Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 106, 1–30.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-3618

Moment bounds for the corrector in stochastic homogenization of a percolation model

Agnes Lamacz* Stefan Neukamm† Felix Otto‡

Abstract

We study the corrector equation in stochastic homogenization for a simplified Bernoulli percolation model on $\mathbb{Z}^d$, $d \geq 3$. The model is obtained from the classical $\{0,1\}$-Bernoulli bond percolation by conditioning all bonds parallel to the first coordinate direction to be open. As a main result we prove (in fact for a slightly more general model) that stationary correctors exist and that all finite moments of the corrector are bounded. This extends a previous result in [18], where uniformly elliptic conductances are treated, to the degenerate case. With regard to the associated random conductance model, we obtain as a side result that the corrector not only grows sublinearly, but slower than any polynomial rate. Our argument combines a quantification of ergodicity by means of a Spectral Gap on Glauber dynamics with regularity estimates on the gradient of the elliptic Green's function.

Keywords: quantitative stochastic homogenization; percolation; corrector.
AMS MSC 2010: 39A70; 60K37; 60F17.
Submitted to EJP on June 22, 2014, final version accepted on September 9, 2015.

1 Introduction

We consider the lattice graph $(\mathbb{Z}^d, \mathbb{B}^d)$, $d \geq 3$, where $\mathbb{B}^d$ denotes the set of nearest-neighbor edges. Given a stationary and ergodic probability measure $\langle\cdot\rangle$ on Ω – the space of conductance fields $\boldsymbol{a} : \mathbb{B}^d \to [0,1]$ – we study the *corrector equation* from stochastic homogenization, i.e. the elliptic difference equation

$$\nabla^*(\boldsymbol{a}(\nabla\phi + e)) = 0, \qquad x \in \mathbb{Z}^d. \tag{1.1}$$

Here, ∇ and ∇^* denote discrete versions of the continuum gradient and (negative) divergence, cf. Section 2, and $e \in \mathbb{R}^d$ denotes a vector of unit length, which is fixed throughout the paper. The corrector equation (1.1) emerges in the homogenization of discrete elliptic equations with random coefficients: For random conductances that are

*Technische Universität Dortmund, Fakultät für Mathematik, Germany.
E-mail: agnes.lamacz@tu-dortmund.de
†Technische Universität Dresden, Department of Mathematics, Germany.
E-mail: stefan.neukamm@tu-dresden.de
‡Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany.
E-mail: felix.otto@mis.mpg.de

stationary and ergodic (with respect to the shifts $\boldsymbol{a}(\cdot) \mapsto \boldsymbol{a}(\cdot + z)$, $z \in \mathbb{Z}^d$, cf. Section 2), and under the assumption of uniform ellipticity (i.e. there exists $\lambda_0 > 0$ such that $\boldsymbol{a} \geq \lambda_0$ on $\mathbb{B}^d$ almost surely), a classical result from stochastic homogenization (e. g. see [23, 25]) shows that the effective behavior of $\nabla^* \boldsymbol{a} \nabla$ on large length scales is captured by the homogenized elliptic operator $\nabla^* \boldsymbol{a}_{\hom} \nabla$ where $\boldsymbol{a}_{\hom}$ is a deterministic, symmetric and positive definite $d \times d$ matrix. It is characterized by the minimization problem

$$e \cdot \boldsymbol{a}_{\hom} e = \inf_{\varphi} \langle (e + \nabla \varphi) \cdot \boldsymbol{a}(e + \nabla \varphi) \rangle, \tag{1.2}$$

where the infimum is taken over random fields φ that are $\langle \cdot \rangle$-stationary in the sense of $\varphi(\boldsymbol{a}, x + z) = \varphi(\boldsymbol{a}(\cdot + z), x)$ for all $x, z \in \mathbb{Z}^d$ and $\langle \cdot \rangle$-almost every $\boldsymbol{a} \in \Omega$. Minimizers to (1.2) are called *stationary correctors* and are characterized as the stationary solutions to the corrector equation (1.1). Due to the lack of a Poincaré inequality for ∇ on the infinite dimensional space of stationary random fields, the elliptic operator $\nabla^* \boldsymbol{a} \nabla$ is highly degenerate and the minimum in (1.2) may not be obtained in general. In fact, it is expected to fail generally for $d = 2$, see Remark 3.4 below. The only existence result of a stationary corrector (in dimensions $d \geq 3$) has been obtained recently in [18] by Gloria and the third author under the assumption that the $\boldsymbol{a}$'s are uniformly elliptic, and that $\langle \cdot \rangle$ satisfies a Spectral Gap Estimate, which is in particular the case for independent and identically distributed coefficients. They also show that $\langle |\phi|^p \rangle \lesssim 1$ for all $p < \infty$.

The goal of the present paper is to extend this result to the case of conductances with degenerate ellipticity. To be definite, consider the probability measure $\langle \cdot \rangle_\lambda$ constructed by the following procedure:

> Take the classical $\{0, 1\}$-Bernoulli-bond percolation on $\mathbb{B}^d$ with parameter $\lambda \in (0, 1]$ and declare all bonds parallel to the coordinate direction e_1 to be open. (1.3)

(We adapt the convention to call a bond "open" if the associated coefficient is "1", while a bond is "closed" if the associated coefficient is "0". The parameter λ denotes the probability that a bond is "open"). As for d-dimensional Bernoulli percolation, $\langle \cdot \rangle_\lambda$ describes a random graph of open bonds, which is locally disconnected with positive probability, i.e. the intersection of the graph with a box of arbitrary size yields a disconnected graph with positive probability. However, as a merit of the modification, any two vertices in the random graph are almost surely connected by some open path. As a main result we show that (1.1) admits a stationary solution, all finite moments of which are bounded:

Theorem (main result). Let $d \geq 3$ and $\lambda \in (0, 1]$. There exits $\phi : \Omega \times \mathbb{Z}^d \to \mathbb{R}$ such that for $\langle \cdot \rangle_\lambda$ almost every $\boldsymbol{a} \in \Omega$ we have

- $\phi(\boldsymbol{a}, \cdot)$ solves (1.1),
- $\phi(\boldsymbol{a}, \cdot + z) = \phi(\boldsymbol{a}(\cdot + z), \cdot)$ for all $z \in \mathbb{Z}^d$,

and for all $1 \leq p < \infty$

$$\langle |\phi|^p \rangle_\lambda^{\frac{1}{p}} \leq C.$$

Here C denotes a constant that only depends on p, λ and d.

The modified Bernoulli percolation model $\langle \cdot \rangle_\lambda$ fits into a slightly more general framework that we introduce in Section 2 below, cf. Lemma 2.3. The result above will then follow as a special case of Theorem 3.2 stated below.

Relation to stochastic homogenization. Consider the decaying solution $u_\varepsilon : \mathbb{Z}^d \to \mathbb{R}$ to the equation

$$\nabla^*(\boldsymbol{a}\nabla u_\varepsilon)(\cdot) = \varepsilon^2 f(\varepsilon\cdot) \qquad \text{in } \mathbb{Z}^d,$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is smooth, compactly supported, and $\boldsymbol{a}$ is distributed according to $\langle\cdot\rangle_\lambda$. Classical results of stochastic homogenization (see [29], [22], [25]) show that for almost every $\boldsymbol{a} \in \Omega$ the (piecewise constant interpolation of the) rescaled function $u_\varepsilon(\frac{\cdot}{\varepsilon})$ converges as $\varepsilon \downarrow 0$ to the unique decaying solution $u_{\rm hom} : \mathbb{R}^d \to \mathbb{R}$ of the deterministic elliptic equation

$$-\nabla \cdot (\boldsymbol{a}_{\rm hom}\nabla u_{\rm hom}) = f \qquad \text{in } \mathbb{R}^d.$$

Moreover, a formal two-scale expansion suggests that

$$u_\varepsilon(x) \approx u_{\rm hom}(\varepsilon x) + \varepsilon \sum_{j=1}^d \phi_j(x)\partial_j u_{\rm hom}(\varepsilon x), \tag{1.4}$$

where ϕ_j denotes the (stationary) solution of (1.1) for $e = e_j$ – the jth coordinate direction. The question how to *quantify* the errors emerging in this limiting process is rather subtle. Note that in the case of deterministic periodic homogenization, the good compactness properties of the d-dimensional “reference cell of periodicity” yield a natural starting point for estimates. In contrast, in the stochastic case the reference cell has to be replaced by the probability space $(\Omega, \langle\cdot\rangle)$, which has infinite dimensions and thus most “periodic technologies” break down. Nevertheless, estimates for the homogenization error $\|u_\varepsilon - u_{\rm hom}\|$ and related quantities have been obtained by [20, 12, 10, 14], see also [11, 3] for recent results on fully nonlinear elliptic equations or equations in non-divergence form.

While the asymptotic result of stochastic homogenization holds for general stationary and ergodic coefficients (at least in the uniformly elliptic case), the derivation of error estimates requires a quantification of ergodicity. In a series of papers (see [16, 17, 18, 19]) two of the authors and Gloria developed a quantitative theory for the corrector equation (1.1) (and regularized versions) based on the assumption that the underlying statistics satisfies a Spectral Gap Estimate (SG) for a Glauber dynamics on the coefficient fields. This assumption is satisfied e.g. in the case of independent and identically distributed (i. i. d. coefficients). In [18, 16] moment bounds for the corrector, similar to the one in the present paper, have been obtained. These bounds are at the basis of various optimal estimates; e.g. [16] contains a complete and optimal analysis of the approximation of $\boldsymbol{a}_{\rm hom}$ via periodic representative volume elements, and [17] establishes optimal estimates for the homogenization error and the expansion in (1.4).

While in the works mentioned above it is always assumed that the coefficients are *uniformly elliptic*, i.e. $\boldsymbol{a} \in [\lambda_0, 1]^{\mathbb{B}^d}$ for some fixed $\lambda_0 > 0$, in the present paper we derive moment bounds for a model with *degenerate* elliptic coefficients. As in [18, 16], a crucial element of our approach is an estimate on the *gradient* of the *elliptic Green’s function* associated with $\nabla^*\boldsymbol{a}\nabla$. The required estimate is pointwise in $\boldsymbol{a}$, but (dyadically) averaged in space, and obtained by a self-contained and short argument, see Proposition 3.8 below. It extends the argument in [18] to the degenerate elliptic case. Since in the degenerate case the elementary inequality $\lambda_0|\nabla u|^2 \leq \nabla u \cdot \boldsymbol{a}\nabla u$ breaks down, we replace it by a weighted, integrated version (see Lemma 3.10 below). Compared to more sophisticated methods that e.g. rely on isoperimetric properties of the graph, an advantage of our approach is that it only invokes simple geometrical properties, namely spatial averages (on balls) of the inverse of the chemical distance between nearest neighbor vertices. We believe that our approach extends (although not in a straight-forward manner) to the case of standard supercritical Bernoulli percolation. This is a question that we study in a work in progress.

Connection to random walks in random environments (RWRE). Although, the main motivation of our work is quantitative homogenization, we would like to comment on the connection to invariance principles for (RWRE). In fact, there is a strong link between stochastic homogenization and (RWRE): The operator $\nabla^*(\boldsymbol{a}\nabla)$ generates a stochastic process, namely the variable-speed random walk $X = (X_{\boldsymbol{a}}(t))_{t\geq 0}$, which is a continuous-time random walk in the random environment $\boldsymbol{a}$. In the early work [21] (see also [25]) the authors considered general stationary and ergodic environments. For uniformly elliptic coefficients they prove an *annealed invariance principal* for X, saying that the law of the rescaled process $\sqrt{\varepsilon}X_{\boldsymbol{a}}(\varepsilon^{-1}t)$ weakly converges to that of a Brownian motion with covariance matrix $2\boldsymbol{a}_{\mathrm{hom}}$. In [31] Sidoravicius and Sznitman prove a stronger *quenched invariance principle* for X, which says that the convergence even holds for almost every environment $\boldsymbol{a}$.

More recently, invariance principles have been obtained for more general environments, see [7] and [24] for recent surveys in this direction. Most prominently, supercritical bond percolation on $\mathbb{Z}^d$ has been considered: Here, the annealed result is due to [15], while quenched results have been obtained in [31] for $d \geq 4$ and in [5, 26] for $d \geq 2$. See also [1, 2] for recent related results on degenerate elliptic, possibly unbounded conductances.

The main difficulty in proving a quenched invariance principle compared to the annealed version is to establish a *quenched sublinear growth* property (see (3.4) below) for a corrector field χ. The latter is closely related to the function ϕ considered in Theorem 3.2, see the discussion below Corollary 3.3 for more details. In the uniformly elliptic case, sublinearity of χ is obtained by soft arguments from ergodic theory combined with a Sobolev embedding, see [31]. For supercritical Bernoulli percolation the argument is more subtle: For $d \geq 3$ the proofs in [31, 5, 26] use heat-kernel upper bounds (as deduced by Barlow [4]) or other “heat-kernel technologies” (e.g. see [8, 1, 2]) that require a detailed understanding of the geometry of the percolation cluster, and thus require the use of sophisticated arguments from percolation theory (e.g. isoperimetry, regular volume growth and comparison of chemical and Euclidean distances). Conceptually, the use of such fine arguments seems not to be necessary in the derivation of quenched invariance principles. Motivated by this in [8] and [2] different methods are employed with a reduced usage of heat-kernel technology.

Our approach yields, as a side-result, an alternative way to achieve this goal: The quenched sublinear growth property can easily be obtained from the moment bound derived in Theorem 3.2. In fact, the estimate of Theorem 3.2 is stronger: As we explain in the discussion following Corollary 3.3, our moment bounds imply that the growth of χ is not only sublinear, but *slower than any polynomial rate*, see (3.5). Of course, the environment considered in the present paper, namely the modified percolation model $\langle\cdot\rangle_\lambda$, is much simpler than supercritical Bernoulli percolation. Nevertheless, it shares some of the “degeneracies” featured by percolation; e.g. for every ball $B \subset \mathbb{Z}^d$ with finite radius Poincaré’s inequality $\sum_{x\in B} u^2(x) \leq C(\boldsymbol{a}, B) \sum_{\mathrm{b}\in B} \boldsymbol{a}(\mathrm{b})|\nabla u(\mathrm{b})|^2$ fails with positive probability. Furthermore, in contrast to the above mentioned results, our argument requires only mild estimates on the Green’s function. More precisely, as already mentioned, we require an estimate on the *gradient* of the *elliptic* Green’s function, which – in contrast to quenched heat kernel estimates – can be obtained by fairly simple arguments, see Proposition 3.8. Of course, as it is well-known, estimates on the gradient of the elliptic Green’s function can also be obtained from estimates on the associated heat kernel by an integration in time, and a subsequent application of Caccioppoli’s inequality. In particular, heat-kernel estimates in the spirit of the one obtained by Barlow in the case of supercritical Bernoulli percolation, see [4, Theorem 1], would be sufficient to make this program work. Yet, since the elliptic estimates that we require are less sensitive to the geometry of the graph, and thus can be obtained by simpler arguments, we opt for a self-contained proof that only relies on elliptic regularity theory. Another interesting, and – as we believe – advantageous property of our approach is that (thanks to the Spectral Gap Estimate) probabilistic and deterministic considerations are well separated, e.g. Proposition 3.8 is pointwise in $\boldsymbol{a}$ and does not involve the ensemble.

Structure of the paper. In Section 2 we gather basic definitions and introduce the slightly more general framework studied in this paper. We then present the main result in the general framework. Section 3.1 is devoted to the proof of the main result: we first discuss the general strategy of the proof and present several auxiliary lemmas needed for the proof of the main theorem – in particular, the coercivity estimates, see Lemmas 3.10 and 3.11, and an estimate for the gradient of the elliptic Green's function, see Proposition 3.8, which play a key role in our argument. The proof of the main result is given at the end of Section 3.1, while the auxiliary results are proven in Section 4.

