Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Annealed estimates on the Green's function

(revised version: April 2013)

by

Daniel Marahrens and Felix Otto

Preprint no.: 69 2012

Annealed estimates on the Green's function

Daniel Marahrens* Felix Otto*

Abstract

We consider a random, uniformly elliptic coefficient field $a(x)$ on the d-dimensional cubic lattice $\mathbb{Z}^d$. We are interested in the spatial decay of the quenched elliptic Green's function $G(a;x,y)$. Next to stationarity, we assume that the spatial correlation of the coefficient field decays sufficiently rapidly to the effect that a Logarithmic Sobolev Inequality holds for the ensemble $\langle\cdot\rangle$. We prove that *all* stochastic moments of the first and second mixed derivatives of the Green's function, that is, $\langle|\nabla_x G(x,y)|^p\rangle$ and $\langle|\nabla_x\nabla_y G(x,y)|^p\rangle$, have the same decay rates in $|x-y|\gg 1$ as for the constant coefficient Green's function, respectively. This result relies on and substantially extends the one by Delmotte and Deuschel [8], which optimally controls second moments for the first derivatives and first moments of the second mixed derivatives of G, that is, $\langle|\nabla_x G(x,y)|^2\rangle$ and $\langle|\nabla_x\nabla_y G(x,y)|\rangle$. As an application, we derive optimal estimates on the random part of the homogenization error.

The outline of this work is as follows: After introducing the discrete setting in Section 1, we present the statistical assumptions and the main result on the annealed moments of the Green's function in Section 2. We present optimal estimates on the random part of the homogenization error in Section 3. Section 4 contains an annealed Hölder-estimate in the spirit of De Giorgi, which we will obtain as a a consequence of our main result. In Section 5 we explain our main assumption, a weakened Logarithmic Sobolev Inequality, which in particular holds for all independent, identically distributed coefficient fields. Section 6 contains the main ingredients of the proof of the annealed Green's function estimates — in particular we recall the result by Delmotte and Deuschel [8]. All proofs are postponed until Section 7.

*Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany, Daniel.Marahrens@mis.mpg.de resp. Felix.Otto@mis.mpg.de
April 16, 2013
MSC2010 subject classifications. Primary: 35B27, Secondary: 35J08, 39A70, 60H25
Key words Stochastic homogenization, elliptic equations, Green's function, annealed estimates

1 Discrete uniformly elliptic equations

In this paper we consider linear second order difference equations with uniformly elliptic, bounded random coefficients of the form

$$\nabla^*(a\nabla u)(x) = f(x) \quad \text{for all } x \in \mathbb{Z}^d.$$

If there is no danger of confusion, we also write $\nabla^* a \nabla u$ for $\nabla^*(a\nabla u)$. In this equation we define the *spatial 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 $i = 1, \ldots, d$ by the expressions

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

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

Here $e_1, \ldots, e_d$ is the canonical basis of $\mathbb{R}^d$. The expressions $\nabla \zeta$ and $-\nabla^* \xi$ are the discrete gradient and divergence for scalar fields and vector fields, respectively, on $\mathbb{Z}^d$. Note that ∇^* is the $\ell^2(\mathbb{Z}^d)$-adjoint of ∇.

Here comes the main assumption on the considered coefficient fields $\{a(x)\}_{x \in \mathbb{Z}^d}$: For every $x \in \mathbb{Z}^d$, $a(x) \in \mathbb{R}^{d \times d}$ is symmetric and *uniformly elliptic* in the sense of $\lambda|\xi|^2 \le \xi \cdot a(x)\xi \le |\xi|^2$ for all $\xi \in \mathbb{R}^d$. Here and below $\lambda \in (0, 1)$ denotes the ellipticity ratio which is fixed throughout the paper. In addition we assume that the coefficients are *diagonal*. We note that the condition $a = \operatorname{diag}(a_{11}, \ldots, a_{dd})$ is diagonal yields that the coefficients can be associated with the bonds of the lattice (rather than the vertices): In fact, this can be seen on the level of the Dirichlet integral, which can be rewritten as

$$\sum_{x \in \mathbb{Z}^d} \zeta(x) \nabla^*(a\nabla\zeta)(x) = \sum_{x \in \mathbb{Z}^d} \sum_{i=1}^{d} a_{ii}(x)(\zeta(x + e_i) - \zeta(x))^2.$$

This allows, for instance, to interpret $\nabla^* a \nabla u$ as the generator of a random walk on $\mathbb{Z}^d$ with jump rates across bonds described by a. Diagonality is known to be a sufficient (but not necessary) condition for the maximum principle to hold for $\nabla^* a \nabla u$. The maximum principle is a crucial ingredient for the estimate (19) on the quenched Green's function. Also, in the proof of Lemma 5, we shall adapt the bond-based point of view.

Our main object is the non-constant coefficient, elliptic, discrete Green's function $G(a; x, x')$ defined through $\nabla_x^* a(x) \nabla_x G(a; x, x') = \delta(x - x')$, where δ stands for the discrete version of the Dirac distribution, i.e.

$$\delta(x) = \begin{Bmatrix} 1 & \text{for } x = 0 \\ 0 & \text{otherwise} \end{Bmatrix}.$$

We usually drop the argument a and just write $G(x,y)$. Often, it is more convenient to appeal to the distributional characterization:

$$\forall\, \zeta(x): \quad \sum_x \nabla\zeta(x)\cdot a(x)\nabla_x G(x,x') = \zeta(x').$$

Throughout the paper, we will work in dimension $d \geq 2$. Dimension $d = 2$ needs a bit more care in terms of the definition of the Green's function. Since we are only interested in *gradient* estimates, this is merely technical. Sometimes, it is more convenient to think of the approximation via a massive term in the sense of

$$T^{-1}G_T(x,x') + \nabla_x^* a(x)\nabla_x G_T(x,x') = \delta(x-x');$$

this is the case in the proof of Proposition 1. At other times, it is more convenient to think in terms of an approximation via periodization in the sense of

$$\nabla_x^* a(x)\nabla_x G_L(x,x') = \sum_{z\in\mathbb{Z}^d} \delta(x-x'-Lz) - L^{-d}; \tag{1}$$

this is the case in the proof of Lemma 6.

2 Assumptions on the ensemble and main result

We are given a probability measure on the space of uniformly elliptic, diagonal coefficient fields (endowed with the product topology), cf. the previous section. Following the convention in statistical mechanics, we call this probability measure an *ensemble* and denote the associated ensemble average (i.e. the expected value) by $\langle\cdot\rangle$. We assume that $\langle\cdot\rangle$ is *stationary* in the sense that for any shift vector $z \in \mathbb{Z}^d$, the shifted coefficient field $a(\cdot+z) := \{\mathbb{Z}^d \ni x \mapsto a(x+z)\}$ has the same distribution as a. We also note that the Green's function is shift-invariant or stationary in the sense that $G(a(\cdot+z);x,y) = G(a;x+z,y+z)$.

Besides stationarity, the main assumption on the ensemble of coefficients and only probabilistic tool will be a weak variant of the Logarithmic Sobolev Inequality (LSI). In Section 5, we will comment on the LSI and the related Spectral Gap Inequality — there we will also describe the relation between this weak LSI and the usual LSI.

Definition 1. [Weak logarithmic Sobolev inequality]. *Let $\langle\cdot\rangle$ be a stationary ensemble of coefficients a.*

i) For a random variable ζ, i.e. a function $\zeta(a)$, a site $y \in \mathbb{Z}^d$, and a direction $i = 1, \cdots, d$, we define the continuum vertical derivative $\frac{\partial \zeta}{\partial a_{ii}(y)}$ as the partial derivative with respect to the variable $a_{ii}(y)$. Furthermore, we denote by $\frac{\partial \zeta}{\partial a(y)}$ the d-dimensional vector with coefficients $\frac{\partial \zeta}{\partial a_{ii}(y)}$ with $i = 1, \ldots, d$. Thus in particular, $(\frac{\partial \zeta}{\partial a(y)})^2 = \sum_{i=1}^d (\frac{\partial \zeta}{\partial a_{ii}(y)})^2$.
ii) We say $\langle\cdot\rangle$ satisfies a weak Logarithmic Sobolev Inequality (LSI) with constant $\rho > 0$ if for all positive random variables $f(a) > 0$ we have

$$\Big\langle f \log \frac{f}{\langle f\rangle} \Big\rangle \leq \frac{1}{2\rho} \Big\langle \sum_{y\in\mathbb{Z}^d} \sup_{a(y)\in[\lambda,1]^d} \frac{1}{f} \Big(\frac{\partial f}{\partial a(y)}\Big)^2 \Big\rangle. \tag{2}$$

Note that the difference between the weakened LSI (2) and the usual LSI, see (15), lies solely in the supremum over $a(y)$. The usual LSI therefore implies (2). The merit of this weakening is that it is satisfied by *any* ensemble of independent, identically distributed coefficients $\{a(y)\}_{y\in\mathbb{Z}^d}$, cf. Lemma 1 below. Our main result is:

Theorem 1. *Let $\langle\cdot\rangle$ be stationary and satisfy the weak version of LSI with constant $\rho > 0$, see Definition 1. Then for all $1 \leq p < \infty$ and $x \in \mathbb{Z}^d$, it holds*

$$\begin{aligned} \langle|\nabla\nabla G(x,0)|^{2p}\rangle^{\frac{1}{2p}} &\leq C(d,\lambda,\rho,p)(|x|+1)^{-d}, &(3)\\ \langle|\nabla_x G(x,0)|^{2p}\rangle^{\frac{1}{2p}} &\leq C(d,\lambda,\rho,p)(|x|+1)^{1-d}. &(4) \end{aligned}$$

Here $\nabla\nabla G(x,y) = \nabla_x\nabla_y G(x,y)$ denote the mixed second derivatives of the Green's function. Also, here and in the sequel, $C(d,\lambda,\rho,p)$ stands for a generic constant that only depends on dimension $d \geq 2$, on the ellipticity ratio $\lambda > 0$, on the LSI constant $\rho > 0$, and on the exponent of integrability $p < \infty$.

Clearly, the spatial decay rates in Theorem 1 are optimal, since those are the decay rates of the constant coefficient Green's function. Note that stationarity of $\langle\cdot\rangle$ and G implies $\langle|\nabla\nabla G(a;x,y)|^{2p}\rangle = \langle|\nabla\nabla G(a(\cdot - y);x,y)|^{2p}\rangle = \langle|\nabla\nabla G(a;x-y,0)|^{2p}\rangle$, so Theorem 1 can be rephrased as

$$\begin{aligned} \langle|\nabla_x\nabla_y G(x,y)|^{2p}\rangle^{\frac{1}{2p}} &\leq C(d,\lambda,\rho,p)(|x-y|+1)^{-d},\\ \langle|\nabla_x G(x,y)|^{2p}\rangle^{\frac{1}{2p}} &\leq C(d,\lambda,\rho,p)(|x-y|+1)^{1-d} \end{aligned}$$

for all $x, y \in \mathbb{Z}^d$. An interesting aspect of Theorem 1 is the following: The *quenched* versions of (3) and (4) are false, i.e. the *uniform* in a and *pointwise* in x estimates $|\nabla\nabla G(x,0)| \leq C(d,\lambda)(|x|+1)^{-d}$ and $|\nabla_x G(x,0)| \leq$

$C(d,\lambda)(|x|+1)^{d-1}$ do *not* hold (while suitably spatially averaged versions of both estimates do hold uniformly in a); see our discussion in Section 4 below.

An easy consequence is the following generalized variance estimate on G itself:

Corollary 1. *Let $\langle\cdot\rangle$ be as in Theorem 1. Then it holds that*

$$\left\langle \left| G(x,0) - \langle G(x,0)\rangle \right|^{2p} \right\rangle^{\frac{1}{p}} \le C(d,\lambda,\rho,p) \begin{Bmatrix} (|x|+1)^{2(1-d)} & d>2 \\ (|x|+1)^{-2}\log(|x|+2) & d=2 \end{Bmatrix} \tag{5}$$

for all $x \in \mathbb{Z}^d$ and $1 \le p < \infty$.

Remark 1. *We note that the estimate in Corollary 1 is optimal in the scaling of the spatial decay. This can be seen by developing to leading order in a small ellipticity ratio $1-\lambda \ll 1$. We expand upon this argument (for the special case of $p=1$) after the proof of Corollary 1.*

3 Homogenization error

In the same vein as Corollary 1, Theorem 1 allows to give optimal estimates on the random part of the homogenization error. These extend the results by Conlon & Naddaf [5, Theorem 1.2, Theorem 1.3] from small ellipticity ratio (i.e. $1-\lambda \ll 1$) to arbitrary ellipticity ratio. For the "strong error" (see below for an explanation of this wording) [5, Theorem 1.2] in $d>3$, this was already achieved by Gloria [11, Theorem 2]. For all other cases, our result appears to be new. Let us be more precise: For a coefficient field $a(x)$ and a right hand side $f(x)$ we consider the solution $u(x)$ of

$$\nabla^* a(x)\nabla u(x) = f(x) \quad \text{on } \mathbb{Z}^d. \tag{6}$$

In order for (6) to have a unique solution that decays (i. e. $\lim_{|x|\uparrow\infty} u(x) = 0$), we assume for simplicity that f is compactly supported (and is of zero spatial average in the case of $d=2$ to ensure decay of u). By the random part of the homogenization error, we understand the "fluctuations" $u(x) - \langle u(x)\rangle$. These are expected to be small (w. r. t. the size of $u(x)$ itself) if $f(x)$ varies only slowly w. r. t. to the lattice spacing. In our notation, the lattice spacing is unity, so that a natural model for a right hand side that has a large characteristic scale $L \gg 1$ is given by $f(x) = L^{-2}\hat{f}(\frac{x}{L})$ for some bounded and compactly supported "mask" $\hat{f}(\hat{x})$, $\hat{x} \in \mathbb{R}^d$. The scaling L^{-2} of the amplitude of f is motivated as follows: In the rescaled variables $\hat{x}$, (6) now assumes the suggestive form of

$$\nabla_\epsilon^* a(\tfrac{\hat{x}}{\epsilon})\nabla_\epsilon u(\hat{x}) = \hat{f}(\hat{x}) \quad \text{on } \epsilon\mathbb{Z}^d, \tag{7}$$

where $\epsilon := L^{-1}$ is the ratio of the lattice spacing to the characteristic scale of the r. h. s. and where ∇_ϵ denote finite differences for the rescaled lattice $\epsilon\mathbb{Z}^d$ (i. e. $\nabla_{\epsilon,i} u(\hat{x}) = \epsilon^{-1}(u(\hat{x} + \epsilon e_i) - u(\hat{x}))$).

The size of the fluctuations will be measured in two different ways.

- Corollary 2: Here, the fluctuations will be controlled in a *strong* way in the sense that we estimate the (discrete, spatial) $\ell^p(\mathbb{Z}^d)$-norm $(\sum_x |u - \langle u\rangle|^p)^{1/p}$ of the fluctuations. This will be done for arbitrary stochastic moments (the role played by rp). Corollary 2 is the generalization of [5, Theorem 1.2] as well as [11, Theorem 2]. For our model right hand side, $f(x) = \epsilon^2 \hat{f}(\epsilon x)$ with bounded and compactly supported $\hat{f}$, the fluctuations are (up to a logarithmic correction for $d = 2$) of the order of ϵ in this measure, see (11).

- Corollary 3: Here, the fluctuations will be controlled in a *weak* way in the sense that we only estimate *spatial averages* $\sum_x (u - \langle u\rangle)g$ of the fluctuations, with *deterministic* averaging function $g(x)$. Again, this will be done for arbitrary stochastic moments (the role played by r). Corollary 3 is the generalization of [5, Theorem 1.3]. For our model right hand side $f(x) = \epsilon^2 \hat{f}(\epsilon x)$ with bounded and compactly supported $\hat{f}$, and an averaging function of the form $g(x) = \hat{g}(\epsilon x)$ with bounded and compactly supported $\hat{g}$, the fluctuations are $O(\epsilon^{d/2})$ in this measure, see (12).

