Sublinear growth of the corrector in stochastic homogenization: Optimal stochastic estimates for slowly decaying correlations

Julian Fischer and Felix Otto

September 5, 2016

Abstract

We establish sublinear growth of correctors in the context of stochastic homogenization of linear elliptic PDEs. In case of weak decorrelation and "essentially Gaussian" coefficient fields, we obtain optimal (stretched exponential) stochastic moments for the minimal radius above which the corrector is sublinear. Our estimates also capture the quantitative sublinearity of the corrector (caused by the quantitative decorrelation on larger scales) correctly. The result is based on estimates on the Malliavin derivative for certain functionals which are basically averages of the gradient of the corrector, on concentration of measure, and on a mean value property for a-harmonic functions.

1   Introduction

In the present work, we are concerned with the stochastic homogenization of linear elliptic equations of the form

$$-\nabla \cdot a\nabla u = f. \tag{1}$$

In stochastic homogenization of elliptic PDEs, a is typically a uniformly elliptic and bounded coefficient field, chosen at random according to some stationary and ergodic ensemble $\langle \cdot \rangle$. On large scales (and for slowly varying f), one may then approximate the solution u to the equation (1) by the solution u_{hom} to the so-called *effective equation*

$$-\nabla \cdot a_{hom}\nabla u_{hom} = f, \tag{2}$$

1

which is a constant-coefficient equation with the so-called effective coefficient a_{hom}. Mathematically, this homogenization effect is encoded in growth properties of the *corrector* (cf. below for a definition of the corrector).

The goal of the present paper is to provide a fairly simple proof of *quantified* sublinear growth of the corrector under very mild assumptions on the decorrelation of the coefficient field a under the ensemble $\langle \cdot \rangle$. We do this in the context of coefficient fields that are essentially Gaussian. More precisely, we consider coefficient fields a which are obtained from a Gaussian random field by pointwise application of some (nonlinear) mapping, the role of the nonlinear map being basically to enforce uniform ellipticity and boundedness of our coefficient field.

The motivation for this result is the following: Gloria, Neukamm, and the second author [11] have shown that *qualitatively* sublinear growth of the (extended) corrector (ϕ, σ) (cf. below for a definition) entails a large-scale intrinsic $C^{1,\alpha}$ regularity theory for a-harmonic functions. In a subsequent work [10], the two authors of the present paper have shown that *slightly quantified* sublinear growth of the corrector even leads to a large-scale intrinsic $C^{k,\alpha}$ regularity theory for any $k \in \mathbb{N}$. Therefore, the results of the present work show that even in case of ensembles with very mild decorrelation, for almost every realization of the coefficient field, a-harmonic functions have arbitrary intrinsic smoothness properties on large scales. Furthermore, our results enable us to estimate the scale above which this happens – a random quantity – in a stochastically optimal way. Indeed, the motivation for the present work was to establish such a (necessarily intrinsic) higher order regularity theory under the weakest possible assumptions on the decay of correlations.

By an *intrinsic* regularity theory we mean that the regularity is measured in terms of objects intrinsic to the Riemannian geometry defined by the coefficient field a, like the dimension of the space of a-harmonic functions of a certain algebraic growth rate, or like estimates on the Hölder modulus of the derivative of a-harmonic functions as measured in terms of their distance to a-linear functions. An *extrinsic* large-scale regularity theory for a-harmonic functions in case of random coefficients was initiated on the level of a $C^{0,\alpha}$ in [7, 21] and pushed to $C^{1,0}$ in [4], which significantly extended qualitative arguments from the periodic case [5] to quantitative arguments in the random case. However, an extrinsic regularity theory is limited to $C^{1,0}$, as can be seen considering the harmonic coordinates: Taking higher order polynomials into account does not increase the local approximation order.

After this motivation, we now return to the discussion of the history on bounds on the corrector, as they depend on assumptions on the stationary

ensemble of coefficient fields. Almost-sure sublinearity (always meant in a spatially averaged sense) of the corrector ϕ under the mere assumption of ergodicity was a key ingredient in the original work on stochastic homogenization by Kozlov [19] and by Papanicolaou & Varadhan [24]. Almost-sure sublinearity of the *extended* corrector (ϕ, σ), as is needed for the large-scale intrinsic $C^{1,\alpha}$-regularity theory, was established in [11] under mere ergodicity.

