Probab. Theory Relat. Fields (2015) 163:527–573
DOI 10.1007/s00440-014-0598-0

Annealed estimates on the Green function

Daniel Marahrens · Felix Otto

Received: 4 January 2014 / Revised: 3 November 2014 / Published online: 25 November 2014

Abstract We consider a random, uniformly elliptic coefficient field $a(x)$ on the d-dimensional integer lattice $\mathbb{Z}^d$. We are interested in the spatial decay of the quenched elliptic Green function $G(a; x, y)$. Next to stationarity, we assume that the spatial correlation of the coefficient field decays sufficiently fast 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 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 function, respectively. This result relies on and substantially extends the one by Delmotte and Deuschel (Probab Theory Relat Fields 133:358–390, 2005), 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 are able to obtain optimal estimates on the random part of the homogenization error even for large ellipticity contrast.

Keywords Stochastic homogenization · Elliptic equations · Green function · Annealed estimates

Mathematics Subject Classification 35B27 · 35J08 · 39A70 · 60H25

D. Marahrens (✉) · F. Otto
Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany
e-mail: daniel.marahrens@mis.mpg.de

F. Otto
e-mail: felix.otto@mis.mpg.de

Outline

The outline of this work is as follows: After introducing the discrete setting in Sect. 1, we present the statistical assumptions and the main result on the annealed moments of the Green function in Sect. 2. The following two sections contain applications of the main result: We present optimal estimates on the random part of the homogenization error in Sects. 3 and 4 contains an annealed Hölder-estimate in the spirit of De Giorgi. In Sect. 5 we explain our main assumption, a logarithmic Sobolev inequality (LSI), which in particular holds for all independent, identically distributed (i.i.d.) coefficient fields. Section 6 contains the main ingredients of the proof of the annealed Green function estimates—in particular we recall the result by Delmotte and Deuschel [9]. We shall prove our main result in Sect. 7 and the corollaries on the homogenization error and the Hölder estimates in Sect. 8.

1 Discrete uniformly elliptic equations

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

$$\nabla^*(a\nabla u)(x) = f(x) \quad \text{for all } x \in \mathbb{Z}^d \tag{1}$$

in $d \geq 2$ dimensions. 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* as follows: Let $\mathbb{E}^d$ denote the set of *edges* of $\mathbb{Z}^d$ consisting of all pairs $[x, x+e_i]$ of neighboring vertices with $x \in \mathbb{Z}^d$, $i = 1, \ldots, d$, where $e_1, \ldots, e_d$ is the canonical basis of $\mathbb{R}^d$. For functions on vertices $\zeta : \mathbb{Z}^d \to \mathbb{R}$ and functions on edges $\xi : \mathbb{E}^d \to \mathbb{R}$ we set

$$\nabla\zeta([x, x+e_i]) = \zeta(x+e_i) - \zeta(x),$$
$$\nabla^*\xi(x) = \sum_{i=1}^{d} \big(\xi([x-e_i, x]) - \xi([x, x+e_i])\big).$$

The spatial derivatives $\nabla\zeta$ and $-\nabla^*\xi$ are the discrete gradient and divergence, respectively, on the lattice $\mathbb{Z}^d$. As our notation suggests, the operators ∇ and ∇^* are adjoint in the sense of

$$\sum_{e\in\mathbb{E}^d} \xi(e)\nabla\zeta(e) = \sum_{x\in\mathbb{Z}^d} \nabla^*\xi(x)\zeta(x).$$

In (1), the coefficient field a is a field on edges $a : \mathbb{E}^d \to \mathbb{R}$. Consequently $\nabla^* a \nabla$ is well-defined as an operator on vertex fields $\mathbb{Z}^d \to \mathbb{R}$. In this paper, we denote edges in $\mathbb{E}^d$ by the letters e and b and vertices in $\mathbb{Z}^d$ by the letters x, y and z.

Throughout this work we consider coefficient fields $a : \mathbb{E}^d \to \mathbb{R}$ in the space Ω of *uniformly elliptic* coefficient fields, i.e. we let

$$\Omega := \left\{ a : \mathbb{E}^d \to \mathbb{R} : \lambda \le a(e) \le 1 \text{ for all } e \in \mathbb{E}^d \right\} = [\lambda, 1]^{\mathbb{E}^d}. \tag{2}$$

Here and below $\lambda \in (0, 1)$ denotes the ellipticity ratio, which is fixed throughout the paper. This allows, for instance, to interpret $\nabla^* a \nabla$ as either the operator of a "conductance model" [i.e. the solution of (1) is a potential on a network of resistors] or the generator of a random walk on $\mathbb{Z}^d$ with jump rates across edges described by a. Note that if we interpreted $\nabla^* a \nabla$ as a discretization of a continuum operator $-\nabla \cdot a\nabla$, the coefficient field $a \in \Omega$ would be diagonal next to being symmetric and uniformly elliptic. In the discrete setting, 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 estimates (24) and (25) on the quenched Green function, on which our results rely.

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

$$\delta(x) = \begin{Bmatrix} 1 & \text{for } x = 0 \\ 0 & \text{otherwise} \end{Bmatrix}. \tag{3}$$

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

$$\sum_e \nabla\zeta(e) a(e) \nabla G(e, x') = \zeta(x') \tag{4}$$

for all compactly supported $\zeta : \mathbb{Z}^d \to \mathbb{R}$.

Remark 1 Here and throughout the paper, derivatives are understood to fall on the edge variable. We will always work in dimension $d \ge 2$. Dimension $d = 2$ needs a bit more care in terms of the definition of the Green function. Since we are only interested in *gradient* estimates, this is merely technical and will be ignored here. Sometimes, it is more convenient to think of ∇G as the limit of ∇G_T as $T \to \infty$ where G_T is the Green's function with a massive term in the sense that

$$T^{-1} G_T(\cdot, x') + \nabla^* a \nabla G_T(\cdot, x') = \delta(\cdot - x'); \tag{5}$$

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

$$\nabla^* a(x) \nabla G_L(\cdot, x') = \sum_{z \in \mathbb{Z}^d} \delta(\cdot - x' - Lz) - L^{-d}; \tag{6}$$

this is the case in the proof of Lemma 5.

2 Assumptions on the ensemble and main result

We are given a probability measure on the space Ω of uniformly elliptic, diagonal coefficient fields (endowed with the product topology), cf. (2) in 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 $\langle\cdot\rangle$. Functions $\zeta : \Omega \to \mathbb{R}$ will also be called *random variables*. Note that $\mathbb{Z}^d$ acts on $\mathbb{E}^d$ by translation and we denote by $b + x \in \mathbb{E}^d$ the edge $b \in \mathbb{E}^d$ shifted by $x \in \mathbb{Z}^d$. With this definition, 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 $a(\cdot + z) := (\mathbb{E}^d \ni e \mapsto a(e + z)) \in \Omega$ has the same distribution as a. We also note that the Green 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 variant of the logarithmic Sobolev inequality (LSI). It constitutes a quantification of ergodicity. In Sect. 5, we will comment on the LSI and the related spectral gap inequality—there we will also describe the relation between this LSI and the usual LSI.

Definition 1 [*Logarithmic Sobolev inequality*]. Let $\langle\cdot\rangle$ be a (not necessarily stationary) ensemble of coefficients a.

We say $\langle\cdot\rangle$ satisfies a logarithmic Sobolev inequality (LSI) with constant $\rho > 0$ if for all random variables $\zeta : \Omega \to \mathbb{R}$, we have that

$$\left\langle \zeta^2 \log \frac{\zeta^2}{\langle \zeta^2 \rangle} \right\rangle \le \frac{1}{2\rho} \left\langle \sum_{e \in \mathbb{E}^d} \left(\operatorname*{osc}_{a(e)} \zeta \right)^2 \right\rangle, \tag{7}$$

where the oscillation is to be taken over all values of $a(e) \in [\lambda, 1]$, i.e. over all coefficient fields $\tilde{a} \in \Omega$ that coincide with a outside of $e \in \mathbb{E}^d$ (i.e. $\tilde{a}(b) = a(b)$ for all $b \neq e$). In formulas:

$$\begin{aligned} \left(\operatorname*{osc}_{a(e)} \zeta \right)(a) = \sup\{\zeta(\tilde{a}) \mid \tilde{a} \in \Omega \text{ s.t. } \tilde{a}(b) = a(b)\ \forall b \neq e\} \\ - \inf\{\zeta(\tilde{a}) \mid \tilde{a} \in \Omega \text{ s.t. } \tilde{a}(b) = a(b)\ \forall b \neq e\}. \end{aligned}$$

Note that the difference between the LSI (7) and the usual LSI, see (20), lies in the use of the oscillation instead of the partial derivative $\frac{\partial \zeta}{\partial a(e)}$. The merit of this form is that it is satisfied by *any* ensemble of independent, identically distributed coefficients $(a(e))_{e \in \mathbb{E}^d}$, cf. Lemma 1 below. Our main result is:

Theorem 1 *Let $\langle\cdot\rangle$ be stationary and satisfy the LSI (7) with constant $\rho > 0$, see Definition 1. Then for all $1 \le p < \infty$, $x \in \mathbb{Z}^d$ and $b, b' \in \mathbb{E}^d$, we have that*

$$\langle |\nabla\nabla G(b, b')|^{2p} \rangle^{\frac{1}{2p}} \le C(d, \lambda, \rho, p)(|b - b'| + 1)^{-d}, \tag{8}$$

$$\langle |\nabla G(b, x)|^{2p} \rangle^{\frac{1}{2p}} \le C(d, \lambda, \rho, p)(|b - x| + 1)^{1-d}. \tag{9}$$

We furthermore let $|b|$ denote the Euclidean distance of the midpoint of the edge b from the origin and $|b - b'|$ the distance between the midpoints of the two edges b and b'. Recall that $b + x$ denotes the edge b shifted by x. 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 $p < \infty$.

We defer the proof of Theorem 1 until Sect. 7.4. Clearly, the spatial decay rates in Theorem 1 are optimal, since those are the decay rates of the constant coefficient Green function, see for instance Theorem 4.3.1 of [19] if $d > 2$ and Corollary 4.4.5 therein if $d = 2$. Note that we may assume without loss of generality that $x = 0$ in (9) since stationarity of $\langle\cdot\rangle$ and G implies

$$\langle|\nabla G(a; b, x)|^{2p}\rangle = \langle|\nabla G(a(\cdot - x); b, x)|^{2p}\rangle = \langle|\nabla G(a; b - x, 0)|^{2p}\rangle.$$

An interesting aspect of Theorem 1 is the following: The *quenched* versions of (8) and (9) are false, i.e. the *uniform* in a and *point-wise* in x estimates $|\nabla\nabla G(a; e, b)| \leq C(d, \lambda)(|e - b| + 1)^{-d}$ and $|\nabla G(a; e, 0)| \leq C(d, \lambda)(|e| + 1)^{d-1}$ do *not* hold (while suitably *spatially* averaged versions of both estimates do hold uniformly in a); see our discussion in Sect. 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 we have that*

$$\left\langle\left|G(x,0) - \langle G(x,0)\rangle\right|^{2p}\right\rangle^{\frac{1}{p}} \leq C(d,\lambda,\rho,p)\begin{Bmatrix}(|x|+1)^{2(1-d)} & d > 2\\ (|x|+1)^{-2}\log(|x|+2) & d = 2\end{Bmatrix} \tag{10}$$

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

The proof of Corollary 1 will be given in Sect. 8.1.

Remark 2 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 $p = 1$) in Sect. 8.2 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 and Naddaf [6, 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) [6, Theorem 1.2] in $d > 3$, this was already achieved by Gloria [12, Theorem 2]. For all other cases, our result appears to be new. Let us be more precise:

For a coefficient field $a : \mathbb{E}^d \to \mathbb{R}$ and a right-hand side $f : \mathbb{Z}^d \to \mathbb{R}$ we consider the solution $u : \mathbb{Z}^d \to \mathbb{R}$ of

$$\nabla^* a \nabla u = f \quad \text{on } \mathbb{Z}^d. \tag{11}$$

In order for (11) to have a unique solution that decays (i. e. $\lim_{|x|\to\infty} u(x) = 0$), we assume for simplicity that f is compactly supported (and furthermore is of zero spatial average in the case of $d = 2$). 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 $\hat{x}$, (11) now assumes the suggestive form of

$$\left(\nabla_\epsilon^* a(\tfrac{\cdot}{\epsilon})\nabla_\epsilon u\right)(\hat{x}) = \hat{f}(\hat{x}) \quad \text{for all } \hat{x} \in \epsilon\mathbb{Z}^d, \tag{12}$$

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. $\nabla_\epsilon u([\hat{x}, \hat{x} + \epsilon e_i]) = \epsilon^{-1}(u(\hat{x} + \epsilon e_i) - u(\hat{x}))$).

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) $\ell^p(\mathbb{Z}^d)$-norm $\left(\sum_x |u - \langle u\rangle|^p\right)^{1/p}$ of the fluctuations. This will be done for arbitrary stochastic moments (the role played by rp). Corollary 2 is the generalization of [6, Theorem 1.2] as well as [12, 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 (16).
- 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 [6, 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 (17). (Here, there is no logarithmic correction even for $d = 2$.)

