&\quad \lesssim \quad \left(\frac{R}{L} \right)^d \left(\left(\frac{R}{R-L} \right)^{d-1/2} T^{3/4} \exp\left(-c\frac{R-L}{\sqrt{T}} \right) \right)^2 . \qquad (3.22)
\end{aligned}$$

We then recall that for all k,

$$\begin{aligned}
\left\langle \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k})\mu_L \right\rangle &= \int_{\mathbb{Z}^d} \langle (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k}) \rangle \mu_L \\
&= \xi \cdot A_T\xi \int_{\mathbb{Z}^d} \mu_L = \xi \cdot A_T\xi
\end{aligned}$$

by stationarity of the energy density, so that the second term of the r. h. s. (3.20) can be split into a variance part and a systematic error, using the independence of the A_k (see [10], (1.8) and (1.9) for details):

$$\begin{aligned}
&\left\langle \left(\frac{1}{N} \sum_{k=1}^N \int_{\mathbb{Z}^d} (\xi + \nabla\phi_{T,k}) \cdot A_k(\xi + \nabla\phi_{T,k})\mu_L - \xi \cdot A_{\mathrm{hom}}\xi \right)^2 \right\rangle \\
&\quad = \quad \frac{1}{N} \mathrm{var}\left[\int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\mu_L \right] \\
&\qquad + \left(\left\langle \int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\mu_L \right\rangle - \xi \cdot A_{\mathrm{hom}}\xi \right)^2 .
\end{aligned}$$

For the variance term, we appeal to [9], Theorem 2.1 and Remark 2.1,

$$\mathrm{var}\left[\int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\mu_L \right] \lesssim \left\{ \begin{array}{lcl} d = 2 & : & (L^{-2} + T^{-2}) \ln^q T \\ d > 2 & : & L^{-d}(1 + T^{-1}L), \end{array} \right.$$

and for the systematic error, we appeal to [10], Theorem 1

$$\left(\left\langle \int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\mu_L \right\rangle - \xi \cdot A_{\mathrm{hom}}\xi \right)^2 \lesssim \left\{ \begin{array}{lcl} d = 2 & : & T^{-2} \ln^q T \\ d = 3 & : & T^{-3} \\ d = 4 & : & T^{-4} \ln^2 T \\ d > 4 & : & T^{-4}, \end{array} \right.$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

so that

$$\left\langle \left(\frac{1}{N} \sum_{k=1}^{N} \int_{\mathbb{Z}^d} (\xi + \nabla \phi_{T,k}) \cdot A_k (\xi + \nabla \phi_{T,k}) \mu_L - \xi \cdot A_{\rm hom} \xi \right)^2 \right\rangle \lesssim \begin{cases} d = 2 & : \quad (N^{-1} L^{-2} + T^{-2}) \ln^q T, \\ d = 3 & : \quad N^{-1} L^{-3} (1 + T^{-1} L) + T^{-3}, \\ d > 4 & : \quad N^{-1} L^{-4} (1 + T^{-1} L) + T^{-4} \ln^2 T, \\ d > 4 & : \quad N^{-1} L^{-d} (1 + T^{-1} L) + T^{-4}. \end{cases} \tag{3.23}$$

The combination of (3.20), (3.22), and (3.23) concludes the proof of (2.7), using in addition the assumption $T \gtrsim L$.

4. Numerical strategy and validation

In this section, we propose a complexity analysis for the computation of $A^N_{T,R,L}$. In particular, we identify the number of realizations $N_{\rm opt}$ (and the associated parameters T, R, L) which minimizes the computational cost to approximate $A_{\rm hom}$ at a given precision. *Precision* is understood here as the *scaling of the error* in terms of the effective number of sites M (see below) of the approximation, as in (2.9) (in particular, we disregard prefactors). The answer depends on the dimension and on the linear solver used. We treat the cases $d = 2, 3$, with a preconditioned conjugate gradient method, and a Cholesky method to solve the linear problems. Whereas the preconditioned conjugate gradient method is the most efficient solution method for this problem, we also provide the analysis of the Cholesky method in view of its application to linear elasticity. Although our techniques of proofs crucially rely on the scalar character of the equation, we believe that the results of this paper are "likely to hold" in the case of linear elasticity considered in [14]. Another application of interest to us is the numerical approximation of the discrete model for rubber studied in [2], which is a nonlinear version of the discrete elliptic equation dealt with here. In those cases, the linear system is ill-conditioned, and direct solvers such as the Cholesky method are to be used. This motivates us to consider direct solvers for the complexity analysis.

In order to illustrate our main result and check the accuracy of this complexity analysis, we have conducted a series of numerical tests on the following problem: $d = 2$, the coefficients a are i. i. d. taking values $\alpha = 1$ and $\beta = 9$ with probability $1/2$. As proved in Appendix A, Dykhne's formula (see the original paper [6], and the monograph [13], Sect. 1.5) holds true in this particular *discrete* case, so that the associated homogenized matrix $A_{\rm hom}$ is given by

$$A_{\rm hom} = \sqrt{\alpha \beta}\, Id = 3\, Id.$$

We then identify $N_{\rm opt}$ for this problem and compare the computation time to the largest (reasonable) N with parallel computing (that is, with N computers). In the last subsection, we compare this method to standard approaches used in the literature. We focus in particular on the importance of the zero-order term in the equation from a numerical point of view.

4.1. Complexity analysis

Let $L \in \mathbb{N}$. If we take $N = 1$, $T = 2L$ and $R = L + b_f \sqrt{T} \ln^2 T$ for some $b_f > 0$ (b_f for buffer zone), Theorem 2.10 ensures that the error on the approximation of $A_{\rm hom}$ is of order $M^{-1/2}$ (up to the logarithmic correction for $d = 2$), with $M := (2L)^d$. Since the associated approximation of $A_{\rm hom}$ is given by a weighted sum of the energy density $(\xi + \nabla \phi_{T,R}(x)) \cdot A(x)(\xi + \nabla \phi_{T,R}(x))$ at exactly $M = (2L)^d$ sites, we recall we shall say that M is the "effective number of sites". Note that it differs from the total number of sites, which is $(2R)^d > (2L)^d = M$.

As discussed in the introduction of this paper, one may distribute the effective number of sites on several smaller independent domains, while keeping the error on the approximation of $A_{\rm hom}$ of order $M^{-1/2}$ (up to a logarithmic correction for $d = 2$). Let N be a number of domains. On the one hand, in order to keep the

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

effective number of sites fixed, we take $L_N := LN^{-1/d}$. On the other hand, in order to keep the precision unchanged, we still need $T \sim L$, and the buffer zone to be of the same order as for $N = 1$. Hence we set $T_N := T = 2L$, and $R_N := L_N + b_f\sqrt{T}\ln^2 T$. The error scales therefore as $M^{-1/2}$ (up to a logarithmic correction for $d = 2$), the effective number of sites is still $N(2L_N)^d = M$, whereas the total number of unknowns is now $N(2R_N)^d > (2R)^d > M$.

In order to make a complexity analysis, one needs to make precise the linear systems to be solved depending on N and M. Let $\mathcal{L}$ denote a symmetric positive definite matrix of order $l \in \mathbb{N}$. Assume further that it has a fixed number δ_d of diagonals (in our discrete case: $\delta_2 = 5$ for $d = 2$, $\delta_3 = 7$ for $d = 3$) and a bandwidth $b \leq l$. Then, solving the system

$$\mathcal{L}X = B$$

in $\mathbb{R}^l$ by a conjugate gradient method requires *approximately* b iterations, and therefore $\mathcal{C}_{\rm CG}(\mathcal{L}) \sim b\delta_d l$ operations. When solved by a Cholesky method, it *exactly* requires $\mathcal{C}_{\rm Chol}(\mathcal{L}) \sim b^2\delta_d l$ operations.

Let us now use these complexity estimates in the homogenization problem under investigation. For our difference operators, l should be replaced by $(2R_N)^d$ and b by $(2R_N)^{d-1}$ (we skip the dependence on δ_d for the comparison). Hence, the overall number of operations to solve the N problems is

$$\begin{aligned}
\Gamma_{\rm CG}(N,2L) &\sim N\left(2LN^{-1/d} + 2b_f\sqrt{2L}\ln^2(2L)\right)^{2d-1} \\
&= N^{-1+1/d}\left(2L + 2b_f N^{1/d}\sqrt{2L}\ln^2(2L)\right)^{2d-1}, \\
\Gamma_{\rm Chol}(N,2L) &\sim N\left(2LN^{-1/d} + 2b_f\sqrt{2L}\ln^2(2L)\right)^{2(d-1)+d} \\
&= N^{-2+2/d}\left(2L + 2b_f N^{1/d}\sqrt{2L}\ln^2(2L)\right)^{3d-2}.
\end{aligned}$$

We have essentially two extreme strategies for the choice of N, the effective number of sites $N(2LN^{-1/d})^d = M$ being fixed. Either we choose $N_{\rm opt}$ which minimizes the cost of solving N problems of size $R_N = 2LN^{-1/d} + b_f\sqrt{2L}\ln^2(2L)$, *i.e.* such that

$$\Gamma(N_{\rm opt},2L) \leq \Gamma(N,2L)$$

for all $N \leq 2L$, where Γ denotes a cost function ($\Gamma_{\rm CG}$ or $\Gamma_{\rm Chol}$). Or, given an arbitrary number of processors, we choose N in order to minimize the effective time to solve the N problems using parallel computing (recall that the N problems are completely independent). The first option consists in minimizing the time on one single processor, the second option, on an arbitrary number of processors.

In order to make the complexity analysis concrete, we fix $b_f = 0.1$, since this is the value we use in the numerical tests. We first treat the case $d = 2$, and then the case $d = 3$. Note that the complexity analysis for the Cholesky method is *exact* since the number of operations involved is known *a priori*. For the conjugate gradient, this is not the case and the number of operations depends on the number of iterations (which in turn does not only depend on the tolerance required but also on the preconditioner used). The present discussion is therefore only *qualitative* for the conjugate gradient method. Numerical tests will complete the discussion for the conjugate gradient method in dimension 2.

Dimension $d = 2$

The ratio $\Gamma_{\rm CG}(N,2L)/\Gamma_{\rm CG}(1,2L)$ is plotted in Figure 1 for $N \in \{1,\dots,10\}$ and $2L \in \{10,10^2,10^3,10^4,10^5\}$ (that is from 100 to 10^{10} effective unknowns). Except for $2L = 10^2$, this ratio is minimal for some $N_{\rm opt} \neq 1$. Hence it seems more advantageous to make several computations (say 3 for a typical number unknowns of $M = 10^6$) on smaller domains. For the Cholesky method (see Fig. 2 for $N \in \{1,\dots,100\}$), this is even more clear, and the gain is much larger.