Throughout this article, we use the following notation, see Section 2 for more details:

- d is the dimension;
- $\mathbb{Z}^d$ is the integer lattice;
- $(e_1, \ldots, e_d)$ is the canonical basis of $\mathbb{Z}^d$;
- $e \in \mathbb{R}^d$, which appears in (1.1), denotes a vector of unit length and is fixed throughout the paper;
- $\mathbb{B}^d := \{\, \mathrm{b} = \{x, x+e_i\} \,:\, x \in \mathbb{Z}^d,\, i = 1, \ldots, d \,\}$ is the set of nearest neighbor bonds of $\mathbb{Z}^d$;
- $B_R(x_0)$ is the cube of vertices $x \in x_0 + ([-R, R] \cap \mathbb{Z})^d$;
- $Q_R(x_0)$ is the cube of bonds $\mathrm{b} = \{x, x+e_i\} \in \mathbb{B}^d$ with $x \in B_R(x_0)$ and $i \in \{1, \ldots, d\}$;
- $|A|$ denotes the number of elements in $A \subset \mathbb{Z}^d$ (resp. $A \subset \mathbb{B}^d$).

2 General framework

In the first part of this section, we introduce the general framework following the presentation of [16]: We introduce a discrete differential calculus, the random conductance model, and finally recall the standard definitions of the corrector and the modified corrector.

2.1 Lattice and discrete differential calculus

We consider the lattice graph $(\mathbb{Z}^d, \mathbb{B}^d)$, where $\mathbb{B}^d := \{\, \mathrm{b} = \{x, x+e_i\} \,:\, x \in \mathbb{Z}^d,\, i = 1, \ldots, d \,\}$ denotes the set of nearest-neighbor bonds. We write $\ell^p(\mathbb{Z}^d)$ and $\ell^p(\mathbb{B}^d)$, $1 \le p \le \infty$, for the usual spaces of p-summable (resp. bounded for $p = \infty$) functions on $\mathbb{Z}^d$ and $\mathbb{B}^d$. For $u : \mathbb{Z}^d \to \mathbb{R}$ the *discrete derivative* $\nabla u(\mathrm{b})$, $\mathrm{b} \in \mathbb{B}^d$, is defined by the expression

$$\nabla u(\mathrm{b}) := u(y_{\mathrm{b}}) - u(x_{\mathrm{b}}).$$

Here x_{b} and y_{b} denote the unique vertices with $\mathrm{b} = \{x_{\mathrm{b}}, y_{\mathrm{b}}\} \in \mathbb{B}^d$ satisfying $y_{\mathrm{b}} - x_{\mathrm{b}} \in \{e_1, \ldots, e_d\}$. We denote by ∇^* the adjoined of ∇, so that we have for $F : \mathbb{B}^d \to \mathbb{R}$

$$\nabla^* F(x) = \sum_{i=1}^{d} \Big(F(\{x - e_i, x\}) - F(\{x, x + e_i\}) \Big).$$

Furthermore, the discrete integration by parts formula reads

$$\sum_{\mathrm{b} \in \mathbb{B}^d} \nabla u(\mathrm{b}) F(\mathrm{b}) = \sum_{x \in \mathbb{Z}^d} u(x) \nabla^* F(x), \tag{2.1}$$

and holds whenever the sums converge absolutely.

2.2 Random conductance field

To each bond $\mathrm{b} \in \mathbb{B}^d$ a *conductance* $\boldsymbol{a}(\mathrm{b}) \in [0,1]$ is attached. Hence, a *configuration* of the lattice is described by a *conductance field* $\boldsymbol{a} \in \Omega$, where $\Omega := [0,1]^{\mathbb{B}^d}$ denotes the *configuration space*. Given $\boldsymbol{a} \in \Omega$ we define the chemical distance between vertices $x, y \in \mathbb{Z}^d$ by

$$\operatorname{dist}_{\boldsymbol{a}}(x,y) := \inf\left\{ \sum_{\mathrm{b}\in\pi} \boldsymbol{a}(\mathrm{b})^{-1} \ :\ \pi \text{ is a path from } x \text{ to } y \right\} \qquad (\text{where } \tfrac{1}{0} := +\infty).$$

We equip Ω with the product topology (i. e. the $\mathbb{Z}^d$-fold product of the Borel-σ-algebra on $[0,1] \subset \mathbb{R}$) and the usual product σ-algebra, and describe *random configurations* by means of a probability measure on Ω, called the *ensemble*. The associated expectation is denoted by $\langle\cdot\rangle$.

Our assumptions on $\langle\cdot\rangle$ are the following:

Assumption 2.1.

(A1) *(Stationarity). The shift operators $\Omega \ni \boldsymbol{a} \mapsto \boldsymbol{a}(\cdot + z) \in \Omega$, $z \in \mathbb{Z}^d$ preserve the measure $\langle\cdot\rangle$. (For a bond $\mathrm{b} = \{x,y\} \in \mathbb{B}^d$ and $z \in \mathbb{Z}^d$ we write $\mathrm{b} + z := \{x+z, y+z\}$ for the shift of b by z.)*

(A2) *(Moment condition). There exists a modulus of integrability $\Lambda : [1,\infty) \to [0,\infty)$ such that the distance of neighbors is finite on average in the sense that*

$$\forall p < \infty\ :\ \max_{i=1,\ldots,d} \langle (\operatorname{dist}_{\boldsymbol{a}}(0,e_i))^p \rangle^{\frac{1}{p}} \le \Lambda(p).$$

(A3) *(Spectral Gap Estimate). There exists a constant $\rho > 0$ such that for all $\zeta \in L^2(\Omega)$ we have*

$$\left\langle (\zeta - \langle\zeta\rangle)^2 \right\rangle \le \frac{1}{\rho} \sum_{\mathrm{b}\in\mathbb{B}^d} \left\langle \left(\frac{\partial\zeta}{\partial\mathrm{b}} \right)^2 \right\rangle,$$

where $\frac{\partial\zeta}{\partial\mathrm{b}}$ denotes the vertical derivative as defined in Definition 2.2 below.

For technical reasons we need to strengthen (A2):

(A2+) *We assume that*

$$\forall p < \infty\ :\ \max_{i=1,\ldots,d} \langle (\operatorname{dist}_{\boldsymbol{a}^{e_i,0}}(0,e_i))^p \rangle^{\frac{1}{p}} \le \Lambda(p),$$

where $\boldsymbol{a}^{e_i,0}$ denotes the conductance field obtained by "deleting" the bond $\{0,e_i\}$ (i. e. $\boldsymbol{a}^{e_i,0}(\mathrm{b}) = \boldsymbol{a}(\mathrm{b})$ for all $\mathrm{b} \neq \{0,e_i\}$ and $\boldsymbol{a}^{e_i,0}(\{0,e_i\}) = 0$).

Let us comment on these properties. A minimal requirement needed for qualitative stochastic homogenization in the uniformly elliptic case is stationarity and ergodicity of the ensemble. The basic example for such an ensemble are i. i. d. coefficients which means that $\langle\cdot\rangle$ is a $\mathbb{B}^d$-fold product of a "single edge" probability measure on $[0,1]$. The assumption (A3) is weaker than assuming i. i. d., but stronger than ergodicity. Indeed, in [16] it is shown that any i. i. d. ensemble satisfies (A3) with constant $\rho = 1$. Moreover, it is shown that (A3) can be seen as a quantification of ergodicity. From the functional analytic point of view the spectral gap estimate is a Poincaré inequality where the derivative is taken in vertical direction, see below. (The terminology "vertical" versus "horizontal" is motivated from viewing $\boldsymbol{a} \in \Omega$ as a "height"-function defined on the "horizontal" plane $\mathbb{B}^d$). We recall from [16] the definition of the vertical derivative:

Definition 2.2. *For $\zeta \in L^1(\Omega)$ the vertical derivative w. r. t. $\mathrm{b} \in \mathbb{B}^d$ is given by*

$$\frac{\partial \zeta}{\partial \mathrm{b}} := \zeta - \langle \zeta \rangle_{\mathrm{b}},$$

where $\langle \zeta \rangle_{\mathrm{b}}$ denotes the conditional expectation where we condition on $\{\boldsymbol{a}(\mathrm{b}')\}_{\mathrm{b}' \neq \mathrm{b}}$. For $\zeta : \Omega \to \mathbb{R}$ sufficiently smooth we denote by $\frac{\partial \zeta}{\partial \boldsymbol{a}(\mathrm{b})}$ the classical partial derivative of ζ w. r. t. the coordinate $\boldsymbol{a}(\mathrm{b})$.

Properties (A2) and (A2+) are crucial assumptions on the connectedness of the graph. In particular they imply that almost surely every pair of vertices can be connected by a path with finite intrinsic length. Note that (A2+) is stronger than (A2), e.g. (A2+) implies that vertices almost surely can be connected by two disjoint paths. It is easy to construct an example which satisfies (A2) but not (A2+): First, one constructs a periodic, deterministic coefficient field that connects all sites of the lattice, but looses this property when deleting a specific edge. The stationary environment is then obtained by randomizing the origin. However, (A2) and (A2+) do not exclude configurations with coefficients that vanish with non-zero probability, as it is the case for $\langle \cdot \rangle_\lambda$ – the model considered in the introduction:

Lemma 2.3. *The modified Bernoulli percolation model $\langle \cdot \rangle_\lambda$ defined via (1.3) satisfies Assumption 1 with $\rho = 1$.*

Proof. Evidently, $\langle \cdot \rangle_\lambda$ can be written as the (infinite) product of probability measures attached to the bonds in $\mathbb{B}^d$. These “single-bond” probability measures only depend on the direction of the bond. Hence, $\langle \cdot \rangle_\lambda$ is stationary. Another consequence of the product structure is that $\langle \cdot \rangle_\lambda$ satisfies (A3) with constant $\rho = 1$ (see [16, Lemma 7] for the argument). It remains to check (A2+). By stationarity and symmetry we may assume that $e_i = e_d$. Consider the (random) set

$$\mathcal{L}(\boldsymbol{a}) := \{\, j \in \mathbb{Z} \,:\, \boldsymbol{a}^{e_d,0}(\{je_1, je_1 + e_d\}) = 1 \,\}.$$

Clearly, each $j \in \mathcal{L}(\boldsymbol{a})$ yields an open path connecting 0 and e_d, for instance the “U-shaped” path through the sites 0, je_1, $je_1 + e_d$ and e_d. Hence, $\operatorname{dist}_{\boldsymbol{a}^{e_d,0}}(0, e_d) \leq 2\operatorname{dist}(0, \mathcal{L}(\boldsymbol{a})) + 1$ almost surely, where $\operatorname{dist}(0, \mathcal{L}(\boldsymbol{a})) := \min_{j \in \mathcal{L}(\boldsymbol{a})} |j|$. Since the random variable $\operatorname{dist}(0, \mathcal{L}(\boldsymbol{a}))$ has the geometric distribution with parameter $(1 - (1-\lambda)^2)$, all of its moments are bounded and (A2+) follows. □

3 Main result

We are interested in stationary solutions to the corrector equation (1.1). Note that we tacitly identify the vector $e \in \mathbb{R}^d$ with the translation invariant vector field $e(\mathrm{b}) := e \cdot (y_{\mathrm{b}} - x_{\mathrm{b}})$. For conciseness we write

$$\mathcal{S} := \Big\{ \varphi \,:\, \Omega \times \mathbb{Z}^d \to \mathbb{R} \,\Big|\, \varphi \text{ is measurable \& stationary, i. e. } \varphi(\boldsymbol{a}(\cdot + z), x) = \varphi(\boldsymbol{a}, x + z) \text{ for all } x, z \in \mathbb{Z}^d \text{ and } \langle \cdot \rangle\text{-almost every } \boldsymbol{a} \in \Omega \Big\}$$

for the space of *stationary random fields*. Thanks to (A1) the expectation $\langle \varphi \rangle = \langle \varphi(\cdot, x) \rangle$ of a stationary random variable does not depend on x. Therefore, $\|\varphi\|_{L^2(\Omega)} := \langle |\varphi|^2 \rangle^{\frac{1}{2}}$ defines a norm on $(\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$.

We are interested in solutions to (1.1) in $(\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$. Thanks to discreteness, the operator $\nabla^*(\boldsymbol{a}\nabla)$ is bounded and linear on $(\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$. However, it is degenerate-elliptic for two-reasons:

- In general the Poincaré inequality does not hold in $(\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$.
- The conductances $\boldsymbol{a}$ may vanish with positive probability.

Therefore, following [29], we regularize the equation by adding a 0th order term and consider for $T > 0$ the modified corrector equation

$$\frac{1}{T}\phi_T(x) + \nabla^*(\boldsymbol{a}(\nabla\phi_T + e))(x) = 0 \qquad \text{for all } x \in \mathbb{Z}^d \text{ and } \boldsymbol{a} \in \Omega. \tag{3.1}$$

Thanks to the regularization and the maximum principle, (3.1) admits (for all $T > 0$ and $\boldsymbol{a} \in \Omega$) a unique bounded solution (see discussion around (4.1)).

Definition 3.1 (modified corrector). *The unique bounded solution ϕ_T to* (3.1) *is called the modified corrector.*

Note that ϕ_T is automatically stationary and thus belongs to $(S, \|\cdot\|_{L^2(\Omega)})$. We think about the modified corrector as an approximation for the stationary corrector and hope to recover a solution to (1.1) in the limit $T \uparrow \infty$. This is possible as soon as we have estimates on (some) moments of ϕ_T that are uniform in T — this is the main result of the paper:

Theorem 3.2 (Moment bounds for the modified corrector). *Let $d \geq 3$ and $\langle\cdot\rangle$ satisfy Assumption 2.1 for some ρ and Λ. Let ϕ_T denote the modified corrector as defined in Definition 3.1. Then for all $T > 0$ and $1 \leq p < \infty$ we have*

$$\langle|\phi_T|^p\rangle^{\frac{1}{p}} \lesssim 1. \tag{3.2}$$

Here $\lesssim$ means $\leq$ up to a constant that only depends on p, Λ, ρ, and d.

Since the estimate in Theorem 3.2 is uniform in T we get as a corollary:

Corollary 3.3. *Let $d \geq 3$ and $\langle\cdot\rangle$ satisfy Assumption 2.1 for some ρ and Λ. Then the corrector equation* (1.1) *has a unique stationary solution $\phi \in (\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$ with $\langle\phi\rangle = 0$. Moreover, we have*

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

for all $1 \leq p < \infty$. Here $\lesssim$ means $\leq$ up to a constant that only depends on p, Λ, ρ and d.

For the proof of Theorem 3.2 and Corollary 3.3 see Section 3.1.

Remark 3.4. In general one cannot expect (3.2) to hold in dimension $d = 2$. In fact, even in the case of uniformly elliptic, independent and identically distributed coefficients one expects that

$$\left\langle\phi_T^2\right\rangle \sim \log T. \tag{3.3}$$

The upper bound has been established in [16]. The lower bound might be established following the lines in [13]. A heuristic argument for the lower bound is the following: In the limit of vanishing ellipticity contrast, the modified corrector equation "linearizes" to

$$\frac{1}{T}\phi_T + \nabla^*\nabla\phi_T = -\nabla^*\boldsymbol{a}e.$$

With the Green's function representation and by appealing to independence of the coefficients we get

$$\left\langle\phi_T^2\right\rangle = (\sum_{\mathrm{b}\in\mathbb{B}^d}(\nabla G_T^0(0,\mathrm{b}))^2\left\langle(\boldsymbol{a}e)^2\right\rangle,$$

where G_T^0 denotes the Green's function associated with the operator $\frac{1}{T} + \nabla^*\nabla$. Now $\sum_{\mathrm{b}\in\mathbb{B}^d}|\nabla G_T^0(0,\mathrm{b})|^2 \sim \log T$ suggests (3.3). Finally, we note that (3.3) implies that (1.1) cannot have a stationary solution with finite second moments. We give an indirect argument: If ϕ is a stationary solution to (1.1) and $\langle \phi^2 \rangle < \infty$, then the difference $\psi_T := \phi_T - \phi$ solves the equation

$$\frac{1}{T}\psi_T + \nabla^*(\boldsymbol{a}\nabla\psi_T) = -\frac{1}{T}\phi,$$

and thus $\langle \psi_T^2 \rangle \leq \langle \phi^2 \rangle$ by a standard energy estimate. But this implies that $\langle \phi_T^2 \rangle$ is bounded uniformly in T, which contradicts (3.3).