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 (11). Let the spatial integrability exponents $\frac{d}{d-1} < p < \infty$ and $1 < q < d$ be related through $\frac{1}{q} = \frac{1}{d} + \frac{1}{p}$.*

In the case $d > 2$, we have for all $r < \infty$:

$$\left\langle \left(\sum_x |u - \langle u\rangle|^p\right)^r \right\rangle^{\frac{1}{rp}} \le C(d, \lambda, \rho, p, r)\left(\sum_x |f|^q\right)^{\frac{1}{q}}. \tag{13}$$

In the case $d = 2$, we additionally require that f is supported in $\{x : |x| \le R\}$ for some $R \ge 1$. Then we have for all $r < \infty$:

$$\left\langle \left(\sum_{x:|x|\le R} \left|u - \langle u\rangle\right|^p \right)^r \right\rangle^{\frac{1}{rp}} \le C(\lambda, \rho, p, r)\,(\log^{\frac{1}{2}} R) \left(\sum_x |f|^q \right)^{\frac{1}{q}}. \tag{14}$$

Corollary 3 *Let $\langle\cdot\rangle$ be as in Theorem 1; for compactly supported right-hand side $f(x)$, consider the decaying solution $u(x)$ to (11). Let the averaging function $g(x)$ be compactly supported. Let the two integrability exponents $\frac{2d}{d+2} < q, \tilde{q} < d$ be related by $\frac{1}{q} + \frac{1}{\tilde{q}} = \frac{2}{d} + \frac{1}{2}$. Then we have for all $r < \infty$:*

$$\left\langle \left| \sum_x (u - \langle u\rangle) g \right|^r \right\rangle^{\frac{1}{r}} \le C(d, \lambda, \rho, r) \left(\sum_x |f|^q \right)^{\frac{1}{q}} \left(\sum_x |g|^{\tilde{q}} \right)^{\frac{1}{\tilde{q}}}. \tag{15}$$

Corollaries 2 and 3 will be proved in Sect. 8.3. 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 $g(x) = \epsilon^d \hat{g}(\epsilon x)$; we also rewrite the solution itself in terms of $u(x) = \hat{u}_\epsilon(\epsilon x)$. In this notation, (13) (multiplied by $\epsilon^{d/p}$) turns into

$$\begin{aligned} &\left\langle \left(\epsilon^d \sum_{\hat{x}\in\epsilon\mathbb{Z}^d} \left|\hat{u}_\epsilon - \langle \hat{u}_\epsilon\rangle\right|^p \right)^r \right\rangle^{\frac{1}{rp}} \\ &\quad \le C(d, \lambda, \rho, p, r)\epsilon \left(\epsilon^d \sum_{\hat{x}\in\epsilon\mathbb{Z}^d} |\hat{f}|^q \right)^{\frac{1}{q}} \le C(d, \lambda, \rho, r, \hat{f})\epsilon. \end{aligned} \tag{16}$$

Note that this can be interpreted as the discrete version of

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

which highlights the $O(\epsilon)$-nature of the "spatially strong" error.

Likewise, (15) turns into

$$\begin{aligned} &\left\langle \left| \epsilon^d \sum_{\hat{x}\in\epsilon\mathbb{Z}^d} (\hat{u}_\epsilon - \langle \hat{u}_\epsilon\rangle)\hat{g} \right|^r \right\rangle^{\frac{1}{r}} \\ &\quad \le C(d, \lambda, \rho, r)\epsilon^{\frac{d}{2}} \left(\epsilon^d \sum_{\hat{x}\in\epsilon\mathbb{Z}^d} |\hat{f}|^q \right)^{\frac{1}{q}} \left(\epsilon^d \sum_{\hat{x}\in\epsilon\mathbb{Z}^d} |\hat{g}|^{\tilde{q}} \right)^{\frac{1}{\tilde{q}}} \\ &\quad \le C(d, \lambda, \rho, r, \hat{f}, \hat{g})\epsilon^{\frac{d}{2}}. \end{aligned} \tag{17}$$

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

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

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

Let us make a couple of further more detailed rems related to Corollaries 2 and 3. We note that the requirement that f has compact support and that u decays can be weakened: All we need is the Green function representation $u(x)=\sum_y G(x,y)f(y)$. It is worth 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 (16) and (17). Finally, since these estimates are of particular interest for stochastic homogenization in the continuum case with an elliptic partial differential equation on $\mathbb{R}^d$ and stationary coefficients satisfying an ergodicity assumption similar to Definition 1, we refer to the work in progress [13] where it is shown how to extend our results to the continuum setting. We just mention that in the continuum setting, the correlation length of the coefficient field presents an additional scale which here coincides with the length scale of the lattice. Due to small-scale regularity effects of the order of the correlation length, we can only expect $L^p_{\langle\cdot\rangle}$-bounds on local L^2-norms of the Green function over balls of radius of the order of the correlation length. Thus we always need additional arguments to treat local terms while the large-scale behavior may be treated exactly as here.

The central limit theorem (CLT) scaling $O(\epsilon^{d/2})$ of the weak error seems to suggests that $u_\epsilon(x)$ behaves like a random field of amplitude $O(1)$ and integrable correlations. In fact, this is misleading, as can be seen by distinguishing the scale $\frac{1}{\epsilon}$ on which f varies from the scale $1 \ll \frac{1}{\delta} \ll \frac{1}{\epsilon}$ on which we take the spatial average with help of the function g. If Corollary 3 were true in the limiting case of $q=d$ (which is not the case since the Hardy-Littlewood-Sobolev inequality in Step 3 in the proof of Corollary 3 requires $q<d$), we would obtain

$$\left\langle\left|\delta^d\sum_{x\in\mathbb{Z}^d}(\hat{u}_\epsilon(\epsilon x)-\langle\hat{u}_\epsilon(\epsilon x)\rangle)\hat{g}(\delta x)\right|^r\right\rangle^{\frac{1}{r}} \le C(d,\lambda,\rho,r,\hat{f},\hat{g})\epsilon\delta^{\frac{d}{2}-1}.$$

This refined estimate does suggest that $\hat{u}_\epsilon(\epsilon x)$ behaves like a random field of amplitude $O(\epsilon)$ and correlations that decay like the Green's function:

$$\left|\left\langle(\hat{u}_\epsilon(\epsilon x)-\langle\hat{u}_\epsilon(\epsilon x)\rangle)(\hat{u}_\epsilon(\epsilon y)-\langle\hat{u}_\epsilon(\epsilon y)\rangle)\right\rangle\right| \le C(d,\lambda,\rho,\hat{f})\epsilon^2(|x-y|+1)^{2-d}$$

for all $x,y\in\mathbb{Z}^d$. This scaling is natural, since it would follow from the (higher-order, two-scale) expansion $\hat{u}_\epsilon(\hat{x})\approx u_{\mathrm{hom}}(\hat{x})+\epsilon\sum_{k=1}^d\phi_k(\frac{\hat{x}}{\epsilon})\partial_k u_{\mathrm{hom}}(\hat{x})$ and the expected—but unproven—estimate on the covariance of this corrector:

$$|\langle\phi_k(x)\phi_k(y)\rangle| \le C(d,\lambda,\rho)(|x-y|+1)^{2-d}$$

for all $x, y \in \mathbb{Z}^d$. In the above, the function ϕ_k is the corrector in direction e_k (which is an a-harmonic function of affine behavior on large scales) and u_{hom} is the solution to the elliptic equation with homogenized coefficients. Here and in the following, we say that the function u is *a-harmonic* if it solves (1) with $f \equiv 0$. We rem here that the above-mentioned expansion for u_ϵ was recently quantified by Gloria, Neukamm and the second author [15] using Theorem 1. Indeed, there it is shown that the error in an H^1-norm in space and L^2-norm in probability for $u_\epsilon - u_{\mathrm{hom}} - \epsilon \sum_{k=1}^{d} \phi_k(\frac{\cdot}{\epsilon})\partial_k u_{\mathrm{hom}}$ is still of order ϵ, cf. (16). In order to obtain this result, the authors also treat the so-called systematic error, which is the difference between $\langle u_\epsilon \rangle$ and u_{hom}.

A more traditional CLT-scaling has been established for the *energy density*. For $g = f$, the weak measure of fluctuations turns into a measure of fluctuations of the energy:

$$\sum_x (u - \langle u \rangle) g = \sum_e a(\nabla u)^2 - \left\langle \sum_e a(\nabla u)^2 \right\rangle.$$

If we set $u = \phi_k$, then the (stationary) energy density defines the homogenized diffusion coefficient. In [16, 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 [14, Proposition 7], that 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 [27] gives a quantitative estimate based on a differential characterization of Gaussian distributions (second-order Poincaré inequality) and relies on the corrector estimates from [16, Theorem 2.1]. Biskup, Salvi and 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 [14]). Rossignol [29] 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 function, in the next section. We now address what we see as a new perspective on these results in the discrete setting, namely on De Giorgi's result on Hölder continuity of a-harmonic functions. (Again, we mention that our results extend to continuous problems with some modifications due to local-regularity effects, see [13].)

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, Corollary16.1.5] shows that for any $\alpha > 0$, there exists an explicit coefficient field $\alpha^2 \le a(z) \le 1$ such that $u(z) = \mathcal{R}e(|z|^{\alpha-1}z)$ is a-harmonic in $z \ni \mathbb{C} \cong \mathbb{R}^2$. 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). A celebrated result of De Giorgi [7, Theorem 2] states that this is the worst-case scenario: 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 $\lambda \le a(x) \le 1$ and any a-harmonic function $u(x)$, a bound of the form $|u(x)| \le C|x|^{\alpha_0}$ for $|x| \gg 1$ implies that u is constant. This result holds both in $\mathbb{R}^d$ and in $\mathbb{Z}^d$; for a discrete result, see [8, 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 is in fact more quantitative and can be rephrased as an inner regularity result in terms of Hölder continuity with Hölder exponent α_0: For any solution $u = u(x)$ to the discrete problem $\nabla^* a \nabla u = 0$ on the ball $\{x \in \mathbb{Z}^d : |x| \le R\}$ of radius $R \ge 1$, the α_0-Hölder modulus of continuity at zero is estimated by the supremum:

$$\sup_{x:|x|\le R} \frac{|u(x)-u(0)|}{(|x|+1)^{\alpha_0}} \le C(d,\lambda) R^{-\alpha_0} \sup_{x:|x|\le R} |u(x)|,$$

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

$$\forall\, \lambda \le a(x) \le 1, \quad \forall\, R < \infty: \quad \sup_u \frac{\sup_{x:|x|\le R} \frac{|u(x)-u(0)|}{(|x|+1)^{\alpha_0}}}{\frac{1}{R^{\alpha_0}} \sup_{x:|x|\le R} |u(x)|} \le C(d,\lambda), \tag{18}$$

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

In this context, we will show in Sect. 8.4 that Theorem 1 has the following Corollary.

Corollary 4 *Let $\langle\cdot\rangle$ be as in Theorem 1; for all $0 < \alpha < 1$, $p < \infty$ and $R < \infty$, we have that*

$$\left\langle \left(\sup_u \frac{\sup_{x:|x|\le R} \frac{|u(x)-u(0)|}{(|x|+1)^{\alpha}}}{\frac{1}{R^{\alpha}} \sup_{x:|x|\le R} |u(x)|} \right)^p \right\rangle \le C(d,\lambda,\rho,p,\alpha), \tag{19}$$

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

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 is estimated by its supremum in the concentric ball of twice the radius with a "quenched" constant $C(a)$ with all moments bounded independently of the radius. Indeed, with the same proof the numerator in Corollary 4 can be chosen as the full Hölder-norm on $\{x : |x| \le \frac{R}{2}\}$. Furthermore it is straight-forward to extend the result to functions $\nabla^* a \nabla u = f$ if we include the ℓ^d-norm of f over $\{x : |x| \le R\}$ in the denominator. The quantitative result of Corollary 4 has the Liouville principle as an easy corollary: For almost every a, any sub-linear a-harmonic function u must be constant. However, surprisingly for us, the (qualitative) Liouville principle 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 $\langle \sum_x |x|^2 G(t, x, 0) \rangle \le Ct$ on the second moments of the parabolic Green function $G(a; t, x, y) \overset{\text{short}}{=} G(t, x, y)$ (cf. [2, (SBD)], see Sect. 6 below for the definition of G), which in our uniformly elliptic context even holds in its stronger "quenched" version, that is, $\sum_x |x|^2 G(t, x, 0) \le Ct$.
– The annealed estimate $-\langle \sum_x G(t, x, 0) \log G(t, x, 0) \rangle \le C \log t$ on the spatial entropy of the parabolic Green 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:

$$\left\langle \sum_y G(1, 0, y) \sum_x \frac{|G(t, 0, x) - G(t-1, y, x)|^2}{G(t, 0, x) + G(t-1, y, x)} \right\rangle \le \frac{C}{t}$$

for some sequence $t \to \infty$.

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 and Spencer [22, Theorem 1] (in form of the Brascamp-Lieb inequality) and used most recently in [16, Lemma 2.3] in an indirect way and in [14] explicitly. The LSI follows like SG 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 [30, Theorem 1.8c)] for a discrete setting. Both SG and LSI quantify ergodicity of the ensemble, see for instance the discussion in [14, Chapter 4]. Recall that the usual LSI in this setting (with continuum derivative) would read

$$\left\langle \zeta^2 \log \frac{\zeta^2}{\langle \zeta^2 \rangle} \right\rangle \le \frac{1}{2\rho} \left\langle \sum_{e \in \mathbb{E}^d} \left(\frac{\partial \zeta}{\partial a(e)} \right)^2 \right\rangle. \tag{20}$$

In the LSI of Definition 1, we have simply changed the derivative by an oscillation in order to capture ensembles whose marginal distribution contains atoms, as we shall explain now.

Both SG and LSI are based on the notion of a *vertical derivative* (here, the oscillation) 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 [23], the version of SG that is based on the *continuum* vertical derivative [as on the r.h.s. of (20)] has been used [22]. However, this assumption rules out the natural example of coefficients with a single-site distribution that only assumes a *finite* number of values (Bernoulli). Hence in order to treat arbitrary single-site distributions, we are forced to consider the version of LSI found in Definition 1. A SG inequality based on the oscillation was already considered in [16, 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 [17]. It implies SG as well as concentration of measure [20, Chapter 5] and is equivalent to the notion of hyper-contractivity, see [17, Theorem 1] or [18, Theorem 4.1]. Incidentally, hyper-contractivity was first observed in the Gaussian context by Nelson [25], see [26] for an improved result. It is thus the older notion and in fact motivated the (somewhat implicit) introduction of LSI by Federbush [11]. We refer to [18] for a recent exposition on LSI.

The result of this section is that any independent, identically distributed coefficient-field satisfies the LSI (7) of Definition 1.

Lemma 1 *Consider an ensemble $\langle\cdot\rangle$ of i. i. d. coefficients on each edge with arbitrary marginal distribution on* $[\lambda, 1]$. *Then* (7) *holds, i.e.*

$$\left\langle \zeta^2 \log \frac{\zeta^2}{\langle \zeta^2 \rangle} \right\rangle \le \frac{1}{2\rho} \left\langle \sum_{e\in\mathbb{E}^d} \left(\operatorname*{osc}_{a(e)\in[\lambda,1]} \zeta \right)^2 \right\rangle$$

for all functions ζ of the coefficient field a. The constant ρ may be taken to be $\rho = \frac{1}{8}$.

Lemma 1 is an immediate consequence of the following two lemas. The first one shows that any single-edge distribution on $[\lambda, 1]$ satisfies the LSI in Definition 1.

Lemma 2 *Let $\langle\cdot\rangle$ be any distribution on* $[\lambda, 1]$. *Then we have that*

$$\left\langle \zeta^2 \log \frac{\zeta^2}{\langle \zeta^2 \rangle} \right\rangle \le \frac{1}{2\rho} \left(\operatorname*{osc}_{a\in[\lambda,1]} \zeta \right)^2 \tag{21}$$

for all functions $\zeta : [\lambda, 1] \to \mathbb{R}$. In fact, the constant $\rho = \frac{1}{8}$ will do.

