Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics

(revised version: August 2014)

by

Antoine Gloria, Stefan Neukamm, and Felix Otto

Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics

A. Gloria [*]   S. Neukamm[†]   F. Otto[‡]

August 20, 2013
(revised: March 20, 2014)

Abstract

We study quantitatively the effective large-scale behavior of discrete elliptic equations on the lattice $\mathbb{Z}^d$ with random coefficients. The theory of stochastic homogenization relates the random, stationary, and ergodic field of coefficients with a deterministic matrix of effective coefficients. This is done via the corrector problem, which can be viewed as a highly degenerate elliptic equation on the infinite-dimensional space of admissible coefficient fields. In this contribution we develop new quantitative methods for the corrector problem based on the assumption that ergodicity holds in the quantitative form of a Spectral Gap Estimate w. r. t. a Glauber dynamics on coefficient fields —as it is the case for independent and identically distributed coefficients. As a main result we prove an optimal decay in time of the semigroup associated with the corrector problem (i. e. of the generator of the process called "random environment as seen from the particle"). As a corollary we recover existence of stationary correctors (in dimensions $d > 2$) and prove new optimal estimates for regularized versions of the corrector (in dimensions $d \geq 2$). We also give a self-contained proof of a new estimate on the gradient of the parabolic, variable-coefficient Green's function, which is a crucial analytic ingredient in our approach.

As an application of these results, we prove the first (and optimal) estimates for the approximation of the homogenized coefficients by the popular periodization method in case of independent and identically distributed coefficients.

Contents

[*]Université Libre de Bruxelles (ULB), Brussels, Belgium & Project-team SIMPAF, Inria Lille - Nord Europe, Villeneuve d'Ascq, France, agloria@ulb.ac.be

[†]Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany, stefan.neukamm@wias-berlin.de

[‡]Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany, felix.otto@mis.mpg.de

1   Introduction and main results

In the 70s Yurinskii [41], Kozlov [28], and Papanicolaou & Varadhan [37] proved the first homogenization results for elliptic equations with random coefficients. They considered the elliptic operator $-\nabla \cdot \boldsymbol{a}(\frac{\cdot}{\varepsilon})\nabla$ (on a domain of $\mathbb{R}^d$) with random, uniformly elliptic coefficients $\boldsymbol{a}(x) \in \mathbb{R}^{d \times d}$ and studied its asymptotic behavior in the macroscopic limit $\varepsilon \downarrow 0$. For coefficients that are stationary and ergodic (w. r. t. the *shifts* $\boldsymbol{a}(\cdot) \mapsto \boldsymbol{a}(\cdot + z)$, $z \in \mathbb{R}^d$) they proved a qualitative homogenization result which says that as $\varepsilon \downarrow 0$ the elliptic operator $-\nabla \cdot \boldsymbol{a}(\frac{\cdot}{\varepsilon})\nabla$ almost surely H-converges to the homogenized elliptic operator $-\nabla \cdot \boldsymbol{a}_{\text{hom}}\nabla$. Morally speaking this means that the rapidly oscillating random coefficients $\boldsymbol{a}(\frac{\cdot}{\varepsilon})$ can be replaced in the macroscopic limit by the *homogenized coefficient matrix* $\boldsymbol{a}_{\text{hom}}$, which is deterministic and characterized by a *homogenization formula*. In the present contribution we consider stochastic homogenization in a *discrete* setting where the continuum domain $\mathbb{R}^d$ is replaced by the lattice $\mathbb{Z}^d$. The qualitative homogenization theory is similar to the one in the continuum setting, see Künnemann [30], Kozlov [29]. This problem corresponds to random conductance models for a network of resistors (see Biskup [4] for a recent review).

We are interested in the homogenization formula. To be precise let $d \geq 2$ denote the dimension and $\lambda > 0$ be a constant of ellipticity which is fixed throughout this article. Let $\Omega \subset (\mathbb{R}^{d \times d})^{\mathbb{Z}^d}$ denote the set of admissible coefficient fields, which is defined as the set of all functions from $\mathbb{Z}^d$ to the set of diagonal matrices with entries in $[\lambda, 1]$. We endow Ω with a stationary ensemble $\langle \cdot \rangle$ that describes the statistics of the random coefficients. With $\langle \cdot \rangle$ we associate the symmetric *matrix of homogenized coefficients* $\boldsymbol{a}_{\text{hom}} \in \mathbb{R}^{d \times d}_{\text{sym}}$ via the minimization problem:

$$(1) \qquad \forall e \in \mathbb{R}^d : \qquad e \cdot \boldsymbol{a}_{\text{hom}} e = \inf_{\overline{\varphi}} \langle (e + \nabla \overline{\varphi}) \cdot \boldsymbol{a}(e + \nabla \overline{\varphi}) \rangle,$$

where the infimum runs over all random fields $\overline{\varphi} : \Omega \times \mathbb{Z}^d \to \mathbb{R}$ that are stationary in the sense of $\overline{\varphi}(\boldsymbol{a}, x + z) = \overline{\varphi}(\boldsymbol{a}(\cdot + z), x)$ for all $x, z \in \mathbb{Z}^d$ and $\langle \cdot \rangle$-almost every $\boldsymbol{a} \in \Omega$. The Euler-Lagrange equation associated with (1) is called *the corrector equation*:

$$(2) \qquad \nabla^* \boldsymbol{a}(\nabla \overline{\phi} + e) = 0 \qquad \text{on } \mathbb{Z}^d, \qquad \overline{\phi} \text{ is a } \langle \cdot \rangle\text{-stationary random field.}$$

We refer to Section 2.1 for the definition of the finite-difference gradient ∇ and its adjoint ∇^*. A solution to (2) is called a *stationary corrector*, as opposed to the non-stationary corrector introduced by Künnemann [30]. In case it exists, the stationary corrector minimizes the Dirichlet energy (1), and the *homogenization formula* reads

$$(3) \qquad \boldsymbol{a}_{\text{hom}} e = \langle \boldsymbol{a}(\nabla \overline{\phi} + e) \rangle.$$

The goal of this paper is to introduce quantitative methods that yield **optimal** estimates on the corrector equation (2) and on approximations of the homogenization formula (3). The methods we present continue and extend earlier ideas of two of the authors in [23, 24].

The quantitative theory for (2) is subtle. As a matter of fact, for general stationary and ergodic ensembles the minimum in (1) may not be attained and stationary correctors may not exist. On top of that, in dimension $d = 2$ stationary correctors do not exist even under the strong assumption of independent and identically distributed (i. i. d.) coefficients. The only existence result of stationary

2

correctors has been obtained recently in [23] by two of the authors in the case of i. i. d. coefficients in dimensions $d > 2$.

Deterministic approaches to (2) that do not exploit the properties of the underlying probability space (e. g. when $\boldsymbol{a} \in \Omega$ is just viewed as a parameter) are too narrow: One generically does not have existence of *bounded* solutions of (2) pointwise in $\boldsymbol{a} \in \Omega$ (as we learn from the small ellipticity-contrast expansion). A natural way to benefit from the underlying probability space relies on the observation that stationary fields $\overline{\varphi}(\boldsymbol{a}, x)$ are fully characterized by the random variable $\varphi(\boldsymbol{a}) := \overline{\varphi}(\boldsymbol{a}, x = 0)$. Based on this, one can introduce a differential calculus for random variables and rewrite (2) as

$$(4) \qquad\qquad D^* \boldsymbol{a}(0)(D\phi + e) = 0 \qquad \phi \text{ is } \langle \cdot \rangle\text{-measurable.}$$

Here D and D^* are what we call *horizontal derivatives* and related to ∇ and ∇^* via stationarity (see Section 2.1 for the details). Moreover, $\boldsymbol{a}(0)$ stands for the coordinate projection $\Omega \ni \boldsymbol{a} \mapsto \boldsymbol{a}(0)$ on diagonal matrices. Thanks to the discrete setting, $D^* \boldsymbol{a}(0) D$ is a bounded linear operator on $L^p(\Omega)$ for any $1 \leq p \leq \infty$. However, it is highly degenerate as an elliptic operator since the horizontal derivative $D = (D_1, \ldots, D_d)$ with its d components typically does not yield a Poincaré inequality for functions on the infinite-dimensional space Ω. Our quantitative analysis is based on the assumption that $\langle \cdot \rangle$ satisfies a *Spectral Gap Estimate* (SG) with respect to a Glauber dynamics on coefficient fields: We assume there exists $\rho > 0$ such that

$$(\text{SG}) \qquad\qquad \forall \zeta \in L^2(\Omega) : \qquad \left\langle (\zeta - \langle \zeta \rangle)^2 \right\rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \left\langle \left(\frac{\partial \zeta}{\partial y} \right)^2 \right\rangle,$$

where $\frac{\partial \zeta}{\partial y}$ denotes the *vertical derivative* and can be viewed as a discrete version of the classical partial derivative $\frac{\partial \zeta(\boldsymbol{a})}{\partial \boldsymbol{a}(y)}$ (see Section 2.1 for details). Note that (SG) can be seen as a Poincaré inequality, this time with respect to the vertical derivatives $\{\frac{\partial}{\partial y}\}_{y \in \mathbb{Z}^d}$. (SG) is stronger than ergodicity, but weaker than the assumption that the coefficients are i. i. d. (cf. Lemma 1 below).

To circumvent the degeneracy of $D^* \boldsymbol{a}(0) D$ one usually considers approximations of (4) that regularize the problem. In this paper we study two natural approximations, which we introduce now.

The modified corrector equation. In the qualitative works by Papanicolaou & Varadhan [37] and Kipnis & Varadhan [27], the corrector equation is regularized by introducing an additional 0th-order ("massive") term. The result is what we call the *modified corrector equation*

$$(5) \qquad\qquad \mu \phi_\mu + D^* \boldsymbol{a}(0)(D\phi_\mu + e) = 0 \qquad \phi_\mu \text{ is } \langle \cdot \rangle\text{-measurable.}$$

Here, the regularization parameter μ is a positive number, sometimes written as an inverse time $\frac{1}{T}$. As a merit of the regularization, the modified corrector equation always admits a unique solution in $L^2(\Omega)$. By standard arguments, $\langle |D\phi_\mu|^2 \rangle$ and $\mu \langle \phi_\mu^2 \rangle$ are bounded uniformly in μ. For general ensembles, bounds on ϕ_μ that are uniform in μ are not available, and thus one can only pass to the limit $\mu \downarrow 0$ on the level of the gradient $D\phi_\mu$, which is enough for the qualitative homogenization theory. However, it is not sufficient for proving existence and estimates for the original problem (4). As an application of our methods, assuming (SG), we shall prove in Proposition 1 that in dimensions $d > 2$ the modified corrector ϕ_μ is bounded in $L^p(\Omega)$ uniformly in μ for every $p < \infty$. In the case of the critical dimension $d = 2$ we obtain the estimate $\langle \phi_\mu^2 \rangle \lesssim (\ln \mu^{-1})|e|^2$, the scaling in μ of which is optimal.

The periodic approximation. Another natural approach to approximate (4) is to replace $\langle \cdot \rangle$ by a stationary ensemble $\langle \cdot \rangle_L$ that concentrates on L-periodic coefficient fields. In that case we can unambiguously find a solution ϕ of (4) (with $\langle \cdot \rangle$ replaced by $\langle \cdot \rangle_L$) by solving for all L-periodic coefficient fields $\boldsymbol{a}$ the *periodic corrector equation* on $\mathbb{Z}^d$:

$$(6) \qquad \nabla^* \boldsymbol{a}(x)(\nabla \overline{\phi}(\boldsymbol{a}, x) + e) = 0, \qquad \overline{\phi}(\boldsymbol{a}, \cdot) \text{ is } L\text{-periodic and} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \overline{\phi}(\boldsymbol{a}, x) = 0.$$

Standard arguments involving the Poincaré estimate on $([0, L) \cap \mathbb{Z})^d$ only show that $\langle |D\phi|^2 \rangle_L$ and $L^{-2}\langle \phi^2 \rangle_L$ are bounded uniformly in L. Assuming an L-periodic version of (SG), we shall prove that $\langle \phi^p \rangle_L \lesssim |e|^p$, $p < \infty$, in dimensions $d > 2$, and that $\langle \phi^2 \rangle_L \lesssim (\ln L)|e|^2$ in dimension $d = 2$ (see Proposition 1). Again the latter is optimal in terms of scaling in L.

A parabolic key estimate. We study both the modified and the periodic corrector equations by a unified approach that is based on the *parabolic* equation

$$(7) \qquad \begin{cases} \partial_t u(t) + D^* \boldsymbol{a}(0) D u(t) = 0 & t > 0, \\ \qquad\qquad\qquad u(t = 0) = \zeta. \end{cases}$$

This equation defines a stochastic process on the space Ω of coefficient fields that can be conveniently interpreted in the context of random walks in random environments (see e.g. [5]): Consider a continuous-in-time random walker (refered to as a particle) on $\mathbb{Z}^d$ whose symmetric jump rate across a bond $\{x, x + e_i\}$ is given by $\boldsymbol{a}_{ii}(x)$. Then the above process describes the "environment $\boldsymbol{a}$ as seen from the particle".

The solution of (7) is unique and given by $u(t) = \exp(-tD^*\boldsymbol{a}(0)D)\zeta$, where $t \mapsto \exp(-tD^*\boldsymbol{a}(0)D)$ denotes the group associated with $D^*\boldsymbol{a}(0)D : L^2(\Omega) \to L^2(\Omega)$. The parabolic equation (7) and the elliptic equation (4) are related by means of the formal identity

$$\phi = \int_0^\infty u(t)\, dt,$$

where for $u(t)$ we choose the initial condition $\zeta = -D^*\boldsymbol{a}(0)e$. This formal relation becomes rigorous (and yields estimates on ϕ) as soon as we have suitable decay estimates on $u(t)$ in t.

Ergodicity of the ensemble implies that $\langle u^2(t) \rangle \to 0$ as $t \uparrow \infty$. (Indeed, since $D^*\boldsymbol{a}(0)D$ is non-negative, $u(t)$ converges to the $L^2(\Omega)$-orthogonal projection of $u(t = 0)$ onto the kernel of $D^*\boldsymbol{a}(0)D$, which by ergodicity only consists of the constant functions). While ergodicity does not yield any rate of convergence, we prove that (SG) yields an algebraic rate, and thus quantifies ergodicity. As a main result we prove an *optimal decay estimate* for $\exp\big(- tD^*\boldsymbol{a}(0)D \big)$ with initial conditions in divergence form.

Theorem 1. *Assume that*

$$(8) \qquad \langle \cdot \rangle \text{ is stationary, } L\text{-periodic and satisfies } SG_L(\rho) \text{ (see Definition 1 below)}.$$

Then there exists an exponent $1 \le p_0 < \infty$ that only depends on the constant of ellipticity λ and the dimension $d \ge 2$ such that the following statement holds for every $p \in (p_0, \infty)$:
For $\xi \in C_b(\Omega)^d$ and $t \ge 0$ define

$$u(t) := \exp(-tD^*\boldsymbol{a}(0)D)D^*\xi.$$

Then we have

$$(9) \qquad \langle |u(t)|^{2p} \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})} \|\partial\xi\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}},$$

where

$$(10) \qquad \|\partial\xi\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}} := \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left| \frac{\partial\xi}{\partial_L y} \right|^{2p} \right\rangle^{\frac{1}{2p}},$$

and $\lesssim$ means $\le$ up to a multiplicative constant that only depends on the integrability exponent p, the spectral gap constant ρ, the constant of ellipticity λ, and the dimension d.

A similar result that is formally obtained by letting $L \uparrow \infty$ holds for the whole-space case, under the assumption

$$(11) \qquad \langle \cdot \rangle \text{ is stationary and satisfies } SG_\infty(\rho) \text{ (see Definition 1 below)}.$$

The exponent $\frac{d}{4}+\frac{1}{2}$ in (9) is optimal, since in the case of vanishing ellipticity contrast (that is, when $\boldsymbol{a}$ is a perturbation of the identity), one may replace at first order the elliptic operator by the discrete Laplacian, in which case

$$\langle |\exp(-tD^*D)D^*\xi|^2\rangle^{\frac{1}{2}} \sim \left(\sup_{x\in\mathbb{Z}^d} |\nabla^2 G(2t,x)|\right)^{\frac{1}{2}} \|\partial\xi\|_{\ell_y^1 L_{\langle\cdot\rangle}^2} \sim (t+1)^{-(\frac{d}{4}+\frac{1}{2})}\|\partial\xi\|_{\ell_y^1 L_{\langle\cdot\rangle}^2}$$

for $\langle\cdot\rangle$ i. i. d. and initial values $\xi(\boldsymbol{a})$ that depend on $\boldsymbol{a}$ only through $\boldsymbol{a}(0)$. Here $\nabla^2 G$ denotes the Hessian of the discrete heat kernel on $\mathbb{Z}^d$.

As a corollary of Theorem 1 we obtain in Section 4 several optimal estimates for the stationary corrector, the modified corrector and the periodic approximation of the corrector. In addition, we obtain optimal estimates of the error introduced by adding the 0th order term in the modified corrector equation. Based on these estimates and spectral theory we derive error estimates for the approximation of the homogenized coefficients by periodization, which we shall present now.

Error estimate for approximations of $\boldsymbol{a}_{\text{hom}}$. In this paper we provide an optimal error estimate for the approximation by periodization of the homogenized coefficients associated with an i. i. d. ensemble. Approximation by periodization is widely used in the mechanical community and also called the "representative volume element method", see [36] for a qualitative analysis in the stationary ergodic case. The homogenized coefficients $\boldsymbol{a}_{\text{hom}}$ associated with an ergodic ensemble $\langle\cdot\rangle$ are approximated by the homogenized coefficients $\boldsymbol{a}_{\text{hom},L}$ associated with a stationary L-periodic ensemble $\langle\cdot\rangle_L$, which we think of as having the same statistical specifications as $\langle\cdot\rangle$. Since $\boldsymbol{a}_{\text{hom},L}$ is still not computable in practice, one replaces $\boldsymbol{a}_{\text{hom},L}$ by a spatial average $\boldsymbol{a}_{\text{av},L}(\boldsymbol{a})$ where $\boldsymbol{a}$ is distributed according to $\langle\cdot\rangle_L$ and $\boldsymbol{a}_{\text{av},L}:\Omega\to\mathbb{R}^{d\times d}_{\text{sym}}$ is given by

$$(12) \qquad e\cdot\boldsymbol{a}_{\text{av},L}(\boldsymbol{a})e := \min_{\varphi\ L\text{-periodic}} \frac{1}{L^d}\sum_{x\in([0,L)\cap\mathbb{Z})^d}(\nabla\varphi(x)+e)\cdot\boldsymbol{a}(x)(\nabla\varphi(x)+e).$$

As an application of our method we quantify the speed of convergence when $\langle\cdot\rangle$ is an i. i. d. ensemble. In that case it is natural to define $\langle\cdot\rangle_L$ as the L-periodic i. i. d. ensemble associated with $\langle\cdot\rangle$ (cf. Definition 3). We estimate the mean square error $\langle |\boldsymbol{a}_{\text{av},L}-\boldsymbol{a}_{\text{hom}}|^2\rangle_L$, which naturally splits into two parts:

$$(13) \qquad \begin{aligned} \langle |\boldsymbol{a}_{\text{av},L}-\boldsymbol{a}_{\text{hom}}|^2\rangle_L &= \text{var}_L[\boldsymbol{a}_{\text{av},L}] + |\langle\boldsymbol{a}_{\text{av},L}\rangle_L - \boldsymbol{a}_{\text{hom}}|^2 \\ &= \langle |\boldsymbol{a}_{\text{av},L}-\langle\boldsymbol{a}_{\text{av},L}\rangle_L|^2\rangle_L + |\boldsymbol{a}_{\text{hom},L}-\boldsymbol{a}_{\text{hom}}|^2 \\ &= (\text{random error})^2 + (\text{systematic error})^2, \end{aligned}$$

where var_L denotes the variance w. r. t. the ensemble $\langle\cdot\rangle_L$. Following the terminology of [23], we call the (square roots of the) first and second terms of the RHS the *random error* and *systematic error*, respectively. The *random error* is due to the lack of ergodicity of the L-periodic ensemble and measures the fluctuation of $\boldsymbol{a}_{\text{av},L}$ around its average. Although $\boldsymbol{a}_{\text{av},L}$ is a spatial average of highly correlated random variables, we obtain in Proposition 2 (under the weaker assumption that $\langle\cdot\rangle_L$ satisfies an L-periodic version of (SG)) that the random error decays with the rate of the central limit theorem, i. e.

$$(14) \qquad \text{var}_L[\boldsymbol{a}_{\text{av},L}] \lesssim L^{-d}.$$

The *systematic error* is of different nature: By replacing the infinite ensemble $\langle\cdot\rangle$ by the L-periodic ensemble $\langle\cdot\rangle_L$, we introduce artificial long-range correlations (in order to enforce the periodicity of $\langle\cdot\rangle_L$). This produces the *systematic error*. In contrast to the *systematic error*, the effect of the *random error* can be reduced by empirical averaging: For $N\in\mathbb{N}$ define the random matrix

$$(15) \qquad \boldsymbol{a}_{\text{av},L,N} := \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{a}_{\text{av},L}(\boldsymbol{a}^i),$$

where $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^N$ denote N independent realizations distributed according to $\langle \cdot \rangle_L$. This is of particular interest since the systematic error is indeed much smaller than the random error. With the help of Theorem 1 we obtain the following optimal error estimate:

Theorem 2. *Let $d \geq 2$, $L, N \in \mathbb{N}$. Let $\langle \cdot \rangle$ be an infinite i. i. d. ensemble and let $\langle \cdot \rangle_L$ be the associated L-periodic i. i. d. ensemble. Then we have*

$$\langle |\boldsymbol{a}_{\mathrm{av},L,N} - \boldsymbol{a}_{\mathrm{hom}}|^2 \rangle_L^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{N}} L^{-\frac{d}{2}} + L^{-d} \ln^d L,$$

where $\lesssim$ means $\leq$ up to a constant that only depends on λ and d.

The proof of Theorem 2 is given in Section 5 and uses new estimates on the modified corrector equation. For a discussion of the literature on error estimates for the approximation of homogenized coefficients (and in particular [42], [6], and [13]), we refer the reader to [23, Section 1.2] and [19]. For a review on applications of the approximation by periodization, see [26]. Theorem 2 is the first quantitative result on this method in a stochastic (yet discrete and scalar) setting. For associated numerical results, we refer the reader to [14].