As mentioned in the introduction the corrector can be used to establish invariance principles for random walks in random environments. Suppose that $\langle\cdot\rangle$ satisfies Assumption 2.1 for some ρ and Λ. Then, thanks to Corollary 3.3, for each coordinate direction e_k there exist stationary correctors $\phi^k \in (\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$ with $\langle \phi^k \rangle = 0$ that solve (1.1) with $e = e_k$. Hence, we can consider the random vector field $\chi = (\chi^1, \ldots, \chi^d) : \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$ defined by

$$\chi^k(\boldsymbol{a}, x) := \phi^k(\boldsymbol{a}, x) - \phi^k(\boldsymbol{a}, x = 0).$$

By construction the map $\mathbb{Z}^d \ni x \mapsto x + \chi(\boldsymbol{a}, x)$ is $\boldsymbol{a}$-harmonic, has finite second moments, and is shift covariant, i.e. $\chi(\boldsymbol{a}, x+y) - \chi(\boldsymbol{a}, x) = \chi(\boldsymbol{a}(\cdot + y), x)$. The field χ is precisely the “corrector” used e.g. in [21, 31] to introduce harmonic coordinates for which the random walk in the random environment is a martingale. In particular, in [31] Sidovaricius and Sznitman use χ to prove a quenched invariance principle for the random walk in a random environment. A key step in their argument is to show that χ has sublinear growth, i.e.

$$\lim_{R\to\infty} \max_{x\in B_R(0)} \frac{|\chi(\boldsymbol{a}, x)|}{R} = 0 \qquad \text{for } \langle\cdot\rangle\text{-almost every } \boldsymbol{a} \in \Omega. \tag{3.4}$$

Variants of this property have been established for supercritical bond percolation on $\mathbb{Z}^d$ in dimension $d \geq 4$ in [31] and for $d \geq 2$ in [5, 26], see also [9] where the uniformly elliptic case in dimension $d = 2$ is treated. The moment bounds established in our work (under the more restrictive Assumption 2.1) are stronger. Indeed, from Theorem 3.2 we get (3.4) in the stronger form

$$\forall\theta \in [0,1): \qquad \lim_{R\uparrow\infty} R^\theta \max_{x\in B_R(0)} \frac{|\chi^k(\boldsymbol{a}, x)|}{R} = 0 \qquad \langle\cdot\rangle\text{-almost surely.} \tag{3.5}$$

This can be seen by the following simple argument (cf. Corollary 4.2 in [27]): Set

$$\zeta_R(\boldsymbol{a}) := R^\theta \max_{x\in B_R(0)} \frac{|\chi^k(\boldsymbol{a}, x)|}{R}. \tag{3.6}$$

By the Borel-Cantelli Lemma, we only need to show that for all $\lambda > 0$ we have

$$\sum_{R=1}^{\infty} \langle \mathbf{1}(\{|\zeta_R| > \lambda\}) \rangle < \infty,$$

where here and below $\mathbf{1}(\cdot)$ denotes the indicator function of the set in the bracktes. Since $\{|\zeta_R| > \lambda\} \subset \bigcup_{x\in B_R(0)} \{|\chi^k(\boldsymbol{a}, x)| > \lambda R^{1-\theta}\}$, for every $p \geq 1$ we get

$$\langle \mathbf{1}(\{|\zeta_R| > \lambda\}) \rangle \leq |B_R(0)| \frac{\langle |\chi^k(\boldsymbol{a}, x)|^p \rangle}{(\lambda R^{1-\theta})^p} \leq C(d) R^{d-p(1-\theta)} \langle |\phi^k|^p \rangle.$$

For $p > \frac{d}{1-\theta}$, the exponent of R is negative and (3.6) follows.

3.1 Outline and Proof of Theorem 3.2

The proof of Theorem 3.2 is inspired by the approach in [18] where uniformly elliptic conductances are treated. All auxiliary results in this section hold for dimension $d \geq 2$. The only place where we use $d > 2$ is in the proof of the theorem itself, where we appeal to the fact that the ℓ^2-norm of $|\nabla G_T|$ is bounded uniformly in T. As it is well-known, this is not the case in dimension $d = 2$. The starting point of our argument is the following p-version of the *Spectral Gap Estimate* (A3), which we recall from [16, Lemma 2]:

Lemma 3.5 (p-version of (SG))**.** *Let $\langle \cdot \rangle$ satisfy (A3) with constant $\rho > 0$. Then for $p \in \mathbb{N}$ and all $\zeta \in L^{2p}(\Omega)$ with $\langle \zeta \rangle = 0$ we have*

$$\left\langle \zeta^{2p} \right\rangle \lesssim \left\langle \left(\sum_{\mathrm{b} \in \mathbb{B}^d} \left(\frac{\partial \zeta}{\partial \mathrm{b}} \right)^2 \right)^p \right\rangle,$$

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

Applied to $\zeta = \phi_T(x = 0)$, this estimate yields a bound on stochastic moments of ϕ_T in terms of the vertical derivatives $\frac{\partial \phi_T(x=0)}{\partial \mathrm{b}}$, $\mathrm{b} \in \mathbb{B}^d$ (see Definition 2.2). Heuristically, we expect the vertical derivative $\frac{\partial \phi_T(x=0)}{\partial \mathrm{b}}$ to behave as the classical partial derivative $\frac{\partial \phi_T(x=0)}{\partial \boldsymbol{a}(\mathrm{b})}$. As we shall see, the latter admits the Green's function representation

$$\frac{\partial \phi_T(x = 0)}{\partial \boldsymbol{a}(\mathrm{b})} = -\nabla G_T(\boldsymbol{a}, \mathrm{b}, 0)(\nabla \phi_T(\mathrm{b}) + e(\mathrm{b})). \tag{3.7}$$

Recall that $e(\mathrm{b}) = e \cdot (y_\mathrm{b} - x_\mathrm{b})$ where $e \in \mathbb{R}^d$ denotes the vector in the modified corrector equation (3.1). Above G_T denotes the Green's function associated with $(\frac{1}{T} + \nabla^* \boldsymbol{a} \nabla)$ and is defined as follows:

Definition 3.6. *For $T > 0$ the Green's function $G_T : \Omega \times \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$ is defined as follows: For each $\boldsymbol{a} \in \Omega$ and $y \in \mathbb{Z}^d$ the function $x \mapsto G_T(\boldsymbol{a}, x, y)$ is the unique solution in $\ell^2(\mathbb{Z}^d)$ to*

$$\frac{1}{T} G_T(\boldsymbol{a}, \cdot, y) + \nabla^* \boldsymbol{a} \nabla G_T(\boldsymbol{a}, \cdot, y) = \delta(\cdot - y). \tag{3.8}$$

For uniformly elliptic conductances we have $\frac{\partial \phi_T(x=0)}{\partial \mathrm{b}} \sim \frac{\partial \phi_T(x=0)}{\partial \boldsymbol{a}(\mathrm{b})}$ up to a constant that only depends on the ratio of ellipticity. In the case of degenerate ellipticity this is no longer true. However, the discrepancy between the vertical and classical partial derivative of ϕ_T can be quantified in terms of weights defined as follows: We introduce the weight function $\omega : \Omega \times \mathbb{B}^d \to [0, \infty]$ as

$$\omega(\boldsymbol{a}, \mathrm{b}) := (\mathrm{dist}_{\boldsymbol{a}}(x_\mathrm{b}, y_\mathrm{b}))^{d+2} \qquad (\boldsymbol{a} \in \Omega,\ \mathrm{b} = \{x_\mathrm{b}, y_\mathrm{b}\} \in \mathbb{B}^d). \tag{3.9}$$

For $\mathrm{b} \in \mathbb{B}^d$ and $\boldsymbol{a} \in \Omega$ we denote by $\boldsymbol{a}^{\mathrm{b},0}$ the conductance field obtained by "deleting" the bond b (i. e. $\boldsymbol{a}^{\mathrm{b},0}(\mathrm{b}') = \boldsymbol{a}(\mathrm{b}')$ for all $\mathrm{b}' \neq \mathrm{b}$ and $\boldsymbol{a}^{\mathrm{b},0}(\mathrm{b}) = 0$), and introduce the modified weight ω_0 as

$$\omega_0(\boldsymbol{a}, \mathrm{b}) := \omega(\boldsymbol{a}^{\mathrm{b},0}, \mathrm{b}). \tag{3.10}$$

Lemma 3.7. *Assume that $\langle \cdot \rangle$ satisfies (A1) and (A2+). For $T > 0$ let ϕ_T denote the modified corrector. Then for all $\mathrm{b} \in \mathbb{B}^d$ we have*

$$\left| \frac{\partial \phi_T(x = 0)}{\partial \mathrm{b}} \right| \lesssim \omega_0^2(\mathrm{b}) \left| \nabla G_T(\mathrm{b}, 0) \right| \left| \nabla \phi_T(\mathrm{b}) + e(\mathrm{b}) \right|.$$

Here $\lesssim$ means $\leq$ up to a constant that only depends on d.

To benefit from (3.7) (in the form of Lemma 3.7) we require an *estimate on the gradient of the Green's function*. As it is well known, the constant coefficient Green's function $G^0_T(x) := G_T(\boldsymbol{a} = \mathbf{1}, x, 0)$ (which is associated with the modified Laplacian $\frac{1}{T} + \nabla^*\nabla$) satisfies the pointwise estimate

$$\forall \mathrm{b} := \{x, x+e_i\} : \qquad |\nabla G^0_T(\mathrm{b})| \lesssim (1+|x|)^{1-d} \qquad \text{uniformly in } T>0. \tag{3.11}$$

We require an estimate that captures the same decay in x. It is known from the continuum, uniformly elliptic case, that such an estimate cannot hold pointwise in x and at the same time pointwise in $\boldsymbol{a}$. In [18, Lemma 2.9], for uniformly elliptic conductances, a spatially averaged version of (3.11) is established, where the averages are taken over dyadic annuli. The constant in this estimate depends on the conductances only through their contrast of ellipticity. In the degenerate elliptic case, the ellipticity contrast is infinite. In order to keep the optimal decay in x, we need to allow the constant in the estimate to depend on $\boldsymbol{a}$. For $x_0 \in \mathbb{Z}^d$, $R > 1$ and $1 \le q < \infty$ consider the spatial average of the weight ω (cf. (3.9))

$$C(\boldsymbol{a}, Q_R(x_0), q) := \left(\frac{1}{|Q_R(x_0)|} \sum_{\mathrm{b} \in Q_R(x_0)} \omega^q(\boldsymbol{a}, \mathrm{b}) \right)^{\frac{1}{q}}. \tag{3.12}$$

We shall prove the following estimate:

Proposition 3.8. *For $R_0 > 1$ and $k \in \mathbb{N}_0$ consider*

$$A_k := \begin{cases} Q_{R_0}(0) & k = 0, \\ Q_{2^k R_0}(0) \setminus Q_{2^{k-1}R_0}(0) & k \ge 1. \end{cases}$$

Then for all $\frac{2d}{d+2} < p < 2$ we have

$$\left(\frac{1}{|A_k|} \sum_{\mathrm{b} \in A_k} |\nabla G_T(\boldsymbol{a}, \mathrm{b}, 0)|^p \right)^{\frac{1}{p}} \lesssim C(\boldsymbol{a})\, 2^{k(1-d)},$$

where $\lesssim$ means $\le$ up to a constant that only depends on R_0, d and p, and

$$C(\boldsymbol{a}) := C^{\frac{\beta}{2}}(\boldsymbol{a}, Q_{2^{k+1}R_0}(0), \tfrac{p}{2-p}) \tag{3.13}$$

with $\beta := 2\frac{p^*-1}{p^*-2} + p^*$ and $p^* := \frac{dp}{d-p}$.

The precise form of the constant C in (3.13) is not important. In fact, in the random setting, when Ω is equipped with a probability measure satisfying (A1) and (A2), we may view C as a random variable with controlled finite moments:

Remark 3.9. Let $\langle \cdot \rangle$ satisfy Assumption (A1). Then the spatial average introduced in (3.12) satisfies

$$\langle C^q(\boldsymbol{a}, Q_R(x_0), q') \rangle = \left\langle \left(\frac{1}{|Q_R(x_0)|} \sum_{\mathrm{b} \in Q_R(x_0)} \omega^{q'}(\boldsymbol{a}, \mathrm{b}) \right)^{\frac{q}{q'}} \right\rangle \le \begin{cases} \left\langle \omega^{q'} \right\rangle^{\frac{q}{q'}} & \text{if } q' \ge q, \\ \langle \omega^q \rangle & \text{if } q' < q, \end{cases}$$

as can be seen by appealing to Jensen's inequality and stationarity. Moreover, if $\langle \cdot \rangle$ additionally fulfills (A2), then C defined in (3.13) satisfies

$$\forall m \in \mathbb{N} : \langle C^m \rangle^{\frac{1}{m}} \lesssim 1,$$

where $\lesssim$ means $\le$ up to a constant that only depends on m, p, Λ and d.

The proof of Proposition 3.8 relies on arguments from elliptic regularity theory, which in the uniformly elliptic case are standard. They typically involve the pointwise inequality

$$\lambda_0|\nabla u(\mathrm{b})|^2 \le \nabla u(\mathrm{b})\,\boldsymbol{a}(\mathrm{b})\nabla u(\mathrm{b}), \qquad (\mathrm{b}\in\mathbb{B}^d), \tag{3.14}$$

where $\lambda_0 > 0$ denotes the constant of ellipticity. In the degenerate case, the conductances $\boldsymbol{a}$ may vanish on a non-negligible set of bonds and (3.14) breaks down. As a replacement we establish estimates which provide a weighted, integrated version of (3.14):

Lemma 3.10. *Let $p > d+1$. For any function $u : \mathbb{Z}^d \to \mathbb{R}$ and all $\boldsymbol{a} \in \Omega$ we have (with the convention $\frac{1}{\infty} = 0$)*

$$\sum_{\mathrm{b}\in\mathbb{B}^d} |\nabla u(\mathrm{b})|^2 \mathrm{dist}_{\boldsymbol{a}}^{-p}(x_\mathrm{b}, y_\mathrm{b}) \le C(p,d) \sum_{\mathrm{b}\in\mathbb{B}^d} \boldsymbol{a}(\mathrm{b})|\nabla u(\mathrm{b})|^2, \tag{3.15}$$

where $C(p,d) := \sum_{x\in\mathbb{Z}^d}(|x|+1)^{1-p}$ and the inequality holds whenever the sums converge.

While Lemma 3.10 is purely deterministic, we also need the following statistically averaged version:

Lemma 3.11. *Let $\langle\cdot\rangle$ be stationary, cf. (A1), and $p > d+1$. Then for any stationary random field u and any bond $\mathrm{b}\in\mathbb{B}^d$ we have (with the convention $\frac{1}{\infty} = 0$)*

$$\left\langle |\nabla u(\mathrm{b})|^2 \mathrm{dist}_{\boldsymbol{a}}^{-p}(x_\mathrm{b}, y_\mathrm{b}) \right\rangle \le C(p,d) \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1\ldots d}} \left\langle \boldsymbol{a}(\mathrm{b}')|\nabla u(\mathrm{b}')|^2 \right\rangle,$$

where $C(p,d) := \sum_{k=0}^\infty 2^{k(1-p)}|B_{2^{k+1}}(0)| < \infty$.