Yurinskii [25] was the first to quantify sublinear growth under general mixing conditions, however only capturing suboptimal rates even in case of finite range of dependence. Very recently, a much improved quantification of sublinear growth of ϕ under finite range assumptions was put forward by Armstrong, Kuusi, and Mourrat [1], relying on a variational approach to quantitative stochastic homogenization introduced by Armstrong and Smart [4], an approach which presumably can be extended to the case of non-symmetric coefficients and more general mixing conditions following [3]. Recently, optimal growth bounds on the corrector ϕ with optimal – i. e. Gaussian – stochastic integrability have been established under the assumption of finite range of dependence [2, 17].

Optimal growth rates have been obtained under a quantification of ergodicity different from finite range or mixing conditions, namely under Spectral Gap assumptions on the ensemble. This functional analytic tool from statistical mechanics was introduced into the field of stochastic homogenization in an unpublished paper by Naddaf and Spencer [23], and further leveraged by Conlon et. al. [8, 9], yielding optimal rates for some errors in stochastic homogenization in case of a small ellipticity contrast. The work of Gloria, Neukamm and the second author extended these results to the present case of arbitrary ellipticity contrast [12, 13, 14], in particular yielding at most logarithmic growth of the corrector (and its stationarity in $d > 2$). Loosely speaking, the assumption of a Spectral Gap Inequality amounts to correlations with integrable tails; in the above-mentioned works it has been used for discrete media (i. e. random conductance models), but has subsequently been extended to the continuum case [15, 16].

A strengthening of the Spectral Gap Inequality is given by the Logarithmic Sobolev Inequality (LSI); it is a slight strengthening in terms of the assumption (still essentially encoding integrable tails of the correlations), but a substantial improvement in its effect, since it implies Gaussian concentration of measure for Lipschitz random variables. The assumption of LSI and implicitly concentration of measure, which will be explicitly used in this work, has been introduced into stochastic homogenization by Marahrens and the second author [21]. In [11], it has been shown that the concept of LSI can be adapted to also capture ensembles with slowly decaying correlations, i. e.

thick non-integrable tails, by adapting the norm of the vertical or Malliavin derivative to the correlation structure. As a result, the stochastic integrability of the optimal rates could be improved from algebraic to (stretched) exponential, but missing the expected Gaussian integrability.

The main merit of the present contribution w. r. t. to [11] is twofold: First, our approach directly provides optimal quantitative sublinearity of the corrector (ϕ, σ) on all scales above a random minimal radius r_*, i.e. in contrast to the estimates of [11] our estimates capture the decorrelation on scales larger than r_* in a single argument. Note that our definition of r_* differs from the one in [11]. Second, in case of weak decorrelation, our simpler arguments are nevertheless sufficient to establish optimal stochastic moments for the minimal radius r_* above which the corrector (ϕ, σ) displays the quantified sublinear growth.

In the present work, we consider the following type of ensembles on λ-uniformly elliptic tensor fields $a = a(x)$ on $\mathbb{R}^d$: Let $\tilde{a} = \tilde{a}(x)$ be a tensor-valued Gaussian random field on $\mathbb{R}^d$ that is centered (i. e. of vanishing expectation) and stationary (i. e. invariant under translation) and thus characterized by the covariance $\langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle$. Our only additional assumption on $\tilde{a}$ is that there exists an exponent $\beta \in (0, d)$ such that

$$\left| \langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle \right| \leq |x|^{-\beta} \quad \text{for all } x \in \mathbb{R}^d. \tag{3}$$

In this work, we are concerned with the case of weak decay of correlation in the sense of $\beta \ll 1$. Let Φ be a 1-Lipschitz map from the space of tensors into the space of λ-elliptic symmetric tensors. Then our ensemble is the distribution of a where a is given by $a(x) := \Phi(\tilde{a}(x))$. Note that the normalization in the constant in (3) and in the Lipschitz constant is not essential, since it can be achieved by a rescaling of x and the amplitude of $\tilde{a}$.

Concerning the mathematical tools of our approach, several ideas are inspired by the work [11]. In particular, a key component of our approach are sensitivity estimates (Malliavin derivative bounds) for certain integral functionals, which basically average the gradient $\nabla(\phi, \sigma)$ over an appropriate cube. Furthermore, we rely on a mean-value property for a-harmonic functions, which has been derived in [11] under appropriate smallness assumptions on the corrector. In our present contribution, we however pursue a conceptually simpler route to estimate the Malliavin derivative: The sensitivity estimate is performed through appropriate L^q-norm bounds and Meyer's estimate, rather than a more involved $\ell^2 - L^1$-norm bound like in [11].