Estimates of the Gradient of the parabolic Green's function.   An important observation in the proof of Theorem 1 is that the vertical derivatives $\frac{\partial u(t)}{\partial y}$ (with $u(t)$ solution of the parabolic problem (7)) can be characterized as the solution of a parabolic equation whose RHS involves $\nabla u(t)$. Hence, it can be represented using the parabolic *Green's function G* associated with the parabolic operator $\partial_t + \nabla^* \boldsymbol{a} \nabla$. The resulting Duhamel formula is a nonlinear identity that involves the gradient of the parabolic Green's function. The quantitative statements we are interested in require estimates on the *gradient* of the parabolic Green's function $G(t, \boldsymbol{a}, x, y)$. For our purposes we need estimates that are *uniform* in $\boldsymbol{a}$, but nevertheless, are optimal in terms of the exponent in t. By optimal we mean that the exponent should be identical to the one of the constant-coefficient Green's function. As a consequence, these estimates cannot be *pointwise* in x, but rather integral estimates. In particular we shall need to capture the decay in x in a better way than the following estimate does: For all $t \geq 0$

$$\left(\sum_{z \in \mathbb{Z}^d} |\nabla G(t, \boldsymbol{a}, z, 0)|^2 \right)^{\frac{1}{2}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})}.$$

We do this by establishing *weighted* integral estimates with weight functions

$$(16) \qquad \omega(t, x) := \left(\frac{|x|^2}{t+1} + 1 \right)^{\frac{1}{2}}, \qquad \omega_L(t, x) := \left(\frac{\mathrm{dist}^2(x, L\mathbb{Z}^d)}{t+1} + 1 \right)^{\frac{1}{2}},$$

where $\mathrm{dist}(x, L\mathbb{Z}^d) := \min_{z \in \mathbb{Z}^d} |x - Lz|$ denotes the distance to 0 on the L-torus. Finally, in order to treat the nonlinear term, we shall need a slightly stronger estimate than the *square* integral estimate. In Section 6 we establish the following estimates for the discrete, whole-space parabolic Green's function G and the discrete, L-periodic parabolic Green's function G_L.

Theorem 3. *There exists an exponent $q_0 > 1$ (only depending on λ and d) such that for all $1 \leq q < q_0$ and all $\alpha < \infty$ we have: For all $L \in \mathbb{N}$ and $\boldsymbol{a} \in \Omega_L$*

$$\left(\sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left(\omega_L^\alpha(t, x-y) |\nabla_x G_L(t, \boldsymbol{a}, x, y)| \right)^{2q} \right)^{\frac{1}{2q}} \lesssim (t+1)^{-(\frac{d}{2} + \frac{1}{2}) + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2} \right),$$

where the multiplicative constant only depends on α, λ, d, and q, and $c_0 > 0$ denotes a constant that only depends on λ and d.
A similar result that is formally obtained by letting $L \uparrow \infty$ holds for the whole-space Green's function.

The estimates of Theorem 3 are optimal in terms of scaling. Indeed, the exponents are the same for the constant coefficient case (i. e. $\boldsymbol{a} = \boldsymbol{id}$) as can be seen by a scaling argument in the continuum case (i. e. $\mathbb{Z}^d$ replaced by $\mathbb{R}^d$). In the proof we make extensive use of the elliptic and parabolic regularity theory by Nash [35] and Meyers [31]. The methods we use rely in particular on the maximum principle, which is the reason why we restrict ourselves, in our discrete setting, to the case where $\boldsymbol{a}$ is diagonal.

Structure of the paper. In Section 2 we introduce the general framework, in particular a discrete differential calculus for functions on the lattice $\mathbb{Z}^d$, a "horizontal" differential calculus for random variables, the description of random coefficients via the ensemble, and introduce our main assumption on the ensemble, namely the Spectral Gap Estimate (SG). In Section 3 we prove Theorem 1, which is the main result of the paper. In Section 4 we apply Theorem 1 to the corrector equation in stochastic homogenization. In Section 5 we prove Theorem 2 which establishes an optimal error estimate for the approximation of the homogenized coefficients by periodization. Finally, in Section 6 we prove the quenched estimates on the gradient of the parabolic Green's function of Theorem 3.

Relation to previous works. The first paper on quantitative stochastic homogenization is due to Yurinskii [42]. Yurinskii considered the dependence of the gradient of the modified corrector $D\phi_\mu$ on μ. For dimension $d > 2$ and under mixing assumptions on the statistics, he obtained non-optimal decay estimates in μ on the $L^2(\Omega)$-distance between the gradient $D\phi_\mu$ of the modified corrector and that of the original corrector ϕ. As a central ingredient, Yurinskii appeals to optimal deterministic estimates on the *parabolic* variable-coefficients Green's function.

The idea of combining stochastic homogenization methods with statistical mechanics methods naturally arises in the study of scaling limits of Gradient Gibbs Measures. Gradient Gibbs Measures can be seen as a model for thermally fluctuating interfaces, as introduced in the mathematical literature by Funaki and Spohn [17], see [16] for a review. Naddaf & Spencer [33] were the first to combine all the three concepts of (discrete) spatial, of horizontal, and of vertical derivatives. They were also the first to use methods of statistical mechanics in stochastic homogenization, cf. the very inspiring unpublished preprint [34]. A main ingredient in [34] are deterministic estimates on the *gradient* of the *elliptic* variable-coefficients Green's function — as opposed to the estimates on the *parabolic* Green's function itself by Yurinskii. Implicitly, Naddaf and Spencer obtain deterministic ℓ^{2q}-estimates (for some q slightly larger than 1) via Meyers' argument. They use it to establish a variance estimate in the spirit of (14) in the case of small ellipticity contrast (that is, when $q \geq 2$). This approach has been further developed by Conlon and Naddaf [7] to obtain estimates on the annealed elliptic and parabolic Green's functions (that is, on the expectation of these Green's functions), and by Conlon and Spencer [8] to quantify the homogenization error, for small ellipticity contrast. Some optimal annealed estimates on the first and mixed second derivatives of the Green's function (that is, pointwise-in-space estimates of moments of order $2q$ in probability) have also been obtained by Delmotte & Deuschel [10] for any ellipticity constrast (yet for $q = 1$ for the first derivative, and $q = \frac{1}{2}$ for the mixed second derivative only).

Another source of inspiring ideas are the works on qualitative stochastic homogenization by Papanicolaou & Varadhan [37] and Kipnis & Varadhan [27], where the modified corrector problem (5) is introduced. The modified corrector yields an approximation $\boldsymbol{a}_{\mathrm{hom},\mu}$ for the homogenized coefficients $\boldsymbol{a}_{\mathrm{hom}}$, see Definition 4 below. In [37, 27] the authors appeal to spectral analysis to treat the original corrector problem (2) with the help of its modified version (5). Furthermore, they devise a spectral representation formula for the homogenized coefficients and its approximation $\boldsymbol{a}_{\mathrm{hom},\mu}$, which we use to develop quantitative estimates on the error $|\boldsymbol{a}_{\mathrm{hom},\mu} - \boldsymbol{a}_{\mathrm{hom}}|$ via Theorem 1.

In [23] and [24], the first and third authors obtained the first optimal quantitative results on the corrector equation, namely the boundedness of correctors, the optimal estimate of the random error, and optimal bounds on the systematic error $|\boldsymbol{a}_{\mathrm{hom},\mu} - \boldsymbol{a}_{\mathrm{hom}}|$ for $d > 2$. In these contributions, an auxiliary result are optimal deterministic estimates on the *gradient* of the variable-coefficients *elliptic* Green's function. In [32], Mourrat obtained a suboptimal version of Theorem 1 for $d \geq 5$ by a

different approach. Then, Mourrat and the first author [20] obtained further quantitative results for the systematic error by appealing to the spectral calculus mentioned above. In particular, in [20], motivated by the spectral representation of $\boldsymbol{a}_{\text{hom}}$, a different "higher order" approximation scheme by extrapolation for $\boldsymbol{a}_{\text{hom}}$ is introduced, and optimal error estimates for the associated systematic error are obtained for dimensions $2 < d \leq 6$, still using the deterministic estimates on the *gradient* of the *elliptic* Green's function. We use this scheme in our analysis of the approximation by periodization.

The interest of the present contribution is threefold. First we obtain optimal results on the corrector and on the systematic error in *any* dimension (and in particular for $d = 2$ and $d > 6$), and give the first complete and optimal quantitative analysis of the very popular approximation by periodization. Second, we introduce a unified method which quantifies optimally the ergodicity of the environment as seen from the particle, result from which all the other estimates follow. Last, the main auxiliary result are new deterministic estimates on the *gradient* of the *parabolic* variable-coefficients Green's function.

This article is the short version of lecture notes [21], which contain in addition detailed proofs of several classical results which are recalled here without a proof.

2 Differential calculus on stationary random fields

In this section, we introduce two differential calculi on stationary random fields, that we call horizontal and vertical.

2.1 Spatial derivatives, stationarity, and horizontal derivatives

Random coefficient fields. We consider linear second-order difference equations with uniformly elliptic, bounded, diagonal random coefficients. We denote the set of admissible coefficient matrices by

$$(17) \qquad \Omega_0 := \left\{ \operatorname{diag}(a_1, \ldots, a_d) \in \mathbb{R}^{d \times d} : \lambda \leq a_j \leq 1 \text{ for } j = 1, \ldots, d \right\},$$

where $\operatorname{diag}(a_1, \ldots, a_d)$ is the diagonal, $d \times d$ matrix with entries $a_1, \ldots, a_d$, and $\lambda > 0$ is an ellipticity constant which is fixed throughout the paper. We equip Ω_0 with the usual topology of $\mathbb{R}^{d \times d}$. A *coefficient field*, denoted by $\boldsymbol{a}$, is a function on $\mathbb{Z}^d$ taking values in Ω_0. We endow $\Omega = (\Omega_0)^{\mathbb{Z}^d}$ with the product topology. For $L \in \mathbb{N}$ we denote by Ω_L the subspace of L-periodic conductivity fields (i. e. $\boldsymbol{a} \in \Omega$ with $\boldsymbol{a}(x + Lz) = \boldsymbol{a}(x)$ for all $x, z \in \mathbb{Z}^d$).

We describe a *random coefficient field* by equipping Ω with a probability measure. Following the convention in statistical mechanics, we call a probability measure on Ω also an *ensemble* and denote the associated ensemble average by $\langle \cdot \rangle$. If $\langle \cdot \rangle$ concentrates on Ω_L we call it an *L-periodic ensemble*.

Unless otherwise stated we always assume that $\langle \cdot \rangle$ is *stationary*. Let $T_z : \Omega \to \Omega$, $\boldsymbol{a}(\cdot) \mapsto \boldsymbol{a}(\cdot + z)$ denote the shift by z. Then $\langle \cdot \rangle$ is stationary if and only if T_z is $\langle \cdot \rangle$-preserving for all shifts $z \in \mathbb{Z}^d$.

Random variables and stationary random fields. A random variable is a measurable function on Ω. We denote by $L^p(\Omega)$, $1 \leq p \leq \infty$, the usual spaces of random variables with finite p-th moment. We denote by $C_b(\Omega)$ the space of bounded continuous functions on Ω equipped with the norm $\|\zeta\|_\infty := \sup_{\boldsymbol{a} \in \Omega} |\zeta(\boldsymbol{a})| < \infty$.

A *random field* $\tilde{\zeta}$ is a measurable function on $\Omega \times \mathbb{Z}^d$. With any random variable $\zeta : \Omega \to \mathbb{R}$ we associate its *$\langle \cdot \rangle$-stationary extension* $\overline{\zeta} : \Omega \times \mathbb{Z}^d \to \mathbb{R}$ via $\overline{\zeta}(\boldsymbol{a}, x) := \zeta(\boldsymbol{a}(\cdot + x))$. Conversely, we say that a random field $\tilde{\zeta}$ is *$\langle \cdot \rangle$-stationary* if there exists a random variable ζ with $\tilde{\zeta}(\boldsymbol{a}, \cdot) = \overline{\zeta}(\boldsymbol{a}, \cdot)$ $\langle \cdot \rangle$-almost surely. Since $\langle \overline{\zeta}(x) \rangle$ does not depend on x by stationarity, we simply write $\langle \overline{\zeta} \rangle$ for the expectation of a stationary field $\overline{\zeta}$.

8

Spatial and horizontal derivatives. For scalar fields $\zeta : \mathbb{Z}^d \to \mathbb{R}$, vector fields $\xi = (\xi_1, \ldots, \xi_d) : \mathbb{Z}^d \to \mathbb{R}^d$, and all $i = 1, \ldots, d$, we define the *spatial derivatives*

$$\nabla_i \zeta(x) := \zeta(x + e_i) - \zeta(x), \quad \nabla_i^* \zeta(x) := \zeta(x - e_i) - \zeta(x),$$

$$\nabla \zeta := (\nabla_1 \zeta, \ldots, \nabla_d \zeta), \qquad \nabla^* \xi := \sum_{i=1}^d \nabla_i^* \xi_i,$$

where $\{e_1, \ldots, e_d\}$ denotes the canonical basis of $\mathbb{R}^d$, ∇ is the discrete gradient for functions on $\mathbb{Z}^d$, and $-\nabla^*$ is the discrete divergence for vector fields on $\mathbb{Z}^d$. By discreteness ∇_i and ∇_i^* are bounded linear operators on $\ell^p(\mathbb{Z}^d)$, $1 \leq p \leq \infty$. They are formally adjoint, as are ∇ and ∇^*.

Next, we recall a similar standard structure for random variables: For scalar random variables $\zeta : \Omega \to \mathbb{R}$, vector-valued random variables $\xi = (\xi_1, \ldots \xi_d) : \Omega \to \mathbb{R}^d$, and $i = 1, \ldots, d$, we define the *horizontal derivatives*

$$D_i \zeta(\boldsymbol{a}) := \zeta(\boldsymbol{a}(\cdot + e_i)) - \zeta(\boldsymbol{a}), \quad D_i^* \zeta(\boldsymbol{a}) := \zeta(\boldsymbol{a}(\cdot - e_i)) - \zeta(\boldsymbol{a}),$$

$$D\zeta := (D_1 \zeta, \ldots, D_d \zeta), \qquad D^* \xi := \sum_{i=1}^d D_i^* \xi_i.$$

The horizontal derivatives D_i and D_i^* are bounded linear operators on $L^p(\Omega)$, $1 \leq p \leq \infty$, and $C_b(\Omega)$. They are formally adjoint, as are D and D^*. An elementary but important observation is the following. Let $\overline{(\cdot)}$ denote the mapping that associates a random variable with its stationary extension, see above. Then

(18) $$\nabla_i \overline{\zeta} = \overline{D_i \zeta}, \qquad \nabla_i^* \overline{\zeta} = \overline{D_i^* \zeta}, \qquad \text{and} \qquad \nabla^* \boldsymbol{a} \nabla \overline{\zeta} = \overline{D^* \boldsymbol{a}(0) D \zeta}.$$

We use $\lesssim$ (resp. $\gtrsim$) for $\leq$ (resp. $\geq$) up to a multiplicative constant whose dependence on the different parameters is made explicit in each statement.

2.2 Spectral gap on Glauber dynamics and vertical derivatives

Definition 1 (spectral gap & vertical derivatives). *We say that $\langle \cdot \rangle$ satisfies a spectral gap for the Glauber dynamics with constant $\rho > 0$, in short $SG_\infty(\rho)$, if for all $\zeta \in L^2(\Omega)$ with $\langle \zeta \rangle = 0$ we have*

(19) $$\langle \zeta^2 \rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \left\langle \left(\frac{\partial \zeta}{\partial y} \right)^2 \right\rangle,$$

where the vertical derivative $\frac{\partial \zeta}{\partial y}$ of ζ w. r. t. y is defined as follows. For all $\zeta \in L^2(\Omega)$ and $y \in \mathbb{Z}^d$,

$$\frac{\partial \zeta}{\partial y} := \zeta - \langle \zeta \rangle_y, \qquad where \qquad \langle \zeta \rangle_y := \left\langle \zeta \,\Big|\, \{\boldsymbol{a}(x)\}_{x \in \mathbb{Z}^d \setminus \{y\}} \right\rangle,$$

i. e. $\langle \zeta \rangle_y$ is the conditional expectation of ζ given $\boldsymbol{a}(x)$ for all $x \in \mathbb{Z}^d \setminus \{y\}$. More analytically, $\langle \zeta \rangle_y$ is the $L^2(\Omega)$-orthogonal projection of ζ on the space of functions of $\boldsymbol{a}$ that do not depend on $\boldsymbol{a}(y)$.

The L-periodic version of (SG) is the following:

Definition 2 (spectral gap & vertical derivatives, periodic case). *We say that $\langle \cdot \rangle$ satisfies a spectral gap with constant $\rho > 0$ on the torus of size $L \in \mathbb{N}$, in short $SG_L(\rho)$, if for all $\zeta \in L^2(\Omega)$ with $\langle \zeta \rangle = 0$ we have*

$$\langle \zeta^2 \rangle \leq \frac{1}{\rho} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left(\frac{\partial \zeta}{\partial_L y} \right)^2 \right\rangle,$$

9

where the (L-periodic) vertical derivative $\frac{\partial \zeta}{\partial_L y}$ of ζ w. r. t. y is defined as follows. For all $\zeta \in L^2(\Omega)$ and $y \in \mathbb{Z}^d$,

$$\frac{\partial \zeta}{\partial_L y} := \zeta - \langle \zeta \rangle_{L,y}, \qquad where \qquad \langle \zeta \rangle_{L,y} := \left\langle \zeta \,\Big|\, \{\boldsymbol{a}(x)\}_{x \in \mathbb{Z}^d \setminus \{y+L\mathbb{Z}^d\}} \right\rangle,$$

i. e. $\langle \zeta \rangle_{L,y}$ is the conditional expectation of ζ given $\boldsymbol{a}(x)$ for all $x \in \mathbb{Z}^d \setminus \{y + L\mathbb{Z}^d\}$.

The vertical derivative does not commute with the shift $T_x : \Omega \to \Omega$, $\boldsymbol{a}(\cdot) \mapsto \boldsymbol{a}(\cdot + x)$. We have $\langle \zeta \circ T_x \rangle_y = \langle \zeta \rangle_{y-x} \circ T_x$, which in terms of the stationary extension $\overline{(\cdot)}$ reads

$$(20) \qquad\qquad \left\langle \overline{\zeta}(x) \right\rangle_y = \overline{\langle \zeta \rangle_{y-x}}(x).$$

Hence,

$$(21) \qquad\qquad \frac{\partial \overline{\zeta}(x)}{\partial y} = \overline{\left(\frac{\partial \zeta}{\partial(y-x)} \right)}(x).$$

Both identities also hold in the L-periodic case.

Remark 1. *Let us comment on the naming of the derivatives D and $\frac{\partial}{\partial y}$. A coefficient field $\boldsymbol{a} \in \Omega$ might be viewed as a "surface" $(x, \boldsymbol{a}(x))$ in the space $\mathbb{Z}^d \times \Omega_0$. We call the directions associated with $\mathbb{Z}^d$ horizontal and the directions associated with Ω_0 vertical. The horizontal derivative $D\zeta$ monitors the sensitivity of ζ w. r. t. to horizontal shifts of the coefficient field. In contrast, the vertical derivative $\frac{\partial}{\partial y}$ is associated with variations of the coefficient field in vertical directions. It can be seen as a discrete version of the classical partial derivatives $\{\frac{\partial}{\partial \boldsymbol{a}_{ii}(y)}\}_{i=1,\dots,d}$, and monitors how sensitively a random variable ζ depends on the value of the coefficient field $\{\mathbb{Z}^d \ni y \mapsto \boldsymbol{a}(y)\}$ at site y.*

Remark 2. *From the functional analytic point of view, (SG) is a Poincaré inequality on $L^2(\Omega)$ for the vertical derivative $\frac{\partial}{\partial y}$, which defines a bounded and symmetric operator $\frac{\partial}{\partial y} : L^2(\Omega) \to L^2(\Omega)$. Since each site $y \in \mathbb{Z}^d$ is endowed with a vertical derivative, the number of degrees of freedom that the vertical gradient $\{\frac{\partial}{\partial y}\}_{y \in \mathbb{Z}^d}$ controls matches the degrees of freedom of the underlying probability space $\Omega = \Omega_0^{\mathbb{Z}^d}$. One can show that $SG_\infty(\rho)$ implies that the underlying ensemble is ergodic (so that the qualitative stochastic homogenization theory holds).*

$SG_\infty(\rho)$ is satisfied in the case of i. i. d. coefficients:

Definition 3 (i. i. d. ensemble)**.** *Let β be a probability measure on Ω_0. The infinite i. i. d. ensemble associated with β is defined as the unique ensemble with the property that the coordinate projections $\Omega \ni \boldsymbol{a} \mapsto \boldsymbol{a}(x) \in \Omega_0$, $x \in \mathbb{Z}^d$, are independent and identically distributed with $\boldsymbol{a}(0) \sim \beta$. Likewise, we define for $L \in \mathbb{N}$ the L-periodic i. i. d. ensemble associated with β as the unique, L-periodic ensemble with the property that the coordinate projections $\Omega \ni \boldsymbol{a} \mapsto \boldsymbol{a}(x) \in \Omega_0$, $x \in ([0, L) \cap \mathbb{Z})^d$, are independent and identically distributed with $\boldsymbol{a}(0) \sim \beta$.*

Lemma 1. *Every infinite (resp. L-periodic) i. i. d. ensemble satisfies $SG_\infty(\rho)$ (resp. $SG_L(\rho)$) with constant $\rho = 1$.*

The proof, which relies on the tensorization principle, is standard (cf. [21]) and is omitted here.

In the statistical mechanics literature, in the context of Ising models, there are several criteria for the validity of the Spectral Gap estimate. In their seminal contributions [11, 12], Dobrushin & Shlosman (DS) introduced ten equivalent conditions that ensure the existence and analyticity of Gibbs fields on $\mathbb{Z}^d$. It was then proved later that the (DS) mixing conditions are in addition equivalent to the validity of a logarithmic-Sobolev inequality (which implies (SG)) for discrete spin spaces [40, Theorem 1.8] and for compact spin spaces [39, Corollary 3.19] & [38, Theorem 1.2]. The most suitable form of the (DS) mixing conditions in our context is given by [40, DSM] (see also [39, Remark 3.23]) and holds, roughly

speaking, provided correlations are integrable. However, the appropriate measure of independence of $\boldsymbol{a}(x)$ from $\boldsymbol{a}(y)$ for $|y - x| \gg 1$ has to be integrable in y under conditioning on $\{\boldsymbol{a}(z)\}_{z \in S}$ *uniformly* in any subset S of the lattice. It is this uniformity condition that puts the (DS) criterion in a different class than the usual mixing conditions in the quantitative theory of stochastic homogenization (cf. the uniformly strong mixing condition used in [42, (2.1)] and [2, Definition 2.1]). We don't know if the condition of integrable correlations on the level of the usual mixing conditions is sufficient for our results.