Corollary 2. *Let $\langle\cdot\rangle$ be as in Theorem 1; for compactly supported right hand side $f(x)$ consider the decaying solution $u(x)$ to (6). Let the spatial integrability exponents $2 \le p < \infty$ and $1 \le q < \infty$ be related through $\frac{1}{q} = \frac{1}{d} + \frac{1}{p}$.*
In case of $d > 2$, we have for all $r < \infty$:

$$\left\langle \Big(\sum_x |u - \langle u\rangle|^p \Big)^r \right\rangle^{\frac{1}{rp}} \le C(d, \lambda, \rho, p, r) \Big(\sum_x |f|^q \Big)^{\frac{1}{q}}. \tag{8}$$

In case of $d = 2$, we additionally require $p > 2$ (so that $q > 1$) and that f is supported in $\{|x| \le R\}$ for some $R \ge 1$. Then we have for all $r < \infty$:

$$\left\langle \Big(\sum_{|x| \le R} |u - \langle u\rangle|^p \Big)^r \right\rangle^{\frac{1}{rp}} \le C(\lambda, \rho, p, r)\, (\log^{\frac{1}{2}} R)\, \Big(\sum_x |f|^q \Big)^{\frac{1}{q}}. \tag{9}$$

Corollary 3. *Let $\langle\cdot\rangle$, $f(x)$, $u(x)$ be as in Corollary 2. Let the averaging function $g(x)$ be compactly supported. Let the two integrability exponents $1 < q, \tilde{q} < \infty$ be related by $\frac{1}{q} + \frac{1}{\tilde{q}} = \frac{2}{d} + \frac{1}{2}$. Then it holds for all $r < \infty$:*

$$\left\langle \left| \sum_x (u - \langle u \rangle) g \right|^r \right\rangle^{\frac{1}{r}} \leq C(d, \lambda, \rho, r) \left(\sum_x |f|^q \right)^{\frac{1}{q}} \left(\sum_x |g|^{\tilde{q}} \right)^{\frac{1}{\tilde{q}}}. \tag{10}$$

For the convenience of the reader we express the results of both corollaries in terms of the rescaled variable $\hat{x} = \epsilon x$, the model right hand side $f(x) = \epsilon^2 \hat{f}(\epsilon x)$ and the model averaging function $g(x) = \hat{g}(\epsilon x)$; we also rewrite the solution itself in terms of $u(x) = \hat{u}_\epsilon(\epsilon x)$. In this notation, (8) (multiplied by $\epsilon^{d/p}$) turns into

$$\begin{aligned} &\left\langle \left(\epsilon^d \sum_{\hat{x} \in \epsilon \mathbb{Z}^d} \left| \hat{u}_\epsilon - \langle \hat{u}_\epsilon \rangle \right|^p \right)^r \right\rangle^{\frac{1}{rp}} \\ &\leq \; C(d, \lambda, \rho, p, r) \epsilon \Big(\epsilon^d \sum_{\hat{x} \in \epsilon \mathbb{Z}^d} |\hat{f}|^q \Big)^{\frac{1}{q}} \; \leq \; C(d, \lambda, \rho, r, \hat{f}) \epsilon. \end{aligned} \tag{11}$$

Note that this can be interpreted as the discrete version of

$$\left\langle \left(\int_{\mathbb{R}^d} \left| \hat{u}_\epsilon - \langle \hat{u}_\epsilon \rangle \right|^p d\hat{x} \right)^r \right\rangle^{\frac{1}{rp}} \leq C(d, \lambda, \rho, p, r) \epsilon \Big(\int_{\mathbb{R}^d} |\hat{f}|^q d\hat{x} \Big)^{\frac{1}{q}},$$

which highlights the $O(\epsilon)$-nature of the "spatially strong" error.
Likewise, (10) (multiplied by ϵ^d) turns into

$$\begin{aligned} &\left\langle \left| \epsilon^d \sum_{\hat{x} \in \epsilon \mathbb{Z}^d} (\hat{u}_\epsilon - \langle \hat{u}_\epsilon \rangle) \hat{g} \right|^r \right\rangle^{\frac{1}{r}} \\ &\leq \; C(d, \lambda, \rho, r) \epsilon^{\frac{d}{2}} \left(\epsilon^d \sum_{\hat{x} \in \epsilon \mathbb{Z}^d} |\hat{f}|^q \right)^{\frac{1}{q}} \left(\epsilon^d \sum_{\hat{x} \in \epsilon \mathbb{Z}^d} |\hat{g}|^{\tilde{q}} \right)^{\frac{1}{\tilde{q}}} \\ &\leq \; C(d, \lambda, \rho, r, \hat{f}, \hat{g}) \epsilon^{\frac{d}{2}}. \end{aligned} \tag{12}$$

As above, this can be seen as the discrete version of

$$\left\langle \left| \int_{\mathbb{R}^d} (\hat{u}_\epsilon - \langle \hat{u}_\epsilon \rangle) \hat{g} \right|^r d\hat{x} \right\rangle^{\frac{1}{r}} \leq C(d, \lambda, \rho, r) \epsilon^{\frac{d}{2}} \left(\int_{\mathbb{R}^d} |\hat{f}|^q d\hat{x} \right)^{\frac{1}{q}} \left(\int_{\mathbb{R}^d} |\hat{g}|^{\tilde{q}} d\hat{x} \right)^{\frac{1}{\tilde{q}}},$$

uncovering the $O(\epsilon^{d/2})$-nature of the "spatially weak" error.

Let us make a couple of further more detailed remarks related to Corollaries 2 and 3. In case of Corollary 2 and $d = 2$, we can use Hölder's inequality to establish an estimate also for $p = 2$. However, in that case we pay the price of an arbitrarily small power of R on the right hand side of (9). We also note that the requirement that f has compact support and that u decays can be weakened: All we need is the Green's function representation $u(x) = \sum_y G(x,y)f(y)$. We conclude by pointing out that our argument does not require any smoothness assumptions on $\hat{f}(\hat{x})$ and $\hat{g}(\hat{x})$ beyond (uniform) boundedness to obtain (11) and (12).

In its ϵ-version (12), the estimate of the weak error (10) hints towards the scaling of the Central Limit Theorem (CLT). CLT-scaling has been established for related quantities, as we shall detail now: For $g = f$, the weak measure of fluctuations turns into a measure of fluctuations of the energy:

$$\sum_x (u - \langle u\rangle)g = \sum_x \nabla u \cdot a\nabla u - \langle \sum_x \nabla u \cdot a\nabla u\rangle.$$

If u is the so-called corrector (which is an a-harmonic function of affine behavior on large scales) the (stationary) energy density defines the homogenized coefficient. In [12, Theorem 2.1], it is shown that in the case of independent, identically distributed (i. i. d.) coefficients, the energy density of the corrector has CLT scaling in the sense that spatial averages behave as if the energy density was independent from site to site; in [13, Proposition 7], that has result has been generalized to ensembles that only satisfy a Spectral Gap condition. The scaling result has been substantially sharpened for i. i. d. ensembles: In this situation, the fluctuations of the energy of the corrector become more and more Gaussian as the box over which the spatial average is taken increases. The latter result has been obtained by three different techniques: Nolen [23] gives a quantitative estimate based on a differential characterization of Gaussian distributions (second order Poincaré inequality) and relies on the corrector estimates from [12, Theorem 2.1]. Biskup, Salvi & Wolff [3] obtain a more qualitative result using a Martingale decomposition of the spatially averaged energy density (their result assumes small ellipticity contrast $1 - \lambda \ll 1$, but presumably could be extended using the results of [13]). Rossignol [25] in turn uses an orthogonal decomposition of the space of coefficients (Walsh decomposition).

4 Relation to De Giorgi's approach to elliptic regularity

While our result heavily relies on the celebrated regularity theory for scalar elliptic operators, connected with the names of De Giorgi, Nash, and Moser, it also gives a new perspective on these results. We will specify the input from regularity theory, namely Nash's (upper) bounds on the parabolic Green's function, in the next section. We now address what we see as a new perspective on these results, namely on De Giorgi's result on Hölder continuity of a-harmonic functions.

An elementary consequence of the mean value property is the following Liouville principle: Harmonic functions that grow sub-linearly must be constant. This holds for the constant-coefficient Laplacian both on $\mathbb{R}^d$ and on $\mathbb{Z}^d$, but is no longer true for *variable* coefficients, even if they are uniformly elliptic. Indeed, a well-known example [1, Corollary 16.1.5] shows that for any $\alpha > 0$, there exists an explicit coefficient field $\alpha^2 \leq a(z) \leq 1$ such that $u(z) = \mathcal{R}e(|z|^{\alpha-1}z)$ is a-harmonic in $z \ni \mathbb{C} = \mathbb{R}^2$. We believe that this example can be adapted to the lattice $\mathbb{Z}^2$ (provided the condition of diagonality is relaxed to the condition that the discrete maximum principle is valid, a setting to which our results presumably can be extended). However, a celebrated result of De Giorgi [6, Theorem 2] states that for any dimension d and any ellipticity ratio λ, there exists an exponent $\alpha_0(d,\lambda) > 0$ with the following property: For any field of coefficients $\lambda \leq a(x) \leq 1$ and any a-harmonic function $u(x)$, $|u(x)| \leq |x|^{\alpha_0}$ for $|x| \gg 1$ implies that u is constant. This result holds both in $\mathbb{R}^d$ and in $\mathbb{Z}^d$ [7, Proposition 6.2]. In this sense, while it is no longer true that "sub-linear implies constant", it remains true that "*very* sub-linear implies constant".

De Giorgi's result can be rephrased as an inner regularity result in terms of Hölder continuity with Hölder exponent α_0: For any harmonic function $u(x)$ on $\{x : |x| \leq R\}$, the Hölder-α_0 modulus of continuity at zero is estimated by the supremum:

$$\sup_{x:|x|\leq R} \frac{|u(x) - u(0)|}{|x|^{\alpha_0}} \leq C(d,\lambda) R^{-\alpha_0} \sup_{x:|x|\leq R} |u(x)|.$$

To contrast De Giorgi's result with our result below, let us rephrase it as follows:

$$\forall\ \lambda \leq a(x) \leq 1,\ \ \forall\ R < \infty : \quad \sup_u \frac{\sup_{x:|x|\leq R} \frac{|u(x)-u(0)|}{|x|^{\alpha_0}}}{\frac{1}{R^{\alpha_0}} \sup_{x:|x|\leq R} |u(x)|} \leq C(d,\lambda), \qquad (13)$$

where the outer supremum is taken over all $u(x)$ that satisfy $\nabla^* a \nabla u = 0$ in $\{x : |x| \leq R\}$.

In this context, we will show that Theorem 1 has the following Corollary.

Corollary 4. *For all* $0 < \alpha < 1$*,* $p < \infty$*, and* $R < \infty$*, we have*

$$\left\langle \left(\sup_u \frac{\sup_{x:|x|\leq R} \frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha} \sup_{x:|x|\leq R} |u(x)|} \right)^p \right\rangle \leq C(d, \lambda, \alpha, p), \tag{14}$$

where the outer supremum is taken over all $u(x)$ *that satisfy* $\nabla^* a \nabla u = 0$ *in* $\{x : |x| \leq R\}$*.*

As a consequence of this annealed Hölder regularity, we obtain the following Liouville principle.

Corollary 5. *For* $\langle \cdot \rangle$*-almost every* a*, if* u *is a solution to* $\nabla^* a \nabla u = 0$ *which is strictly sub-linear in the sense that there exists* $0 < \alpha < 1$ *such that*

$$\lim_{R\to\infty} \frac{1}{R^\alpha} \sup_{|x|\leq R} |u(x)| = 0,$$

then u *is constant.*

Loosely speaking, Corollary 4 implies that for "most" coefficient fields, an a-harmonic function $u(x)$ is Hölder continuous with an exponent *arbitrarily close to one*. More precisely, the modulus of near-Lipschitz continuity of $u(x)$ in some large ball $\{x : |x| \leq R\}$ only depends on its supremum in the ball of double radius. This quantitative result has the Liouville principle as an easy corollary: For almost every a, any sub-linear a-harmonic function must be constant, see Corollary 5 for the version of this result we can derive from Corollary 4. For the convenience of the reader, we display the short proof of how Corollary 4 yields Corollary 5. However, surprisingly for us, the qualitative Corollary 5 holds *without any assumption on the ensemble* $\langle \cdot \rangle$ besides stationarity! This is established in a very inspiring paper [2, Theorem 3]. The main ingredients for the short and elegant argument are

- The "annealed" estimate $\langle \sum_x |x|^2 G(t, x, 0) \rangle \lesssim t$ on the second moments of the parabolic Green's function $G(a; t, x, y) \overset{\text{short}}{=} G(t, x, y)$ (cf. [2, (SBD)], see Subsection 6 below for the definition of G), which in our uniformly elliptic context even holds in its stronger "quenched" version, that is, $\sum_x |x|^2 G(t, x, 0) \lesssim t$.

- The annealed estimate $-\langle\sum_x G(t,x,0)\ln G(t,x,0)\rangle \lesssim \log t$ on the spatial entropy of the parabolic Green's function G (cf. [2, p.12]), which in our context is an immediate consequence of the second moments estimate. This ingredient is shown to imply the following annealed continuity property of G:

$$\Big\langle \sum_y G(1,0,y) \sum_x \frac{|G(t,0,x)-G(t-1,y,x)|^2}{G(t,0,x)+G(t-1,y,x)} \Big\rangle \lesssim \frac{1}{t}.$$

5 Logarithmic Sobolev inequality

In the following we give a more detailed description of our use of the logarithmic Sobolev inequality and prove that any i. i. d. ensemble satisfies Definition 1. LSI substitutes the Spectral Gap Inequality (SG) in prior work on quantitative stochastic homogenization. SG has been introduced into the field by Naddaf & Spencer [18, Theorem 1] (in form of the Brascamp-Lieb inequality) and used most recently in [12, Lemma 2.3] in an indirect way and in [13] explicitly. Like SG, LSI follows from the property that there is an integrable fall-off of correlations in the sense of a uniform mixing condition à la Dobrushin-Shlosman, see for instance [26, Theorem 1.8 c)] for a discrete setting. Both SG and LSI quantify ergodicity of the ensemble, see for instance the discussion in [13, Chapter 4]. Recall that the usual LSI in this setting (with continuum derivative) would read

$$\Big\langle f \log \frac{f}{\langle f\rangle} \Big\rangle \le \frac{1}{2\rho} \Big\langle \sum_{y\in\mathbb{Z}^d} \frac{1}{f} (\frac{\partial f}{\partial a(y)})^2 \Big\rangle. \tag{15}$$

The difference to Definition 1 lies solely in the presence of a supremum over $a(y)$ inside the expectation, see (2).

Both SG and LSI are based on the notion of a *vertical derivative* that defines a Dirichlet form and thus a reversible dynamics, namely Glauber dynamics, on the space of coefficient fields (the word "vertical" is used to distinguish this derivative from the "horizontal" derivative naturally arising in stochastic homogenization, but not used in this paper). In the earlier work on stochastic homogenization and motivated by field theories, see [19], the version of SG that is based on the *continuum* vertical derivative (as on the r.h.s. of (15)) has been used [18]. However, this assumption rules out the natural example of coefficients with a single-site distribution that only assumes a *finite* number of values (Bernoulli). In [13], it has been shown that SG based on a *discrete* version of vertical derivative can be used instead — without much more work and allowing for any single-site distribution. Indeed, it is known that SG for the discrete vertical derivative holds for *any* independently and identically distributed field of coefficients: This follows from the tensorization principle and the fact that any single-site distribution satisfies SG for the *discrete* vertical derivative with constant one.

Here lies a difference between SG and LSI: While the tensorization principle still holds for LSI, see for instance [15, Theorem 4.4], it is *not* true that the appropriate discrete version of the usual LSI holds for any single-site distribution — in fact, the discrete LSI does not hold for a continuum distribution. Hence in order to treat arbitrary single-site distributions, we are forced to consider the weakened version of LSI of Definition 1. A similarly weakened version of SG was already considered in [12, Lemma 2.3].

The LSI has been of great use in the setting of stochastic processes and diffusion semi-groups, for the first time introduced in generality by Gross [14]. It implies SG and is equivalent to the notion of hyper-contractivity, see [14, Theorem 1] as well as [15, Theorem 4.1] for a recent exposition. Incidentally, hyper-contractivity was first observed in the Gaussian context by Nelson [21], see [22] for an improved result. It is thus the older notion and in fact motivated the (somewhat implicit) introduction of LSI by Federbush [10]. We refer to [15] for a recent exposition on LSI.

The result of this section is that any independent, identically distributed coefficient-field satisfies the LSI (2) in Definition 1.

Lemma 1. *Consider an ensemble $\langle\cdot\rangle$ of i. i. d. coefficients with arbitrary single-site distribution on $[\lambda,1]^d$. Then (2) holds, i.e.*

$$\Big\langle f\log\frac{f}{\langle f\rangle}\Big\rangle \le \frac{1}{2\rho}\Big\langle\sum_y \sup_{a(y)\in[\lambda,1]^d}\frac{1}{f}\Big(\frac{\partial f}{\partial a(y)}\Big)^2\Big\rangle$$

for all (continuously differentiable) positive functions f of the coefficient field a. The constant ρ can be taken to be $\rho=\frac{1}{2d}$.

Note that even for discrete distributions, we have to take the supremum over the whole box $[\lambda,1]^d$, even though the coefficients attain only a countable number of values. Lemma 1 is an immediate consequence of the following two lemmas. The first one shows that any single-site distribution on $[\lambda,1]^d$ satisfies the LSI in Definition 1.

Lemma 2. *Let $\langle\cdot\rangle$ be any distribution on $[\lambda,1]^d$. Then it holds that*

$$\Big\langle f\log\frac{f}{\langle f\rangle}\Big\rangle \le \frac{1}{2\rho}\sup_{a\in[\lambda,1]^d}\frac{1}{f(a)}\Big(\frac{\partial f(a)}{\partial a}\Big)^2 \tag{16}$$

for all functions $f : [\lambda, 1]^d \to (0, \infty)$. In fact, the constant $\rho = \frac{1}{2d}$ will do.

The next lemma shows that the LSI in Definition 1 satisfies the tensorization principle.

