Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

An optimal error estimate in stochastic homogenization of discrete elliptic equations

by

Antoine Gloria, and Felix Otto

Preprint no.: 28 2010

AN OPTIMAL ERROR ESTIMATE IN STOCHASTIC HOMOGENIZATION OF DISCRETE ELLIPTIC EQUATIONS

ANTOINE GLORIA & FELIX OTTO

Abstract. This paper is the companion article of [5]. We consider a discrete elliptic equation on the d-dimensional lattice $\mathbb{Z}^d$ with random coefficients A of the simplest type: They are identically distributed and independent from edge to edge. On scales large w. r. t. the lattice spacing (i. e. unity), the solution operator is known to behave like the solution operator of a (continuous) elliptic equation with constant deterministic coefficients. This symmetric "homogenized" matrix $A_{\rm hom} = a_{\rm hom}\mathrm{Id}$ is characterized by $\xi \cdot A_{\rm hom}\xi = \langle(\xi + \nabla\phi) \cdot A(\xi + \nabla\phi)\rangle$ for any direction $\xi \in \mathbb{R}^d$, where the random field ϕ (the "corrector") is the unique solution of $-\nabla^* \cdot A(\xi + \nabla\phi) = 0$ in $\mathbb{Z}^d$ such that $\phi(0) = 0$, $\nabla\phi$ is stationary and $\langle\nabla\phi\rangle = 0$, $\langle\cdot\rangle$ denoting the ensemble average (or expectation).

In order to approximate the homogenized coefficients $A_{\rm hom}$, the corrector problem is usually solved in a box $Q_L = [-L, L)^d$ of size $2L$ with periodic boundary conditions, and the space averaged energy on Q_L defines an approximation A_L of $A_{\rm hom}$. Although the statistics is modified (independence is replaced by periodic correlations) and the ensemble average is replaced by a space average, the approximation A_L converges almost surely to $A_{\rm hom}$ as $L \uparrow \infty$. In this paper, we give estimates on both errors. To be more precise, we do not consider periodic boundary conditions on a box of size $2L$, but replace the elliptic operator by $T^{-1} - \nabla^* \cdot A\nabla$ with (typically) $T \sim L^2$, as standard in the homogenization literature. We then replace the ensemble average by a space average on Q_L, and estimate the overall error on the homogenized coefficients in terms of L and T.

Keywords: stochastic homogenization, effective coefficients, difference operator.

2000 Mathematics Subject Classification: 35B27, 39A70, 60H25, 60F99.

1. Introduction

1.1. **Motivation.** In this article, we continue the analysis we began in [5] 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

$$-\nabla^* \cdot (A\nabla u)(x) := \sum_{y\in\mathbb{Z}^d,|x-y|=1} a(e)(u(x) - u(y)),$$

Date: May 25, 2010.

where the sum is over the $2d$ sites y which are connected by an edge $e = [x, y]$ to the site x. It is sometimes more convenient to think in terms of the associated Dirichlet form, i. e.

$$\begin{aligned}\sum \nabla v \cdot A\nabla u \;\; &:= \;\; \sum_{x\in\mathbb{Z}^d} v(x)\,(-\nabla^* \cdot (A\nabla u)(x)) \\ &= \;\; \sum_e (v(x) - v(y))a(e)(u(x) - u(y)),\end{aligned}$$

where the last sum is over all edges e and (x, y) denotes the two sites connected by e, i. e. $e = [x, y] = [y, x]$ (with the convention that an edge is not oriented). We assume the conductivities a to be uniformly elliptic in the sense of

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

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

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

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

Hence the statistics are described by a distribution on the finite interval $[\alpha, \beta]$. We'd like to see this discrete elliptic operator with random coefficients as a good model problem for continuum elliptic operators with random coefficients of correlation length unity.

Classical results in stochastic homogenization of linear elliptic equations (see [8] and [12] for the continuous case, and [10] and [9] for the discrete case) state that there exist *homogeneous and deterministic* coefficients $A_{\rm hom}$ such that the solution operator of the continuum 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 means 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). 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 [14] and Naddaf & Spencer

[11], we refer to [5, Section 1.2]. A standard approach used in practice consists in solving (1.1) in a box $Q_L = [-L, L)^d \cap \mathbb{Z}^d$ 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

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

Such an approach is consistent in the sense that

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

almost surely, as proved for instance in [1] for the continuous case, and in [2] for the discrete case. 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 [13].

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

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

Numerical experiments on such a method are reported on in [6].

An important question for practical purposes is to quantify the dependence of the error $\left\langle |A_{\text{hom}} - A^N_{L,\#}|^2 \right\rangle^{1/2}$ in terms of L and N. Let us give another interpretation of (1.3): This equation on Q_L is equivalent to (1.1) on $\mathbb{Z}^d$ with a modified conductivity matrix $\tilde{A}_L$, that is the periodization of $A_{|Q_L}$ on $\mathbb{Z}^d$. Doing this, we have replaced independent coefficients A by Q_L-periodically correlated coefficients $\tilde{A}$. Since A and $\tilde{A}$ are not jointly stationary (see Definition 4), it may be difficult to compare $\nabla\phi$ to $\nabla\phi_{L,\#}$. To circumvent this difficulty, and following the route of [12], [10], [14], & [11], and as in [5], we slightly depart from (1.3) by introducing a zero-order term in (1.1):

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

As for the periodization, this localizes the dependence of $\phi_T(z)$ upon $A(z')$ to those points $z' \in \mathbb{Z}^d$ such that $|z - z'| \lesssim \sqrt{T}$ (at first order). Yet, unlike the periodization, $\nabla\phi_T$ and $\nabla\phi$ are jointly stationary. In terms of random walk interpretation, the lifetime of the random walker is of order T, and the distance to the origin of order $\sqrt{T}$. Hence, up to taking $T \sim L^2$, in first approximation, the function $\phi_{T|Q_L}$ only depends on the coefficients $A(z)$ for $z \in Q_L$, as it is the case for $\phi_{L,\#}$.

We'd like to view $\phi_{T|Q_L}$ as a variant of $\phi_{L,\#}$ which is convenient for our analysis. We then define

$$\xi \cdot A_{T,L}\xi \;\;=\;\; \int_{\mathbb{Z}^d} (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\eta_L dx, \tag{1.6}$$

where η_L is a smooth mask with unit mass and support Q_L. The aim of this paper is to determine the scaling of the error $\left\langle |A_{\rm hom}-A_{T,L}^N|^2\right\rangle^{1/2}$ in terms of L, T and N, where

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

and $A_{T,L,k}$ are N independent realizations of $A_{T,L}$. Eventually this will allow us to make a reasonable choice for N,T,L at fixed computational complexity (recall that the computational cost to solve a general linear problem with M unknowns is superlinear in M, no matter what the solution method is).

1.2. **Informal statement of the results.** When approximating $A_{\rm hom}$ by $A_{T,L}^N$, we make two types of errors: A "systematic error" and a "random error". The systematic error (see [5, (1.10)])

$$\begin{aligned}
\mathrm{Error}_{\rm sys}(T) \;&:=\; |\,\langle(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\rangle-\langle(\xi+\nabla\phi)\cdot A(\xi+\nabla\phi)\rangle\,| \\
&=\; \langle(\nabla\phi_T-\nabla\phi)\cdot A(\nabla\phi_T-\nabla\phi)\rangle
\end{aligned} \tag{1.7}$$

measures the fact that the coefficient $a(e)$ at bond e does (up to exponentially small terms) not influence $\phi_T(x)$ if $|x-e|\gg\sqrt{T}$. This error vanishes for $T=L^2\uparrow\infty$. The random error to the square can be written as

$$\begin{aligned}
&\mathrm{Error}_{\rm rand}(T,L,N)^2 \\
&:=\; \left\langle\left(\langle(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\rangle-\frac{1}{N}\sum_{k=1}^N\int_{\mathbb{Z}^d}(\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\,dx\right)^2\right\rangle \\
&=\; \left\langle\left(\frac{1}{N}\sum_{k=1}^N\left(\left\langle\int_{\mathbb{Z}^d}(\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\eta_L\,dx\right\rangle\right.\right.\right. \\
&\qquad\qquad \left.\left.\left.-\int_{\mathbb{Z}^d}(\xi+\nabla\phi_{T,k})\cdot A_k(\xi+\nabla\phi_{T,k})\eta_L\,dx\right)\right)^2\right\rangle \\
&=\; \frac{1}{N}\mathrm{var}\left[\int_{\mathbb{Z}^d}(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\eta_L\,dx\right]
\end{aligned} \tag{1.8}$$

using the independence of the $\phi_{T,k}$. 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 [5, Theorem 1], we have proved that

$$\mathrm{var}\left[\int_{\mathbb{Z}^d}(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\eta_L\,dx\right]^{1/2} \;\lesssim\; \left\{\begin{array}{lcl} d=2 &:& L^{-1}\ln^q T, \\ d>2 &:& L^{-d/2}, \end{array}\right. \tag{1.9}$$

for some q depending only on α,β, where "$\lesssim$" stands for "$\le$" up to a multiplicative constant depending only on α,β, and d. Hence,

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

We have also identified the systematic error in the limit of vanishing conductivity contrast, i. e. $1 - \beta/\alpha \ll 1$, and found

$$\mathrm{Error_{sys}}(T) \sim \begin{cases} d=2 & : \quad T^{-1}, \\ d=3 & : \quad T^{-3/2}, \\ d=4 & : \quad T^{-2}\ln T, \\ d>4 & : \quad T^{-2}, \end{cases}$$

where "$\sim$" means that both terms have the same scaling (in T). In this paper, we shall actually prove that for general α and β (see Theorem 1)

$$\mathrm{Error_{sys}}(T) \lesssim \begin{cases} d=2 & : \quad T^{-1}\ln^q T, \\ d=3 & : \quad T^{-3/2}, \\ d=4 & : \quad T^{-2}\ln T, \\ d>4 & : \quad T^{-2}, \end{cases} \tag{1.10}$$

where there is a logarithmic correction for $d = 2$ when compared to the vanishing conductivity asymptotics.