Roughly speaking, in particular disregarding the above-mentioned difference between these mixing conditions, it is clear that the condition of *integrability of correlations* is necessary for some of our results, as we shall explain now. Consider for instance Proposition 2. Suppose that we are in the case of small ellipticity contrast, i. e. $1 - \lambda \ll 1$. In this situation, an expansion of the definition (12) of the approximate homogenized coefficient in $1 - \lambda$ reveals that to leading order, $\boldsymbol{a}_{\mathrm{av},L}(\boldsymbol{a})$ is a simple spatial average

$$\boldsymbol{a}_{\mathrm{av},L}(\boldsymbol{a}) \approx L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x).$$

This approximate representation indeed implies that the variance of $\boldsymbol{a}_{\mathrm{av},L}(\boldsymbol{a})$ only scales as L^{-d} if the covariances $\langle (\boldsymbol{a}(x) - \langle \boldsymbol{a} \rangle_L)(\boldsymbol{a}(y) - \langle \boldsymbol{a} \rangle_L) \rangle_L = \langle (\boldsymbol{a}(x - y) - \langle \boldsymbol{a} \rangle_L)(\boldsymbol{a}(0) - \langle \boldsymbol{a} \rangle_L) \rangle_L$ are integrable in $x - y$ (i. e. summable over the periodic box with a sum that stays finite for $L \uparrow \infty$). A main contribution of this paper is to show that this scaling is preserved in the case of arbitrary ellipticity contrast $\lambda > 0$ when $\boldsymbol{a}_{\mathrm{av},L}(\boldsymbol{a})$ is a bigly nonlinear function of $\boldsymbol{a}$.

3   Decay of the variable-coefficient semigroup

This section is devoted to the proof of Theorem 1. The proof essentially relies on the vertical differential calculus and the optimal decay estimates of the parabolic Green's function from Theorem 3. In order to exploit the vertical differential calculus on the solution u of (7), we shall work with its stationary extension $\overline{u}$, as explained in the upcoming section. Section 3.2 is then devoted to the proof of Theorem 1.

3.1   The parabolic equation and Green's function, and Duhamel's formula

Solutions of (7) – the parabolic problem in the probability space – are characterized by a corresponding parabolic equation in the physical space $\mathbb{Z}^d$, and thus admit a representation via the parabolic Green's function on $\mathbb{Z}^d$. To make that precise, recall that the operator $X \ni \zeta \mapsto D^* \boldsymbol{a}(0) D \zeta \in X$ is bounded and linear with X denoting any of the spaces $L^p(\Omega)$, $1 \le p \le \infty$, or $C_b(\Omega)$. Hence, by standard semigroup theory, for all $\zeta \in X$ the function $u(t) := \exp\left(-t D^* \boldsymbol{a}(0) D\right)\zeta$ is a $C^\infty(\mathbb{R}, X)$-solution of (7). Thanks to (18) it is elementary to see that for all $\zeta \in C_b(\Omega)$, the stationary extension $\overline{u}$ of u solves (for all $\boldsymbol{a} \in \Omega$) the following parabolic equation in physical space:

$$(22) \qquad \begin{cases} \partial_t \overline{u}(t,x) + \nabla^* \boldsymbol{a}(x) \nabla \overline{u}(t,x) = g(t,x) & \text{for all } t > 0, \ x \in \mathbb{Z}^d, \\ \overline{u}(t = 0, x) = g_0(x) & \text{for all } x \in \mathbb{Z}^d, \end{cases}$$

with $g(t,x) = 0$ and $g_0(x) := \overline{\zeta}(\boldsymbol{a}, x)$. The solution of (22) can be represented via Duhamel's formula:

$$(23) \qquad \overline{u}(t,x) = \sum_{y \in \mathbb{Z}^d} G(t, \boldsymbol{a}, x, y) g_0(y) + \int_0^t \sum_{y \in \mathbb{Z}^d} G(t - s, \boldsymbol{a}, x, y) g(s, y) \, ds,$$

where G denotes the parabolic Green's function and is defined for all $\boldsymbol{a} \in \Omega$ and $y \in \mathbb{Z}^d$ as the function $(t,x) \mapsto G(t, \boldsymbol{a}, x, y)$ in $C^\infty(\mathbb{R}_+, \ell^1(\mathbb{Z}^d))$ given by $G(t, \boldsymbol{a}, \cdot, y) := \exp(-t \nabla^* \boldsymbol{a}(\cdot) \nabla) \delta(\cdot - y)$, where $\delta(x) = 1$ if $x = 0$ and $\delta(x) = 0$ if $x \in \mathbb{Z}^d \setminus \{0\}$. Likewise, the L-periodic parabolic Green's function G_L, $L \in \mathbb{N}$, is defined for all $\boldsymbol{a} \in \Omega_L$ and $y \in \mathbb{Z}^d$ as the function $(t,x) \mapsto G_L(t, \boldsymbol{a}, x, y)$ in

$C^\infty(\mathbb{R}^+, \ell^\infty(\mathbb{Z}^d))$ given by $G_L(t, \boldsymbol{a}, \cdot, y) := \exp(-t\nabla^*\boldsymbol{a}(\cdot)\nabla)\delta_L(\cdot - y)$, where $\delta_L(x) := \sum_{z \in \mathbb{Z}^d} \delta(x + Lz)$ denotes the Dirac function on the discrete torus of size L.

Remark 3. *Thanks to the assumption that $\boldsymbol{a}$ is diagonal (and elliptic) we have a discrete mean value property in the following sense: If $\nabla^*\boldsymbol{a}\nabla u = 0$ then $u(\cdot) \leq \max\{u(\cdot \pm e_1), \ldots, u(\cdot \pm e_d)\}$, which yields a weak maximum principle for $\partial_t + \nabla^*\boldsymbol{a}(\cdot)\nabla$. In particular, this implies that G and G_L are non-negative. This is crucial, since our results heavily rely on estimates of the Green's function that are based on elliptic and parabolic regularity theory.*

3.2 Auxiliary lemmas and proof of Theorem 1

We split the proof of Theorem 1 into several lemmas (which we prove in Section 3.3). The starting point is the spectral gap estimate. Since we have to estimate higher moments, we need the following version of (SG):

Lemma 2 (*p*-version of (SG)). *Let $\langle \cdot \rangle$ satisfy (11). Then for all $p \geq 1$ and all $\zeta \in C_b(\Omega)$ with $\langle \zeta \rangle = 0$ we have*

$$\left\langle (\zeta^2)^p \right\rangle^{\frac{1}{2p}} \lesssim \|\partial\zeta\|_{L^{2p}_{\langle\cdot\rangle}\ell^2_y},$$

where

$$\|\partial\zeta\|_{L^{2p}_{\langle\cdot\rangle}\ell^2_y} := \left\langle \left(\sum_{y \in \mathbb{Z}^d} \left(\frac{\partial\zeta}{\partial y} \right)^2 \right)^p \right\rangle^{\frac{1}{2p}},$$

and the constant only depends on p and ρ.

In the L-periodic case, i. e. when $\langle \cdot \rangle$ satisfies (8), the statement remains valid for $\sum_{y \in \mathbb{Z}^d}$ and $\frac{\partial}{\partial y}$ replaced by $\sum_{y \in ([0,L) \cap \mathbb{Z})^d}$ and $\frac{\partial}{\partial_L y}$, respectively.

In the course of proving Theorem 1 via Lemma 2, we have to estimate the vertical derivative $\frac{\partial u}{\partial y}$, which can be conveniently characterized with the help of the stationary extension $\overline{u}$ of u. Indeed, since $\overline{u}$ solves (22) with $g \equiv 0$ and $g_0(x) = \nabla^*\overline{\xi}(\boldsymbol{a}, x)$, application of $\frac{\partial}{\partial y}$ to (22) shows that the non-stationary random field $\frac{\partial\overline{u}(t,\boldsymbol{a},x)}{\partial y}$ is the unique solution of the parabolic equation (with time variable t and space variable x)

$$\begin{cases} \partial_t \frac{\partial\overline{u}(t,\boldsymbol{a},x)}{\partial y} + \nabla^*\boldsymbol{a}(x)\nabla\frac{\partial\overline{u}(t,\boldsymbol{a},x)}{\partial y} = \nabla^*g(t, \boldsymbol{a}, x, y) & t > 0, \, x \in \mathbb{Z}^d, \\ \frac{\partial\overline{u}(t=0,\boldsymbol{a},x)}{\partial y} = \nabla^*\overline{\left(\frac{\partial\xi}{\partial(y-x)}\right)}(x) & x \in \mathbb{Z}^d, \end{cases}$$

where by Leibniz' rule

$$g(t, \boldsymbol{a}, x, y) \overset{\text{formally}}{=} -\frac{\partial\boldsymbol{a}(x)}{\partial y}\overline{Du(t, \boldsymbol{a})}(x).$$

Since $\frac{\partial\boldsymbol{a}(x)}{\partial y}$ vanishes for $x \neq y$, Duhamel's formula (23) yields after integrating by parts

$$(24) \quad \frac{\partial\overline{u}(t,\boldsymbol{a},x)}{\partial y} \overset{\text{formally}}{=} \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, \boldsymbol{a}, x, z) \cdot \overline{\left(\frac{\partial\xi}{\partial(y-z)}\right)}(z) - \int_0^t \nabla_y G(t-s, \boldsymbol{a}, x, y) \cdot \frac{\partial\boldsymbol{a}(y)}{\partial y}\overline{Du(s, \boldsymbol{a})}(y) \, ds.$$

Note that the formula for g (and thus the identity above) is only formal, since Leibniz' rule does not hold for the (discrete) vertical derivative. However, we obtain precisely the same estimate as if (24) was correct:

Lemma 3. *Let $\langle \cdot \rangle$ be stationary. Consider*

$$u(t) := \exp\left(-tD^*\boldsymbol{a}(0)D\right)D^*\xi, \qquad \xi \in C_b(\Omega)^d.$$

12

We then have

$$(25) \qquad \frac{\partial u(t)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, 0, z) \cdot \overline{\left(\frac{\partial \xi}{\partial(y-z)}\right)}(z) + \int_0^t \nabla_y G(t-s, 0, y) \cdot \bar{g}(s, y) \, ds$$

$\langle \cdot \rangle$-*almost surely, where $\bar{g}$ is the stationary extension of some $g(t, \boldsymbol{a})$ that satisfies*

$$(26) \qquad \left\langle |g(t)|^{2p} \right\rangle^{\frac{1}{2p}} \leq 2 \left\langle |Du(t)|^{2p} \right\rangle^{\frac{1}{2p}}$$

for all $p < \infty$.

Likewise, the statement holds for $\sum_{z \in \mathbb{Z}^d}$, $\frac{\partial}{\partial(y-z)}$, and G replaced by $\sum_{z \in ([0,L) \cap \mathbb{Z})^d}$, $\frac{\partial}{\partial_L(y-z)}$, and G_L, respectively.

Lemma 3 shows that we need time-decay estimates of ∇G. These are given by Theorem 3. As we shall see in the proof of Lemma 4 below, the fact that in Theorem 3 we obtain optimal decay for ∇G for exponents up to some $2q_0$ slightly larger than 2 is crucial. In fact, the exponent p_0 in the statement of Theorem 1 is the dual exponent of q_0. This explains why we are forced to estimate high moments of u even if ultimately we are mostly interested in the second and fourth moments, cf. Section 4.

Combining the Spectral Gap Estimate in its p-version with the representation formula of Lemma 3 we get:

Lemma 4. *In the situation of Theorem 1 there exists an exponent $1 \leq p_0 < \infty$ (only depending on λ and d), such that for every $p \in (p_0, \infty)$ we have*

$$(27) \qquad \left\langle u^{2p}(t) \right\rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})} \|\partial \xi\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}} + \int_0^t (t-s+1)^{-(\frac{d}{4}+\frac{1}{2})} \left\langle |Du(s)|^{2p} \right\rangle^{\frac{1}{2p}} \, ds,$$

where $\|\partial \xi\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}}$ is defined as in Theorem 1, and the constant only depends on p, ρ, λ, and d.

To complete the proof of Theorem 1 we have to gain control over the nonlinear term $\int_0^t (t-s+1)^{-(\frac{d}{4}+\frac{1}{2})} \langle |Du(s)|^{2p} \rangle^{\frac{1}{2p}} \, ds$. This is done by using Caccioppoli's inequality in probability combined with an ODE-argument. More precisely, the following lemma shows that $\langle |Du(t)|^{2p} \rangle$ has better decay than $\langle u^{2p}(t) \rangle$.

Lemma 5 (Caccioppoli). *Let $\langle \cdot \rangle$ be stationary. Consider $u(t) := \exp(-tD^* \boldsymbol{a}(0) D)\zeta$ with $\zeta \in C_b(\Omega)$. Then for all $p \geq 1$ we have*

$$\left\langle |Du(t)|^{2p} \right\rangle \lesssim -\frac{d}{dt} \left\langle (u^2)^p(t) \right\rangle,$$

where the constant only depends on d and p.

The following lemma shows how to absorb the nonlinear term into the LHS of (27).

Lemma 6 (ODE-argument). *Let $1 \leq p, \gamma < \infty$ and $a(t), b(t) \geq 0$. Suppose that there exists $C_1 < \infty$ such that for all $t \geq 0$,*

$$(28a) \qquad a(t) \leq C_1 \left((t+1)^{-\gamma} + \int_0^t (t-s+1)^{-\gamma} b(s) \, ds \right),$$

$$(28b) \qquad b^p(t) \leq C_1 \left(-\frac{d}{dt} a^p(t) \right).$$

Then there exists $C_2 < \infty$ depending only on C_1, p and γ such that

$$a(t) \leq C_2(t+1)^{-\gamma}.$$

We are in position to prove Theorem 1.

Proof of Theorem 1. Let p_0 be given by Lemma 4, and fix an exponent $p \in (p_0, \infty)$. By homogeneity we may assume that $\|\xi\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}} = 1$, so that the desired estimate reduces to

$$(29) \qquad \langle u^{2p}(t) \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})}.$$

Set

$$a(t) := \langle u^{2p}(t) \rangle^{\frac{1}{2p}}, \qquad b(t) := \langle |Du(t)|^{2p} \rangle^{\frac{1}{2p}}, \qquad \gamma = \frac{d}{4} + \frac{1}{2}.$$

By Lemma 4 and Lemma 5 we have

$$a(t) \;\lesssim\; (t+1)^{-\gamma} + \int_0^t (t-s+1)^{-\gamma} b(s)\, ds,$$

$$b^{2p}(t) \;\lesssim\; -\frac{d}{dt} a^{2p}(t).$$

Hence, Lemma 6 yields (29) in the form of $a(t) \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})}$. $\qquad\square$

We finally state a slight improvement of Theorem 1 that yields an additional exponential decay (for times $t \gtrsim L^2$) in the L-periodic case:

Lemma 7. *In the situation of Theorem 1 the following holds: If $\langle \cdot \rangle$ satisfies (8), then there exists $c_0 > 0$ depending only on d such that*

$$\langle |\exp(-tD^* \boldsymbol{a}(0)D)D^* \xi|^2 \rangle^{\frac{1}{2}} \lesssim \exp\big(-\frac{c_0 \lambda}{L^2} t\big) \langle |\xi|^2 \rangle^{\frac{1}{2}},$$

where the constant only depends on ρ, λ, and d.

3.3   Proofs of the auxiliary lemmas

We prove the auxiliary lemmas in the case (11). The argument in the periodic case (8) is similar.

Proof of Lemma 2. The only technicality is due to the failure of the Leibniz rule for the vertical derivative. In what follows, for all $a \in \mathbb{R}$ and $p > 0$, we use the notation $a^{2p} := (a^2)^p$.
Step 1. Substitute for the Leibniz rule: For all $p \geq 1$,

$$(30) \qquad \left\langle \big(\frac{\partial(\zeta|\zeta|^{p-1})}{\partial y}\big)^2 \right\rangle \lesssim \left\langle \zeta^{2(p-1)}\big(\frac{\partial \zeta}{\partial y}\big)^2 + \big(\frac{\partial \zeta}{\partial y}\big)^{2p} \right\rangle,$$

where the constant only depends on p and d.
By definition of $\frac{\partial}{\partial y}$, (30) can be rewritten as

$$\langle (\zeta|\zeta|^{p-1} - \langle \zeta|\zeta|^{p-1} \rangle_y)^2 \rangle \lesssim \left\langle \zeta^{2(p-1)}(\zeta - \langle \zeta \rangle_y)^2 + (\zeta - \langle \zeta \rangle_y)^{2p} \right\rangle.$$

Since $\langle \langle \cdot \rangle_y \rangle = \langle \cdot \rangle$, it suffices to show that

$$\langle (\zeta|\zeta|^{p-1} - \langle \zeta|\zeta|^{p-1} \rangle_y)^2 \rangle_y \lesssim \left\langle \zeta^{2(p-1)}(\zeta - \langle \zeta \rangle_y)^2 + (\zeta - \langle \zeta \rangle_y)^{2p} \right\rangle_y \qquad \langle \cdot \rangle\text{-almost surely.}$$

Since the conditional expectation is an orthogonal projection in $L^2(\Omega)$, we have

$$\forall \tilde{\zeta} \in L^2(\Omega), \forall c \in \mathbb{R}: \qquad \left\langle (\tilde{\zeta} - \langle \tilde{\zeta} \rangle_y)^2 \right\rangle_y \leq \left\langle (\tilde{\zeta} - c)^2 \right\rangle_y \qquad \langle \cdot \rangle\text{-almost surely.}$$

In particular, with $\tilde{\zeta} = \zeta|\zeta|^{p-1}$ and $c = \langle \zeta \rangle_y |\langle \zeta \rangle_y|^{p-1}$ we get $\langle \cdot \rangle$-almost surely

$$\langle (\zeta|\zeta|^{p-1} - \langle \zeta|\zeta|^{p-1} \rangle_y)^2 \rangle_y \leq \langle (\zeta|\zeta|^{p-1} - \langle \zeta \rangle_y |\langle \zeta \rangle_y|^{p-1})^2 \rangle_y.$$

14

Hence, it suffices to argue that

$$\left\langle (\zeta|\zeta|^{p-1} - \langle\zeta\rangle_y|\langle\zeta\rangle_y|^{p-1})^2\right\rangle_y \lesssim \left\langle \zeta^{2(p-1)}(\zeta - \langle\zeta\rangle_y)^2 + (\zeta - \langle\zeta\rangle_y)^{2p}\right\rangle_y \qquad \langle\cdot\rangle\text{-almost surely.}$$

The latter follows from the elementary inequality:

$$\forall a, b \in \mathbb{R}: \qquad (a|a|^{p-1} - b|b|^{p-1})^2 \lesssim a^{2(p-1)}(a-b)^2 + (a-b)^{2p}.$$

Step 2. Application of the Spectral Gap Estimate.
For all $p \geq 1$ we claim that

$$\left\langle \zeta^{2p}\right\rangle \lesssim \left\langle \zeta^2\right\rangle^p + \left\langle \Big(\sum_{y\in\mathbb{Z}^d} \big(\tfrac{\partial\zeta}{\partial y}\big)^2\Big)^p\right\rangle.$$

Assume that $p > 1$. The application of $\mathrm{SG}_\infty(\rho)$ to $\zeta|\zeta|^{p-1} - \langle\zeta|\zeta|^{p-1}\rangle$ yields

$$\left\langle (\zeta|\zeta|^{p-1} - \langle\zeta|\zeta|^{p-1}\rangle)^2\right\rangle \leq \frac{1}{\rho}\sum_{y\in\mathbb{Z}^d}\left\langle \big(\tfrac{\partial}{\partial y}(\zeta|\zeta|^{p-1})\big)^2\right\rangle,$$

which, by Step 1 and the triangle inequality, turns into

(31) $$\left\langle \zeta^{2p}\right\rangle \lesssim \left\langle |\zeta|^p\right\rangle^2 + \left\langle \zeta^{2(p-1)}\sum_{y\in\mathbb{Z}^d}\big(\tfrac{\partial\zeta}{\partial y}\big)^2 + \sum_{y\in\mathbb{Z}^d}\big(\tfrac{\partial\zeta}{\partial y}\big)^{2p}\right\rangle.$$

We treat each of the three terms on the RHS of (31) separately. For the third term we appeal to the discrete ℓ^{2p}-ℓ^2 estimate. For the second term on the RHS of (31) we use Hölder's inequality with exponents $(\frac{p}{p-1}, p)$, combined with Young's inequality. We turn to the first term. For $p = 2$ there is nothing to do, whereas for $p > 2$ we may apply Hölder's inequality with exponents $(2\frac{p-1}{p-2}, 2\frac{p-1}{p})$ to $\langle|\zeta|^p\rangle = \left\langle |\zeta|^{p\frac{p-2}{p-1}}|\zeta|^{\frac{p}{p-1}}\right\rangle$:

$$\left\langle |\zeta|^p\right\rangle^2 \leq \left\langle \zeta^{2p}\right\rangle^{\frac{p-2}{p-1}}\left\langle \zeta^2\right\rangle^{\frac{p}{p-1}},$$

and we absorb the first factor into the LHS of (31) by Young's inequality.
Step 3. Conclusion.
Application of $\mathrm{SG}_\infty(\rho)$ to ζ combined with Jensen's inequality yields

$$\left\langle \zeta^2\right\rangle^p \leq \frac{1}{\rho^p}\left\langle \Big(\sum_{y\in\mathbb{Z}^d}\big(\tfrac{\partial\zeta}{\partial y}\big)^2\Big)^p\right\rangle,$$

so that the claim of Lemma 2 follows from Step 2. $\qquad\square$

Proof of Lemma 3. Recall that the stationary extension $\overline{u}(t, \cdot)$ of u satisfies (for all $\boldsymbol{a} \in \Omega$) the spatial parabolic equation (22) with $g \equiv 0$ and $g_0(x) := \nabla^*\overline{\xi}(x)$. We take the vertical derivative $\frac{\partial}{\partial y}$ of this equation:

(32) $$\begin{cases} \partial_t \frac{\partial\overline{u}(t,x)}{\partial y} + \nabla^*\boldsymbol{a}(x)\nabla\frac{\partial\overline{u}(t,x)}{\partial y} = \nabla^*\xi_1(t,x,y) & \text{for all } t > 0,\ x \in \mathbb{Z}^d, \\ \frac{\partial\overline{u}(t=0,x)}{\partial y} = \nabla^*\xi_0(x,y) & \text{for all } x \in \mathbb{Z}^d, \end{cases}$$

where

$$\begin{aligned} \xi_1(t,x,y) &:= \boldsymbol{a}(x)\nabla\frac{\partial\overline{u}(t,x)}{\partial y} - \frac{\partial}{\partial y}\big(\boldsymbol{a}(x)\nabla\overline{u}(t,x)\big), \\ \xi_0(x,y) &:= \frac{\partial\overline{\xi}(x)}{\partial y}. \end{aligned}$$

15

Duhamel's formula, cf. (23), and two integrations by parts yield

$$
(33) \qquad \frac{\partial \overline{u}(t,0)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t,0,z) \cdot \xi_0(z,y) + \int_0^t \sum_{z \in \mathbb{Z}^d} \nabla_z G(t-s,0,z) \cdot \xi_1(s,z,y)\, ds.
$$

We then claim that

$$
(34) \qquad \xi_0(x,y) = \overline{\left(\frac{\partial \xi}{\partial(y-x)}\right)}(x),
$$
$$
(35) \qquad \xi_1(t,x,y) = \delta(y-x)\overline{g}(t,x),
$$

where $g(t) := \langle \boldsymbol{a}(0) D u(t)\rangle_0 - \boldsymbol{a}(0)\langle Du(t)\rangle_0$.