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

Lemma 3 *Let $\langle\cdot\rangle$ be an ensemble consisting of independent distributions on the edges such that each single-edge distribution satisfies the LSI* (21) *with the same constant ρ. Then $\langle\cdot\rangle$ itself satisfies the LSI* (7) *with constant ρ.*

The proofs of Lemmas 2 and 3 will be given in Sect. 8.5.

6 Main ingredients of the proof

Loosely speaking, our approach consists in upgrading the (optimal) annealed estimates of Delmotte and Deuschel [9, Theorem 1.1] in terms of the integrability exponent p.

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

$$\langle|\nabla\nabla G(b,b')|\rangle \leq C(d,\lambda)(|b-b'|+1)^{-d}, \tag{22}$$

$$\langle|\nabla G(b,x)|\rangle \leq C(d,\lambda)(|b-x|+1)^{1-d}. \tag{23}$$

More precisely, we refer to the estimates (1.4) and (1.5a) in [9, Theorem 1.1] on the discrete *parabolic* Green function $G(t,x,y) = G(a;t,x,y)$ (i.e. the solution of $\partial_t G(t,\cdot,y) + \nabla^* a\nabla G(t,\cdot,y) = 0$ with $G(t=0,x,y) = \delta(x-y)$) that in our notation imply for any weight exponent $\alpha < \infty$:

$$\langle|\nabla\nabla G(t,b,b')|\rangle \leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-1}\left(\frac{|b-b'|^2}{t+1}+1\right)^{-\frac{\alpha}{2}}, \tag{24}$$

$$\langle|\nabla G(t,b,x)|\rangle \leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-\frac{1}{2}}\left(\frac{|b-x|^2}{t+1}+1\right)^{-\frac{\alpha}{2}}. \tag{25}$$

(In fact, [9] establishes (24) and (25) with exponentially decaying weights instead of just algebraically decaying ones.) Since the elliptic Green function can be inferred from the parabolic one via $G(x,y) = \int_0^\infty G(t,x,y)dt$, these estimates imply (22) and (23) (by fixing some $\alpha > d$ and performing the change of variables $\hat{t} = |x|^{-2}(t+1)$). Actually, [9] establishes (25) and thus (23) in the stronger form where the L^1-norm $\langle|\cdot|\rangle$ is replaced by the L^2-norm $\langle|\cdot|^2\rangle^{1/2}$, i.e.

$$\langle|\nabla G(t,b,x)|^2\rangle^{\frac{1}{2}} \leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-\frac{1}{2}}\left(\frac{|b-x|^2}{t+1}+1\right)^{-\frac{\alpha}{2}}.$$

Let us point out that the *spatially point-wise annealed* estimates (24) and (25) are consequences of the following *spatially averaged quenched* estimates

$$\sum_{x\in\mathbb{Z}^d}\left(\left(\frac{|x|^2}{t+1}+1\right)^{\frac{\alpha}{2}} G(t,x,0)\right)^2 \leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}}, \tag{26}$$

$$\sum_{b\in\mathbb{E}^d}\left(\left(\frac{|b|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla G(t,b,0)|\right)^2 \leq C(d,\lambda,\alpha)(t+1)^{-\frac{d}{2}-1}. \tag{27}$$

The first estimate (26) is the (upper, off-diagonal part of the) celebrated Nash estimate [24, Appendix]. The discrete case was treated in full generality in [4, Corollary 3.28]. The second estimate (27) is a consequence of the first one. For an elementary proof of both, we refer to [14, Lemmas 24 and 25], with the Nash inequality as only noteworthy ingredient. Let us point out how (27) is implies (24). Using the semi-group property in form of $\nabla\nabla G(t,b,b') = \sum_y \nabla G(\frac{t}{2},b,y)\nabla G(\frac{t}{2},y,b')$ 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:

$$\begin{aligned}
&\left(\frac{|b-b'|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla\nabla G(t,b,b')| \\
&\le \sum_{y\in\mathbb{Z}^d} \left(\frac{2|b-y|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla G(\tfrac{t}{2},b,y)| \left(\frac{2|b'-y|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla G(\tfrac{t}{2},y,b')| \\
&\le \left(\sum_{y\in\mathbb{Z}^d} \left(\left(\frac{2|b-y|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla G(\tfrac{t}{2},b,y)|\right)^2 \right. \\
&\quad \left. \times \sum_{y\in\mathbb{Z}^d} \left(\left(\frac{2|b'-y|^2}{t+1}+1\right)^{\frac{\alpha}{2}} |\nabla G(\tfrac{t}{2},b',y)|\right)^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Note that the right-hand side of the last inequality does not allow for application of (27), since the sum is not in the variable in which the derivative is taken. However, we take the expectation, use the Cauchy-Schwarz inequality in $\langle\cdot\rangle$ and stationarity and symmetry in form of $\langle|\nabla G(\frac{t}{2},b,y)|^2\rangle = \langle|\nabla G(\frac{t}{2},b-y,0)|^2\rangle$ to obtain

$$\begin{aligned}
&\left(\frac{|b-b'|^2}{t+1}+1\right)^{\frac{\alpha}{2}} \langle|\nabla\nabla G(t,b,b')|\rangle \\
&\le \left(\sum_{y\in\mathbb{Z}^d} \left(\frac{2|b-y|^2}{t+1}+1\right)^{\alpha} \langle|\nabla G(\tfrac{t}{2},b,y)|^2\rangle \right. \\
&\quad \left. \times \sum_{y\in\mathbb{Z}^d} \left(\frac{2|b'-y|^2}{t+1}+1\right)^{\alpha} \langle|\nabla G(\tfrac{t}{2},b',y)|^2\rangle\right)^{\frac{1}{2}} \\
&\le \left(\left\langle \sum_{y\in\mathbb{Z}^d} \left(\frac{2|b-y|^2}{t+1}+2\right)^{\alpha} |\nabla G(\tfrac{t}{2},b-y,0)|^2\right\rangle \right. \\
&\quad \left. \times \left\langle \sum_{y\in\mathbb{Z}^d} \left(\frac{2|b'-y|^2}{t+1}+2\right)^{\alpha} |\nabla G(\tfrac{t}{2},b'-y,0)|^2\right\rangle\right)^{\frac{1}{2}}.
\end{aligned}$$

We now see that (27) implies (24). The estimate (25) is derived via the semi-group property in form of $\nabla G(t,b,x) = \sum_y \nabla G(\frac{t}{2},b,y)G(\frac{t}{2},y,x)$ from the combination of (26) and (27) 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 lema, which we shall prove in Sect. 7.1.

Lemma 4 *Let $\langle\cdot\rangle$ satisfy LSI (7) with constant $\rho > 0$. Then for arbitrary $\delta > 0$ and $1 \le p < \infty$ and for any $\zeta : \Omega \to \mathbb{R}$, we have that*

$$\langle|\zeta|^{2p}\rangle^{\frac{1}{2p}} \le C(d,\rho,p,\delta)\langle|\zeta|\rangle + \delta\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle^{\frac{1}{2p}}. \tag{28}$$

The preceding lema may be seen as a reverse Hölder inequality in probability: If one controls a bit (recall that $\delta > 0$ may be arbitrarily small) of the vertical derivative of a random variable ζ, then its $L^1_{\langle\cdot\rangle}$-norm bounds its $L^{2p}_{\langle\cdot\rangle}$-norm. It can be seen as a softening of the concentration of measure phenomenon, which requires Lipschitz continuity of ζ, cf. [20, Theorem 5.3].

In order to make use of Lemma 4, we need to estimate the vertical derivatives (the oscillation) of $\nabla\nabla G$ and ∇G. A corresponding estimate (see Lemma 6 below) stands at the heart of our result and and it is worthwhile sketching its derivation.

Let us discuss the estimation of $\nabla\nabla G$ only—the first derivative is then bounded in a similar fashion. As a stand-in for the oscillation, let us consider the conceptually slightly simpler derivative $\frac{\partial}{\partial a(e)}$. Formally taking $\frac{\partial}{\partial a(e)}$ in the defining equation for G yields

$$\nabla^* a\nabla\frac{\partial G(\cdot,x')}{\partial a(e)} = -\nabla^*\delta(\cdot-e)\nabla G(\cdot,x'),$$

where as before δ denotes the discrete version of the Dirac distribution given in equation (3). This is an elliptic equation of the form (1) and the Green function representation for its solution yields

$$\frac{\partial G(x,x')}{\partial a(e)} = -\nabla G(x,e)\nabla G(e,x')$$

and upon taking spatial derivatives along the edges b and b' we obtain:

$$\frac{\partial\nabla\nabla G(b,b')}{\partial a(e)} = -\nabla\nabla G(b,e)\nabla\nabla G(e,b').$$

(This is the infinitesimal version of Step 1 of the proof of Lemma 6 in Sect. 7.3.) Substituting this result in place of $\operatorname{osc}_{a(e)}\nabla\nabla G(b,b')$ yields a quadratic term of the form

$$\left\langle\left(\sum_e(\nabla\nabla G(b,e))^2\left(\nabla\nabla G(e,b')\right)^2\right)^p\right\rangle^{\frac{1}{2p}},$$

which we need to bound in terms of $\langle|\nabla\nabla G|^{2p}\rangle^{\frac{1}{2p}}$—then we can absorb this term on the left-hand side by chosing δ in Lemma 4 sufficiently small. Now we notice that, for instance, $(\nabla\nabla G(b,e))^2$ localizes strongly around $e \approx b$ (it decays like $|b-e|^{-2d}$ in the constant-coefficient case where $-d$ is the critical exponent). In the general variable-coefficient case, this decay behaviour is quantified in the following suboptimal, but *quenched* estimates on the (elliptic) Green function:

Lemma 5 [Gloria and Otto] *There exists an exponent* $\alpha_0 = \alpha_0(d,\lambda) > 0$ *such that for all* $R > 0$ *and* $b \in \mathbb{E}^d$, *we have that*

$$R^{2\alpha_0} \sum_{e: R \le |e-b| < 2R} |\nabla\nabla G(e,b)|^2 \le C(d,\lambda), \tag{29}$$

$$\sum_{e: R \le |e| < 2R} |\nabla G(e,0)|^2 \le C(d,\lambda). \tag{30}$$

The estimate (30) was established in the stronger (dimensionally optimal) form of $\sum_{R \le |e| < 2R} |\nabla G(e,0)|^2 \lesssim R^{2-d}$ in [16, Lemma 2.9]; in its weaker form of (30), it is straight forward for $d > 2$. The proof of estimate (30) in [16] in case of $d = 2$ is subtle and relied on an adaptation of [10]. In Sect. 7.2, we will give an elementary argument for the estimate (29) which we could not find in the literature. In fact, the proof we give shows (29) in the following equivalent (up to redefinition of α_0) form without upper bound on the area of summation:

$$R^{2\alpha_0} \sum_{e: R \le |e-b|} |\nabla\nabla G(e,b)|^2 \le C(d,\lambda). \tag{31}$$

We rem that the proof presented here does not make use of the maximum principle (directly or indirectly). However, as pointed out above, the present proof of Proposition 1 relies on Nash's estimates (24) and (25) which are only available in the case of scalar equations. In a work in preparation [5], we obtain the result of Proposition 1 with help of arguments that are also applicable in case of systems. Hence we expect Theorem 1 to also hold in case of systems.

Using the result of Lemma 5, we may introduce a weight $|b-e|^{2\alpha_0}$ (and corresponding negative weight $|b-e|^{-2\alpha_0}$ on the other term $(\nabla\nabla G(e,b'))^2$) and apply Hölder's inequality with exponents q and p such that $-2p\alpha_0 < -d$ in order to localize the remaining term $\sum_e |b-e|^{-2\alpha_0} \langle |\nabla\nabla G(e,b')|^{2p} \rangle$ and we will deduce in Sect. 7.3 the following result which then allows us to absorb the oscillation-term in Lemma 4 with $\zeta = \nabla\nabla G(b,b')$:

Lemma 6 *There exists an integrability exponent* $p_0 = p_0(d,\lambda) < \infty$ *such that for all* $p \ge p_0$, *we have that*

$$\sup_{b,b' \in \mathbb{E}^d} \left\{ (|b-b'|+1)^d \left\langle \left(\sum_{e \in \mathbb{E}^d} \left(\operatorname*{osc}_{a(e)} \nabla\nabla G(b,b') \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \right\}$$
$$\le C(d,\lambda,p) \sup_{b,b' \in \mathbb{E}^d} \left\{ (|b-b'|+1)^d \langle |\nabla\nabla G(b,b')|^{2p} \rangle^{\frac{1}{2p}} \right\} \tag{32}$$

and

$$\sup_{b\in\mathbb{E}^d, x\in\mathbb{Z}^d}\left\{(|b-x|+1)^{d-1}\left\langle\left(\sum_{e\in\mathbb{E}^d}\left(\operatorname*{osc}_{a(e)}\nabla G(b,x)\right)^2\right)^p\right\rangle^{\frac{1}{2p}}\right\}$$

$$\leq C(d,\lambda,p)\Bigg(\sup_{b\in\mathbb{E}^d, x\in\mathbb{Z}^d}\left\{(|b-x|+1)^{d-1}\langle|\nabla G(b,x)|^{2p}\rangle^{\frac{1}{2p}}\right\}$$

$$+\sup_{b,b'\in\mathbb{E}^d}\left\{(|b-b'|+1)^{d}\langle|\nabla\nabla G(b,b')|^{2p}\rangle^{\frac{1}{2p}}\right\}\Bigg). \tag{33}$$

Note that in contrast to Proposition 1, here the only assumption on the ensemble is LSI (7)—in particular, *Lemmas* 4 *and* 6 *do not require stationarity* and stationarity enters the proof of Theorem 1 only through Proposition 1. The above sketch and the formulation of Lemma 6 show that with our method, we first have to estimate the mixed *second* derivatives $\langle|\nabla\nabla G(b,b')|^{2p}\rangle$ before we can tackle the *first* derivatives $\langle|\nabla G(b,0)|^{2p}\rangle$. 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 $\langle|\nabla G(b,0)|^4\rangle$ that is needed in the proof of Corollary 1.

Remark 3 We mention that with the same proof, one obtains a periodic version of Theorem 1 (with constants uniform in L) for the Green function defined in (6). In that case, one just replaces the Euclidean distance $|x|$ on $\mathbb{Z}^d$ by its periodic version $\operatorname{dist}(x, L\mathbb{Z}^d)$ 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 [14, Theorem 3(b)]. The same is true in the presence of a massive term, cf. (5).