A last ingredient required for the proof of Theorem 3.2 is a *Caccioppoli inequality in probability* that yields a gain of stochastic integrability and helps to treat the $\nabla\phi_T$-term on the right-hand side in (3.7). In the uniformly elliptic case, i. e. when $0 < \lambda_0 \le \boldsymbol{a} \le 1$, the Caccioppoli inequality

$$\left\langle |\nabla\phi_T|^{2p+2} \right\rangle^{\frac{1}{2p+2}} \lesssim \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}\frac{p}{p+1}} \tag{3.16}$$

holds for any integer exponents p (see [18, Lemma 2.7]). The inequality follows from combining the elementary discrete inequality

$$|\nabla u(\mathrm{b})| = |u(y_\mathrm{b}) - u(x_\mathrm{b})| \le |u(y_\mathrm{b})| + |u(x_\mathrm{b})|, \tag{3.17}$$

with the estimate

$$\left\langle \phi_T^{2p}|\nabla\phi_T|^2 \right\rangle \lesssim \frac{1}{\lambda_0} \left\langle \phi_T^{2p}|\nabla\phi_T| \right\rangle. \tag{3.18}$$

The latter is obtained by testing the modified corrector equation (3.1) with ϕ_T^{2p+1} and uses the uniform ellipticity of $\boldsymbol{a}$. In the degenerate elliptic case (3.18) is not true any longer. However, by appealing to Lemma 3.11 the following weaker version of (3.16) survives:

$$\left\langle |\nabla\phi_T|^{(2p+2)\theta} \right\rangle^{\frac{1}{(2p+2)\theta}} \lesssim \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}\frac{p}{p+1}} \tag{3.19}$$

for any factor $0 < \theta < 1$. Hence, we only gain an increase of integrability by exponents strictly smaller than two. As a matter of fact, in the proof of our main result we only need the estimate in the following form:

Lemma 3.12 (Caccioppoli estimate in probability)**.** *Let $\langle\cdot\rangle$ satisfy (A1) and (A2). Let ϕ_T denote the corrector associated with $e\in\mathbb{R}^d$, $|e| = 1$, $T > 0$. For every even integer p and $\mathrm{b}\in\mathbb{B}^d$ we have*

$$\left\langle |\nabla\phi_T(\mathrm{b})|^{2p+1} \right\rangle^{\frac{1}{2p+1}} \lesssim \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}\frac{p}{p+1}}, \tag{3.20}$$

where $\lesssim$ means $\le$ up to a constant that only depends on p, Λ and d.

Now we are ready to prove our main result:

Proof of Theorem 3.2. It suffices to consider exponents $p \in 2\mathbb{N}$ that are larger than a threshold depending only on d – the threshold is determined by (3.22) below.

Further, we only need to prove

$$\left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}} \lesssim \max_{\substack{\mathrm{b}'=\{0,e_i\} \\ i=1,\dots,d}} \left\langle |\nabla\phi_T(\mathrm{b}')|^{2p+1} \right\rangle^{\frac{1}{2p+1}} + 1. \tag{3.21}$$

Indeed, in combination with the Caccioppoli estimate in probability, cf. Lemma 3.12, estimate (3.21) yields $\left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}} \lesssim \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{2p}\frac{p}{p+1}} + 1$. Since $\frac{p}{p+1} < 1$ the first term can be absorbed and the desired estimate follows.

We will now prove (3.21). For reasons that will become clear at the end of the argument we fix an exponent $\frac{2d}{d+2} < q < 2$ such that

$$d(\frac{1}{q} + \frac{1}{2p} - 1) + 1 < 0. \tag{3.22}$$

This is always possible for $p \gg 1$ and $0 < 2 - q \ll 1$, since

$$\lim_{q\uparrow 2, p\uparrow\infty} d(\frac{1}{q} + \frac{1}{2p} - 1) = -\frac{d}{2} < -1 \qquad \text{for } d \geq 3.$$

Our argument for (3.21) starts with the p-version of the spectral gap estimate, see Lemma 3.5, that we combine with Lemma 3.7:

$$\begin{aligned} \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{p}} &= \left\langle \phi_T^{2p}(x=0) \right\rangle^{\frac{1}{p}} \lesssim \left\langle \left(\sum_{\mathrm{b}\in\mathbb{B}^d} \left(\frac{\partial\phi_T(x=0)}{\partial\mathrm{b}} \right)^2 \right)^p \right\rangle^{\frac{1}{p}} \\ &\lesssim \left\langle \left(\sum_{\mathrm{b}\in\mathbb{B}^d} (\nabla G_T(\mathrm{b},0))^2 (\nabla\phi_T(\mathrm{b}) + e(\mathrm{b}))^2 \omega_0^4(\mathrm{b}) \right)^p \right\rangle^{\frac{1}{p}}. \end{aligned}$$

Now we wish to benefit from the decay estimate for ∇G_T in Proposition 3.8, and therefore decompose $\mathbb{B}^d$ into dyadic annuli: Let the dyadic annuli A_k, $k \in \mathbb{N}_0$ be defined as in Proposition 3.8 with initial radius $R_0 = 2$. Note that $\mathbb{B}^d$ can be written as the disjoint union of $A_0, A_1, A_2, \dots$. With the triangle inequality w. r. t. $\langle (\cdot)^p \rangle^{\frac{1}{p}}$ and Hölder's inequality in b-space with exponents $(\frac{p}{p-1}, p)$ we get

$$\begin{aligned} \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{p}} &\lesssim \sum_{k\in\mathbb{N}_0} \left\langle \left(\sum_{\mathrm{b}\in A_k} (\nabla G_T(\mathrm{b},0))^2 (\nabla\phi_T(\mathrm{b}) + e(\mathrm{b}))^2 \omega_0^4(\mathrm{b}) \right)^p \right\rangle^{\frac{1}{p}} \\ &\lesssim \sum_{k\in\mathbb{N}_0} \left\langle \left(\sum_{\mathrm{b}\in A_k} |\nabla G_T(\mathrm{b},0)|^{\frac{2p}{p-1}} \right)^{p-1} \left(\sum_{\mathrm{b}\in A_k} (\nabla\phi_T(\mathrm{b}) + e(\mathrm{b}))^{2p} \omega_0^{4p}(\mathrm{b}) \right) \right\rangle^{\frac{1}{p}}. \end{aligned} \tag{3.23}$$

Since $\frac{2d}{d+2} < q < 2 < \frac{2p}{p-1}$, we have $\|\cdot\|_{\ell^{\frac{2p}{p-1}}} \leq \|\cdot\|_{\ell^q}$, which combined with the decay estimate of Proposition 3.8 yields

$$\left(\sum_{\mathrm{b}\in A_k} |\nabla G_T(\mathrm{b},0)|^{\frac{2p}{p-1}} \right)^{p-1} \leq \left(\sum_{\mathrm{b}\in A_k} |\nabla G_T(\mathrm{b},0)|^q \right)^{\frac{2p}{q}} \leq C 2^{k(2p(1-(1-\frac{1}{q})d))}. \tag{3.24}$$

Here and below, C denotes a generic, non-negative random variable with the property that $\langle C^m \rangle \lesssim 1$ for all $m < \infty$, where $\lesssim$ means $\leq$ up to a constant that only depends on m, p, q, Λ and d. Combining (3.23) and (3.24) yields

$$\left\langle \phi_T^{2p} \right\rangle^{\frac{1}{p}} \lesssim \sum_{k\in\mathbb{N}_0} 2^{2k(1-(1-\frac{1}{q})d)} \left(\sum_{\mathrm{b}\in A_k} \left\langle C\ (\nabla\phi_T(\mathrm{b})+e(\mathrm{b}))^{2p}\ \omega_0^{4p}(\mathrm{b}) \right\rangle \right)^{\frac{1}{p}}. \tag{3.25}$$

Next we apply a triple Hölder inequality in probability with exponents (θ,θ',θ'), where we choose $\theta = \frac{2p+1}{2p}$ (so that $2p\theta = 2p+1$). We have

$$\left\langle C\ (\nabla\phi_T(\mathrm{b})+e(\mathrm{b}))^{2p}\ \omega_0^{4p}(\mathrm{b}) \right\rangle \leq \left\langle (\nabla\phi_T(\mathrm{b})+e(\mathrm{b}))^{2p+1} \right\rangle^{\frac{2p}{2p+1}} \left\langle C^{\theta'} \right\rangle^{\frac{1}{\theta'}} \left\langle \omega_0^{4p\theta'}(\mathrm{b}) \right\rangle^{\frac{1}{\theta'}}.$$

The first term is estimated by stationarity of $\nabla\phi_T$ and the assumption $|e| = 1$ as

$$\left\langle (\nabla\phi_T(\mathrm{b})+e(\mathrm{b}))^{2p+1} \right\rangle^{\frac{2p}{2p+1}} \lesssim \max_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle |\nabla\phi_T(\mathrm{b}')|^{2p+1} \right\rangle^{\frac{2p}{2p+1}} + 1.$$

For the second term we have $\left\langle C^{\theta'} \right\rangle^{\frac{1}{\theta'}} \left\langle \omega_0^{4p\theta'}(\mathrm{b}) \right\rangle^{\frac{1}{\theta'}} \lesssim 1$ due to (A2+), so that we obtain

$$\left\langle C\ (\nabla\phi_T(\mathrm{b})+e(\mathrm{b}))^{2p}\ \omega_0^{4p}(\mathrm{b}) \right\rangle \lesssim \max_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle |\nabla\phi_T(\mathrm{b}')|^{2p+1} \right\rangle^{\frac{2p}{2p+1}} + 1. \tag{3.26}$$

Combined with (3.25) we get

$$\begin{aligned} \left\langle \phi_T^{2p} \right\rangle^{\frac{1}{p}} &\lesssim \left(\max_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle |\nabla\phi_T(\mathrm{b}')|^{2p+1} \right\rangle^{\frac{2}{2p+1}} + 1 \right) \times \sum_{k\in\mathbb{N}_0} 2^{2k(1-(1-\frac{1}{q})d)} |A_k|^{\frac{1}{p}} \\ &\lesssim \left(\max_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle |\nabla\phi_T(\mathrm{b}')|^{2p+1} \right\rangle^{\frac{2}{2p+1}} + 1 \right). \end{aligned}$$

In the last line we used that

$$\sum_{k\in\mathbb{N}_0} 2^{2k(1-(1-\frac{1}{q})d)} |A_k|^{\frac{1}{p}} \lesssim \sum_{k\in\mathbb{N}_0} 2^{2k(1-(1-\frac{1}{2p}-\frac{1}{q})d)} \lesssim 1,$$

which holds since the exponent is negative, cf. (3.22). This proves (3.21). □

Proof of Corollary 3.3. Since the estimate of Theorem 3.2 is uniform in T, we recover $\phi \in (\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$ from ϕ_T by taking the limit $T \uparrow \infty$. To prove uniqueness, we only need to show that every $u \in (\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$ that satisfies

$$\nabla^* \boldsymbol{a} \nabla u = 0, \qquad \langle u \rangle = 0, \tag{3.27}$$

must be zero. This can be seen as follows: Fix $p > d+1$. Then for any $\mathrm{b} \in \mathbb{B}^d$ we have

$$\langle |\nabla u(\mathrm{b})| \rangle \leq \left\langle |\nabla u(\mathrm{b})|^2 \mathrm{dist}_{\boldsymbol{a}}^{-p}(x_{\mathrm{b}}, y_{\mathrm{b}}) \right\rangle^{\frac{1}{2}} \left\langle \mathrm{dist}_{\boldsymbol{a}}^{p}(x_{\mathrm{b}}, y_{\mathrm{b}}) \right\rangle,$$

which combined with Lemma (3.15) and assumption (A2) turns into

$$\langle |\nabla u(\mathrm{b})| \rangle \leq C \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1\ldots d}} \left\langle \boldsymbol{a}(\mathrm{b}') |\nabla u(\mathrm{b}')|^2 \right\rangle$$

for some positive constant C. Thanks to (3.27) and stationarity, on the right-hand side we may integrate by parts and get

$$\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1\ldots d}} \left\langle \boldsymbol{a}(\mathrm{b}') |\nabla u(\mathrm{b}')|^2 \right\rangle = \langle u \nabla^* \boldsymbol{a} \nabla u \rangle = 0.$$

Hence, for all $\mathrm{b} \in \mathbb{B}^d$ we have $\langle |\nabla u(\mathrm{b})| \rangle = 0$, and thus ergodicity implies that $u(\boldsymbol{a}, \cdot)$ is a constant for almost every $\boldsymbol{a} \in \Omega$. Now, $\langle u \rangle = 0$ implies $u \equiv 0$ in $(\mathcal{S}, \|\cdot\|_{L^2(\Omega)})$. □

4 Proofs of the auxiliary lemmas

4.1 Proof of Lemma 3.7

The argument for Lemma 3.7 is split into three lemmas. In the first lemma we gather a couple of standard formulas for derivatives of ϕ_T and G_T, where the derivative is taken w. r. t. to a coefficent $\boldsymbol{a}(\mathrm{b})$. Taking derivatives w.r.t. coefficients is a standard method in the context of quantitative stochastic homogenization. It is also used to study central limit theorems and related questions for random conductance models. Formulas, similar to those in Lemma 4.1 below, have been obtained e.g. in [28, 30, 6]. In the form stated below, the formulas can be found in [18].

A pecularity of the situation considered below, is the fact that we treat coefficients $\boldsymbol{a} \in \Omega$ that might vanish for an infinite number of edges, including the extreme case of $\boldsymbol{a} \equiv 0$. Nevertheless, we can differentiate w. r. t. coefficients thanks to the regularizing effect of the 0th order term in the modified corrector equation (3.1) and the equation for G_T, respectively. The reason is the following: As a consequence of the maximum principle we have the estimate

$$\forall \boldsymbol{a} \in \Omega : \qquad \sum_{x \in \mathbb{Z}^d} |G_T(\boldsymbol{a}, x, 0)| = T, \tag{4.1}$$

which follows from the non-negativity of G_T and testing the equation for G_T with test-functions that approximate the constant function $x \mapsto 1$. Thanks to (4.1) for $T > 0$ and all $\boldsymbol{a} \in \Omega$ the modified corrector can be expressed via the convolution

$$\phi_T(\boldsymbol{a}, x) := \sum_{y \in \mathbb{Z}^d} G_T(x, y) \nabla^*(\boldsymbol{a} e)(y).$$

Since (4.1) is uniform in $\boldsymbol{a} \in \Omega$, a direct calculation shows that for all $\boldsymbol{a} \in \Omega$, $\mathrm{b} \in \mathbb{B}^d$ and $x \in \mathbb{Z}^d$ the function

$$[0, 1] \ni a \mapsto \phi_T(\boldsymbol{a}^{\mathrm{b},a}, x) \in \mathbb{R}$$

belongs to $C^2([0, 1])$ (in fact it is smooth up to the boundary). By discreteness, the same is true for $\nabla\phi_T$ and similar arguments show that G_T, ∇G_T, $\nabla\nabla G_T$ feature the same regularity.

Lemma 4.1. *Let $\mathrm{b} \in \mathbb{B}^d$ be fixed. For $T > 0$ let ϕ_T and G_T denote the modified corrector and the Green's function, respectively. Then*

$$\frac{\partial \phi_T(x=0)}{\partial \boldsymbol{a}(\mathrm{b})} = -\nabla G_T(\mathrm{b}, 0)(\nabla\phi_T(\mathrm{b}) + e(\mathrm{b})), \tag{4.2}$$

$$\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})} \frac{\partial \phi_T(x=0)}{\partial \boldsymbol{a}(\mathrm{b})} = -2\nabla\nabla G_T(\mathrm{b}, \mathrm{b}) \frac{\partial \phi_T(x=0)}{\partial \boldsymbol{a}(\mathrm{b})}, \tag{4.3}$$

$$\frac{\partial}{\partial \boldsymbol{a}(b)} \nabla\nabla G_T(\mathrm{b}, \mathrm{b}) = -\left(\nabla\nabla G_T(\mathrm{b}, \mathrm{b})\right)^2. \tag{4.4}$$

Moreover, $\nabla\nabla G_T(\mathrm{b}, \mathrm{b})$ and $1 - \boldsymbol{a}(\mathrm{b})\nabla\nabla G_T(\mathrm{b}, \mathrm{b})$ are strictly positive.

Proof of Lemma 4.1. For simplicity we write ϕ and G instead of ϕ_T and G_T.

Step 1. Argument for (4.4).