For the second strategy, the objective is to minimize the effective computational time given a fixed number S of processors. Roughly speaking, it is reasonable to take $N \geq S$, whatever the method. There are then two

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

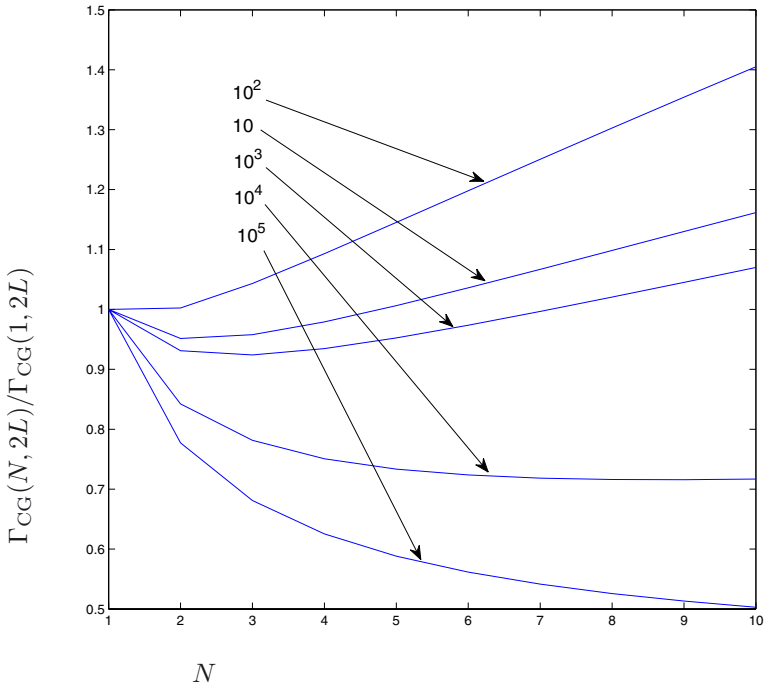


FIGURE 1. Complexity for $d = 2$ with a conjugate gradient method, and $2L = 10^1, \ldots, 10^5$.

regimes. In the first one (small number of processors), the gain is linear in the number of processors: If the number of processors is doubled, the effective time is divided by two. This is the scalable regime. Eventually, it saturates and the relative gain in effective time decreases. In Figures 3 and 4, the effective time is plotted in function of the number of processors (for which N is optimized) in logarithmic scale, for the two methods and with $2L = 10^3$ (which corresponds to $M = 10^6$ unknowns). The thick line is the effective time, whereas the thin straight line is the perfect scalable regime. It is a lower bound. For the conjugate gradient method, the problem is scalable up to 15 processors, whereas for the Cholesky method, the problem can be considered scalable up to 50 processors.

Dimension $d = 3$

The conclusions are essentially the same as for $d = 2$. The ratios $\Gamma_{\mathrm{CG}}(N, 2L)/\Gamma_{\mathrm{CG}}(1, 2L)$ and $\Gamma_{\mathrm{Chol}}(N, 2L)/\Gamma_{\mathrm{Chol}}(1, 2L)$ are plotted in Figures 5 and 6, respectively, for $N \in \{1, \ldots, 100\}$. Note that $N = 1$ is never the optimal choice (especially for the Cholesky method, for which the cost can be reduced by a factor 3 for $M = 10^6$ by taking $N = 31$). The effective time in function of the number of processors is plotted in logarithmic scale for $M = 10^6$ unknows in Figures 7 and 8, for both the conjugate gradient and Cholesky methods. For the conjugate gradient method, although the gain in time remains important, the problem is only scalable up to 8 processors. For the Cholesky method, the method is perfectly scalable in the regime considered.

4.2. Numerical tests with the conjugate gradient method in dimension $d = 2$

The first series of tests illustrates Theorem 2.10 for $N = 1$, with

$$\begin{cases} 2L = \sqrt{M} \\ T = 2L + 3, \\ R = \left(1 + 0.1\dfrac{\ln^2(2L)}{\sqrt{2L}}\right)(L+1), \end{cases}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

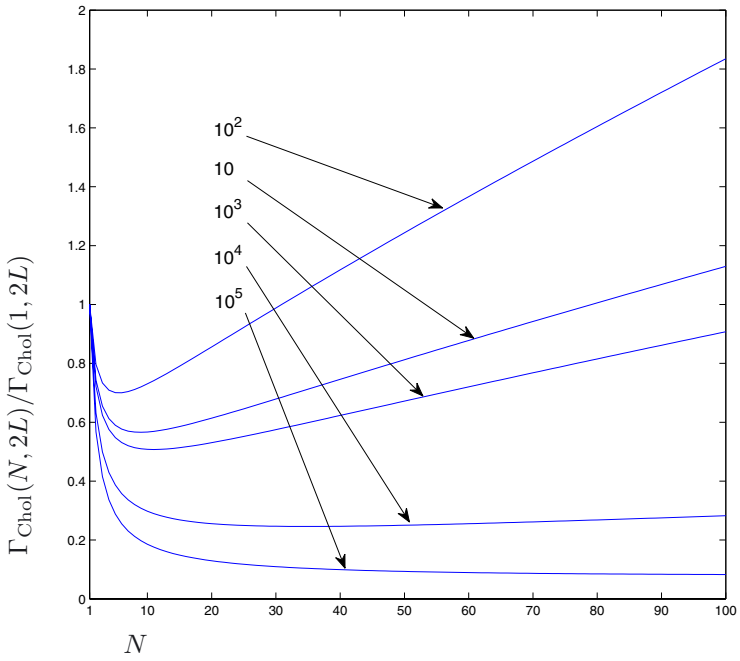


FIGURE 2. Complexity for $d = 2$ with a Cholesky method, and $2L = 10^1, \ldots, 10^5$.

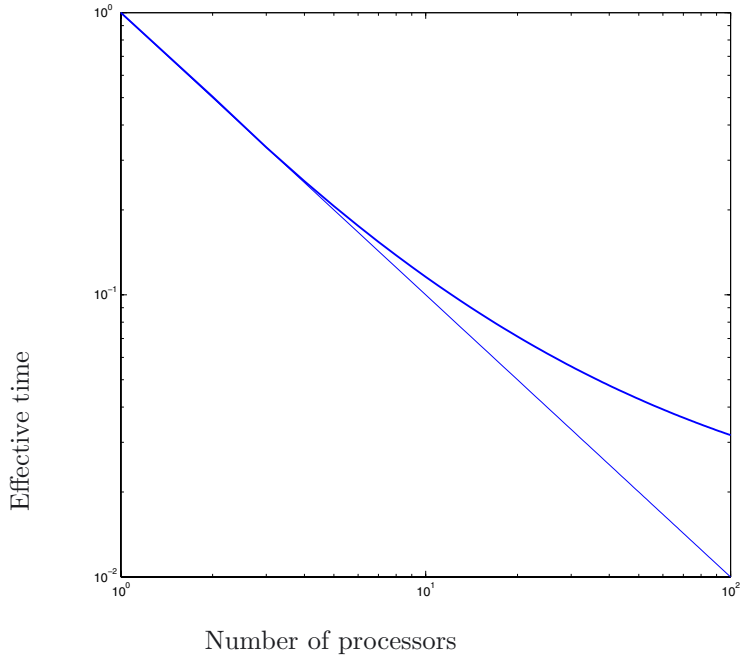


FIGURE 3. Effective time in function of the number of processors for $d = 2$ with a conjugate gradient method (dots), and $2L = 10^3$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

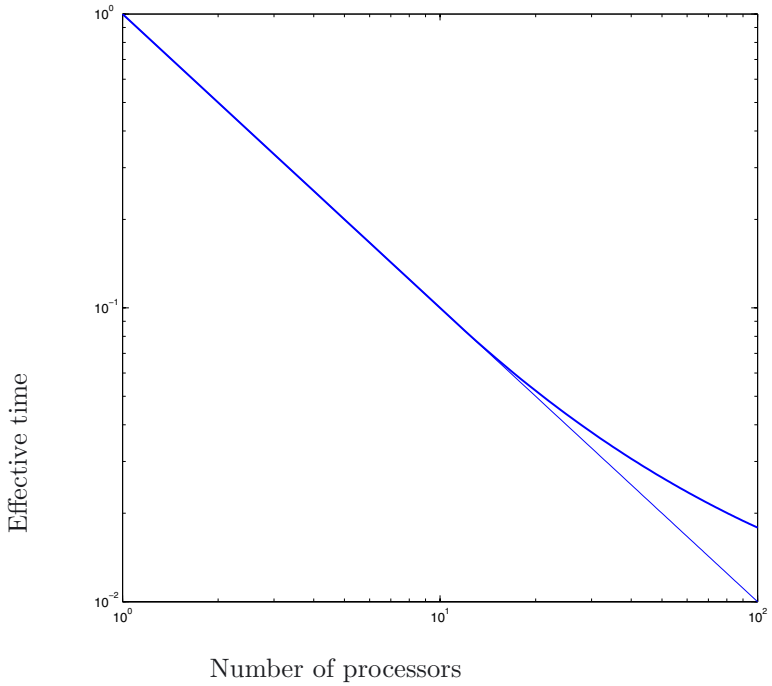


FIGURE 4. Effective time in function of the number of processors for $d = 2$ with a Cholesky method (dots), and $2L = 10^3$.