Before stating our main results, let us recall the concept of correctors in homogenization and introduce some notation. The basic idea underlying

the concept of correctors in homogenization is the observation that the oscillations in the gradient ∇u_{hom} of solutions to the homogenized (constant-coefficient) problem (2) occur on a much larger scale than the oscillations in the gradient ∇u of solutions to the original problem (1). Thus, it is important to understand how to add oscillations to an affine map (an affine map being always a_{hom}-harmonic) to obtain an a-harmonic map. In the context of stochastic homogenization, one is therefore interested in constructing random scalar fields $\phi_i = \phi_i(a, x)$ subject to the equation

$$-\nabla \cdot a(e_i + \nabla \phi_i) = 0 \tag{4}$$

which almost surely display sublinear growth in x: The ϕ_i then facilitate the transition from the a_{hom}-harmonic (Euclidean) coordinates $x \mapsto x_i$ to the "a-harmonic coordinates" $x \mapsto x_i + \phi_i(x)$. Since any affine map may be represented in the form $b + \sum_i \xi_i x_i$ for $b, \xi_i \in \mathbb{R}$, the ϕ_i also facilitate the construction of associated a-harmonic "corrected affine maps" $b + \sum_i \xi_i(x_i + \phi_i)$.

With the help of the corrector, one may characterize the effective coefficient a_{hom}: In our setting of stochastic homogenization, the effective coefficient is given by the formula

$$a_{hom} e_i = \big\langle a(e_i + \nabla \phi_i) \big\rangle, \tag{5}$$

where $\langle \cdot \rangle$ refers to the expectation with respect to our ensemble (i.e. probability measure).

In the language of a conducting medium with conductivity tensor a – note that in this picture, one has $f \equiv 0$ in (1) – , the quantity $E_i := e_i + \nabla \phi_i$ corresponds to the (curl-free) "microscopic" electric field associated with a "macroscopic" electric field e_i (and, therefore, ϕ_i corresponds to the "microscopic" correction to the "macroscopic" electric potential x_i). The corresponding (divergence-free) "microscopic" current density is given by

$$q_i := a(e_i + \nabla \phi_i), \tag{6}$$

while the "macroscopic" current density associated with the "macroscopic" electric field e_i is given by the "average" of this quantity, i.e. by the expression (5).

In periodic homogenization of linear elliptic PDEs, it turns out to be convenient to introduce a dual quantity to the corrector ϕ_i (cf. e.g. [18, p.27]): One constructs a tensor field σ_{ijk}, skew-symmetric in the last two indices, which is a potential for the flux correction $q_i - a_{hom} e_i$ in the sense

$$\nabla \cdot \sigma_i = a(e_i + \nabla \phi_i) - a_{hom} e_i, \tag{7}$$

where we have set $(\nabla \cdot \sigma_i)_j := \sum_{k=1}^d \partial_k \sigma_{ijk}$. With the help of this "extended corrector" (ϕ, σ), it is possible to give a bound on the homogenization error (in terms of appropriate norms of ϕ and σ).

One of the main merits of [11] is the discovery of the usefulness of this extended corrector (ϕ, σ) in the context of stochastic homogenization. For stationary and ergodic ensembles $\langle \cdot \rangle$ of λ-uniformly elliptic and symmetric coefficient fields $a = a(x)$ on $\mathbb{R}^d$, in [11] correctors ϕ_i and σ_{ijk} such that

$$\nabla\phi_i, \nabla\sigma_{ijk} \quad \begin{array}{l} \text{are stationary,} \\ \text{of bounded second moment,} \\ \text{and of vanishing expectation,} \end{array} \tag{8}$$

have been constructed. As a consequence of this and of ergodicity, the ϕ_i and σ_{ijk} almost surely display sublinear growth. Note that in case of σ_i, the choice of the appropriate gauge is important for the property (8) and for our work, as the equation (7) determines σ_i (which by its skew-symmetry and its behavior under change of coordinates may be identified with a $d - 1$-form) only up to the exterior derivative of a $d - 2$-form. In fact, the choice of the gauge in [11] is such that

$$-\triangle\sigma_{ijk} = \partial_j q_{ik} - \partial_k q_{ij}, \tag{9}$$

which in view of (4) and (6) is clearly compatible with (7).