We first claim that

$$\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})} G(x, y) = -\nabla G(\mathrm{b}, y)\nabla G(\mathrm{b}, x), \tag{4.5a}$$

$$\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})} \nabla G(x, \mathrm{b}) = -\nabla\nabla G(\mathrm{b}, \mathrm{b})\nabla G(\mathrm{b}, x). \tag{4.5b}$$

Indeed, since ∇ and $\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}$ commute, an application of $\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}$ to (3.8) yields

$$\left(\frac{1}{T}+\nabla^*\boldsymbol{a}\nabla\right)\frac{\partial G(\cdot,y)}{\partial \boldsymbol{a}(\mathrm{b})} = -\nabla^*\frac{\partial \boldsymbol{a}(\cdot)}{\partial \boldsymbol{a}(\mathrm{b})}\nabla G(\cdot,y). \tag{4.6}$$

We test this identity with $G(\cdot,x)$:

$$\begin{aligned}\frac{\partial G(x,y)}{\partial \boldsymbol{a}(\mathrm{b})} \quad &= \quad \sum_{y'\in\mathbb{Z}^d}\frac{\partial G(y',y)}{\partial \boldsymbol{a}(\mathrm{b})}\delta(x-y') \\ &\overset{(3.8)}{=} \quad \sum_{y'\in\mathbb{Z}^d}\frac{\partial G(y',y)}{\partial \boldsymbol{a}(\mathrm{b})}\left(\frac{1}{T}+\nabla^*\boldsymbol{a}\nabla\right)G(y',x) \\ &\overset{(2.1)}{=} \quad \sum_{y'\in\mathbb{Z}^d}G(y',x)\left(\frac{1}{T}+\nabla^*\boldsymbol{a}\nabla\right)\frac{\partial G(y',y)}{\partial \boldsymbol{a}(\mathrm{b})} \\ &\overset{(4.6),(2.1)}{=} \quad -\sum_{\mathrm{b}'\in\mathbb{B}^d}\frac{\partial \boldsymbol{a}(\mathrm{b}')}{\partial \boldsymbol{a}(\mathrm{b})}\nabla G(\mathrm{b}',y)\,\nabla G(\mathrm{b}',x).\end{aligned} \tag{4.7}$$

Since $\frac{\partial \boldsymbol{a}(\mathrm{b}')}{\partial \boldsymbol{a}(\mathrm{b})}$ is equal to 1 if $\mathrm{b}'=\mathrm{b}$ and 0 else, the sum on the right-hand side reduces to $\nabla G(\mathrm{b},y)\nabla G(\mathrm{b},x)$ and we get (4.5a). An application of ∇ to (4.5a) yields (4.5b), and an application of ∇ to (4.5b) finally yields (4.4).

Step 2. Argument for (4.2) and (4.3).

We apply $\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}$ to the modified corrector equation (3.1):

$$\frac{1}{T}\frac{\partial \phi}{\partial \boldsymbol{a}(\mathrm{b})}+\nabla^*\boldsymbol{a}\nabla\frac{\partial \phi}{\partial \boldsymbol{a}(\mathrm{b})} = -\nabla^*\frac{\partial \boldsymbol{a}(\cdot)}{\partial \boldsymbol{a}(\mathrm{b})}(\nabla\phi+e(\mathrm{b})). \tag{4.8}$$

As in (4.7) testing with $G(\cdot,x)$ yields

$$\frac{\partial \phi(x)}{\partial \boldsymbol{a}(\mathrm{b})} = -(\nabla\phi(\mathrm{b})+e(\mathrm{b}))\nabla G(\mathrm{b},x), \tag{4.9}$$

and (4.2) follows. By applying $\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}$ and ∇ to (4.9) we obtain the two identities

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}\frac{\partial \phi(x)}{\partial \boldsymbol{a}(\mathrm{b})} \quad &= \quad -\frac{\partial(\nabla\phi(\mathrm{b})+e(\mathrm{b}))}{\partial \boldsymbol{a}(\mathrm{b})}\nabla G(\mathrm{b},x)-(\nabla\phi(\mathrm{b})+e(\mathrm{b}))\frac{\partial \nabla G(\mathrm{b},x)}{\partial \boldsymbol{a}(\mathrm{b})}, \\ \nabla\frac{\partial \phi(\mathrm{b})}{\partial \boldsymbol{a}(\mathrm{b})} \quad &= \quad -(\nabla\phi(\mathrm{b})+e(\mathrm{b}))\nabla\nabla G(\mathrm{b},\mathrm{b}).\end{aligned}$$

By combining the first with the second identity, (4.5b) and (4.9) we get

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{a}(\mathrm{b})}\frac{\partial \phi(x)}{\partial \boldsymbol{a}(\mathrm{b})} \quad &= \quad 2(\nabla\phi(\mathrm{b})+e(\mathrm{b}))\nabla\nabla G(\mathrm{b},\mathrm{b})\nabla G(\mathrm{b},x) \\ &= \quad -2\frac{\partial \phi(x)}{\partial \boldsymbol{a}(\mathrm{b})}\nabla\nabla G(\mathrm{b},\mathrm{b}),\end{aligned}$$

and thus (4.3).

Step 3. Positivity of $\nabla\nabla G(\mathrm{b},\mathrm{b})$ and $1-\boldsymbol{a}(\mathrm{b})\nabla\nabla G(\mathrm{b},\mathrm{b})$.

Let $\mathrm{b}=(x_{\mathrm{b}},y_{\mathrm{b}})\in\mathbb{B}^d$ be fixed. An application of ∇ (w. r. t. the y-component) to (3.8) yields

$$\left(\frac{1}{T}+\nabla^*\boldsymbol{a}\nabla\right)\nabla G(\cdot,\mathrm{b}) = \delta(\cdot-y_{\mathrm{b}})-\delta(\cdot-x_{\mathrm{b}}). \tag{4.10}$$

We test this equation with $\nabla G(\cdot,\mathrm{b})$ and get

$$\frac{1}{T}\sum_{x\in\mathbb{Z}^d}|\nabla G(x,\mathrm{b})|^2+\sum_{\mathrm{b}'\in\mathbb{B}^d}\boldsymbol{a}(\mathrm{b}')\,|\nabla\nabla G(\mathrm{b}',\mathrm{b})|^2=\nabla\nabla G(\mathrm{b},\mathrm{b}). \tag{4.11}$$

This identity implies that $\nabla\nabla G(\mathrm{b},\mathrm{b})$ and $1-\boldsymbol{a}(\mathrm{b})\nabla\nabla G(\mathrm{b},\mathrm{b})$ are strictly positive. Indeed, first notice that $\sum_{x\in\mathbb{Z}^d}|\nabla G(x,\mathrm{b})|^2$ must be strictly positive, since otherwise $\nabla G(\cdot,\mathrm{b})=0$ in contradiction to (4.10). Combined with (4.11) we deduce that $\nabla\nabla G(\mathrm{b},\mathrm{b})$ must also be strictly positive. The strict positivity of $1-\boldsymbol{a}(\mathrm{b})\nabla\nabla G(\mathrm{b},\mathrm{b})$ follows from the strict positivity of $\nabla\nabla G(\mathrm{b},\mathrm{b})-\boldsymbol{a}(\mathrm{b})\,|\nabla\nabla G(\mathrm{b},\mathrm{b})|^2$. The latter can be seen by the following argument:

$$\begin{aligned}\nabla\nabla G(\mathrm{b},\mathrm{b})-\boldsymbol{a}(\mathrm{b})\,|\nabla\nabla G(\mathrm{b},\mathrm{b})|^2\quad&\geq\quad\nabla\nabla G(\mathrm{b},\mathrm{b})-\sum_{\mathrm{b}'\in\mathbb{B}^d}\boldsymbol{a}(\mathrm{b}')\,|\nabla\nabla G(\mathrm{b}',\mathrm{b})|^2\\&\overset{(4.11)}{=}\quad\frac{1}{T}\sum_{x\in\mathbb{Z}^d}|\nabla G(x,\mathrm{b})|^2>0.\end{aligned}$$

□

The next lemma establishes a (quantitative) link between the vertical and classical partial derivative of ϕ_T.

Lemma 4.2. *Let $\mathrm{b}\in\mathbb{B}^d$ be fixed. For $T>0$ let ϕ_T and G_T denote the modified corrector and the Green's function. Then*

$$\left|\frac{\partial\phi_T(x=0)}{\partial\mathrm{b}}\right|\leq\left(1+\frac{\boldsymbol{a}(\mathrm{b})}{1-\boldsymbol{a}(\mathrm{b})\nabla\nabla G_T(\mathrm{b},\mathrm{b})}\right)\left|\frac{\partial\phi_T(x=0)}{\partial\boldsymbol{a}(\mathrm{b})}\right|. \tag{4.12}$$

Proof of Lemma 4.2. Fix $\boldsymbol{a}\in\Omega$ and $\mathrm{b}\in\mathbb{B}^d$. Set $a_0:=\boldsymbol{a}(\mathrm{b})$. We shall use the following shorthand notation

$$\varphi(a):=\frac{\partial\phi_T(\boldsymbol{a}^{\mathrm{b},a},x=0)}{\partial\boldsymbol{a}(\mathrm{b})},\qquad g(a):=\nabla\nabla G_T(\boldsymbol{a}^{\mathrm{b},a},\mathrm{b},\mathrm{b}),\qquad(a\in[0,1]), \tag{4.13}$$

where $\boldsymbol{a}^{\mathrm{b},a}$ denotes the coefficient field obtained from $\boldsymbol{a}$ by setting $\boldsymbol{a}^{\mathrm{b},a}(\mathrm{b}')=a$ if $\mathrm{b}'=\mathrm{b}$ and $\boldsymbol{a}^{\mathrm{b},a}(\mathrm{b}'):=\boldsymbol{a}(\mathrm{b}')$ else. With that notation (4.3) and (4.4) turn into

$$\varphi'=-2g\varphi, \tag{4.14}$$
$$g'=-g^2. \tag{4.15}$$

Since we have $\left|\frac{\partial\phi_T(x=0)}{\partial\mathrm{b}}\right|\leq\int_0^1|\varphi(a)|\,da$, it suffices to show

$$\int_0^1|\varphi(a)|\,da\leq\left(1+\frac{a_0}{1-a_0g(a_0)}\right)|\varphi(a_0)|. \tag{4.16}$$

The positivity of g and (4.14) imply that φ is either strictly positive, strictly negative or that it vanishes identically. In the latter case, the claim is trivial. In the other cases we have

$$\varphi(a)=\exp(h(a))\varphi(a_0),\qquad\text{where } h(a):=\ln\frac{\varphi(a)}{\varphi(a_0)},$$

and (4.16) reduces to the inequality

$$\int_0^1\exp(h(a))\,da\leq1+\frac{a_0}{1-a_0g(a_0)}. \tag{4.17}$$

From (4.14) we learn that $h' = -2g$. Since $g > 0$, h is decreasing. Combined with the identity $h(a_0) = 0$ we get

$$h(a) \leq \begin{cases} 2\int_a^{a_0} g(a')\,da' & \text{for } a \in [0, a_0), \\ 0 & \text{for } a \in [a_0, 1]. \end{cases} \tag{4.18}$$

On the other hand, we learn from integrating (4.15) that $g(a') = \frac{g(a_0)}{1+(a'-a_0)g(a_0)}$. Hence, for $a < a_0$ the right-hand side in (4.18) turns into

$$2\int_a^{a_0} g(a')\,da' = -2\ln(1 + (a - a_0)g(a_0)),$$

which in combination with (4.18) yields (4.17). □

Lemma 3.7 is a direct consequence of (4.12), (4.2) and the following estimate:

Lemma 4.3. *Let G_T denote the Green's function. Assume that (A1) is satisfied. Then for all $T > 0$, $\boldsymbol{a} \in \Omega$ and $\mathrm{b} \in \mathbb{B}^d$ we have*

$$1 + \frac{\boldsymbol{a}(b)}{1 - \boldsymbol{a}(b)\nabla\nabla G_T(\mathrm{b},\mathrm{b})} \lesssim \omega_0^2(\boldsymbol{a}, \mathrm{b}), \tag{4.19}$$

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

Proof of Lemma 4.3. Step 1. Reduction to an estimate for $\boldsymbol{a}^{\mathrm{b},0}$.

We claim that

$$\frac{\boldsymbol{a}(\mathrm{b})}{1 - \boldsymbol{a}(\mathrm{b})\nabla\nabla G_T(\boldsymbol{a},\mathrm{b},\mathrm{b})} \leq (1 + \nabla\nabla G_T(\boldsymbol{a}^{\mathrm{b},0},\mathrm{b},\mathrm{b}))^2$$

For the argument let $\boldsymbol{a} \in \Omega$ and $\mathrm{b} \in \mathbb{B}^d$ be fixed. With the shorthand notation introduced in (4.13), the claim reads

$$\frac{a_0}{1 - a_0 g(a_0)} \leq (1 + g(0))^2. \tag{4.20}$$

For $a_0 = 0$ the statement is trivial. For $a_0 > 0$ consider the function

$$f(a) := \frac{1}{a}g(a) - g^2(a),$$

with help of which the left-hand side in (4.20) can be written as $\frac{g(a_0)}{f(a_0)}$. The function f is non-negative and decreasing, as can be seen by combining the inequality $0 < g(a) < \frac{1}{a}$ from Lemma 4.1 with the identity $f'(a) = g(a)(g^2(a) - \frac{1}{a^2} + g^2(a) - \frac{1}{a}g(a))$ which follows from (4.15). The latter also implies that $g(1) = \frac{g(0)}{1+g(0)}$ and thus $f(1) = g(1)(1 - g(1)) = \frac{g(0)}{(1+g(0))^2}$. Hence,

$$\frac{a_0}{1 - a_0 g(a_0)} = \frac{g(a_0)}{f(a_0)} \leq \frac{g(a_0)}{f(1)} = (1 + g(0))^2 \frac{g(a_0)}{g(0)} \leq (1 + g(0))^2;$$

in the last step we used in addition that $g(a_0) \leq g(0)$ which is a consequence of (4.15).

Step 2. Conclusion.

To complete the argument we only need to show that

$$\nabla\nabla G_T(\boldsymbol{a}^{\mathrm{b},0},\mathrm{b},\mathrm{b}) \lesssim \omega_0(\boldsymbol{a},\mathrm{b}). \tag{4.21}$$

For simplicity set $\boldsymbol{a}_0 := \boldsymbol{a}^{\mathrm{b},0}$. Note that $\omega_0(\boldsymbol{a},\mathrm{b}) = \omega(\boldsymbol{a}_0,\mathrm{b})$. From (4.11) and Lemma 3.10 (which we prove below) we obtain

$$\begin{aligned}\nabla\nabla G_T(\boldsymbol{a}_0,\mathrm{b},\mathrm{b}) &\overset{(4.11)}{\geq} \sum_{\mathrm{b}'\in\mathbb{B}^d} \boldsymbol{a}_0(\mathrm{b}')\left(\nabla\nabla G_T(\boldsymbol{a}_0,\mathrm{b}',\mathrm{b})\right)^2 \\ &\overset{(3.15)}{\gtrsim} \sum_{\mathrm{b}'\in\mathbb{B}^d} \omega^{-1}(\boldsymbol{a}_0,\mathrm{b}')\left(\nabla\nabla G_T(\boldsymbol{a}_0,\mathrm{b}',\mathrm{b})\right)^2 \\ &\geq \omega^{-1}(\boldsymbol{a}_0,\mathrm{b})\left(\nabla\nabla G_T(\boldsymbol{a}_0,\mathrm{b},\mathrm{b})\right)^2.\end{aligned}$$

Dividing both sides by $\omega^{-1}(\boldsymbol{a}_0,\mathrm{b})\nabla\nabla G_T(\boldsymbol{a}_0,\mathrm{b},\mathrm{b})$ yields (4.21). □

4.2 Proof of Lemma 3.10 and Lemma 3.11

Proof of Lemma 3.10. Fix for a moment $\boldsymbol{a}\in\Omega$. For every $\mathrm{b}\in\mathbb{B}^d$ with $\mathrm{dist}_{\boldsymbol{a}}(x_\mathrm{b},y_\mathrm{b})<\infty$, let $\pi_{\boldsymbol{a}}(\mathrm{b})$ denote a shortest open path (arbitrary but fixed from now on) that connects x_b and y_b, i.e.

$$\mathrm{dist}_{\boldsymbol{a}}(x_\mathrm{b},y_\mathrm{b}) = \sum_{\mathrm{b}'\in\pi_{\boldsymbol{a}}(\mathrm{b})}\frac{1}{\boldsymbol{a}(\mathrm{b}')}.$$