7 Proof of Theorem 1

We begin by proving the necessary Lemmas 4, 5 and 6 and finish by combining these results to obtain Theorem 1. (Lemma 5 only enters in the proof of Lemma 6.)

7.1 Proof of Lemma 4

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

$$\langle\zeta^2\rangle^{\frac{1}{2}} \leq \left(\exp\left(\frac{2}{\rho\delta^2}\right)+\frac{\rho\delta^2}{2e}\right)\langle|\zeta|\rangle+\delta\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right\rangle^{\frac{1}{2}}, \tag{34}$$

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 we have that

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

Since $x\log x$ is bounded from below by $\frac{1}{e}$, we have that $\frac{2}{e}|\zeta| + \zeta^2\log\zeta^2 \ge 0$ for all ζ. It follows that

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

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

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

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

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

Combining the last two estimates, we deduce

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

Hence LSI yields

$$\langle\zeta^2\rangle \le \left(\exp\left(\frac{2}{\rho\delta^2}\right) + \frac{\rho\delta^2}{2e}\right)^2\langle|\zeta|\rangle^2 + \delta^2\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right\rangle$$

and estimate (34) follows from taking the square root and applying the inequality $\sqrt{\zeta+\xi} \le \sqrt{\zeta} + \sqrt{\xi}$ for all numbers $\zeta, \xi \ge 0$.

Step 2. We finish the proof of (28), i.e. show that

$$\langle\zeta^{2p}\rangle^{\frac{1}{2p}} \le C(\rho, p, \delta)\langle|\zeta|\rangle + \delta\left(\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle\right)^{\frac{1}{2p}}$$

for general $p \ge 1$. To that end, we apply (34) to ζ replaced by $|\zeta|^p$:

$$\langle|\zeta|^{2p}\rangle \le C(\rho, p, \delta)\langle|\zeta|^p\rangle^2 + \delta\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}|\zeta|^p\right)^2\right\rangle,$$

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

$$\langle|\zeta|^{2p}\rangle \leq C(\rho, p, \delta)\langle|\zeta|\rangle^{2p} + 2\delta\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}|\zeta|^p\right)^2\right\rangle. \tag{35}$$

Now we use that

$$\operatorname*{osc}_{a(e)}|\zeta|^p \leq C(p)\left(|\zeta|^{p-1}\operatorname*{osc}_{a(e)}\zeta + \left(\operatorname*{osc}_{a(e)}\zeta\right)^p\right),$$

which follows from the elementary inequality $|\zeta^p - \xi^p| \leq C(p)(\zeta^{p-1}|\zeta - \xi| + |\zeta - \xi|^p)$ for all numbers $\zeta, \xi > 0$ and the triangle inequality in form of $\operatorname{osc}_{a(e)}|\zeta| \leq \operatorname{osc}_{a(e)}\zeta$. Hence (35) yields

$$\begin{aligned}\langle|\zeta|^{2p}\rangle \leq C(\rho, p, \delta)\langle|\zeta|\rangle^{2p} &+ 2C(p)\delta\left\langle|\zeta|^{2p-2}\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right\rangle \\ &+ 2C(p)\delta\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^{2p}\right\rangle.\end{aligned} \tag{36}$$

The last term on the right-hand side may be estimated by the obvious relation $\ell^2 \subset \ell^{2p}$ between discrete spaces:

$$\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^{2p}\right\rangle \leq \left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle. \tag{37}$$

Furthermore, Hölder's inequality followed by Young's inequality yields

$$\begin{aligned}\left\langle|\zeta|^{2p-2}\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right\rangle &\leq \langle|\zeta|^{2p}\rangle^{1-\frac{1}{p}}\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle \\ &\leq \frac{1}{4C(p)\delta}\langle|\zeta|^{2p}\rangle + (4C(p)\delta)^{p-1}\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle.\end{aligned} \tag{38}$$

Hence collecting (36), (37) and (38) yields

$$\langle|\zeta|^{2p}\rangle \leq C(\rho, p, \delta)\langle|\zeta|\rangle^{2p} + 2\left(2C(p)\delta + (4C(p)\delta)^p\right)\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\zeta\right)^2\right)^p\right\rangle,$$

where we have absorbed the second term of (38) on the left-hand side.

By redefining δ, we obtain (28).

7.2 Proof of Lemma 5

We just give the proof of (29); for (30), we refer to [16, Lemma 2.9]. Note that in the stronger form $\sum_{e:R\le|b-e|<2R}|\nabla\nabla G(e,b)|^2 \le C(d,\lambda)R^{2-d-2\alpha_0}$, Estimate (29) can also be seen as a consequence of the following classical ingredients (which however would not hold in the systems case):

- The optimal decay of $G(x,y)$ itself, that is just needed in a spatially averaged sense of $R^{-d}\sum_{y:R\le|x-y|<2R}|G(x,y)-\bar{G}| \le C(d,\lambda)R^{2-d}$ (thanks to subtracting the average $\bar{G}$ over the annulus $\{y : R \le |x-y| \le 2R\}$, this estimate also holds in $d=2$, cf. [16, Lemma 2.8]),
- De Giorgi's Hölder continuity estimate, that then yields for some $\alpha_0=\alpha_0(d,\lambda)>0$ that $\sup_{x:R\le|b-x|<2R}|\nabla G(x,b)| \le C(d,\lambda)R^{2-d-\alpha_0}$,
- Caccioppoli's estimate, that then yields

$$\sum_{e:R\le|b-e|<2R}|\nabla\nabla G(e,b)|^2 \le C(d,\lambda)R^{2-d-2\alpha_0}.$$

Sketch of proof: In the following, we will give a proof for (31), i.e.

$$R^{2\alpha_0}\sum_{e:R\le|e-b|}|\nabla\nabla G(e,b)|^2 \le C(d,\lambda), \tag{31}$$

which also holds in the case of systems and which we could not find in the literature. The proof is based on a bound on $\nabla\nabla G$ in $\ell^2(\mathbb{Z}^d)$ (which crucially uses discreteness), see Step 1, and which is then upgraded in Step 2 to an improved decay estimate (31) using the fact that $\nabla G(\cdot,b)$ is a-harmonic away from the endpoints of b. We will proceed formally at first, assuming for simplicity that $\nabla G(\cdot,b)$ is a valid test function for (4) and that the Leibniz rule holds. This is the content of Steps 1 and 2. In Step 3, we will provide a discrete argument that replaces the use of the Leibniz rule in Step 2. (Essentially, we will establish a discrete version of the Leibniz rule.) Finally, in the discrete setting, differentiability is not an issue and therefore admissibility is just a question of sufficiently fast decay at infinity. In Step 4, we will show how to avoid this problem by considering the *periodic* Green function G_L on a torus of size L and then discuss convergence of G_L to G as $L\to\infty$.

Step 1. In this step, we derive the a priori estimate

$$\sum_e|\nabla\nabla G(e,b)|^2 \le C(d,\lambda). \tag{39}$$

Indeed, recall the weak formulation (4) of the defining equation for G, i.e.

$$\forall\,\zeta(x): \sum_e\nabla\zeta(e)a(e)\nabla G(e,x)=\zeta(x).$$

Taking the derivative w. r. t. the variable x along some edge b yields

$$\forall\ \zeta(x): \quad \sum_e \nabla\zeta(e)a(e)\nabla\nabla G(e,b) = \nabla\zeta(b). \tag{40}$$

The choice of $\zeta(x) = \nabla G(x,b)$ formally yields

$$\sum_e a(e)(\nabla\nabla G(e,b))^2 = \nabla\nabla G(b,b).$$

Since $a(b) \geq \lambda$, this implies (39) in the explicit form of

$$\sum_e |\nabla\nabla G(e,b)|^2 \leq \lambda^{-2}. \tag{41}$$

However, $\zeta(x) = \nabla G(x,b)$ is not compactly supported so one needs another argument to arrive at (41). This is described in Step 3.

Step 2. In this step, we will use the usual Leibniz rule. Since this is only possible in the continuum setting, we will suggestively use continuous notation and formally derive a continuum version of (31), that is

$$\int_{\{x:|x|\geq R\}} |\nabla u|^2 dx \leq C(d,\lambda)R^{-2\alpha_0} \int_{\{x:|x|\geq 1\}} |\nabla u|^2 dx \tag{42}$$

for $R \geq 1$ and a function u (which in (31) is taken to be $u = \nabla G(\cdot,b)$) satisfying

$$-\nabla\cdot a\nabla u(x) = 0 \quad \text{in } \{x : |x| > 1\}. \tag{43}$$

Indeed, let $\eta(x)$ be a cut-off function for $\{x : |x| \geq 2R\}$ in $\{x : |x| \geq R\}$. We test (43) with $\zeta = \eta^2(u - \bar{u})$, where $\bar{u}$ is the spatial average of u on the annulus $\{x : R \leq |x| \leq 2R\}$. It is a priori not clear that this is an admissible test function for (43); we shall address this in the next step. We appeal to the identity

$$\nabla(\eta^2(u-\bar{u}))\cdot a\nabla u = \nabla(\eta(u-\bar{u}))\cdot a\nabla(\eta(u-\bar{u})) - (u-\bar{u})^2\nabla\eta\cdot a\nabla\eta, \tag{44}$$

which in view of ellipticity in form of $\lambda|\xi|^2 \leq \xi\cdot a(x)\xi \leq |\xi|^2$ for all $\xi \in \mathbb{R}^d$ turns into the inequality

$$\nabla(\eta^2(u-\bar{u}))\cdot a\nabla u \geq \lambda|\nabla(\eta(u-\bar{u}))|^2 - (u-\bar{u})^2|\nabla\eta|^2. \tag{45}$$

Hence from testing (43) we obtain

$$\lambda\int_{\mathbb{R}^d} |\nabla(\eta(u-\bar{u}))|^2 dx \leq \int_{\mathbb{R}^d} (u-\bar{u})^2|\nabla\eta|^2 dx,$$

which by the choice of η yields the Caccioppoli estimate

$$\int_{\{x:|x|\geq 2R\}} |\nabla u|^2 dx \leq C(d,\lambda) R^{-2} \int_{\{x:R\leq|x|\leq 2R\}} (u-\bar{u})^2 dx. \tag{46}$$

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

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

which can be reformulated as

$$\int_{\{x:|x|\geq R\}} |\nabla u|^2 dx \leq C(d,\lambda) \int_{\{x:R\leq|x|\leq 2R\}} |\nabla u|^2 dx. \tag{47}$$

A standard iteration argument now leads from (47) to (42): Introducing the notation $I_k := \int_{\{x:|x|\geq 2^k\}} |\nabla u|^2 dx$, estimate (47) reads

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

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

$$\forall k \in \{0,1,\ldots\} \quad I_{k+1} \leq \theta I_k,$$

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

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

In the original notation, this implies (42) in form of

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

Step 3. Discrete Leibniz rule: In this step, we indicate the modifications in Step 2 that are necessary to treat the discrete case. The first modification results from the fact that Leibniz rule and thus the neat identity (44) does not hold anymore. However, we claim that the estimate (45) survives in form of

$$\nabla(\eta^2(u-\bar{u}))(e)a(e)\nabla u(e) \geq \lambda(\nabla(\eta(u-\bar{u}))(e))^2 - ([u](e)-\bar{u})^2(\nabla\eta(e))^2, \tag{48}$$

where we denote by $[u]([x, x+e_i]) = \frac{1}{2}(u(x) + u(x+e_i))$ the local average of u along each edge $e = [x, x+e_i]$. Indeed, since $\lambda \leq a(e) \leq 1$ is elliptic, this follows from the simple inequality on 4 *numbers* $\eta = \eta(x)$, $\tilde{\eta} = \eta(x+e_i)$, $v = u(x) - \bar{u}$ and $\tilde{v} = u(x+e_i) - \bar{u}$:

$$(\eta^2 v - \tilde{\eta}^2\tilde{v})(v-\tilde{v}) - (\eta v - \tilde{\eta}\tilde{v})^2 = -(\eta-\tilde{\eta})^2 v\tilde{v} \geq -(\eta-\tilde{\eta})^2(\tfrac{1}{2}(v+\tilde{v}))^2.$$

Hence, if $\eta(x)$ denotes the (slightly narrower) cut-off function for $\{x : |x| \geq 2R-2\}$ in $\{x : |x| \geq R+2\}$ (which is possible for $R \geq 5$), from (48) we obtain the following substitute of (46)

$$\sum_{e:|e-b|\geq 2R} |\nabla u(e)|^2 \leq C(d,\lambda)R^{-2} \sum_{e:R+1\leq|e-b|\leq 2R-1} ([u](e)-\bar{u})^2$$
$$\leq C(d,\lambda)R^{-2} \sum_{x:R\leq|b-x|\leq 2R} (u(x)-\bar{u})^2, \qquad (49)$$

as long as $\eta^2(u-\bar{u})$ is an admissible test function; we show how to avoid the question of admissibility in the next step.