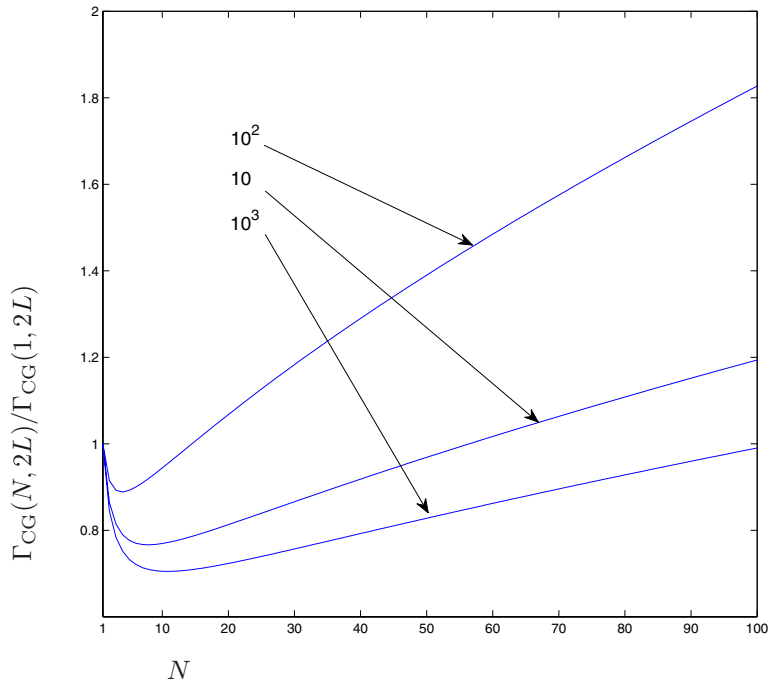


FIGURE 5. Complexity for $d = 3$ with a conjugate gradient method, and $2L = 10^1, 10^2, 10^3$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

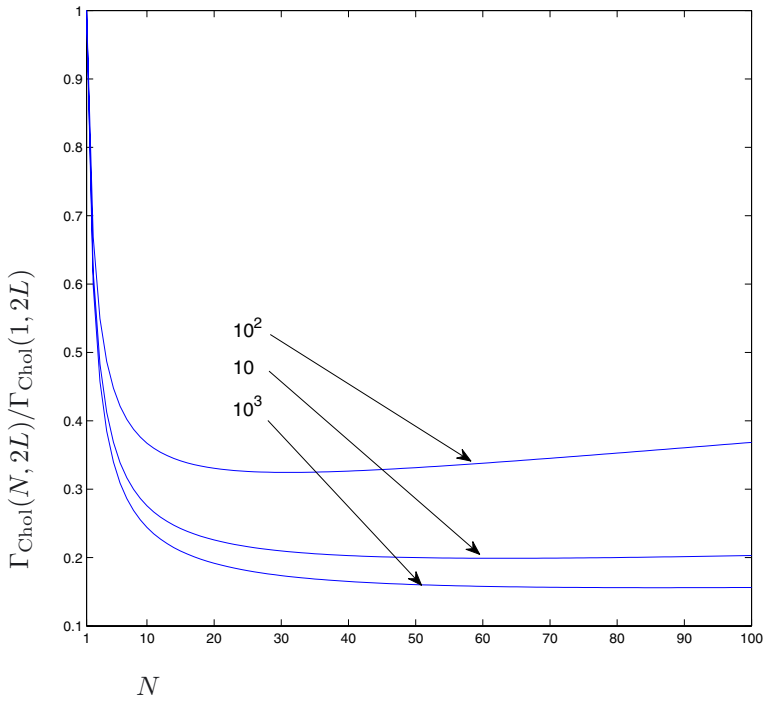


FIGURE 6. Complexity for $d = 3$ with a Cholesky method, and $2L = 10^1, 10^2, 10^3$.

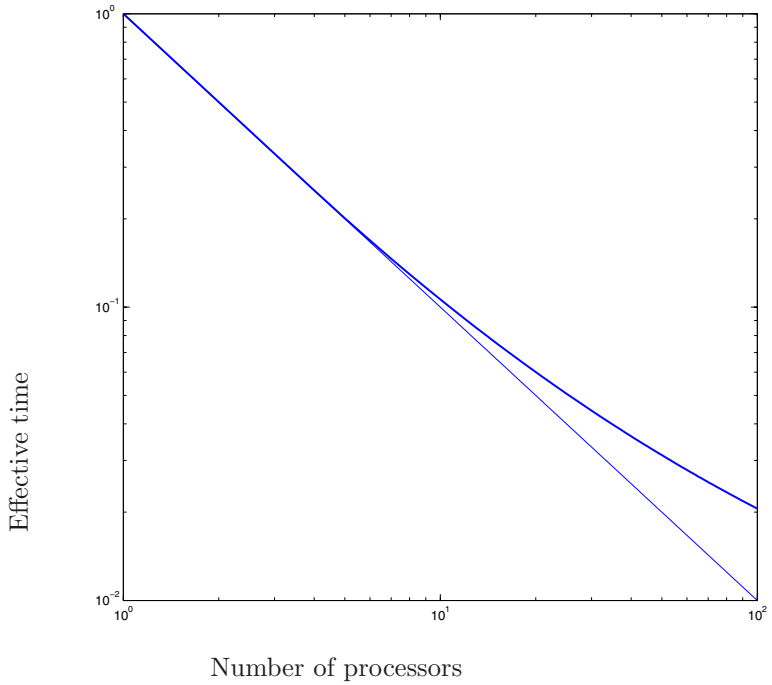


FIGURE 7. Effective time in function of the number of processors for $d = 3$ with a conjugate gradient method (dots), and $2L = 10^2$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

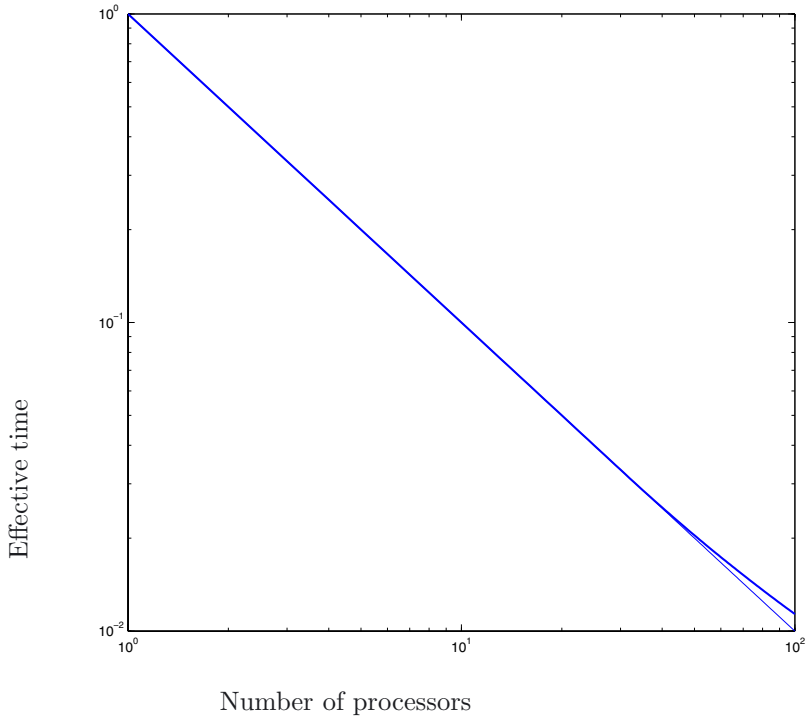


FIGURE 8. Effective time in function of the number of processors for $d = 3$ with a Cholesky method (dots), and $2L = 10^2$.

The mask has been chosen as follows:

$$\mu_L(x) = \tilde{\mu}_L(x_1)\tilde{\mu}_L(x_2),$$

where $x = (x_1, x_2) \in \mathbb{Z}^2$, and $\tilde{\mu}_L : \mathbb{Z} \to \mathbb{R}^+$ is given by

$$\tilde{\mu}_L(t) = \gamma_L \begin{cases} L \leq & |t| & : & 0, \\ \frac{L}{3} \leq & |t| \leq L & : & \frac{3}{2} - \frac{3}{2L}|t|, \\ & |t| \leq \frac{L}{3} & : & 1, \end{cases}$$

and γ_L is such that $\int_{\mathbb{Z}} \tilde{\mu}_L(t)\mathrm{d}t = 1$. The linear system (2.5) is solved by a preconditioned conjugate gradient method, whose preconditioner is the incomplete Cholesky factorization $IC(2)$ (see [18]). For a uniform sampling of $\log L$, $L \in [10, 2000]$, we approximate the expectation

$$\left\langle \left(\int_{\mathbb{Z}^d} (\mathbf{e}_1 + \nabla\phi_{T,R}(x)) \cdot A(x)(\mathbf{e}_1 + \nabla\phi_{T,R}(x))\mu_L(x)\mathrm{d}x - 3 \right)^2 \right\rangle$$

by an empirical average over $r(M)$ realizations (this number is chosen large enough so that the error between the empirical average and the expectation is negligible with respect to the approximation error $\left\langle |A^N_{T,R,L} - A_{\mathrm{hom}}|^2 \right\rangle^{1/2}$), and define the error by

$$\mathrm{Error}(M) := \sqrt{\frac{1}{r(M)} \sum_{j=1}^{r(M)} \left(\int_{\mathbb{Z}^d} (\mathbf{e}_1 + \nabla\phi^j_{T,R}(x)) \cdot A_j(x)(\mathbf{e}_1 + \nabla\phi^j_{T,R}(x))\mu_L(x)\mathrm{d}x - 3 \right)^2},$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

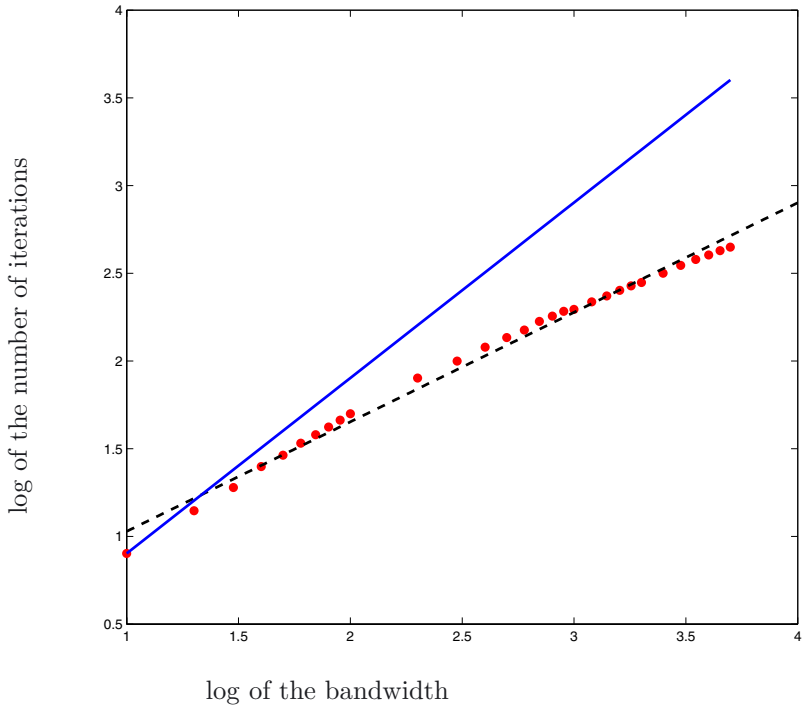


FIGURE 9. Number of iterations of the conjugate gradient method in function of the bandwidth of the matrix.