Thanks to the triangle inequality and the Cauchy-Schwarz inequality we have

$$\begin{aligned}|\nabla u(\mathrm{b})| &\leq \sum_{\mathrm{b}'\in\pi(\mathrm{b})}|\nabla u(\mathrm{b}')| \leq \left(\sum_{\mathrm{b}'\in\pi_{\boldsymbol{a}}(\mathrm{b})}\frac{1}{\boldsymbol{a}(\mathrm{b}')}\right)^{\frac12}\left(\sum_{\mathrm{b}'\in\pi_{\boldsymbol{a}}(\mathrm{b})}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right)^{\frac12} \\ &= \mathrm{dist}_{\boldsymbol{a}}^{\frac12}(x_\mathrm{b},y_\mathrm{b})\left(\sum_{\mathrm{b}'\in\pi_{\boldsymbol{a}}(\mathrm{b})}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right)^{\frac12}.\end{aligned}$$

Hence, using the convention $\frac{1}{\infty}=0$, we conclude that for all $\mathrm{b}\in\mathbb{B}^d$ and $\boldsymbol{a}\in\Omega$:

$$\mathrm{dist}_{\boldsymbol{a}}^{-p}(x_\mathrm{b},y_\mathrm{b})|\nabla u(\mathrm{b})|^2 \leq \mathrm{dist}_{\boldsymbol{a}}^{1-p}(x_\mathrm{b},y_\mathrm{b})\sum_{\mathrm{b}'\in\pi_{\boldsymbol{a}}(\mathrm{b})}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}'). \tag{4.22}$$

We drop the "$\boldsymbol{a}$" in the notation from now on. Summation (4.22) over $\mathrm{b}\in\mathbb{B}^d$ yields

$$\begin{aligned}\sum_{\mathrm{b}\in\mathbb{B}^d}\mathrm{dist}^{-p}(x_\mathrm{b},y_\mathrm{b})|\nabla u(\mathrm{b})|^2 &\leq \sum_{\mathrm{b}\in\mathbb{B}^d}\sum_{\mathrm{b}'\in\pi(\mathrm{b})}\mathrm{dist}^{1-p}(x_\mathrm{b},y_\mathrm{b})|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}') \\ &= \sum_{\mathrm{b}'\in\mathbb{B}^d}\sum_{\substack{\mathrm{b}\in\mathbb{B}^d\text{ with}\\ \pi(\mathrm{b})\ni\mathrm{b}'}}\mathrm{dist}^{1-p}(x_\mathrm{b},y_\mathrm{b})|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}').\end{aligned}$$

Since $\pi(\mathrm{b})$ is a shortest path, and because $\boldsymbol{a}\leq 1$, we have $\mathrm{dist}(x_\mathrm{b},y_\mathrm{b})\geq|x_\mathrm{b}-x_{\mathrm{b}'}|+1$ for all $\mathrm{b},\mathrm{b}'\in\mathbb{B}^d$ with $\mathrm{b}'\in\pi(\mathrm{b})$. Combined with the previous estimate we get

$$\begin{aligned}\sum_{\mathrm{b}\in\mathbb{B}^d}\mathrm{dist}^{-p}(x_\mathrm{b},y_\mathrm{b})|\nabla u(\mathrm{b})|^2 &\leq \sum_{\mathrm{b}'\in\mathbb{B}^d}\sum_{\substack{\mathrm{b}\in\mathbb{B}^d\text{ with}\\ \pi(\mathrm{b})\ni\mathrm{b}'}}(|x_\mathrm{b}-x_{\mathrm{b}'}|+1)^{1-p}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}') \\ &\leq C(d,p)\sum_{\mathrm{b}'\in\mathbb{B}^d}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}').\end{aligned}$$

□

Proof of Lemma 3.11. Fix $\mathrm{b}\in\mathbb{B}^d$. For $L\in\mathbb{N}$ consider the indicator function

$$\chi_L(\boldsymbol{a}) := \begin{cases}1 & \text{if } L\leq\mathrm{dist}_{\boldsymbol{a}}(x_\mathrm{b},y_\mathrm{b})<2L,\\ 0 & \text{else.}\end{cases} \tag{4.23}$$

With the convention $\frac{1}{\infty}=0$, we have

$$\sum_{k=0}^{\infty}\chi_{2^k}(\boldsymbol{a})\mathrm{dist}_{\boldsymbol{a}}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})=\mathrm{dist}_{\boldsymbol{a}}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}}) \tag{4.24}$$

for all $\boldsymbol{a}\in\Omega$. In the following we drop "$\boldsymbol{a}$" in the notation. We recall (4.22) in the form of

$$\chi_L\mathrm{dist}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})|\nabla u(\mathrm{b})|^2 \;\le\; \chi_L\mathrm{dist}^{1-p}(x_{\mathrm{b}},y_{\mathrm{b}})\sum_{\mathrm{b}'\in\pi(\mathrm{b})}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}'). \tag{4.25}$$

From $\boldsymbol{a}\le 1$ and $\mathrm{dist}(x_{\mathrm{b}},y_{\mathrm{b}})<2L$ for $\chi_L\neq 0$, cf. (4.23), we learn that $\pi(\mathrm{b})$ is contained in the box $Q_{2L}(x_{\mathrm{b}})$. Hence, (4.25) turns into

$$\chi_L\mathrm{dist}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})|\nabla u(\mathrm{b})|^2 \overset{(4.23)}{\le} \chi_L L^{1-p}\sum_{\mathrm{b}'\in Q_{2L}(x_{\mathrm{b}})}|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}').$$

We take the expectation on both sides and appeal to stationarity:

$$\begin{aligned}
\left\langle\chi_L\mathrm{dist}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})|\nabla u(\mathrm{b})|^2\right\rangle &\le L^{1-p}\sum_{\mathrm{b}'\in Q_{2L}(x_{\mathrm{b}})}\left\langle\chi_L|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle\\
&\overset{\chi_L\le 1}{\le} L^{1-p}\sum_{x\in B_{2L}(x_{\mathrm{b}})}\sum_{\substack{\mathrm{b}'=\{x,x+e_i\}\\ i=1,\ldots,d}}\left\langle|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle\\
&\overset{\text{stationarity}}{\le} L^{1-p}|B_{2L}(0)|\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}}\left\langle|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle.
\end{aligned}$$

Using $1+d-p<0$ we get

$$\begin{aligned}
\left\langle\mathrm{dist}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})|\nabla u(\mathrm{b})|^2\right\rangle &\overset{(4.24)}{=} \sum_{k=0}^{\infty}\left\langle\chi_{2^k}\mathrm{dist}^{-p}(x_{\mathrm{b}},y_{\mathrm{b}})|\nabla u(\mathrm{b})|^2\right\rangle\\
&\le C(p,d)\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}}\left\langle|\nabla u(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle.
\end{aligned}$$

□

4.3 Proof of Proposition 3.8 – Green's function estimates

We first establish an estimate for the Green's function itself:

Lemma 4.4. *Let $d\ge 2$ and consider $u,f\in\ell^1(\mathbb{Z}^d)$ with*

$$\nabla^*\boldsymbol{a}\nabla u=f \qquad \text{in } \mathbb{Z}^d. \tag{4.26}$$

Then for all $\frac{2d}{d+2}<p<2$, $R\ge 1$ and $x_0\in\mathbb{Z}^d$ we have

$$\sum_{x\in B_R(x_0)}|u(x)-\bar u|\lesssim C\,R^2\sum_{x\in\mathbb{Z}^d}|f(x)|. \tag{4.27}$$

Here, $\bar u:=\frac{1}{|B_R(x_0)|}\sum_{x\in B_R(x_0)}u(x)$ denotes the average of u on $B_R(x_0)$, $C:=C(\boldsymbol{a},Q_R(x_0),\frac{p}{2-p})$, and $\lesssim$ means $\le$ up to a constant that only depends on d and p.

Proof of Lemma 4.4. W. l. o. g. we assume $\sum_{\mathbb{Z}^d}|f|=1$ and $R\in\mathbb{N}$. To shorten the notation we write B_R and Q_R for $B_R(x_0)$ and $Q_R(x_0)$, respectively. We start with a small reduction step. Let $M(u)$ denote a median of u on B_R, i. e.

$$|\{u\ge M(u)\}\cap B_R|,|\{u\le M(u)\}\cap B_R|\ge\frac{1}{2}|B_R|.$$

By Jensen's inequality we have $|\bar{u} - M(u)| \leq \frac{1}{|B_R|}\sum_{B_R}|u - M(u)|$, so that it suffices to prove for $v := u - M(u)$ the estimate

$$\sum_{B_R}|v| \lesssim C\,R^2 \sum_{\mathbb{Z}^d}|f| = C\,R^2.$$

Note that by construction v satisfies

$$|\{v \geq 0\} \cap B_R|, |\{v \leq 0\} \cap B_R| \geq \frac{1}{2}|B_R|, \tag{4.28}$$

which is the only reason why we introduced the shift by $M(u)$ (which we may forget about from now on). Also notice that (4.28) is symmetric in the sense that we might replace v by $-v$ without changing the statement.

The proof of the sought for estimate invokes a truncation of peaks of v above a threshold that we denote by M. Moreover, we consider the positive and negative part of v separately. In fact, by symmetry of (4.28), it suffices to consider the positive part of v (for the negative part consider $-v$). In conclusion, for $0 \leq M < \infty$ we consider the cut-off version of v

$$v_M := \max\{\min\{v, M\}, 0\}.$$

Then v_M satisfies

$$\sum_{\mathbb{B}^d}\nabla v_M\,\boldsymbol{a}\nabla v_M = \sum_{\mathbb{B}^d}\nabla u\,\boldsymbol{a}\nabla v_M,$$

since either $\nabla v_M(\mathrm{b}) = \nabla u(\mathrm{b})$ or $\nabla v_M(\mathrm{b}) = 0$. Since $u \in \ell^1(\mathbb{Z}^d)$ (by assumption) and $v_M \in \ell^\infty(\mathbb{Z}^d)$ (by construction), we may integrate by parts:

$$\sum_{\mathbb{B}^d}\nabla u\,\boldsymbol{a}\nabla v_M = \sum_{\mathbb{Z}^d} v_M\,\nabla^*\boldsymbol{a}\nabla u = \sum_{\mathbb{Z}^d} f v_M \leq M\sum_{\mathbb{Z}^d}|f| = M.$$

Hence,

$$\sum_{\mathbb{B}^d}\nabla v_M\,\boldsymbol{a}\nabla v_M \leq M. \tag{4.29}$$

Set $p^* = \frac{pd}{d-p}$ and $q^* := \frac{p^*}{p^*-1}$. By (4.28) and the definition of v_M, we have $|\{v_M = 0\} \cap B_R| = |\{v_M \leq 0\} \cap B_R| \geq \frac{1}{2}|B_R|$. Hence, the Sobolev-Poincaré inequality yields

$$\left(R^{-d}\sum_{B_R}|v_M|^{p^*}\right)^{\frac{1}{p^*}} \lesssim R\left(R^{-d}\sum_{Q_R}|\nabla v_M|^p\right)^{\frac{1}{p}}.$$

Lemma 3.10 combined with Hölder's inequality with exponents $(\frac{2}{2-p}, \frac{2}{p})$ yields

$$\begin{aligned}
\left(R^{-d}\sum_{Q_R}|\nabla v_M|^p\right)^{\frac{1}{p}} &= \left(R^{-d}\sum_{Q_R}\omega^{\frac{p}{2}}|\nabla v_M|^p\omega^{-\frac{p}{2}}\right)^{\frac{1}{p}} \\
&\leq \left(R^{-d}\sum_{Q_R}\omega^{\frac{p}{2-p}}\right)^{\frac{2-p}{2p}}\left(R^{-d}\sum_{Q_R}|\nabla v_M|^2\omega^{-1}\right)^{\frac{1}{2}} \\
&\overset{\text{Lemma 3.10}}{\lesssim} C^{\frac{1}{2}}\left(R^{-d}\sum_{\mathbb{B}^d}\nabla v_M\,\boldsymbol{a}\nabla v_M\right)^{\frac{1}{2}},
\end{aligned} \tag{4.30}$$

so that

$$\left(R^{-d}\sum_{B_R}|v_M|^{p^*}\right)^{\frac{1}{p^*}} \lesssim C^{\frac{1}{2}}R\left(R^{-d}\sum_{\mathbb{B}^d}\nabla v_M\,\boldsymbol{a}\nabla v_M\right)^{\frac{1}{2}} \overset{(4.29)}{\lesssim} (CR^{2-d}M)^{\frac{1}{2}}. \tag{4.31}$$

Next we use Chebyshev's inequality in the form of

$$M\left(R^{-d}|\{\,v>M\,\}\cap B_R|\right)^{\frac{1}{p^*}} \lesssim \left(R^{-d}\sum_{B_R}|v_M|^{p^*}\right)^{\frac{1}{p^*}}.$$

With (4.31) we get

$$R^{-d}|\{\,v>M\,\}\cap B_R| \lesssim C^{\frac{p^*}{2}}\,R^{(2-d)\frac{p^*}{2}}M^{-\frac{p^*}{2}}.$$

By the symmetry in (4.28) we get the same estimate for $-v$ and thus arrive at

$$R^{-d}|\{\,|v|>M\,\}\cap B_R| \lesssim C^{\frac{p^*}{2}}\,R^{(2-d)\frac{p^*}{2}}M^{-\frac{p^*}{2}}.$$

Since $p>\frac{2d}{d+2}$ (by assumption), we have $\frac{p^*}{2}>1$ and the "wedding cake formula" for $M:=CR^{2-d}$ yields

$$\begin{aligned} R^{-d}\sum_{B_R}|v| &= \int_0^\infty R^{-d}|\{\,|v|>M'\,\}\cap B_R|\,dM' \lesssim M+\int_M^\infty R^{-d}|\{\,|v|>M'\,\}\cap B_R|\,dM' \\ &\lesssim M+C^{\frac{p^*}{2}}R^{(2-d)\frac{p^*}{2}}M^{1-\frac{p^*}{2}} \lesssim CR^{2-d}. \end{aligned}$$

□

A careful Caccioppoli estimate combined with the previous lemma yields:

Lemma 4.5. *Let $d\geq 2$, $x_0\in\mathbb{Z}^d$ and $R\geq 1$. Consider $f\geq 0$ and u related as*

$$\nabla^*\boldsymbol{a}\nabla u=-f \qquad \text{in } B_{2R}(x_0). \tag{4.32}$$

Then for $\frac{2d}{d+2}<p<2$ we have

$$\left(R^{-d}\sum_{Q_R(x_0)}|R\nabla u|^p\right)^{\frac{1}{p}} \lesssim C^{\frac{\alpha}{2}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|+\left(R^{2-d}\sum_{B_{2R}(x_0)}fu_-\right)^{\frac{1}{2}}\right), \tag{4.33}$$

where $u_-:=\max\{-u,0\}$ denotes the negative part of u, $C:=C(\boldsymbol{a},Q_{2R}(x_0),\frac{p}{2-p})$, $\alpha:=2\frac{p^*-1}{p^*-2}$ and $p^*:=\frac{dp}{d-p}$. Here $\lesssim$ stands for $\leq$ up to a constant that only depends on p and d

Proof of Lemma 4.5. Step 1. Caccioppoli estimate.

We claim that for every cut-off function η that is supported in $B_{2R-1}(x_0)$ (so that in particular $\nabla\eta=0$ outside of $Q_{2R}(x_0)$) we have

$$\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^p\right)^{\frac{1}{p}} \lesssim C^{\frac{1}{2}}\left(R^{2-d}\sum_{\mathbb{Z}^d}fu_-\eta^2+R^{-d}\sum_{\mathrm{b}\in\mathbb{B}^d}u(x_{\mathrm{b}})u(y_{\mathrm{b}})|R\nabla\eta(\mathrm{b})|^2\boldsymbol{a}(\mathrm{b})\right)^{\frac{1}{2}}. \tag{4.34}$$

Indeed, we get with Lemma 3.10 (using an argument similar to (4.30)):

$$\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^p\right)^{\frac{1}{p}} = \left(R^{-d}\sum_{Q_{2R}(x_0)}|R\nabla(u\eta)|^p\right)^{\frac{1}{p}} \lesssim C^{\frac{1}{2}}\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^2\boldsymbol{a}\right)^{\frac{1}{2}},$$

Combined with the elementary identity

$$|\nabla(u\eta)(\mathrm{b})|^2 = \nabla u(\mathrm{b})\nabla(u\eta^2)(\mathrm{b}) + u(x_\mathrm{b})u(y_\mathrm{b})|\nabla\eta(\mathrm{b})|^2,$$

the equation for u, and the fact that $-fu\eta^2 \le fu_-\eta^2$ (here we use $f \ge 0$), the claimed estimate (4.34) follows.