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

Step 4. In this step, we discuss the approximation of the Green function by the periodic problem, see also Remark 1 above. This way we avoid the issue of admissibility of ∇G in Steps 1 and 2. We consider the *periodic* discrete elliptic Green function $G_L(x,x') = G_L(a,x,x')$ of period L. Up to additive constants, it is characterized by the weak equation

$$\sum_{e\in\mathbb{E}_L^d} \nabla\zeta_L(e)a(e)\nabla G_L(e,x') = \zeta_L(x') - L^{-d}\sum_{x\in\mathbb{Z}_L^d}\zeta_L(x) \qquad (50)$$

for all L-periodic $\zeta_L(x)$. Here $x \in \mathbb{Z}_L^d = \mathbb{Z}^d \cap [-\frac{L}{2},\frac{L}{2})^d$ stands short for the (discrete) torus of length L and $b \in \mathbb{E}_L^d = \mathbb{E}^d \cap [-\frac{L}{2},\frac{L}{2})^d$ stands short for the set of all edges b whose midpoint is contained in the box $[-\frac{L}{2},\frac{L}{2})^d$. The additive constant may be determined by the requirement that all solutions have vanishing average over the torus, i.e.

$$\sum_{x\in\mathbb{Z}_L^d} G_L(x,x') = 0.$$

Letting $\zeta_L(\cdot) = G_L(\cdot,x')$ in (50) yields

$$\frac{1}{\lambda}\sum_{e\in\mathbb{E}_L^d}|\nabla G_L(e,x')|^2 \leq \sum_{e\in\mathbb{E}_L^d}\nabla G_L(e,x')a(e)\nabla G_L(e,x') = G_L(x,x').$$

Extending G_L by zero outside the torus $\mathbb{Z}_L^d$ to a function on the whole of $\mathbb{Z}^d$, we may consider these sums to be over all of $\mathbb{Z}^d$ resp. $\mathbb{E}^d$. (We lose a constant factor 2 in the process since we count some edges twice.) If $d > 2$, we may apply a Sobolev embedding to obtain that

$$\left(\sum_{x\in\mathbb{Z}^d}|G_L(x,x')|^{\frac{2d}{d-2}}\right)^{\frac{d-2}{d}} \leq C(d)\sum_{e\in\mathbb{E}^d}|\nabla G_L(e,x')|^2 \leq C(d,\lambda)|G_L(x',x')|.$$

Since the right-hand side is bounded by $\|G_L(\cdot,x')\|_{\ell^{\frac{2d}{d-2}}(\mathbb{Z}^d)}$, we obtain that

$$\|G_L(\cdot,x')\|_{\ell^{\frac{2d}{d-2}}(\mathbb{Z}^d)} \le C(d,\lambda) \quad \text{and} \quad \|\nabla G_L(\cdot,x')\|_{\ell^2(\mathbb{E}^d)} \le C(d,\lambda)$$

as long as $d > 2$. Since these bounds are uniform in L and the Green function $G(x,x')$ vanishes as $|x| \to \infty$, a standard diagonal argument shows that there exists a subsequence $L \to \infty$ such that G_L and hence ∇G_L as well as $\nabla\nabla G_L$ converge point-wise. These limits are indeed G, ∇G and $\nabla\nabla G$, since letting $\zeta_L(x) = \sum_{k\in\mathbb{Z}^d} \zeta(x - kL)$ in (50) with some compactly supported $\zeta : \mathbb{Z}^d \to \mathbb{R}$ yields

$$\sum_{e\in\mathbb{E}^d} \nabla\zeta(e)a(e)\nabla G_L(e,x') = \sum_{k\in\mathbb{Z}^d} \zeta(x'-kL) - \frac{1}{L^d}\sum_{x\in\mathbb{Z}^d}\zeta(x)$$

and we may take the limit $L \to \infty$. In $d = 2$, this argument breaks down and indeed the Green function is not unambiguously defined—indeed, the corresponding random walk is transient and the expected occupation time of the random walk (which equals the Green function in $d > 2$) becomes infinite. However, we may let $\lim_{L\to\infty} \nabla G_L(e,x')$ to be the definition of $\nabla G(e,x')$ in the case of $d = 2$. (The limit of $\nabla\nabla G_L$ exists thanks to (39)—to deduce a limit of ∇G_L in some $\ell^p(\mathbb{Z}^d)$ for large p, we may use that

$$\sup_{x:R\le|b-x|<2R} |\nabla G_L(x,b)| \le C(d,\lambda)R^{2-d-\alpha_0}$$

for some $\alpha_0 = \alpha_0(d,\lambda) > 0$, which follows as outlined at the beginning of this subsection.) We now discuss the consequences for Steps 1 and 2.

With the same argument as in Step 1, we obtain

$$\sum_{e\in\mathbb{E}^d_L} |\nabla\nabla G_L(e,b)|^2 \le \lambda^{-2}. \tag{51}$$

(Note that the torus average does not interfere since it vanishes when taking the derivative with respect to x' in (50).) Since $\nabla\nabla G_L$ converges point-wise to $\nabla\nabla G$, we may conclude by Fatou's lema.

Finally we deal with the issue in Step 2 that we do not know a priori that $\eta^2(u-\bar{u})$ is an admissible test function for $\nabla^* a\nabla u = 0$. We note that $u_L(x) = \nabla G_L(x,b)$ is characterized by

$$\sum_{e\in\mathbb{E}^d_L} \nabla\zeta(e)a(e)\nabla u_L(e) = \nabla\zeta(b). \tag{52}$$

for all L-periodic $\zeta(x)$ and we may take $\zeta = \eta^2(u_L - \bar{u}_L)$. As in Steps 2 and 3, we thus arrive at (49) with u replaced by u_L for all $C(d) \le R \le \frac{1}{C(d)}L$ and finally

$$\sum_{e\in\mathbb{E}_L^d\cap\{e:|e|\geq R\}} |\nabla u_L(e)|^2 \leq C(d,\lambda)R^{-2\alpha_0} \sum_{e\in\mathbb{E}_L^d\cap\{e:|e|\geq 2\}} |\nabla u_L(e)|^2.$$

Since $\nabla\nabla G_L$ converges point-wise to $\nabla\nabla G$ as $L\to\infty$, we can estimate the right-hand side using (51) and apply weak lower semi-continuity to take the limit as $L\to\infty$ on the left-hand side to obtain (31).

7.3 Proof of Lemma 6

Step 1. In this first step, we consider two coefficient fields $\tilde{a}, a \in \Omega$ and their associated Green functions $\tilde{G} = G(\tilde{a};\cdot,\cdot)$ and $G = G(a;\cdot,\cdot)$, respectively. We claim that if $\tilde{a}$ and a differ only at some edge $e\in\mathbb{E}^d$, then we have that:

$$\tilde{G}(x,x') - G(x,x') = (a(e)-\tilde{a}(e))\nabla\tilde{G}(x,e)\nabla G(e,x'), \tag{53}$$

$$\nabla\tilde{G}(b,x') - \nabla G(b,x') = (a(e)-\tilde{a}(e))\nabla\nabla\tilde{G}(b,e)\nabla G(e,x'), \tag{54}$$

$$\nabla\nabla\tilde{G}(b,b') - \nabla\nabla G(b,b') = (a(e)-\tilde{a}(e))\nabla\nabla\tilde{G}(b,e)\nabla\nabla G(e,b'). \tag{55}$$

Indeed, the difference satisfies the equation

$$\nabla^*\tilde{a}\nabla(\tilde{G}-G)(\cdot,x') = \nabla^*(a-\tilde{a})\nabla G(\cdot,x')$$

Since by assumption $\tilde{a}(b) = a(b)$ for all edges $b\neq e$, the Green function representation (4) immediately yields (53). Differentiating (53) then yields (54) and (55).

Step 2. In this step, we derive the following estimate on the oscillations:

$$\operatorname*{osc}_{a(e)} G(x,x') \leq 4\left(1+\tfrac{1}{\lambda}\right)|\nabla G(x,e)||\nabla G(e,x')|, \tag{56}$$

$$\operatorname*{osc}_{a(e)} \nabla G(b,x') \leq 4\left(1+\tfrac{1}{\lambda}\right)|\nabla\nabla G(b,e)||\nabla G(e,x)|, \tag{57}$$

$$\operatorname*{osc}_{a(e)} \nabla\nabla G(b,b') \leq 4\left(1+\tfrac{1}{\lambda}\right)|\nabla\nabla G(b,e)||\nabla\nabla G(e,b')|. \tag{58}$$

To do so, we first show that for any edge e, the dependence of $\nabla G(e,\cdot)$ on the value of $a(e)$ of the conductivity is mild in the sense that

$$|\nabla\tilde{G}(e,x') - \nabla G(e,x')| \leq \frac{1}{\lambda}|\nabla G(e,x')|, \tag{59}$$

$$|\nabla\nabla\tilde{G}(e,b') - \nabla\nabla G(e,b')| \leq \frac{1}{\lambda}|\nabla\nabla G(e,b')|, \tag{60}$$

where $\tilde{G}$ and G are given in Step 1. This indeed follows from letting $b=e$ in (54) and (55) and recalling the a priori estimate $|\nabla\nabla\tilde{G}(e,e)|\leq\lambda^{-1}$ from (41). We turn to the proof of (57). It is clear that for any $a\in\Omega$, there exist $\tilde{a}_1,\tilde{a}_2\in\Omega$ with $\tilde{a}_1(b) = a(b) = \tilde{a}_2(b)$ for all $b\neq e$ and associated Green functions $\tilde{G}_1$ and $\tilde{G}_2$ such that

$$\begin{aligned}\operatorname*{osc}_{a(e)} G(x,x') &\le 2|\tilde{G}_1(x,x') - \tilde{G}_2(x,x')| \\ &\le 2|\tilde{G}_1(x,x') - G(x,x')| + 2|G(x,x') - \tilde{G}_2(x,x')|.\end{aligned}$$

We insert (53) with $\tilde{a} := \tilde{a}_i$, $i = 1, 2$, into this estimate to obtain that

$$\operatorname*{osc}_{a(e)} G(x,x') \le 2|\nabla\tilde{G}_1(x,e)||\nabla G(e,x')| + 2|\nabla\tilde{G}_2(x,e)||\nabla G(e,x')|$$

Consequently, symmetry $\nabla\tilde{G}_i(x,e) = \nabla\tilde{G}_i(e,x)$ and estimate (59) yield

$$\operatorname*{osc}_{a(e)} G(x,x') \le 4\left(1 + \tfrac{1}{\lambda}\right)|\nabla G(x,e)||\nabla G(e,x')|.$$

This proves (56). The estimates (57) and (58) follow similarly using (60).

Step 3. In this step, we rephrase Lemma 5, more precisely (29), 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

$$\sup_{a\in\Omega} \sum_e \left|(|e-b|+1)^\alpha \nabla\nabla G(e,b)\right|^{2q} \le C(d,\lambda,q), \tag{61}$$

for all $q \ge 1$ and all $b \in \mathbb{E}^d$. In fact, we claim that

$$\alpha := \frac{1}{2}\alpha_0 \tag{62}$$

does the job. Because of $q \ge 1$ and thus $\ell^2(\mathbb{E}^d) \subset \ell^{2q}(\mathbb{E}^d)$, we have

$$\sum_e \left|(|e-b|+1)^\alpha \nabla\nabla G(e,b)\right|^{2q} \le \left(\sum_e \left|(|e-b|+1)^\alpha \nabla\nabla G(e,b)\right|^2\right)^q.$$

Using a dyadic decomposition, we see

$$\begin{aligned}&\sum_e \left|(|e-b|+1)^\alpha \nabla\nabla G(e,b)\right|^2 \\ &= |\nabla\nabla G(b,b)|^2 + \sum_{n=0}^\infty \sum_{e:2^{n-1}\le|e-b|<2^n} \left|(|e-b|+1)^\alpha \nabla\nabla G(e,b)\right|^2 \\ &\le |\nabla\nabla G(b,b)|^2 + \sum_{n=0}^\infty 2^{2\alpha(n+1)} \sum_{e:2^{n-1}\le|e|<2^n} |\nabla\nabla G(e,b)|^2.\end{aligned}$$

We now may appeal to (29) to obtain

$$\sum_{e}\left|(|e-b|+1)^{\alpha}\nabla\nabla G(e,b)\right|^{2}\leq C(d,\lambda)\left(1+\sum_{n=0}^{\infty}2^{2\alpha(n+1)}2^{-2\alpha_{0}n}\right)\overset{(62)}{\leq}C(d,\lambda). \tag{63}$$

Step 4. In this step, we establish the first statement of Lemma 6, namely (32). More precisely, we claim that for $p\geq\max\{\frac{d}{\alpha},1\}$ with α chosen in Step 3 and all $b,b'\in\mathbb{E}^{d}$:

$$(|b-b'|+1)^{2pd}\left\langle\left(\sum_{e}\left(\underset{a(e)}{\operatorname{osc}}\nabla\nabla G(b,b')\right)^{2}\right)^{p}\right\rangle \leq C(d,\lambda,p)\sup_{e,e'}\left\{(|e-e'|+1)^{2pd}\langle|\nabla\nabla G(e,e')|^{2p}\rangle\right\}. \tag{64}$$

Indeed, we first square (58) and sum over e:

$$\sum_{e}\left(\underset{a(e)}{\operatorname{osc}}\nabla\nabla G(b,b')\right)^{2}\leq C(\lambda)\sum_{e}|\nabla\nabla G(b,e)|^{2}|\nabla\nabla G(e,b')|^{2}.$$

After taking the p-th power, we split the sum into its contributions over $\{e:|e-b|\leq|e-b'|\}$ and $\{e:|e-b|>|e-b'|\}$ to obtain

$$\left(\sum_{e}\left(\underset{a(e)}{\operatorname{osc}}\nabla\nabla G(b,b')\right)^{2}\right)^{p} \leq C(\lambda,p)\left(\left(\sum_{e:|e-b|\leq|e-b'|}|\nabla\nabla G(b,e)|^{2}|\nabla\nabla G(e,b')|^{2}\right)^{p} + \left(\sum_{e:|e-b|\geq|e-b'|}|\nabla\nabla G(b,e)|^{2}|\nabla\nabla G(e,b')|^{2}\right)^{p}\right). \tag{65}$$

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

$$\left(\sum_{e:|e-b|\leq|e-b'|}|\nabla\nabla G(b,e)|^{2}|\nabla\nabla G(e,b')|^{2}\right)^{p} \leq\left(\sum_{e:|e-b|\leq|e-b'|}\left((|e-b|+1)^{\alpha}|\nabla\nabla G(b,e)|\right)^{2q}\right)^{p-1} \times\sum_{e:|e-b|\leq|e-b'|}\left((|e-b|+1)^{-\alpha}|\nabla\nabla G(e,b)|\right)^{2p}.$$

The first term on the right-hand side is bounded by Step 3, that is (61). After taking the expectation, we smuggle in another weight $(|e-b'|+1)^{2pd}$ and take the supremum over appropriate terms to obtain

$$\begin{aligned}&\left\langle \sum_{e:|e-b|\le|e-b'|} \big((|e-b|+1)^{-\alpha}|\nabla\nabla G(e,b')|\big)^{2p}\right\rangle\\&\quad\le \left(\sum_{e:|e-b|\le|e-b'|}(|e-b|+1)^{-2p\alpha}(|e-b'|+1)^{-2pd}\right)\\&\qquad\times\sup_{e'}\left\{(|e'-b'|+1)^{2pd}\left\langle|\nabla\nabla G(e',b')|^{2p}\right\rangle\right\}.\end{aligned}$$

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

$$\begin{aligned}&\sum_{e:|e-b|\le|e-b'|}(|e-b|+1)^{-2p\alpha}(|e-b'|+1)^{-2pd}\\&\quad\le\left(\frac{1}{2}|b-b'|+1\right)^{-2pd}\sum_e(|e-b|+1)^{-2p\alpha}.\end{aligned}$$

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

$$\sum_{e\in\mathbb{E}^d}(|e-b|+1)^{-2p\alpha}\le C(d).$$

Combining these estimates yields the bound

$$\begin{aligned}&\left\langle\left(\sum_{e:|e-b|\le|e-b'|}|\nabla\nabla G(b,e)|^2|\nabla\nabla G(e,b')|^2\right)^p\right\rangle\\&\quad\le\left(C(d,\lambda,p)(|b-b'|+1)^{-d}\sup_{e,e'}\left\{(|e-e'|+1)^d\left\langle|\nabla\nabla G(e,e')|^{2p}\right\rangle^{\frac{1}{2p}}\right\}\right)^{2p},\end{aligned}$$

i.e. the expectation of the first term on the right-hand side of (65) is bounded as desired. The second term in (65) can be dealt with exactly as the first term by simply exchanging the roles of b and b'.