These two estimates then give a clear suggestion on the choice of N and L (and therefore T via $T = L^2$). Let us compare the two extreme strategies:

(A) compute $N_A = N$ realizations of a system of size $L_A = L$, $T_A = L^2$;

(B) compute $N_B = 1$ realization of a system of size $L_B = N^{1/d}L$, $T_B = N^{2/d}L^2$.

For both methods, the number of unknowns is $M = NL^d$ and the overall errors are:

$$\begin{aligned} \mathrm{Error_A}(N, M) \;&:=\; \mathrm{Error_{sys}}(L^2) + \mathrm{Error_{rand}}(L^2, L, N) \\ &\lesssim \begin{cases} d=2 & : \quad NM^{-1}\ln^q(NM^{-1}) + M^{-1/2}\ln^q M, \\ d=3 & : \quad NM^{-1} + M^{-1/2}, \\ d=4 & : \quad NM^{-1}\ln(NM^{-1}) + M^{-1/2}, \\ d>4 & : \quad (NM^{-1})^{4/d} + M^{-1/2}, \end{cases} \end{aligned}$$

and

$$\begin{aligned} \mathrm{Error_B}(M) \;&:=\; \mathrm{Error_{sys}}(N^{2/d}L^2) + \mathrm{Error_{rand}}(N^{2/d}L^2, N^{1/d}L, 1) \\ &\lesssim \begin{cases} d=2 & : \quad M^{-1/2}\ln^q M, \\ 2<d\le 8 & : \quad M^{-1/2}, \\ d>8 & : \quad M^{-4/d}. \end{cases} \end{aligned}$$

For $d = 2, 3, 4$ and $N = \sqrt{M}$, $\mathrm{Error_A}(\sqrt{M}, M)$ and $\mathrm{Error_B}(M)$ are of the same order (up to a logarithm for $d = 4$). Since the cost of solving a linear system is superlinear in terms of the number of unknows, it is more favourable to solve N systems of L^d unknowns than 1 system of NL^d unknowns. Hence, strategy (A) may be preferable for $d \le 4$. *In particular, for $d \le 4$ it seems best to evenly split a given number M of unknowns into the number N of realizations and the number L^d of unknowns per realization, i. e. $N = L^d = \sqrt{M}$.* In practice however, such a statement has to be taken with caution, since boundary terms come into play (ϕ_T has to be approximated on a finite domain Q_R, $R \ge L$) and the conclusion also strongly depends on the linear solver used. We refer the reader to [4] for a more complete discussion of such a numerical strategy.

The article is organized as follows: In Section 2, we introduce the general framework and state the main results of this paper, i. e. the systematic error actually scales as in (1.10). The last two sections are dedicated to its proof.

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

- $d \geq 2$ is the dimension;
- $\int_{\mathbb{Z}^d} dx$ denotes the sum over $x \in \mathbb{Z}^d$, and $\int_D dx$ denotes the sum over $x \in \mathbb{Z}^d$ such that $x \in D$, D subset 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;
- $\mathrm{cov}[\cdot;\cdot]$ is the covariance associated with the ensemble average;
- $\lesssim$ and $\gtrsim$ stand for $\leq$ and $\geq$ up to a multiplicative constant which only depends on the dimension d and the constants α, β (see Definition 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;
- $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ denotes the canonical basis of $\mathbb{Z}^d$.

2. Main result

2.1. General framework.

Definition 1. We say that a is a conductivity function if there exist $0 < \alpha \leq \beta < \infty$ such that for every edge e of $\mathbb{Z}^d$, one has $a(e) \in [\alpha, \beta]$. We denote by $\mathcal{A}_{\alpha\beta}$ the set of such conductivity functions.

Definition 2. The elliptic operator $L : L^2_{\mathrm{loc}}(\mathbb{Z}^d) \to L^2_{\mathrm{loc}}(\mathbb{Z}^d), u \mapsto Lu$ associated with a conductivity function $a \in \mathcal{A}_{\alpha\beta}$ is defined for all $x \in \mathbb{Z}^d$ by

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

where

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

and

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

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

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

Definition 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 4. The conductivity matrix A is obviously stationary in the sense that for all $z \in \mathbb{Z}^d$, $A(\cdot + z)$ and $A(\cdot)$ have the same statistics, so that for all $x, z \in \mathbb{Z}^d$,

$$\langle A(x+z) \rangle = \langle A(x) \rangle\,.$$

Therefore, any translation invariant function of A, such as the modified corrector ϕ_T (see Lemma 2), is jointly stationary with A. In particular, not only are ϕ_T and its gradient $\nabla\phi_T$ stationary, but also 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 $\nabla\phi_T$.

Another translation invariant function of A is the Green functions G_T of Definition 6. In this case, stationarity means that $G_T(\cdot+z,\cdot+z)$ has the same statistics as $G_T(\cdot,\cdot)$ for all $z\in\mathbb{Z}^d$, so that in particular, for all $x,y,z\in\mathbb{Z}^d$,

$$\langle G_T(x+z,y+z)\rangle=\langle G_T(x,y)\rangle\,.$$

Lemma 1 (corrector)**.** [10, Theorem 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*

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

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

We also define an "approximation" of the corrector as follows:

Lemma 2 (approximate corrector)**.** [10, 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 "approximate" corrector equation*

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

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

Definition 5 (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 1. We define the homogenized $d\times d$-matrix A_{hom} as

$$\xi\cdot A_{\mathrm{hom}}\xi\;=\;\langle(\xi+\nabla\phi)\cdot A(\xi+\nabla\phi)(0)\rangle\,. \tag{2.4}$$

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

2.2. **Statement of the main results.** The main result of the article is the following estimate of the systematic error introduced in Section 1.

Theorem 1. *Let $a\in\mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function, and let ϕ_T denote the approximate corrector associated with the conductivity function a and direction $\xi\in\mathbb{R}^d$, $|\xi|=1$. We then define for all $T\gg1$ the symmetric matrix A_T characterized by*

$$\xi\cdot A_T\xi:=\langle(\xi+\nabla\phi_T)\cdot A(\xi+\nabla\phi_T)\rangle\,. \tag{2.5}$$

Then, there exists an exponent $q>0$ depending only on α,β such that

$$\begin{array}{ll} d=2: & |A_{\mathrm{hom}}-A_T|\lesssim T^{-1}(\ln T)^q,\\ d=3: & |A_{\mathrm{hom}}-A_T|\lesssim T^{-3/2},\\ d=4: & |A_{\mathrm{hom}}-A_T|\lesssim T^{-2}\ln T,\\ d>4: & |A_{\mathrm{hom}}-A_T|\lesssim T^{-2}.\end{array} \tag{2.6}$$

As a by-product of the proof of Theorem 1, we obtain the following

Corollary 1. *Let $a\in\mathcal{A}_{\alpha\beta}$ be an i. i. d. conductivity function, $d>2$, $T>0$, and let ϕ_T and ϕ denote the approximate corrector and stationary corrector (see* [5, Corollary 1]*) associated with the conductivity function a and direction $\xi\in\mathbb{R}^d$, $|\xi|=1$, respectively. Then*

$$T^{-1}\left\langle(\phi_T-\phi)^2\right\rangle+\left\langle|\nabla\phi_T-\nabla\phi|^2\right\rangle\lesssim\left\{\begin{array}{lll} d=3 & : & T^{-3/2},\\ d=4 & : & T^{-2}\ln T,\\ d>4 & : & T^{-2}.\end{array}\right. \tag{2.7}$$

In particular,

$$\lim_{T\to\infty}\left(\left\langle(\phi_T-\phi)^2\right\rangle+\left\langle|\nabla\phi_T-\nabla\phi|^2\right\rangle\right) \;=\; 0.$$

This corollary gives a full characterization of the convergence of the regularized corrector to the exact corrector for $d>2$.

Remark 1. Note that the definition (2.5) of A_T does not include the zero-order term $T^{-1}\left\langle\phi_T^2\right\rangle$, so that $\xi\cdot A_T\xi$ does not coincide with the energy associated with the equation. Surprisingly, the addition of the zero-order term in the definition of A_T would make the estimate (2.6) saturate at T^{-1} for $d>2$.

Remark 2. For $d=2$, although we lose control of ϕ_T we may still quantify the rate of convergence of $\nabla\phi_T$ to $\nabla\phi$, the gradient of the corrector of Definition 1. In particular, (2.7) is replaced by

$$\left\langle|\nabla\phi_T-\nabla\phi|^2\right\rangle \;\lesssim\; T^{-1}\ln^q T$$

for some $q>0$ depending only on α,β.

2.3. **Auxiliary lemmas.** In order to prove Theorem 1 and Corollary 1, we need three auxiliary lemmas in addition to the results of [5]: The first one is a covariance estimate very similar to the variance estimate in [5, Lemma 2.3], the next one is a refined version of the decay estimates of [5, Lemma 2.8], whereas the last one is a generalization of the convolution estimate of [5, Lemma 2.10].

Lemma 3 (covariance estimate)**.** *Let $a=\{a_i\}_{i\in\mathbb{N}}$ be a sequence of i. i. d. random variables with range $[\alpha,\beta]$. Let X and Y be two Borel measurable functions of $a\in\mathbb{R}^{\mathbb{N}}$ (i. e. measurable w. r. t. the smallest σ-algebra on $\mathbb{R}^{\mathbb{N}}$ for which all coordinate functions $\mathbb{R}^{\mathbb{N}}\ni a\mapsto a_i\in\mathbb{R}$ are Borel measurable, cf.* [7, Definition 14.4]*). Then we have*

$$\operatorname{cov}[X;Y] \;\le\; \sum_{i=1}^{\infty}\left\langle\sup_{a_i}\left|\frac{\partial X}{\partial a_i}\right|^2\right\rangle^{1/2}\left\langle\sup_{a_i}\left|\frac{\partial Y}{\partial a_i}\right|^2\right\rangle^{1/2}\operatorname{var}[a_1]\,, \tag{2.8}$$

where $\sup_{a_i}\left|\frac{\partial Z}{\partial a_i}\right|$ denotes the supremum of the modulus of the i-th partial derivative

$$\frac{\partial Z}{\partial a_i}(a_1,\cdots,a_{i-1},a_i,a_{i+1},\cdots)$$

of Z with respect to the variable $a_i\in[\alpha,\beta]$, for $Z=X,Y$.