Indeed, (34) is a consequence of (21). Let us prove (35). Since $\xi_1(t,x,y) = \langle \boldsymbol{a}(x)\nabla \overline{u}(t,x)\rangle_y - \boldsymbol{a}(x)\nabla \langle \overline{u}(t,x)\rangle_y$ vanishes for all $y \neq x$, (35) follows from the properties of the stationary extension:

$$
\begin{aligned}
\langle \boldsymbol{a}(x)\nabla \overline{u}(t,x)\rangle_y - \boldsymbol{a}(x)\langle \nabla \overline{u}(t,x)\rangle_y \; &= \; \left\langle \overline{\boldsymbol{a}(0)Du(t)}(x)\right\rangle_y - \boldsymbol{a}(x)\left\langle \overline{Du(t)}(x)\right\rangle_y \\
&\overset{(20)}{=} \; \delta(y-x)\overline{\left(\langle \boldsymbol{a}(0)Du(t)\rangle_{y-x} - \boldsymbol{a}(0)\langle Du(t)\rangle_{y-x}\right)}(x) \\
&= \; \delta(y-x)\overline{g}(t,x).
\end{aligned}
$$

Since $u(t) = \overline{u}(t,0)$, we have $\frac{\partial \overline{u}(t,0)}{\partial y} = \frac{\partial u(t)}{\partial y}$, and the combination of (33)—(35) yields (25), whereas (26) follows from the triangle inequality in $L^{2p}(\Omega)$ and Jensen's inequality in probability. $\qquad \square$

Proof of Lemma 4. Since $\langle D^* \xi \rangle = 0$, we have $\langle u(t)\rangle = 0$ for all $t > 0$, so that Lemma 2 yields

$$
\langle u^{2p}(t)\rangle \lesssim \left\langle \left(\sum_{y \in \mathbb{Z}^d} \left(\tfrac{\partial u(t)}{\partial y}\right)^2\right)^p\right\rangle.
$$

The representation formula of Lemma 3 for $\frac{\partial u(t)}{\partial y}$ and the triangle inequality w. r. t. $\left\langle \left(\sum_{y \in \mathbb{Z}^d}(\cdot)^2\right)^p\right\rangle^{\frac{1}{2p}}$ and in the time integral show that

$$
\begin{aligned}
(36) \qquad \langle u^{2p}(t)\rangle^{\frac{1}{2p}} \; \lesssim \; & \left\langle \left(\sum_{y \in \mathbb{Z}^d}\left(\sum_{z \in \mathbb{Z}^d}|\nabla_z G(t,0,z)|\left|\overline{\left(\tfrac{\partial \xi}{\partial(y-z)}\right)}(z)\right|\right)^2\right)^p\right\rangle^{\frac{1}{2p}} \\
& + \int_0^t \left\langle \left(\sum_{y \in \mathbb{Z}^d}|\nabla_y G(t-s,0,y)|^2 |\overline{g}(s,y)|^2\right)^p\right\rangle^{\frac{1}{2p}} ds.
\end{aligned}
$$

Recall that $\langle |\overline{g}(t,x)|^{2p}\rangle^{\frac{1}{2p}} \leq 2\langle |Du(t)|^{2p}\rangle^{\frac{1}{2p}}$. To estimate the first term of the RHS we change variables and use the triangle inequality w. r. t. $\left\langle \left(\sum_{y \in \mathbb{Z}^d}(\cdot)^2\right)^p\right\rangle^{\frac{1}{2p}}$:

$$
\begin{aligned}
& \left\langle \left(\sum_{y \in \mathbb{Z}^d}\left(\sum_{z \in \mathbb{Z}^d}|\nabla_z G(t,0,z)|\left|\overline{\left(\tfrac{\partial \xi}{\partial(y-z)}\right)}(z)\right|\right)^2\right)^p\right\rangle^{\frac{1}{2p}} \\
& \overset{x:=y-z}{=} \left\langle \left(\sum_{y \in \mathbb{Z}^d}\left(\sum_{x \in \mathbb{Z}^d}|\nabla_y G(t,0,y-x)|\left|\overline{\tfrac{\partial \xi}{\partial x}}(y-x)\right|\right)^2\right)^p\right\rangle^{\frac{1}{2p}} \\
& \overset{\triangle\text{-inequality}}{\leq} \sum_{x \in \mathbb{Z}^d}\left\langle \left(\sum_{y \in \mathbb{Z}^d}\left(|\nabla_y G(t,0,y-x)|\left|\overline{\tfrac{\partial \xi}{\partial x}}(y-x)\right|\right)^2\right)^p\right\rangle^{\frac{1}{2p}} \\
& \overset{x':=y-x}{=} \sum_{x \in \mathbb{Z}^d}\left\langle \left(\sum_{x' \in \mathbb{Z}^d}\left(|\nabla_{x'} G(t,0,x')|\left|\overline{\tfrac{\partial \xi}{\partial x}}(x')\right|\right)^2\right)^p\right\rangle^{\frac{1}{2p}}.
\end{aligned}
$$

16

Hence, (36) turns into

$$
\langle u^{2p}(t) \rangle^{\frac{1}{2p}} \leq \sum_{x \in \mathbb{Z}^d} \left\langle \left(\sum_{y \in \mathbb{Z}^d} \left(|\nabla_y G(t,0,y)| |\overline{\tfrac{\partial \xi}{\partial x}}(y)| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}}
$$
$$
+ \int_0^t \left\langle \left(\sum_{y \in \mathbb{Z}^d} |\nabla_y G(t-s,0,y)|^2 |\overline{g}(s,y)|^2 \right)^p \right\rangle^{\frac{1}{2p}} ds.
$$

It remains to show that

$$
(37) \qquad \left\langle \left(\sum_{y \in \mathbb{Z}^d} \left(|\nabla_y G(t,0,y)| |\overline{\tfrac{\partial \xi}{\partial x}}(y)| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})} \left\langle |\tfrac{\partial \xi}{\partial x}|^{2p} \right\rangle^{\frac{1}{2p}},
$$

$$
(38) \qquad \left\langle \left(\sum_{y \in \mathbb{Z}^d} |\nabla_y G(t-s,0,y)|^2 |\overline{g}(s,y)|^2 \right)^p \right\rangle^{\frac{1}{2p}} \lesssim (t-s+1)^{-(\frac{d}{4}+\frac{1}{2})} \left\langle |Du(s)|^{2p} \right\rangle^{\frac{1}{2p}}.
$$

We only give the argument for (37), the argument for (38) being similar. Let $q := \frac{p}{p-1}$ be the dual exponent of p, let $\alpha > 0$ be some exponent to be fixed later, and let $\omega(t,x)$ be the weight defined in (16). By Hölder's inequality with exponents (q,p) and the symmetry $G(t,\boldsymbol{a},x,y) = G(t,\boldsymbol{a},y,x)$ of G, we have

$$
\sum_{y \in \mathbb{Z}^d} \left(|\nabla_y G(t,0,y)| |\overline{\tfrac{\partial \xi}{\partial x}}(y)| \right)^2 \leq \left(\sum_{y \in \mathbb{Z}^d} \left(\omega^\alpha(t,y) |\nabla_y G(t,y,0)| \right)^{2q} \right)^{\frac{1}{q}} \left(\sum_{y \in \mathbb{Z}^d} \left(\omega^{-\alpha}(t,y) |\overline{\tfrac{\partial \xi}{\partial x}}(y)| \right)^{2p} \right)^{\frac{1}{p}}.
$$

Hence, by stationarity of $\overline{\tfrac{\partial \xi}{\partial x}}$,

$$
(39) \qquad [\text{LHS of } (37)] \leq \sup_{\boldsymbol{a} \in \Omega} \left(\sum_{y \in \mathbb{Z}^d} \left(\omega^\alpha(t,y) |\nabla_y G(t,y,0)| \right)^{2q} \right)^{\frac{1}{2q}} \left(\sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t,y) \right)^{\frac{1}{2p}} \left\langle |\tfrac{\partial \xi}{\partial x}|^{2p} \right\rangle^{\frac{1}{2p}}.
$$

We now address the choice of p and α. First, set p_0 to be the dual exponent of q_0 defined in Theorem 3, and let $p > p_0$, so that its dual exponent q lies in the range $[1, q_0)$ of applicability of Theorem 3. We then choose α so large that $2p\alpha > d$. Theorem 3 thus implies

$$
\sup_{\boldsymbol{a} \in \Omega} \left(\sum_{y \in \mathbb{Z}^d} \left(\omega^\alpha(t,y) |\nabla_y G(t,y,0)| \right)^{2q} \right)^{\frac{1}{2q}} \lesssim (t+1)^{-(\frac{d}{2}+\frac{1}{2})+\frac{d}{2}\frac{1}{2q}},
$$

whereas

$$
\sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t,y) = \sum_{y \in \mathbb{Z}^d} \left(\frac{|y|^2}{t+1} + 1 \right)^{-p\alpha} \lesssim (t+1)^{\frac{d}{2}}.
$$

Combined with (39), this yields the desired estimate (37). $\qquad \square$

Proof of Lemma 5. For all $t > 0$ and $p \geq 1$

$$
\lambda \left\langle D(u|u|^{2p-2}) \cdot Du \right\rangle \leq -\frac{d}{dt} \frac{1}{2p} \left\langle (u^2)^p \right\rangle.
$$

Indeed, since $u(t) = \exp(-tD^*\boldsymbol{a}(0)D)\zeta$,

$$
(40) \qquad \frac{\partial}{\partial t} u + D^* \boldsymbol{a}(0) D u = 0 \qquad t > 0,
$$

whose weak formulation with test-function $u|u|^{2p-2} \in C^\infty(\mathbb{R}_+, C_b(\Omega))$ reads

$$\frac{d}{dt}\frac{1}{2p}\left\langle (u^2)^p \right\rangle = -\left\langle D(u|u|^{2p-2}) \cdot \boldsymbol{a}(0)Du \right\rangle \le -\lambda \left\langle D(u|u|^{2p-2}) \cdot Du \right\rangle,$$

where we used the diagonality of $\boldsymbol{a}(0)$ in the last inequality. The claim follows from the inequality $D(u|u|^{2p-2}) \cdot Du \gtrsim |Du|^{2p}$, which is a consequence of the elementary estimate: For all $a, b \in \mathbb{R}$ we have $(a|a|^{2p-2} - b|b|^{2p-2})(a-b) \gtrsim (a-b)^{2p}$. $\qquad\square$

Proof of Lemma 6. The claim can be reformulated as

$$(41) \qquad\qquad \Lambda(t) := \sup_{0 \le s \le t} (1+s)^\gamma a(s) \lesssim 1.$$

Step 1. Two auxiliary estimates.
We claim that

$$(42) \qquad\qquad \int_{\tau_1}^{\tau_2} b(s)\,ds \lesssim \begin{cases} (\tau_2 - \tau_1)^{1-\frac{1}{p}} a(\tau_1) & \text{for all } 0 \le \tau_1 \le \tau_2, \\ \tau_1^{1-\gamma-\frac{1}{p}} \Lambda(\tau_2) & \text{for all } 1 \le \tau_1 \le \tau_2. \end{cases}$$

The first estimate follows from Hölder's inequality, equation (28b), and the non-negativity of a:

$$(43) \qquad \int_{\tau_1}^{\tau_2} b(s)\,ds \le (\tau_2 - \tau_1)^{1-\frac{1}{p}} \left(-\int_{\tau_1}^{\tau_2} \frac{d}{dt} a^p(t)\,dt \right)^{\frac{1}{p}} \le (\tau_2 - \tau_1)^{1-\frac{1}{p}} a(\tau_1).$$

The second inequality can be deduced from the first one as follows: Let $N \in \mathbb{N}$ satisfy $2^{N-1}\tau_1 < \tau_2 \le 2^N \tau_1$. We then have

$$
\begin{aligned}
\int_{\tau_1}^{\tau_2} b(s)\,ds \quad &\le \quad \sum_{n=0}^{N-1} \int_{2^n \tau_1}^{2^{n+1}\tau_1} b(s)\,ds \overset{(43)}{\lesssim} \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}} a(2^n\tau_1) \\
&\le \quad \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}} (1+2^n\tau_1)^{-\gamma}\Lambda(2^n\tau_1) \\
&\overset{\Lambda \text{ monotone}}{\le} \quad \Lambda(\tau_2)\sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}}(1+2^n\tau_1)^{-\gamma} \\
&\le \quad \Lambda(\tau_2)\sum_{n=0}^{N-1} (2^n\tau_1)^{1-\gamma-\frac{1}{p}} \overset{\gamma \ge 1}{\lesssim} \Lambda(\tau_2)\tau_1^{1-\gamma-\frac{1}{p}}.
\end{aligned}
$$

Step 2. A threshold estimate.
Let $1 \le \tau \le \frac{1}{4}t$. We claim that

$$(t+1)^\gamma a(t) \lesssim 1 + \tau^{1-\frac{1}{p}} + \left(\tau^{1-\gamma-\frac{1}{p}} + \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}} \right)\Lambda(t).$$

First notice that by (28b) the function $a(\cdot)$ is non-increasing. Hence,

$$
\begin{aligned}
(44) \qquad a(t) \quad &\le \quad \frac{2}{t}\int_{\frac{t}{2}}^t a(t')\,dt' \\
&\overset{(28a)}{\lesssim} \quad \frac{1}{t}\int_{\frac{t}{2}}^t (t'+1)^{-\gamma}\,dt' + \frac{1}{t}\int_{\frac{t}{2}}^t \int_0^{t'} (t'-s+1)^{-\gamma} b(s)\,ds\,dt'..
\end{aligned}
$$

The first term of the RHS is estimated by

$$(45) \qquad \frac{1}{t} \int_{\frac{t}{2}}^{t} (t'+1)^{-\gamma} \, dt' \lesssim (t+1)^{-\gamma}.$$

For the second term of the RHS of (44), we split the inner integral into three contributions that we estimate separately. More precisely, we shall prove that

$$(46) \qquad \frac{1}{t} \int_{\frac{t}{2}}^{t} \int_{0}^{\tau} (t'-s+1)^{-\gamma} b(s) \, ds \, dt' \lesssim (t+1)^{-\gamma} \tau^{1-\frac{1}{p}},$$

$$(47) \qquad \frac{1}{t} \int_{\frac{t}{2}}^{t} \int_{\tau}^{\frac{t'}{2}} (t'-s+1)^{-\gamma} b(s) \, ds \, dt' \lesssim (t+1)^{-\gamma} \tau^{1-\gamma-\frac{1}{p}} \Lambda(t),$$

$$(48) \qquad \frac{1}{t} \int_{\frac{t}{2}}^{t} \int_{\frac{t'}{2}}^{t'} (t'-s+1)^{-\gamma} b(s) \, ds \, dt' \lesssim (t+1)^{-\gamma} \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}} \Lambda(t)..$$

Argument for (46): Since $\tau \leq \frac{t}{4} \leq \frac{t'}{2}$, we have $t'-s+1 \geq \frac{t'}{2}+1$. Hence,

$$\int_{0}^{\tau} (t'-s+1)^{-\gamma} b(s) \, ds \lesssim (t'+1)^{-\gamma} \int_{0}^{\tau} b(s) \, ds \overset{(42)}{\lesssim} (t'+1)^{-\gamma} \tau^{1-\frac{1}{p}} a(0),$$

and (46) follows by (45) and (28a) for $t = 0$ in the form of $a(0) \lesssim 1$.

Argument for (47): As above we have

$$\int_{\tau}^{\frac{t'}{2}} (t'-s+1)^{-\gamma} b(s) \, ds \lesssim (t'+1)^{-\gamma} \int_{\tau}^{\frac{t'}{2}} b(s) \, ds \overset{(42)}{\lesssim} (t'+1)^{-\gamma} \tau^{1-\gamma-\frac{1}{p}} \Lambda(\tfrac{t'}{2}),$$

and (47) follows by (45) and the monotonicity of Λ.

Argument for (48): Since

$$\frac{t}{2} \leq t' \leq t \qquad \text{and} \qquad \frac{t'}{2} \leq s \leq t'$$
$$\Leftrightarrow$$
$$\frac{t}{4} \leq s \leq t \qquad \text{and} \qquad \max\{s, \tfrac{t}{2}\} \leq t' \leq \min\{2s, t\},$$

we obtain by switching the order of the integrals:

$$
\begin{aligned}
[\text{LHS of (48)}] \quad &\leq \quad \frac{1}{t} \int_{\frac{t}{4}}^{t} \int_{s}^{t} (t'-s+1)^{-\gamma} \, dt' \, b(s) \, ds = \frac{1}{t} \int_{\frac{t}{4}}^{t} \int_{0}^{t-s} (t''+1)^{-\gamma} \, dt'' \, b(s) \, ds \\
&\overset{\gamma \geq 1}{\lesssim} \quad \int_{\frac{t}{4}}^{t} b(s) \, ds \times \frac{1}{t} \int_{0}^{t} (t''+1)^{-1} \, dt'' \overset{(42)}{\lesssim} (t+1)^{1-\gamma-\frac{1}{p}} \Lambda(t) \frac{\ln(t+1)}{t} \\
&\overset{t \geq 1}{\lesssim} \quad (t+1)^{-\gamma} \Lambda(t) \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}}.
\end{aligned}
$$

The claim of Step 2 follows from the combination of (44)—(48).

Step 3. Proof of (41).

For $\tau \gg 1$, $\frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}}$ is monotone decreasing for $t \geq \tau$. Hence for $\tau \gg 1$, Step 2 can be upgraded to

$$(t+1)^{\gamma} a(t) \lesssim 1 + \tau^{1-\frac{1}{p}} + \left(\tau^{1-\gamma-\frac{1}{p}} + \frac{\ln(\tau+1)}{(\tau+1)^{\frac{1}{p}}} \right) \Lambda(t).$$

19

Because of $\gamma \geq 1$ and $p < \infty$, the expressions $\tau^{1-\gamma-\frac{1}{p}}$ and $\frac{\ln(\tau+1)}{(\tau+1)^{\frac{1}{p}}}$ tend to zero as $\tau \to \infty$. Hence, by Step 2 we can find a threshold $\tau_0 > 0$ only depending on p, γ and C_1 such that for all $t \geq 4\tau_0$

$$(t+1)^\gamma a(t) \leq C_2 \left(1 + \tau_0^{1-\frac{1}{p}}\right) + \frac{1}{2}\Lambda(t),$$

where C_2 is a constant only depending on p, γ and C_1. For all $t > 0$ we then have

$$
\begin{aligned}
\Lambda(t) &\leq \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + \sup_{4\tau_0 \leq s \leq t} (1+s)^\gamma a(s) \\
&\leq \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + C_2 \left(1 + \tau_0^{1-\frac{1}{p}}\right) + \frac{1}{2}\Lambda(t),
\end{aligned}
$$

that is

$$\Lambda(t) \leq 2 \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + 2C_2(1 + \tau_0^{1-\frac{1}{p}}).$$

Since

$$\sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) \overset{(28b)}{\lesssim} (1+4\tau_0)^{\gamma+1-\frac{1}{p}} a(0) \overset{(28a)}{\lesssim} (1+4\tau_0)^{\gamma+1-\frac{1}{p}},$$

we deduce that $\Lambda(t)$ is bounded for all $t > 0$ by a constant that only depends on p, γ, and C_1. $\qquad \square$

Proof of Lemma 7. Let $\overline{v}(x) := \overline{D^*\xi}(x)$ denote the stationary extension of the initial data. Since $\overline{\xi}$ is almost surely L-periodic, we have

$$\sum_{x \in ([0,L) \cap \mathbb{Z})^d} \overline{v}(x) = \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \nabla^*\overline{\xi}(x) = 0 \qquad \langle \cdot \rangle\text{-almost surely,}$$

and consequently

$$\sum_{x \in ([0,L) \cap \mathbb{Z})^d} \overline{u}(t,x) = 0 \qquad \langle \cdot \rangle\text{-almost surely.}$$

From the Poincaré inequality for mean free L-periodic functions, we deduce that

$$\sum_{x \in ([0,L) \cap \mathbb{Z})^d} |\overline{u}(t,x)|^2 \leq \frac{L^2}{c_0} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} |\nabla \overline{u}(t,x)|^2 \qquad \langle \cdot \rangle\text{-almost surely,}$$

and thus, by stationarity,

$$\langle u^2(t) \rangle \leq \frac{L^2}{c_0} \langle |Du(t)|^2 \rangle \leq \frac{L^2}{c_0\lambda} \langle Du(t) \cdot \boldsymbol{a}(0) Du(t) \rangle = -\frac{L^2}{2c_0\lambda} \frac{d}{dt} \langle u^2(t) \rangle,$$

so that the claim follows by Gronwall's lemma. $\qquad \square$

4 Estimates on the corrector in stochastic homogenization

The corrector ϕ can be formally recovered by integrating in time the function

$$(49) \qquad\qquad u(t) := \exp(-tD^*\boldsymbol{a}(0)D)\mathfrak{d},$$

where $\mathfrak{d} := -D^*\boldsymbol{a}(0)e$ is the RHS of the corrector equation (4). Making this connection rigorous for the modified and periodic correctors yields the following moment bounds.

Proposition 1. *Let $d \geq 2$ and let $e \in \mathbb{R}^d$ be an arbitrary direction with $|e| = 1$.*

(a) *(Modified corrector).* *Let $\mu > 0$. Assume that $\langle\cdot\rangle$ satisfies either* (11) *or* (8) *(for some $L \in \mathbb{N}$). Then the unique solution $\phi_\mu \in L^2(\Omega)$ of the modified corrector equation* (5) *satisfies*

$$\langle |\phi_\mu|^p \rangle^{\frac{1}{p}} \lesssim \begin{cases} \ln^{\frac{1}{2}}(\frac{1}{\mu} + 1) & \text{for } d = 2 \text{ and } 1 \le p \le 2, \\ \ln(\frac{1}{\mu} + 1) & \text{for } d = 2 \text{ and } p > 2, \\ 1 & \text{for } d > 2, \end{cases}$$

$$\langle |D\phi_\mu|^p \rangle^{\frac{1}{p}} \lesssim 1,$$

for all $1 \le p < \infty$.