TABLE 1. Error (4.1) for different M, and $r(M)$ realizations.

M	1.6E+02	4.0E+02	9.0E+02	2.1E+03	5.6E+03	1.5E+04	4.0E+04
$r(M)$	10 000	10 000	10 000	10 000	5000	3000	1600
Error(M)	2.39E-01	1.48E-01	9.35E-02	5.97E-02	3.65E-02	2.28E-02	1.41E-02
M	1.1E+05	3.0E+05	8.1E+05	2.2E+06	6.0E+06	1.6E+07	
$r(M)$	1000	600	360	200	150	100	
Error(M)	9.11E-03	5.95E-03	3.95E-03	2.46E-03	1.55E-03	8.78E-04	

where $\{\phi^j_{T,R}\}$ are the solutions of (2.5) for the $r(M)$ different realizations A_j of the coefficients A. The number of realizations for each M considered and the associated error (4.1), are reported in Table 1.

The error is also plotted in function of M in logarithmic scale in Figure 12. The dots (which indicate calculations) are in very good agreement with the straight line of slope $-1/2$ corresponding to the decay provided by Theorem 2.10.

In order to determine N_{opt}, one needs to know the number of iterations of the conjugate gradient method (for the fixed tolerance 10^{-9}). The number of iterations is plotted in function of the bandwidth of the matrix in Figure 9 (in log-log). The dots represent the numerical experiments, the straight line represents a linear dependence as assumed in Section 4.1 (b times a matrix-vector multiplication which costs $O(l)$ operations), whereas the dashed line is a linear fitting of the numerical experiments (equation: $y(x) = 0.64\,x + 0.36$)). In the range of effective unknowns considered (M from 10^2 to 10^7), this implies that $N_{\text{opt}} = 1$. The reason for this is the efficiency of the preconditioner. To illustrate this fact, we have plotted in Figure 10 the ratio $\Gamma_{\text{CG}}(N, 2L)/\Gamma_{\text{CG}}(1, 2L)$ using this time the number of iterations of the conjugate gradient method obtained in the numerical tests for $N \in \{1, \ldots, 10\}$. As can be seen, $2L = 10^{3.5}$ (that is $M = 10^7$) is the critical number of unknowns under which $N_{\text{opt}} = 1$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

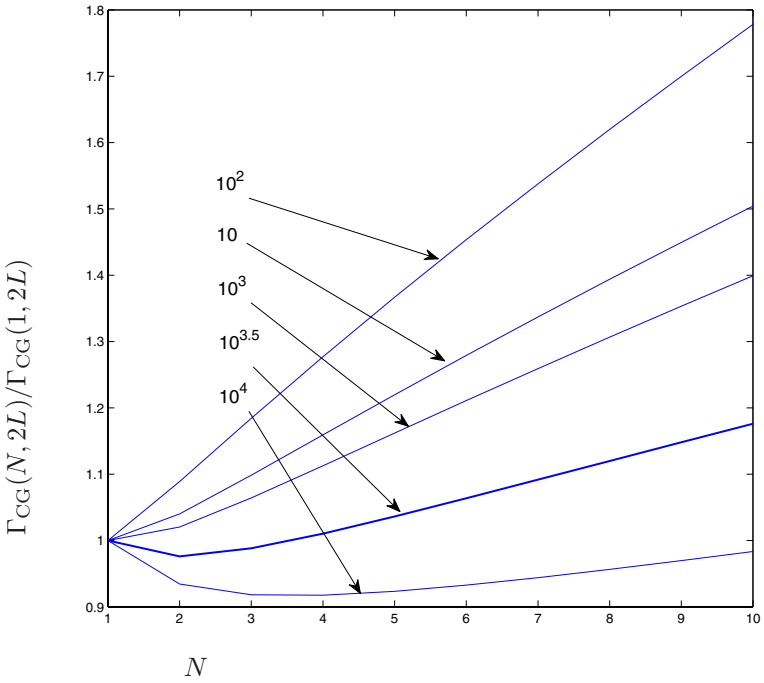


FIGURE 10. Complexity for $d = 2$ with a conjugate gradient method (number of iterations observed numerically), for $2L = 10, 10^2, 10^3, 10^{3.5}$.

We now turn to the interest of parallel computing and assume we have an arbitrary number of processors at our disposal. We distribute the number of unknowns over K^2 subdomains, $K \in \mathbb{N}$. We then set

$$\begin{cases} 2L_K &= \sqrt{M}/K, \\ T &= \sqrt{M} + 3, \\ R_K &= L_K + 1 + 0.1 \ln^2(\sqrt{M}) M^{1/4}, \end{cases}$$

where the number of subdomains is $N = K^2$, $K \in \{1, \dots, K_{\max}\}$, and $K_{\max}$ is given by

$$K_{\max} := \frac{\sqrt{M}}{0.1 M^{1/4} \ln^2(\sqrt{M})} = \frac{40 M^{1/4}}{\ln^2 M}. \tag{4.1}$$

The values of $K_{\max}$ are plotted in Figure 11 for M in $[4 \times 10^2, 3.8 \times 10^8]$. We have also gathered in Table 2 the description of the tests in function of the values of M: the sizes of the domain $R_{\max}$ and $L_{\max}$, the number $N_{\max} = K_{\max}^2$ of such domains, the number of realizations $r(M)$, and the associated error

$$\text{Error}(M) := \sqrt{\frac{1}{r(M)} \sum_{j=1}^{r(M)} \left(\frac{1}{N_{\max}} \sum_{k=1}^{N_{\max}} \int_{\mathbb{Z}^d} (\mathbf{e}_1 + \nabla \phi_{T,R_{\max}}^{j,k}(x)) \cdot A_{j,k}(x) (\mathbf{e}_1 + \nabla \phi_{T,R_{\max}}^{j,k}(x)) \mu_{L_{\max}}(x) \mathrm{d}x - 3 \right)^2}, \tag{4.2}$$

where the $\phi_{T,R_{\max}}^{j,k}$ are the solutions of (2.5) for the $N_{\max} \times r(M)$ different realizations $A_{j,k}$ of the coefficients A.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

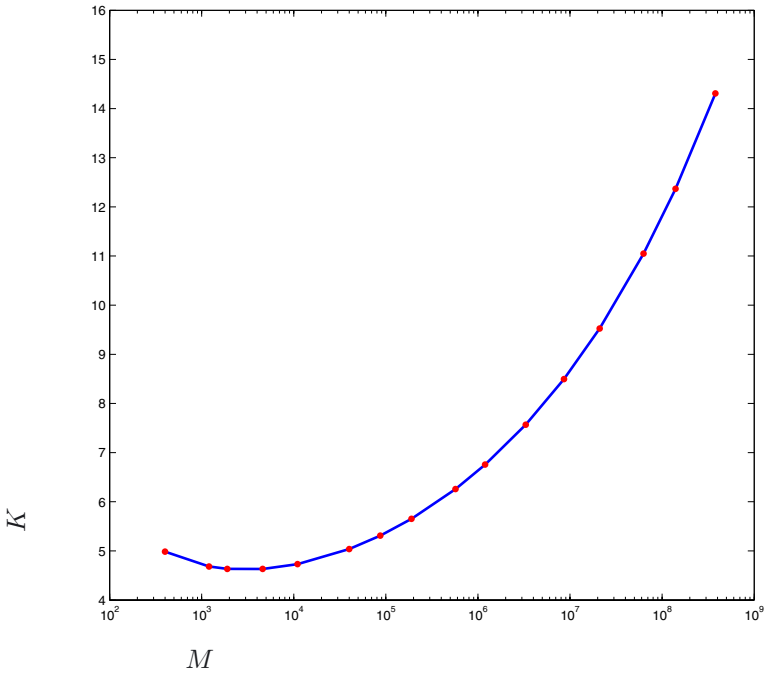


FIGURE 11. Number of subdomains K per dimension in function of the total number of unknowns M according to formula (4.1).

TABLE 2. Error (4.2) for different M, and $r(M)$ realizations.

M	4.0E+02	1.2E+03	1.9E+03	4.6E+03	1.1E+04	4.0E+04	8.7E+04	1.9E+05
$r(M)$	10 000	10 000	10 000	10 000	5000	3000	1600	1000
$N_{\max}$	5×5	5×5	4×4	4×4	4×4	5×5	5×5	5×5
$2L_{\max}$	4	7	11	17	26	40	59	87
$2R_{\max}$	12	21	33	51	78	120	177	261
Error(M)	4.83E-01	2.26E-01	2.02E-01	1.22E-01	8.45E-02	3.41E-02	2.31E-02	1.53E-02
M	5.7E+05	1.2E+06	3.3E+06	8.6E+06	2.1E+07	6.3E+07	1.4E+08	3.8E+08
$r(M)$	600	360	200	150	100	30	10	4
$N_{\max}$	6×6	6×6	7×7	8×8	9×9	11×11	12×12	14×14
$2L_{\max}$	126	181	258	366	514	720	1001	1358
$2R_{\max}$	378	543	774	1098	1542	2160	3003	4074
Error(M)	8.18E-03	6.13E-03	3.41E-03	2.06E-03	1.28E-03	8.57E-04	5.53E-04	1.69E-04

As expected, the convergence rate has the scaling of the central limit theorem $-1/2$, as can be seen in Figure 13, where the logarithm of the error is plotted in function of the logarithm of M.

To complete the comparison, we have plotted in Figure 14 the error in function of the computational time for $N = 1$ and $N = N_{\max}$ (using $N_{\max}$ processors). In particular, splitting the M effective number of sites into $N_{\max}$ subdomains is cheaper as soon as $M \geq 5 \times 10^3$. Note however that if the $N_{\max}$ problems are solved sequentially then the computational time is approximately 2.5 times larger than with $N = 1$, and the error is approximately 5 times larger (in other words, the prefactor in front of $M^{-1/2}$ is 5 times larger for $N = N_{\max}$ than for $N = 1$). Hence, splitting the way proposed here is effective provided parallel computing is used.