We define discrete Green's functions as follows:

Definition 6 (discrete Green's function)**.** Let $d\ge 2$. For all $T>0$, the Green 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

$$\int_{\mathbb{Z}^d}T^{-1}G_T(x,y;a)v(x)\,dx+\int_{\mathbb{Z}^d}\nabla v(x)\cdot A(x)\nabla_x G_T(x,y;a)\,dx=v(y),\qquad \forall v\in L^2(\mathbb{Z}^d), \tag{2.9}$$

where A is as in (2.1).

Throughout this paper, when no confusion occurs, we use the shorthand notation $G_T(x,y)$ for $G_T(x,y;a)$. We need a decay of the Green function $G_T(x,y)$ and its (discrete) gradient $\nabla_x G_T(x,y)$ in $|x-y|\gg 1$ that is *uniform* in a but nevertheless coincides (in terms of scaling) with the decay of the *constant-coefficient* Green function. The constant-coefficient Green function in the continuous case is known to decay as

$$|x-y|^{2-d}\exp(-\text{const.}\frac{|x-y|}{\sqrt{T}}) \quad \text{for} \quad d>2 \ \text{ and}$$
$$(\ln\frac{\sqrt{T}}{|x-y|})\exp(-\text{const.}\frac{|x-y|}{\sqrt{T}}) \quad \text{for} \quad d=2;$$

its gradient decays as the first derivative of these expressions. Note the cross-over of the decay at distances $|x-y|$ of the order of the intrinsic length scale $\sqrt{T}\gg 1$ from algebraic (or logarithmic in case of $d=2$) to exponential.

In the class of a-uniform estimates, these decay properties survive as *pointwise* in (x,y) estimates on the level of the discrete Green function $G_T(x,y)$ itself, but only as *averaged* estimates on the level of its discrete gradient $\nabla_x G_T(x,y)$. More precisely, $\nabla_x G_T(x,y)$ has to be averaged in x on dyadic annuli centered at $x=y$. It will be important that the average can be (at least slightly) stronger than a *square* average (see [5, Lemma 2.9]). On the other hand, we do not need the exponential decay: Super algebraic decay is sufficient for our purposes.

Lemma 4 (pointwise decay estimate on G_T)**.** *Let $a\in\mathcal{A}_{\alpha\beta}$, and G_T be the associated Green function. For $d>2$, we have for all $k>0$, and all $x,y\in\mathbb{Z}^d$*

$$G_T(x,y) \ \lesssim\ (1+|x-y|)^{2-d}\min\{1,(\frac{|x-y|}{\sqrt{T}})^{-k}\}, \tag{2.10}$$

where the constant in "$\lesssim$" depends on k. For $d=2$, we have for all $k>0$

$$G_T(x,y)\ \lesssim\ \left\{\begin{array}{ll}\ln(\frac{\sqrt{T}}{1+|x-y|}) & \text{for} \quad |x-y|\ll\sqrt{T}\\ (\frac{|x-y|}{\sqrt{T}})^{-k} & \text{for} \quad |x-y|\gtrsim\sqrt{T}\end{array}\right\}, \tag{2.11}$$

where the constant in "$\lesssim$" depends on k.

Finally, for the proof of Theorem 1, we need to know that also the *convolution* of the gradient of the Green's function with itself decays at the optimal rate, i. e.

Lemma 5 (convolution estimate)**.** *Let $h_T, g_T:\mathbb{Z}^d\to\mathbb{R}^+$ satisfy the following properties.* Assumptions on h_T (estimate of $|\nabla_x G_T(y+z,y)|$): *For all $R\gg 1$ and $T>0$,*

$$d>2: \qquad \int_{R<|z|\le 2R} h_T(z)^2dz\ \lesssim\ R^{2-d}, \tag{2.12}$$

$$d=2: \qquad \int_{R<|z|\le 2R} h_T(z)^2dz \lesssim \min\{1,\sqrt{T}R^{-1}\}^2, \tag{2.13}$$

$$\tag{2.14}$$

and for $R\sim 1$

$$d\ge 2: \qquad \int_{|z|\le R} h_T(z)^2dz\ \lesssim\ 1\,. \tag{2.15}$$

Assumptions on g_T (estimate of $G_T(y+z,y)$): *For $d>2$, and for all $z\in\mathbb{Z}^d$,*

$$g_T(z)\ =\ (1+|z|)^{2-d}\min\{1,(\frac{|z|}{\sqrt{T}})^{-3}\}, \tag{2.16}$$

and for $d=2$,

$$g_T(z) \;=\; \left\{\begin{array}{ll} \ln(\frac{1+|z|}{\sqrt{T}}) & \textit{for} \quad |z|\le\sqrt{T} \\ (\frac{|z|}{\sqrt{T}})^{-3} & \textit{for} \quad |z|>\sqrt{T} \end{array}\right\}. \tag{2.17}$$

Then we have

$$\int_{\mathbb{Z}^d} g_T(z)\int_{\mathbb{Z}^d} h_T(w)h_T(z-w)\,dw\,dz \;\lesssim\; \left\{\begin{array}{ll} d=2: & T, \\ d=3: & \sqrt{T}, \\ d=4: & \ln T, \\ d>4: & 1. \end{array}\right. \tag{2.18}$$

3. Proof of the main results

Throughout this section, we let $\xi\in\mathbb{R}^d$ be such that $|\xi|=1$.

3.1. **Proof of Theorem 1.** In view of (1.7), in order to estimate $|A_T-A_{\mathrm{hom}}|$, we need to estimate how close the modified corrector ϕ_T is to the original corrector ϕ (in terms of $\langle|\nabla\phi_T-\nabla\phi|^2\rangle$). Therefore, it is natural to introduce $\psi_T=T^2\frac{\partial\phi_T}{\partial T}$ (the prefactor T^2 is such that ψ_T is properly renormalized in the limit $T\uparrow\infty$ at least for large d). Considering ψ_T is also convenient since for $d=2$, the corrector ϕ is not known to be stationary (only its gradient is known to be stationary) so that working with the modified correctors ϕ_T, which are known to be stationary, avoids technical subtleties. In fact, we opt for a dyadically discrete version of ψ_T defined via

$$\psi_T \;:=\; T(\phi_{2T}-\phi_T). \tag{3.1}$$

This discrete version has the technical advantage that we don't have to think about the differentiability of ϕ_T in T. Moreover, its dyadic nature is in line with the dyadic decomposition of the T-axis according to

$$|A_T-A_{\mathrm{hom}}| \;\le\; \sum_{i=0}^{\infty}|A_{2^iT}-A_{2^{i+1}T}| \tag{3.2}$$

forced upon us in the case of $d=2$. In order to get (3.2), we used the fact that

$$\lim_{T\to\infty} A_T \;=\; A_{\mathrm{hom}}, \tag{3.3}$$

which is proved in [5, Proof of Theorem 1, Step 8]. We shall also use that ψ_T solves

$$T^{-1}\psi_T-\nabla^*\cdot A\nabla\psi_T \;=\; \frac{1}{2}\phi_{2T}. \tag{3.4}$$

We split the proof in eight steps.

Step 1. Derivation of

$$|\xi\cdot(A_{2T}-A_T)\xi| \;\le\; T^{-2}|\langle\phi_T\psi_T\rangle|+\frac{T^{-2}}{2}|\langle\phi_{2T}\psi_T\rangle|. \tag{3.5}$$

We recall the following consequence of (2.3) which is proved in [5, Proof of Theorem 1, Step 8]:

$$T^{-1}\langle\phi_T\chi\rangle+\langle(\xi+\nabla\phi_T)\cdot A\nabla\chi\rangle \;=\; 0 \tag{3.6}$$

for every field $\chi : \mathbb{Z}^d \to \mathbb{R}$ that is jointly stationary with A and such that $\langle \chi^2 \rangle < \infty$. From formally differentiating the definition (2.5) of A_T w. r. t. T and using (3.6) for $\chi = \frac{\partial \phi_T}{\partial T}$, we obtain

$$\xi \cdot \frac{\partial A_T}{\partial T}\xi \;=\; -2T^{-1}\langle \frac{\partial \phi_T}{\partial T}\phi_T\rangle.$$

We claim that the corresponding discrete-in-T version reads

$$\xi \cdot (A_{2T} - A_T)\xi \;=\; -T^{-2}\left(\langle \psi_T\phi_T\rangle + \frac{1}{2}\langle \psi_T\phi_{2T}\rangle\right). \tag{3.7}$$

Indeed, by definition of A_T, by expanding the square, by symmetry of A, by definition of ψ_T, and (3.6), we have

$$\begin{aligned}
&\xi \cdot (A_{2T} - A_T)\xi \\
&= \langle (\xi + \nabla\phi_{2T}) \cdot A(\xi + \nabla\phi_{2T})\rangle - \langle (\xi + \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\rangle \\
&= \langle (\nabla\phi_{2T} - \nabla\phi_T) \cdot A(\xi + \nabla\phi_{2T})\rangle + \langle (\nabla\phi_{2T} - \nabla\phi_T) \cdot A(\xi + \nabla\phi_T)\rangle \\
&\overset{(3.1)}{=} T^{-1}\left(\langle \nabla\psi_T \cdot A(\xi + \nabla\phi_{2T})\rangle + \langle \nabla\psi_T \cdot A(\xi + \nabla\phi_T)\rangle\right) \\
&\overset{(3.6)}{=} -T^{-1}\left((2T)^{-1}\langle \psi_T\phi_{2T}\rangle + T^{-1}\langle \psi_T\phi_T\rangle\right).
\end{aligned}$$

In the next four steps, we focus on the first term of the r. h. s. of (3.5). The second term will be dealt with the same way in Step 7.