Lemma 3. *Let $\langle\cdot\rangle$ be an ensemble consisting of independent single-site distributions such that each single-site distribution satisfies the single-site LSI* (16) *with the same constant ρ. Then $\langle\cdot\rangle$ itself satisfies the LSI* (2) *with constant ρ.*

6 Main ingredients of the proof

Loosely speaking, our approach consists in upgrading the (optimal) annealed estimates of Delmotte & Deuschel [8, Theorem 1.1] in terms of the integrability p:

Proposition 1. [Delmotte & Deuschel]*. Let $\langle\cdot\rangle$ be stationary. Then we have for all $x \in \mathbb{Z}^d$*

$$\begin{aligned} \langle|\nabla\nabla G(x,0)|\rangle &\leq C(d,\lambda)(|x|+1)^{-d}, && (17)\\ \langle|\nabla_x G(x,0)|\rangle &\leq C(d,\lambda)(|x|+1)^{1-d}. && (18) \end{aligned}$$

More precisely, we refer to the estimates (1.4) and (1.5a) in [8, Theorem 1.1] on the discrete *parabolic* Green's function $G(t,x,y) = G(a;t,x,y)$ (i.e. the solution of $\partial_t G(t,x,y) + \nabla_x^* a(x)\nabla_x G(t,x,y) = 0$ with $G(t=0,x,y) = \delta(x-y)$) that in our notation imply for any weight exponent $\alpha < \infty$:

$$\begin{aligned} \langle|\nabla\nabla G(t,x,0)|\rangle &\leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-1}\Big(\frac{|x|^2}{t+1}+1\Big)^{-\frac{\alpha}{2}}, && (19)\\ \langle|\nabla_x G(t,x,0)|\rangle &\leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-\frac{1}{2}}\Big(\frac{|x|^2}{t+1}+1\Big)^{-\frac{\alpha}{2}}. && (20) \end{aligned}$$

(In fact, [8] establishes (19) & (20) with exponentially decaying weights instead of just algebraically decaying ones.) Since the elliptic Green's function can be inferred from the parabolic one via $G(x,y) = \int_0^\infty G(t,x,y)dt$, these estimates imply (17) & (18) (by fixing some $\alpha > d$ and performing the change of variables $\hat{t} = |x|^{-2}(t+1)$). Actually, [8] establishes (20) and thus (18) in the stronger form where the L^1-norm $\langle|\cdot|\rangle$ is replaced by the L^2-norm $\langle|\cdot|^2\rangle^{1/2}$: $\langle|\nabla_x G(x,0)|^2\rangle^{1/2} \leq C(d,\lambda,\alpha)(|x|+1)^{1-d}(\frac{|x|^2}{t+1}+1)^{-\frac{\alpha}{2}}$.

Let us point out that the *spatially point-wise annealed* estimates (19) & (20) are consequences of the following *spatially averaged quenched* estimates

$$\sum_x \Big((\frac{|x|^2}{t+1}+1)^{\frac{\alpha}{2}} G(t,x,0) \Big)^2 \;\leq\; C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}}, \tag{21}$$

$$\sum_x \Big((\frac{|x|^2}{t+1}+1)^{\frac{\alpha}{2}} |\nabla_x G(t,x,0)| \Big)^2 \;\leq\; C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-1}. \tag{22}$$

The first estimate (21) is the (upper, off-diagonal part of the) celebrated Nash estimate [20, Appendix]. The discrete case was treated in full generality in [4, Corollary 3.28]. The second estimate (22) is a consequence of the first one. For an elementary proof of both, we refer to [13, Lemmas 24 & 25], with the Nash inequality as only noteworthy ingredient. Let us point out how (22) is implies (19): Using the semi group property in form of $\nabla\nabla G(t,x,x') = \sum_y \nabla_x G(\frac{t}{2},x,y)\nabla_{x'} G(\frac{t}{2},y,x')$ we obtain by the triangle inequality for the weight, Cauchy Schwarz in $\sum_y$, and the symmetry of $G(t,x,y)$ in x and y:

$$\begin{aligned}
&\Big(\frac{|x-x'|^2}{t+1}+1\Big)^{\frac{\alpha}{2}} |\nabla\nabla G(t,x,x')| \\
&\leq \sum_y \Big(\frac{2|x-y|^2}{t+1}+1\Big)^{\frac{\alpha}{2}} |\nabla_x G(\frac{t}{2},x,y)| \Big(\frac{2|y-x'|^2}{t+1}+1\Big)^{\frac{\alpha}{2}} |\nabla_{x'} G(\frac{t}{2},y,x')| \\
&\leq \Bigg(\sum_y \Big((\frac{2|x-y|^2}{t+1}+1)^{\frac{\alpha}{2}} |\nabla_x G(\frac{t}{2},x,y)| \Big)^2 \\
&\qquad \times \sum_y \Big((\frac{2|x'-y|^2}{t+1}+1)^{\frac{\alpha}{2}} |\nabla_{x'} G(\frac{t}{2},x',y)| \Big)^2 \Bigg)^{\frac{1}{2}}.
\end{aligned}$$

Note that the right-hand side of the last inequality does not allow for application of (22), since the sum is not in the variable the derivative is taken. However, taking the expectation, using Cauchy Schwarz in $\langle\cdot\rangle$, and using stationarity and symmetry in form of $\langle |\nabla_x G(\frac{t}{2},x,y)|^2\rangle = \langle |\nabla_2 G(\frac{t}{2},x-y,0)|^2\rangle = \langle |\nabla_3 G(\frac{t}{2},0,x-y)|^2\rangle = \langle |\nabla_y G(\frac{t}{2},-x,-y)|^2\rangle$ (where ∇_2 stands for the derivative w. r. t. to the second variable), yields

$$
\begin{aligned}
&(\frac{|x-x'|^2}{t+1}+1)^{\frac{\alpha}{2}}\langle|\nabla\nabla G(t,x,x')|\rangle\\
\leq\;& \Big(\sum_y(\frac{2|x-y|^2}{t+1}+1)^{\alpha}\langle|\nabla_x G(\frac{t}{2},x,y)|^2\rangle\\
&\times\sum_y(\frac{2|x'-y|^2}{t+1}+1)^{\alpha}\langle|\nabla_{x'} G(\frac{t}{2},x',y)|^2\rangle\Big)^{\frac{1}{2}}\\
=\;& \Big(\Big\langle\sum_y(\frac{2|x-y|^2}{t+1}+1)^{\alpha}|\nabla_y G(\frac{t}{2},-x,-y)|^2\Big\rangle\\
&\times\Big\langle\sum_y(\frac{2|x'-y|^2}{t+1}+1)^{\alpha}|\nabla_y G(\frac{t}{2},-x',-y)|^2\Big\rangle\Big)^{\frac{1}{2}}.
\end{aligned}
$$

We now see that (22) implies (19). The estimate (20) is derived via the semi group property in form of $\nabla_x G(t,x,x')=\sum_y\nabla_x G(\frac{t}{2},x,y)G(\frac{t}{2},y,x')$ from the combination of (21) and (22) by an analogous argument.

Note that the estimates of Proposition 1 make *no* assumptions on the ensemble besides stationarity. In order to pass from Proposition 1 to Theorem 1, we need the assumption on the ensemble from Definition 1. In fact, LSI enters only through the following lemma.

Lemma 4. *Let $\langle\cdot\rangle$ be stationary and satisfy LSI with constant $\rho>0$. Then for arbitrary $\delta>0$ and $1\leq p<\infty$ and for any $\zeta(a)$ it holds that*

$$
\langle|\zeta|^{2p}\rangle^{\frac{1}{2p}} \;\leq\; C(d,\rho,p,\delta)\langle|\zeta|\rangle+\delta\Big\langle\Big(\sum_x\sup_{a(x)}(\frac{\partial\zeta}{\partial a(x)})^2\Big)^p\Big\rangle^{\frac{1}{2p}}. \tag{23}
$$

In order to make use of Lemma 4, we need to estimate the vertical derivatives of $\nabla\nabla G$ and $\nabla_x G$. The following lemma is at the core of our result.

Lemma 5. *There exists an integrability exponent $p_0=p_0(d,\lambda)<\infty$ such that for all $p\geq p_0$ and any $x\in\mathbb{Z}^d$, it holds*

$$
\begin{aligned}
&(|x|+1)^d\Big\langle\Big(\sum_y\sup_{a(y)}\big|\frac{\partial}{\partial y}\nabla\nabla G(x,0)\big|^2\Big)^p\Big\rangle^{\frac{1}{2p}}\\
&\leq C(d,\lambda,p)\sup_z\Big\{(|z|+1)^d\langle|\nabla\nabla G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big\}
\end{aligned} \tag{24}
$$

and

$$
\begin{aligned}
&(|x|+1)^{d-1}\Big\langle\Big(\sum_y \sup_{a(y)}\big|\frac{\partial}{\partial y}\nabla_x G(x,0)\big|^2\Big)^p\Big\rangle^{\frac{1}{2p}} \\
&\leq C(d,\lambda,p)\bigg(\sup_z\Big\{(|z|+1)^{d-1}\langle|\nabla_z G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big\} \qquad (25)\\
&\qquad + \sup_z\Big\{(|z|+1)^{d}\langle|\nabla\nabla G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big\}\bigg). \qquad (26)
\end{aligned}
$$

The formulation of Lemma 5 shows that with our method, we first have to estimate the mixed *second* derivatives $\langle|\nabla\nabla G(x,0)|^{2p}\rangle$ before we can tackle the *first* derivatives $\langle|\nabla_x G(x,0)|^{2p}\rangle$. It also reveals that it is necessary to estimate *high moments* $p \geq p_0$ in $\langle\cdot\rangle$ in order to estimate *moderately low moments* like the fourth moment $\langle|\nabla_x G(x,0)|^4\rangle$ that is needed in the proof of Corollary 1.

The preceding lemma relies on the following suboptimal, but *quenched* estimates on the (elliptic) Green's function:

Lemma 6. [Gloria & Otto] *There exists* $\alpha_0 = \alpha_0(d,\lambda) > 0$ *such that for all* $R > 0$

$$
|\nabla\nabla G(0,0)|^2 + R^{2\alpha_0}\sum_{R\leq|x|<2R}|\nabla\nabla G(x,0)|^2 \;\leq\; C(d,\lambda), \qquad (27)
$$

$$
|\nabla_x G(0,0)|^2 + \sum_{R\leq|x|<2R}|\nabla_x G(x,0)|^2 \;\leq\; C(d,\lambda). \qquad (28)
$$

The estimate (28) was established in the stronger (dimensionally optimal) form of $\sum_{R\leq|x|\leq 2R}|\nabla_x G(x,0)|^2 \lesssim R^{2-d}$ in [12, Lemma 2.9]; in its weaker form of (28), it is straight forward for $d > 2$. The proof of estimate (28) in [12] in case of $d = 2$ is subtle and relied on an adaptation of [9]. In this paper, we will give an elementary argument for the estimate (27), which we could not find in the literature. However, in the stronger form of $\sum_{R\leq|x|<2R}|\nabla\nabla G(x,0)|^2 \leq C(d,\lambda)R^{2-d-2\alpha_0}$, this estimate can also be seen as a consequence of the following classical ingredients:

- the optimal decay of $G(x,y)$ itself, that is just needed in a spatially averaged sense of $R^{-d}\sum_{y:R\leq|x-y|<2R}|G(x,y)-\bar{G}| \leq C(d,\lambda)R^{2-d}$ (thanks to subtracting the average $\bar{G}$ over the annulus $\{y : R \leq |x-y| \leq 2R\}$, this estimate also holds in $d = 2$),

- De Giorgi's Hölder continuity estimate, that then yields for some $\alpha_0 = \alpha_0(d,\lambda) > 0$ that $\sup_{y:R\leq|x-y|<2R}|\nabla_x G(x,y)| \leq C(d,\lambda)R^{2-d-\alpha_0}$,

- Caccioppoli's estimate, that then yields $\sum_{x:R\le|x-y|<2R}|\nabla_y\nabla_x G(x,y)|^2 \le C(d,\lambda)R^{2-d-2\alpha_0}$.

Remark 2. *We remark that with the same proof, one obtains a periodic version of Theorem 1 (with constants uniform in L) for the Green's function defined in* (1)*. In that case, one just replaces the Euclidean distance* $|x|$ *on* $\mathbb{Z}^d$ *by its periodic version* $\mathrm{dist}(x, L\mathbb{Z}^d)$ *on the torus* $\mathbb{R}/L\mathbb{Z}^d$*. The periodic version of Proposition 1 follows as above from the quenched spatially averaged estimates of [13, Theorem 3(b)].*

7 Proofs

Proof of Lemma 4.

Step 1. Result for $p=1$. We claim that for any $\delta>0$ and all $\zeta(a)$, it holds

$$\langle\zeta^2\rangle^{\frac12} \le \left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)\langle|\zeta|\rangle + \delta\Big\langle\sum_x \sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big\rangle^{\frac12}, \tag{29}$$

where ρ denote the constant in the LSI, see Definition 1. By homogeneity, we may assume $\langle\zeta^2\rangle=1$. For all real-valued ζ it holds that

$$\zeta^2 \le \left\{\begin{matrix} \exp(\frac{2}{\rho\delta^2})|\zeta| & \text{if } |\zeta|\le\exp\frac{2}{\rho\delta^2} \\ \frac{\rho\delta^2}{4}\zeta^2\log\zeta^2 & \text{if } |\zeta|\ge\exp\frac{2}{\rho\delta^2}\end{matrix}\right\}.$$

Since $x\log x$ is bounded from below by $\frac1e$, it is elementary to verify that $\frac2e|\zeta|+\zeta^2\log\zeta^2\ge 0$ whenever $|\zeta|\le 1$. Since furthermore $\zeta^2\log\zeta^2>0$ whenever $|\zeta|>1$, we find that

$$\zeta^2 \le \left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)|\zeta| + \frac{\rho\delta^2}{4}\zeta^2\log\zeta^2.$$

Hence taking the expectation $\langle\cdot\rangle$ yields

$$\langle\zeta^2\rangle \le \left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)\langle|\zeta|\rangle + \frac{\rho\delta^2}{4}\Big\langle\zeta^2\log\zeta^2\Big\rangle.$$

Since $\langle\zeta^2\rangle=1$, Young's inequality yields

$$\begin{aligned}\langle|\zeta|\rangle &\le \frac12\left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)\langle|\zeta|\rangle^2 + \frac12\left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)^{-1} \\ &= \frac12\left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)\langle|\zeta|\rangle^2 + \frac12\left(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\right)^{-1}\langle\zeta^2\rangle.\end{aligned}$$

Combining the last two estimates, we deduce

$$\langle\zeta^2\rangle \le \Bigg(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\Bigg)^2\langle|\zeta|\rangle^2+\frac{\rho\delta^2}{2}\Big\langle\zeta^2\log\frac{\zeta^2}{\langle\zeta^2\rangle}\Big\rangle.$$

Hence LSI yields

$$\langle\zeta^2\rangle \le \Bigg(\exp\Big(\frac{2}{\rho\delta^2}\Big)+\frac{\rho\delta^2}{2e}\Bigg)^2\langle|\zeta|\rangle^2+\delta^2\Big\langle\sum_x\sup_{a(x)}\big(\frac{\partial\zeta}{\partial a(x)}\big)^2\Big\rangle$$

and estimate (29) follows from taking the square root and applying the estimate $\sqrt{a+b}\le\sqrt{a}+\sqrt{b}$.