Step 5. Like in Step 3, we rephrase Lemma 5, this time (30), in a way more suitable for its application in Step 6. We claim that for any integrability exponent $q\ge 1$ and any weight exponent $\beta>0$ we have

$$\sup_{a\in\Omega}\sum_e\left|(|e|+1)^{-\beta}\nabla G(e,0)\right|^{2q}\le C(d,\lambda,q,\beta) \tag{66}$$

We note that by (30) we have as soon as $\beta > 0$:

$$\sum_e |(|e|+1)^{-\beta}\nabla G(e,0)|^{2q} \le \left(\sum_e |(|e|+1)^{-\beta}\nabla G(e,0)|^2\right)^q$$
$$\le \left(\sum_{i=1}^d |\nabla G([0,e_i],0)|^2 + \sum_{n=0}^\infty 2^{-q\beta n} \sum_{e:2^n\le|e|<2^{n+1}} |\nabla G(e,0)|^2\right)^q$$
$$\overset{(30)}{\le} C(d,\lambda,\beta). \tag{67}$$

Step 6. In this step we establish the second conclusion of Lemma 6, namely (33). More precisely, we show that for any integrability exponent $p < \infty$ at least as large as in Step 3 and for any weight exponent $\beta > 0$ sufficiently small such that

$$2p(\beta-d)+d<0 \tag{68}$$

we have

$$(|b-x|+1)^{d-1}\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\nabla G(b,x)\right)^2\right)^p\right\rangle^{\frac{1}{2p}}$$
$$\le C(d,\lambda,p,\beta)\left(\sup_{e,x'}\left\{(|e-x'|+1)^{d-1}\langle|\nabla G(e,x')|^{2p}\rangle^{\frac{1}{2p}}\right\}\right.$$
$$\left.+(|b-x|+1)^{\beta-1+\frac{d}{2p}}\sup_{e,e'}\left\{(|e-e'|+1)^d\langle|\nabla\nabla G(e,e')|^{2p}\rangle^{\frac{1}{2p}}\right\}\right), \tag{69}$$

for all $x \in \mathbb{Z}^d$ and $b \in \mathbb{E}^d$, 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 (33). In order to establish (69), we first square (57) and sum over $e \in \mathbb{E}^d$ to obtain that

$$\sum_e\left(\operatorname*{osc}_{a(e)}\nabla G(b,x)\right)^2 \le C(\lambda)\sum_e|\nabla\nabla G(b,e)|^2|\nabla G(e,x)|^2.$$

We now split the sum over e:

$$\sum_e|\nabla\nabla G(b,e)|^2|\nabla G(e,x)|^2$$
$$\le C(d,\lambda)\left(\sum_{e:|e-x|\ge\frac12|b-x|}+\sum_{e:|e-x|<\frac12|b-x|}\right)|\nabla\nabla G(b,e)|^2|\nabla G(e,x)|^2.$$

Since $|e-x| < \frac{1}{2}|b-x|$ implies $|e-b| > \frac{1}{2}|b-x|$, it follows

$$\begin{aligned}
&\sum_e |\nabla\nabla G(b,e)|^2 |\nabla G(e,x)|^2 \\
&\le C(d,\lambda)\left(\sum_{e:|e-x|\ge\frac{1}{2}|b-x|} |\nabla\nabla G(b,e)|^2 |\nabla G(e,x)|^2\right. \\
&\quad \left. + \sum_{e:|e-b|>\frac{1}{2}|b-x|} |\nabla G(e,x)|^2 |\nabla\nabla G(b,e)|^2 \right). \qquad (70)
\end{aligned}$$

We start by treating the first term on the r.-h. s. of (70) in an analogous way to Step 4. For that purpose, let α be as in Step 3. We smuggle in the weight $(|e-b|+1)^\alpha$ and apply Hölder's inequality with p and q such that $\frac{1}{p}+\frac{1}{q}=1$:

$$\begin{aligned}
&\left(\sum_{e:|e-x|\ge\frac{1}{2}|b-x|} |\nabla\nabla G(b,e)|^2 |\nabla G(e,x)|^2\right)^p \\
&\le \left(\sum_e \left|(|e-b|+1)^\alpha \nabla\nabla G(b,e)\right|^{2q}\right)^{p-1} \\
&\quad \times \sum_{e:|e-x|\ge\frac{1}{2}|b-x|} \left|(|e-b|+1)^{-\alpha}\nabla G(e,x)\right|^{2p}.
\end{aligned}$$

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 $(|e-x|+1)^{2p(d-1)}$ to obtain as desired:

$$\begin{aligned}
&\left\langle\left(\sum_{e:|e-x|\ge\frac{1}{2}|b-x|} |\nabla\nabla G(b,e)|^2 |\nabla G(e,x)|^2\right)^p\right\rangle \\
&\le \left(\sum_{e:|e-x|\ge\frac{1}{2}|b-x|} (|e-b|+1)^{-2p\alpha}(|e-x|+1)^{-2p(d-1)}\right. \\
&\quad \left. \times \sup_{e'}\left\{(|e'-x|+1)^{2p(d-1)}\langle|\nabla G(e',x)|^{2p}\rangle\right\}\right) \\
&\overset{(61)}{\le} C(d,\lambda,p)\,(|b-x|+1)^{-2p(d-1)} \sup_{e'}\left\{(|e'-x|+1)^{2p(d-1)}\langle|\nabla G(e',x)|^{2p}\rangle\right\}, \qquad (71)
\end{aligned}$$

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

We now address the second term on the r.-h. s. of (70) in a similar way, just exchanging the roles of ∇G and $\nabla\nabla G$, of b and x, and of α and $-\beta$, where the weight exponent $\beta > 0$ needs to satisfy (68). By Hölder's inequality we obtain:

$$\left(\sum_{e:|e-b|\geq\frac{1}{2}|b-x|} |\nabla G(e,x)|^2 |\nabla\nabla G(b,e)|^2\right)^p$$
$$\leq \left(\sum_e \left|(|e-x|+1)^{-\beta}\nabla G(e,x)\right|^{2q}\right)^{p-1}$$
$$\times \sum_{e:|e-b|\geq\frac{1}{2}|b-x|} |(|e-x|+1)^{\beta}\nabla\nabla G(b,e)|^{2p}.$$

The first term is bounded by Step 5 in form of (66).
Taking the expectation and smuggling in a weight $(|e-b|+1)^{2pd}$ yields

$$\left\langle \sum_{e:|e-b|\geq\frac{1}{2}|b-x|} |(|e-x|+1)^{\beta}\nabla\nabla G(b,e)|^{2p} \right\rangle$$
$$\leq \sum_{e:|e-b|\geq\frac{1}{2}|b-x|} (|e-x|+1)^{2p\beta}(|e-b|+1)^{-2pd}$$
$$\times \sup_{e'} \left\{(|e'-b|+1)^{2pd}\langle|\nabla\nabla G(b,e')|^{2p}\rangle\right\}.$$

We note that by the triangle inequality in form of $|e-x| \leq |e-b| + |b-x|$, in the range (68) the remaining sum is bounded as follows:

$$\sum_{e:|e-b|\geq\frac{1}{2}|b-x|} (|e-x|+1)^{2p\beta}(|e-b|+1)^{-2pd}$$
$$\leq C(p,\beta)\Bigg((|b-x|+1)^{2p\beta} \sum_{e:|e-b|\geq\frac{1}{2}|b-x|} (|e-b|+1)^{-2pd}$$
$$+ \sum_{e:|e-b|\geq\frac{1}{2}|b-x|} (|e-b|+1)^{2p(\beta-d)}\Bigg)$$
$$\leq C(d,p,\beta)(|b-x|+1)^{2p(\beta-d)+d}.$$

Hence we have obtained

$$\left\langle\left(\sum_{e:|e-b|\geq\frac{1}{2}|b-x|}|\nabla G(e,x)|^2|\nabla\nabla G(b,e)|^2\right)^p\right\rangle$$
$$\leq C(d,\lambda,p,\beta)(|b-x|+1)^{2p(\beta-d)+d}\sup_{e,e'}\left\{(|e-e'|+1)^{2pd}\langle|\nabla\nabla G(e,e')|^{2p}\rangle\right\}. \tag{72}$$

In view of (70), the combination of (71) and (72) as well as taking the $2p$-th root yields (69).

7.4 Proof of Theorem 1

We start with the proof of (8). For this purpose, we fix $b, b' \in \mathbb{E}^d$ and $p < \infty$; by Jensen's inequality, we may assume that $p \geq p_0$ with p_0 from Lemma 6. Applying Lemma 4 to $\zeta(a) = \nabla\nabla G(a; b, b')$ and inserting the estimate (32) of Lemma 6 yields (after redefining δ)

$$(|b-b'|+1)^d\langle|\nabla\nabla G(b,b')|^{2p}\rangle^{\frac{1}{2p}} \leq C(d,\lambda,\rho,p,\delta)(|b-b'|+1)^d\langle|\nabla\nabla G(b,b')|\rangle$$
$$+\delta\sup_{e,e'}\left\{(|e-e'|+1)^d\langle|\nabla\nabla G(e,e')|^{2p}\rangle^{\frac{1}{2p}}\right\}.$$

We now insert (22) and take the supremum over b and b':

$$\sup_{b,b'}\left\{(|b-b'|+1)^d\langle|\nabla\nabla G(b,b')|^{2p}\rangle^{\frac{1}{2p}}\right\}$$
$$\leq C(d,\lambda,\rho,p,\delta)+\delta\sup_{e,e'}\left((|e-e'|+1)^d\langle|\nabla\nabla G(e,e')|^{2p}\rangle^{\frac{1}{2p}}\right).$$

Choosing $\delta = 1/2$, we obtain (8). We deal with the objection that $\sup_{e,b}\{(|e-b|+1)^d\langle|\nabla\nabla G(e,b)|^{2p}\rangle^{1/(2p)}\}$ may be infinite by first working with the periodic Green function G_L as in the proof of Lemma 5 and then letting $L \to \infty$.

We now turn to the proof of (9). With help of the just established (8), we may upgrade the result of Lemma 6, cf. (33), to

$$(|b-x|+1)^{d-1}\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}\nabla G(b,x)\right)^2\right)^p\right\rangle^{\frac{1}{2p}}$$
$$\leq C(d,\lambda,\rho,p)\left(\sup_{e,x'}\left\{(|e-x'|+1)^{d-1}\langle|\nabla G(e,x')|^{2p}\rangle^{\frac{1}{2p}}\right\}+1\right). \tag{73}$$

We apply Lemma 4 to $\zeta = \nabla G(b,x)$ and insert (73) (after redefining δ):

$$
\begin{aligned}
&(|b-x|+1)^{d-1}\langle|\nabla G(b,x)|^{2p}\rangle^{\frac{1}{2p}}\\
&\quad\le C(d,\lambda,\rho,p,\delta)(|b-x|+1)^{d-1}\langle|\nabla G(b,x)|\rangle\\
&\qquad\quad+\delta\left(\sup_{e,x'}\left\{(|e-x'|+1)^{d-1}\langle|\nabla G(e,x')|^{2p}\rangle^{\frac{1}{2p}}\right\}+1\right).
\end{aligned}
$$

We now insert (23) and take the supremum over b and x:

$$
\begin{aligned}
&\sup_{b,x}\left((|b-x|+1)^{d-1}\langle|\nabla G(b,x)|^{2p}\rangle^{\frac{1}{2p}}\right)\\
&\quad\le C(d,\lambda,\rho,p,\delta)+\delta\sup_{e,x'}\left\{(|e-x'|+1)^{d-1}\langle|\nabla G(e,x')|^{2p}\rangle^{\frac{1}{2p}}\right\}.
\end{aligned}
$$

As before, letting $\delta=1/2$ yields (9).

8 Proofs of the corollaries

8.1 Proof of Corollary 1

It is well known that an LSI implies a corresponding SG, see for instance [18, Theorem 4.9]. Indeed, using $\zeta^2=1+\epsilon f$ for some $f(a)$ in (7) and expanding to second order in $\epsilon\ll 1$ one obtains

$$
\langle(f-\langle f\rangle)^2\rangle\le\frac{1}{\rho}\left\langle\sum_e\left(\operatorname*{osc}_{a(e)}f\right)^2\right\rangle.
$$

As in Step 2 of the proof of Lemma 4, see also [14, Lemma 11], it follows that

$$
\langle|f-\langle f\rangle|^{2p}\rangle\le C(\rho,p)\left\langle\left(\sum_e\left(\operatorname*{osc}_{a(e)}f\right)^2\right)^p\right\rangle. \tag{74}
$$

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 6, i.e.

$$
\operatorname*{osc}_{a(e)}G(x,0)\le C(\lambda)|\nabla G(x,e)||\nabla G(e,0)|,
$$

to obtain

$$
\left\langle\left|G(x,0)-\langle G(x,0)\rangle\right|^{2p}\right\rangle^{\frac{1}{p}}\le C(\lambda,\rho,p)\left\langle\left(\sum_e|\nabla G(x,e)|^2|\nabla G(e,0)|^2\right)^p\right\rangle^{\frac{1}{p}}.
$$

The triangle inequality in $\langle(\cdot)^p\rangle^{1/p}$ yields

$$\left\langle\left|G(x,0)-\langle G(x,0)\rangle\right|^{2p}\right\rangle^{\frac{1}{p}} \le C(\lambda,\rho,p)\sum_e \left\langle|\nabla G(x,e)|^{2p}|\nabla G(e,0)|^{2p}\right\rangle^{\frac{1}{p}}.$$

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

$$\begin{aligned}
&\left\langle\left|G(x,0)-\langle G(x,0)\rangle\right|^{2p}\right\rangle^{\frac{1}{p}}\\
&\quad\le C(\lambda,\rho,p)\sum_e \langle|\nabla G(x,e)|^{4p}\rangle^{\frac{1}{2p}}\langle|\nabla G(e,0)|^{4p}\rangle^{\frac{1}{2p}}\\
&\quad= C(\lambda,\rho,p)\sum_e \langle|\nabla G(e-x,0)|^{4p}\rangle^{\frac{1}{2p}}\langle|\nabla G(e,0)|^{4p}\rangle^{\frac{1}{2p}},
\end{aligned}$$

where we recall that $e-x\in\mathbb{E}^d$ is the edge e shifted by x and ∇ always falls on the edge variable. Into this estimate, we insert the result of Theorem 1:

$$\left\langle\left|G(x,0)-\langle G(x,0)\rangle\right|^{2p}\right\rangle^{\frac{1}{p}} \le C(d,\lambda,\rho,p)\sum_e ((|e-x|+1)(|e|+1))^{2(1-d)}. \tag{75}$$