Step 2. Proof of

$$|\langle \phi_T\psi_T\rangle| \;\lesssim\; \sum_e \left\langle \sup_{a(e)} \left|\frac{\partial \phi_T(0)}{\partial a(e)}\right|^2 \right\rangle^{1/2} \left\langle \sup_{a(e)} \left|\frac{\partial \psi_T(0)}{\partial a(e)}\right|^2 \right\rangle^{1/2}, \tag{3.8}$$

where the sum runs over the edges e, and proof of the representation formulas

$$\frac{\partial \phi_T(0)}{\partial a(e)} \;=\; -(\xi_i + \nabla_i\phi_T(z))\nabla_{z_i}G_T(z,0), \tag{3.9}$$

$$\begin{aligned}\frac{\partial \psi_T(0)}{\partial a(e)} \;=\;& -\nabla_i\psi_T(z)\nabla_{z_i}G_T(z,0) \\ &-\frac{1}{2}\int_{\mathbb{Z}^d} G_T(0,w)(\xi_i + \nabla_i\phi_{2T}(z))\nabla_{z_i}G_{2T}(z,w)dw,\end{aligned} \tag{3.10}$$

where the edge is $e = [z, z + \mathbf{e}_i]$.

Due to [5, Lemma 2.6] the functions ϕ_T and ψ_T are measurable with respect to the coefficients a. Hence (3.8) is a consequence of the covariance estimate of Lemma 3: Since $\langle \phi_T\rangle = \langle \psi_T\rangle = 0$,

$$\begin{aligned}\langle \phi_T\psi_T\rangle \;&=\; \langle (\phi_T - \langle \phi_T\rangle)(\psi_T - \langle \psi_T\rangle)\rangle \\ &=\; \mathrm{cov}\left[\phi_T; \psi_T\right].\end{aligned}$$

Formula (3.9) is identical to [5, Lemma 2.4, (2.12)]. To prove (3.10), we first make use of the Green representation formula for the solution to (3.4):

$$\psi_T(x) \;=\; \frac{1}{2}\int_{\mathbb{Z}^d} G_T(x,w)\phi_{2T}(w)dw, \tag{3.11}$$

for all $x \in \mathbb{Z}^d$. Since $a(e) \mapsto \phi_T(\cdot;a(e))$ and $a(e) \mapsto \phi_{2T}(\cdot;a(e))$ are continuously differentiable by [5, Lemma 2.4], we deduce by formula (3.1) that $a(e) \mapsto \psi_T(\cdot;a(e))$ is also continuously differentiable. Using then the formulas [5, Lemma 2.5, (2.15)] and [5, Lemma 2.4, (2.12)] for the derivatives of G_T and ϕ_T with respect to $a(e)$, and the fact that $G_T \in L^1(\mathbb{Z}^d)$ (see [5, Corollary 2.2]), we may switch the order of the differentiation and the integration to obtain for all $x \in \mathbb{Z}^d$

$$\begin{aligned}
\frac{\partial \psi_T(x)}{\partial a(e)} &\\
=\;& \frac{1}{2}\int_{\mathbb{Z}^d} \frac{\partial G_T(x,w)}{\partial a(e)}\phi_{2T}(w)dw + \frac{1}{2}\int_{\mathbb{Z}^d} G_T(x,w)\frac{\partial \phi_{2T}(w)}{\partial a(e)}dw \\
\overset{[5,\ (2.12)\&(2.15)]}{=}\;& -\frac{1}{2}\int_{\mathbb{Z}^d} \nabla_{z_i} G_T(x,z)\nabla_{z_i} G_T(z,w)\phi_{2T}(w)dw \\
& -\frac{1}{2}\int_{\mathbb{Z}^d} G_T(x,w)(\xi_i + \nabla_i \phi_{2T}(z))\nabla_{z_i} G_{2T}(z,w)dw \\
\overset{(3.11)}{=}\;& -\nabla_{z_i} G_T(x,z)\nabla_i \psi_T(z) - \frac{1}{2}\int_{\mathbb{Z}^d} G_T(x,w)(\xi_i + \nabla_i \phi_{2T}(z))\nabla_{z_i} G_{2T}(z,w)dw,
\end{aligned} \tag{3.12}$$

which is (3.10) taking $x = 0$.

From now on in the proof, we let g_T be defined as in Lemma 5 (that is, g_T decays as the Green function G_T).

Step 3. In this step, we shall prove that

$$|\langle \phi_T \psi_T \rangle| \;\lesssim\; \mathcal{L} + \mathcal{N}, \tag{3.13}$$

where

$$\begin{aligned}
\mathcal{L} \;:=\;& \int_{\mathbb{Z}^d} \left\langle (1+|\nabla \phi_T(z)|^2)|\nabla_z G_T(z,0)|^2 \right\rangle^{1/2} \\
& \left\langle \left(\int_{\mathbb{Z}^d} g_T(w)(1+|\nabla \phi_{2T}(z)|)|\nabla_z G_{2T}(z,w)|dw \right)^2 \right\rangle^{1/2} dz,
\end{aligned} \tag{3.14}$$

and $\mathcal{N} = \mathcal{N}_1 + \mathcal{N}_2$,

$$\mathcal{N}_1 \;:=\; \int_{\mathbb{Z}^d} \left\langle (1+|\nabla \phi_T(z)|^2)|\nabla_z G_T(z,0)|^2 \right\rangle^{1/2} \left\langle |\nabla \psi_T(z)|^2 |\nabla_z G_T(z,0)|^2 \right\rangle^{1/2} dz \tag{3.15}$$

$$\begin{aligned}
\mathcal{N}_2 \;:=\;& \mu_d(T)\int_{\mathbb{Z}^d} \left\langle (1+|\nabla \phi_T(z)|^2)|\nabla_z G_T(z,0)|^2 \right\rangle^{1/2} \\
& \left\langle (1+|\nabla \phi_{2T}(z)|^2)|\nabla_z G_T(z,0)|^2 \right\rangle^{1/2} dz,
\end{aligned} \tag{3.16}$$

with

$$\mu_d(T) \;:=\; \begin{cases} d=2 & : \ \ln T, \\ d>2 & : \ 1. \end{cases}$$

The term $\mathcal{L}$ is a linear error: It is of the same type as for the analysis in the limit of vanishing ellipticity contrast (see [5, Appendix]). On the contrary, the term $\mathcal{N}$ is nonlinear and does not appear in the limit of vanishing ellipticity contrast. As we shall prove, it is of lower order. The terms $\mathcal{L}$ and $\mathcal{N}_1$ in estimate (3.13) would be direct consequences of (3.8), and (3.9) & (3.10), disregarding the suprema in $a(e)$ in (3.8). Taking the suprema in $a(e)$ into account actually brings the second nonlinear term $\mathcal{N}_2$, which turns out to be of lower order than $\mathcal{N}_1$.

According to [5, Lemma 2.4, (2.13)] we have for (3.9)

$$\sup_{a(e)}\left|\frac{\partial\phi_T(0)}{\partial a(e)}\right| \;\lesssim\; (1+|\nabla_i\phi_T(z)|)|\nabla_z G_T(z,0)|. \tag{3.17}$$

It remains to deal with (3.10). Using the pointwise decay of G_T in Lemma 4 combined with the susceptibility estimates [5, Lemma 2.4, (2.14)] and [5, Lemma 2.5, (2.16)] of $\nabla\phi_T$ and ∇G_T w. r. t. $a(e)$, we obtain

$$\begin{aligned}&\sup_{a(e)}\left|\frac{1}{2}\int_{\mathbb{Z}^d} G_T(0,w)(\xi_i+\nabla_i\phi_{2T}(z))\nabla_{z_i}G_{2T}(z,w)dw\right|\\ &\lesssim\; \int_{\mathbb{Z}^d} g_T(w)(1+|\nabla_i\phi_{2T}(z)|)|\nabla_z G_{2T}(z,w)|dw,\end{aligned} \tag{3.18}$$

which together with (3.17) gives the linear term $\mathcal{L}$.

To treat the first term of the r. h. s. of (3.10), we need to deal with the supremum of $|\nabla_i\psi_T(z)|$ over $a(e)$. We appeal to (3.12) that we rewrite in the form

$$\frac{\partial\psi_T(x)}{\partial a(e)} \;=\; -\nabla_i\psi_T(z)G_T(x,e)-\frac{1}{2}(\xi_i+\nabla_i\phi_{2T}(z))\int_{\mathbb{Z}^d}G_T(x,w)G_T(e,w)dw,$$

where $G_T(x,e):=G_T(x,z+\mathbf{e}_i)-G_T(x,z)$ and $G_T(e,w):=G_T(z+\mathbf{e}_i,w)-G_T(z,w)$. Hence,

$$\frac{\partial\nabla_i\psi_T(z)}{\partial a(e)} \;=\; -\nabla_i\psi_T(z)G_T(e,e)-\frac{1}{2}(\xi_i+\nabla_i\phi_{2T}(z))\int_{\mathbb{Z}^d}G_T(e,w)G_{2T}(e,w)dw, \tag{3.19}$$

where $G_T(e,e):=G_T(z+\mathbf{e}_i,z+\mathbf{e}_i)+G_T(z,z)-G_T(z+\mathbf{e}_i,z)-G_T(z,z+\mathbf{e}_i)$. On the one hand, the uniform bound [5, Corollary 2.3] on ∇G_T yields $|G_T(e,e)|\lesssim 1$. On the other hand, as we shall argue, the integrability of ∇G_T and ∇G_{2T} from [5, Lemma 2.9] (combined with the uniform bound [5, Corollary 2.3] on gradients) implies

$$\int_{\mathbb{Z}^d}G_T(e,w)G_{2T}(e,w)dw \;\lesssim\; \mu_d(T)=\left\{\begin{array}{lcl} d=2 &:& \ln T\,,\\ d>2 &:& 1\,.\end{array}\right. \tag{3.20}$$

Hence if we regard (3.19) as an ordinary differential equation for $\nabla_i\psi_T(z)$ in the variable $a(e)$, we obtain

$$\sup_{a(e)}|\nabla_i\psi_T(z)| \;\lesssim\; |\nabla_i\psi_T(z)|+\mu_d(T)(1+|\nabla_i\phi_{2T}(z)|) \tag{3.21}$$

since $a(e)$ lies in a bounded domain $[\alpha,\beta]$, and $\sup_{a(e)}|\nabla_i\phi_{2T}(z)|\lesssim 1+|\nabla_i\phi_{2T}(z)|$ according to [5, Lemma 2.4, (2.14)] with $2T$ instead of T. Note that (3.17), (3.21) and $\sup_{a(e)}|\nabla_{z_i}G_T(z,0)|\lesssim|\nabla_{z_i}G_T(z,0)|$ give the nonlinear terms $\mathcal{N}_1$ and $\mathcal{N}_2$.