Notation. To quantify the ellipticity and boundedness of our coefficient fields, throughout the paper we shall work with the assumptions

$$av \cdot v \geq \lambda|v|^2 \quad \text{for all } v \in \mathbb{R}^d, \tag{10}$$

$$|av| \leq |v| \quad \text{for all } v \in \mathbb{R}^d, \tag{11}$$

where $\lambda \in (0, 1)$. Note that in view of rescaling, the upper bound (11) on a does not induce a loss of generality of our results.

For our convenience, throughout the paper we shall assume that our coefficient field a is symmetric. The arguments however easily carry over to the case of non-symmetric coefficient fields by simultaneously considering the correctors for the dual equation (i.e. the PDE with coefficient field a^*, a^* denoting the transpose of a).

The expression $s \lesssim t$ is an abbreviation for $s \leq Ct$ with C a generic constant only depending on the dimension d, the exponent $\beta > 0$, and the ellipticity ratio $\lambda > 0$.

The expression $s \ll t$ stands for $s \leq \frac{1}{C}t$ with C a generic sufficiently large constant only depending on the dimension d, the exponent $\beta > 0$, and the ellipticity ratio $\lambda > 0$.

By $I(E)$ we denote the characteristic function of an event E.

The notation $\fint_A f$ refers to the average integral over the set A, i.e. we have $\fint_A f\, dx = \int_A f\, dx / \int_A 1\, dx$.

In the sequel, (ϕ, σ) stands for any component ϕ_i, σ_{ijk} for $i, j, k = 1, \cdots, d$.

2 Main Results and Structure of Proof

Let us now state our main theorem. To quantify the sublinear growth of the extended corrector (ϕ, σ), we first quantify the decay of spatial averages of $\nabla(\phi, \sigma)$ over larger scales. In view of the decorrelation assumption (3) for our ensemble of coefficient fields, we expect that, up to logarithms, it is the exponent $\frac{\beta}{2}$ that governs the decay of averages of $\nabla(\phi, \sigma)$ and the improvement over linear growth for (ϕ, σ). Indeed, this exponent is reflected in the theorem.

Theorem 1. *Let $\tilde{a} = \tilde{a}(x)$ be a tensor-valued Gaussian random field on $\mathbb{R}^d$ that is centered (i. e. of vanishing expectation) and stationary (i. e. invariant under translation); assume that the covariance of $\tilde{a}$ satisfies the estimate*

$$\left| \langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle \right| \leq |x|^{-\beta} \quad \text{for all } x \in \mathbb{R}^d \qquad (12)$$

for some $\beta \in (0, d)$. Let $\Phi : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ be a Lipschitz map with Lipschitz constant ≤ 1; suppose that Φ takes values in the set of symmetric matrices subject to the ellipticity and boundedness assumptions (10), (11). Define the ensemble $\langle \cdot \rangle$ as the probability distribution of a, where a is the image of $\tilde{a}$ under pointwise application of the map Φ, i.e. $a(x) := \Phi(\tilde{a}(x))$.

Assume in addition on the ensemble $\langle \cdot \rangle$ that β in (12) is sufficiently small in the sense of

$$\beta \leq \frac{1}{C}, \qquad (13)$$

where C denotes a generic constant only depending on d and λ.

   i) Consider a linear functional $F = Fh$ on vector fields $h = h(x)$ satisfying the boundedness property

$$|Fh| \leq \Big(\fint_{|x| \leq r} |h|^{\frac{2d}{d+\beta}} dx \Big)^{\frac{d+\beta}{2d}} \qquad (14)$$

   for some radius $r > 0$. Then the random variable $F\nabla(\phi, \sigma)$ satisfies uniform Gaussian bounds in the sense of

$$\langle I(|F\nabla(\phi, \sigma)| \geq M) \rangle \leq C \exp(-\frac{1}{C} r^\beta M^2) \quad \text{for all } M \leq 1. \qquad (15)$$

ii) There exists a (random) radius r_* for which the "iterated logarithmic" bound

$$\frac{1}{r^2}\fint_{|x|\leq r}|(\phi,\sigma)-\fint_{|x|\leq r}(\phi,\sigma)|^2 dx \leq (\frac{r_*}{r})^\beta \log(e+\log(\frac{r}{r_*})) \quad for\ r\geq r_* \tag{16}$$