We now turn to the sum on the r.-h. s. of (75): By symmetry, we have

$$\sum_e ((|e-x|+1)(|e|+1))^{2(1-d)} \le 2\sum_{e:|e-x|\le|e|}((|e-x|+1)(|e|+1))^{2(1-d)}. \tag{76}$$

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

$$\sum_e (|e|+1)^{2(1-d)} \le C(d)<\infty. \tag{77}$$

Since $|e-x|\le|e|$ implies $|e|\ge\frac{1}{2}|x|$ we thus have as desired for (76)

$$\begin{aligned}
&\sum_{e:|e-x|\le|e|}(|e-x|+1)^{2(1-d)}(|e|+1)^{2(1-d)}\\
&\quad\le (\tfrac{1}{2}|x|+1)^{2(1-d)}\sum_e(|e-x|+1)^{2(1-d)}\\
&\quad\overset{(77)}{\le} C(d)(|x|+1)^{2(1-d)}.
\end{aligned} \tag{78}$$

We now turn to the case of $d=2$. In this case, we split the sum on the r. h. s. of (76) according to

$$\sum_{e:|e-x|\le|e|} = \sum_{e:|e-x|\le|e| \text{ and } |e|\ge 2|x|} + \sum_{e:|e-x|\le|e| \text{ and } |e|<2|x|}$$
$$\le \sum_{e:|e-x|\ge\frac{1}{2}|e| \text{ and } |e|\ge 2|x|} + \sum_{e:|e-x|\le 2|x| \text{ and } |e|\ge\frac{1}{2}|x|},$$

so that

$$\sum_{e:|e-x|\le|e|} (|e-x|+1)^{-2}(|e|+1)^{-2}$$
$$\le \sum_{e:|e|\ge 2|x|} (\tfrac{1}{2}|e|+1)^{-4} + (\tfrac{1}{2}|x|+1)^{-2} \sum_{e:|e-x|\le 2|x|} (|e-x|+1)^{-2}$$
$$\le C(|x|+1)^{-2} + C(|x|+1)^{-2}\log(|x|+2). \tag{79}$$

Combining (78) and (79), we gather

$$\sum_{e}(|e-x|+1)^{2(1-d)}(|e|+1)^{2(1-d)}$$
$$\le C(d)(|x|+1)^{2(1-d)} \begin{Bmatrix} 1 & \text{for } d>2 \\ \log(|x|+2) & \text{for } d=2 \end{Bmatrix}, \tag{80}$$

which we insert into (75) to obtain (10).

8.2 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 function satisfies $\nabla^* a \nabla G(\cdot, x') = \delta(\cdot - x')$. Now let $a(e) = 1 + \epsilon\tilde{a}(e)$ for $\epsilon \ll 1$, where $\tilde{a}$ is i. i. d. with values at each edge taken in $[-1, 1]$. Furthermore we assume $\langle \tilde{a}(e)\rangle = 0$ as well as $\langle \tilde{a}(e)^2\rangle = 1$. 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 function corresponding to a in powers of ϵ:

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

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

$$\nabla^*\nabla G_0(\cdot, x') = \delta(\cdot - x'),$$

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

$$\nabla^*\nabla G_1(\cdot, x') + \nabla^*\tilde{a}\nabla G_0(\cdot, x') = 0.$$

Hence we have that

$$G_1(x,x') = -\sum_e \nabla G_0(x,e)\tilde{a}(e)\nabla G_0(e,x').$$

Since $\langle \tilde{a}(e)\rangle = 0$, we deduce $\langle G_1\rangle = 0$ and consequently

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

Since the coefficients $\tilde{a}(x)$ are i. i. d. with variance 1, it follows

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

The behavior of the constant-coefficient Green function G_0 is well-known, cf. [19, Theorem 4.3.1], and yields that $(\nabla G_0(e,0))^2$ scales like $(|e|+1)^{1-d}$ with a similar expression for $(\nabla G_0(x,e))^2$. Hence we find that

$$\begin{aligned}&\langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle \le C(d)\sum_e ((|e-x|+1)(|e|+1))^{2(1-d)} \quad \text{and} \\ &\langle (G_1(x,0) - \langle G_1(x,0)\rangle)^2\rangle \ge \frac{1}{C(d)}\sum_e ((|e-x|+1)(|e|+1))^{2(1-d)}. \end{aligned} \tag{81}$$

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

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

If $d>2$, a lower bound can be obtained by considering only the summand $e=[0,e_i]$ in (81). If $d=2$, we restrict the sum to all e such that $|e|\le|x|$ and use $|e-x|\le 2|x|$ in that region to obtain

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

Thus Corollary 1 is indeed optimal in scaling.

8.3 Proof of Corollaries 2 and 3

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

$$\left\langle \left(\sum_x \left| u(x) - \langle u(x) \rangle \right|^p \right)^r \right\rangle^{\frac{1}{rp}} \leq \left(\sum_x \left\langle \left| u(x) - \langle u(x) \rangle \right|^{rp} \right\rangle^{\frac{1}{r}} \right)^{\frac{1}{p}}. \tag{82}$$

Since u is the decaying solution of (11) with compactly supported right-hand side f, it can be represented via the Green function:

$$u(x) = \sum_y G(x, y) f(y). \tag{83}$$

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

$$\left\langle \left| u(x) - \langle u(x) \rangle \right|^{rp} \right\rangle^{\frac{1}{rp}} \leq C(d, \lambda, \rho, r, p) \sum_y (|x - y| + 1)^{1-d} |f(y)|. \tag{84}$$

We now insert (84) in (82) to obtain

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

Now let us recall the Hardy-Littlewood-Sobolev inequality in $\mathbb{R}^d$, see [21, 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 constant functions. We use the discrete version for $\alpha = d - 1$, that is,

$$\left(\sum_x\left(\sum_y(|x-y|+1)^{1-d}|f(y)|\right)^p\right)^{\frac{1}{p}} \leq C(d,p)\left(\sum_y|f(y)|^q\right)^{\frac{1}{q}}, \quad (86)$$

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

Step 2. Changes if $d=2$. In this case, using that $f(y)$ is supported in $\{y:|y|\leq R\}$, (85) assumes the form

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

As in Step 1, it remains to apply the discrete Hardy-Littlewood-Sobolev inequality, where $p>2$ and $q<2$ are admissible.

Proof of Corollary 3
Step 1. In this step, we derive the estimate

$$\begin{aligned}
&\left\langle\left|\sum_x(u(x)-\langle u(x)\rangle)g(x)\right|^r\right\rangle^{\frac{1}{r}}\\
&\quad\leq C(\rho,r)\left\langle\left(\sum_e\left(\sum_x\sum_y\left(\operatorname*{osc}_{a(e)}G(x,y)\right)|f(y)||g(x)|\right)^2\right)^{\frac{r}{2}}\right\rangle^{\frac{1}{r}}. \quad (87)
\end{aligned}$$

Indeed, it follows from the representation (83) that

$$\begin{aligned}
&\left\langle\left|\sum_x(u(x)-\langle u(x)\rangle)\,g(x)\right|^r\right\rangle^{\frac{1}{r}}\\
&\quad=\left\langle\left|\sum_x\sum_y(G(x,y)-\langle G(x,y)\rangle)\,f(y)g(x)\right|^r\right\rangle^{\frac{1}{r}}.
\end{aligned}$$

Hence the L^p-version of SG (74), with $2p$ replaced by r (w. l. o. g. we may assume $r\geq 2$), yields

$$\left\langle \left| \sum_x (u(x) - \langle u(x) \rangle)\, g(x) \right|^r \right\rangle^{\frac{1}{r}}$$
$$\leq C(\rho, r) \left\langle \left(\sum_e \left(\operatorname*{osc}_{a(e)} \sum_x \sum_y G(x, y) f(y) g(x) \right)^2 \right)^{\frac{r}{2}} \right\rangle^{\frac{1}{r}}.$$

Since the only dependence on the coefficients a is through G, we may use sub-linearity of the oscillation to obtain (87).

Step 2. In this step, we estimate the right-hand side of (87) as follows:

$$\left\langle \left(\sum_e \left(\sum_x \sum_y \left(\operatorname*{osc}_{a(e)} G(x, y) \right) |f(y)||g(x)| \right)^2 \right)^{\frac{r}{2}} \right\rangle^{\frac{1}{r}}$$
$$\leq C(d, \lambda, \rho, r) \left(\sum_e \left(\sum_x (|e - x| + 1)^{1-d} |g(x)| \right)^2 \right.$$
$$\left. \times \left(\sum_y (|e - y| + 1)^{1-d} |f(y)| \right)^2 \right)^{\frac{1}{2}}. \tag{88}$$

Indeed, the triangle inequality with respect to $\langle | \cdot |^{r/2} \rangle^{2/r}$ allows us to take out the sum over e as follows:

$$\left\langle \left(\sum_e \left(\sum_x \sum_y \left(\operatorname*{osc}_{a(e)} G(x, y) \right) |f(y)||g(x)| \right)^2 \right)^{\frac{r}{2}} \right\rangle^{\frac{1}{r}}$$
$$\leq \left(\sum_e \left\langle \left(\sum_x \sum_y \left(\operatorname*{osc}_{a(e)} G(x, y) \right) |f(y)||g(x)| \right)^r \right\rangle^{\frac{2}{r}} \right)^{\frac{1}{2}}.$$

Now, the triangle inequality with respect to $\langle | \cdot |^r \rangle^{1/r}$ allows us to take out the sum over x and y and we obtain that

$$\left(\sum_e \left\langle \left(\sum_x \sum_y \left(\operatorname*{osc}_{a(e)} G(x, y) \right) |f(y)||g(x)| \right)^r \right\rangle^{\frac{2}{r}} \right)^{\frac{1}{2}}$$
$$\leq \left(\sum_e \left(\sum_x \sum_y \left\langle \left(\operatorname*{osc}_{a(e)} G(x, y) \right)^r \right\rangle^{\frac{1}{r}} |f(y)||g(x)| \right)^2 \right)^{\frac{1}{2}}.$$

Theorem 1 and the oscillation estimate (56) yield by Hölder's inequality that

$$\left\langle \left(\operatorname*{osc}_{a(e)} G(x, y) \right)^r \right\rangle^{\frac{1}{r}} \leq C(d, \lambda, \rho, r)(|e - x| + 1)^{1-d} (|e - y| + 1)^{1-d},$$

which completes this step since we may now write the sum over x and y as a product.

Step 3. Conclusion. An application of Hölder's inequality w. r. t. the sum over e on the r.-h. s. of (88) yields a bound by

$$\Big(\sum_e\Big(\sum_x(|e-x|+1)^{1-d}|g(x)|\Big)^{\tilde p}\Big)^{\frac{1}{\tilde p}} \times\Big(\sum_e\Big(\sum_y(|e-y|+1)^{1-d}|f(y)|\Big)^{p}\Big)^{\frac{1}{p}}, \tag{89}$$

with $p,\tilde p>2$ such that $\frac{2}{p}+\frac{2}{\tilde p}=1$ to be chosen later. We recall the Hardy-Littlewood-Sobolev inequality (86), i.e.

$$\Big(\sum_e\Big(\sum_x(|e-x|+1)^{1-d}|f(x)|\Big)^{p}\Big)^{\frac{1}{p}}\le C(d,q)\Big(\sum_x|f(x)|^q\Big)^{\frac{1}{q}},$$

if we choose p such that $\frac{1}{q}=\frac{1}{d}+\frac{1}{p}$. (Here we require $\frac{2d}{d+2}<q<d$ so that in particular $2<p<\infty$.) The Hardy-Littlewood-Sobolev inequality likewise yields

$$\Big(\sum_e\Big(\sum_x(|e-x|+1)^{1-d}|g(x)|\Big)^{\tilde p}\Big)^{\frac{1}{\tilde p}}\le C(d,q)\Big(\sum_x|g(x)|^{\tilde q}\Big)^{\frac{1}{\tilde q}},$$

where $\frac{1}{\tilde q}=\frac{1}{d}+\frac{1}{\tilde p}=\frac{1}{d}+\frac{1}{2}-\frac{1}{p}$ and we require $\frac{2d}{d+2}<\tilde q<d$. Inserting these estimates into (89) and then into Steps 2 and 1 yields Corollary 3.

8.4 Proof of Corollary 4

Step 1. Let u satisfy $\nabla^*a\nabla u=0$ in $\{x:|x|\le R\}$. We claim that for any function $\eta(x)$ supported in $\{x:|x|<R\}$, we obtain the representation

$$(\eta u)(x)=\sum_{e=[y,y']:|e|\le R}\big(u(y)\nabla G(e,x)a(e)\nabla\eta(e)-G(y,x)\nabla\eta(e)a(e)\nabla u(e)\big)\,, \tag{90}$$

where we sum over all edges in $\mathbb{E}^d$ of the form $[y,y']$ such that their midpoint is of distance at most R from the origin. We start by noting that even on the discrete level, some aspects of Leibniz rule survive, such as

$$\begin{aligned}&\nabla\zeta(e)a(e)\nabla(\eta u)(e)-\nabla(\eta\zeta)(e)a(e)\nabla u(e)\\&\quad=u(y)\nabla\zeta(e)a(e)\nabla\eta(e)-\zeta(y)\nabla\eta(e)a(e)\nabla u(e)\end{aligned} \tag{91}$$

for any function $\zeta:\mathbb{Z}^d\to\mathbb{R}$ and $e=[y,y']\in\mathbb{E}^d$. Indeed, (91) 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 (91):

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

and use it for $\zeta = G(\cdot, x)$. By definition of G, the first term on the l.-h. s. of (92) yields $(\eta u)(x)$. Since $\eta G(\cdot, x)$ is supported in $\{y : |y| \le R\}$, the second term on the l.-h. s. of (92) vanishes. This completes the step.