We now give the argument for (3.20). We first use Cauchy-Schwarz' inequality

$$\begin{aligned}\int_{\mathbb{Z}^d}G_T(e,w)G_{2T}(e,w)dw \;&\le\; \left(\int_{\mathbb{Z}^d}G_T(e,w)^2dw\right)^{1/2}\left(\int_{\mathbb{Z}^d}G_{2T}(e,w)^2dw\right)^{1/2}\\ &\le\; \left(\int_{\mathbb{Z}^d}|\nabla_z G_T(z,w)|^2dw\right)^{1/2}\left(\int_{\mathbb{Z}^d}|\nabla_z G_{2T}(z,w)|^2dw\right)^{1/2}\end{aligned}$$

and then make a decomposition of $\mathbb{Z}^d$ into the ball of radius $R \sim 1$, and dyadic annuli $\{w : 2^i R < |z-w| \le 2^{i+1}R\}$ for $i \in \mathbb{N}$. On the ball of radius R we use the uniform estimate of [5, Corollary 2.3] on ∇G_T, whereas on the dyadic annuli we appeal to the decay estimate in [5, Lemma 2.9] for the gradient of the Green function, which requires R to be sufficiently large although still of order 1. Both terms in the r. h. s. scale the same way and we only treat the first one:

$$\begin{aligned}
&\int_{\mathbb{Z}^d} |\nabla_z G_T(z,w)|^2 dw \\
&= \int_{|z-w|\le R} |\nabla_z G_T(z,w)|^2 dw + \sum_{i=0}^{\infty} \int_{2^i R<|z-w|\le 2^{i+1}R} |\nabla_z G_T(z,w)|^2 dw \\
&\lesssim 1 + \sum_{i=1}^{\infty} (2^i)^{d+2(1-d)} \min\{1, \sqrt{T}(2^i R)^{-1}\}^2 \\
&\lesssim \mu_d(T),
\end{aligned}$$

using [5, Corollary 2.3] and [5, Lemma 2.9] for $k=2$, respectively. This concludes Step 3.

Step 4. Suboptimal estimate of the nonlinear term $\mathcal{N}$:

$$\mathcal{N}_1 \lesssim \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2} \begin{cases} d=2 & : \sqrt{T}\ln^q T, \\ d=3 & : \ln T, \\ d>3 & : 1, \end{cases} \tag{3.22}$$

$$\mathcal{N}_2 \lesssim \mu_d(T)^q, \tag{3.23}$$

where q is a generic exponent which only depends on α, β. We first deal with $\mathcal{N}_1$, and begin with the second factor of the r. h. s. of (3.15). The pointwise estimate (2.10) of Lemma 4 for $d>2$ on the Green function gives the *suboptimal pointwise* estimate on the gradient of the Green function

$$|\nabla G_T(z,0)| \le G_T(z,0) + \sum_{i=1}^{d} G_T(z+\mathbf{e}_i,0) \lesssim (1+|z|)^{2-d}. \tag{3.24}$$

This estimate coincides for $d=2$ with the uniform bound of [5, Corollary 2.3]. The coercivity of A thus yields

$$\begin{aligned}
&\left\langle |\nabla G_T(z,0)|^2 |\nabla\psi_T(z)|^2 \right\rangle^{1/2} \\
&\lesssim (1+|z|)^{2-d} \langle \nabla\psi_T(z)\cdot A(z)\nabla\psi_T(z)\rangle^{1/2} \\
&= (1+|z|)^{2-d} \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2}
\end{aligned}$$

by joint stationarity of $\nabla\psi_T$ and A. Hence, (3.15) turns into

$$\mathcal{N}_1 \lesssim \langle \nabla\psi_T\cdot A\nabla\psi_T\rangle^{1/2} \int_{\mathbb{Z}^d} (1+|z|)^{2-d} \left\langle (1+|\nabla\phi_T(z)|^2)|\nabla G_T(z,0)|^2\right\rangle^{1/2} dz.$$

We then let $p>2$ be a Meyers' exponent as in [5, Lemma 2.9] and use Hölder's inequality in probability with exponents $(p/(p-2), p/2)$, the stationarity of $\nabla\phi_T$, the fact that the gradient of ϕ_T is estimated by ϕ_T as in (3.24), and the bounds on the stochastic moments of ϕ_T in [5, Proposition 1]

$$\begin{aligned}
\mathcal{N}_1 &\lesssim \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2} \\
&\qquad \int_{\mathbb{Z}^d} (1+|z|)^{2-d} \left\langle 1+|\nabla\phi_T(z)|^{2p/(p-2)}\right\rangle^{(p-2)/(2p)} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \\
&= \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2} \left\langle 1+|\nabla\phi_T|^{2p/(p-2)}\right\rangle^{(p-2)/(2p)} \\
&\qquad \int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \\
&\lesssim \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2} \left\langle 1+|\phi_T|^{2p/(p-2)}\right\rangle^{(p-2)/(2p)} \int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \\
&\lesssim \mu_d(T)^q \langle \nabla\psi_T \cdot A\nabla\psi_T\rangle^{1/2} \int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz, \qquad (3.25)
\end{aligned}$$

for some generic q depending only on α,β. Hölder's inequality with exponents $(p,p/(p-1))$ in $\mathbb{Z}^d$, combined with the same dyadic decomposition of $\mathbb{Z}^d$ as for the proof of (3.20), (and the uniform bound on ∇G_T from [5, Corollary 2.3]) yields

$$\begin{aligned}
&\int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \\
&\lesssim 1+\sum_{i=0}^{\infty} \left(\left\langle \int_{2^iR<|z|\le 2^{i+1}R} |\nabla G_T(z,0)|^p dz\right\rangle\right)^{1/p} \\
&\qquad \left(\int_{2^iR<|z|\le 2^{i+1}R} (1+|z|)^{(2-d)p/(p-1)} dz\right)^{(p-1)/p}.
\end{aligned}$$

Using the optimal decay of ∇G_T on dyadic annuli in L^p norm from [5, Lemma 2.9] with $k=2p$, this turns into

$$\begin{aligned}
&\int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \\
&\lesssim 1+\sum_{i=0}^{\infty} \left((2^iR)^d (2^iR)^{(1-d)p} \min\{1,\frac{\sqrt{T}}{2^iR}\}^{2p}\right)^{1/p} \left((2^iR)^d (2^iR)^{(2-d)p/(p-1)}\right)^{(p-1)/p} \\
&= 1+\sum_{i=0}^{\infty} (2^iR)^{3-d} \min\{1,\frac{\sqrt{T}}{2^iR}\}^2.
\end{aligned}$$

Recalling that $R\sim 1$, this implies

$$\int_{\mathbb{Z}^d} (1+|z|)^{2-d} \langle |\nabla G_T(z,0)|^p\rangle^{1/p}\,dz \lesssim \begin{cases} d=2 & : \sqrt{T}, \\ d=3 & : \ln T, \\ d>3 & : 1. \end{cases}$$

Combined with (3.25) it proves (3.22).

We now turn to $\mathcal{N}_2$. Proceeding as above to deal with the terms $\nabla\phi_T$ and $\nabla\phi_{2T}$ in $\mathcal{N}_2$, we obtain as desired

$$\begin{aligned}\mathcal{N}_2 &\lesssim \mu_d(T)\mu_d(T)^{2q}\int_{\mathbb{Z}^d}\langle|\nabla_z G_T(z,0)|^p\rangle^{2/p}\,dz\\ &\lesssim \mu_d(T)^{2q+2},\end{aligned}$$

using the same dyadic decomposition of $\mathbb{Z}^d$ as for the proof of (3.20) together with the higher integrability of gradients of [5, Lemma 2.9], and [5, Corollary 2.3].