(b) *(L-periodic corrector).* *Let $L \in \mathbb{N}$ and assume that $\langle\cdot\rangle$ satisfies* (8). *Then the unique solution $\phi \in L^2(\Omega)$ of the corrector equation* (4) *in the form of* (6) *satisfies*

$$\langle |\phi|^p \rangle^{\frac{1}{p}} \lesssim \begin{cases} \ln^{\frac{1}{2}}(L + 1) & \text{for } d = 2 \text{ and } 1 \le p \le 2, \\ \ln(L + 1) & \text{for } d = 2 \text{ and } p > 2, \\ 1 & \text{for } d > 2, \end{cases}$$

$$\langle |D\phi|^p \rangle^{\frac{1}{p}} \lesssim 1,$$

for all $1 \le p < \infty$.

For both (a) and (b), the multiplicative constants only depend on p, ρ, λ, and d.

(The proof is postponed to the end of this section). For $d > 2$ in the case of (11), the boundedness of the corrector was obtained originally in [23, Proposition 2.1] (the spectral gap estimate $\text{SG}_\infty(\rho)$ indeed implies the version the spectral gap estimate used in [23, Lemma 2.3]). For $d = 2$ in the case of (11), and for $d \ge 2$ in the case of (8), these results are new. In terms of scaling in μ and L, these estimates are optimal except for $d = 2$ and $p > 2$. Yet, using the spectral gap estimate on the corrector equation and bounds on the elliptic Green's function, one can prove that for $d = 2$ and for all $p \ge 1$ (see e.g. [25] in the case of continuum elliptic equations),

$$\langle |\phi_\mu|^p \rangle^{\frac{1}{p}} \lesssim \ln^{\frac{1}{2}}(\frac{1}{\mu} + 1) \langle |D\phi_\mu|^p \rangle^{\frac{1}{p}},$$

so that the optimal bound for $p > 2$ follows from the bounds on $\langle |D\phi_\mu|^p \rangle$ in Proposition 1. The same argument holds for periodic ensembles.

Remark 4. *In dimensions $d > 2$ Proposition 1 implies that the corrector equation* (4) *admits a unique solution $\phi \in L^2(\Omega)$ with $\langle \phi \rangle = 0$. In addition, for all $1 \le p < \infty$,*

$$\langle |\phi|^p \rangle^{\frac{1}{p}} \lesssim 1.$$

In dimension $d = 2$ stationary correctors do not exist. However, as $\mu \downarrow 0$, $D\phi_\mu$ converges to some potential field Ψ in $L^2(\Omega)^d$ which, in view of Proposition 1, satisfies for all $1 \le p < \infty$

$$\langle |\Psi|^p \rangle^{\frac{1}{p}} \lesssim 1.$$

Proof of Proposition 1. Both statements (a) and (b) follow from Theorem 1 using the same strategy, and we only prove the former. Let u be given by (49). Since for any $1 \le p \le \infty$ we have $\sum_{y \in \mathbb{Z}^d} \left\langle \left(\frac{\partial \mathfrak{d}}{\partial y}\right)^{2p} \right\rangle^{\frac{1}{2p}} \le 2$, Theorem 1 combined with Jensen's inequality in probability yields

(50) $$\langle |u(t)|^p \rangle^{\frac{1}{p}} \lesssim (t + 1)^{-(\frac{d}{4} + \frac{1}{2})} \qquad \text{for all } 1 \le p < \infty,$$

from which we deduce that

$$\phi_\mu := \int_0^\infty \exp(-\mu t) u(t) \, dt$$

21

defines a function in $L^p(\Omega)$. By construction ϕ_μ satisfies the identity:

$$
\begin{aligned}
\mu\phi_\mu &= \mu \int_0^\infty \exp(-\mu t) u(t)\, dt \\
&= -\int_0^\infty \frac{\partial}{\partial t}(\exp(-\mu t)u(t)) + \exp(-\mu t)\frac{\partial}{\partial t}u(t)\, dt \\
&= u(0) - \int_0^\infty D^*\boldsymbol{a}(0)D(\exp(-\mu t)u(t))\, dt \\
&= \mathfrak{d} - D^*\boldsymbol{a}(0)D\phi_\mu.
\end{aligned}
$$

It remains to establish the desired estimates. By the triangle inequality in $L^p(\Omega)$ and (50) we have

$$
\begin{aligned}
\langle \phi_\mu^p \rangle^{\frac{1}{p}} &\leq \int_0^\infty \exp(-\mu t)\, \langle |u(t)|^p \rangle^{\frac{1}{p}}\, dt \lesssim \int_0^\infty \exp(-\mu t)(t+1)^{-(\frac{d}{4}+\frac{1}{2})}\, dt \\
&\lesssim \begin{cases} \ln(\frac{1}{\mu}+1) & \text{for } d=2, \\ 1 & \text{for } d>2. \end{cases}
\end{aligned}
$$

For the sharper estimates on ϕ_μ for $d=2$ and $p\leq 2$, it is enough to use the semi-group property for $p=2$. Since $D^*\boldsymbol{a}(0)D$ is symmetric, we have for all $t,s\geq 0$, $\langle u(s)u(t)\rangle = \langle u^2(\frac{s+t}{2})\rangle$. Hence,

$$
\begin{aligned}
\left\langle \left(\int_0^\infty u(t)\exp(-\mu t)\, dt \right)^2 \right\rangle &= \int_0^\infty \int_0^\infty \langle u^2(\tfrac{s+t}{2})\exp(-\mu(s+t))\rangle\, dt\, ds \\
&= 2\int_0^\infty \int_{\frac{s}{2}}^\infty \langle u^2(\tau)\exp(-2\mu\tau)\rangle\, d\tau\, ds \\
&\overset{(50)}{\lesssim} \int_0^\infty \int_{\frac{s}{2}}^\infty \exp(-2\mu\tau)(\tau+1)^{-2}\, d\tau\, ds \\
&\lesssim \ln(\tfrac{1}{\mu}+1),
\end{aligned}
$$

as desired.

It remains to prove the estimates on the gradient $D\phi_\mu$ for $d=2$ (for $d>2$ this follows by discreteness and the estimate on ϕ_μ itself). By Jensen's inequality in probability, it is enough to prove the claim for $p=2\tilde{p}$ with $\tilde{p}\geq 1$ so that one can appeal to the Caccioppoli inequality of Lemma 5. We fix $\gamma \in (p-1,p)$ and note that this yields both

$$
(51) \qquad \int_0^\infty (t+1)^{-\frac{\gamma}{p-1}}\, dt < \infty \qquad \text{and} \qquad \int_0^\infty (t+1)^{\gamma-1-p}\, dt < \infty.
$$

As announced, we then have:

$$
\begin{aligned}
\langle |D\phi_\mu|^p \rangle^{\frac{1}{p}} &\overset{\triangle\text{-inequality}}{\leq} \int_0^\infty \langle |\exp(-\mu t)Du(t)|^p \rangle^{\frac{1}{p}}\, dt \\
&\leq \int_0^\infty \langle |Du(t)|^p \rangle^{\frac{1}{p}} (t+1)^{\frac{\gamma}{p}}(t+1)^{-\frac{\gamma}{p}}\, dt \\
&\overset{\text{Hölder}}{\leq} \left(\int_0^\infty \langle |Du(t)|^p \rangle (t+1)^\gamma\, dt \right)^{\frac{1}{p}} \left(\int_0^\infty (t+1)^{-\frac{\gamma}{p-1}}\, dt \right)^{\frac{p-1}{p}} \\
&\overset{\underset{\text{Lemma 5}}{(51)}}{\lesssim} \left(\int_0^\infty -\frac{d}{dt}\langle |u(t)|^p \rangle (t+1)^\gamma\, dt \right)^{\frac{1}{p}} \\
&\overset{\langle |u(0)|^p\rangle \lesssim 1}{\lesssim} \left(\int_0^\infty \langle |u(t)|^p \rangle (t+1)^{\gamma-1}\, dt \right)^{\frac{1}{p}} + 1 \\
&\overset{(50),d=2}{\lesssim} \left(\int_0^\infty (t+1)^{\gamma-1-p}\, dt \right)^{\frac{1}{p}} + 1 \overset{(51)}{\lesssim} 1.
\end{aligned}
$$

$\square$

5   Approximation of the homogenized coefficients by periodization

In this section we prove Theorem 2 which establishes an optimal error estimate for the approximation of $\boldsymbol{a}_{\mathrm{hom}}$ by periodization. We present a complete analysis for i. i. d. coefficients and partial results for ensembles satisfying (SG).

5.1   Auxiliary results and proof of Theorem 2

The mathematical version of the periodization method is the following: Let $\langle\cdot\rangle$ denote a stationary and ergodic ensemble and $\boldsymbol{a}_{\mathrm{hom}}$ the homogenized matrix associated with $\langle\cdot\rangle$ via (1). To approximate $\boldsymbol{a}_{\mathrm{hom}}$ we

- approximate $\boldsymbol{a}_{\mathrm{hom}}$ by $\boldsymbol{a}_{\mathrm{hom},L}$, where $\boldsymbol{a}_{\mathrm{hom},L}$ is the homogenized coefficient associated via (1) with a suitable stationary L-periodic ensemble $\langle\cdot\rangle_L$, which we think of as having the same "specifications" as $\langle\cdot\rangle$. This introduces a systematic error, which we will only fully study in the case of the i. i. d. ensembles;

- approximate $\boldsymbol{a}_{\mathrm{hom},L}$ by $\boldsymbol{a}_{\mathrm{av},L}$, cf. (12), for some realization $\boldsymbol{a}$ distributed according to $\langle\cdot\rangle_L$. By stationarity and periodicity we have

$$
e \cdot \boldsymbol{a}_{\mathrm{av},L}(\boldsymbol{a}) e = L^{-d} \sum_{x \in ([0,L)\cap\mathbb{Z})^d} (\nabla\overline{\phi}(x) + e) \cdot \boldsymbol{a}(x)(\nabla\overline{\phi}(x) + e) \quad \text{and} \quad \langle\boldsymbol{a}_{\mathrm{av},L}\rangle_L = \boldsymbol{a}_{\mathrm{hom},L},
$$

where $\overline{\phi}$ is defined via (6). Replacing $\boldsymbol{a}_{\mathrm{hom},L}$ by $\boldsymbol{a}_{\mathrm{av},L}$ introduces a random error.

Remark 5. *In case of the i. i. d. ensemble studied in this paper, the definition of the stationary periodic ensemble $\langle\cdot\rangle_L$ and its coupling to $\langle\cdot\rangle$ is natural, see Definition 3. For a general ensemble, this is more subtle, and we shall now sketch two situations in which it is clear what to do.*

1) Suppose that $\langle\cdot\rangle$ is the push-forward of an ensemble $\langle\cdot\rangle_0$ under a transformation $\Phi\colon \Omega \mapsto \Omega$, which is stationary in the sense of $\Phi(\boldsymbol{a})(x + z) = \Phi(\boldsymbol{a}(\cdot + z))(x)$ (in order to preserve stationarity of the ensemble), and which we think of being short-ranged (e. g. a convolution operator with integrable kernel). Suppose further that the base measure $\langle\cdot\rangle_0$ is such that there exists a natural stationary L-periodic version $\langle\cdot\rangle_{0,L}$ and that it is naturally coupled to $\langle\cdot\rangle_0$. The latter means that there exists an ensemble $\langle\cdot\rangle_{0,c}$ on the product space $(\boldsymbol{a}, \boldsymbol{b}) \in ([0,L)\cap\mathbb{Z})^d \times \mathbb{Z}^d$ such that the distribution of $\boldsymbol{a}$ is that of $\langle\cdot\rangle_{0,L}$, the distribution of $\boldsymbol{b}$ is that of $\langle\cdot\rangle_0$, and that $\boldsymbol{a} \approx \boldsymbol{b}$ in (the middle of) $([0,L)\cap\mathbb{Z})^d$ (as would be the case for the i. i. d. ensemble). Then a suitable L-periodic ensemble $\langle\cdot\rangle_L$ and coupling $\langle\cdot\rangle_c$ is given by the push forward of $\langle\cdot\rangle_{0,L}$ and $\langle\cdot\rangle_{0,c}$ under Φ.

2) Suppose that, in the jargon of statistical mechanics, the ensemble $\langle\cdot\rangle$ is an infinite-volume Gibbs measure, with a translation invariant formal Hamiltonian coming from finite-range specifications. To fix ideas, we consider the one-dimensional lattice (i. e. $d = 1$), a two-valued spin space (i. e. $\boldsymbol{a}(x) \in \{\lambda, 1\}$), and a nearest neighbor interaction (i. e. given by the specification $H(\boldsymbol{a}(x), \boldsymbol{a}(x+1))$). In this case, the natural definition of a stationary L-periodic ensemble is the following: The probability of a configuration $\boldsymbol{a} = (\boldsymbol{a}(0), \cdots, \boldsymbol{a}(L-1))$ is given by

$$
P_L(\boldsymbol{a}) = \frac{1}{Z} \exp\big(-\sum_{x=0}^{L-1} H(\boldsymbol{a}(x), \boldsymbol{a}(x+1)) - H(\boldsymbol{a}(L-1), \boldsymbol{a}(0))\big),
$$

where Z is a normalization constant. The coupling is more subtle: consider $\langle\cdot|\boldsymbol{a}(0)\rangle_L$, i. e. the L-periodic ensemble $\langle\cdot|\boldsymbol{a}(0)\rangle_L$ conditioned on $\boldsymbol{a}(0)$, and $\langle\cdot|\boldsymbol{b}(0), \boldsymbol{b}(L)\rangle$, the infinite ensemble conditioned on $(\boldsymbol{b}(0), \boldsymbol{b}(L))$ (we use the latter $\boldsymbol{b}$ for the random variable of the infinite ensemble to avoid confusion). Because the interaction is nearest neighbor, $(\boldsymbol{b}(1), \cdots, \boldsymbol{b}(L-1))$ and $\{\boldsymbol{b}(x)\}_{x\notin\{0,\cdots,L\}}$ are independent under $\langle\cdot|\boldsymbol{b}(0), \boldsymbol{b}(L)\rangle$. Because the interaction is stationary, $(\boldsymbol{a}(1), \cdots, \boldsymbol{a}(L-1))$ and $(\boldsymbol{b}(1), \cdots, \boldsymbol{b}(L-1))$

are identically distributed under $\langle \cdot | \boldsymbol{a}(0) \rangle_L$ and $\langle \cdot | \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle$ provided $\boldsymbol{a}(0) = \boldsymbol{b}(0) = \boldsymbol{b}(L)$. Hence it is straightforward to couple the conditional measures $\langle \cdot | \boldsymbol{a}(0) \rangle_L$ and $\langle \cdot | \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle$ in such a way that $(\boldsymbol{a}(1), \cdots, \boldsymbol{a}(L-1)) = (\boldsymbol{b}(1), \cdots, \boldsymbol{b}(L-1))$ provided $\boldsymbol{a}(0) = \boldsymbol{b}(0) = \boldsymbol{b}(L)$. Let $\langle \cdot | \boldsymbol{a}(0), \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle_c$ denote this coupling. Then a coupling $\langle \cdot \rangle_c$ of $\langle \cdot \rangle_L$ and $\langle \cdot \rangle$ is given as follows: Let $P_L(\boldsymbol{a}(0))$ denote the probability of $\boldsymbol{a}(0)$ under $\langle \cdot \rangle_L$ and let $P(\boldsymbol{b}(0), \boldsymbol{b}(L))$ denote the probability of $(\boldsymbol{b}(0), \boldsymbol{b}(1))$ under $\langle \cdot \rangle$. For any function $\zeta = \zeta(\boldsymbol{a}; \boldsymbol{b})$ we set

$$\langle \zeta \rangle_c = \sum_{(\boldsymbol{b}(0), \boldsymbol{b}(L)) \in \{\lambda, 1\}^2} \sum_{\boldsymbol{a}(0) \in \{\lambda, 1\}} \langle \zeta | \boldsymbol{a}(0), \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle_c P_L(\boldsymbol{a}(0)) P(\boldsymbol{b}(0), \boldsymbol{b}(L)).$$

This is a good coupling: Since the distribution of $\boldsymbol{a}(L/2)$ under $\langle \cdot | \boldsymbol{a}(0) \rangle_L$ depends only weakly (exponentially weakly in $L \gg 1$) on $\boldsymbol{a}(0)$ and likewise the distribution of $\boldsymbol{b}(L/2)$ under $\langle \cdot | \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle$ depends only weakly on $(\boldsymbol{b}(0), \boldsymbol{b}(L))$, we have $\boldsymbol{a}(L/2) \approx \boldsymbol{b}(L/2)$ under the distribution of $\langle \cdot | \boldsymbol{a}(0), \boldsymbol{b}(0), \boldsymbol{b}(L) \rangle_c$ and thus under $\langle \cdot \rangle_c$.

We first discuss the random error, which is the variance of $\boldsymbol{a}_{\mathrm{av},L}$. Using the bound on the quartic moment of the gradient of the corrector in Proposition 1 (b), we shall prove the following optimal estimate:

Proposition 2 (Optimal variance estimate). *Let $d \geq 2$, $\langle \cdot \rangle$ be stationary, L-periodic and satisfy $SG_L(\rho)$. Then for all $e \in \mathbb{R}^d$, $|e| = 1$, we have*

$$\mathrm{var}[e \cdot \boldsymbol{a}_{\mathrm{av},L} e] \lesssim L^{-d},$$

where the constant only depends on ρ, λ, and d.

Proposition 2 shows that the random error decays at the rate $L^{-\frac{d}{2}}$ of the central limit theorem. Since this error is due to fluctuations, its effect can be reduced by empirical averaging: For $N \in \mathbb{N}$ consider the random matrix $\boldsymbol{a}_{\mathrm{av},L,N}$ defined via (15) where $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^N$ denote N independent realizations of the coefficient field distributed according to $\langle \cdot \rangle$. Under the assumption of Proposition 2, for all $e \in \mathbb{R}^d$, $|e| = 1$, we then have

$$(52) \qquad \mathrm{var}[e \cdot \boldsymbol{a}_{\mathrm{av},L,N} e] = \frac{1}{N} \mathrm{var}[e \cdot \boldsymbol{a}_{\mathrm{av},L} e] \lesssim \frac{1}{N} L^{-d},$$

where the constant only depends on ρ, λ and d.

For the systematic error we prove the following estimate under the assumption that the coefficients are i. i. d.:

Proposition 3 (Optimal estimate of the systematic error). *Let $d \geq 2$, $\langle \cdot \rangle$ and $\langle \cdot \rangle_L$ be the infinite and L-periodic i. i. d. ensembles associated with the same measure β on Ω_0 (cf. Definition 3). Then we have*

$$|\boldsymbol{a}_{\mathrm{hom}} - \boldsymbol{a}_{\mathrm{hom},L}| \lesssim L^{-d} \ln^d L,$$

where the constant only depends on λ and d.

Evidently, Theorem 2 is a direct consequence of Proposition 2 in the form of (52) and of Proposition 3.

In order to estimate the systematic error we require additional "inner" approximations that rely on the modified corrector equation (5). To that end notice that – as a merit of the μ-regularization – the modified corrector ϕ_μ (associated with a direction e via (5)) can be defined independently of the ensemble: Indeed, pointwise in $\boldsymbol{a}$ we have $\phi_\mu(\boldsymbol{a}) = \overline{\phi}_\mu(\boldsymbol{a}, x = 0)$ where $\overline{\phi}_\mu(\boldsymbol{a}, \cdot)$ is the unique bounded solution of

$$(53) \qquad \mu \overline{\phi}_\mu(\boldsymbol{a}, x) + \nabla^* \boldsymbol{a}(x) \nabla \overline{\phi}_\mu(\boldsymbol{a}, x) = -\nabla^* \boldsymbol{a}(x) e, \qquad x \in \mathbb{Z}^d.$$

The resulting function $\phi_\mu : \Omega \to \mathbb{R}$ is continuous and thus a measurable solution of (5) (see [23, Lemma 2.6]).

24

Definition 4. (*μ-approximation*). *Let $\langle \cdot \rangle$ be a stationary ensemble. For $k \in \mathbb{N}_0$ and $\mu > 0$ we inductively define symmetric matrices $\boldsymbol{a}_{\mathrm{hom},\mu}^k$ via: For all $e \in \mathbb{R}^d$,*

$$
\begin{aligned}
e \cdot \boldsymbol{a}_{\mathrm{hom},\mu}^0 e \;&:=\; e \cdot \boldsymbol{a}_{\mathrm{hom},\mu} e := \langle (D\phi_\mu + e) \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \rangle, \\
\boldsymbol{a}_{\mathrm{hom},\mu}^k \;&:=\; \frac{1}{2^{k+1} - 1}\left(2^{k+1} \boldsymbol{a}_{\mathrm{hom},\mu}^{k-1} - \boldsymbol{a}_{\mathrm{hom},2\mu}^{k-1} \right),
\end{aligned}
$$

(54)

where ϕ_μ denotes the modified corrector (for direction e).