Step 2. We finish the proof of (23), i.e. show that

$$\langle\zeta^{2p}\rangle^{\frac{1}{2p}} \le C(\rho,p,\delta)\langle|\zeta|\rangle+\delta\Big(\Big\langle\big(\sum_x\sup_{a(x)}(\frac{\partial\zeta}{\partial a(x)})^2\big)^p\Big\rangle\Big)^{\frac{1}{2p}}$$

for general $p\ge1$. To that end, we apply (29) to ζ replaced by $|\zeta|^p$ and take the $2p$-th root:

$$\langle|\zeta|^{2p}\rangle^{\frac{1}{2p}} \le C(\rho,p,\delta)\langle|\zeta|^p\rangle^{\frac{1}{p}}+\Big(\delta\Big\langle\sum_x\sup_{a(x)}\big(\frac{\partial}{\partial a(x)}|\zeta|^p\big)^2\Big\rangle\Big)^{\frac{1}{2p}},$$

where $C(\rho,p,\delta)$ denotes a generic constant only depending on ρ, p, and δ. Since $p<2p$, an application of Hölder's inequality in $\langle\cdot\rangle$ and Young's inequality on the first r. h. s. term yields

$$\langle|\zeta|^{2p}\rangle^{\frac{1}{2p}} \le C(\rho,p,\delta)\langle|\zeta|\rangle+2\Big(\delta\Big\langle\sum_x\sup_{a(x)}\big(\frac{\partial}{\partial a(x)}|\zeta|^p\big)^2\Big\rangle\Big)^{\frac{1}{2p}}. \tag{30}$$

Now we apply the chain rule $\frac{\partial}{\partial a(x)}|\zeta|^p=p|\zeta|^{p-2}\zeta\frac{\partial}{\partial a(x)}\zeta$, since $p>1$. Furthermore we note that for every coefficient field $a\in[\lambda,1]^{\mathbb{Z}^d}$, it holds that $\sup_{a(x)}|\zeta|\le|\zeta|+\sqrt{d}\sup_{a(x)}|\frac{\partial\zeta}{\partial a(x)}|$ and hence

$$\sup_{a(x)}\Big(|\zeta|^{2p-2}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big)\le C(d,p)\Bigg(|\zeta|^{2p-2}\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2+\Big(\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big)^p\Bigg).$$

Hence

$$\begin{aligned}
&\Big\langle\sum_x\sup_{a(x)}\Big(\frac{\partial}{\partial a(x)}|\zeta|^p\Big)^2\Big\rangle\\
&\le C(d,p)\Bigg(\Big\langle\sum_x|\zeta|^{2p-2}\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big\rangle+\Big\langle\sum_x\Big(\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big)^p\Big\rangle\Bigg)\\
&\le C(d,p)\Bigg(\langle|\zeta|^{2p}\rangle^{1-\frac{1}{p}}\Big\langle\Big(\sum_x\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big)^p\Big\rangle^{\frac{1}{p}}+\Big\langle\Big(\sum_x\sup_{a(x)}\Big(\frac{\partial\zeta}{\partial a(x)}\Big)^2\Big)^p\Big\rangle\Bigg).
\end{aligned}$$

Here in the last line, we have used Hölder's inequality in $\langle\cdot\rangle$ with exponents $(\frac{p-1}{p}, p)$ on the first term and the embedding $\ell^2(\mathbb{Z}^d) \subset \ell^{2p}(\mathbb{Z}^d)$ on the second term. Another application of Young's inequality thus allows us to pass from (30) to

$$\langle|\zeta|^{2p}\rangle^{\frac{1}{2p}} \leq C(d,\rho,p,\delta)\langle|\zeta|\rangle + C(p)\Big((\delta+\delta^p)\Big\langle\Big(\sum_x \sup_{a(x)} \big(\frac{\partial\zeta}{\partial a(x)}\big)^2\Big)^p\Big\rangle\Big)^{\frac{1}{2p}}.$$

By redefining δ, this yields (23).

Proof of Lemma 6.
We just give the proof of (27); for (28), we refer to [12, Lemma 2.9].
Step 1. In this step, we derive the a priori estimate

$$\sum_x |\nabla\nabla G(x,0)|^2 \leq C(d,\lambda). \tag{31}$$

Indeed, note that the weak formulation of the defining equation for G reads

$$\forall\, \zeta(x): \quad \sum_x \nabla\zeta(x)\cdot a(x)\nabla_x G(x,x') = \zeta(x').$$

Taking the derivative w. r. t. the variable x' in some direction $i = 1,\cdots,d$ yields

$$\forall\, \zeta(x): \quad \sum_x \nabla\zeta(x)\cdot a(x)\nabla_x\nabla_{i,x'} G(x,x') = \nabla_i\zeta(x'). \tag{32}$$

The choice of $\zeta(x) = \nabla_{i,x'}G(x,x')$ (we address the question of admissibility of this test function below) yields

$$\sum_x \nabla_x\nabla_{i,x'}G(x,x')\cdot a(x)\nabla_x\nabla_{i,x'}G(x,x') = \nabla_{i,x}\nabla_{i,x'}G(x,x')|_{x=x'}.$$

Since $a(x) \geq \lambda$, this implies (31) in the explicit form of

$$\sum_x |\nabla_x\nabla_{i,x'}G(x,x')|^2 \leq \lambda^{-2}. \tag{33}$$

We now turn to the question of admissibility of $\zeta(x) = \nabla_{i,x'}G(x,x')$ as a test function for (32), i.e. the question of decay as $|x| \uparrow \infty$ of this function and its gradient. This issue can be circumvented as in Step 3 below through approximation by the *periodic* problem. More precisely, we consider the *periodic* discrete elliptic Green's function $G_L(x,x') = G_L(a,x,x')$ of period L. Up to additive constants, it is characterized by the weak equation

$$\sum_{x\in[-\frac{L}{2},\frac{L}{2})^d} \nabla\zeta(x)\cdot a(x)\nabla_x G_L(x,x') = \zeta(x') - L^{-d} \sum_{x\in[-\frac{L}{2},\frac{L}{2})^d} \zeta(x)$$

for all periodic $\zeta(x)$. With the same argument as above, we obtain

$$\sum_{x\in[-\frac{L}{2},\frac{L}{2})^d} |\nabla_x\nabla_{i,x'}G_L(x,x')|^2 \le \lambda^{-2}. \tag{34}$$

Since $G_L(x,x')$ converges point-wise to $G(x,x')$, the latter implies (33) in the limit $L\uparrow\infty$ by Fatou's lemma. Incidentally, $\lim_{L\uparrow\infty}\nabla_x G_L(x,x')$ may be taken as a definition of $\nabla_x G(x,x')$ in the case of $d=2$, where G itself is not unambiguously defined.

In the following steps, we use the fact that $u(x) = \nabla_{i,x'}G(x,x')|_{x'=0}$, $i = 1,\cdots,d$, is a-harmonic away from $x=0$ to show that there exists a decay exponent $\alpha_0(d,\lambda)>0$ such that for all $R\ge C(d)$ we have

$$\sum_{x:|x|\ge R} |\nabla u|^2 \le C(d,\lambda)R^{-2\alpha_0} \sum_{x:|x|\ge 1} |\nabla u|^2. \tag{35}$$

Together with (31), this implies (27). In Step 2, we'll formally treat the continuum whole-space case. In Step 3, we will show how to make the continuum case rigorous by approximation through the continuum periodic case. More precisely, using (31), we will directly prove the estimate (27) in form of

$$\sum_{x:|x|\ge R} |\nabla u|^2 \le C(d,\lambda)R^{-2\alpha_0}. \tag{36}$$

In Step 4, we indicate the changes necessary to treat the discrete case.

Step 2. Formal derivation of the continuum version of (35), that is

$$\int_{\{|x|\ge R\}} |\nabla u|^2 dx \le C(d,\lambda)R^{-2\alpha_0} \int_{\{|x|\ge 1\}} |\nabla u|^2 dx \tag{37}$$

for $R\ge 1$ and a function $u(x)$ satisfying

$$-\nabla_x\cdot a(x)\nabla_x u = 0 \quad \text{in } \{|x|>1\}. \tag{38}$$

Indeed, let $\eta(x)$ be a cut-off function for $\{x:|x|\ge 2R\}$ in $\{x:|x|\ge R\}$. We test (38) with $\zeta = \eta^2(u-\bar{u})$, where $\bar{u}$ is the spatial average of u on the annulus $\{x : R \le |x| \le 2R\}$. It is a priori not clear that this is an admissible test function for (38); we shall address this in the next step. We appeal to the identity

$$\nabla(\eta^2(u-\bar{u}))\cdot a\nabla u = \nabla(\eta(u-\bar{u}))\cdot a\nabla(\eta(u-\bar{u})) - (u-\bar{u})^2\nabla\eta\cdot a\nabla\eta, \quad (39)$$

which in view of $\lambda \le a(x) \le 1$ turns into the inequality

$$\nabla(\eta^2(u-\bar{u}))\cdot a\nabla u \ge \lambda|\nabla(\eta(u-\bar{u}))|^2 - (u-\bar{u})^2|\nabla\eta|^2. \quad (40)$$

Hence from testing (38) we obtain

$$\lambda\int|\nabla(\eta(u-\bar{u}))|^2dx \le \int(u-\bar{u})^2|\nabla\eta|^2dx,$$

which by the choice of η yields the Caccioppoli estimate

$$\int_{\{x:|x|\ge 2R\}}|\nabla u|^2dx \le C(d,\lambda)R^{-2}\int_{\{x:R\le|x|\le 2R\}}(u-\bar{u})^2dx. \quad (41)$$

By Poincaré's estimate on $\{x : R \le |x| \le 2R\}$ with mean value zero, this turns into

$$\int_{\{x:|x|\ge 2R\}}|\nabla u|^2dx \le C(d,\lambda)\int_{\{x:R\le|x|\le 2R\}}|\nabla u|^2dx,$$

which can be reformulated as

$$\int_{\{x:|x|\ge R\}}|\nabla u|^2dx \le C(d,\lambda)\int_{\{x:R\le|x|\le 2R\}}|\nabla u|^2dx, \quad (42)$$

where $C(d,\lambda)$ as always denotes a *generic* constant only depending on d and on λ whose value may change from line to line.

A standard iteration argument now leads from (42) to (37): Introducing the notation $I_k := \int_{\{x:|x|\ge 2^k\}}|\nabla u|^2dx$, (42) reads

$$\forall k\in\{0,1,\cdots\}\quad I_k \le C(d,\lambda)(I_k - I_{k+1}),$$

which with help of $\theta = \theta(d,\lambda) := 1 - \frac{1}{C} < 1$ can be reformulated

$$\forall k\in\{0,1,\cdots\}\quad I_{k+1} \le \theta I_k,$$

or with help of $\alpha_0 = \alpha_0(d,\lambda) := \frac{-\log\theta}{2\log 2}$ as

$$\forall k\in\{0,1,\cdots\}\quad I_k \le \theta^k I_0 = (2^k)^{-2\alpha_0}I_0.$$

In the original notation, this implies (37) in form of

$$\forall R \geq 1 \quad \int_{\{x:|x|\geq R\}} |\nabla u|^2 dx \leq (\frac{R}{2})^{-2\alpha_0} \int_{\{x:|x|\geq 1\}} |\nabla u|^2 dx.$$

Step 3. Rigorous derivation of the continuum version (36) for $R \geq 1$, and where u is now specified to be a partial derivative of the Green's function, i.e. $u(x) = \nabla_{i,x'} G(x, x')|_{x'=0}$ with $i = 1, \cdots, d$. In this step, as opposed to the previous step, we deal with the issue that we don't know a priori that $\eta^2(u - \bar{u})$ is an admissible test function for (38). More precisely, we worry about the *decay* at $|x| \uparrow \infty$ — we don't worry about local smoothness since anyway, we'll apply the argument to the discrete case in the next step. As in Step 1, we circumvent the problem of decay through approximation by the *periodic* problem. More precisely, we consider the *periodic* continuum elliptic Green's function $G_L(x, x') = G_L(a, x, x')$ of period L. Up to additive constants, it is characterized by the weak equation

$$\int_{[-\frac{L}{2},\frac{L}{2})^d} \nabla\zeta(x) \cdot a(x) \nabla_x G_L(x, x') dx = \zeta(x') - L^{-d} \int_{[-\frac{L}{2},\frac{L}{2})^d} \zeta(x) dx \qquad (43)$$

for all periodic $\zeta(x)$. We note that $u_L(x) = \nabla_{i,x'} G_L(x, x')|_{x'=0}$ thus is characterized by

$$\int_{[-\frac{L}{2},\frac{L}{2})^d} \nabla\zeta(x) \cdot a(x) \nabla_x u_L(x) dx = \nabla_i \zeta(0). \qquad (44)$$

Since $\nabla\nabla G_L$ distributionally converges to $\nabla\nabla G$ as $L \uparrow \infty$, it is enough to show (44) implies

$$\int_{[-\frac{L}{2},\frac{L}{2})^d \cap \{|x|\geq R\}} |\nabla u_L|^2 dx \leq C(d, \lambda) R^{-2\alpha_0} \int_{[-\frac{L}{2},\frac{L}{2})^d \cap \{|x|\geq 1\}} |\nabla u_L|^2 dx \qquad (45)$$

for $1 \leq R \leq C(d)L$. Indeed we can estimate the right-hand side of (45) using (34) and apply weak lower semi-continuity to take the limit as $L \to \infty$ on the left-hand side to obtain (36). Now, disregarding smoothness issues, $\eta^2(u_L - \bar{u}_L)$ is an admissible test function for (44). The argument for (45) is identical to the one in Step 2.

Step 4. Rigorous derivation of (35) for $R \geq C(d)$. In this step, we indicate the modifications in Step 2 (or rather Step 3) that are necessary to treat the discrete case. The first modification results from the fact that Leibniz' rule and thus the neat identity (39) does not hold anymore. However, we claim that the estimate (40) survives in form of

$$\nabla(\eta^2(u - \bar{u})) \cdot a\nabla u \geq \sum_{i=1}^{d} \Big(\lambda(\nabla_i(\eta(u - \bar{u})))^2 - ([u]_i - \bar{u})^2(\nabla_i\eta)^2\Big), \qquad (46)$$

where $[u]_i(x) = \frac{1}{2}(u(x) + u(x+e_i))$ denotes a local average of u. Indeed, since $\lambda \le a(x) \le 1$ is diagonal, it is enough to show for any two functions $\eta(x)$ and $v(x)$

$$\nabla_i(\eta^2 v)\nabla_i v \ge (\nabla_i(\eta v))^2 - [v]_i^2(\nabla_i \eta)^2.$$

This in turn follows from the simple inequality on 4 *numbers* $\eta, \tilde{\eta}, v, \tilde{v}$:

$$\begin{aligned} &(\eta^2 v - \tilde{\eta}^2 \tilde{v})(v - \tilde{v}) - (\eta v - \tilde{\eta}\tilde{v})^2 \\ &\quad = \; -(\eta - \tilde{\eta})^2 v\tilde{v} \;\ge\; -(\eta - \tilde{\eta})^2(\tfrac{1}{2}(v + \tilde{v}))^2. \end{aligned}$$

Hence, if $\eta(x)$ denotes the (slightly narrower) cut-off function for $\{x : |x| \ge 2R - 2\}$ in $\{x : |x| \ge R + 2\}$ (which is OK for $R \ge 5$), from (46) we obtain the following substitute of (41)

$$\begin{aligned} \sum_{x:|x|\ge 2R} |\nabla u|^2 \;&\le\; C(d,\lambda)R^{-2}\sum_{i=1}^{d}\sum_{x:R+1\le|x|\le 2R-1}([u]_i - \bar{u})^2 \\ &\le\; C(d,\lambda)R^{-2}\sum_{x:R\le|x|\le 2R}(u - \bar{u})^2. \end{aligned} \tag{47}$$

The second modification comes from the fact that we need a *discrete* version of the Poincaré estimate with mean value zero on the annulus $\mathbb{Z}^d \cap \{R \le |x| \le 2R\}$, which obviously holds with a constant $C(d)R^2$ provided that $R \ge C(d)$.

Proof of Lemma 5.

Step 1. In this step, we derive the following formulas for the continuum vertical derivative $\frac{\partial}{\partial a_{kk}(y)}$ of spatial derivatives of the Green's function:

$$\frac{\partial}{\partial a_{kk}(y)} G(x,x') \;=\; -\nabla_{k,y}G(x,y)\nabla_{k,y}G(y,x'), \tag{48}$$

$$\frac{\partial}{\partial a_{kk}(y)} \nabla_{i,x} G(x,x') \;=\; -\nabla_{i,x}\nabla_{k,y}G(x,y)\nabla_{k,y}G(y,x'), \tag{49}$$

$$\frac{\partial}{\partial a_{kk}(y)} \nabla_{i,x}\nabla_{i',x'} G(x,x') \;=\; -\nabla_{i,x}\nabla_{k,y}G(x,y)\nabla_{k,y}\nabla_{i',x'}G(y,x'). \tag{50}$$

These formulas highlight the fact that the directional spatial derivative $\nabla_{i,x}$ is associated with the bond $b = x + e_i$; likewise, the continuum vertical derivative $\frac{\partial}{\partial a_{kk}(y)}$ is associated with the bond $e = y + e_k$. In the next step, it will thus be more convenient to introduce the notation $\nabla_{i,x}\zeta(x) = \nabla\zeta(b)$ and $\frac{\partial}{\partial a_{kk}(y)}\zeta = \frac{\partial}{\partial a(e)}\zeta$. Equipped with this notation, we rewrite (48) - (50) in the more compact form

$$\begin{aligned}\frac{\partial}{\partial a(e)}G(x,x') &= -\nabla G(x,e)\nabla G(e,x'), &(51)\\ \frac{\partial}{\partial a(e)}\nabla G(b,x') &= -\nabla\nabla G(b,e)\nabla G(e,x'), &(52)\\ \frac{\partial}{\partial a(e)}\nabla\nabla G(b,b') &= -\nabla\nabla G(b,e)\nabla\nabla G(e,b'). &(53)\end{aligned}$$

We turn to the argument for (48) - (50): From the defining equation

$$\nabla_x^* a(x)\nabla_x G(x,x') = \delta(x-x'),$$

we obtain by Leibniz' rule from applying $\frac{\partial}{\partial a_{kk}(y)}$ that

$$\nabla_x^* a(x)\nabla_x \frac{\partial}{\partial a_{kk}(y)}G(x,x') + \nabla_{k,x}^*\delta(x-y)\nabla_{k,x}G(x,x') = 0$$

and thus

$$\begin{aligned}\frac{\partial}{\partial a_{kk}(y)}G(x,x') &= -\sum_z G(x,z)\nabla_{k,z}^*\delta(z-y)\nabla_{k,z}G(z,x')\\ &= -\sum_z \delta(z-y)\nabla_{k,z}G(x,z)\nabla_{k,z}G(z,x')\\ &= -\nabla_{k,y}G(x,y)\nabla_{k,y}G(y,x'),\end{aligned}$$

yielding the representation (48). An application of $\nabla_{i,x}$ (with $i = 1,\cdots,d$) yields (49), a further application of $\nabla_{i',x'}$ (with $i' = 1,\cdots,d$) yields (50).