Concerning the issue of memory, it is worth noticing that the larger N the less memory needed. Hence, larger effective numbers of sites can be reached using larger N. In the present case, the computation for 3.8×10^8 effective sites could not have been done on one computer without taking $N = 196$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

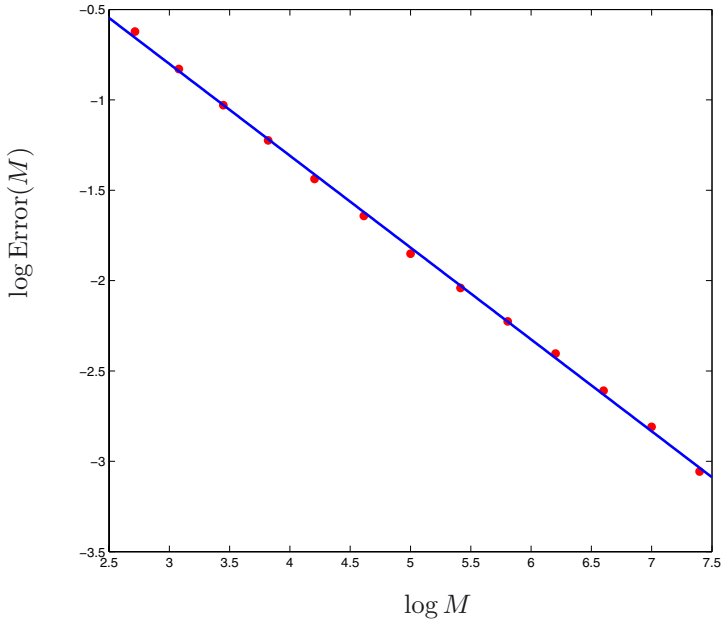


FIGURE 12. Error (4.1) in log scale in function of M.

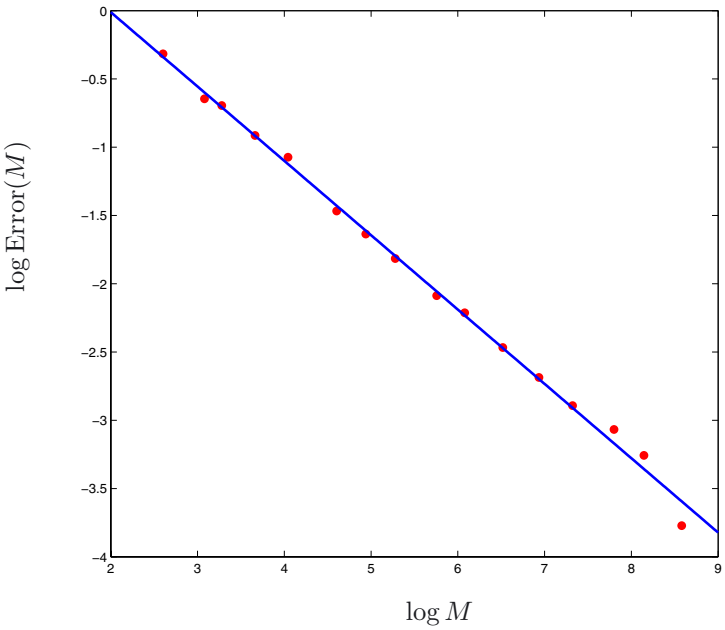


FIGURE 13. Error (4.2) in log scale in function of M.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

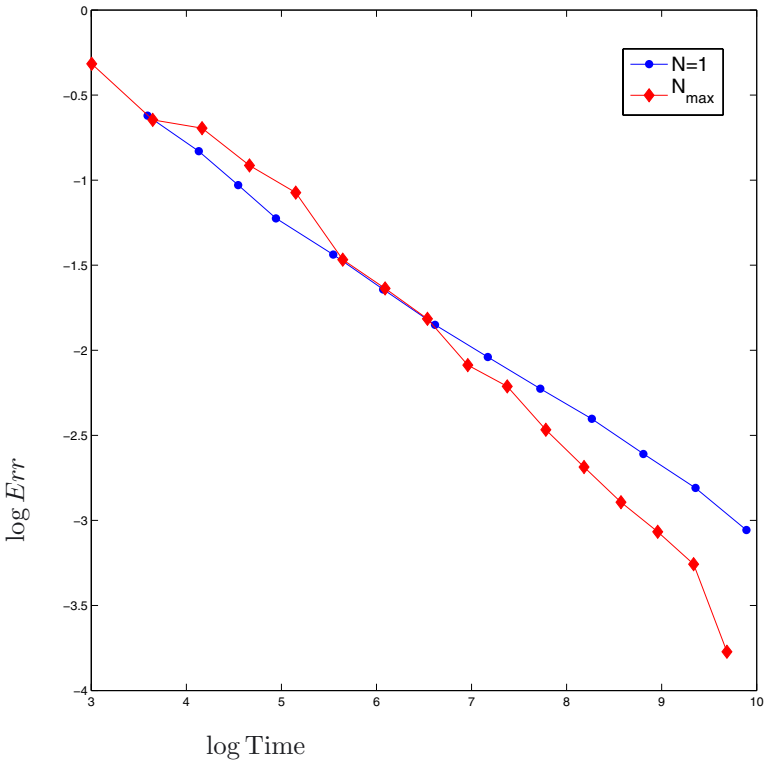


FIGURE 14. Errors (4.1) and (4.2) in function of computation time (log-log scale).

4.3. **Comparison with standard approaches and comments**

In order to approximate effective coefficients in stochastic homogenization, it is standard to replace the abstract corrector field by the solution to

$$\left\{ \begin{array}{rcl} -\nabla^* \cdot A(\xi + \nabla\phi_R) & = & 0 \quad \text{in } Q_R, \\ \phi_R & = & 0 \quad \text{on } \mathbb{Z}^d \setminus Q_R, \end{array} \right. \tag{4.3}$$

that is (2.5) without the zero-order term. This is typically the case in numerical homogenization methods applies to stationary stochastic problems (see for instance [8,12,22], in the continuous case). Let $M = (2R)^d$ be the effective number of sites. The formula for the approximation of the homogenized coefficients is then

$$\mathrm{Error}(M) := \sqrt{\frac{1}{r(M)} \sum_{j=1}^{r(M)} \left(\fint_{Q_R} (\mathbf{e}_1 + \nabla\phi_R^j(x)) \cdot A_j(x)(\mathbf{e}_1 + \nabla\phi_R^j(x)) \mathrm{d}x - 3 \right)^2}, \tag{4.4}$$

where $r(M)$ is the number of independent realizations. Let us make a *formal* error analysis of such an approach.

Assuming that the corrector field ϕ is uniformly bounded (which we do not know *a priori* since we only control $\langle |\phi|^q \rangle$ for all $q < \infty$ in [9], Prop. 1), the error we make by replacing ϕ by ϕ_R is due to the use of the Dirichlet boundary conditions, which are not exact. This error typically scales as a surface term in the energy, that is $R^{d-1}/R^d = 1/R$ in any dimension. In dimension 2, the effect due to the boundary conditions and the central limit theorem have the same scaling $R/R^2 = 1/R \sim M^{-1/2}$. Hence the addition of the zero order term may not be crucial in dimension 2 to obtain the optimal scaling $M^{-1/2}$ (although we are not able to turn this

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

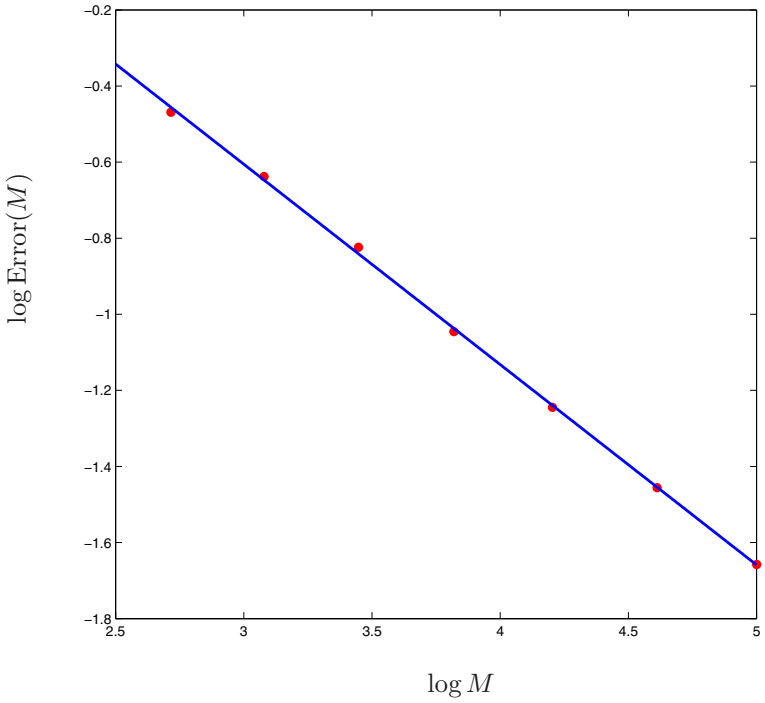


FIGURE 15. Error (4.4) in log scale.

into a rigorous argument). This is confirmed by numerical tests, as illustrated in Figure 15, where the proxy for ϕ is the solution to (4.3) (note that the prefactor is larger than in Fig. 12).

On the contrary, in dimension 3 (and more), the effect of the boundary conditions now scales as $1/R = M^{-1/3}$, whereas the central limit theorem scaling is still $M^{-1/2} \ll M^{-1/3}$. Hence the use of the zero-order term is crucial to observe the optimal scaling. Another case for which the zero-order term is crucial is when $N > 1$, even in dimension 2. If the proxy for ϕ on the domain Q_R is approximated independently on subdomains of size R^γ (with $1 \geq \gamma \geq 1/2$), the error due to the boundary conditions scales as $1/R^\gamma$ without the zero-order term, whereas it will remain of order $1/R$ with the zero-order term.

Let us now discuss the use of periodic boundary conditions. In the case of an i. i. d. conductivity function, we indeed expect the systematic $|\langle A_{L,\#}\rangle - A_{\rm hom}|$ to be of order $L^{-d/2}$ in any dimension (with a logarithmic correction in dimension $d = 2$), where $A_{L,\#}$ is defined in (1.4). Yet, the picture is much less clear when the coefficients display correlations. In particular, the use of the zero order term seems to be much more flexible in terms of applicability (it requires no knowledge on the structure of the correlations, as can be seen on the extreme case of periodic coefficients, [8]). Another practical advantage of the approach is the type of boundary conditions used in (2.5), that is homogeneous Dirichlet boundary conditions. Compared to periodic boundary conditions (which are "widely recognised" as less perturbative than Dirichlet boundary conditions for the cell problem without the zero-order term), the former has the advantage not to destroy the band structure of the stiffness matrix. Periodic boundary conditions change the profile of the matrix, making the triangle matrix in the Cholesky factorization less sparse (only the band structure is preserved by the algorithm) and the factorization more expensive. In addition, there is no optimal preconditioner for periodic boundary conditions. With this respect, the proposed strategy with the zero-order term and Dirichlet boundary conditions allows us to use efficient methods to solve the linear systems, without sacrificing the convergence rate.