By elementary spectral theory one can show that $\boldsymbol{a}_{\mathrm{hom},\mu} \to \boldsymbol{a}_{\mathrm{hom}}$ as $\mu \downarrow 0$ for general ergodic ensembles. The matrix $\boldsymbol{a}_{\mathrm{hom},\mu}^k$ is a "higher-order" approximation for $\boldsymbol{a}_{\mathrm{hom}}$ based on a Richardson extrapolation. With these approximations at hand we split the systematic error $|\boldsymbol{a}_{\mathrm{hom},L} - \boldsymbol{a}_{\mathrm{hom}}|$ into two *systematic sub-errors* and a *coupling error*: For arbitrary $\mu > 0$ and $k \in \mathbb{N}_0$ we have

(55)
$$
\begin{aligned}
|\boldsymbol{a}_{\mathrm{hom},L} - \boldsymbol{a}_{\mathrm{hom}}| \;\leq\; & |\boldsymbol{a}_{\mathrm{hom},L,\mu}^k - \boldsymbol{a}_{\mathrm{hom},L}| + |\boldsymbol{a}_{\mathrm{hom},\mu}^k - \boldsymbol{a}_{\mathrm{hom}}| \quad \text{(systematic sub-errors)} \\
& + |\boldsymbol{a}_{\mathrm{hom},L,\mu}^k - \boldsymbol{a}_{\mathrm{hom},\mu}^k| \quad\quad\quad\quad\quad\quad \text{(coupling error)},
\end{aligned}
$$

where $\boldsymbol{a}_{\mathrm{hom},\mu}^k$ and $\boldsymbol{a}_{\mathrm{hom},L,\mu}^k$ are associated with $\langle \cdot \rangle$ and $\langle \cdot \rangle_L$ respectively, through the induction (54). We first discuss the systematic sub-errors, the estimate of which only requires (SG). Following [20], we estimate the systematic sub-errors by appealing to spectral calculus. Indeed, since the elliptic operator $D^* \boldsymbol{a}(0) D$ is bounded, symmetric, and non-negative on $L^2(\Omega)$, the spectral theorem yields the existence of a spectral measure $P(d\nu)$ on $[0, +\infty)$ such that for all $\zeta \in L^2(\Omega)$, and suitable continuous functions f, we have

$$
f(D^* \boldsymbol{a}(0) D)\zeta = \int_0^\infty f(\nu) P(d\nu)\zeta.
$$

As already used by Mourrat in [32], estimates on the semigroup allow one to quantify the bottom of the spectrum of $D^* \boldsymbol{a}(0) D$. In particular, by [20, Theorem 7] (see also [32, Theorem 2.4]), Theorem 1 answers positively a conjecture of [20] and yields:

Corollary 1. *Let $d \geq 2$. Let $\langle \cdot \rangle$ denote either a stationary ensemble that satisfies $SG_\infty(\rho)$, or a stationary, L-periodic ensemble that satisfies $SG_L(\rho)$. We let $P(d\nu)$ denote the spectral measure of the operator $D^* \boldsymbol{a}(0) D$, and let $\mathfrak{d} = -D^* \boldsymbol{a}(0) e$ denote the RHS of the corrector equation in some fixed direction $e \in \mathbb{R}^d$ with $|e| = 1$. Then, for all $\tilde{\nu} \geq 0$, we have*

$$
\int_0^{\tilde{\nu}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \lesssim \tilde{\nu}^{\frac{d}{2}+1},
$$

where the constant only depends on ρ, λ and d.

Based on this, we shall prove the following estimate of the systematic sub-error:

Lemma 8. *Let $d \geq 2$. Let $\langle \cdot \rangle$ denote either a stationary ensemble that satisfies $SG_\infty(\rho)$, or a stationary, L-periodic ensemble that satisfies $SG_L(\rho)$. Let $\boldsymbol{a}_{\mathrm{hom}}$ and $\boldsymbol{a}_{\mathrm{hom},\mu}^k$ be associated with $\langle \cdot \rangle$ via (1) and (54), respectively. Then for all non-negative integers $k > \frac{d}{2} - 2$ and $\mu > 0$ we have*

$$
|\boldsymbol{a}_{\mathrm{hom},\mu}^k - \boldsymbol{a}_{\mathrm{hom}}| \lesssim \mu^{\frac{d}{2}},
$$

where the constant only depends on k, ρ, λ, and d.

Note that this lemma applies to both errors $|\boldsymbol{a}_{\mathrm{hom},\mu}^k - \boldsymbol{a}_{\mathrm{hom}}|$ and $|\boldsymbol{a}_{\mathrm{hom},L,\mu}^k - \boldsymbol{a}_{\mathrm{hom},L}|$.

Finally, we study the *coupling error*. It is the only term that relates the original ensemble $\langle \cdot \rangle$ with the periodic proxy $\langle \cdot \rangle_L$. Hence we need to specify the approximation mechanism used to construct $\langle \cdot \rangle_L$ from $\langle \cdot \rangle$. In the present contribution we discuss this issue rigorously only for i. i. d. coefficients, for which the L-periodic proxy $\langle \cdot \rangle_L$ can unambiguously be obtained by using the same base measure β. For our purpose the following estimate on the *coupling error* is sufficient:

Lemma 9. *Let $d \geq 2$, $\langle \cdot \rangle$ and $\langle \cdot \rangle_L$ be the infinite and L-periodic i. i. d. ensembles associated with the same measure β on Ω_0 (cf. Definition 3). Then there exist $\alpha > 0$ only depending on d and $c_0 > 0$ only depending on λ and d, such that for all $\mu \in (0,1]$ and any direction $e \in \mathbb{R}^d$, $|e| = 1$, we have*

$$
(56) \qquad |e \cdot \boldsymbol{a}_{\mathrm{hom},L,\mu} e - e \cdot \boldsymbol{a}_{\mathrm{hom},\mu} e| \lesssim \sqrt{\mu}^{-\alpha} \exp(-c_0 \sqrt{\mu} L),
$$

where the constant only depends on λ and d.

5.2 Proofs of the auxiliary results

Proof of Proposition 2. Set $\mathcal{E} := e \cdot \boldsymbol{a}_{\mathrm{av},L} e = L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left(\nabla \overline{\phi}(x) + e \right) \cdot \boldsymbol{a}(x) \left(\nabla \overline{\phi}(x) + e \right)$. In the following we drop the subindex L in the notation for $\frac{\partial}{\partial_L y}$ and $\langle \cdot \rangle_{L,y}$.

Step 1. Estimate of $\nabla \frac{\partial \overline{\phi}}{\partial y}$:

$$
(57) \qquad \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left| \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \right|^2 \lesssim \left\langle |e + \nabla \overline{\phi}(y)|^2 \right\rangle_y.
$$

We apply the vertical derivative $\frac{\partial}{\partial y}$ to (6):

$$
0 = \nabla^* \boldsymbol{a}(x) \nabla \frac{\partial \overline{\phi}(x)}{\partial y} + \nabla^* \left(\boldsymbol{a}(x) \left\langle \nabla \overline{\phi}(x) + e \right\rangle_y - \left\langle \boldsymbol{a}(x)(\nabla \overline{\phi}(x) + e) \right\rangle_y \right).
$$

Hence, $\frac{\partial \overline{\phi}(\boldsymbol{a},\cdot)}{\partial y}$ is an L-periodic function and satisfies

$$
(58) \qquad \nabla^* \boldsymbol{a}(x) \nabla \frac{\partial \overline{\phi}(\boldsymbol{a},x)}{\partial y} = \nabla^* \xi(\boldsymbol{a},x,y) \qquad \text{for all } x, y \in \mathbb{Z}^d \text{ and } \langle \cdot \rangle\text{-almost every } \boldsymbol{a} \in \Omega,
$$

with a RHS given by

$$
\xi(\boldsymbol{a},x,y) = \left\langle \boldsymbol{a}(x)(\nabla \overline{\phi}(x) + e) \right\rangle_y - \boldsymbol{a}(x) \left\langle \nabla \overline{\phi}(x) + e \right\rangle_y.
$$

This yields $\xi(\boldsymbol{a},x,y) = 0$ whenever $x - y \notin L\mathbb{Z}^d$. The weak formulation of (58) with (periodic) test-function $\frac{\partial \overline{\phi}}{\partial y}$ yields the a priori estimate

$$
\lambda \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left| \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \right|^2 \leq \nabla \frac{\partial \overline{\phi}}{\partial y}(y) \cdot \xi(y,y) \leq 2 \left| \nabla \frac{\partial \overline{\phi}}{\partial y}(y) \right| \left\langle |e + \nabla \overline{\phi}(y)| \right\rangle_y.
$$

Combined with $\left| \nabla \frac{\partial \overline{\phi}}{\partial y}(y) \right|^2 \leq \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left| \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \right|^2$, we get (57).

Step 2. Estimate of $\frac{\partial \mathcal{E}}{\partial y}$:

$$
\left| \frac{\partial \mathcal{E}}{\partial y} \right| \lesssim L^{-d} \left(|e + \nabla \overline{\phi}(y)|^2 + \left\langle |e + \nabla \overline{\phi}(y)|^2 \right\rangle_y \right).
$$

Indeed, by the following vertical Leibniz rule

$$
\frac{\partial(\zeta_1 \zeta_2)}{\partial y} = \zeta_1 \frac{\partial \zeta_2}{\partial y} + \frac{\partial \zeta_1}{\partial y} \zeta_2 - \frac{\partial \zeta_1}{\partial y} \frac{\partial \zeta_2}{\partial y} - \left\langle \frac{\partial \zeta_1}{\partial y} \frac{\partial \zeta_2}{\partial y} \right\rangle_y,
$$

we have

$$
\begin{aligned}
\frac{\partial \mathcal{E}}{\partial y} = \ & L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x) : \frac{\partial}{\partial y} \left(\left(e + \nabla \overline{\phi}(x) \right) \otimes \left(e + \nabla \overline{\phi}(x) \right) \right) \\
& + L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \frac{\partial \boldsymbol{a}(x)}{\partial y} : \left(e + \nabla \overline{\phi}(x) \right) \otimes \left(e + \nabla \overline{\phi}(x) \right) \\
& - L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \frac{\partial \boldsymbol{a}(x)}{\partial y} : \frac{\partial}{\partial y} \left(\left(e + \nabla \overline{\phi}(x) \right) \otimes \left(e + \nabla \overline{\phi}(x) \right) \right) \\
& - L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \left\langle \frac{\partial \boldsymbol{a}(x)}{\partial y} : \frac{\partial}{\partial y} \left(\left(e + \nabla \overline{\phi}(x) \right) \otimes \left(e + \nabla \overline{\phi}(x) \right) \right) \right\rangle_y.
\end{aligned}
$$

26

For convenience, we denote by I_1 the first term of the RHS, and by I_2 the sum of the other three terms. Since

$$
\left|\frac{\partial \boldsymbol{a}(x)}{\partial y}\right| \leq \begin{cases} 2 & \text{if } x - y \in L\mathbb{Z}^d, \\ 0 & \text{else}, \end{cases}
$$

and $|\frac{\partial \zeta}{\partial y}| \leq |\zeta| + \langle |\zeta| \rangle_y$ for all $\zeta \in L^2(\Omega)$, we have

$$
|I_2| \lesssim L^{-d} \left(|\nabla \overline{\phi}(y) + e|^2 + \langle |\nabla \overline{\phi}(y) + e|^2 \rangle_y \right).
$$

To estimate I_1 we appeal again to the vertical Leibniz rule:

$$
\begin{aligned}
I_1 &= L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x) : \frac{\partial}{\partial y} \Big((e + \nabla \overline{\phi}(x)) \otimes (e + \nabla \overline{\phi}(x)) \Big) \\
&= 2L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x) : (e + \nabla \overline{\phi}(x)) \otimes \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \\
&\quad - L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x) : \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \otimes \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \\
&\quad - L^{-d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \boldsymbol{a}(x) : \left\langle \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \otimes \nabla \frac{\partial \overline{\phi}}{\partial y}(x) \right\rangle_y.
\end{aligned}
$$

The first term of the RHS vanishes identically by (6), whereas the last two terms are controlled by Step 1.

Step 3. Conclusion via Spectral Gap Estimate.

We apply the Spectral Gap Estimate to $e \cdot \boldsymbol{a}_{L,\text{hom}} e$, use Step 2, and then Jensen's inequality to bound the variance of $e \cdot \boldsymbol{a}_{L,\text{hom}} e$ by the quartic moment of $D\phi$:

$$
\text{var}[e \cdot \boldsymbol{a}_{L,\text{hom}} e] = \langle (\mathcal{E} - \langle \mathcal{E} \rangle)^2 \rangle \leq \frac{1}{\rho} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left(\frac{\partial \mathcal{E}}{\partial y} \right)^2 \right\rangle
$$

$$
\overset{\text{Step 2}}{\lesssim} L^{-2d} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle |e + \nabla \overline{\phi}(y)|^4 \right\rangle + \left\langle \langle |e + \nabla \overline{\phi}(y)|^2 \rangle_y^2 \right\rangle \lesssim L^{-d} \left\langle |e + D\phi|^4 \right\rangle,
$$

so that the claim follows from Proposition 1 (b). $\qquad \square$

We only display the proofs of Corollary 1 and Lemma 8 in the case (11) since the argument in the periodic case (8) is similar.

Proof of Corollary 1. We simply apply the semigroup $t \mapsto \exp(-tD^* \boldsymbol{a}(0)D)$ of Theorem 1 to $\mathfrak{d}$, and define for all $t \geq 0$

$$
u(t) := \exp(-tD^* \boldsymbol{a}(0)D)\mathfrak{d}.
$$

By the spectral theorem, $u(t) = \int_0^\infty \exp(-t\nu) P(d\nu) \mathfrak{d}$, so that for all $t > 0$,

$$
\langle u^2(t) \rangle = \int_0^\infty \exp(-2t\nu) \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \geq \exp(-2) \int_0^{\frac{1}{t}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.
$$

Corollary 1 thus follows from the estimate $\langle u^2(t) \rangle \lesssim (t+1)^{-(\frac{d}{2}+1)}$ of Theorem 1. $\qquad \square$

Proof of Lemma 8. We start with writing the Richardson extrapolation in spectral variables in the spirit of [37] and [27]. First, notice that by (4) we have

$$
\begin{aligned}
(59) \quad e \cdot \boldsymbol{a}_{\text{hom},\mu} e &= \langle (D\phi_\mu + e) \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \rangle \\
&= \langle e \cdot \boldsymbol{a}(0)e \rangle + 2\langle D\phi_\mu \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \rangle - \langle D\phi_\mu \cdot \boldsymbol{a}(0)D\phi_\mu \rangle \\
&= \langle e \cdot \boldsymbol{a}(0)e \rangle - 2\mu \langle \phi_\mu^2 \rangle - \langle \phi_\mu (D^* \boldsymbol{a}(0)D)\phi_\mu \rangle.
\end{aligned}
$$

27

Since $\phi_\mu = (\mu + D^*\boldsymbol{a}(0)D)^{-1}\mathfrak{d}$, the spectral theorem yields $\phi_\mu = \int_0^\infty \frac{1}{\mu+\nu} P(d\nu)\mathfrak{d}$, and thus (59) turns into the spectral representation formula:

$$e \cdot \boldsymbol{a}_{\mathrm{hom},\mu} e = \langle e \cdot \boldsymbol{a}(0)e \rangle - \int_0^\infty \frac{2\mu + \nu}{(\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

By the Lebesgue dominated convergence theorem, we obtain in the limit $\mu \downarrow 0$

$$(60) \qquad e \cdot \boldsymbol{a}_{\mathrm{hom}} e = \langle e \cdot \boldsymbol{a}(0)e \rangle - \int_0^\infty \frac{1}{\nu} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

The combination of both yields

$$e \cdot (\boldsymbol{a}_{\mathrm{hom},\mu} - \boldsymbol{a}_{\mathrm{hom}})e = \mu^2 \int_0^\infty \frac{1}{\nu(\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

From (54) we conclude by an elementary induction argument that for all $k \in \mathbb{N}_0$,

$$e \cdot (\boldsymbol{a}_{\mathrm{hom},\mu}^k - \boldsymbol{a}_{\mathrm{hom}})e = \mu^{k+2} \int_0^\infty \frac{p_k(\nu,\mu)}{\nu(2^0\mu + \nu)^2 \cdots (2^k\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle,$$

where p_k denotes a linear combination of monomials $\mu^i \nu^j$ of total degree $i + j = k$. Since $|p_k(\nu,\mu)| \lesssim (\mu + \nu)^k$ for all $\nu, \mu \geq 0$,

$$|e \cdot (\boldsymbol{a}_{\mathrm{hom},\mu}^k - \boldsymbol{a}_{\mathrm{hom}})e| \lesssim \mu^{k+2} \int_0^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

Corollary 1 implies for all monotone decreasing functions $f(\nu)$:

$$\int_0^1 f(\nu)\langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \lesssim \int_0^1 f(\nu)\nu^{\frac{d}{2}} d\nu,$$

and thus

$$\int_0^1 \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \lesssim \int_0^1 \nu^{\frac{d}{2}-1} \frac{1}{(\mu + \nu)^{k+2}} d\nu = \mu^{\frac{d}{2}-2-k} \int_0^{\frac{1}{\mu}} \frac{\tilde{\nu}^{\frac{d}{2}-1}}{(1+\tilde{\nu})^{k+2}} d\tilde{\nu}$$

$$\lesssim \mu^{\frac{d}{2}-2-k},$$

where we used that $\tilde{\nu} \mapsto \frac{\tilde{\nu}^{\frac{d}{2}-1}}{(1+\tilde{\nu})^{k+2}}$ is integrable on $(0,\infty)$ for $k > \frac{d}{2}-2$. Combined with $\int_0^\infty \frac{1}{\nu}\langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \leq 1$ (cf. (60)) the claim of the lemma follows. $\qquad\square$

Proof of Lemma 9. We split the proof into four steps.
Step 1. Definition of the coupling.
For $L \in \mathbb{N}$ consider the periodization mapping

$$T_L : \Omega \to \Omega_L, \qquad \boldsymbol{a} \mapsto T_L \boldsymbol{a},$$

where $T_L \boldsymbol{a}$ denotes the unique element in Ω_L such that $\boldsymbol{a}(x) = T_L \boldsymbol{a}(x)$ for all $x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d$. Since $\langle \cdot \rangle$ is a product measure, $\langle \cdot \rangle_L$ is obviously the pushforward of $\langle \cdot \rangle$ under T_L, i. e.

$$(61) \qquad \langle f \rangle_L = \langle f \circ T_L \rangle \qquad \text{for all } \langle \cdot \rangle_L\text{-measurable } f.$$

Step 2. Reduction to an estimate on the corrector.
We claim that

$$(62) \qquad |e \cdot \boldsymbol{a}_{\mathrm{hom},L,\mu} e - e \cdot \boldsymbol{a}_{\mathrm{hom},\mu} e| \lesssim \sum_{x_0 \in \{0,e_1,\ldots,e_d\}} \langle |\phi_\mu \circ T_{x_0} \circ T_L \circ T_{-x_0} - \phi_\mu|^2 \rangle^{\frac{1}{2}},$$

28

where $T_x : \Omega \to \Omega$, $\boldsymbol{a} \mapsto \boldsymbol{a}(\cdot + x)$ is the shift operator. Thanks to (61) we have

$$
\begin{aligned}
e \cdot \boldsymbol{a}_{\mathrm{hom},L,\mu} e \quad &= \quad \langle (D\phi_\mu + e) \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \rangle_L \\
&= \quad \left\langle \Big((D\phi_\mu + e) \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \Big) \circ T_L \right\rangle \\
\overset{\boldsymbol{a}(0) \circ T_L = \boldsymbol{a}(0)}{=} \quad & \langle (D\phi_\mu \circ T_L + e) \cdot \boldsymbol{a}(0)(D\phi_\mu \circ T_L + e) \rangle,
\end{aligned}
$$

so that

$$
\begin{aligned}
e \cdot (\boldsymbol{a}_{\mathrm{hom},L,\mu} - \boldsymbol{a}_{\mathrm{hom},\mu}) e \quad &= \quad \langle (D\phi_\mu \circ T_L + e) \cdot \boldsymbol{a}(0)(D\phi_\mu \circ T_L + e) - (D\phi_\mu + e) \cdot \boldsymbol{a}(0)(D\phi_\mu + e) \rangle \\
&\lesssim \quad \langle |D\phi_\mu \circ T_L - D\phi_\mu|^2 \rangle^{\frac{1}{2}} \left(\langle |D\phi_\mu|^2 \rangle^{\frac{1}{2}} + \langle |D\phi_\mu|^2 \rangle_L^{\frac{1}{2}} + 1 \right) \\
&\lesssim \quad \langle |D\phi_\mu \circ T_L - D\phi_\mu|^2 \rangle^{\frac{1}{2}},
\end{aligned}
$$

using the elementary a priori estimates $\langle |D\phi_\mu|^2 \rangle, \langle |D\phi_\mu|^2 \rangle_L \lesssim 1$. Combined with

$$
\langle |D\phi_\mu \circ T_L - D\phi_\mu|^2 \rangle^{\frac{1}{2}} \quad \leq \quad d \langle |\phi_\mu \circ T_L - \phi_\mu|^2 \rangle^{\frac{1}{2}} + \sum_{i=1}^{d} \langle |\phi_\mu \circ T_{e_i} \circ T_L \circ T_{-e_i} - \phi_\mu|^2 \rangle^{\frac{1}{2}},
$$

this yields (62).

Step 3. Representation formula for the difference of the correctors.

Let $x_0 \in \{0, e_1, \ldots, e_d\}$ be fixed. For $\boldsymbol{a} \in \Omega$ define $\tilde{\boldsymbol{a}} := (T_{x_0} \circ T_L \circ T_{-x_0})(\boldsymbol{a})$. We claim that

(63)
$$
\left| \phi_\mu(\tilde{\boldsymbol{a}}) - \phi_\mu(\boldsymbol{a}) \right| \leq \sum_{\substack{z \in \mathbb{Z}^d \\ |z| \geq \frac{L}{4}}} |\nabla_z G_\mu(\tilde{\boldsymbol{a}}, 0, z)| |\nabla \overline{\phi}_\mu(\boldsymbol{a}, z) + e|.
$$

where $G_\mu(\tilde{\boldsymbol{a}}, \cdot, \cdot)$ is the Green's function of the elliptic operator $\mu + \nabla^* \tilde{\boldsymbol{a}} \nabla$. Indeed, since $\overline{\phi}_\mu(\tilde{\boldsymbol{a}}, \cdot) - \overline{\phi}_\mu(\boldsymbol{a}, \cdot)$ is an $\ell^\infty(\mathbb{Z}^d)$-solution of

$$
(\mu + \nabla^* \tilde{\boldsymbol{a}} \nabla)\big(\overline{\phi}_\mu(\tilde{\boldsymbol{a}}, \cdot) - \overline{\phi}_\mu(\boldsymbol{a}, \cdot) \big) = \nabla^*(\boldsymbol{a} - \tilde{\boldsymbol{a}})(\nabla \overline{\phi}_\mu(\boldsymbol{a}, \cdot) + e),
$$

the Green representation formula reads

$$
\overline{\phi}_\mu(\tilde{\boldsymbol{a}}, 0) - \overline{\phi}_\mu(\boldsymbol{a}, 0) = \sum_{z \in \mathbb{Z}^d} \nabla_z G_\mu(\tilde{\boldsymbol{a}}, 0, z) \cdot f(z), \qquad \text{with } f(z) := (\boldsymbol{a}(z) - \tilde{\boldsymbol{a}}(z))(\nabla \overline{\phi}_\mu(\boldsymbol{a}, z) + e).
$$

By construction, we have $\tilde{\boldsymbol{a}}(z) = \boldsymbol{a}(z)$ for $z + x_0 \in \big([-\frac{L}{2}, \frac{L}{2}) \cap \mathbb{Z} \big)^d$. Hence,

$$
|f(z)| \leq \begin{cases} |\nabla \overline{\phi}_\mu(\boldsymbol{a}, \cdot) + e| & \text{for } |z| \geq \frac{L}{4}, \\ 0 & \text{else}, \end{cases}
$$

and the desired estimate (63) follows.