Step 2. In this step, we argue that for any two sites x and y, the dependence of the derivatives $\nabla_y G(y,x)$ and $\nabla\nabla G(y,x)$ on the value $a(y)$ of the conductivity is mild in the sense of

$$\begin{aligned}\sup_{a(y)}|\nabla_y G(y,x)| &\leq \exp(d^{\frac{1}{2}}\lambda^{-1})\inf_{a(y)}|\nabla_y G(y,x)|, &(54)\\ \sup_{a(y)}|\nabla\nabla G(y,x)| &\leq \exp(d^{\frac{1}{2}}\lambda^{-1})\inf_{a(y)}|\nabla\nabla G(y,x)|. &(55)\end{aligned}$$

An immediate consequence of (48) - (50), (54) and (55) is

$$\begin{aligned}\sup_{a(y)}|\frac{\partial}{\partial a(y)}G(x,x')| &\leq C(d,\lambda)|\nabla_y G(x,y)||\nabla_y G(y,x')|, &(56)\\ \sup_{a(y)}|\frac{\partial}{\partial a(y)}\nabla_x G(x,x')| &\leq C(d,\lambda)|\nabla\nabla G(x,y)||\nabla_y G(y,x')|, &(57)\\ \sup_{a(y)}|\frac{\partial}{\partial a(y)}\nabla\nabla G(x,x')| &\leq C(d,\lambda)|\nabla\nabla G(x,y)||\nabla\nabla G(y,x')|. &(58)\end{aligned}$$

Coordinate-wise, (54) turns into

$$\begin{aligned}
&\sup_{\{a_{kk}(y)\}_{k=1,\cdots,d}} \sum_{j=1}^{d} (\nabla_{j,y} G(y,x))^2 \\
&\le \ \exp(2d^{\frac{1}{2}}\lambda^{-1}) \inf_{\{a_{kk}(y)\}_{k=1,\cdots,d}} \sum_{j=1}^{d} (\nabla_{j,y} G(y,x))^2
\end{aligned}$$

and by summation over $i = 1, \cdots, d$, (55) is a consequence of

$$\begin{aligned}
&\sup_{\{a_{kk}(y)\}_{k=1,\cdots,d}} \sum_{j=1}^{d} (\nabla_{j,y}\nabla_{i,x} G(y,x))^2 \\
&\le \ \exp(2d^{\frac{1}{2}}\lambda^{-1}) \inf_{\{a_{kk}(y)\}_{k=1,\cdots,d}} \sum_{j=1}^{d} (\nabla_{j,y}\nabla_{i,x} G(y,x))^2.
\end{aligned}$$

In terms of the notation introduced introduced at the end of Step 1, we thus have to show for any pair of site x, y and any bond b

$$\begin{aligned}
\sup_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla G(e',x))^2 \ &\le \ \exp(2d^{\frac{1}{2}}\lambda^{-1}) \inf_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla G(e',x))^2, \\
\sup_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla\nabla G(e',b))^2 \ &\le \ \exp(2d^{\frac{1}{2}}\lambda^{-1}) \inf_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla\nabla G(e',b))^2,
\end{aligned}$$

where S is the set of bonds adjacent to y, i.e. $\{y+e_1, \cdots, y+e_d\}$. In fact, we shall show for *any* finite set S of bonds that

$$\max_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla u(e'))^2 \le \exp(2|S|^{\frac{1}{2}}\lambda^{-1}) \min_{\{a(e)\}_{e\in S}} \sum_{e'\in S} (\nabla u(e'))^2, \qquad (59)$$

where the field $u = u(x')$ either denotes $u(x') = G(x', x)$ or $u(x') = \nabla G(x', b)$ and $|S|$ denotes the number of bonds in the set S. At the basis of the argument is the a priori estimate (33) from Lemma 6, which in the edge-based notation reads

$$\sum_{b} (\nabla\nabla G(b,e))^2 \le \lambda^{-2}. \qquad (60)$$

We now derive the following inequality:

$$\sum_{e\in S} \left(\frac{\partial}{\partial a(e)} \Big(\log \big(\sum_{e'\in S} (\nabla u(e'))^2 \big)^{\frac{1}{2}} \Big) \right)^2 \le \sup_{e\in S} \sum_{e'\in S} (\nabla\nabla G(e',e))^2. \qquad (61)$$

Indeed, by Leibniz's rule we obtain from (52) & (53) (that in our notation both take the form of $\frac{\partial}{\partial a(e)}\nabla u(e') = -\nabla\nabla G(e',e)\nabla u(e)$):

$$\frac{\partial}{\partial a(e)}\frac{1}{2}\sum_{e'\in S}(\nabla u(e'))^2 = -\nabla u(e)\sum_{e'\in S}\nabla\nabla G(e',e)\nabla u(e'),$$

and thus by Cauchy-Schwarz

$$\big(\frac{\partial}{\partial a(e)}\frac{1}{2}\sum_{e'\in S}(\nabla u(e'))^2\big)^2 \quad\le\quad (\nabla u(e))^2\sum_{e'\in S}(\nabla\nabla G(e',e))^2\sum_{e'\in S}(\nabla u(e'))^2,$$

which, using the chain rule, we rewrite as

$$\Big(\frac{\partial}{\partial a(e)}\big(\sum_{e'\in S}(\nabla u(e'))^2\big)^{\frac{1}{2}}\Big)^2 \le (\nabla u(e))^2\sum_{e'\in S}(\nabla\nabla G(e',e))^2.$$

We sum over $e\in S$ and relabel:

$$\begin{aligned}&\sum_{e\in S}\Big(\frac{\partial}{\partial a(e)}\big(\sum_{e'\in S}(\nabla u(e'))^2\big)^{\frac{1}{2}}\Big)^2\\ &\quad\le\ \sum_{e'\in S}(\nabla u(e'))^2\sup_{e\in S}\sum_{e'\in S}(\nabla\nabla G(e',e))^2;\end{aligned}$$

by the chain rule again, this can be reformulated as (61).

Inserting (60) into (61) we obtain

$$\left(\sum_{e\in S}\Big(\frac{\partial}{\partial a(e)}\log\big(\sum_{e'\in S}(\nabla u(e'))^2\big)^{\frac{1}{2}}\Big)^2\right)^{\frac{1}{2}} \le \lambda^{-1},$$

which we rewrite as

$$\left(\sum_{e\in S}\Big(\frac{\partial}{\partial a(e)}\log\big(\sum_{e'\in S}(\nabla u(e'))^2\big)\Big)^2\right)^{\frac{1}{2}} \le 2\lambda^{-1}, \tag{62}$$

We note that (62) amounts to a uniform bound of the Euclidean gradient of $X := \log\big(\sum_{e'\in S}(\nabla u(e'))^2\big)$ with respect to the variables $\{a(e)\}_{e\in S}\in[\lambda,1]^{|S|}$. Hence X is Lipschitz continuous on this convex set with Lipschitz constant $2\lambda^{-1}$. Since the set of $\{a(e)\}_{e\in S}$ in $[0,1]^{|S|}$ has diameter bounded by $|S|^{1/2}$, we obtain that the oscillation of X on this set is bounded by $2|S|^{1/2}\lambda^{-1}$:

$$\max_{\{a(e)\}_{e\in S}}\log\big(\sum_{e'\in S}(\nabla u(e'))^2\big) \le \min_{\{a(e)\}_{e\in S}}\log\big(\sum_{e'\in S}(\nabla u(e'))^2\big) + 2|S|^{\frac{1}{2}}\lambda^{-1},$$

which turns into (59).

Step 3. In this step, we rephrase Lemma 6, more precisely (27), in a way more suitable for its application in Step 4. More specifically, we claim that there exists a weight exponent $\alpha(d,\lambda) > 0$ such that

$$\sum_y |(|x-y|+1)^\alpha \nabla\nabla G(x,y)|^{2q} \le C(d,\lambda,q), \tag{63}$$

for all $q \ge 1$. In fact, we claim that

$$\alpha := \frac{1}{2}\alpha_0 \tag{64}$$

does the job. We start by noting that due to the symmetry of $\nabla\nabla G(x,y)$ under the interchange of x and y and to translational invariance (since we can replace a by some translation of it), it suffices to establish

$$\sum_y |(|y|+1)^\alpha \nabla\nabla G(y,0)|^{2q} \le C(d,\lambda,q). \tag{65}$$

Because of $q \ge 1$, and thus $\ell^2(\mathbb{Z}^d) \subset \ell^{2q}(\mathbb{Z}^d)$, we have

$$\sum_y |(|y|+1)^\alpha \nabla\nabla G(y,0)|^{2q} \le \Big(\sum_y |(|y|+1)^\alpha \nabla\nabla G(y,0)|^2\Big)^q.$$

Using a dyadic decomposition, we see

$$\begin{aligned}
&\sum_y |(|y|+1)^\alpha \nabla\nabla G(y,0)|^2 \\
&\quad = |\nabla\nabla G(0,0)|^2 + \sum_{n=0}^\infty \sum_{y:2^n\le|y|<2^{n+1}} |(|y|+1)^\alpha \nabla\nabla G(y,0)|^2 \\
&\quad \le |\nabla\nabla G(0,0)|^2 + \sum_{n=0}^\infty 2^{2\alpha(n+2)} \sum_{y:2^n\le|y|<2^{n+1}} |\nabla\nabla G(y,0)|^2.
\end{aligned}$$

We now may appeal to (27) to obtain

$$\begin{aligned}
&|\nabla\nabla G(0,0)|^2 + \sum_y |(|y|+1)^\alpha \nabla\nabla G(y,0)|^2 \\
&\quad \le C(d,\lambda)\Big(1+\sum_{n=0}^\infty 2^{2\alpha(n+2)}2^{-2\alpha_0 n}\Big) \overset{(64)}{\le} C(d,\lambda).
\end{aligned}$$

Step 4. In this step we establish the first statement of Lemma 5, namely (24). More precisely, we claim that for $p \geq \max\{\frac{d}{\alpha}, 1\}$ with α chosen in Step 3 and all $x \in \mathbb{Z}^d$ it holds

$$(|x|+1)^{2pd}\Big\langle\Big(\sum_y \sup_{a(y)} |\frac{\partial}{\partial a(y)} \nabla\nabla G(x,0)|^2\Big)^p\Big\rangle \\ \leq C(d,\lambda,p) \sup_y \Big\{(|y|+1)^{2pd}\langle|\nabla\nabla G(y,0)|^{2p}\rangle\Big\}, \tag{66}$$

where $C(d,\lambda,p)$ denotes a generic constant that only depends on d, λ, and p. Indeed, we first square (58) taken at $x' = 0$ and sum over y:

$$\sum_y \sup_{a(y)} \Big|\frac{\partial}{\partial a(y)} \nabla\nabla G(x,0)\Big|^2 \leq C(d,\lambda) \sum_y |\nabla\nabla G(x,y)|^2 |\nabla\nabla G(y,0)|^2.$$

After taking the p-th power, we split the sum into its contributions over $|y-x| \leq |y|$ and $|y-x| > |y|$ to obtain

$$\Big(\sum_y \sup_{a(y)} |\frac{\partial}{\partial a(y)} \nabla\nabla G(x,0)|^2\Big)^p \\ \leq C(d,\lambda,p)\bigg(\Big(\sum_{y:|y-x|\leq|y|} |\nabla\nabla G(x,y)|^2 |\nabla\nabla G(y,0)|^2\Big)^p \\ + \Big(\sum_{y:|y-x|\geq|y|} |\nabla\nabla G(x,y)|^2 |\nabla\nabla G(y,0)|^2\Big)^p\bigg). \tag{67}$$

We first bound the first term. To this end, smuggle in a weight $(|y-x|+1)^{2\alpha}$ with $\alpha = \alpha(d,\lambda)$ from Step 3 and apply Hölder's inequality with p and its dual exponent q such that $\frac{1}{p} + \frac{1}{q} = 1$:

$$\Big(\sum_{y:|y-x|\leq|y|} |\nabla\nabla G(x,y)|^2 |\nabla\nabla G(y,0)|^2\Big)^p \\ \leq \bigg(\sum_{y:|y-x|\leq|y|} \Big((|y-x|+1)^{\alpha}|\nabla\nabla G(x,y)|\Big)^{2q}\bigg)^{p-1} \\ \times \sum_{y:|y-x|\leq|y|} \Big((|y-x|+1)^{-\alpha}|\nabla\nabla G(y,0)|\Big)^{2p}.$$

The first term on the right hand side is bounded by Step 3, that is (63). After taking the expectation, we smuggle in another weight $(|y|+1)^{2pd}$ and take the supremum over appropriate terms to obtain

$$\Big\langle \sum_{y:|y-x|\le|y|} \Big((|y-x|+1)^{-\alpha}|\nabla\nabla G(y,0)|\Big)^{2p}\Big\rangle$$

$$\le \Big(\sum_{y:|y-x|\le|y|}(|y-x|+1)^{-2p\alpha}(|y|+1)^{-2pd}\Big)\sup_z\Big((|z|+1)^{2pd}\big\langle|\nabla\nabla G(z,0)|^{2p}\big\rangle\Big).$$

Since $|y-x|\le|y|$ implies $|y|\ge\frac{1}{2}|x|$, we find for the first r. h. s. factor that

$$\sum_{y:|y-x|\le|y|}(|y-x|+1)^{-2p\alpha}(|y|+1)^{-2pd}$$
$$\le\big(\tfrac{1}{2}|x|+1\big)^{-2pd}\sum_y(|y-x|+1)^{-2p\alpha}.$$

Since by assumption $2p\alpha\ge 2d>d$, we obtain for the last factor

$$\sum_y(|y|+1)^{-2p\alpha}\le C(d).$$

Combining the last three estimates yields the bound

$$\Big\langle\Big(\sum_{y:|y-x|\le|y|}|\nabla\nabla G(x,y)|^2|\nabla\nabla G(y,0)|^2\Big)^p\Big\rangle$$
$$\le\Big(C(d,\lambda,p)(|x|+1)^{-d}\sup_z\Big((|z|+1)^d\big\langle|\nabla\nabla G(z,0)|^{2p}\big\rangle^{\frac{1}{2p}}\Big)\Big)^{2p},$$

i.e. the expectation of the first term on the right hand side of (67) is bounded as desired. The second term in (67) can be dealt with exactly as the first term by simply exchanging the roles of x and 0.

Step 5. Like in Step 3, we rephrase Lemma 6, this time (28), in a way more suitable for its application in Step 6. We claim that for any integrability exponent $q\ge1$ and any weight exponent $\beta>0$ we have

$$\sup_a\sum_y\big|(|y|+1)^{-\beta}\nabla_y G(y,0)\big|^{2q}\le C(d,\lambda,q,\beta) \tag{68}$$

We note that by (28) we have as soon as $\beta>0$:

$$\sum_y|(|y|+1)^{-\beta}\nabla_y G(y,0)|^{2q}$$
$$\le\Big(\sum_y|(|y|+1)^{-\beta}\nabla_y G(y,0)|^2\Big)^q$$
$$\le\Big(|\nabla_y G(0,0)|^2+\sum_{n=0}^{\infty}2^{-q\beta n}\sum_{y:2^n\le|y|<2^{n+1}}|\nabla_y G(y,0)|^2\Big)^q\overset{(28)}{\le}C(d,\lambda,\beta). \tag{69}$$

Step 6. In this step we establish the second conclusion of Lemma 5, namely (26). More precisely, we show that for any integrability exponent $p < \infty$ at least as large as in Step 3 and for any weight exponent $\beta > 0$ such that

$$2p(\beta - d) + d < 0 \tag{70}$$

we have

$$\begin{aligned}
(|x|+1)^{d-1}\Big\langle\Big(\sum_y \sup_{a(y)} |\frac{\partial}{\partial a(y)}\nabla_x G(x,0)|^2\Big)^p\Big\rangle^{\frac{1}{2p}} \\
\leq C(d,\lambda,p,\beta)\Big(\sup_y\Big((|y|+1)^{d-1}\langle|\nabla_y G(y,0)|^{2p}\rangle^{\frac{1}{2p}}\Big) \\
+ (|x|+1)^{\beta-1+\frac{d}{2p}}\sup_z\Big((|z|+1)^d\langle|\nabla\nabla G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big)\Big),
\end{aligned} \tag{71}$$

for all $x \in \mathbb{Z}^d$, where $C(d,\lambda,p,\beta)$ denotes a generic constant that only depends on d, λ, p, and β. We note that by choosing β small and p large, the exponent $\beta - 1 + \frac{d}{2p}$ can be made to be non-positive (in fact, as close to -1 as we want), which proves (26). In order to establish (71), we first square (57) and sum over y:

$$\sum_y \sup_{a(y)} |\frac{\partial}{\partial a(y)}\nabla_x G(x,x')|^2 \leq C(d,\lambda)\sum_y |\nabla\nabla G(x,y)|^2|\nabla_y G(y,x')|^2.$$