Step 5. Estimate of the linear term $\mathcal{L}$:

$$\mathcal{L}\lesssim\left\{\begin{array}{lcl} d=2 &:& T\,\ln^q T\,,\\ d=3 &:& \sqrt{T}\,,\\ d=4 &:& \ln T\,,\\ d>4 &:& 1\,.\end{array}\right. \tag{3.26}$$

We first treat the second factor of (3.14). We proceed as in Step 4 to deal with the expectation of the corrector term, and let $p>2$ be a Meyers' exponent as in [5, Lemma 2.9]. We obtain by Hölder's inequality in probability with exponents $(p/(p-2),p,p)$ and the bounds on the stochastic moments of ϕ_T from [5, Proposition 1]:

$$\begin{aligned}&\left\langle\left(\int_{\mathbb{Z}^d}g_T(w)(1+|\nabla\phi_{2T}(z)|)|\nabla_{z_i}G_{2T}(z,w)|dw\right)^2\right\rangle\\ &= \int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}g_T(w)g_T(w')\left\langle(1+|\nabla\phi_{2T}(z)|)^2|\nabla_{z_i}G_{2T}(z,w)||\nabla_{z_i}G_{2T}(z,w')|\right\rangle dwdw'\\ &\lesssim \left(1+\left\langle|\phi_{2T}|^{2p/(p-2)}\right\rangle^{(p-2)/p}\right)\\ &\qquad\int_{\mathbb{Z}^d}\int_{\mathbb{Z}^d}g_T(w)g_T(w')\left\langle|\nabla_{z_i}G_{2T}(z,w)|^p\right\rangle^{1/p}\left\langle|\nabla_{z_i}G_{2T}(z,w')|^p\right\rangle^{1/p}dwdw'\\ &\lesssim \mu_d(T)^q\left(\int_{\mathbb{Z}^d}g_T(w)\left\langle|\nabla_z G_{2T}(z,w)|^p\right\rangle^{1/p}dw\right)^2.\end{aligned}$$

We thus have

$$\mathcal{L}\;\lesssim\;\mu_d(T)^q\int_{\mathbb{Z}^d}\left\langle(1+|\nabla\phi_T(z)|^2)|\nabla G_T(z,0)|^2\right\rangle^{1/2}\int_{\mathbb{Z}^d}g_T(w)\left\langle|\nabla_z G_{2T}(z,w)|^p\right\rangle^{1/p}dwdz.$$

Appealing once more to Hölder's inequality in probability with exponents $(p/(p-2),p/2)$ and to [5, Proposition 1], this turns into

$$\begin{aligned}\mathcal{L} &\lesssim \mu_d(T)^{2q}\int_{\mathbb{Z}^d}g_T(w)\int_{\mathbb{Z}^d}\left\langle|\nabla_z G_{2T}(z,w)|^p\right\rangle^{1/p}\left\langle|\nabla_z G_T(z,0)|^p\right\rangle^{1/p}dzdw\\ &= \mu_d(T)^{2q}\int_{\mathbb{Z}^d}g_T(w)\int_{\mathbb{Z}^d}h_{2T}(z-w)h_T(z)dzdw,\end{aligned}$$

where, by stationarity, we have set

$$\begin{aligned}h_T(w) &= \left\langle|\nabla_w G_T(w,0)|^p\right\rangle^{1/p},\\ h_{2T}(w) &= \left\langle|\nabla_w G_{2T}(w,0)|^p\right\rangle^{1/p}.\end{aligned}$$

By the optimal decay estimate of ∇G_T on dyadic annuli from [5, Lemma 2.9] (and by the uniform bounds on ∇G_T from [5, Corollary 2.3]), and by definition of g_T, we are in position to apply Lemma 5. Estimate (3.26) is thus proved.

Step 6. Proof of

$$\langle \nabla\psi_T \cdot A\nabla\psi_T\rangle \;\le\; |\langle \phi_T\psi_T\rangle|. \tag{3.27}$$

Using (3.1), we rewrite (3.4) as

$$\begin{aligned}(2T)^{-1}\psi_T - \nabla^*\cdot A\nabla\psi_T &= \frac{1}{2}\phi_{2T} - (2T)^{-1}\psi_T \\ &= \frac{1}{2}\phi_T. \end{aligned}\tag{3.28}$$

We now multiply (3.28) by ψ_T:

$$(2T)^{-1}\psi_T^2 - (\nabla^*\cdot A\nabla\psi_T)\psi_T \;=\; \frac{1}{2}\phi_T\psi_T.$$

By integration by parts and joint stationarity of ψ_T, $\nabla\psi_T$ and A (see [5, Proof of Theorem 1, Step 8] for details), this turns into

$$(2T)^{-1}\left\langle \psi_T^2\right\rangle + \langle \nabla\psi_T\cdot A\nabla\psi_T\rangle \;=\; \frac{1}{2}\langle\phi_T\psi_T\rangle\,.$$

We then conclude by the non-negativity of the first term.

Step 7. Proof of

$$|\langle\phi_T\psi_T\rangle| \lesssim \begin{cases} d=2 &: \; T\ln^q T, \\ d=3 &: \; \sqrt{T}, \\ d=4 &: \; \ln T, \\ d>4 &: \; 1, \end{cases} \tag{3.29}$$

and

$$|\langle\phi_{2T}\psi_T\rangle| \lesssim \begin{cases} d=2 &: \; T\ln^q T, \\ d=3 &: \; \sqrt{T}, \\ d=4 &: \; \ln T, \\ d>4 &: \; 1. \end{cases} \tag{3.30}$$

From Steps 3, 4 & 5, and Young's inequality, we deduce that

$$|\langle\phi_T\psi_T\rangle| - \frac{1}{2}\langle\nabla\psi_T\cdot A\nabla\psi_T\rangle \lesssim \begin{cases} d=2 &: \; T\ln^q T, \\ d=3 &: \; \sqrt{T}, \\ d=4 &: \; \ln T, \\ d>4 &: \; 1. \end{cases}$$

Combined with Step 6, this shows (3.29).

For (3.30), we proceed exactly as for (3.29) in Steps 2-6. In particular, with obvious notation, we have

$$|\langle\phi_{2T}\psi_T\rangle| \lesssim \mathcal{N}' + \mathcal{L}',$$

where

$$\mathcal{N}' - \frac{1}{2}\langle\nabla\psi_T\cdot A\nabla\psi_T\rangle \lesssim \begin{cases} d=2 &: \; T\ln^q T, \\ d=3 &: \; \ln^2 T, \\ d>3 &: \; 1, \end{cases}$$

and

$$\mathcal{L}' \lesssim \begin{cases} d=2 & : \quad T\ln^q T, \\ d=3 & : \quad \sqrt{T}, \\ d=4 & : \quad \ln T, \\ d>4 & : \quad 1. \end{cases}$$

We then conclude as above.

Step 8. Proof of (2.6).
Steps 1 & 7 yield

$$\begin{aligned} |\xi\cdot(A_T-A_{2T})\xi| &\leq T^{-2}|\,\langle\phi_T\psi_T\rangle\,|+(2T^2)^{-1}|\,\langle\phi_{2T}\psi_T\rangle\,| \\ &\lesssim T^{-2}\begin{cases} d=2 & : \quad T\ln^q T\,, \\ d=3 & : \quad \sqrt{T}\,, \\ d=4 & : \quad \ln T\,, \\ d>4 & : \quad 1\,. \end{cases} \end{aligned} \tag{3.31}$$

We finally appeal to the dyadic decomposition of the T-axis (3.2), which, combined with (3.31), turns into

$$\begin{aligned} |\xi\cdot(A_T-A_{\text{hom}})\xi| &\lesssim \sum_{i=1}^{\infty}\begin{cases} d=2 & : \quad (2^iT)^{-1}\ln^q(2^iT) \\ d=3 & : \quad (2^iT)^{-3/2} \\ d=4 & : \quad (2^iT)^{-2}\ln(2^iT) \\ d>4 & : \quad (2^iT)^{-2} \end{cases} \\ &\lesssim \begin{cases} d=2 & : \quad T^{-1}\ln^q T\,, \\ d=3 & : \quad T^{-3/2} \\ d=4 & : \quad T^{-2}\ln T\,, \\ d>4 & : \quad T^{-2}\,. \end{cases} \end{aligned}$$

This concludes the proof of the theorem.

3.2. **Proof of Corollary 1.** By Steps 6 and 7 in the proof of Theorem 1 and by the definition (3.1) of ψ_T, we learn that

$$\begin{aligned} \left\langle|\nabla\phi_{2T}-\nabla\phi_T|^2\right\rangle &\overset{(3.1)}{=} T^{-2}\left\langle|\nabla\psi_T|^2\right\rangle \\ &\overset{(3.27)\&(3.29)}{\lesssim} \begin{cases} d=3 & : \quad T^{-3/2}, \\ d=4 & : \quad T^{-2}\ln T, \\ d>4 & : \quad T^{-2}. \end{cases} \end{aligned}$$

In particular, $\nabla\phi_T$ is a Cauchy sequence in L^2 in probability. Hence, $\nabla\phi_T$ converges in L^2 to its weak limit $\nabla\phi$, and by a dyadic decomposition of the T-axis the above estimate yields

$$\left\langle|\nabla\phi_T-\nabla\phi|^2\right\rangle \lesssim \begin{cases} d=3 & : \quad T^{-3/2}, \\ d=4 & : \quad T^{-2}\ln T, \\ d>4 & : \quad T^{-2}, \end{cases}$$

which gives the second term of the l. h. s. of (2.7).

Likewise, from Step 7 in the proof of Theorem 1, we learn that

$$\begin{aligned}\left\langle(\phi_{2T}-\phi_T)^2\right\rangle &\overset{(3.1)}{=} T^{-1}\left\langle(\phi_{2T}-\phi_T)\psi_T\right\rangle\\ &\le T^{-1}(\langle|\phi_{2T}\psi_T|\rangle+\langle|\phi_T\psi_T|\rangle)\\ &\overset{(3.29)\&(3.30)}{\lesssim} \begin{cases} d=3 &:\ T^{-1/2},\\ d=4 &:\ T^{-1}\ln T,\\ d>4 &:\ T^{-1},\end{cases}\end{aligned}$$

so that ϕ_T is a Cauchy sequence in L^2 in probability and ϕ_T converges in L^2 to its weak limit ϕ provided by [5, Corollary 1]. In particular, by a dyadic decomposition of the T-axis the above estimate yields

$$\left\langle(\phi_T-\phi)^2\right\rangle = \begin{cases} d=3 &:\ T^{-1/2},\\ d=4 &:\ T^{-1}\ln T,\\ d>4 &:\ T^{-1},\end{cases}$$

which is the first term of the l. h. s. of (2.7). This concludes the proof of the corollary.