Step 4. Conclusion.

We note that $0 \leq G_\mu(\boldsymbol{a}, 0, z) \lesssim \mu^{-1} \exp\big(- c_0 \sqrt{\mu} |z| \big)$, for some $c_0 > 0$ only depending on λ and d, as can be seen by an elementary argument (see e. g. [21, Lemma 23]). Thanks to discreteness, we get

(64)
$$
|\nabla_z G_\mu(\boldsymbol{a}, 0, z)| \lesssim \mu^{-1} \exp\big(- c_0 \sqrt{\mu} |z| \big).
$$

The combination of (62), (63), and (64) then yields

$$
e \cdot (\boldsymbol{a}_{\mathrm{hom},L,\mu} - \boldsymbol{a}_{\mathrm{hom},\mu}) e \qquad \lesssim \qquad \left\langle \left(\sum_{|z| \geq \frac{L}{4}} \mu^{-1} \exp\big(- c_0 \sqrt{\mu} |z| \big) |\nabla \overline{\phi}_\mu(z) + e| \right)^2 \right\rangle^{\frac{1}{2}}
$$

$$
\overset{\triangle\text{-ineq. \& stationarity of } \nabla \overline{\phi}_\mu}{\leq} \quad \big(\langle |D\phi_\mu|^2 \rangle + 1 \big)^{\frac{1}{2}} \sum_{|z| \geq \frac{L}{4}} \mu^{-1} \exp\big(- c_0 \sqrt{\mu} |z| \big),
$$

and (56) follows from the energy estimate $\langle |D\phi_\mu|^2 \rangle \lesssim 1$ and the evaluation of the sum on the RHS. $\square$

Proof of Proposition 3. Fix a non-negative integer $\frac{d}{2} - 2 < k < \frac{d}{2}$, and let $\boldsymbol{a}^k_{\mathrm{hom},\mu}$ (resp. $\boldsymbol{a}^k_{\mathrm{hom},L,\mu}$) be the approximate homogenized coefficients of Definition 4 associated with $\langle \cdot \rangle$ (resp. $\langle \cdot \rangle_L$). By combining (55), Lemma 8, and Lemma 9 in the form of

$$|\boldsymbol{a}^k_{\mathrm{hom},L,\mu} - \boldsymbol{a}^k_{\mathrm{hom},\mu}| \lesssim \sup_{\mu \le \tilde{\mu} \le 2^k \mu} |\boldsymbol{a}_{\mathrm{hom},L,\tilde{\mu}} - \boldsymbol{a}_{\mathrm{hom},\tilde{\mu}}|,$$

we get

(65) $$|\boldsymbol{a}_{\mathrm{hom},L} - \boldsymbol{a}_{\mathrm{hom}}| \lesssim \mu^{\frac{d}{2}} + \sqrt{\mu}^{-\alpha} \exp\left(-c_0 \sqrt{\mu} L\right).$$

It remains to optimize (65) in μ. To this aim, note that for all $\alpha' > 0$, $c_0 > 0$ and $L \gg 1$, we have

$$\left(\frac{\ln L}{L}\right)^d \lesssim \min_{\mu > 0} \left(\sqrt{\mu}^d + \sqrt{\mu}^{-\alpha'} \exp(-c_0 \sqrt{\mu} L)\right) \lesssim \left(\frac{\ln L}{L}\right)^d,$$

where the constant only depends on α', c_0 and d. Combined with (65), Proposition 3 follows. □

6 Estimates on the gradient of the parabolic Green's function

6.1 Auxiliary lemmas and outline of the proof of Theorem 3

Unless stated otherwise, $\boldsymbol{a}$ denotes in this section an L-periodic coefficient field in Ω_L (cf. Section 2.1) and $G_L(t, \boldsymbol{a}, x, y)$ is the associated L-periodic Green's function. When no confusion occurs, we suppress the argument $\boldsymbol{a}$ in the notation. Note that $\bar{G}_L := L^{-d}$ is the spatial average on $([0, L) \cap \mathbb{Z})^d$ of $G_L(t, \cdot, y)$ for all $t > 0$ and $y \in \mathbb{Z}^d$. Throughout this section we shall write $\int dx$ for the sum over $x \in ([0, L) \cap \mathbb{Z})^d$. We denote by d_L the L-periodic weight $d_L(x) := \mathrm{dist}(x, L\mathbb{Z}^d) + 1$ and recall that

$$\omega_L(t, x) = \left(\frac{d_L^2(x)}{t + 1} + 1\right)^{\frac{1}{2}}.$$

We first recall standard pointwise bounds on the Green's function itself:

Lemma 10 (Estimate of G_L). *For any weight exponent $\beta < \infty$ we have*

(66) $$G_L(t, x, y) \lesssim (t + 1)^{-\frac{d}{2}} \omega_L^{-\beta}(t, x - y) \quad \text{for } t \lesssim L^2,$$

(67) $$|G_L(t, x, y) - \bar{G}_L| \lesssim L^{-d} \exp\left(-c_0 \frac{t}{L^2}\right) \quad \text{for } t \gtrsim L^2,$$

where the constant depends only on β, next to λ and d, and $c_0 > 0$ is a constant that only depends on λ and d.

Lemma 10 essentially reflects the exponentially decaying tail of the Green's function. It is well-known that in the continuum case the exponential decay can be upgraded to a Gaussian decay:

$$G^{\mathrm{cont}}(t, x, y) \lesssim t^{-\frac{d}{2}} \exp\left(-c_0 \frac{|x - y|^2}{t}\right),$$

as was first established for the continuum Green's function G^{cont} by Nash [35] and Aronson [1]; see also Fabes & Stroock [15] for a stream-lined approach. In the spatially discrete case the Gaussian behavior of the tail does not hold. A lower bound with the exact (exponential) tail behavior has been obtained by Delmotte [9]. For the upper bound and a partial Aronson lower bound see also [18, Propositions B.3 and B.4].

Lemma 10 treats the L-periodic Green's function, not the whole-space Green's function as addressed in [9]. Its proof is a slight refinement of the classical approach by Nash, smuggling in the weight function when needed. Since the argument is standard, we omit the proof and refer to [21, Section 7] for details.

In addition, we need for the proof of Theorem 3 the following Meyers estimate for discrete L-periodic parabolic equations:

Lemma 11 (discrete, L-periodic, parabolic Meyers' estimate). *There exists $\bar{q} > 1$ depending only on λ and d such that for all $u : \mathbb{R} \times \mathbb{Z}^d \to \mathbb{R}$, $f : \mathbb{R} \times \mathbb{Z}^d \to \mathbb{R}$ and $g : \mathbb{R} \times \mathbb{Z}^d \to \mathbb{R}^d$ compactly supported in t and L-periodic in x, with u smooth in time and related to f and g via the equation*

$$(68) \qquad \partial_t u + \nabla^* \boldsymbol{a} \nabla u = \nabla^* g + f$$

for some $\boldsymbol{a} \in \Omega$, we have for all $1 < q \leq \bar{q}$

$$(69) \qquad \left(\int_{-\infty}^{\infty} \int |\nabla u|^{2q} dx dt \right)^{\frac{1}{2q}} \lesssim \left(\int_{-\infty}^{\infty} \int |g|^{2q} dx dt \right)^{\frac{1}{2q}} + \left(\int_{-\infty}^{\infty} \left(\int |f|^{\frac{2qd}{2q+d}} dx \right)^{\frac{2q+d}{d}} dt \right)^{\frac{1}{2q}},$$

where the constant only depends on q, λ, and d.

This result is the discrete counterpart of the well-known parabolic version of the original Meyers estimates for elliptic equations [31]. Since the proof is standard we only sketch the argument and refer to [21, Section 7] for details. The argument relies on the Calderón-Zygmund estimate for discrete parabolic and elliptic equations (with periodic boundary conditions), and a perturbation argument where $\boldsymbol{a}$ is viewed as a perturbation of the constant matrix $\frac{1+\lambda}{2}\boldsymbol{id}$ – recall that we assume the uniform ellipticity in the form of $\lambda \leq \boldsymbol{a} \leq 1$. Indeed, a direct calculation shows that we may rewrite (68) in terms of the function $\tilde{u} := u - \bar{u}$ (where $\bar{u}$ denotes the mean of u) as

$$\partial_t \tilde{u} + \tfrac{1+\lambda}{2} \nabla^* \nabla \tilde{u} = \nabla^* \tilde{g}$$

with RHS

$$\tilde{g} := g + \nabla v - (\boldsymbol{a} - \tfrac{1+\lambda}{2}\boldsymbol{id}) \nabla \tilde{u}$$

and v given as the unique solution with $\sum_{([0,L)\cap\mathbb{Z})^d} v = 0$ of the L-periodic Poisson equation

$$\nabla^* \nabla v = f - \bar{f} \qquad \text{(with } \bar{f} \text{ the mean of } f\text{)}.$$

In the constant-coefficient case, i. e. when $\boldsymbol{a} = \frac{1+\lambda}{2}\boldsymbol{id}$ and therefore $\tilde{g} = g + \nabla v$, the estimate (69) directly follows from

$$(70) \qquad \int_{-\infty}^{\infty} \int |\nabla \tilde{u}(t,x)|^{2q} \, dx dt \lesssim \int_{-\infty}^{\infty} \int |\tilde{g}(t,x)|^{2q} \, dx dt,$$

$$(71) \qquad \int |\nabla^2 v(x)|^{2q} \, dx \lesssim \left(\int |f(x) - \bar{f}|^{\frac{2qd}{2q+d}} \, dx \right)^{\frac{2q+d}{d}}.$$

Estimate (70) is a discrete, parabolic Calderón-Zygmund estimate, while (71) is obtained by combining a discrete elliptic Calderón-Zygmund estimate with a Sobolev-Poincaré inequality. Both discrete Calderón-Zygmund estimates can for instance be obtained from their well-known continuum counterparts by a direct comparison of the associated discrete and continuum Fourier multipliers (see [21, Section 7.4] for details). The additional $\nabla \tilde{u}$-term that appears in the variable-coefficient case in the definition of $\tilde{g}$ can be absorbed into the LHS. Indeed, the multiplicative constant in (70) tends to 1 as $q \to 1$, whereas $|(\boldsymbol{a} - \frac{1+\lambda}{2}\boldsymbol{id})\nabla \tilde{u}| \leq \frac{1-\lambda}{2}|\nabla \tilde{u}|$ by uniform ellipticity. For the details we refer to [21, Step 3 and 4, Proof of Lemma 24].

The proof of Theorem 3 is organized as follows. We only address the L-periodic case, that is, statement (b). Based on Lemma 10, we shall derive the following pointwise-in-time estimate

$$(72) \qquad \int \left(\omega_L^\alpha(t, x - y) |\nabla G_L(t, x, y)| \right)^2 dx \lesssim (t+1)^{-\frac{d}{2}-1} \exp\left(- c_0 \frac{t}{L^2} \right),$$

see Steps 1 and 2 of the proof. The most delicate part is to increase the exponent of integrability from 2 to $2q > 2$ by appealing to Lemma 11. Statement (a) of Theorem 3 follows from statement (b) by soft arguments in the limit $L \uparrow \infty$. Indeed, by the Arzelà-Ascoli theorem, the periodic Green's function $G_L(\boldsymbol{a}_L)$ converges pointwise to the whole-space Green's function $\bar{G}(\boldsymbol{a})$ if $\boldsymbol{a}_L \in \Omega$ is the periodic extension on $\mathbb{Z}^d$ of the restriction $\boldsymbol{a}_{|([-L/2,L/2)\cap\mathbb{Z})^d}$.

6.2 Proof of Theorem 3 (b)

We shall need a Leibniz rule and a chain rule for discrete spatial derivatives:

$$(73) \qquad \forall u, v : \mathbb{Z}^d \to \mathbb{R}, \quad \begin{cases} \nabla_i(uv) &=& (\nabla_i u)v + u(\cdot + e_i)\nabla_i v, \\ \nabla_i^*(uv) &=& (\nabla_i^* u)v + u(\cdot - e_i)\nabla_i^* v, \end{cases}$$

and for all $\beta \geq 1$

$$(74) \qquad \forall a, b \geq 0 : |a^\beta - b^\beta| \leq C(a^{\beta-1} + b^{\beta-1})|a-b|,$$

where C only depends on β.

For further reference we note that

$$(75) \qquad |\nabla \omega_L(t,x)| \leq (t+1)^{-\frac{1}{2}},$$
$$(76) \qquad \omega_L(t,x) \lesssim 1 \quad \text{for } t \gtrsim L^2,$$

and recall that the weight d_L satisfies

$$(77a) \qquad d_L(x \pm e_i) \lesssim d_L(x),$$
$$(77b) \qquad |\nabla d_L(x)| \leq 1.$$

so that for $2\alpha \geq 1$, we have

$$(77c) \qquad |\nabla_i d_L^{2\alpha}(x)| \overset{(74),(77b),(77a)}{\lesssim} d_L^{2\alpha-1}(x).$$

Step 1. $L_t^2 \ell_x^2$ estimate of ∇G_L:

$$\frac{1}{T} \int_T^{2T} \int \left(\omega_L^\alpha(t, x-y)|\nabla_x G_L(t,x,y)|\right)^2 dx dt \lesssim (T+1)^{-\frac{d}{2}-1} \exp\left(-c_0 \frac{T}{L^2}\right).$$

In what follows, the spatial gradient of the Green's function is always taken w. r. t. the first variable, and we write $\nabla G_L(t,x,y)$ for $\nabla_x G_L(t,x,y)$. Since the estimates are uniform in $\boldsymbol{a} \in \Omega$, we may assume $y = 0$. In view of (76), this statement follows from the two inequalities:

$$(78) \qquad \frac{1}{T} \int_T^{2T} \int \omega_L^{2\alpha}(t,x)|\nabla G_L(t,x,0)|^2 dx dt \lesssim (T+1)^{-\frac{d}{2}-1} \quad \text{for } T \lesssim L^2,$$

$$(79) \qquad \frac{1}{T} \int_T^{2T} \int |\nabla G_L(t,x,0)|^2 dx dt \lesssim L^{-d} T^{-1} \exp\left(-c_0 \frac{T}{L^2}\right) \quad \text{for } T \gtrsim L^2,$$

up to changing c_0 when bounding $L^{-d}T^{-1} \exp\left(-c_0 \frac{T}{L^2}\right)$ by $T^{-\frac{d}{2}-1} \exp\left(-c_0 \frac{T}{L^2}\right)$ for $T \gtrsim L^2$. We first note that integrating the inequality

$$\frac{d}{dt} \int \frac{1}{2}(G_L(t,x,0) - \bar{G}_L)^2 dx = \int (G_L(t,x,0) - \bar{G}_L)\partial_t G_L(t,x,0) dx$$

$$= -\int \nabla G_L(t,x,0) \cdot \boldsymbol{a} \nabla G_L(t,x,0) dx$$

$$\leq -\lambda \int |\nabla G_L(t,x,0)|^2 dx$$

from T to $2T$ yields

$$(80) \qquad \int_T^{2T} \int |\nabla G_L(t,x,0)|^2 dx dt \lesssim \int (G_L(T,x,0) - \bar{G}_L)^2 dx.$$

Combined with (67), it yields (79). We now turn to (78). For $T \lesssim 1$ this reduces to

$$\frac{1}{T} \int_T^{2T} \int \omega_L^{2\alpha}(t,x) |\nabla G_L(t,x,0)|^2 dxdt \lesssim 1,$$

which follows from the discrete estimate $|\nabla_i G_L(t,x,0)| \leq |G_L(t,x,0)| + |G_L(t,x+e_i,0)|$ and (66). It remains to address the case $1 \lesssim T \lesssim L^2$. By definition of ω_L,

$$\frac{1}{T} \int_T^{2T} \int \omega_L^{2\alpha}(t,x) |\nabla G_L(t,x,0)|^2 dxdt \lesssim \frac{1}{T} \int_T^{2T} \int |\nabla G_L(t,x,0)|^2 dxdt$$
$$+ \frac{1}{(T+1)^{\alpha+1}} \int_T^{2T} \int d_L^{2\alpha}(x) |\nabla G_L(t,x,0)|^2 dxdt.$$

Hence, we need to show that for all $\alpha \geq 0$,

(81)
$$\int_T^{2T} \int d_L^{2\alpha}(x) |\nabla G_L(t,x,0)|^2 dxdt \lesssim T^{\alpha - \frac{d}{2}} \quad \text{for } T \lesssim L^2.$$

By Hölder's inequality, it is enough to show (81) for $\alpha = 0$ and $\alpha \geq 1$. For $\alpha = 0$, this is a consequence of (80) and (66), where the latter is used for a β with $\beta > \frac{d}{2}$. The starting point for $\alpha \geq 1$ is the identity

$$\frac{d}{dt} \int d_L^{2\alpha}(x) \frac{1}{2} G_L^2(t,x,0) dx \;=\; -\int \nabla(d_L^{2\alpha}(x) G_L(t,x,0)) \cdot \boldsymbol{a}(x) \nabla G_L(t,x,0) dx$$
$$\overset{(73)}{=}\; -\int d_L^{2\alpha}(x) \nabla G_L(t,x,0) \cdot \boldsymbol{a}(x) \nabla G_L(t,x,0) dx$$
$$-\int \sum_{i=1}^d G_L(t,\cdot+e_i,0) \nabla_i d_L^{2\alpha}(x) \, \boldsymbol{a}_{ii}(x) \nabla_i G_L(t,x,0) dx,$$

which, combined with (77c), yields

$$\frac{d}{dt} \int d_L^{2\alpha}(x) \frac{1}{2} G_L^2(t,x,0) dx \;\leq\; -\lambda \int d_L^{2\alpha}(x) |\nabla G_L(t,x,0)|^2 dx$$
$$+ C \int d_L^{2\alpha-1}(x) \, G_L(t,x,0) \, |\nabla G_L(t,x,0)| dx$$

for some constant C only depending on α and d. Young's inequality then yields

$$\int d_L^{2\alpha}(x) |\nabla G_L(t,x,0)|^2 dx \lesssim -\frac{d}{dt} \int d_L^{2\alpha}(x) \frac{1}{2} G_L^2(t,x,0) dx + \int d_L^{2\alpha-2}(x) G_L^2(t,x,0) dx,$$

which we integrate between T and $2T$:

$$\int_T^{2T} \int d_L^{2\alpha}(x) |\nabla G_L(t,x,0)|^2 dxdt \lesssim \int d_L^{2\alpha}(x) G_L^2(T,x,0) dx + \int_T^{2T} \int d_L^{2\alpha-2}(x) G_L^2(t,x,0) dxdt$$
$$\lesssim (T+1)^\alpha \int \omega_L^{2\alpha}(T,x) G_L^2(T,x,0) dx + \int_T^{2T} (t+1)^{\alpha-1} \int \omega_L^{2\alpha-2}(t,x) G_L^2(t,x,0) dxdt.$$

We combine this inequality with (66) for a β with $2\alpha - 2\beta < -d$, and the estimate

(82)
$$\int \omega_L^{-r}(t,x) dx \leq \int_{\mathbb{R}^d} \left(\frac{|x|^2}{t+1} + 1 \right)^{-\frac{r}{2}} dx \lesssim (t+1)^{\frac{d}{2}}$$

which holds for $T \lesssim L^2$ and any $r > d$, and which we apply with $r = 2(\beta - \alpha)$ and $r = 2(\beta - (\alpha-1))$. In conclusion we get (81) for $\alpha \geq 1$.

Step 2. $L_t^\infty \ell_x^2$ estimate of ∇G_L:

$$(83) \qquad \int \left(\omega_L^\alpha(t, x - y) |\nabla G_L(t, x, y)| \right)^2 dx \;\lesssim\; (t+1)^{-\frac{d}{2}-1} \exp\left(- c_0 \frac{t}{L^2} \right).$$

This upgrade of Step 1 follows from the semigroup property $G_L(T, x, y) = \int G_L(t, x, z) G_L(T - t, z, y) \, dz$, which we use in the form of

$$G_L(T, x, y) \;=\; \frac{3}{T} \int_{T/3}^{2T/3} \int G_L(t, x, z) G_L(T - t, z, y) dz dt.$$

We differentiate this identity w. r. t. x and use that $\frac{3}{T} \int_{T/3}^{2T/3} \int G_L(T - t, z, y) dz dt = 1$, so that by Jensen's inequality

$$(84) \qquad |\nabla G_L(T, x, y)|^2 \leq \frac{3}{T} \int_{T/3}^{2T/3} \int G_L(T - t, z, y) |\nabla G_L(t, x, z)|^2 dz dt.$$

Note that for all $\theta \in [0, 1]$, we have by the triangle inequality for all $t \geq 0$, $x, y, z \in \mathbb{Z}^d$,

$$(85) \quad \omega_L^2(t, x - y) = \frac{d_L^2(x - y)}{t + 1} + 1 \leq 2 \frac{d_L^2(x - z) + d_L^2(z - y)}{t + 1} + 1$$

$$\leq 2 \left(\frac{d_L^2(x - z)}{t + 1} + 1 \right) \left(\frac{d_L^2(z - y)}{t + 1} + 1 \right) \leq 2 \omega_L^2(\theta t, x - z) \omega_L^2((1 - \theta) t, z - y).$$

Thus, integrating (84) on $([0, L] \cap \mathbb{Z})^d$ and using (85) with $\theta = \frac{t}{T} \in [0, 1]$ yields

$$(86) \quad \int \omega_L^{2\alpha}(T, x - y) |\nabla G_L(T, x, y)|^2 dx \;\lesssim\; \frac{3}{T} \int_{T/3}^{2T/3} \int \omega_L^{2\alpha}(T - t, y - z) G_L(T - t, z, y)$$

$$\times \int \omega_L^{2\alpha}(t, x - z) |\nabla G_L(t, x, z)|^2 dx \, dz dt.$$

Next, we use Lemma 10 to estimate the term $\omega_L^{2\alpha}(t', z - y) G_L(t', z, y)$ for $t' = T - t$ in (86):

Case $t' \lesssim L^2$: Estimate (66) shows that $G_L(t', z, y) \lesssim (t' + 1)^{-\frac{d}{2}} \omega_L^{-\gamma}(t', z - y)$ for any weight exponent $\gamma < \infty$. Taking $\gamma = 2\alpha + r$ for some $r > d$ yields

$$\omega_L^{2\alpha}(t', z - y) G_L(t', z, y) \lesssim (t' + 1)^{-\frac{d}{2}} \omega_L^{-r}(t', z - y).$$

Case $t' \gtrsim L^2$: Estimate (67) shows that $G_L(t', x, y) \lesssim L^{-d}$. Combined with (76), this yields

$$\omega_L^{2\alpha}(t', z - y) G_L(t', z, y) \lesssim L^{-d}.$$

Since $t' = T - t \in [T/3, 2T/3]$, we get

$$(87) \qquad \omega_L^{2\alpha}(T - t, z - y) G_L(T - t, z, y) \lesssim \left\{ \begin{array}{ll} (T + 1)^{-\frac{d}{2}} \omega_L^{-r}(T, z - y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\}.$$

We note for further reference that the integral of the RHS is of order 1:

$$(88) \qquad \left\{ \begin{array}{lll} T \leq L^2 & : & \int (T + 1)^{-\frac{d}{2}} \omega_L^{-r}(T, z - y) dz \\ T \geq L^2 & : & \int L^{-d} dz \end{array} \right\} \lesssim 1,$$

which follows from (82).