We now specify to $x' = 0$ and split the sum over y:

$$\begin{aligned}
&\sum_y \sup_{a(y)} |\frac{\partial}{\partial a(y)}\nabla_x G(x,0)|^2 \\
&\leq C(d,\lambda)\Big(\sum_{y:|y|\geq\frac{1}{2}|x|} + \sum_{y:|y|<\frac{1}{2}|x|}\Big)|\nabla\nabla G(x,y)|^2|\nabla_y G(y,0)|^2 \\
&\leq C(d,\lambda)\Big(\sum_{y:|y|\geq\frac{1}{2}|x|}|\nabla\nabla G(x,y)|^2|\nabla_y G(y,0)|^2 \\
&\qquad + \sum_{y:|y-x|>\frac{1}{2}|x|}|\nabla_y G(y,0)|^2|\nabla\nabla G(x,y)|^2\Big).
\end{aligned} \tag{72}$$

We start by treating the first term on the r. h. s. of (72) in an analogous way to Step 4. For that purpose, let α be as in Step 3. We smuggle in the weight $(|x-y|+1)^\alpha$ and apply Hölder's inequality with p and q such that

$\frac{1}{p}+\frac{1}{q}=1$:

$$
\begin{aligned}
&\Big(\sum_{y:|y|\geq\frac{1}{2}|x|}|\nabla\nabla G(x,y)|^2|\nabla_y G(y,0)|^2\Big)^p\\
&\quad\leq\ \Big(\sum_y|(|x-y|+1)^{\alpha}\nabla\nabla G(x,y)|^{2q}\Big)^{p-1}\\
&\qquad\times\sum_{y:|y|\geq\frac{1}{2}|x|}|(|x-y|+1)^{-\alpha}\nabla_y G(y,0)|^{2p}.
\end{aligned}
$$

The first term was bounded by a constant $C(d,\lambda,p)$ in Step 3. Now we take the expectation $\langle\cdot\rangle$ w. r. t. a and then smuggle in a weight $(|y|+1)^{2p(d-1)}$ to obtain as desired:

$$
\begin{aligned}
&\Big\langle\Big(\sum_{y:|y|\geq\frac{1}{2}|x|}|\nabla\nabla G(x,y)|^2|\nabla_y G(y,0)|^2\Big)^p\Big\rangle\\
&\leq\Bigg(\sum_{y:|y|\geq\frac{1}{2}|x|}(|x-y|+1)^{-2p\alpha}(|y|+1)^{-2p(d-1)}\\
&\qquad\times\sup_z\Big((|z|+1)^{2p(d-1)}\langle|\nabla_z G(z,0)|^{2p}\rangle\Big)\Bigg)\\
&\overset{(63)}{\leq} C(d,\lambda,p)\big(|x|+1\big)^{-2p(d-1)}\sup_z\Big((|z|+1)^{2p(d-1)}\langle|\nabla_z G(z,0)|^{2p}\rangle\Big), \qquad (73)
\end{aligned}
$$

where we have used that $2p\alpha>d$.

We now address the second term on the r. h. s. of (72) in a similar way, just exchanging the roles of ∇G and $\nabla\nabla G$, of y and $y-x$, and of α and $-\beta$, where the weight exponent $\beta>0$ needs to satisfy (70). By Hölder's inequality we obtain:

$$
\begin{aligned}
&\Big(\sum_{y:|y-x|\geq\frac{1}{2}|x|}|\nabla_y G(y,0)|^2|\nabla\nabla G(x,y)|^2\Big)^p\\
&\quad\leq\ \Big(\sum_y|(|y|+1)^{-\beta}\nabla_y G(y,0)|^{2q}\Big)^{p-1}\\
&\qquad\times\sum_{y:|y-x|\geq\frac{1}{2}|x|}|(|y|+1)^{\beta}\nabla\nabla G(x,y)|^{2p}.
\end{aligned}
$$

The first term is bounded by Step 5 in form of (68). By stationarity in form of $\langle|\nabla\nabla G(x,y)|^{2p}\rangle=\langle|\nabla\nabla G(x-y,0)|^{2p}\rangle$, taking expectation and smuggling in a weight $(|y-x|+1)^{2pd}$ yields

$$\Big\langle\Big(\sum_{y:|y-x|\geq\frac{1}{2}|x|}|\nabla_y G(y,0)|^2|\nabla\nabla G(x,y)|^2\Big)^p\Big\rangle$$
$$\leq \sum_{y:|y-x|\geq\frac{1}{2}|x|}(|y|+1)^{2p\beta}(|y-x|+1)^{-2pd}\sup_z\big((|z|+1)^{2pd}\langle|\nabla\nabla G(z,0)|^{2p}\rangle\big).$$

We note that by the triangle inequality in form of $|y|\leq|x|+|y-x|$, in the range (70) the remaining sum is bounded as follows:

$$\begin{aligned}&\sum_{y:|y-x|\geq\frac{1}{2}|x|}(|y|+1)^{2p\beta}(|y-x|+1)^{-2pd}\\&\leq C(p,\beta)\Big((|x|+1)^{2p\beta}\sum_{y:|y-x|\geq\frac{1}{2}|x|}(|y-x|+1)^{-2pd}\\&\qquad+\sum_{y:|y-x|\geq\frac{1}{2}|x|}(|y-x|+1)^{2p(\beta-d)}\Big)\\&\leq C(d,p,\beta)(|x|+1)^{2p(\beta-d)+d}.\end{aligned}$$

Hence we have obtained

$$\Big\langle\Big(\sum_{y:|y-x|\geq\frac{1}{2}|x|}|\nabla_y G(y,0)|^2|\nabla\nabla G(x,y)|^2\Big)^p\Big\rangle$$
$$\leq C(d,\lambda,p,\beta)(|x|+1)^{2p(\beta-d)+d}\sup_z\big((|z|+1)^{2pd}\langle|\nabla\nabla G(z,0)|^{2p}\rangle\big). \quad (74)$$

In view of (72), the combination of (73) and (74) and taking the $2p$-th root yields (71).

Proof of Theorem 1.
We start with the proof of (3). To this purpose, we fix $x\in\mathbb{Z}^d$ and $p<\infty$; by Jensen's inequality, we may assume that $p\geq p_0$ with p_0 from Lemma 5. Applying Lemma 4 to $\zeta(a)=\nabla\nabla G(a;x,0)$ (component-wise) and inserting the estimate (24) of Lemma 5 yields (after redefining δ)

$$\begin{aligned}(|x|+1)^d\langle|\nabla\nabla G(x,0)|^{2p}\rangle^{\frac{1}{2p}}&\leq C(d,\lambda,\rho,p,\delta)(|x|+1)^d\langle|\nabla\nabla G(x,0)|\rangle\\&\quad+\delta\sup_z\Big((|z|+1)^d\langle|\nabla\nabla G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big).\end{aligned}$$

We now insert (17) and take the supremum over x:

$$\begin{aligned}&\sup_x\Big((|x|+1)^d\langle|\nabla\nabla G(x,0)|^{2p}\rangle^{\frac{1}{2p}}\Big)\\&\qquad\leq C(d,\lambda,\rho,p,\delta)+\delta\sup_z\Big((|z|+1)^d\langle|\nabla\nabla G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big).\end{aligned}$$

Choosing $\delta=\frac{1}{2}$, we obtain (3). We deal with the objection that $\sup_x\big((|x|+1)^d\langle|\nabla\nabla G(x,0)|^{2p}\rangle^{1/(2p)}\big)$ may be infinite by first working with the periodic Green's function G_L as in the proof of Lemma 6 and then letting $L\uparrow\infty$.

We now turn to the proof of (4). With help of the just established (3), we may upgrade the result of Lemma 5, cf. (26), to

$$
\begin{aligned}
&(|x|+1)^{d-1}\Big\langle\Big(\sum_y\sup_{a(y)}|\frac{\partial}{\partial a(y)}\nabla_x G(x,0)|^2\Big)^p\Big\rangle^{\frac{1}{2p}}\\
&\qquad\le C(d,\lambda,\rho,p)\bigg(\sup_z\Big((|z|+1)^{d-1}\langle|\nabla_z G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big)+1\bigg). \qquad (75)
\end{aligned}
$$

Applying Lemma 4 to $\zeta=\nabla_x G(x,0)$ (component-wise) and inserting (75) yields (after redefining δ)

$$
\begin{aligned}
&(|x|+1)^{d-1}\langle|\nabla_x G(x,0)|^{2p}\rangle^{\frac{1}{2p}}\\
&\le C(d,\lambda,\rho,p,\delta)(|x|+1)^{d-1}\langle|\nabla_x G(x,0)|\rangle\\
&\qquad+\delta\bigg(\sup_z\Big((|z|+1)^{d-1}\langle|\nabla_z G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big)+1\bigg).
\end{aligned}
$$

We now insert (18) and take the supremum over x:

$$
\begin{aligned}
&\sup_x\Big((|x|+1)^{d-1}\langle|\nabla_x G(x,0)|^{2p}\rangle^{\frac{1}{2p}}\Big)\\
&\qquad\le C(d,\lambda,\rho,p,\delta)+\delta\sup_z\Big((|z|+1)^{d-1}\langle|\nabla_z G(z,0)|^{2p}\rangle^{\frac{1}{2p}}\Big).
\end{aligned}
$$

Proof of Corollary 1.
By stationarity, it suffices to prove the result for $y=0$. It is well known that an LSI implies a corresponding SG, see for instance [15, Theorem 4.9]. Indeed, using $\zeta^2=1+\epsilon f$ for some $f(a)$ in (2) and expanding to second order in $\epsilon\ll 1$ one obtains

$$
\langle(f-\langle f\rangle)^2\rangle\le\frac{1}{\rho}\Big\langle\sum_y\sup_{a(y)}\big(\frac{\partial f}{\partial a(y)}\big)^2\Big\rangle.
$$

As in Step 2 of the proof of Lemma 4, see also [13, Lemma 11], it follows that

$$
\langle|f-\langle f\rangle|^{2p}\rangle\le C(\rho,p)\Big\langle\Big(\sum_y\sup_{a(y)}\big(\frac{\partial f}{\partial a(y)}\big)^2\Big)^p\Big\rangle. \qquad (76)
$$

We fix $x \in \mathbb{Z}^d$ and apply this inequality to $f(a) = G(a; x, 0)$ and use (56) from the proof of Lemma 5

$$\sup_{a(y)} \Big| \frac{\partial G(x,0)}{\partial a(y)} \Big| \leq C(d,\lambda) |\nabla_y G(x,y)| |\nabla_y G(y,0)|$$

to obtain

$$\Big\langle \big| G(x,0) - \langle G(x,0) \rangle \big|^{2p} \Big\rangle^{\frac{1}{p}} \leq C(d,\lambda,\rho,p) \Big\langle \Big(\sum_y |\nabla_y G(x,y)|^2 |\nabla_y G(y,0)|^2 \Big)^p \Big\rangle^{\frac{1}{p}}.$$

The triangle inequality yields

$$\Big\langle \big| G(x,0) - \langle G(x,0) \rangle \big|^{2p} \Big\rangle^{\frac{1}{p}} \leq C(d,\lambda,\rho,p) \sum_y \Big\langle |\nabla_y G(x,y)|^{2p} |\nabla_y G(y,0)|^{2p} \Big\rangle^{\frac{1}{p}}.$$

Using Cauchy Schwarz's inequality in $\langle \cdot \rangle$ and appealing to stationarity, we obtain

$$\begin{aligned} &\Big\langle \big| G(x,0) - \langle G(x,0) \rangle \big|^{2p} \Big\rangle^{\frac{1}{p}} \\ &\leq C(d,\lambda,\rho,p) \sum_y \langle |\nabla_y G(x,y)|^{4p} \rangle^{\frac{1}{2p}} \langle |\nabla_y G(y,0)|^{4p} \rangle^{\frac{1}{2p}} \\ &= C(d,\lambda,\rho,p) \sum_y \langle |\nabla_1 G(x-y,0)|^{4p} \rangle^{\frac{1}{2p}} \langle |\nabla_1 G(y,0)|^{4p} \rangle^{\frac{1}{2p}}, \end{aligned}$$

where ∇_1 denotes the derivative w. r. t. the first variable. Into this estimate, we insert the result of Theorem 1:

$$\Big\langle \big| G(x,0) - \langle G(x,0) \rangle \big|^{2p} \Big\rangle^{\frac{1}{p}} \leq C(d,\lambda,\rho,p) \sum_y (|x-y|+1)^{2(1-d)} (|y|+1)^{2(1-d)}. \tag{77}$$

We now turn to the sum on the r. h. s. of (77): By symmetry, we have

$$\sum_y (|x-y|+1)^{2(1-d)} (|y|+1)^{2(1-d)} \leq 2 \sum_{y : |x-y| \leq |y|} (|x-y|+1)^{2(1-d)} (|y|+1)^{2(1-d)}. \tag{78}$$

We note that in the case of $d > 2$ we have $2(1-d) < -d$ so that

$$\sum_z (|z|+1)^{2(1-d)} \leq C(d) < \infty. \tag{79}$$

Since $|x-y| \le |y|$ implies $|y| \ge \frac{1}{2}|x|$ we thus have as desired for (78)

$$\begin{aligned}
&\sum_{y:|x-y|\le|y|} (|x-y|+1)^{2(1-d)}(|y|+1)^{2(1-d)} \\
&\le \quad (\frac{1}{2}|x|+1)^{2(1-d)} \sum_y (|x-y|+1)^{2(1-d)} \\
&\overset{(79)}{\le} \quad C(d)(|x|+1)^{2(1-d)}. \qquad (80)
\end{aligned}$$

We now turn to the case of $d=2$. In this case, we split the sum on the r. h. s. of (78) according to

$$\begin{aligned}
\sum_{y:|x-y|\le|y|} &= \sum_{y:|x-y|\le|y| \text{ and } |y|\ge 2|x|} + \sum_{y:|x-y|\le|y| \text{ and } |y|<2|x|} \\
&\le \sum_{y:|x-y|\ge\frac{1}{2}|y| \text{ and } |y|\ge 2|x|} + \sum_{y:|x-y|\le 2|x| \text{ and } |y|\ge\frac{1}{2}|x|},
\end{aligned}$$

so that

$$\begin{aligned}
&\sum_{y:|x-y|\le|y|} (|x-y|+1)^{-2}(|y|+1)^{-2} \\
&\le \sum_{y:|y|\ge 2|x|} (\frac{1}{2}|y|+1)^{-4} + (\frac{1}{2}|x|+1)^{-2} \sum_{z:|z|\le 2|x|} (|z|+1)^{-2} \\
&\le C(|x|+1)^{-2} + C(|x|+1)^{-2}\log(|x|+2). \qquad (81)
\end{aligned}$$

Combining (80) and (81), we gather

$$\begin{aligned}
&\sum_y (|x-y|+1)^{2(1-d)}(|y|+1)^{2(1-d)} \\
&\le C(d)(|x|+1)^{2(1-d)} \left\{ \begin{array}{ccc} 1 & \text{for} & d>2 \\ \log(|x|+2) & \text{for} & d=2 \end{array} \right\},
\end{aligned}$$

which we insert into (77) to obtain (5).

Optimality of Corollary 1 for $p=1$.
In this section we will show by formal calculations that Corollary 1 is optimal by considering the regime $1-\lambda \ll 1$. Recall that the Green's function satisfies $\nabla_x^* a(x)\nabla_x G(x,x') = \delta(x-x')$. Recall that in edge-based notation introduced in Step 1 of the proof of Lemma 5, the components of the (diagonal) matrix $a(x)$ on sites can be identified with scalars on edges via $a(e) = a_{ii}(x)$ for the edge $e = x + e_i$. Now let $a(e) = 1 + \epsilon b(e)$, where b is i. i. d. with

values at each edge taken in $[-1,1]$. Of course in site-based notation, $b(x)$ is given by a diagonal matrix with diagonal elements $b(x+e_i)$, $i=1,\ldots,d$. Furthermore we assume $\langle b(e)\rangle = 0$ as well as $\langle b(e)^2\rangle = 1$. Note that this implies $a \in [1-\epsilon, 1+\epsilon] \subset [1/2, 3/2]$ (w. l. o. g. $\epsilon < 1/2$), but (by linearity of the equation in a) all results remain true with this new upper bound on a. Let us expand the Green's function corresponding to a in powers of ϵ:

$$G(x,y) = G_0(x,y) + \epsilon G_1(x,y) + \ldots\ .$$

Substituting into the defining equation for G, we find that to zeroth order in ϵ, it holds

$$\nabla_x^* \nabla_x G_0(x,y) = \delta(x,y),$$

i.e. G_0 is the constant-coefficient Green's function. Then to first order, it follows

$$\nabla_x^* \nabla_x G_1(x,y) + \nabla_x^* b(x) \nabla_x G_0(x,y) = 0.$$

Hence it holds that

$$\begin{aligned} G_1(x,y) &= -\sum_{z\in\mathbb{Z}^d} G_0(x,z)\nabla_z^* b(z)\nabla_z G_0(z,y) \\ &= -\sum_e \nabla_2 G_0(x,e) b(e) \nabla_1 G_0(e,y), \end{aligned}$$

where the last sum is taken over all edges e. Since $\langle b(e)\rangle = 0$, we deduce $\langle G_1\rangle = 0$ and consequently

$$\begin{aligned} \langle G_1(x,0)^2\rangle &= \langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle \\ &= \sum_{e,e'} \nabla_2 G_0(x,e)\nabla_2 G_0(x,e')\langle b(e)b(e')\rangle \nabla_1 G_0(e,0)\nabla_1 G_0(e',0). \end{aligned}$$

Since the coefficients $b(x)$ are i. i. d. with variance 1, it follows

$$\langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle = \sum_e (\nabla_2 G_0(x,e))^2 (\nabla_1 G_0(e,0))^2.$$

The behavior of the constant-coefficient Green's function is well-known, see for instance [16, Theorem 4.3.1], and yields that at the edge $e = z + e_i$, the term $(\nabla_1 G_0(e,0))^2$ behaves as $(|z|+1)^{1-d}$ with a similar expression for $(\nabla_2 G_0(x,e))^2$. Hence we find that

$$\begin{aligned} \langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle &\le C(d) \sum_{z\in\mathbb{Z}^d} \big((|x-z|+1)(|z|+1)\big)^{2(1-d)} \quad \text{and} \\ \langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle &\ge \frac{1}{C(d)} \sum_{z\in\mathbb{Z}^d} \big((|x-z|+1)(|z|+1)\big)^{2(1-d)}. \qquad (82) \end{aligned}$$

Thus (80), (81), and (82) yield the upper bound

$$\langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle \leq \frac{1}{C(d)}(|x|+1)^{2(1-d)} \begin{Bmatrix} 1 & \text{for } d > 2 \\ \log(|x|+2) & \text{for } d = 2 \end{Bmatrix}.$$

If $d > 2$, a lower bound can be obtained by considering only the summand $z = 0$ in (82). If $d = 2$, we restrict the sum to all z such that $|z| \leq |x|$ to obtain

$$\begin{aligned} \langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle &\geq \frac{1}{C(d)} \sum_{z:|z|\leq|x|} (|x-z|+1)^{-2}(|z|+1)^{-2} \\ &\geq \frac{1}{C(d)}(|x|+1)^{-2} \sum_{z:|z|\leq|x|} (|z|+1)^{-2} \\ &\geq \frac{1}{C(d)}(|x|+1)^{-2}\log(2+|x|). \end{aligned}$$

Thus Corollary 1 is indeed optimal in scaling.