Step 2. Conclusion.

Set $\theta := \frac{\alpha-1}{\alpha}$ and note that α is defined in such a way that for the considered range of p we have

$$\frac{1}{2} = \theta\frac{1}{p^*} + (1-\theta) \qquad \text{and} \qquad 2(1-\theta) < 1. \tag{4.35}$$

As we shall see below in Step 3, there exists a cut-off function η with $\eta = 1$ in $B_{R+1}(x_0)$ and $\eta = 0$ outside of $B_{2R-1}(x_0)$, such that

$$\begin{aligned}\left(R^{-d}\sum_{\mathrm{b}\in\mathbb{B}^d}|u(x_\mathrm{b})||u(y_\mathrm{b})||R\nabla\eta(\mathrm{b})|^2\right)^{\frac12} \lesssim{}& \left(R^{-d}\sum_{\mathbb{Z}^d}|u\eta|^{p^*}\right)^{\frac{\theta}{p^*}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{1-\theta}\\ &+\left(R^{-d}\sum_{\mathbb{Z}^d}|u\eta|^{p^*}\right)^{\frac{1}{2p^*}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{\frac12}.\end{aligned} \tag{4.36}$$

Let us explain the right-hand side of this estimate. While the first term on the right-hand side would also appear in the continuum case (i.e. when $\mathbb{Z}^d$ is replaced by $\mathbb{R}^d$), the second term is an error term coming from discreteness. In fact, it is of lower order: A sharp look at (4.40) below shows that (4.36) holds with the vanishing factor $R^{-\epsilon}$ (for some $\epsilon > 0$ only depending on p and d) in front of the second term on the right-hand side.

To continue our estimate, we appeal to the Gagliardo-Nirenberg-Sobolev inequality on $\mathbb{Z}^d$, i.e. $\left(R^{-d}\sum_{\mathbb{Z}^d}|u\eta|^{p^*}\right)^{\frac{1}{p^*}} \lesssim \left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^p\right)^{\frac1p}$, which we apply to the right-hand side of (4.36):

$$\begin{aligned}&\left(R^{-d}\sum_{\mathbb{Z}^d}|u\eta|^{p^*}\right)^{\frac{\theta}{p^*}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{1-\theta}\\ &\quad+\left(R^{-d}\sum_{\mathbb{Z}^d}|u\eta|^{p^*}\right)^{\frac{1}{2p^*}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{\frac12}\\ \lesssim{}&\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^{p}\right)^{\frac{\theta}{p}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{1-\theta}\\ &\quad+\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^{p}\right)^{\frac{1}{2p}}\left(R^{-d}\sum_{B_{2R}(x_0)}|u|\right)^{\frac12}.\end{aligned}$$

We estimate both products on the right-hand side by appealing to Youngs' inequality (with exponents $(\frac{1}{\theta}, \frac{1}{1-\theta})$ and $(2,2)$, respectively), and find that for all $\delta > 0$ there exists a constant $C(\delta) > 0$ only depending on δ, p and d, such that

$$\left(CR^{-d}\sum_{\mathrm{b}\in\mathbb{B}^d}|u(x_{\mathrm{b}})||u(y_{\mathrm{b}})|(\nabla\eta(\mathrm{b}))^2\boldsymbol{a}(\mathrm{b})\right)^{\frac{1}{2}}$$
$$\leq \delta\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^p\right)^{\frac{1}{p}} + C(\delta)\left(C^{\frac{1}{2(1-\theta)}}R^{-d}\sum_{B_{2R}(x_0)}|u| + CR^{-d}\sum_{B_{2R}(x_0)}|u|\right)$$
$$\overset{2(1-\theta)<1}{\leq} \delta\left(R^{-d}\sum_{\mathbb{B}^d}|R\nabla(u\eta)|^p\right)^{\frac{1}{p}} + 2C(\delta)C^{\frac{1}{2(1-\theta)}}R^{-d}\sum_{B_{2R}(x_0)}|u|.$$

We combine this estimate with (4.34) and absorb the first term on the right-hand side of the previous estimate into the left-hand side of (4.34). Since $\nabla(\eta u) = \nabla u$ in $Q_R(x_0)$ this yields (4.33).

Step 3. Proof of (4.36).

We first construct a suitable cut-off function η for $B_{R+1}(x_0)$ in $B_{2R-1}(x_0)$. W. l. o. g. we assume that $x_0 = 0$. Recall that $\alpha = 2\frac{p^*-1}{p^*-2}$. For $t \geq 0$ set

$$\tilde{\eta}(t) := \max\{1 - 2\max\{\tfrac{t}{R+1} - 1, 0\}, 0\}^{\alpha},$$

and define

$$\eta(x) := \prod_{i=1}^{d}\tilde{\eta}(|x_i|). \tag{4.37}$$

Using the relation $\alpha - 1 = \theta\alpha$, cf. (4.35), it is straightforward to check that for all edges b with $|\nabla\eta(\mathrm{b})| > 0$ the function η satisfies:

$$R|\nabla\eta(\mathrm{b})| \lesssim \begin{cases}\min\{\eta^{\theta}(x_{\mathrm{b}}), \eta^{\theta}(y_{\mathrm{b}})\} & \text{if } \min\{\eta(x_{\mathrm{b}}), \eta(y_{\mathrm{b}}) > 0\},\\ R^{1-\alpha} & \text{if } \min\{\eta(x_{\mathrm{b}}), \eta(y_{\mathrm{b}})\} = 0.\end{cases} \tag{4.38}$$

Now we turn to (4.36). We split the sum into a “interior” and a “boundary” contribution:

$$\sum_{\mathrm{b}\in\mathbb{B}^d}|u(x_{\mathrm{b}})||u(y_{\mathrm{b}})|(\nabla\eta(\mathrm{b}))^2$$
$$= \sum_{\mathrm{b}\in A_{\text{int}}}|u(x_{\mathrm{b}})||u(y_{\mathrm{b}})|(\nabla\eta(\mathrm{b}))^2 + \sum_{\mathrm{b}\in A_{\text{bound}}}|u(x_{\mathrm{b}})||u(y_{\mathrm{b}})|(\nabla\eta(\mathrm{b}))^2,$$

where

$$A_{\text{int}} := \{\,\mathrm{b} : |\nabla\eta(\mathrm{b})| > 0 \text{ and } \min\{\eta(x_{\mathrm{b}}), \eta(y_{\mathrm{b}})\} > 0\,\},$$
$$A_{\text{bound}} := \{\,\mathrm{b} : |\nabla\eta(\mathrm{b})| > 0 \text{ and } \min\{\eta(x_{\mathrm{b}}), \eta(y_{\mathrm{b}})\} = 0\,\}.$$

For A_{int} we get with (4.38), Young's inequality, and Hölder's inequality with exponents $(p^*\frac{1}{2\theta}, \frac{1}{2(1-\theta)})$:

$$\begin{aligned} R^{-d}\sum_{\mathrm{b}\in A_{\text{int}}}|u(x_{\mathrm{b}})||u(y_{\mathrm{b}})||R\nabla\eta(\mathrm{b})|^2 &\lesssim R^{-d}\sum_{\mathbb{Z}^d}u^2\eta^{2\theta}\\ = R^{-d}\sum_{\mathbb{Z}^d}(u\eta)^{2\theta}u^{2(1-\theta)} &\leq \left(R^{-d}\sum_{\mathbb{Z}^d}(u\eta)^{p^*}\right)^{\frac{2\theta}{p^*}}\left(R^{-d}\sum_{B_{2R}}|u|\right)^{2(1-\theta)}. \end{aligned} \tag{4.39}$$

Next we treat A_{bound}, which is an error term coming from the discrete nature of the gradient. By the definition of A_{bound} the cut-off function η vanishes at one and only one of the two sites adjacent to $\mathrm{b} \in A_{\text{bound}}$. Given $\mathrm{b} \in A_{\text{bound}}$ we denote by $\tilde{x}_{\mathrm{b}}$ (resp. $\tilde{y}_{\mathrm{b}}$) the site adjacent to b with $\eta(\tilde{x}_{\mathrm{b}}) = 0$ (resp. $\eta(\tilde{y}_{\mathrm{b}}) \neq 0$), so that

$$R^{-d} \sum_{\mathrm{b}\in A_{\text{bound}}} |u(x_{\mathrm{b}})||u(y_{\mathrm{b}})||R\nabla\eta(\mathrm{b})|^2 = R^{1-d} \sum_{\mathrm{b}\in A_{\text{bound}}} |u(\tilde{x}_{\mathrm{b}})||u(\tilde{y}_{\mathrm{b}})|\eta(\tilde{y}_{\mathrm{b}})|R\nabla\eta(\mathrm{b})|.$$

We combine this with (4.38), Hölder's inequality with exponents $(p^*, q^* := \frac{p^*}{p^*-1})$, and the discrete ℓ^1-ℓ^{q^*}-estimate:

$$\begin{aligned}
R^{-d} \sum_{\mathrm{b}\in A_{\text{bound}}} |u(x_{\mathrm{b}})||u(y_{\mathrm{b}})||R\nabla\eta(\mathrm{b})|^2 &\lesssim R^{2-d-\alpha} \sum_{\mathrm{b}\in A_{\text{bound}}} |u(\tilde{x}_{\mathrm{b}})||u(\tilde{y}_{\mathrm{b}})|\eta(\tilde{y}_{\mathrm{b}}) \\
&\leq R^{2-d-\alpha} \left(\sum_{B_{2R}} |u\eta|^{p^*}\right)^{\frac{1}{p^*}} \left(\sum_{B_{2R}} |u|^{q^*}\right)^{\frac{1}{q^*}} \leq R^{2-d-\alpha} \left(\sum_{B_{2R}} |u\eta|^{p^*}\right)^{\frac{1}{p^*}} \sum_{B_{2R}} |u| \\
&= R^{\frac{d}{p^*}-2-\alpha} \left(R^{-d} \sum_{B_{2R}} |u\eta|^{p^*}\right)^{\frac{1}{p^*}} \left(R^{-d} \sum_{B_{2R}} |u|\right). \qquad (4.40)
\end{aligned}$$

From the definition of α and p^*, and the fact that $\alpha > 2$, we deduce that the exponent $\frac{d}{p^*} - 2 - \alpha$ is negative. Together with (4.39) the desired estimate (4.36) follows. ☐

Now we are ready to prove Proposition 3.8. We distinguish the cases $k \geq 1$ and $k = 0$.

Proof of Proposition 3.8. Step 1. Argument for $k \geq 1$.

For brevity set $R := 2^{k-1}R_0$ and recall that $A_k = Q_{2R}(0) \setminus Q_R(0)$. We cover the annulus A_k by boxes $Q_{\frac{R}{2}}(x_0)$, $x_0 \in X_R \subset \mathbb{Z}^d$, such that

$$A_k \subset \bigcup_{x_0\in X_R} Q_{\frac{R}{2}}(x_0) \subset \bigcup_{x_0\in X_R} Q_R(x_0) \subset Q_{3R}(0) \setminus \{0\}. \qquad (4.41)$$

Since the diameter of the annulus and the side length of the boxes are comparable, we may choose X_R such that its cardinality is bounded by a constant only depending on d. Since in addition we have for $x_0 \in X_R$ the inequality $C(\boldsymbol{a}, Q_R(x_0), \frac{p}{2-p}) \lesssim C(\boldsymbol{a}, Q_{3R}(0), \frac{p}{2-p})$ (thanks to the third inclusion in (4.41)), it suffices to prove

$$\left(R^{-d} \sum_{\mathrm{b}\in Q_{\frac{R}{2}}(x_0)} |\nabla G_T(\boldsymbol{a}, \mathrm{b}, 0)|^p\right)^{\frac{1}{p}} \lesssim C^{\frac{\beta}{2}} R^{1-d}, \qquad \text{where } C := C(\boldsymbol{a}, Q_R(x_0), \tfrac{p}{2-p}),$$

for each $x_0 \in X_R$ separately. We use the shorthand $G_T(x) := G_T(\boldsymbol{a}, x, 0)$ and set $\bar{G}_T := \frac{1}{|B_R(x_0)|}\sum_{x\in B_R(x_0)} G_T(x)$. In view of (3.8), $u(x) := G_T(x) - \bar{G}_T$ satisfies (4.26) with $f = \delta - \frac{1}{T}G_T$. Since

$$\sum_{\mathbb{Z}^d} |\delta - \frac{1}{T}G_T| \leq 1 + \frac{1}{T}\sum_{\mathbb{Z}^d} G_T(x) = 2, \qquad (4.42)$$

Lemma 4.4 yields

$$R^{-d} \sum_{B_R(x_0)} |u| \lesssim C^{\frac{1}{2}p^*} R^{2-d}. \qquad (4.43)$$

Thanks to the last inclusion in (4.41) we have $0 \notin B_R(x_0)$, and thus u satisfies (4.32) with $f = \frac{1}{T} G_T$ (with $B_{2R}(x_0)$ replaced by $B_R(x_0)$). Hence, Lemma 4.5 yields

$$\begin{aligned} \left(R^{p-d} \sum_{Q_{\frac{R}{2}}(x_0)} |\nabla G_T|^p \right)^{\frac{1}{p}} &= \left(R^{p-d} \sum_{Q_{\frac{R}{2}}(x_0)} |\nabla u|^p \right)^{\frac{1}{p}} \\ &\lesssim C^{\frac{1}{2}\alpha} R^{-d} \sum_{B_R(x_0)} |u| + C^{\frac{1}{2}\alpha} \left(R^{2-d} \sum_{B_R(x_0)} \frac{1}{T} G_T u_- \right)^{\frac{1}{2}} \\ &\overset{(4.43)}{\lesssim} C^{\frac{1}{2}(\alpha+p^*)} R^{2-d} + C^{\frac{1}{2}\alpha} \left(R^{2-d} \sum_{B_R(x_0)} \frac{1}{T} G_T u_- \right)^{\frac{1}{2}}. \end{aligned} \tag{4.44}$$

Regarding the second term on the right-hand side we only need to show

$$\frac{1}{T} \sum_{B_R(x_0)} G_T u_- \lesssim C^{p^*} R^{2-d}. \tag{4.45}$$

We note that $(G_T - \bar{G}_T)(G_T - \bar{G}_T)_- \leq 0$, so that

$$\frac{1}{T} \sum_{B_R(x_0)} G_T u_- = \frac{1}{T} \sum_{B_R(x_0)} (G_T - \bar{G}_T + \bar{G}_T)(G_T - \bar{G}_T)_- \leq \frac{1}{T} \bar{G}_T \sum_{B_R(x_0)} |G_T - \bar{G}_T|.$$

Combined with (4.43) and the inequality $\frac{1}{T} \bar{G}_T \lesssim R^{-d} \frac{1}{T} \sum_{B_R(x_0)} G_T \leq R^{-d}$, (4.45) follows.