Inserting (87) into (86) yields by Fubini's theorem

$$
\int \omega_L^{2\alpha}(T, x-y)|\nabla G_L(T,x,y)|^2 dx \;\lesssim\; \int \left\{ \begin{array}{ll} (T+1)^{-\frac{d}{2}}\omega_L^{-r}(T, z-y) & \text{for} \quad T \le L^2 \\ L^{-d} & \text{for} \quad T \ge L^2 \end{array} \right\}
$$

$$
\times \frac{1}{T}\int_{T/3}^{2T/3} \int \omega_L^{2\alpha}(t, x-z)|\nabla G_L(t,x,z)|^2 dx dt \, dz.
$$

By Step 1 (with T replaced by $T/3$ and y replaced by z), this estimate turns into

$$
\int \omega_L^{2\alpha}(T, x-y)|\nabla G_L(T,x,y)|^2 dx
$$

$$
\lesssim \int \left\{ \begin{array}{ll} (T+1)^{-\frac{d}{2}}\omega_L^{-r}(T, z-y) & \text{for} \quad T \le L^2 \\ L^{-d} & \text{for} \quad T \ge L^2 \end{array} \right\} dz \times (T+1)^{-\frac{d}{2}-1}\exp\left(-c_0\frac{T}{L^2}\right),
$$

which by (88) yields (83).

Step 3. Weighted $L_t^{2q}\ell_x^{2q}$-estimate of ∇G_L: For all $1 \le q < q_0 := \min\{\bar{q}, \frac{d}{d-2}\}$, where $\bar{q}$ is the Meyers exponent of Lemma 11, we have

$$
(89) \qquad \left(\frac{1}{T}\int_T^{2T}\int \left(\omega_L^\alpha(t, x-y)|\nabla G_L(t,x,y)|\right)^{2q} dx dt\right)^{\frac{1}{2q}} \lesssim (T+1)^{-\frac{d}{2}-\frac{1}{2}+\frac{d}{2}\frac{1}{2q}}\exp\left(-c_0\frac{T}{L^2}\right),
$$

where the constant only depends on λ, d, α, and q. In order to treat the cases of $T \le L^2$ and $T \ge L^2$ simultaneously, we introduce the notation

$$
G_L'(t,x,y) := \left\{ \begin{array}{ll} G_L(t,x,y) & \text{for} \quad T \le L^2 \\ G_L(t,x,y) - \bar{G}_L & \text{for} \quad T \ge L^2 \end{array} \right\}.
$$

W. l. o. g. we assume that $y = 0$. We first treat the case $T \lesssim 1$. Since $|\nabla_i G_L(t,x,0)| \le |G_L(t,x,0)| + |G_L(t, x+e_i, 0)|$, (77a) and the discrete ℓ_x^{2q}-ℓ_x^2-estimate yield

$$
\int \left(\omega_L^\alpha(t,x)|\nabla G_L(t,x,0)|\right)^{2q} dx \;\lesssim\; \int \left(\omega_L^\alpha(t,x)|G_L(t,x,0)|\right)^{2q} dx \le \left(\int \left(\omega_L^\alpha(t,x)|G_L(t,x,0)|\right)^2 dx\right)^q,
$$

so that (89) follows from (66) for $T \lesssim 1$.

We now assume that $T \gtrsim 1$, so that $T \sim (T+1)$. Let $\eta : \mathbb{R}^+ \to [0,1]$ be a smooth temporal cut-off function for the interval $[T, 2T]$ such that $\eta(t) = 0$ for $t \le \frac{T}{2}$ and $t \ge 4T$, $\eta(t) = 1$ for $T \le t \le 2T$, and $|\frac{d\eta}{dt}| \lesssim \frac{1}{T}$. We shall apply the Meyers estimate of Lemma 11 to $u(t,x) = \eta(t)\omega_L^\alpha(T,x)G_L'(t,x,0)$ (note that $\omega_L^\alpha(T,x)$ does not depend on t). By applying the (continuum) Leibniz rule to the time derivative, by adding and substracting the term $\eta(\nabla^*\boldsymbol{a}\nabla G_L')\omega_L^\alpha$, and by the defining equation of the Green's function

$$
\partial_t G_L(t,\boldsymbol{a},x,0) + \nabla_x^*\boldsymbol{a}(x)\nabla_x G_L(t,\boldsymbol{a},x,0) = 0,
$$

we have

$$
\begin{aligned}
\partial_t u + \nabla^*\boldsymbol{a}\nabla u &= (\partial_t G_L' + \nabla^*\boldsymbol{a}\nabla G_L')\eta\omega_L^\alpha + \eta\nabla^*\boldsymbol{a}\nabla(\omega_L^\alpha G_L') - \eta(\nabla^*\boldsymbol{a}\nabla G_L')\omega_L^\alpha + (\partial_t\eta)\omega_L^\alpha G_L' \\
(90) \qquad &= \eta\nabla^*\boldsymbol{a}\nabla(\omega_L^\alpha G_L') - \eta(\nabla^*\boldsymbol{a}\nabla G_L')\omega_L^\alpha + (\partial_t\eta)\omega_L^\alpha G_L'.
\end{aligned}
$$

For the first term on the RHS we use the discrete Leibniz rule (73) for all $i = 1, \ldots, d$:

$$
\begin{aligned}
\nabla_i^*\big(\boldsymbol{a}_{ii}\nabla_i(\omega_L^\alpha G_L')\big) &= \nabla_i^*\big(\boldsymbol{a}_{ii}(\nabla_i G_L')\omega_L^\alpha\big) + \nabla_i^*\big(\boldsymbol{a}_{ii}(\nabla_i\omega_L^\alpha)G_L'(\cdot + e_i)\big) \\
&= \big(\nabla_i^*(\boldsymbol{a}_{ii}\nabla_i G_L')\big)\omega_L^\alpha + (\nabla_i^*\omega_L^\alpha)\big(\boldsymbol{a}_{ii}\nabla_i G_L'\big)(\cdot - e_i) + \nabla_i^*\big(\boldsymbol{a}_{ii}(\nabla_i\omega_L^\alpha)G_L'(\cdot + e_i)\big).
\end{aligned}
$$

Hence (90) turns into

$$
(91) \qquad\qquad \partial_t u + \nabla^*\boldsymbol{a}\nabla u = \nabla^*(\boldsymbol{a}g_0) + f_1 + f_2
$$

with

$$g_0(t,x) \;\; := \;\; \sum_{i=1}^{d} \left(\nabla_i \omega_L^\alpha(T,x)\right) \eta(t) G_L'(t, \cdot + e_i, 0) e_i,$$

$$f_1(t,x) \;\; := \;\; \tfrac{d\eta}{dt}(t)\, \omega_L^\alpha(T,x) G_L'(t,x,0),$$

$$f_2(t,x) \;\; := \;\; \eta(t) \sum_{i=1}^{d} \left(\nabla_i^* \omega_L^\alpha(T,x)\right) \boldsymbol{a}_{ii}(x - e_i) \nabla_i G_L'(t, x - e_i, 0).$$

Since we have written (91) in the form of (68) we may apply Lemma 11. We thus need to estimate $\boldsymbol{a}g_0$, f_1 and f_2. It suffices to establish the pointwise-in-time estimates

(92a)
$$\left(\int |\boldsymbol{a}(x) g_0(t,x)|^{2q} dx\right)^{\frac{1}{2q}} \;\; \lesssim \;\; \eta(t) t^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

(92b)
$$\left(\int |f_1(t,x)|^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{2qd}} \;\; \lesssim \;\; |\tfrac{d\eta}{dt}(t)| t^{-\frac{d}{2} + \frac{1}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

(92c)
$$\left(\int |f_2(t,x)|^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{2qd}} \;\; \lesssim \;\; t^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right).$$

Indeed, (92a)—(92c) yield

$$\left(\int_{-\infty}^{\infty} \int |\boldsymbol{a}(x) g_0(t,x)|^{2q} dx\, dt\right)^{\frac{1}{2q}} + \left(\int_{-\infty}^{\infty} \left(\int |f_1(t,x)|^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{d}} dt\right)^{\frac{1}{2q}}$$

$$+ \left(\int_{-\infty}^{\infty} \left(\int |f_2(t,x)|^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{d}} dt\right)^{\frac{1}{2q}} \;\; \lesssim \;\; T^{\frac{1}{2q}} T^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{T}{L^2}\right),$$

so that the desired estimate (89) follows from Lemma 11 and the identity $\nabla u = \eta \omega_L^\alpha \nabla G_L + g_0$.
It remains to prove (92a)—(92c). From (74), (77a), (75), and the definitions of g_0, f_1 and f_2, we learn that

(93)
$$\begin{cases}
\int |\boldsymbol{a}(x) g_0(t,x)|^{2q} dx \;\; \lesssim \;\; \eta^{2q}(t) (t+1)^{-\frac{1}{2} 2q} \int |\omega_L^{\alpha-1}(T,x) G_L'(t,x,0)|^{2q} dx, \\[2mm]
\int |f_1(t,x)|^{\frac{2qd}{2q+d}} dx \;\; \lesssim \;\; |\tfrac{d\eta}{dt}(t)|^{\frac{2qd}{2q+d}} \int |\omega_L^\alpha(T,x) G_L'(t,x,0)|^{\frac{2qd}{2q+d}} dx, \\[2mm]
\int |f_2(t,x)|^{\frac{2qd}{2q+d}} dx \;\; \lesssim \;\; \eta^{\frac{2qd}{2q+d}}(t) (T+1)^{-\frac{1}{2}\frac{2qd}{2q+d}} \int \left(\omega_L^{\alpha-1}(T,x) |\nabla G_L(t,x,0)|\right)^{\frac{2qd}{2q+d}} dx.
\end{cases}$$

By definition of η, we only need to consider $t \geq \frac{T}{2} \gtrsim 1$. Since $\omega_L(t,x) \sim \omega_L(T,x)$ (uniformly in x) for all $\frac{T}{2} \leq t \leq 4T$, estimates (92a)—(92c) follow from (93) combined with

(94a)
$$\left(\int |\omega_L^{\alpha-1}(t,x) G_L'(t,x,0)|^{2q} dx\right)^{\frac{1}{2q}} \;\; \lesssim \;\; t^{-\frac{d}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

(94b)
$$\left(\int |\omega_L^\alpha(t,x) G_L'(t,x,0)|^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{2qd}} \;\; \lesssim \;\; t^{-\frac{d}{2} + \frac{1}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

(94c)
$$\left(\int \left(\omega_L^{\alpha-1}(t,x) |\nabla G_L(t,x,0)|\right)^{\frac{2qd}{2q+d}} dx\right)^{\frac{2q+d}{2qd}} \;\; \lesssim \;\; t^{-\frac{d}{2} + \frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

which we prove next.
Note that (94a) and (94b) can be combined into the single statement: For all $1 \leq q' < \infty$ and $0 \leq \alpha < \infty$,

(95)
$$\left(\int |\omega_L^\alpha(t,x) G_L'(t,x,0)|^{q'} dx\right)^{\frac{1}{q'}} \;\; \lesssim \;\; t^{-\frac{d}{2} + \frac{d}{2}\frac{1}{q'}} \exp\left(-c_0 \frac{t}{L^2}\right).$$

To prove (95) we distinguish the cases $T \leq L^2$ and $T \geq L^2$. In the latter case we have $t \geq \frac{T}{2} \geq \frac{1}{2}L^2$, so that (67) in Lemma 10 combined with (76) yields

$$\left(\int |\omega_L^\alpha(t,x)(G_L(t,x,0) - \bar{G}_L)|^{q'} dx \right)^{\frac{1}{q'}} \lesssim L^{-d+d\frac{1}{q'}} \exp\left(-c_0 \frac{t}{L^2} \right) \lesssim t^{-\frac{d}{2}+\frac{d}{2}\frac{1}{q'}} \exp\left(-c_0 \frac{t}{L^2} \right),$$

up to redefining c_0. This proves (95) for $T \geq L^2$.

For $T \leq L^2$, we have $t \leq 4L^2$, so that (66) in Lemma 10 (with $\beta = \alpha + r/q'$ for some $r > d$) yields

$$\left(\int \left(\omega_L^\alpha(t,x) G_L(t,x,0) \right)^{q'} dx \right)^{\frac{1}{q'}} \lesssim t^{-\frac{d}{2}} \left(\int \omega_L^{-r}(t,x) dx \right)^{\frac{1}{q'}} \overset{r>d,(82)}{\lesssim} t^{-\frac{d}{2}} t^{\frac{d}{2}\frac{1}{q'}}.$$

This establishes (95) for $t \leq L^2$.

We finally turn to (94c). We note that for $q < \frac{d}{d-2}$ in dimensions $d > 2$ (and all $q < \infty$ in dimension $d = 2$) we have $\frac{qd}{2q+d} < 1$. For all $1 \leq q < q_0 := \min\{\bar{q}, \frac{d}{d-2}\}$, we then have by Hölder's inequality with exponents $(\frac{2q+d}{2q+d-qd}, \frac{2q+d}{qd})$,

$$\left(\int \left(\omega_L^{\alpha-1}(t,x) |\nabla G_L(t,x,0)| \right)^{\frac{2qd}{2q+d}} dx \right)^{\frac{2q+d}{2qd}}$$

$$= \left(\int \omega_L^{-\frac{2qd}{2q+d}}(t,x) \left(\omega_L^\alpha(t,x) |\nabla G_L(t,x,0)| \right)^{\frac{2qd}{2q+d}} dx \right)^{\frac{2q+d}{2qd}}$$

$$\overset{\frac{2qd}{2q+d}<2, \text{Hölder}}{\leq} \left(\int \omega_L^{-\frac{2qd}{2q+d-qd}}(t,x) dx \right)^{\frac{2q+d-qd}{2qd}} \left(\int \left(\omega_L^\alpha(t,x) |\nabla G_L(t,x,0)| \right)^2 dx \right)^{\frac{1}{2}}$$

$$\overset{\frac{2qd}{2q+d(1-q)}>d,(82)}{\lesssim} t^{\frac{d}{2}\frac{2q+d-qd}{2qd}} \left(\int \left(\omega_L^\alpha(t,x) |\nabla G_L(t,x,0)| \right)^2 dx \right)^{\frac{1}{2}}$$

$$\overset{(83)}{\lesssim} t^{\frac{d}{2}\frac{2q+d-qd}{2qd}} \left(t^{-\frac{d}{2}-1} \exp\left(-c_0 \frac{t}{L^2} \right) \right)^{\frac{1}{2}}$$

$$\lesssim t^{-\frac{d}{2}+\frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{t}{L^2} \right).$$

This establishes (94c).

Step 4. $L_t^\infty \ell_x^{2q}$-estimate for ∇G_L: For all $1 \leq q < q_0 = \min\{\bar{q}, \frac{d}{d-2}\}$, and all $0 \leq \alpha < \infty$, we have

$$\left(\int \left(\omega_L^\alpha(T,x-y) |\nabla G_L(T,x,y)| \right)^{2q} dx \right)^{\frac{1}{2q}} \lesssim (T+1)^{-\frac{d}{2}-\frac{1}{2}+\frac{d}{2}\frac{1}{2q}} \exp\left(-c_0 \frac{T}{L^2} \right).$$

We essentially repeat the argument of Step 2. The only differences are:

- In (84), we use Jensen's inequality applied to $\mathbb{R}^d \ni g \mapsto |g|^{2q}$:

$$|\nabla G_L(T,x,y)|^{2q} \leq \frac{3}{T} \int_{T/3}^{2T/3} \int G_L(T-t,z,y) |\nabla G_L(t,x,z)|^{2q} dz dt.$$

- In (86), we multiply by $\omega^{2q\alpha}$:

$$(96) \quad \int \omega_L^{2q\alpha}(T,x-y) |\nabla G_L(T,x,y)|^{2q} dx \leq \frac{3}{T} \int_{T/3}^{2T/3} \int \omega_L^{2q\alpha}(T-t,y-z) G_L(T-t,z,y)$$

$$\times \int \omega_L^{2q\alpha}(t,x-z) |\nabla G_L(t,x,z)|^{2q} dx dz dt.$$

- In (87), we replace 2α by $2q\alpha$:

$$(97) \qquad \omega_L^{2q\alpha}(T-t, z-y)G_L(T-t, z, y) \lesssim \left\{ \begin{array}{ll} T^{-\frac{d}{2}}\omega_L^{-r}(T, z-y) & \text{for} \quad T \le L^2 \\ L^{-d} & \text{for} \quad T \ge L^2 \end{array} \right\}.$$

Inserting (97) into (96) then yields

$$
\int \omega_L^{2q\alpha}(T, x-y)|\nabla G_L(T, x, y)|^{2q}dx \quad \lesssim \quad \int \left\{ \begin{array}{ll} T^{-\frac{d}{2}}\omega_L^{-r}(T, z-y) & \text{for} \quad T \le L^2 \\ L^{-d} & \text{for} \quad T \ge L^2 \end{array} \right\}
$$
$$
\times \frac{1}{T}\int_{T/3}^{2T/3}\int \omega_L^{2q\alpha}(t, x-z)|\nabla G_L(t, x, z)|^{2q}dxdt\, dz,
$$

and the conclusion follows from Step 3 and estimate (88).

Acknowledgement

References

[1] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73(6):890–896, 1967.

[2] A. Bourgeat and A. Piatnitski. Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator. Asympt. Anal., 21(3-4):303–315, 1999.

[3] A. Bensoussan, J.-L. Lions and G. Papanicolaou. Asymptotic analysis for periodic structures. *Studies in Mathematics and its Applications, Vol. 5, North-Holland Publishing Co.*, 1978.

[4] M. Biskup. Recent progress on the Random Conductance Model. *Probability Surveys*, 8:294–373, 2011.

[5] E. Bolthausen and A-S. Sznitman. Ten lectures on random media. DMV Seminar, Vol. 32. Basel: Birkhäuser Verlag, 2002.

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

[7] J. G. Conlon and A. Naddaf. Greens Functions for Elliptic and Parabolic Equations with Random Coefficients. *New York J. Math.*, 6:153-225, 2000.

[8] J. G. Conlon and T. Spencer. Strong convergence to the homogenized limit of elliptic equations with random coefficients. *Trans. Amer. Math. Soc.*, in press.

[9] T. Delmotte. Estimations pour les chaînes de Markov réversibles. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1053–1058, 1997.

[10] T. Delmotte and J.-D. Deuschel. On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab. Theory Relat. Fields*, 133:358–390, 2005.

[11] R. L. Dobrushin and S. B. Shlosman. Completely analytical Gibbs fields. In Statistical physics and dynamical systems (Köszeg, 1984), volume 10 of Progr. Phys., pages 371–403. Birkhäuser Boston, Boston, MA, 1985.

[12] R. L. Dobrushin and S. B. Shlosman. Constructive criterion for the uniqueness of Gibbs field. In Statistical physics and dynamical systems (Köszeg, 1984), volume 10 of Progr. Phys., pages 347–370. Birkhäuser Boston, Boston, MA, 1985.

[13] W. E, P.B. Ming, and P.W. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18:121–156, 2005.

[14] A.-C. Egloffe, A. Gloria, J.-C. Mourrat, and T. N. Nguyen. Random walk in random environment, corrector equation, and homogenized coefficients: from theory to numerics, back and forth. Preprint, http://hal.archives-ouvertes.fr/hal-00749667.

[15] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic harnack inequality using the old ideas of Nash. *Archive for Rational Mechanics and Analysis*, 96(4):327–338, 1986.

[16] T. Funaki. Stochastic interface models. *Lectures on Probability Theory and Statistics*, Lecture Notes in Mathematics Volume 1869:103–274, 2005.

[17] T. Funaki and H. Spohn. Motion by mean curvature from the GinzburgLandau $\nabla\varphi$ interface models. *Commun. Math. Phys.*, 185:1–36, 1997.

[18] G. Giacomin, S. Olla and H. Spohn. Equilibrium Fluctuations for $\nabla\varphi$ Interface Model. *Ann. Probab.*, 29(3):1138–1172, 2001.

[19] A. Gloria. Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *M2AN Math. Model. Numer. Anal.*, 46(1):1–38, 2012.

[20] A. Gloria and J.-C. Mourrat. Spectral measure and approximation of homogenized coefficients. *Probab. Theory Relat. Fields*, 154(1):287–326, 2012.

[21] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics – long version. MPI Preprint 3/2013.

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

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

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

[25] A. Gloria and F. Otto. Quantitative estimates on the corrector equation in stochastic homogenization. In preparation.

[26] 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.

[27] C. Kipnis and S.R.S. Varadhan. Central limit theorem for additive functional of reversible Markov processes and applications to simple exclusion. *Commun. Math. Phys.*, 104:1–19, 1986.

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

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

[30] 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.

[31] N. Meyers. An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 17(3):189–206, 1963.

[32] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(11): 294–327, 2011.

[33] A. Naddaf and T. Spencer. On homogenization and scaling limits of some gradient perturbations of a massless free field. *Comm. Math Phys.*, 183:55–84, 1997.

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

[35] J. Nash. Continuity of Solutions of Parabolic and Elliptic Equations *American Journal of Mathematics*, 80(4):931–954, 1958.

[36] H. Owhadi. Approximation of the effective conductivity of ergodic media by periodization. *Probab. Theory Relat. Fields*, 125:225–258, 2003.

[37] 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.

[38] D. W. Stroock and B. Zegarliński. The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition. Comm. Math. Phys., 144(2):303–323, 1992.

[39] D. W. Stroock and B. Zegarliński. The logarithmic Sobolev inequality for continuous spin systems on a lattice. J. Funct. Anal., 104(2):299–326, 1992.

[40] D. W. Stroock and B. Zegarliński. The logarithmic Sobolev inequality for discrete spin systems on a lattice. Comm. Math. Phys., 149(1):175–193, 1992.

[41] V.V. Yurinskii. *Vilnius Conference Abstracts*, Sept. 1978, 54.

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