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

$$\begin{aligned}
&\left(\frac{\sup_{x:|x|\le\frac{R}{8}} \frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha}\sup_{x:|x|\le R}|u(x)|}\right)^p \\
&\quad\le C(d,\lambda,p) R^{\alpha p} R^{-p}\Big(R^{d(p-1)} \sum_{e:|e|\le\frac{R}{8}}\ \sum_{b:\frac{R}{4}\le|b|\le\frac{R}{2}} |\nabla\nabla G(e,b)|^p \\
&\qquad + R^{d(p-1)-p} \sum_{e:|e|\le\frac{R}{8}}\ \sum_{x:\frac{R}{4}\le|x|\le\frac{R}{2}} |\nabla G(e,x)|^p\Big),
\end{aligned} \tag{93}$$

if $\alpha < 1$ and $p > d$ are related by $\alpha p = p - d$. To this end, we choose a cut-off function η for $\{x : |x| \le \frac{R}{4} + 1\}$ in $\{x : |x| \le \frac{R}{2} - 1\}$ (w. l. o. g. $R > 8$). We restrict to $|x| \le \frac{R}{4}$ and take the derivative of (90) along the edge b to obtain

$$\nabla u(b) = \sum_{e=[y,y']\in\mathbb{E}^d} \left(u(y)\nabla\nabla G(e,b)a(e)\nabla\eta(e) - (\nabla\eta(e)a(e)\nabla u(e))\nabla G(y,b)\right).$$

This implies

$$|\nabla u(b)| \le \frac{C(d)}{R} \sum_{e=[y,y']:\frac{R}{4}\le|y|,|y'|\le\frac{R}{2}} \left(|u(y)||\nabla\nabla G(e,b)| + |\nabla G(y,b)||\nabla u(e)|\right). \tag{94}$$

Applying Hölder's inequality and summing the p-th power of (94), we obtain

$$\begin{aligned}
&\sum_{b:|b|\le\frac{R}{8}} |\nabla u(b)|^p \\
&\quad\le C(d,p) R^{-p}\left(\left(\sum_{y:|y|\le\frac{R}{2}} |u(y)|^q\right)^{p-1} \sum_{b:|b|\le\frac{R}{8}}\ \sum_{e:\frac{R}{4}\le|e|\le\frac{R}{2}} |\nabla\nabla G(e,b)|^p \right.\\
&\qquad \left. + \left(\sum_{e:|e|\le\frac{R}{2}} |\nabla u(e)|^q\right)^{p-1} \sum_{b:|b|\le\frac{R}{8}}\ \sum_{y:\frac{R}{4}\le|y|\le\frac{R}{2}} |\nabla G(y,b)|^p\right),
\end{aligned} \tag{95}$$

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

$$\alpha = 1 - \frac{d}{p}, \tag{96}$$

then we have that

$$\sup_{x:|x|\le\frac{R}{8}} \frac{|u(x)-u(0)|}{|x|^\alpha} \le C(d,p)\left(\sum_{b:|b|\le\frac{R}{8}} |\nabla u(b)|^p\right)^{\frac{1}{p}}. \tag{97}$$

(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 (95) bounds the α-Hölder norm as desired, albeit over a smaller ball.

Let us now turn to the right-hand side of (95). We trivially have that

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

To estimate the second summand on the right-hand side, we note that Caccioppoli's estimate (49) implies

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

Together with Jensen's inequality (here we need $q \le 2$, that is $p \ge 2$, which is obvious since even $p > d$ from (96)), we obtain that

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

Substituting (98) and (99) into (95) yields the claim of this step.

Step 3. Using (97) and bounding the Green function, we conclude that

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

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 $\{x : |x| \leq R\}$. Indeed, Theorem 1 applied to the result (93) of Step 2 yields

$$\left\langle \left(\sup_u \frac{\sup_{x:|x|\leq \frac{R}{8}} \frac{|u(x)-u(0)|}{|x|^\alpha}}{\frac{1}{R^\alpha} \sup_{x:|x|\leq R} |u(x)|} \right)^p \right\rangle \leq C(d, \lambda, \rho, p) R^{\alpha p} \Bigg(R^{d(p-1)-p} \sum_{e:|e|\leq \frac{R}{8}} \sum_{b:\frac{R}{4}\leq |b| \leq \frac{R}{2}} (|e-b|+1)^{-pd} + R^{d(p-1)-2p} \sum_{e:|e|\leq \frac{R}{8}} \sum_{x:\frac{R}{4}\leq |x| \leq \frac{R}{2}} (|e-x|+1)^{p(1-d)} \Bigg)$$

if α and p are related by (96). In the domains of e and b, we have $|e-b|+1 \geq |b|-|e| \geq R/8$. 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 (96) holds, we thus 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 (96) holds. Since in (96), $\alpha \to 1$ 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 $\alpha < 1$.

8.5 Proof of Lemmas 2 and 3

Lemma 1 is a direct consequence of Lemmas 2 and 3.

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

$$\langle \zeta^2 \log \zeta^2 \rangle = \langle \zeta^2 \log \zeta^2 - \zeta^2 + 1 \rangle \leq \langle (\zeta^2 - 1)^2 \rangle.$$

Since $(\zeta^2-1)^2=(|\zeta|-1)^2(|\zeta|+1)^2$, we find that

$$\langle \zeta^2 \log \zeta^2\rangle \le \langle (|\zeta|+1)^2\rangle \sup_a (|\zeta|-1)^2.$$

Since $\langle \zeta^2\rangle = 1$, there exists $a_* \in [\lambda, 1]$ such that $|\zeta(a_*)| \le 1$. It follows that

$$|\zeta(a)|-1 \le |\zeta(a)|-|\zeta(a_*)| \le |\zeta(a)-\zeta(a_*)| \le \operatorname*{osc}_a \zeta(a).$$

Likewise there exists $a^* \in [\lambda, 1]$ such that $|\zeta(a^*)| \ge 1$ and therefore

$$1-|\zeta(a)| \le |\zeta(a^*)|-|\zeta(a)| \le |\zeta(a^*)-\zeta(a)| \le \operatorname*{osc}_a \zeta(a).$$

Hence it follows that

$$\langle \zeta^2 \log \zeta^2\rangle \le \langle (|\zeta|+1)^2\rangle \left(\operatorname*{osc}_a \zeta\right)^2.$$

Finally we have that

$$\langle (|\zeta|+1)^2\rangle \le \langle 2\zeta^2+2\rangle = 4,$$

and the combination of the previous two inequalities yields (21) with constant $\rho = \frac{1}{8}$.

Proof of Lemma 3 The following is a simple adaptation of the usual tensorization proof, cf. [18, Theorem 4.4]. Take any enumeration $(e_n)_{n\ge 1}$ of the edge set $\mathbb{E}^d$ and denote by $\langle\cdot\rangle_n$ the e_n-marginal of the (product) ensemble $\langle\cdot\rangle$. We assume that every marginal satisfies the LSI

$$\left\langle \zeta^2 \log \frac{\zeta^2}{\langle \zeta^2\rangle_n}\right\rangle_n \le \frac{2}{\rho}\left(\operatorname*{osc}_{a\in[\lambda,1]} \zeta\right)^2$$

for all $\zeta : [\lambda, 1] \to \mathbb{R}$. Replacing ζ^2 by f in the definition of the LSI, it suffices to prove

$$\left\langle f \log \frac{f}{\langle f\rangle}\right\rangle \le \frac{2}{\rho}\sum_{n=1}^{\infty}\left\langle \left(\operatorname*{osc}_{a(e_n)} \sqrt{f}\right)^2\right\rangle$$

for all positive random variables $f : \Omega \to (0,\infty)$. By a simple density argument, it suffices to consider *local* random variables, i.e. f that depend on a only through a finite number of sites so that the above sum is finite. 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 edges. Then the l.-h. s. of (7) can be expressed as a telescope sum (a finite sum for local random variables):

$$\langle f\log f\rangle-\langle f\rangle\log\langle f\rangle=\sum_{n=1}^{\infty}\big\langle f_{n-1}\log f_{n-1}-f_n\log f_n\big\rangle$$
$$=\sum_{n=1}^{\infty}\big\langle\langle f_{n-1}\log f_{n-1}\rangle_n-\langle f_{n-1}\rangle_n\log\langle f_{n-1}\rangle_n\big\rangle. \quad (101)$$

The assumption of single-edge LSI yields

$$\langle f_{n-1}\log f_{n-1}\rangle_n-\langle f_{n-1}\rangle_n\log\langle f_{n-1}\rangle_n\le\frac{2}{\rho}\big(\operatorname*{osc}_{a(e_n)}\sqrt{f_{n-1}}\big)^2. \quad (102)$$

Notice that the definition of f_{n-1} immediately yields $f_{n-1}=\langle f\rangle_{<n}$, where we have abbreviated the ensemble average over the first $n-1$ edges as $\langle\cdot\rangle_{<n}$. We clearly have

$$\operatorname*{osc}_{a(e_n)}\sqrt{f_{n-1}}=\left(\sup_{a(e_n)}\langle f\rangle_{<n}\right)^{\frac12}-\left(\inf_{a(e_n)}\langle f\rangle_{<n}\right)^{\frac12}$$
$$\le\left\langle\sup_{a(e_n)}f\right\rangle_{<n}^{\frac12}-\left\langle\inf_{a(e_n)}f\right\rangle_{<n}\Big)^{\frac12}.$$

By monotonicity of the square root, it follows

$$\operatorname*{osc}_{a(e_n)}\sqrt{f_{n-1}}\le\left\langle\left(\sup_{a(e_n)}\sqrt{f}\right)^2\right\rangle_{<n}^{\frac12}-\left\langle\left(\inf_{a(e_n)}\sqrt{f}\right)^2\right\rangle_{<n}^{\frac12}.$$

Consequently the triangle inequality w. r. t. $\langle(\cdot)^2\rangle_{<n}^{\frac12}$ on the right-hand side yields

$$\operatorname*{osc}_{a(e_n)}\sqrt{f_{n-1}}\le\left\langle\left(\sup_{a(e_n)}\sqrt{f}-\inf_{a(e_n)}\sqrt{f}\right)^2\right\rangle_{<n}^{\frac12},$$

which by definition of $\operatorname*{osc}_{a(e_n)}$ can be written as

$$\operatorname*{osc}_{a(e_n)}\sqrt{f_{n-1}}\le\left\langle\left(\operatorname*{osc}_{a(e_n)}\sqrt{f}\right)^2\right\rangle_{<n}^{\frac12}. \quad (103)$$

Finally we collect (101), (102) and (103) to obtain

$$\langle f\log f\rangle-\langle f\rangle\log\langle f\rangle\le\frac{1}{2\rho}\sum_{n=1}^{\infty}\left\langle\left(\operatorname*{osc}_{a(e_n)}\sqrt{f}\right)^2\right\rangle,$$

which is the LSI (7) for $f=\zeta^2$.

References

1. Astala, K., Iwaniec, T., Martin, G.: Elliptic partial differential equations and quasiconformal mappings in the plane, Princeton Mathematical Series 48. Princeton University Press, Princeton (2009)
2. Benjamini, I., Duminil-Copin, H., Kozma, G., Yadin, A.: Disorder, entropy and harmonic functions, Preprint (2011). arXiv:1111.4853
3. Biskup, M., Salvi, M., Wolff, T.: A central limit theorem for the effective conductance: I. Linear boundary data and small ellipticity contrasts, Preprint (2012). arXiv:1210.2371
4. Carlen, E.A., Kusuoka, S., Stroock, D.W.: Upper bounds for symmetric Markov transition functions. Ann. Inst. H. Poincaré Probab. Statist. **23**(2), 245–287 (1987)
5. Conlon, J.C., Giunti, A., Otto, F.: Green's function for elliptic systems: Delmotte–Deuschel bounds, in preparation
6. Conlon, J.C., Naddaf, A.: On homogenization of elliptic equations with random coefficients. Electron. J. Probab. **9**(5), 1–58 (2000)
7. De Giorgi, E.: 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)
8. Delmotte, T.: Inégalité de Harnack elliptique sur les graphes. Colloq. Math. **72**(1), 19–37 (1997)
9. Delmotte, T., Deuschel, J.-D.: 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)
10. Dolzmann, G., Hungerbühler, N., Müller, S.: 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)
11. Federbush, P.: Partially alternate derivation of a result by Nelson. J. Math. Phys. **10**(1), 50–52 (1969)
12. Gloria, A.: Fluctuation of Solutions to Linear Elliptic Equations with Noisy Diffusion Coefficients. Comm. PDE **38**(2), 304–338 (2012)
13. Gloria, A., Marahrens, D.: Improved De Giorgi-Nash-Moser theory in the large, annealed estimates on the Green function and application to fluctuation estimates, in preparation
14. Gloria, A., Neukamm, S., Otto, F.: 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)
15. Gloria, A., Neukamm, S., Otto, F.: An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. Max Planck Institute for Mathematics in the Sciences Preprint 41/(2013)
16. Gloria, A., Otto, F.: An optimal variance estimate in stochastic homogenization of discrete elliptic equations. Ann. Probab. **39**(3), 779–856 (2011)
17. Gross, L.: Logarithmic Sobolev inequalities. Am. J. Math. **97**(4), 1061–1083 (1975)
18. Guionnet, A., Zegarlinski, B.: Lecture notes on Logarithmic Sobolev Inequalities. Lecture Notes Math. **1801**, 1–134 (2003)
19. Lawler, G. F., Limic, V.: Random walk: a modern introduction, Cambridge Stud. in Adv. Math. 123, CUP, Cambridge (2010)
20. Ledoux, M.: The concentration of measure phenomenon, Math. Surveys and Monographs 89, AMS, Providence, RI (2001)
21. Lieb, E. H., Loss, M.: Analysis, graduate stud. in Math. 14, 2nd edn, AMS, Providence, RI (2001)
22. Naddaf, A., Spencer, T.: Estimates on the variance of some homogenization problems, unpublished
23. Naddaf, A., Spencer, T.: On homogenization and scaling limit of some gradient perturbation of a massless free field. Commun. Math. Phys. **183**, 55–84 (1997)
24. Nash, J.: Continuity of solutions of parabolic and elliptic equations. Am. J. Math. **80**, 931–954 (1958)
25. Nelson, F.: A quartic interaction in two dimensions, in Mathematical theory of elementary particles. (eds.) Goodman, R., Segal, I. MIT Press, Cambridge, pp. 69–73 (1966)
26. Nelson, F.: The free Markoff field. J. Funct. Anal. **12**, 211–227 (1973)
27. Nolen, J., Normal approximation for a random elliptic equation, Preprint (2011). http://math.duke.edu/nolen/preprints/ellipfluctper_rev
28. Riesz, F., Sz.-Nagy, B.: Functional Analysis, Dover Books on Adv. Math. Dover Publ. Inc., New York (1990)

29. Rossignol, R.: Noise-stability and central limit theorems for effective resistance of random electric networks. arXiv:1206.3856 (2012)
30. Stroock, D., Zegarlinski, B.: The logarithmic Sobolev inequality for discrete spin systems on a lattice. Commun. Math. Phys. **149**, 175–193 (1992)