4. Proof of the auxiliary lemmas

4.1. **Proof of Lemma 3.** W. l. o. g. we may assume

$$\sum_{i=1}^{\infty}\left\langle\sup_{a_i}\left|\frac{\partial X}{\partial a_i}\right|^2\right\rangle,\sum_{i=1}^{\infty}\left\langle\sup_{a_i}\left|\frac{\partial Y}{\partial a_i}\right|^2\right\rangle < \infty. \tag{4.1}$$

Let Z_n denote the expected value of Z conditioned on $a_1,\cdots,a_n$, that is

$$Z_n(a_1,\cdots,a_n) := \langle Z\,|\,a_1,\cdots,a_n\,\rangle.$$

From [5, (5.2) & (5.3)] in the proof of [5, Lemma 2.3], we learn that

$$\lim_{n\uparrow\infty}\left\langle(Z-Z_n)^2\right\rangle = 0\,,$$

for $Z=X, Z_n=X_n$ and $Z=Y, Z_n=Y_n$, respectively, so that, by Cauchy-Schwarz' inequality in probability,

$$\begin{aligned}\lim_{n\uparrow\infty}\langle X_n\rangle &= \langle X\rangle\,,\\ \lim_{n\uparrow\infty}\langle Y_n\rangle &= \langle Y\rangle\,,\\ \lim_{n\uparrow\infty}\langle X_nY_n\rangle &= \langle XY\rangle\,.\end{aligned}$$

Hence,

$$\begin{aligned}\lim_{n\uparrow\infty}\operatorname{cov}[X_n;Y_n] &= \lim_{n\uparrow\infty}(\langle X_nY_n\rangle-\langle X_n\rangle\langle Y_n\rangle)\\ &= \langle XY\rangle-\langle X\rangle\langle Y\rangle\\ &= \operatorname{cov}[X;Y]\,.\end{aligned}\tag{4.2}$$

Note also that

$$\operatorname{cov}[X_n;Y_n] = \sum_{i=1}^{n}\big(\langle X_iY_i\rangle-\langle X_{i-1}Y_{i-1}\rangle\big) \tag{4.3}$$

with the notation $X_0 = \langle X\rangle$ and $Y_0 = \langle Y\rangle$, so that $\langle X_n\rangle = X_0$ and $\langle Y_n\rangle = Y_0$. Inequality (2.8) then follows from (4.1), (4.2), (4.3), and

$$\langle X_iY_i\rangle - \langle X_{i-1}Y_{i-1}\rangle \;\lesssim\; \left\langle \sup_{a_i}\left|\frac{\partial X}{\partial a_i}\right|^2\right\rangle^{1/2}\left\langle \sup_{a_i}\left|\frac{\partial Y}{\partial a_i}\right|^2\right\rangle^{1/2}, \tag{4.4}$$

that we prove now. By our assumption that $\{a_i\}_{i\in\mathbb{N}}$ are i. i. d., we have

$$\begin{aligned} Z_{i-1}(a_1,\cdots,a_{i-1}) &= \int Z_i(a_1,\cdots,a_{i-1},a_i'')\,\beta(da_i''),\\ \langle X_i(a_1,\cdots,a_i)Y_i(a_1,\cdots,a_i)\rangle &= \left\langle\int X_i(a_1,\cdots,a_{i-1},a_i')Y_i(a_1,\cdots,a_{i-1},a_i')\,\beta(da_i')\right\rangle, \end{aligned}$$

where β denotes the distribution of a_1. Hence we obtain

$$\begin{aligned} &\langle X_iY_i\rangle - \langle X_{i-1}Y_{i-1}\rangle\\ &= \left\langle\int X_i(a_1,\cdots,a_{i-1},a_i')Y_i(a_1,\cdots,a_{i-1},a_i')\,\beta(da_i')\right\rangle\\ &\quad -\left\langle\int X_i(a_1,\cdots,a_{i-1},a_i')\,\beta(da_i')\int Y_i(a_1,\cdots,a_{i-1},a_i'')\,\beta(da_i'')\right\rangle\\ &= \left\langle\int\int \tfrac12\left(X_i(a_1,\cdots,a_{i-1},a_i')-X_i(a_1,\cdots,a_{i-1},a_i'')\right)\right.\\ &\qquad\qquad \left.\times\left(Y_i(a_1,\cdots,a_{i-1},a_i')-Y_i(a_1,\cdots,a_{i-1},a_i'')\right)\,\beta(da_i')\,\beta(da_i'')\right\rangle\\ &\le \left\langle\int\int \tfrac12\left(X_i(a_1,\cdots,a_{i-1},a_i')-X_i(a_1,\cdots,a_{i-1},a_i'')\right)^2\,\beta(da_i')\,\beta(da_i'')\right\rangle^{1/2}\\ &\quad\times\left\langle\int\int \tfrac12\left(Y_i(a_1,\cdots,a_{i-1},a_i')-Y_i(a_1,\cdots,a_{i-1},a_i'')\right)^2\,\beta(da_i')\,\beta(da_i'')\right\rangle^{1/2}. \end{aligned}$$

We then conclude the proof of (4.4) as in the proof of [5, Lemma 2.3].

4.2. **Proof of Lemma 4.** We divide the proof in two main parts and deal with $|z|\le\sqrt{T}$ and $|z|>\sqrt{T}$ separately. The proof relies on the Harnack inequality on graphs. We refer to Zhou [15] for $\mathbb{Z}^d$, and to Delmotte [3] for other graphs. We recall here the easy part of Harnack's inequality (see [3, Proposition 5.3] or [15, Proof of Theorem 3.3, (3.11)]).

Lemma 6 (Harnack's inequality)**.** *Let $a\in\mathcal{A}_{\alpha\beta}$ and $R\gg 1$. If $g:\mathbb{Z}^d\to\mathbb{R}^+$ satisfies*

$$-\nabla^*\cdot A\nabla g(x)\;\le\;0 \tag{4.5}$$

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

$$\sup_{R<|x|\le 2R} g(x)\;\lesssim\;\left(R^{-d}\int_{R/2<|x|\le 4R} g(x)^2dx\right)^{1/2}. \tag{4.6}$$

Step 1. Proof of (2.10) for $|x-y|\le\sqrt{T}$.
Since G_T satisfies

$$-\nabla_x^*\cdot A\nabla_x G_T(x,y)\;=\;-T^{-1}G_T(x,y)\;\le\;0 \tag{4.7}$$

for $|x-y| \gg 1$, one may apply Lemma 6. For $R \gg 1$, we then have

$$\sup_{x:R<|x-y|\le 2R} G_T(x,y) \lesssim \left(R^{-d}\int_{R/2<|x-y|\le 4R} G_T(x,y)^2dx\right)^{1/2}.$$

Combined with [5, Lemma 2.8, (2.21)] for $q=2$ (which is uniform in $T>0$ and $y\in\mathbb{Z}^d$), this yields

$$\sup_{R<|x-y|\le 2R} G_T(x,y) \lesssim R^{2-d},$$

from which we deduce (2.10) for $\sqrt{T}\ge |x-y| \gg 1$. For $|x-y|\sim 1$, we appeal to [5, Proof of Lemma 2.8, (4.4)] with $R\sim 1$ and $q=1$, which yields $\sup_{|x-y|\le R} G_T(x,y) \lesssim 1$ by the discrete L^1-L^∞ estimate.

Step 2. Proof of (2.11) for $|x-y|\le\sqrt{T}$.
Let N be a positive integer such that $2^N\sim\sqrt{T}$ and $2^{-N}\sqrt{T}\gg 1$. For all $i\in\{1,\dots,N\}$, we first show that

$$\left((2^{-i}\sqrt{T})^{-2}\int_{2^{-i-1}\sqrt{T}<|x-y|\le 2^{-i+2}\sqrt{T}} G_T(x,y)^2dx\right)^{1/2} \lesssim i \sim \ln(\frac{\sqrt{T}}{1+2^{-i}\sqrt{T}}). \quad (4.8)$$

Estimate (4.8) follows from the triangle inequality and the BMO estimate of [5, Lemma 2.8 (2.20)] provided we show that

$$\overline{G_T}_{\{|x-y|\le 2^{-i+2}\sqrt{T}\}} \lesssim i, \tag{4.9}$$

where $\overline{G_T}_{\{|x-y|\le 2^{-i+2}\sqrt{T}\}}$ denotes the average of $G_T(x,y)$ on the set $\{|x-y|\le 2^{-i+2}\sqrt{T}\}$. By the triangle inequality and the BMO estimate of [5, Lemma 2.8 (2.20)], we have

$$\begin{aligned}
&\overline{G_T}_{\{|x-y|\le 2^{-i+2}\sqrt{T}\}}\\
&\quad\le \overline{G_T}_{\{|x-y|\le 2^{-i+3}\sqrt{T}\}}\\
&\qquad +2\left(\frac{1}{|\{|x-y|\le 2^{-i+3}\sqrt{T}\}|}\int_{|x-y|\le 2^{-i+3}\sqrt{T}}(G_T(x,y)-\overline{G_T}_{\{|x-y|\le 2^{-i+3}\sqrt{T}\}})^2dx\right)^{1/2}\\
&\quad\le \overline{G_T}_{\{|x-y|\le 2^{-i+3}\sqrt{T}\}}+C,
\end{aligned}$$

where C is a universal constant independent of i. Combined with the estimate for $i=1$

$$\overline{G_T}_{\{|x-y|\le 4\sqrt{T}\}} \lesssim 1,$$

which is a consequence of [5, Lemma 2.8 (2.22)], this implies (4.9) by induction.