Proof of Corollary 2

Step 1. Proof in dimension $d > 2$. First of all, the triangle inequality in $\langle (\cdot)^r\rangle^{1/r}$ yields

$$\left\langle \left(\sum_x \left|u(x) - \langle u(x)\rangle\right|^p \right)^r \right\rangle^{\frac{1}{rp}} \leq \left(\sum_x \left\langle \left|u(x) - \langle u(x)\rangle\right|^{rp} \right\rangle^{\frac{1}{r}} \right)^{\frac{1}{p}}. \quad (83)$$

Since u is the decaying solution of (6) with compactly supported right hand side f, it can be represented via the Green's function:

$$u(x) = \sum_y G(x,y)f(y), \quad (84)$$

Consequently, an application of the triangle inequality in $\langle (\cdot)^{rp}\rangle^{1/(rp)}$ yields

$$\left\langle \left|u(x) - \langle u(x)\rangle\right|^{rp} \right\rangle^{\frac{1}{rp}} \leq \sum_y \left\langle \left|G(x,y) - \langle G(x,y)\rangle\right|^{rp} \right\rangle^{\frac{1}{rp}} |f(y)|,$$

so that we may use Corollary 1 to the effect of

$$\left\langle \left|u(x) - \langle u(x)\rangle\right|^{rp} \right\rangle^{\frac{1}{rp}} \leq C(d,\lambda,\rho,r,p) \sum_y (|x-y|+1)^{1-d}|f(y)|. \quad (85)$$

We now insert (85) in (83) to obtain

$$\left\langle \left(\sum_x \left| u(x) - \langle u(x) \rangle \right|^p \right)^r \right\rangle^{\frac{1}{rp}}$$
$$\leq C(d, \lambda, \rho, r, p) \left(\sum_x \left(\sum_y (|x-y|+1)^{1-d} |f(y)| \right)^p \right)^{\frac{1}{p}}. \quad (86)$$

Now let us recall the Hardy-Littlewood-Sobolev inequality in $\mathbb{R}^d$, see [17, Section 4.3] for a proof:

$$\left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |x-y|^{-\alpha} f(y)\, dy \right)^p \right)^{\frac{1}{p}} \leq C(d, \alpha, p) \left(\int_{\mathbb{R}^d} |f(y)|^q\, dy \right)^{\frac{1}{q}}$$

for all weight exponents $0 < \alpha < d$ and for all integrability exponents $1 < p, q < \infty$ related by $1 + \frac{1}{p} = \frac{\alpha}{d} + \frac{1}{q}$. A discrete version can easily be obtained by applying the continuum version to piecewise linear functions on a triangulation subordinate to the lattice $\mathbb{Z}^d$. We use the discrete version for $\alpha = d - 1$, that is,

$$\left(\sum_x \left(\sum_y (|x-y|+1)^{1-d} |f(y)| \right)^p \right)^{\frac{1}{p}} \leq C(d, p) \left(\sum_y |f(y)|^q \right)^{\frac{1}{q}}, \quad (87)$$

in which case the relation turns as desired into $\frac{1}{p} + \frac{1}{d} = \frac{1}{q}$. Our assumption $p \geq 2$ and $d > 2$ ensure that q is indeed admissible for Hardy-Littlewood-Sobolev in the sense of the strict inequality $q > 1$.

Step 2. Changes if $d = 2$. In this case, using that $f(y)$ is supported in $\{|y| \leq R\}$, (86) assumes the form

$$\left\langle \left(\sum_{x:|x|\leq R} \left| u(x) - \langle u(x) \rangle \right|^p \right)^r \right\rangle^{\frac{1}{rp}}$$
$$\leq \; C(d, \lambda, \rho, r, p) \left(\sum_{x:|x|\leq R} \left(\sum_{y:|y|\leq R} (|x-y|+1)^{1-d} (\log^{\frac{1}{2}} |x-y|) |f(y)| \right)^p \right)^{\frac{1}{p}}$$
$$\leq \; C(d, \lambda, \rho, r, p)(\log^{\frac{1}{2}} R) \left(\sum_x \left(\sum_y (|x-y|+1)^{1-d} |f(y)| \right)^p \right)^{\frac{1}{p}}.$$

As in Step 1, it remains to apply the discrete Hardy-Littlewood-Sobolev inequality, where we note that our assumption $p > 2$ now ensures $q > 1$ even for $d = 2$.

Proof of Corollary 3

Step 1. In this step, we derive the estimate

$$\Big\langle\Big|\sum_x(u(x)-\langle u(x)\rangle)g(x)\Big|^r\Big\rangle^{\frac{1}{r}}$$
$$\leq C(\rho,r)\Big\langle\Big(\sum_z\sup_{a(z)}\Big(\sum_x\sum_y\frac{\partial G(x,y)}{\partial a(z)}f(y)g(x)\Big)^2\Big)^{\frac{r}{2}}\Big\rangle^{\frac{1}{r}}. \quad (88)$$

Indeed, it follows from the representation (84) that

$$\Big\langle\Big|\sum_x\big(u(x)-\langle u(x)\rangle\big)g(x)\Big|^r\Big\rangle^{\frac{1}{r}}$$
$$=\Big\langle\Big|\sum_x\sum_y\big(G(x,y)-\langle G(x,y)\rangle\big)f(y)g(x)\Big|^r\Big\rangle^{\frac{1}{r}}.$$

Since the only dependence on the coefficients a is through G, the L^p-version of SG (76), with $2p$ replaced by r, yields (88).

Step 2. In this step, we estimate the right hand side of (88) as follows:

$$\Big\langle\Big(\sum_z\sup_{a(z)}\Big(\sum_x\sum_y\frac{\partial G(x,y)}{\partial a(z)}f(y)g(x)\Big)^2\Big)^{\frac{r}{2}}\Big\rangle^{\frac{1}{r}}$$
$$\leq C(d,\lambda,\rho,r)\left(\sum_z\Big(\sum_x(|x-z|+1)^{1-d}|g(x)|\Big)^2\Big(\sum_y(|y-z|+1)^{1-d}|f(y)|\Big)^2\right)^{\frac{1}{2}}. \quad (89)$$

Indeed, expanding the square (of a vector) on the l. h. s. of (89) and inserting (56) yields

$$\sup_{a(z)}\Big(\sum_x\sum_y\frac{\partial G(x,y)}{\partial a(z)}f(y)g(x)\Big)^2$$
$$=\sup_{a(z)}\sum_{x,x',y,y'}\frac{\partial G(x,y)}{\partial a(z)}\cdot\frac{\partial G(x',y')}{\partial a(z)}g(x)g(x')f(y)f(y')$$
$$\leq C(d,\lambda)\sum_{x,x',y,y'}|\nabla_z G(x,z)||\nabla_z G(x',z)||\nabla_z G(z,y)||\nabla_z G(z,y')|$$
$$\times|g(x)g(x')f(y)f(y')|.$$

Consequently we obtain by the triangle inequality w. r. t. $\langle\sum_{z,x,x',y,y'}|\cdot|^{r/2}\rangle^{2/r}$ (w. l. o. g. we may assume $r\geq 2$)

$$\Big\langle\Big(\sum_z\Big(\sum_x\sum_y\frac{\partial G(x,y)}{\partial a(z)}f(y)g(x)\Big)^2\Big)^{\frac{r}{2}}\Big\rangle^{\frac{1}{r}}$$
$$\leq C(d,\lambda)\Bigg(\sum_{z,x,x',y,y'}\langle|\nabla_z G(x,z)|^{\frac{r}{2}}|\nabla_z G(x',z)|^{\frac{r}{2}}|\nabla_z G(z,y)|^{\frac{r}{2}}|\nabla_z G(z,y')|^{\frac{r}{2}}\rangle^{\frac{1}{r}}$$
$$\times|g(x)||g(x')||f(y)||f(y')|\Bigg)^{\frac12}.$$

Hölder's inequality with respect to the ensemble $\langle\cdot\rangle$ and Theorem 1 yield

$$\langle|\nabla_z G(x,z)|^{\frac{r}{2}}|\nabla_z G(x',z)|^{\frac{r}{2}}|\nabla_z G(z,y)|^{\frac{r}{2}}|\nabla_z G(z,y')|^{\frac{r}{2}}\rangle^{\frac{2}{r}}$$
$$\leq C(d,\lambda,\rho,r)\Big((|x-z|+1)(|x'-z|+1)(|y-z|+1)(|y'-z|+1)\Big)^{1-d}.$$

Hence (89) follows from partly undoing the expansion of the square:

$$\Bigg(\sum_{z,x,x',y,y'}(|x-z|+1)^{1-d}(|x'-z|+1)^{1-d}(|y-z|+1)^{1-d}$$
$$\times(|y'-z|+1)^{1-d}|g(x)||g(x')||f(y)||f(y')|\Bigg)^{\frac12}$$
$$=\Bigg(\sum_z\Big(\sum_x(|x-z|+1)^{1-d}|g(x)|\Big)^2\Big(\sum_y(|y-z|+1)^{1-d}|f(y)|\Big)^2\Bigg)^{\frac12}.$$

Step 3. Conclusion. An application of Hölder's inequality w. r. t. the sum over z on the r. h. s. of (89) yields a bound by

$$\Bigg(\sum_z\Big(\sum_x(|x-z|+1)^{1-d}|g(x)|\Big)^{\tilde p}\Bigg)^{\frac{1}{\tilde p}}$$
$$\times\Bigg(\sum_z\Big(\sum_y(|y-z|+1)^{1-d}|f(y)|\Big)^{p}\Bigg)^{\frac{1}{p}},\quad(90)$$

with $\tilde p$ and p such that $\frac{2}{p}+\frac{2}{\tilde p}=1$ to be chosen later. We recall the Hardy-Littlewood-Sobolev inequality (87), i.e.

$$\Bigg(\sum_z\Big(\sum_x(|x-z|+1)^{1-d}|f(x)|\Big)^{p}\Bigg)^{\frac{1}{p}}\leq C(d,q)\Big(\sum_x|f(x)|^q\Big)^{\frac1q},$$

if we choose p such that $\frac{1}{q} = \frac{1}{d} + \frac{1}{p}$. The Hardy-Littlewood-Sobolev inequality likewise yields

$$\left(\sum_z \left(\sum_x (|x-z|+1)^{1-d} |g(x)| \right)^{\tilde{p}} \right)^{\frac{1}{\tilde{p}}} \leq C(d,q) \left(\sum_x |g(x)|^{\tilde{q}} \right)^{\frac{1}{\tilde{q}}},$$

where $\frac{1}{\tilde{q}} = \frac{1}{d} + \frac{1}{\tilde{p}} = \frac{1}{d} + \frac{1}{2} - \frac{1}{p}$. Inserting these estimates into (90) and then into Steps 2 and 1 yields Corollary 3.

Proof of Corollary 4.

Step 1. We claim that for any function $\eta(x)$ supported in $\{|x| < R\}$, we obtain the representation formula

$$(\eta u)(x) = \sum_{y:|y|\leq R} \big(u(y) \nabla_y G(y,x) \cdot a(y) \nabla \eta(y) - G(y,x) \nabla \eta(y) \cdot a(y) \nabla u(y) \big). \tag{91}$$

We start by noting that even on the discrete level, some aspects of Leibniz' rule survive, such as

$$\nabla \zeta \cdot a \nabla (\eta u) - \nabla (\eta \zeta) \cdot a \nabla u = u \nabla \zeta \cdot a \nabla \eta - \zeta \nabla \eta \cdot a \nabla u \tag{92}$$

for any function ζ. Indeed, since a is diagonal, (92) reduces to

$$\nabla_i \zeta \nabla_i (\eta u) - \nabla_i (\eta \zeta) \nabla_i u = u \nabla_i \zeta \nabla_i \eta - \zeta \nabla_i \eta \nabla_i u,$$

which in turn reduces to the elementary identity

$$(\tilde{\zeta} - \zeta)(\tilde{\eta}\tilde{u} - \eta u) - (\tilde{\eta}\tilde{\zeta} - \eta\zeta)(\tilde{u} - u) = u(\tilde{\zeta} - \zeta)(\tilde{\eta} - \eta) - \zeta(\tilde{\eta} - \eta)(\tilde{u} - u).$$

We integrate (92):

$$\sum_y \nabla \zeta \cdot a \nabla (\eta u) - \sum_y \nabla (\eta \zeta) \cdot a \nabla u = \sum_y \big(u \nabla \zeta \cdot a \nabla \eta - \zeta \nabla \eta \cdot a \nabla u \big) \tag{93}$$

and use it for $\zeta = G(\cdot, x)$. By definition of G, the first term on the l. h. s. of (93) yields $(\eta u)(x)$; since $\eta G(\cdot, x)$ is supported in $\{|y| \leq R\}$, the second term on the l. h. s. of (93) vanishes. This yields (91).

Step 2. We now use the representation obtained in Step 1 to obtain bounds on the gradient of u and consequently on the α-Hölder norm of u. Specifically, we claim that

$$\left(\frac{\sup_{x:|x|\le\frac{R}{8}}\frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha}\sup_{x:|x|\le R}|u(x)|}\right)^p$$
$$\le C(d,\lambda,p)R^{\alpha p}R^{-p}\Big(R^{d(p-1)}\sum_{x:|x|\le\frac{R}{8}}\ \sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}|\nabla\nabla G(y,x)|^p$$
$$+R^{d(p-1)-p}\sum_{x:|x|\le\frac{R}{8}}\ \sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}|\nabla_x G(y,x)|^p\Big),\quad (94)$$

where α and p are related by $\alpha p=p-d$. To this end, we choose a cut-off function η for $\{|x|\le\frac{R}{4}+1\}$ in $\{|x|\le\frac{R}{2}-1\}$ (w. l. o. g. $R>8$). If we restrict to $|x|\le\frac{R}{4}$ and take the derivative in x, we obtain

$$\nabla u(x)=\sum_y\big(u(y)\nabla\nabla G(y,x)a(y)\nabla\eta(y)-(\nabla\eta(y)\cdot a(y)\nabla u(y))\nabla_x G(y,x)\big),$$

which implies

$$|\nabla u(x)|\le C(d)R^{-1}\sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}\big(|u(y)||\nabla\nabla G(y,x)|+|\nabla_x G(y,x)||\nabla u(y)|\big).\quad (95)$$

Summing the p-th power of (95) and applying Hölder's inequality, we obtain

$$\sum_{x:|x|\le\frac{R}{8}}|\nabla u(x)|^p$$
$$\le C(d,p)R^{-p}\bigg(\Big(\sum_{y:|y|\le\frac{R}{2}}|u(y)|^q\Big)^{p-1}\sum_{x:|x|\le\frac{R}{8}}\ \sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}|\nabla\nabla G(y,x)|^p$$
$$+\Big(\sum_{y:|y|\le\frac{R}{2}}|\nabla u(y)|^q\Big)^{p-1}\sum_{x:|x|\le\frac{R}{8}}\ \sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}|\nabla_x G(y,x)|^p\bigg),\quad (96)$$

where q is the dual Hölder exponent of p. Now we apply the following (discrete) Sobolev inequality: If $\alpha<1$ and $p>d$ are related by

$$\alpha=1-\frac{d}{p},\quad (97)$$

then it holds that

$$\sup_{x:|x|\le\frac{R}{8}}\frac{|u(x)-u(0)|}{|x|^\alpha}\le C(d,p)\Big(\sum_{x:|x|\le\frac{R}{8}}|\nabla u(x)|^p\Big)^{\frac{1}{p}}.$$

(This discrete version can easily be derived from its continuum version by extending u to a piecewise linear function on a triangulation subordinate to the lattice.) Therefore the left-hand side of (96) bounds the α-Hölder norm as desired, albeit over a smaller ball.