holds and which satisfies the stretched exponential bound

$$\langle\exp(\frac{1}{C}r_*^\beta)\rangle \leq C. \tag{17}$$

Morally speaking, Theorem 1 converts statistical information on the coefficient field a (or rather $\tilde{a}$) into statistical information on the coefficient field $\nabla\phi := \nabla(\phi_1,\cdots,\phi_d)$ related by (4). Despite the nonlinearity of the map $a \mapsto \nabla\phi$, which only in its linearization around $a = $ id turns into the Helmholtz projection, Theorem 1 states that $\nabla\phi$ essentially inherits the statistics of a: (15) implies in particular that spatial averages $F = \fint_{|x|\leq r}\nabla\phi dx$ of $\nabla\phi$ satisfy the same bounds as if $\nabla\phi$ itself was Gaussian with correlation decay (3). On the level of these Gaussian bounds, the only price to pay for the nonlinearity is the restriction $M \lesssim 1$ in (15) on the threshold.

Incidentally, the way we obtain ii) from i) bears similarities with an argument in [1] in the sense that a decomposition into Haar wavelets is implicitly used.

To see the optimality of the corrector bound (16) and the stochastic integrability (17), consider a perturbative setting: Let $\tilde{a}$ be a scalar, centered, stationary Gaussian field with covariance given by

$$\langle\tilde{a}(x)\tilde{a}(0)\rangle = |x|^{-\beta} \tag{18}$$

and consider the infinitesimally perturbed Laplace operator

$$a := \left(1 + \frac{1}{\pi}\tan^{-1}(\delta\pi\tilde{a})\right)\mathrm{Id} = (1 + \delta\,\tilde{a})\,\mathrm{Id}, \tag{19}$$

where $\delta > 0$ is infinitesimal. In this setting, by (4) the gradient of the corrector ϕ is given by

$$\nabla\phi_i(x) = -\delta(\nabla\partial_i\Phi * \tilde{a})(x),$$

where Φ denotes the fundamental solution of the (negative) Laplacian (note that this identity is just formal, as we silently pass over integrability issues, but it may be given a rigorous meaning). In particular, $\nabla\phi_i$ is a stationary

centered Gaussian random field satisfying

$$
\begin{aligned}
&Cov(\partial_j \phi_i(x), \partial_k \phi_i(0)) \\
&= \delta^2 \big\langle (\partial_j \partial_i \Phi * \tilde{a})(x)(\partial_k \partial_i \Phi * \tilde{a})(0) \big\rangle \\
&= \delta^2 \int \int \partial_j \partial_i \Phi(x - z) \partial_k \partial_i \Phi(0 - w) Cov(\tilde{a}(z), \tilde{a}(w))\, dz\, dw \\
&= \delta^2 \int \int \partial_j \partial_i \Phi(x - z) \partial_k \partial_i \Phi(0 - w)|z - w|^{-\beta}\, dz\, dw \\
&= C(d, \beta, i, j, k)\delta^2 |x|^{-\beta}
\end{aligned}
\tag{20}
$$

with an explicitly computable (and for general i, j, k nonzero) constant $C(d, \beta, i, j, k)$. The difference $\fint_{|x| \leq r} \phi - \fint_{|x| \leq r/2} \phi$ is therefore a centered Gaussian random variable with variance $\sim r^{2-\beta}$, which entails that the moment bound (17) for the factor $r_*^\beta \log(e + \log(r/r_*))$ in the estimate (16) is (almost) optimal. The scaling in r of the bound (16) is optimal in view of the law of the iterated logarithm; details are provided in the appendix.

To obtain an estimate like (15), the starting point of our proof is the Gaussian concentration of measure applied to $\tilde{a}$. Recall the notion of the covariance operator Cov, which in our setting of a stationary centered Gaussian random field $\tilde{a}$ is given as the convolution with the tensor field $\langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle$.