As a conclusion, we have proposed and fully analyzed a numerical method to approximate effective coefficients for the stochastic homogenization of discrete elliptic equations. The analysis is sharp, and the numerical method

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

effective. It crucially relies on the introduction of a zero-order term in the corrector equation, which is not only essential for the analysis, but also for the numerical practice (at least in dimension $d \geq 3$).

Acknowledgements. The research of the author was partly supported by INRIA, under the grant "Action de Recherche Collaborative" DISCO. It is a pleasure to thank Felix Otto for stimulating discussions on the subject. Special thanks are due to the anonymous referees for their very insightful comments, which helped improving the quality of the manuscript.

A. Proof of Dykhne's formula in the discrete stochastic case

In periodic homogenization of elliptic differential operators in dimension two, Dykhne's formula is as follows. Let $A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ be a periodic function taking values in a subspace of uniformly bounded and elliptic matrices, and such that $A(x)A(R\cdot x) = \gamma \mathrm{Id}$, where R denotes the rotation by $\pi/2$ in $\mathbb{R}^2$ and $\gamma > 0$. Then the homogenized matrix associated with A is $A_{\rm hom} = \sqrt{\gamma}\,\mathrm{Id}$. This formula is a particular case of general duality relations. Its proof (see for instance [13], Sect. 1.5) makes use of the following two facts:

(i) if $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ is a gradient field, then $x \mapsto \chi(R\cdot x)$ is divergence-free;
(ii) if $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ is divergence-free, then $x \mapsto \chi(R\cdot x)$ is a gradient field.

In particular, let ϕ be periodic and satisfy the corrector equation $\nabla \cdot A(x)\nabla\phi(x) = 0$. We further set $\xi = \langle \nabla\phi \rangle$ (where $\langle\cdot\rangle$ denotes the average in this periodic case). Since $\nabla\phi$ is a gradient, (i) implies that

$$\nabla \cdot A(x)A(R\cdot x)\nabla\phi(R\cdot x) = \gamma \nabla\cdot\nabla\phi(R\cdot x) = 0. \tag{A.1}$$

On the other hand $\nabla \cdot A(x)\nabla\phi(x) = 0$ implies by (ii) that $x \mapsto A(R\cdot x)\nabla\phi(R\cdot x)$ is a gradient field, so that, by definition of the homogenized coefficients, (A.1) yields

$$\langle A(x)A(R\cdot x)\nabla\phi(R\cdot x)\rangle \;=\; A_{\rm hom}\,\langle A(R\cdot x)\nabla\phi(R\cdot x)\rangle\,.$$

Using now that the average does not change by rotation, this turns into

$$\gamma\xi \;=\; \langle A(x)A(R\cdot x)\nabla\phi(R\cdot x)\rangle \;=\; A_{\rm hom}\,\langle A\nabla\phi\rangle \;=\; A_{\rm hom}^2\xi$$

from which we deduce the claim by symmetry and coercivity of $A_{\rm hom}$.

In the two-dimensional discrete case, the following counterparts to (i) and (ii) hold:

(i′) if $\chi : \mathbb{Z}^2 \to \mathbb{R}^2$ is a gradient field, then $x \mapsto \chi(R\cdot x)$ is divergence free, that is $\nabla\cdot\chi \equiv 0$;
(ii′) if $\chi : \mathbb{Z}^2 \to \mathbb{R}^2$ is divergence-free, then $x \mapsto \chi(R\cdot x)$ is a gradient field $\nabla\phi$.

However, if one considers ϕ such that $\nabla^*\cdot A(x)\nabla\phi(x) \equiv 0$, one has

$$\nabla^*\cdot A(x)A(R\cdot x)\nabla\phi(R\cdot x) = \gamma\nabla^*\cdot\nabla\phi(R\cdot x) \not\equiv \gamma\nabla\cdot\nabla\phi(R\cdot x) \equiv 0$$

in general. Similarly, $x \mapsto A(R\cdot x)\nabla\phi(R\cdot x)$ is not a gradient field either. Hence, the arguments of the proof do not carry out to difference operators in the periodic case. Actually, counterexamples to Dykhne's formula are easily constructed in this case (see for instance the discrete example in [8], Sect. 4.1 for which $\gamma = 10 < 26.240099009901\ldots = a_{\rm hom}$).

As opposed to the periodic case, the proof of Dykhne's formula carries out from the stochastic continuous case to the stochastic discrete case, as we quickly show now. Let $d = 2$ and $A \in \mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity matrix whose entries, denoted by $x \mapsto a_1(x)$ and $x \mapsto a_2(x)$, take values $\alpha > 0$ with probability $1/2$ and $\beta > 0$ with probability $1/2$. We then introduce the following two auxiliary conductivity matrices: $\bar{A} := \alpha\beta A^{-1}$, and $\tilde{A}$ defined by $\tilde{A} : x \mapsto \alpha\beta\,\mathrm{diag}\left[a_2(x+\mathbf{e}_2)^{-1}, a_1(x+\mathbf{e}_1)^{-1}\right]$. Note that $A, \bar{A}$ and $\tilde{A}$ have the same law, so that they yield the same homogenized matrix $A_{\rm hom}$. Let then $\xi = \xi_1\mathbf{e}_1 + \xi_2\mathbf{e}_2 \in \mathbb{R}^2$ and $\tilde{\phi}$ be the corrector associated

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

with $\tilde{A}$ and ξ. We introduce the vector field

$$W : x \mapsto \bar{A}(x) \left[\begin{array}{c} \xi_2 + \nabla_2^* \tilde{\phi}(x) \\ -\xi_1 - \nabla_1^* \tilde{\phi}(x) \end{array} \right].$$

The field W satisfies the following three properties.

Property 1.

$$\nabla^* \cdot A(x) W(x) \; = \; 0. \tag{A.2}$$

A direct computation actually shows

$$\begin{aligned} \nabla^* \cdot A(x) W(x) &= \nabla^* \cdot A(x) \bar{A}(x) \left[\begin{array}{c} \xi_2 + \nabla_2^* \tilde{\phi}(x) \\ -\xi_1 - \nabla_1^* \tilde{\phi}(x) \end{array} \right] \\ &= \alpha\beta \nabla^* \cdot \left[\begin{array}{c} \nabla_2^* \tilde{\phi}(x) \\ -\nabla_1^* \tilde{\phi}(x) \end{array} \right] \\ &\equiv 0. \end{aligned}$$

Property 2.

$$\nabla \times W(x) \; = \; 0. \tag{A.3}$$

Using the defining equation for the corrector, and the definitions of $\bar{A}$ and $\tilde{A}$, one has

$$\begin{aligned} \nabla \times W(x) &= \nabla \times \bar{A}(x) \left[\begin{array}{c} \xi_2 + \nabla_2^* \tilde{\phi}(x) \\ -\xi_1 - \nabla_1^* \tilde{\phi}(x) \end{array} \right] \\ &= \nabla_2 (\bar{a}_1(x)(\xi_2 + \nabla_2^* \tilde{\phi}(x))) + \nabla_1 (\bar{a}_2(x)(\xi_1 + \nabla_1^* \tilde{\phi}(x))) \\ &= \nabla_2 (\tilde{a}_2(x - \mathbf{e}_2)(\xi_2 + \nabla_2^* \tilde{\phi}(x))) + \nabla_1 (\tilde{a}_1(x - \mathbf{e}_1)(\xi_1 + \nabla_1^* \tilde{\phi}(x))) \\ &= \nabla_2^* (\tilde{a}_2(x)(\xi_2 + \nabla_2 \tilde{\phi}(x))) + \nabla_1^* (\tilde{a}_1(x)(\xi_1 + \nabla_1 \tilde{\phi}(x))) \\ &= \nabla^* \cdot \tilde{A}(x)(\xi + \tilde{\phi}) \\ &\equiv 0. \end{aligned}$$

From (A.3) we deduce that $x \mapsto W(x)$ is a gradient field.

Property 3.

$$\langle W(x) \rangle \; = \; A_{\rm hom}(\xi_2 \mathbf{e}_1 - \xi_1 \mathbf{e}_2). \tag{A.4}$$

To prove this property, we use the definitions of $\bar{A}$ and $\tilde{A}$, and the joint stationarity of $\tilde{A}$ and $\nabla\tilde{\phi}$.

$$\begin{aligned} \langle W(x) \rangle &= \left\langle \bar{A}(x) \left[\begin{array}{c} \xi_2 + \nabla_2^* \tilde{\phi}(x) \\ -\xi_1 - \nabla_1^* \tilde{\phi}(x) \end{array} \right] \right\rangle \\ &= \left\langle \tilde{a}_2(x - \mathbf{e}_2)(\xi_2 + \nabla_2^* \tilde{\phi}(x)) \mathbf{e}_1 - \tilde{a}_1(x - \mathbf{e}_1)(\xi_1 + \nabla_1^* \tilde{\phi}(x)) \mathbf{e}_2 \right\rangle \\ &= \left\langle \tilde{a}_2(x)(\xi_2 + \nabla_2 \tilde{\phi}(x)) \mathbf{e}_1 - \tilde{a}_1(x)(\xi_1 + \nabla_1 \tilde{\phi}(x)) \mathbf{e}_2 \right\rangle \\ &= A_{\rm hom}(\xi_2 \mathbf{e}_1 - \xi_1 \mathbf{e}_2). \end{aligned}$$

In the last equality, we have used that $\tilde{a}_1$ and $\tilde{a}_2$ have the same law.