We are now in position to prove (2.11) for $|x-y|\le\sqrt{T}$. Since $x\mapsto G_T(x,y)$ satisfies

$$-\nabla_x^*\cdot A\nabla_x G_T(x,y) = -T^{-1}G_T(x,y) \le 0$$

in the annulus $\{x, 2^{-i-1}\sqrt{T}<|x-y|\le 2^{-i+2}\sqrt{T}\}$, Lemma 6 implies

$$\begin{aligned}
\sup_{x:2^{-i}\sqrt{T}<|x-y|\le 2^{-i+1}\sqrt{T}} G_T(x,y) &\lesssim \left((2^{-i}\sqrt{T})^{-2}\int_{2^{-i-1}\sqrt{T}<|x-y|\le 2^{-i+2}\sqrt{T}} G_T(x,y)^2dx\right)^{1/2}\\
&\lesssim \ln(\frac{\sqrt{T}}{1+2^{-i}\sqrt{T}})
\end{aligned}$$

using (4.8) for $2^{-i}\sqrt{T} \gg 1$. For $|x-y| \le R \sim 1$, we appeal to (4.9) and to the discrete $L^1 - L^\infty$ estimate

$$G_T(x,y) \;\le\; R^2\overline{G_T}_{\{|x-y|\le R\}} \;\lesssim\; \ln T.$$

This completes the proof of (2.11) for $|x-y| \le \sqrt{T}$.

Step 3. Proof of (2.10) & (2.11) for $|x-y| > \sqrt{T}$.
Let $R \ge \sqrt{T}$. Since G_T satisfies

$$-\nabla_x^* \cdot A\nabla_x G_T(x,y) \;=\; -T^{-1}G_T(x,y) \;\le\; 0, \qquad \text{for } |x-y| \ge 1,$$

Lemma 6 implies

$$\sup_{x:R<|x-y|\le 2R} G_T(x,y) \;\;\lesssim\;\; \left(R^{-d}\int_{R/2<|x-y|\le 4R} G_T(x,y)^2 dx\right)^{1/2}.$$

Combined with [5, Lemma 2.8, (2.23)] for $q=2$ and $r=k$, i. e.

$$\int_{R/2<|x-y|\le 4R} G_T(x,y)^2 dx \;\;\lesssim\;\; R^{d+(2-d)2}(\sqrt{T}R^{-1})^k,$$

this yields the desired pointwise bound.

4.3. **Proof of Lemma 5.** First note that by symmetry,

$$\int_{|z|\le|z-x|} h_T(z)h_T(z-x)dz = \int_{|z|\ge|z-x|} h_T(z)h_T(z-x)dz \ge \frac{1}{2}\int_{\mathbb{Z}^d} h_T(z)h_T(z-x)dz.$$

Hence, it is enough to consider

$$\int_{\mathbb{Z}^d} g_T(x)\int_{|z|\le|z-x|} h_T(z)h_T(z-x)dzdx.$$

In this proof, we essentially combine the pointwise decay of g_T with the results of [5, Lemma 2.10] that we recall here for the reader's convenience (see [5, Proof of Lemma 2.10, Steps 1, 2 & 4]): There exists $\tilde{R} \sim 1$ such that for all $R \ge \tilde{R}/2$,

$$\int_{R<|x|\le 2R}\int_{|z|\le|z-x|} h_T(z)h_T(z-x)dzdx \lesssim \begin{cases} d=2 & : \quad R^2\max\{1,\ln(\sqrt{T}R^{-1})\}, \\ d>2 & : \quad R^2, \end{cases} \tag{4.10}$$

$$\int_{|x|\le 4\tilde{R}}\int_{|z|\le|z-x|} h_T(z)h_T(z-x)dzdx \lesssim \begin{cases} d=2 & : \quad \ln T, \\ d>2 & : \quad 1. \end{cases} \tag{4.11}$$

In view of (4.11) & (4.10), it will be convenient to make a dyadic decomposition of space. In order to also benefit from the decay of $g_T(x)$ for $|x| \gg \sqrt{T}$, we make the following decomposition of $\mathbb{Z}^d$:

$$\mathbb{Z}^d = \qquad \{|x| \;\le\; 2^{-I}\sqrt{T}\} \tag{4.12}$$

$$\cup \quad \bigsqcup_{i=-I,\dots,-1} \{2^i\sqrt{T} \;<\; |x| \;\le\; 2^{i+1}\sqrt{T}\} \tag{4.13}$$

$$\cup \quad \bigsqcup_{i\in\mathbb{N}} \{2^i\sqrt{T} \;<\; |x| \;\le\; 2^{i+1}\sqrt{T}\}, \tag{4.14}$$

where I is characterized by $2\tilde{R} \;<\; 2^{-I}\sqrt{T} \;\le\; 4\tilde{R}$.

For the integral over the r. h. s. of (4.12), we appeal to (4.11) and to the definitions (2.16) and (2.17) of $g_T(x)$ for $|x| \lesssim 1$:

$$\int_{|x|\leq 2^{-I}\sqrt{T}} g_T(x) \int_{|z|\leq|z-x|} h_T(z)h_T(z-x)dzdx \quad \lesssim \quad \left\{ \begin{array}{lcl} d=2 & : & \ln^2 T, \\ d>2 & : & 1. \end{array} \right. \tag{4.15}$$

For the integral over (4.14), we use this time (4.10) for $R \geq \sqrt{T}$ and the definitions (2.16) and (2.17) of $g_T(x)$ for $|x| \geq \sqrt{T}$, so that for all $i \in \mathbb{N}$ we have

$$\begin{aligned} \int_{2^i\sqrt{T}<|x|\leq 2^{i+1}\sqrt{T}} g_T(x) \int_{|z|\leq|z-x|} h_T(z)h_T(z-x)dzdx \quad &\lesssim \quad (2^i\sqrt{T})^{2-d}(2^i)^{-3}(2^i\sqrt{T})^2 \\ &= \quad \sqrt{T}^{4-d}(2^i)^{1-d}. \end{aligned}$$

Summing this inequality on $i \in \mathbb{N}$ then yields the estimate

$$\int_{\sqrt{T}<|x|} g_T(x) \int_{|z|\leq|z-x|} h_T(z)h_T(z-x)dzdx \quad \lesssim \quad \sqrt{T}^{4-d}. \tag{4.16}$$

We now deal with the integral over the last part (4.13) of $\mathbb{Z}^d$. To this aim, we combine (4.10) for $R \leq \sqrt{T}$ with the definitions (2.16) and (2.17) of $g_T(x)$ for $|x| \leq \sqrt{T}$. In particular, for all $i \in \{-I, \dots, -1\}$, we have

$$\begin{aligned} &\int_{2^i\sqrt{T}<|x|\leq 2^{i+1}\sqrt{T}} g_T(x) \int_{|z|\leq|z-x|} h_T(z)h_T(z-x)dzdx \\ &\lesssim \left\{ \begin{array}{lcl} d=2 & : & \ln(2^{-i})\,(2^i\sqrt{T})^2\ln(2^{-i}) \sim i^2(2^i\sqrt{T})^2, \\ d>2 & : & (2^i\sqrt{T})^2\,(2^i\sqrt{T})^{2-d} = (2^i\sqrt{T})^{4-d}. \end{array} \right. \end{aligned}$$

Summing this inequality over $i \in \{-I, \dots, -1\}$ and using that $2^I \sim \sqrt{T}$ then yield

$$\begin{aligned} &\int_{2^{-I}\sqrt{T}<|x|\leq\sqrt{T}} g_T(x) \int_{|z|\leq|z-x|} h_T(z)h_T(z-x)dzdx \\ &\lesssim \left\{ \begin{array}{lcl} d=2 & : & 1+T\displaystyle\sum_{i=-I}^{-1} i^2 4^i \\ d>2 & : & 1+\sqrt{T}^{4-d}\displaystyle\sum_{i=-I}^{-1}(2^{4-d})^i \end{array} \right. \\ &\overset{2^I\sim\sqrt{T}}{\lesssim} \left\{ \begin{array}{lcl} d=2 & : & T, \\ d=3 & : & \sqrt{T}, \\ d=4 & : & \ln T, \\ d>4 & : & 1. \end{array} \right. \end{aligned} \tag{4.17}$$

The combination of (4.15), (4.16) & (4.17) finally proves (2.18).

Acknowledgments

The authors acknowledge partial support of the Hausdorff Center for Mathematics, Bonn, Germany.

References

[1] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. *Ann. I. H. Poincaré*, 40:153–165, 2005.

[2] 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(3):505–525, 2003.

[3] T. Delmotte. Inégalité de Harnack elliptique sur les graphes. *Colloq. Math.*, 72(1):19–37, 1997.

[4] A. Gloria and F. Otto. Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. In preparation.

[5] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. 2009. Preprint. Available at http://hal.archives-ouvertes.fr/hal-00383953/en/.

[6] 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.*, 40:3647–3679, 2003.

[7] K. Klenke. *Wahrscheinlichkeitstheorie*. Springer-Verlag, Heidelberg-Berlin, 2006. English version appeared as: Probability theory. A comprehensive course, Universitext, Springer-Verlag London, Ltd., London, 2008.

[8] S.M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.

[9] S.M. Kozlov. Averaging of difference schemes. *Math. USSR Sbornik*, 57(2):351–369, 1987.

[10] R. Künnemann. The diffusion limit for reversible jump processes on $\mathbb{Z}^d$ with ergodic random bond conductivities. *Commun. Math. Phys.*, 90:27–68, 1983.

[11] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. Preprint, 1998.

[12] G.C. Papanicolaou and S.R.S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[13] X. Yue and W. E. The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys.*, 222(2):556–572, 2007.

[14] V.V. Yurinskii. Averaging of symmetric diffusion in random medium. *Sibirskii Matematicheskii Zhurnal*, 27(4):167–180, 1986.

[15] X. Y. Zhou. Green function estimates and their applications to the intersections of symmetric random walks. *Stochastic Process. Appl.*, 48(1):31–60, 1993.

(Antoine Gloria) Projet SIMPAF, INRIA Lille-Nord Europe, France
E-mail address: antoine.gloria@inria.fr

(Felix Otto) Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
E-mail address: otto@mis.mpg.de