Let us now turn to the right-hand side of (96). It holds that

$$\Big(\sum_{y:|y|\le\frac{R}{2}}|u(y)|^q\Big)^{p-1}\le C(d,p)R^{d(p-1)}\Big(\sup_{x:|x|\le\frac{R}{2}}|u(y)|\Big)^p. \tag{98}$$

To estimate the second summand on the right-hand side, we note that Caccioppoli's estimate (47) implies

$$\sum_{y:|y|\le\frac{R}{2}}|\nabla u(y)|^2\le C(d,\lambda)R^{-2}\sum_{y:|y|\le R}|u(y)|^2\le C(d,\lambda)R^{d-2}\Big(\sup_{x:|x|\le R}|u(x)|\Big)^2.$$

Together with Jensen's inequality (here we need $p\ge 2$, which is obvious since even $p>d$), we obtain that

$$\begin{aligned}\Big(\sum_{y:|y|\le\frac{R}{2}}|\nabla u(y)|^q\Big)^{p-1}&\le C(d,p)R^{d(\frac{p}{2}-1)}\Big(\sum_{y:|y|\le\frac{R}{2}}|\nabla u(y)|^2\Big)^{\frac{p}{2}}\\&\le C(d,\lambda,p)R^{d(p-1)-p}\Big(\sup_{x:|x|\le R}|u(x)|\Big)^p.\end{aligned} \tag{99}$$

Substituting (98) and (99) into (96) yields the claim of this step.

Step 3. We end by bounding the Green's function to arrive at the conclusion

$$\Big\langle\Big(\sup_u\frac{\sup_{x:|x|\le R}\frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha}\sup_{x:|x|\le R}|u(x)|}\Big)^p\Big\rangle\le C(d,\lambda,\rho,p,\alpha) \tag{100}$$

for all $\alpha<1$, $p<\infty$, and $R<\infty$, where the outer supremum is taken over all solutions $u(x)$ to $\nabla^*a\nabla u=0$ in $\{x:|x|\le R\}$. Indeed, Theorem 1 applied to the result (94) of Step 2 yields

$$\begin{aligned}&\Big\langle\Big(\sup_u\frac{\sup_{x:|x|\le\frac{R}{8}}\frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha}\sup_{x:|x|\le R}|u(x)|}\Big)^p\Big\rangle\\&\quad\le C(d,\lambda,\rho,p)R^{\alpha p}\Big(R^{d(p-1)-p}\sum_{x:|x|\le\frac{R}{8}}\sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}(|x-y|+1)^{-pd}\\&\qquad\qquad+R^{d(p-1)-2p}\sum_{x:|x|\le\frac{R}{8}}\sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}}(|x-y|+1)^{p(1-d)}\Big)\end{aligned}$$

if α and p are related by (97). Here we have used symmetry and stationarity to replace $\langle|\nabla_x G(y,x)|^p\rangle$ by $\langle|\nabla_1 G(x-y,0)|^p\rangle$ and thus bring the terms involving G into the form of Theorem 1. In the domains of x and y, it holds $|x-y|+1 \geq |y|/2 \geq R/8$. Therefore the first double-sum on the right-hand side is bounded by

$$C(d,p)R^{2d-pd}.$$

Likewise the second double-sum is bounded by

$$C(d,p)R^{2d+p(1-d)}.$$

If (97) holds, we conclude that

$$\left\langle \left(\sup_u \frac{\sup_{x:|x|\leq \frac{R}{8}} \frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha}\sup_{x:|x|\leq R}|u(x)|} \right)^p \right\rangle \leq C(d,\lambda,\rho,p).$$

In the region $\{x : \frac{R}{8} \leq |x| \leq R\}$, it obviously holds

$$\frac{|u(x)-u(0)|}{|x|^\alpha} \leq 2\frac{8^\alpha}{R^\alpha}\sup_{x:|x|\leq R}|u(x)|.$$

Thus we have obtained (100) for p and α such that (97) holds. Since in (97), $\alpha \to 1$ as $p \to \infty$ and since we can always decrease p and α in the conclusion (100) (in p this follows from Jensen's inequality), the estimate (100) indeed holds for arbitrary $p < \infty$ and $\alpha < 1$.

Proof of Corollary 5.

Step 1. First we write Corollary 4 in a point-wise (in a) form. The corollary implies that for ($\langle\cdot\rangle$-almost) every a, $R > 0$, $\alpha < 1$, and all solutions u to $\nabla^* a \nabla u = 0$, it holds that

$$\sup_{x:|x|\leq R} \frac{|u(x)-u(0)|}{|x|^\alpha} \leq K(a,R)\frac{1}{R^\alpha}\sup_{x:|x|\leq R}|u(x)|$$

with a constant $K(a,R)$ such that $\langle K(\cdot,R)^p\rangle \leq C(d,\lambda,\alpha,p)$ for all $p < \infty$.

Step 2. We argue that for every $R_0 > 0$, there exists a sequence of increasing radii $(R_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} R_n = \infty$ such that

$$\frac{1}{N}\sum_{n=0}^{N} K(a,R_n) \leq C(a) < \infty.$$

Indeed, Step 1 yields that the set of random variables $(K(a,R))_{R>1}$ is bounded in $L^2(\langle\cdot\rangle)$. By the Banach-Saks theorem [24, Chapter II.38], we obtain a subsequence $(K(a,R_n))_{n\in\mathbb{N}}$ with increasing, unbounded radii R_n such that $N^{-1}\sum_{n=0}^{N} K(a,R_n)$ converges strongly in $L^2(\langle\cdot\rangle)$. Thus we have obtained a subsequence with the desired property.

Step 3. In this step, we deduce the estimate

$$\sup_{|x|=R_0}|u(x)-u(0)| \le C(a)\sup_{n>M}\Big\{\frac{1}{R_n^\alpha}\sup_{|x|\le R_n}|u(x)|\Big\} + \frac{1}{N}\sum_{n=1}^{M}K(a,R_n)\frac{1}{R_n^\alpha}\sup_{|x|\le R_n}|u(x)| \quad (101)$$

for all $N>M>0$. Indeed, since the radii $R_n \ge R_0$ are increasing, it holds

$$\sup_{|x|=R_0}|u(x)-u(0)| \le \frac{1}{N}\sum_{n=0}^{N}\sup_{|x|\le R_n}\frac{|u(x)-u(0)|}{|x|^\alpha}.$$

Inserting Step 1 and splitting the sum yields

$$\sup_{|x|=R_0}|u(x)-u(0)| \le \frac{1}{N}\sum_{n=M+1}^{N}K(a,R_n)\frac{1}{R_n^\alpha}\sup_{|x|\le R_n}|u(x)| + \frac{1}{N}\sum_{n=1}^{M}K(a,R_n)\frac{1}{R_n^\alpha}\sup_{|x|\le R_n}|u(x)|.$$

Upon noting the positivity of $K(a,R_n)$ and the monotonicity in n of the suprema, Step 2 yields estimate (101).

Step 4. We conclude by taking the limit as first $N\to\infty$ and then $M\to\infty$. First, letting $N\to\infty$ in (101) from Step 3 yields

$$\sup_{|x|=R_0}|u(x)-u(0)| \le C(a)\sup_{n>M}\Big\{\frac{1}{R_n^\alpha}\sup_{|x|\le R_n}|u(x)|\Big\}.$$

Since $R_n^{-\alpha}\sup_{|x|\le R_n}|u(x)|$ vanishes in the limit as $n\to\infty$ by our boundedness assumption, letting $M\to\infty$ yields $u(x)=u(0)$ for all $|x|\le R_0$. Since R_0 was arbitrary, this implies that u is indeed constant.

Proof of Lemma 2.

Without loss of generality, we may assume $\langle f\rangle = 1$. The elementary inequality $f\log f - f + 1 \le (f-1)^2$ then yields

$$\langle f\log f\rangle = \langle f\log f - f + 1\rangle \le \langle (f-1)^2\rangle.$$

Since $(f-1)^2 = (\sqrt{f}-1)^2(\sqrt{f}+1)^2$, we find that

$$\langle f \log f \rangle \leq \langle (\sqrt{f}+1)^2 \rangle \sup_a (\sqrt{f}-1)^2.$$

Since $\langle f \rangle = 1$, there exists $a_* \in [\lambda, 1]^d$ such that $f(a_*) \leq 1$. Since the diameter of $[\lambda, 1]^d$ is bounded by $d^{1/2}$, it follows that

$$\sqrt{f(a)} - 1 \leq \sqrt{f(a)} - \sqrt{f(a_*)} \leq d^{\frac{1}{2}} \sup_a \Big| \frac{\partial \sqrt{f}}{\partial a} \Big|.$$

Likewise there exists $a^* \in [\lambda, 1]^d$ such that $f(a^*) \geq 1$ and therefore

$$1 - \sqrt{f(a)} \leq \sqrt{f(a^*)} - \sqrt{f(a)} \leq d^{\frac{1}{2}} \sup_a \Big| \frac{\partial \sqrt{f}}{\partial a} \Big|.$$

Hence it follows that

$$\langle f \log f \rangle \leq \langle (\sqrt{f}+1)^2 \rangle \, d \sup_a \Big(\frac{\partial \sqrt{f}}{\partial a} \Big)^2 = \langle (\sqrt{f}+1)^2 \rangle \frac{d}{4} \sup_a \frac{1}{f} \Big(\frac{\partial f}{\partial a} \Big)^2.$$

Finally it holds that

$$\langle (\sqrt{f}+1)^2 \rangle \leq \langle 2f + 2 \rangle = 4,$$

and the combination of the previous two inequalities yields (16) with constant $\rho = \frac{1}{2d}$.

Proof of Lemma 3.
The following is a simple adaptation of the usual tensorization proof, cf. [15, Theorem 4.4]. Take any enumeration $(x_n)_{n \in \mathbb{N}}$ of the lattice $\mathbb{Z}^d$ and denote by $\langle \cdot \rangle_n$ the x_n-marginal of the (product) ensemble $\langle \cdot \rangle$. Furthermore we denote iteratively $f_0 := f$ and $f_n := \langle f_{n-1} \rangle_n$. Thus f_n is the average of f over the first n sites. Then the l. h. s. of (2) can be expressed as the following telescope sum:

$$\begin{aligned}\langle f \log f \rangle - \langle f \rangle \log \langle f \rangle &= \sum_{n=1}^{\infty} \big\langle f_{n-1} \log f_{n-1} - f_n \log f_n \big\rangle \\ &= \sum_{n=1}^{\infty} \big\langle \langle f_{n-1} \log f_{n-1} \rangle_n - \langle f_{n-1} \rangle_n \log \langle f_{n-1} \rangle_n \big\rangle. \quad (102)\end{aligned}$$

The assumption of single-site LSI yields

$$\langle f_{n-1} \log f_{n-1} \rangle_n - \langle f_{n-1} \rangle_n \log \langle f_{n-1} \rangle_n \leq \frac{1}{2\rho} \sup_{a(x_n)} \frac{1}{f_{n-1}} \Big(\frac{\partial f_{n-1}}{\partial a(x_n)} \Big)^2. \quad (103)$$

Notice that the definition of f_{n-1} immediately yields $f_{n-1} = \langle f \rangle_{<n}$, where we have denoted shortly by $\langle \cdot \rangle_{<n}$ the ensemble average over the first $n-1$ sites. By Cauchy-Schwarz, it holds that

$$\Big(\frac{\partial \langle f \rangle_{<n}}{\partial a(x_n)}\Big)^2 = \Big\langle \frac{\partial f}{\partial a(x_n)} \Big\rangle_{<n}^2 \le \langle f \rangle_{<n} \Big\langle \frac{1}{f} \Big(\frac{\partial f}{\partial a(x_n)}\Big)^2 \Big\rangle_{<n}.$$

Hence we find that

$$\sup_{a(x_n)} \frac{1}{f_{n-1}} \Big(\frac{\partial f_{n-1}}{\partial a(x_n)}\Big)^2 \le \sup_{a(x_n)} \Big\langle \frac{1}{f} \Big(\frac{\partial f}{\partial a(x_n)}\Big)^2 \Big\rangle_{<n} \le \Big\langle \sup_{a(x_n)} \frac{1}{f} \Big(\frac{\partial f}{\partial a(x_n)}\Big)^2 \Big\rangle_{<n}. \tag{104}$$

Finally we collect (102), (103), and (104) to obtain

$$\langle f \log f \rangle - \langle f \rangle \log \langle f \rangle \le \frac{1}{2\rho} \sum_{n=1}^{\infty} \Big\langle \sup_{a(x_n)} \frac{1}{f} \Big(\frac{\partial f}{\partial a(x_n)}\Big)^2 \Big\rangle,$$

which is the LSI (2).

References

[1] K. Astala, T. Iwaniec, G. Martin, Elliptic partial differential equations and quasiconformal mappings in the plane, *Princeton Mathematical Series 48*, Princeton University Press, Princeton, NJ (2009).

[2] I. Benjamini, H. Duminil-Copin, G. Kozma and A. Yadin, Disorder, entropy and harmonic functions, Preprint `arXiv:1111.4853` (2011)

[3] M. Biskup, M. Salvi and T. Wolff, A central limit theorem for the effective conductance: I. Linear boundary data and small ellipticity contrasts, Preprint `arXiv:1210.2371` (2012)

[4] E. A. Carlen, S. Kusuoka, D. W. Stroock, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré Probab. Statist.* **23 (2)**, 245-287 (1987)

[5] J. C. Conlon, A. Naddaf, On homogenization of elliptic equations with random coefficients, *Electron. J. Probab.* **9 (5)**, 1-58 (2000)

[6] E. De Giorgi, Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari, *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)* **3**, 25-43 (1957)

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

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

[9] G. Dolzmann, N. Hungerbühler, S. Müller, Uniqueness and maximal regularity for nonlinear elliptic systems of n-Laplace type with measure valued right hand side. *J. Reine Angew. Math.* **520** 1-35 (2000)

[10] P. Federbush, Partially alternate derivation of a result by Nelson, *J. Math. Phys.* **10 (1)**, 50-52 (1969)

[11] A. Gloria, Fluctuation of Solutions to Linear Elliptic Equations with Noisy Diffusion Coefficients, *Comm. PDE* **38 (2)**, 304-338 (2012)

[12] A. Gloria, F. Otto, An optimal variance estimate in stochastic homogenization of discrete elliptic equations, *Annals of Probability* **39 (3)**, 779-856 (2011)

[13] A. Gloria, S. Neukamm, F. Otto, Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics, Max Planck Institute for Mathematics in the Sciences Preprint 3/2013

[14] L. Gross, Logarithmic Sobolev inequalities, *American J. Math.* **97 (4)**, 1061-1083 (1975)

[15] A. Guionnet, B. Zegarlinski, Lecture notes on Logarithmic Sobolev Inequalities, *Lecture Notes Math.* **1801**, 1-134 (2003)

[16] G. F. Lawler and V. Limic, Random walk: a modern introduction, *Cambridge Stud. in Adv. Math.* **123**, CUP, Cambridge, UK (2010).

[17] E. H. Lieb and M. Loss, Analysis, *Graduate Stud. in Math.* **14**, 2nd ed., AMS, Providence, RI (2001)

[18] A. Naddaf, T. Spencer, Estimates on the variance of some homogenization problems, unpublished

[19] A. Naddaf, T. Spencer, On homogenization and scaling limit of some gradient perturbation of a massless free field, *Commun. Math. Phys.* **183**, 55-84 (1997)

[20] J. Nash, Continuity of solutions of parabolic and elliptic equations, *American J. Math.* **80**, 931-954 (1958)

[21] F. Nelson, A quartic interaction in two dimensions, in *Mathematical theory of elementary particles* (edited by R. Goodman, I. Segal), M. I. T. Press (Cambridge, MA), 69-73 (1966)

[22] F. Nelson, The free Markoff field, *J. Funct. Anal.* **12**, 211-227 (1973)

[23] J. Nolen, Normal approximation for a random elliptic equation, Preprint (2011). [Available online at `http://math.duke.edu/~nolen/preprints/ellipfluctper_rev.pdf`.]

[24] F. Riesz and B. Sz.-Nagy, Functional Analysis, *Dover Books on Adv. Math.*, Dover Publ. Inc., New York (1990)

[25] R. Rossignol, Noise-stability and central limit theorems for effective resistance of random electric networks, Preprint `arXiv:1206.3856`

[26] D. Stroock, B. Zegarlinski, The logarithmic Sobolev inequality for discrete spin systems on a lattice, *Commun. Math. Phys.* **149**, 175-193 (1992)