We are now in position to conclude. Since $x \mapsto W(x)$ is a gradient field and satisfies (A.2), one has by definition of the homogenized matrix

$$A_{\rm hom} \langle W \rangle \; = \; \langle AW \rangle . \tag{A.5}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

We use (A.4) to rewrite the l. h. s. of (A.5) as

$$A_{\rm hom}\langle W\rangle \;=\; A_{\rm hom}^2(\xi_2\mathbf{e}_1-\xi_1\mathbf{e}_2), \tag{A.6}$$

and the definition of $\bar{A}$ to rewrite the r. h. s. of (A.5) as

$$\langle AW\rangle \;=\; \alpha\beta\left\langle\left[\begin{array}{c}\xi_2+\nabla_2^*\tilde{\phi}\\ -\xi_1-\nabla_1^*\tilde{\phi}\end{array}\right]\right\rangle = \alpha\beta(\xi_2\mathbf{e}_1-\xi_1\mathbf{e}_2), \tag{A.7}$$

by stationarity of $\tilde{\phi}$. The combination of (A.5), (A.6) & (A.7) proves that

$$A_{\rm hom}^2 \;=\; \alpha\beta Id,$$

from which we deduce Dykhne's formula

$$A_{\rm hom} \;=\; \sqrt{\alpha\beta}\,Id$$

since $A_{\rm hom}$ is symmetric positive definite.

B. Proof of Lemma 3.2

The proof of Lemma 3.2 is a discrete version of the corresponding proof in the continuous case in [8], Appendix. The discreteness compels us to slightly modify the definitions of $g_{T,j,k}$ and $\chi_{T,j}$ (see below). Likewise, we have to use a discrete Leibniz' rule, which complexifies notation. As for the continuous case, we combine the decay estimates of [9], Lemma 8, with Harnack's inequality and Agmon's positivity method (see [1]). We recall here the Harnack inequality on graphs due to Delmotte (see [5], Prop. 5.3).

Lemma B.1 (Harnack's inequality)**.** *Let $a\in\mathcal{A}_{\alpha\beta}$ and $R\gg 1$. If $g:\mathbb{Z}^d\to\mathbb{R}^+$ satisfies*

$$-\nabla^*\cdot A\nabla g(x) \;\le\; 0 \tag{B.1}$$

in the annulus $\{R/2<|x|\le 4R\}$ (that is g_T is a nonnegative subsolution), then

$$\sup_{R<|x|\le 2R} g(x) \;\lesssim\; \left(R^{-d}\int_{R/2<|x|\le 4R} g(x)^2{\rm d}x\right)^{1/2}, \tag{B.2}$$

where the multiplicative constant does not depend on R, nor on A.

For $|x-y|\le\sqrt{T}$, (3.2) and (3.2) coincide with [10], Lemma 5, (2.10) and (2.11), and we only treat the case $|x-y|\ge\sqrt{T}$.

Step 1. Operator positivity method. We first prove the exponential decay using the classical Leibniz rule, for which the algebra is simpler. We shall deal with the modifications due to the discrete Leibniz rule in the last step of this proof.

For $b>0$, $j,k\in\mathbb{N}$, let $\tilde{\chi}_{T,j}:\mathbb{Z}\to[0,1]$ and $\tilde{g}_{T,j,k}:\mathbb{Z}\to\mathbb{R}^+$ be given by

$$\tilde{\chi}_{T,j}(x) \;=\; \left\{\begin{array}{lcl}\text{for } |x|\le 2^j\sqrt{T} & : & 0\\ \text{for } 2^j\sqrt{T}\le|x|\le 2^{j+1}\sqrt{T} & : & (2^j\sqrt{T})^{-1}(|x|-2^j\sqrt{T})\\ \text{for } |x|\ge 2^{j+1}\sqrt{T} & : & 1,\end{array}\right.$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

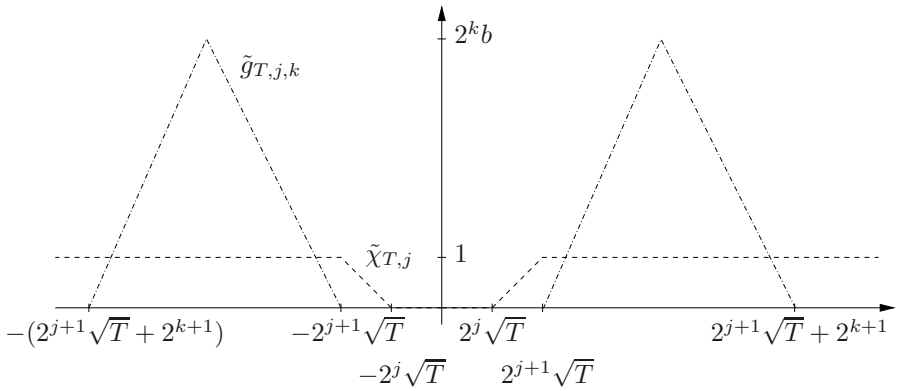


FIGURE 16. Functions $\tilde{g}_{T,j,k}$ and $\tilde{\chi}_{T,j}$.

and

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

These functions are plotted for convenience in Figure 16. We then set

$$\begin{aligned} g_{T,j,k}: \quad \mathbb{Z}^d &\to \mathbb{R}^+ \\ x &\mapsto \sum_{i=1}^{d} \tilde{g}_{T,j,k}(x_i), \end{aligned}$$

and

$$\begin{aligned} \chi_{T,j}: \quad \mathbb{Z}^d &\to [0,1] \\ x &\mapsto 1-\prod_{i=1}^{d}(1-\tilde{\chi}_{T,j}(x_i)). \end{aligned}$$

Note that $|\nabla g_{T,j,k}(x)| \leq \sqrt{d}b$ for all $x \in \mathbb{Z}^d$, $j \in \{1,\dots,d\}$, $k \in \mathbb{N}$. With the notation $|x|_\infty := \max\{x_1,\dots,x_d\}$, $\chi_{T,j}$ satisfies: $\chi_{T,j}{}_{||x_\infty| \leq 2^j\sqrt{T}} \equiv 0$ and $\chi_{T,j}{}_{||x_\infty| \geq 2^{j+1}\sqrt{T}} \equiv 1$.

Let $s \in \mathbb{N} \setminus \{0\}$. We multiply the defining equation for G_T by the test function

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

Since this function is in $L^2(\mathbb{Z}^d)$ due to [9], Lemma 9, (2.23), we may integrate by parts the equation on $\mathbb{Z}^d$, obtaining

$$\begin{aligned} T^{-1}\int_{\mathbb{Z}^d} \Big(\chi_{T,j}(x)\exp(g_{T,j,k}(x))G_T(x)\Big)^2 \mathrm{d}x \\ + \int_{\mathbb{Z}^d} \nabla\Big(\chi_{T,j}(x)^2\exp(2g_{T,j,k}(x))G_T(x)\Big)\cdot A(x)\nabla G_T(x)\mathrm{d}x \;=\; 0. \end{aligned} \tag{B.3}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

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

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

We rewrite the second term of th r. h. s. as follows:

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

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

$$\begin{aligned}
&\nabla\Big(\chi^2\exp(2g)G\Big)\cdot A\nabla G\\
&\quad=\quad \nabla\Big(\chi\exp(g)G\Big)\cdot A\nabla\Big(\chi\exp(g)G\Big)-G^2\nabla\Big(\chi\exp(g)\Big)\cdot A\nabla\Big(\chi\exp(g)\Big)\\
&\quad\geq\quad -\beta G^2\left|\nabla\Big(\chi\exp(g)\Big)\right|^2 \qquad\qquad (B.4)
\end{aligned}$$

by the uniform bounds on A.

Setting $\psi_{T,j,k}:x\mapsto\exp(g_{T,j,k}(x))G_T(x)$, we insert (B.4) into (B.3) to get

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

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

$$\begin{aligned}
&\int_{|x|_\infty\geq 2^j\sqrt{T}}\Big(T^{-1}\psi_{T,j,k}(x)^2-\beta G_T(x)^2\left|\nabla\exp(g_{T,j,k}(x))\right|^2\Big)\chi_{T,j}(x)^2\mathrm{d}x\\
&\quad\leq\quad \int_{2^j\sqrt{T}\leq|x|_\infty<2^{j+1}\sqrt{T}}\beta\left|\nabla\chi_{T,j}(x)\right|^2G_T(x)^2\mathrm{d}x,
\end{aligned}$$

and finally

$$\begin{aligned}
&\int_{|x|_\infty\geq 2^j\sqrt{T}}\Big(T^{-1}-\beta\left|\nabla g_{T,j,k}(x)\right|^2\Big)\psi_{T,j,k}(x)^2\chi_{T,j}(x)^2\mathrm{d}x\\
&\quad\leq\quad \beta T^{-1}2^{-2j}\int_{2^j\sqrt{T}\leq|x|_\infty<2^{j+1}\sqrt{T}}G_T(x)^2\mathrm{d}x.
\end{aligned}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

Choosing $b = (2d\beta T)^{-1/2}$ and using the definition of $\chi_{T,j}$, this turns into

$$\begin{aligned}\int_{|x|_\infty \geq 2^{j+1}\sqrt{T}} \psi_{T,j,k}(x)^2 \mathrm{d}x &\leq \int_{|x|_\infty \geq 2^{j}\sqrt{T}} \psi_{T,j,k}(x)^2 \chi_{T,j}(x)^2 \mathrm{d}x \\ &\lesssim 2^{-2j} \int_{2^j\sqrt{T}\leq |x|_\infty < 2^{j+1}\sqrt{T}} G_T(x)^2 \mathrm{d}x.\end{aligned}$$

We then pass to the limit $k \to \infty$ by the monotone convergence theorem and use the definition of $\psi_{T,j,k}$ for $|x|_\infty \geq 2^{j+2}\sqrt{T}$ to prove the main result of this step:

$$\int_{|x|_\infty \geq 2^{j+2}\sqrt{T}} G_T(x)^2 \mathrm{d}x \lesssim 2^{-2j} \exp(-b2^{j+1}\sqrt{T}) \int_{2^j\sqrt{T}\leq |x|_\infty < 2^{j+1}\sqrt{T}} G_T(x)^2 \mathrm{d}x \tag{B.5}$$

for all $s \in \mathbb{N} \setminus \{0\}$.