Proposition 1 (Concentration of Measure, cf. e.g. [20, Proposition 2.18])).
Let $\tilde{a} = \tilde{a}(x)$ be a tensor-valued Gaussian random field on $\mathbb{R}^d$ that is centered and stationary; denote its covariance operator by Cov. *Consider a random variable F, that is, a function(al) $F = F(\tilde{a})$. Suppose that F is 1-Lipschitz in the sense that its functional derivative, or rather its Fréchet derivative with respect to $L^2(\mathbb{R}^d; \mathbb{R}^{d \times d})$, $\frac{\partial F}{\partial \tilde{a}} = \frac{\partial F}{\partial \tilde{a}}(\tilde{a}, x)$, which can be considered a random tensor field and assimilated with a Malliavin derivative, satisfies*

$$
\int_{\mathbb{R}^d} \frac{\partial F}{\partial \tilde{a}}(\tilde{a}, x)(\text{Cov} \frac{\partial F}{\partial \tilde{a}}(\tilde{a}, \cdot))(x)dx \leq 1 \quad \text{for almost every } \tilde{a}.
\tag{21}
$$

Then F has Gaussian moments in the sense of

$$
\langle \exp(M(F - \langle F \rangle)) \rangle \leq \exp(\frac{M^2}{2}) \quad \text{for all } M \geq 0.
\tag{22}
$$

Furthermore, for any $M \geq 0$ we have the estimate

$$
\langle I(|F - \langle F \rangle| \geq M) \rangle \leq 2 \exp(-\frac{M^2}{2}).
\tag{23}
$$

9

We now substitute our assumption (21) on the Fréchet derivative by a stronger but more tractable condition.

Lemma 1. *Let $\tilde{a} = \tilde{a}(x)$ be a tensor-valued Gaussian random field on $\mathbb{R}^d$ that is centered and stationary; denote its covariance operator as* Cov *and suppose that for some $\beta \in (0, d)$ we have the bound*

$$\left| \langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle \right| \le |x|^{-\beta} \quad \text{for all } x \in \mathbb{R}^d.$$

Let $\Phi : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ be a 1-Lipschitz map; denote the probability distribution of $\Phi(\tilde{a})$ as $\langle \cdot \rangle$. Consider a functional F on the space of tensor fields $\tilde{a}$ of the form $F = F(a)$ with $a(x) := \Phi(\tilde{a}(x))$; we shall use the abbreviation $F(\tilde{a})$ for $F(\Phi(\tilde{a}))$. Let $q \in (1, 2)$ be given by

$$\frac{1}{q} = 1 - \frac{\beta}{2d} \tag{24}$$

and suppose that the Fréchet derivative of F with respect to $L^2(\mathbb{R}^d; \mathbb{R}^{d \times d})$ satisfies

$$\left(\int |\frac{\partial F}{\partial a}|^q dx \right)^{\frac{2}{q}} \ll 1 \quad \text{for } \langle \cdot \rangle\text{-almost every } a. \tag{25}$$

Then the estimate (21) is satisfied, i.e. we have

$$\int_{\mathbb{R}^d} \frac{\partial F}{\partial \tilde{a}}(\tilde{a}, x)(\mathrm{Cov}\, \frac{\partial F}{\partial \tilde{a}}(\tilde{a}, \cdot))(x) dx \le 1 \quad \text{for almost every } \tilde{a}.$$

We observe that if q and β are related by (24), as $\beta \uparrow d$ we have $q \uparrow 2$ and for $\beta \downarrow 0$ we have $q \downarrow 1$.

For linear functionals of (the gradient of) the corrector (which are therefore nonlinear functionals of the coefficient field a), we now establish an explicit representation of the Fréchet derivative; this will aid us in verifying the Lipschitz condition (25) and thus ultimately the concentration of measure statements (22) and (23) for (an appropriate modification of) such functionals.

Lemma 2. *Consider a linear functional on $L^{\frac{p}{p-1}}(\mathbb{R}^d; \mathbb{R}^d)$ of the form*

$$Fh := \int_{\mathbb{R}^d} g \cdot h \, dx, \tag{26}$$

where $g \in L^p(\mathbb{R}^d; \mathbb{R}^d)$, $p \ge 2$, and $\mathrm{supp}\, g \subset \{|x| \le r\}$ for some $r \ge 1$. Let a be some coefficient field subject to the ellipticity and boundedness conditions (10), (11). Then the following two assertions hold:

1) Consider the Fréchet derivative $\frac{\partial F}{\partial a}$ of the functional $F := F\nabla\sigma_{ijk}$ (note that this functional is nonlinear in a, although it is linear in σ_{ijk}) at a (for some fixed i, j, k). Introduce the decaying solutions $v, \tilde{v}_{jk}$ to the equations

$$-\triangle v = \nabla \cdot g \qquad (27)$$

and (where a^* denotes the transpose of a)

$$-\nabla \cdot a^*\big(\nabla\tilde{v}_{jk} + (\partial_j ve_k - \partial_k ve_j)\big) = 0. \qquad (28)$$