Step 2. Argument for $k = 0$. Fix $\boldsymbol{a} \in \Omega$. For brevity set $G_T(x) := G_T(\boldsymbol{a}, x, 0)$ and $\bar{G}_T := \frac{1}{|B_{2R_0}(0)|} \sum_{x \in B_{2R_0}(0)} G_T(x)$. By the discrete ℓ^1-ℓ^p-estimate and the elementary inequality $|\nabla G_T(\mathrm{b})| \leq |G_T(x_\mathrm{b}) - \bar{G}_T| + |G_T(y_\mathrm{b}) - \bar{G}_T|$ we have

$$\left(\frac{1}{|Q_{R_0}(0)|} \sum_{\mathrm{b} \in Q_{R_0}(0)} |\nabla G_T(\mathrm{b})|^p \right)^{\frac{1}{p}} \lesssim \sum_{B_{2R_0}(0)} |G_T - \bar{G}_T|.$$

As in Step 1 an application of Lemma 4.4 yields

$$\sum_{B_{2R_0}(0)} |G_T - \bar{G}_T| \lesssim C^{\frac{p^*}{2}}(\boldsymbol{a}, Q_{2R_0}(0), \tfrac{p}{2-p}) R_0^2.$$

Since $R_0^2 \sim R_0^{1-d}$ and because the exponent of the constant satisfies $\frac{p^*}{2} \leq \frac{\beta}{2}$, the desired estimate follows. □

4.4 Proof of Lemma 3.12

In order to deal with the failure of the Leibniz rule we will appeal to a number of discrete estimates, which are stated in Lemma A.1 below. As already mentioned, we replace the missing uniform ellipticity of $\boldsymbol{a}$ by the coercivity estimate of Lemma 3.11 which makes use of the weight ω defined in (3.9). Morally speaking it plays the role of $\frac{1}{\lambda_0}$ in (3.18). In view of Assumption (A2) all moments of ω are bounded, i. e. $\langle \omega^k \rangle \lesssim 1$, where $\lesssim$ means $\leq$ up to a constant that only depends on k, p, Λ and d. We split the proof of Lemma 3.12 into the following two inequalities:

$$\left\langle|\nabla\phi(\mathrm{b})|^{2p+1}\right\rangle^{\frac{2p+2}{2p+1}} \lesssim \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{p+1})(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle, \tag{4.46}$$

$$\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{p+1})(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle \lesssim \left\langle\phi^{2p}\right\rangle. \tag{4.47}$$

Here and below we write ϕ instead of ϕ_T for simplicity. We start with (4.46). We smuggle in ω by appealing to Hölder's inequality with exponent $\frac{2p+2}{2p+1}$ and exploit that all moments of ω are bounded by Assumption (A2):

$$\left\langle|\nabla\phi(\mathrm{b})|^{2p+1}\right\rangle^{\frac{2p+2}{2p+1}} \lesssim \left\langle|\nabla\phi(\mathrm{b})|^{2p+2}\omega^{-1}(\mathrm{b})\right\rangle.$$

We combine (3.17) in the form of $|\nabla\phi(\mathrm{b})|^{2p+2} \lesssim (\frac{\phi^p(x_\mathrm{b})+\phi^p(y_\mathrm{b})}{2})^2|\nabla\phi(\mathrm{b})|^2$ (where we use that p is even) with the discrete version of the Leibniz rule $F^p\nabla F = \frac{1}{p+1}\nabla(F^{p+1})$, see (A.3) in Corollary A.2 below:

$$\left\langle|\nabla\phi(\mathrm{b})|^{2p+2}\omega^{-1}(\mathrm{b})\right\rangle \lesssim \left\langle|\nabla(\phi^{p+1})(\mathrm{b})|^2\omega^{-1}(\mathrm{b})\right\rangle. \tag{4.48}$$

Now (4.46) follows from the coercivity estimate of Lemma 3.11.

Next we prove (4.47). The discrete version of the Leibniz rule $|\nabla(F^{p+1})|^2 = \frac{(p+1)^2}{(2p+1)}\nabla F\nabla(F^{2p+1})$ (see Lemma A.1 (ii)) yields

$$\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{p+1})(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle \lesssim \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle\nabla\phi(\mathrm{b}')\boldsymbol{a}(\mathrm{b}')\nabla(\phi^{2p+1})(\mathrm{b}')\right\rangle.$$

By stationarity and the modified corrector equation (3.1) we have

$$\begin{aligned}
&\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle\nabla\phi(\mathrm{b}')\boldsymbol{a}(\mathrm{b}')\nabla(\phi^{2p+1})(\mathrm{b}')\right\rangle = \left\langle(\nabla^*\boldsymbol{a}\nabla\phi)\ \phi^{2p+1}\right\rangle \\
&= \quad -\frac{1}{T}\left\langle\phi^{2(p+1)}\right\rangle - \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle\nabla\phi^{2p+1}(\mathrm{b}')\boldsymbol{a}(\mathrm{b}')e(\mathrm{b}')\right\rangle \\
&\le \quad \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{2p+1})(\mathrm{b}')|\boldsymbol{a}(\mathrm{b}')\right\rangle,
\end{aligned}$$

where for the last inequality we use that $\left\langle\phi^{2(p+1)}\right\rangle \ge 0$ and $|e| = 1$. By Corollary A.2 and Young's inequality we get for any $\epsilon > 0$

$$\begin{aligned}
\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{2p+1})(\mathrm{b}')|\boldsymbol{a}(\mathrm{b}')\right\rangle \overset{\text{(A.2)}}{\lesssim}\ & \epsilon \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla\phi(\mathrm{b}')|^2\left(\tfrac{\phi^p(x_{\mathrm{b}'})+\phi^p(y_{\mathrm{b}'})}{2}\right)^2\boldsymbol{a}(\mathrm{b}')\right\rangle \\
&+\frac{1}{\epsilon}\sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle\left(\tfrac{\phi^p(x_{\mathrm{b}'})+\phi^p(y_{\mathrm{b}'})}{2}\right)^2\right\rangle \\
\overset{\text{(A.3)}}{\lesssim}\ & \epsilon \sum_{\substack{\mathrm{b}'=\{0,e_i\}\\ i=1,\ldots,d}} \left\langle|\nabla(\phi^{p+1})(\mathrm{b}')|^2\boldsymbol{a}(\mathrm{b}')\right\rangle + \frac{1}{\epsilon}\left\langle\phi^{2p}\right\rangle.
\end{aligned}$$

Since we may choose $\epsilon > 0$ as small as we wish, the first term on the right-hand side can be absorbed into the left-hand side of (4.47) and the claim follows.

A Appendix: Replacements of the Leibniz rule for the discrete derivative

Lemma A.1. *Let F be a scalar function on $\mathbb{Z}^d$ and $\mathrm{b} \in \mathbb{B}^d$.*

[(i)]

1. *Assume that $p \in 2\mathbb{N}$. Then we have*

$$|\nabla(F^{p+1})(\mathrm{b})| \sim |\nabla F(\mathrm{b})| \frac{F^p(x_\mathrm{b}) + F^p(y_\mathrm{b})}{2}.$$

2. *For every integer p we have*

$$|\nabla(F^{p+1})(\mathrm{b})|^2 \lesssim \nabla F(\mathrm{b}) \nabla(F^{2p+1})(\mathrm{b}).$$

Here $\lesssim$ (resp. $\sim$) means up to a constant that only depends on p.

Proof of Lemma A.1. Let $x, y \in \mathbb{Z}^d$ denote the vertices with $\mathrm{b} = \{x, y\}$ and $y - x \in \{e_1, \ldots, e_d\}$ so that $\nabla F(\mathrm{b}) = F(y) - F(x)$.

Proof of part (i). The statement " $\lesssim$ " is equivalent to [18, Equation (5.29)] and is proven there. Concerning $\gtrsim$ we appeal to [18, Equation (5.28)]. From that equation we learn that

$$\nabla(F^{p+1})(\mathrm{b}) \nabla F(b) \gtrsim \frac{F^p(x_\mathrm{b}) + F^p(y_\mathrm{b})}{2} |\nabla F(\mathrm{b})|^2.$$

By dividing by $|\nabla F(\mathrm{b})|$ one immediately finds the claimed result.

Proof of part (ii). We have to distinguish two cases.

First case: $F(x), F(y) \geq 0$ or $F(x), F(y) \leq 0$. It suffices to show the statement for $F(x), F(y) \geq 0$, since then the case $F(x), F(y) \leq 0$ follows by symmetry. We have to prove that

$$(F^{p+1}(y) - F^{p+1}(x))^2 \lesssim (F(y) - F(x))(F^{2p+1}(y) - F^{2p+1}(x)).$$

By symmetry and and scale invariance, it suffices to show the elementary inequality

$$\forall f \geq 0 : \quad (1 - f^{p+1})^2 \leq c(1 - f)(1 - f^{2p+1}), \tag{A.1}$$

where $c > 0$ only depends on p. We omit its proof for the sake of brevity.

Second case: $F(x) \leq 0, F(y) \geq 0$ or $F(x) \geq 0, F(y) \leq 0$. It suffices to show the statement for $F(x) \leq 0, F(y) \geq 0$, since then the case $F(x) \geq 0, F(y) \leq 0$ follows by symmetry. We have to prove that

$$(F^{p+1}(y) - F^{p+1}(x))^2 \lesssim (F(y) - F(x))(F^{2p+1}(y) - F^{2p+1}(x))$$

or equivalently

$$\begin{aligned} &F^{2(p+1)}(y) + F^{2(p+1)}(x) - 2F^{p+1}(y)F^{p+1}(x) \\ &\qquad \lesssim F^{2p+2}(y) + F^{2p+2}(x) - F(x)F^{2p+1}(y) - F(y)F^{2p+1}(x). \end{aligned}$$

Note that since $2p + 1$ is an odd integer, the last two terms on the right hand side of the above inequality are positive. Hence, it suffices to prove that

$$F^{2(p+1)}(y) + F^{2(p+1)}(x) - 2F^{p+1}(y)F^{p+1}(x) \lesssim F^{2p+2}(y) + F^{2p+2}(x),$$

which follows due to $-2F^{p+1}(y)F^{p+1}(x) \leq F^{2p+2}(y) + F^{2p+2}(x)$. □

In the course of proving our main result we will use the discrete Leibniz rule, (i) in the above lemma, in the following form.

Corollary A.2. *For every scalar function F, every bond b and every even integer p we have*

$$|\nabla(F^{2p+1})(\mathrm{b})| \lesssim |\nabla F(\mathrm{b})| \left(\frac{F^p(x_\mathrm{b}) + F^p(y_\mathrm{b})}{2}\right)^2, \tag{A.2}$$

$$|\nabla F(\mathrm{b})|^2 \left(\frac{F^p(x_\mathrm{b}) + F^p(y_\mathrm{b})}{2}\right)^2 \lesssim |\nabla(F^{p+1})(\mathrm{b})|^2. \tag{A.3}$$

Here $\lesssim$ means up to a constant that only depends on p.

References

[1] S. Andres, M. T. Barlow, J.-D. Deuschel, and B. M. Hambly, *Invariance principle for the random conductance model*, Probab. Theory Related Fields **156** (2013), no. 3-4, 535–580. MR-3078279

[2] Sebastian Andres, Jean-Dominique Deuschel, and Martin Slowik, *Invariance principle for the random conductance model in a degenerate ergodic environment*, Ann. Probab. **43** (2015), no. 4, 1866–1891. MR-3353817

[3] Scott N. Armstrong and Charles K. Smart, *Quantitative stochastic homogenization of elliptic equations in nondivergence form*, Arch. Ration. Mech. Anal. **214** (2014), no. 3, 867–911. MR-3269637

[4] Martin T. Barlow, *Random walks on supercritical percolation clusters*, Ann. Probab. **32** (2004), no. 4, 3024–3084. MR-2094438

[5] Noam Berger and Marek Biskup, *Quenched invariance principle for simple random walk on percolation clusters*, Probab. Theory Related Fields **137** (2007), no. 1-2, 83–120. MR-2278453

[6] M. Biskup, M. Salvi, and T. Wolff, *A central limit theorem for the effective conductance: linear boundary data and small ellipticity contrasts*, Comm. Math. Phys. **328** (2014), no. 2, 701–731. MR-3199997

[7] Marek Biskup, *Recent progress on the random conductance model*, Probab. Surv. **8** (2011), 294–373. MR-2861133

[8] Marek Biskup and Timothy M. Prescott, *Functional CLT for random walk among bounded random conductances*, Electron. J. Probab. **12** (2007), no. 49, 1323–1348. MR-2354160

[9] Daniel Boivin, *Weak convergence for reversible random walks in a random environment*, Ann. Probab. **21** (1993), no. 3, 1427–1440. MR-1235423

[10] Alain Bourgeat and Andrey Piatnitski, *Approximations of effective coefficients in stochastic homogenization*, Ann. Inst. H. Poincaré Probab. Statist. **40** (2004), no. 2, 153–165. MR-2044813

[11] Luis A. Caffarelli and Panagiotis E. Souganidis, *Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media*, Invent. Math. **180** (2010), no. 2, 301–360. MR-2609244

[12] Pietro Caputo and Dmitry Ioffe, *Finite volume approximation of the effective diffusion matrix: the case of independent bond disorder*, Ann. Inst. H. Poincaré Probab. Statist. **39** (2003), no. 3, 505–525. MR-1978989

[13] Joseph G. Conlon and Thomas Spencer, *A strong central limit theorem for a class of random surfaces*, Comm. Math. Phys. **325** (2014), no. 1, 1–15. MR-3182484

[14] ______, *Strong convergence to the homogenized limit of elliptic equations with random coefficients*, Trans. Amer. Math. Soc. **366** (2014), no. 3, 1257–1288. MR-3145731

[15] A. De Masi, P. A. Ferrari, S. Goldstein, and W. D. Wick, *An invariance principle for reversible Markov processes. Applications to random motions in random environments*, J. Statist. Phys. **55** (1989), no. 3-4, 787–855. MR-1003538

[16] A. Gloria, S. Neukamm, and F. Otto, *Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics*, Invent. Math. **199** (2015), no. 2, 455–515. MR-3302119

[17] Antoine Gloria, Stefan Neukamm, and Felix Otto, *An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations*, ESAIM Math. Model. Numer. Anal. **48** (2014), no. 2, 325–346. MR-3177848

[18] Antoine Gloria and Felix Otto, *An optimal variance estimate in stochastic homogenization of discrete elliptic equations*, Ann. Probab. **39** (2011), no. 3, 779–856. MR-2789576

[19] ______, *An optimal error estimate in stochastic homogenization of discrete elliptic equations*, Ann. Appl. Probab. **22** (2012), no. 1, 1–28. MR-2932541

[20] V. V. Jurinskiĭ, *On a Dirichlet problem with random coefficients*, Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978), Lecture Notes in Control and Information Sci., vol. 25, Springer, Berlin-New York, 1980, pp. 344–353. MR-609200

[21] C. Kipnis and S. R. S. Varadhan, *Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions*, Comm. Math. Phys. **104** (1986), no. 1, 1–19. MR-834478

[22] S.M. Kozlov, *The averaging of random operators*, Mat. Sb. (N.S.) **109(151)** (1979), no. 2, 188–202.

[23] ______, *Averaging of difference schemes*, Math. USSR Sbornik **57(2)** (1987), no. 2, 351–369.

[24] Takashi Kumagai, *Random walks on disordered media and their scaling limits*, Lecture Notes in Mathematics, vol. 2101, Springer, Cham, 2014, Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR-3156983

[25] Rolf Künnemann, *The diffusion limit for reversible jump processes on* $\mathbf{Z}^d$ *with ergodic random bond conductivities*, Comm. Math. Phys. **90** (1983), no. 1, 27–68. MR-714611

[26] P. Mathieu and A. Piatnitski, *Quenched invariance principles for random walks on percolation clusters*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **463** (2007), no. 2085, 2287–2307. MR-2345229

[27] Jean-Christophe Mourrat and Felix Otto, *Correlation structure of the corrector in stochastic homogenization*, arXiv:1402.1924.

[28] Ali Naddaf and Thomas Spencer, *On homogenization and scaling limit of some gradient perturbations of a massless free field*, Comm. Math. Phys. **183** (1997), no. 1, 55–84. MR-1461951

[29] G. C. Papanicolaou and S. R. S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, Random fields, Vol. I, II (Esztergom, 1979), Colloq. Math. Soc. János Bolyai, vol. 27, North-Holland, Amsterdam-New York, 1981, pp. 835–873. MR-712714

[30] Raphael Rossignol, *Noise-stability and central limit theorems for effective resistance of random electric networks*, To appear in Ann. Probab.

[31] Vladas Sidoravicius and Alain-Sol Sznitman, *Quenched invariance principles for walks on clusters of percolation or among random conductances*, Probab. Theory Related Fields **129** (2004), no. 2, 219–244. MR-2063376

Acknowledgments. We thank Artem Sapozhnikov for stimulating discussions on percolation models. SN was partially supported by ERC-2010-AdG no.267802 AnaMultiScale. SN furthermore appreciates funding by the DFG in the context of the Dresden Institutional Strategy "The Synergetic University". Most of this work was done while all three authors were employed at the Max-Planck-Institute for Mathematics in the Sciences, Leipzig.