Step 2. Decay estimate and Harnack inequality. We now use the decay estimates of [9], Lemma 8, (2.23), with $q = 2$: For all $j \in \mathbb{N}$ and $d \geq 2$, we have (recall that $|\cdot|$ and $|\cdot|_\infty$ are equivalent norms)

$$\int_{2^j\sqrt{T}\leq |x|_\infty < 2^{j+1}\sqrt{T}} G_T(x)^2 \mathrm{d}x \lesssim (2^j\sqrt{T})^{d+(2-d)2}. \tag{B.6}$$

Since

$$-\nabla^* \cdot A\nabla G_T(x) = -TG_T(x) \leq 0$$

for $|x| \gg 1$, Lemma B.1 implies

$$\begin{aligned}&\sup_{2^{j+3}\sqrt{T}\leq |x|_\infty \leq 2^{j+4}\sqrt{T}} G_T(x) \\ &\lesssim \left((2^j\sqrt{T})^{-d} \int_{2^{j+2}\sqrt{T}\leq |x|_\infty \leq 2^{j+5}\sqrt{T}} G_T(x)^2 \mathrm{d}x\right)^{1/2} \\ &\overset{(B.5)}{\lesssim} \left((2^j\sqrt{T})^{-d} 2^{-2j} \exp(-b2^{j+1}\sqrt{T}) \int_{2^j\sqrt{T}\leq |x|_\infty < 2^{j+1}\sqrt{T}} G_T(x)^2 \mathrm{d}x\right)^{1/2} \\ &\overset{(B.6)}{\lesssim} \left((2^j\sqrt{T})^{-d} 2^{-2j} \exp(-b2^{j+1}\sqrt{T}) (2^j\sqrt{T})^{d+(2-d)2}\right)^{1/2} \\ &\leq (2^j\sqrt{T})^{2-d} \exp(-b2^j\sqrt{T}),\end{aligned}$$

which is the claim (3.2) for $d > 2$ and (3.2) for $d = 2$.

Step 3. Modifications due to the discreteness. In Step 1, we have used the standard Leibniz rule. In this step, we complete the proof by turning to the discrete Leibniz rule, and prove the following inequality corresponding to (B.4):

$$\begin{aligned}&\nabla(\chi_{T,j}^2 \exp(2g_{T,j,k})\, G_T(x)) \cdot A\nabla G_T(x) \\ &\quad \geq -4\beta(G_T(x+\mathbf{e}_i)^2 + G_T(x)^2) \left|\nabla\Big(\chi_{T,j}(x)^2 \exp(2g_{T,j,k}(x))\Big)\right|^2.\end{aligned} \tag{B.7}$$

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

This inequality is a consequence of the following version of [9], Proof of Lemma 8, Step 5, (4.27)

$$\begin{aligned}\nabla(\eta^2\, G_T^{q-1})\cdot A\nabla G_T(x)\\ \geq\ (1-2C^{-1})\sum_{i=1}^d a_i(x)\frac{\eta^2(x+\mathbf{e}_i)+\eta^2(x)}{2}\underbrace{(G_T^{q-1}(x+\mathbf{e}_i)-G_T^{q-1}(x))\nabla_i G_T(x)}_{\geq 0}\\ -C\sum_{i=1}^d a_i(x)(G_T(x+\mathbf{e}_i)^q+G_T(x)^q)|\nabla_i\eta(x)|^2,\end{aligned}$$

with the values $q=2$, $C=4$, $\eta(x)=\chi_{T,j}(x)\exp(g_{T,j,k}(x))$, and using that $a_i\leq\beta$. To prove this version of [9], (4.26), one just needs to keep track of a_i and not use the lower and upper bounds on a_i in [9], Proof of Lemma 9, Step 5. Setting $\psi_{T,j,k}: x\mapsto\exp(g_{T,j,k}(x))G_T(x)^{2s}$, we insert (B.7) into (B.3) to get

$$\begin{aligned}\int_{\mathbb{Z}^d}\Big(T^{-1}\chi_{T,j}(x)^2\psi_{T,j,k}(x)^2-4\beta\sum_{i=1}^d(G_T(x+\mathbf{e}_i)^2+G_T(x)^2)\\ \big(\nabla_i\big(\chi_{T,j}(x)\exp(g_{T,j,k}(x))\big)\big)^2\Big)\mathrm{d}x\quad\leq\quad 0.\end{aligned}$$

By the properties of $\chi_{T,j}$ and $g_{T,j,k}$, this turns into

$$\begin{aligned}\int_{|x|_\infty\geq 2^j\sqrt{T}}\Big(T^{-1}\psi_{T,j,k}(x)^2-4\beta\sum_{i=1}^d(G_T(x+\mathbf{e}_i)^2+G_T(x)^2)\\ \times(\nabla_i\exp(g_{T,j,k}(x)))^2\Big)\chi_{T,j}(x)^2\mathrm{d}x\\ \leq\quad\int_{2^j\sqrt{T}\leq|x|_\infty\leq 2^{j+1}\sqrt{T}}4\beta(1+\exp(2b))\sum_{i=1}^d(\nabla_i\chi_{T,j}(x))^2(G_T(x+\mathbf{e}_i)^2+G_T(x)^2)\mathrm{d}x.\end{aligned}\tag{B.8}$$

We then use the specific structures of $g_{T,j,k}$: For all $x_i\in[-2^{j+1}\sqrt{T}-2^{k+1},-2^{j+1}\sqrt{T}-2^k)\cup[2^{j+1}\sqrt{T},2^{j+1}\sqrt{T}+2^k)$, one has $\tilde g_{T,j,k}(x_i+1)=\tilde g_{T,j,k}(x_i)+b$, and

$$\begin{aligned}\nabla_i\exp(g_{T,j,k}(x))&=\exp\Big(\sum_{l\neq i}\tilde g_{T,j,k}(x_l)\Big)\Big(\exp(\tilde g_{T,j,k}(x_i+1))-\exp(\tilde g_{T,j,k}(x_i))\Big)\\ &=\exp\Big(\sum_{l\neq i}\tilde g_{T,j,k}(x_l)\Big)\exp(\tilde g_{T,j,k}(x_i))(\exp(b)-1)\\ &=\exp\Big(\sum_{l\neq i}\tilde g_{T,j,k}(x_l)\Big)\exp(\tilde g_{T,j,k}(x_i+1))(1-\exp(-b)),\end{aligned}$$

so that

$$\Big(\nabla_i\exp(g_{T,j,k}(x))\Big)^2=\exp(g_{T,j,k}(x))^2(\exp(b)-1)^2\tag{B.9}$$

$$=\exp(g_{T,j,k}(x+\mathbf{e}_i))^2(1-\exp(-b))^2\tag{B.10}$$

for $x_i\in[-2^{j+1}\sqrt{T}-2^{k+1},-2^{j+1}\sqrt{T}-2^k)\cup[2^{j+1}\sqrt{T},2^{j+1}\sqrt{T}+2^k)$.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

For $x_i \in [-2^{j+1}\sqrt{T}-2^k, -2^{j+1}\sqrt{T}) \cup [2^{j+1}\sqrt{T}+2^k, 2^{j+1}\sqrt{T}+2^{k+1})$, the same chain of arguments also yields (B.9) and (B.10).

Finally, for

$$\begin{array}{l} x_i \geq 2^{j+1}\sqrt{T}+2^{k+1}, \\ x_i < -(2^{j+1}\sqrt{T}+2^{k+1}), \\ -2^{j+1}\sqrt{T} \leq x_i < 2^{j+1}\sqrt{T}, \end{array}$$

one has $\Big(\nabla_i \exp(g_{T,j,k}(x))\Big)^2 = 0$.

We then deduce the following bound for all $|x_i| \geq 2^j\sqrt{T}$:

$$\begin{aligned} &\Big(\nabla_i \exp(g_{T,j,k}(x))\Big)^2 (G_T(x+\mathbf{e}_i)^2 + G_T(x)^2) \\ &\quad \overset{\text{(B.9)\&(B.10)}}{=} \psi_{T,j,k}(x+\mathbf{e}_i)^2(1-\exp(-b))^2 + \psi_{T,j,k}(x)^2(\exp(b)-1)^2 \\ &\quad \lesssim \Big(\psi_{T,j,k}(x+\mathbf{e}_i)^2 + \psi_{T,j,k}(x)^2\Big) b^2, \end{aligned} \tag{B.11}$$

for all $0 < b \leq 1$. Inserting (B.11) into (B.8) yields

$$\int_{|x|_\infty \geq 2^{j+1}\sqrt{T}} (T^{-1} - 4\beta d b^2)\psi_{T,j,k}(x)^2 \mathrm{d}x \lesssim \int_{2^j\sqrt{T} \leq |x|_\infty \leq 2^{j+1}\sqrt{T}+1} G_T(x)^2 \mathrm{d}x,$$

and we may conclude as before, taking this time $b = (8\beta d T)^{-1/2}$.

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* **29**. Princeton University Press, Princeton, NJ (1982).

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

[3] A. Bourgeat and A. Piatnitski, Approximations of effective coefficients in stochastic homogenization. *Ann. Inst. H. Poincaré* **40** (2004) 153–165.

[4] P. Caputo and D. Ioffe, Finite volume approximation of the effective diffusion matrix: the case of independent bond disorder. *Ann. Inst. H. Poincaré Probab. Statist.* **39** (2003) 505–525.

[5] T. Delmotte, Inégalité de Harnack elliptique sur les graphes. *Colloq. Math.* **72** (1997) 19–37.

[6] A. Dykhne, Conductivity of a two-dimensional two-phase system. *Sov. Phys. JETP* **32** (1971) 63–65. Russian version: *Zh. Eksp. Teor. Fiz.* **59** (1970) 110–5.

[7] 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.

[8] A. Gloria, Reduction of the resonance error – Part 1: Approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.*, to appear.

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

[10] A. Gloria and F. Otto, An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, to appear.

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

[12] 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.

[13] V.V. Jikov, S.M. Kozlov and O.A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, Berlin (1994).

[14] 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. *Int. J. Sol. Struct.* **40** (2003) 3647–3679.

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

[16] S.M. Kozlov, Averaging of difference schemes. *Mat. Sb.* **57** (1987) 351–369.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press

[17] R. Künnemann, The diffusion limit for reversible jump processes on $\mathbb{Z}^d$ with ergodic random bond conductivities. *Commun. Math. Phys.* **90** (1983) 27–68.
[18] J.A. Meijerink and H.A. van der Vorst, An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Math. Comp.* **31** (1977) 148–162.
[19] A. Naddaf and T. Spencer, Estimates on the variance of some homogenization problems. Preprint (1998).
[20] H. Owhadi, Approximation of the effective conductivity of ergodic media by periodization. *Probab. Theory Relat. Fields* **125** (2003) 225–258.
[21] G.C. Papanicolaou and S.R.S. Varadhan, Boundary value problems with rapidly oscillating random coefficients, in *Random fields I, II (Esztergom, 1979)*, *Colloq. Math. Soc. János Bolyai* **27**. North-Holland, Amsterdam (1981) 835–873.
[22] 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.
[23] V.V. Yurinskii, Averaging of symmetric diffusion in random medium. *Sibirskii Matematicheskii Zhurnal* **27** (1986) 167–180.

https://doi.org/10.1051/m2an/2011018 Published online by Cambridge University Press