We then have the representation

$$\frac{\partial F}{\partial a}(a) = \big(\partial_j ve_k - \partial_k ve_j + \nabla\tilde{v}_{jk}\big) \otimes (\nabla\phi_i + e_i). \qquad (29)$$

2) Consider the Fréchet derivative $\frac{\partial F}{\partial a}$ of the functional $F := F\nabla\phi_i$ at a. Introduce the decaying solution $\overline{v}$ to the equation (again, a^* denoting the transpose of a)

$$-\nabla \cdot a^*\nabla\overline{v} = \nabla \cdot g. \qquad (30)$$

We then have the representation

$$\frac{\partial F}{\partial a}(a) = \nabla\overline{v} \otimes (\nabla\phi_i + e_i). \qquad (31)$$

The previous explicit representation of the Fréchet derivative for certain linear functionals of (the gradient of) the corrector (ϕ, σ) enables us to verify the bound (25) for the Malliavin derivative, provided that a certain mean value property is satisfied for a-harmonic functions. Note that the latter requirement is a condition on the coefficient field a; in Lemma 4 below we shall provide a sufficient condition for this property to hold.

As the functionals which the next lemma shall be applied to are basically averages of $\nabla\phi$ or $\nabla\sigma$ over cubes of a certain scale r, we state the lemma in a form which makes it directly applicable in such a setting. In particular, the boundedness assumption (32) for the linear functional is motivated by these considerations.

Lemma 3. Let $r \geq 1$ and consider a linear functional $h \mapsto Fh$ on $L^{\frac{p}{p-1}}(\mathbb{R}^d; \mathbb{R}^d)$ (with $p \in (2, \infty)$) satisfying the support and boundedness condition

$$|Fh| \leq \Big(\fint_{|x|\leq r} |h|^{\frac{p}{p-1}} dx\Big)^{\frac{p-1}{p}}. \qquad (32)$$

11

Suppose that the constraint

$$2 < p < 2 + c(d, \lambda) \tag{33}$$

holds (with $c(d, \lambda) > 0$ to be fixed in the proof below). Let $q \in (1, 2)$ be related to p through

$$\frac{1}{p} = \frac{1}{q} - \frac{1}{2}. \tag{34}$$

Consider the Fréchet derivative $\frac{\partial F}{\partial a}$ of the functional $F := F \nabla \sigma_{ijk}$ (or the functional $F := F \nabla \phi_i$; note that these functionals are nonlinear functionals of a) at some symmetric coefficient field a subject to the conditions (10), (11).
Provided that the coefficient field a is such that the mean value property

$$\fint_{|x| \leq \rho} |\nabla u|^2 dx \lesssim \fint_{|x| \leq R} |\nabla u|^2 dx \quad \text{for any } R \geq r \text{ and any } \rho \in [r, R] \tag{35}$$

holds for any a-harmonic function u and provided that furthermore a is such that

$$\lim_{R \to \infty} \frac{1}{R} \left(\fint_{|x| \leq R} |(\phi, \sigma) - \fint_{|x| \leq R} (\phi, \sigma)|^2 \, dx \right)^{1/2} = 0 \tag{36}$$

is satisfied, we have the estimate

$$\left(\int |\frac{\partial F}{\partial a}|^q dx \right)^{\frac{2}{q}} \lesssim r^{-\frac{(p-2)d}{p}}. \tag{37}$$

Note that for q related to β through (24) and p related to q through (34), we have $r^{-\frac{(p-2)d}{p}} = r^{-\beta}$, i.e. by (37) the L^q-norm of the Malliavin derivative decays like $r^{-\frac{\beta}{2}}$. This demonstrates that for functionals like our averages of $\nabla(\phi, \sigma)$ – note that these functionals have vanishing expectation due to the vanishing expectation of $\nabla(\phi, \sigma)$ – , the concentration of measure indeed improves on large scales with the desired exponent: The "typical value" of the average of $\nabla(\phi, \sigma)$ on some scale r decays like $r^{-\frac{\beta}{2}}$.

We now have to provide a sufficient condition for the mean value property for a-harmonic functions (35). To do so, we make use of the following result from [11], which provides the mean-value property assuming just an appropriate sublinearity condition on the corrector (ϕ, σ).

Proposition 2 (see [11, Lemma 2]). *There exists a constant C_0 only depending on dimension d and ellipticity ratio $\lambda > 0$ with the following property:*

Suppose that for an elliptic coefficient field a subject to the ellipticity and boundedness conditions (10) and (11) the scalar and vector potentials (ϕ, σ), cf. (4) and (7), satisfy

$$\left(\fint_{|x| \leq R} |(\phi, \sigma) - \fint_{|x| \leq R} (\phi, \sigma)|^2 dx \right)^{\frac{1}{2}} \leq \frac{1}{C_0} R \quad \text{for all } R \geq r. \qquad (38)$$

Then for any two radii $R \geq r$ and $\rho \in [r, R]$ and any a-harmonic function u in $\{|x| \leq R\}$ we have

$$\fint_{|x| \leq \rho} |\nabla u|^2 dx \lesssim \fint_{|x| \leq R} |\nabla u|^2 dx.$$

We shall show in the proof of the next lemma that the quantitative sublinearity condition on the corrector (38) may be reduced to a smallness assumption on a certain family of linear functionals of the gradient of the corrector. This reduction relies on the compactness of the left-hand side of (38) with respect to the L^2-norm of $\nabla(\phi, \sigma)$, which in turn may be estimated via Caccioppoli's estimate by the left-hand side. It appeals to a quantitative version of inequalities in functional analysis where an intermediate norm is estimated by a bit of a stronger norm and a lot of a weaker (semi-)norm, the role of which is played by the expression in (39). A slight subtlety follows from the fact that the use of Caccioppoli's inequality increases the radius (by a factor of two, say), so that one has to buckle on the level of all dyadic radii R larger than the given radius r, cf. the expression in (39). This requires the qualitative a priori information (36). One has a lot of flexibility in the choice of the functionals F_n; for pure convenience we choose the same functionals, of Haar wavelet-type, that play a prominent role in the proof of Assertion ii) of Theorem 1. Other natural choices would be the first N eigenfunctions of the Neumann-Laplacian, like in Step 7 of the proof of Theorem 2 in [6] or the proof of Lemma 2.6 in [15].

Lemma 4. *There exist $N \lesssim 1$ linear functionals $F_1, \cdots, F_N$ satisfying the support and boundedness condition*

$$|F_n h| \lesssim \left(\int_{|x| \leq 1} |h|^{\frac{p}{p-1}} \, dx \right)^{\frac{p-1}{p}}$$

such that the following holds: Denote by $F_{n,R}$ the rescaling of F_n given by $F_{n,R} h = F_n \left(h(\frac{\cdot}{R}) \right)$. Let $r \geq 0$. Provided that the condition

$$\sup_{R \geq r \; dyadic; n=1,\cdots,N} F_{n,R} \nabla(\phi, \sigma) \ll 1 \qquad (39)$$

13

and the condition (36) are satisfied, we have the smallness estimate (38) for the corrector; in particular, by Proposition 2 the mean value property (35) holds for any a-harmonic function u on scales ≥ r, i.e. we have

$$\fint_{|x| \leq \rho} |\nabla u|^2 dx \lesssim \fint_{|x| \leq R} |\nabla u|^2 dx \quad \text{for any } R \geq r \text{ and any } \rho \in [r, R].$$

With these preparations, we are able to establish our main theorem. The main technical difficulty in the proof below is that our estimate

$$\left(\int |\frac{\partial F}{\partial a}|^q dx \right)^{\frac{2}{q}} \lesssim r^{-\beta}$$

for the Malliavin derivative of linear functionals of the gradient of the corrector (cf. (37)) is a conditional bound: It relies on the assumption that the mean-value property (35) holds for a-harmonic functions on scales larger than r. For the concentration of measure estimate (22), however, an unconditional estimate of the form (21) or (25) (the latter being a proxy for (21)) is needed. By Lemma 4 we know that the mean-value property holds, provided that for a certain family of linear functionals of the corrector the smallness estimate

$$\sup_{R \geq r \text{ dyadic}; n=1,\cdots,N} F_{n,R} \nabla(\phi, \sigma) \leq \frac{1}{C_0}$$

is satisfied (C_0 being a universal constant). To circumvent this problem, in the proof below we therefore introduce the family of functionals

$$\bar{F}_r := \min \left\{ \sup_{R \geq r \text{ dyadic}; n=1,\cdots,N} F_{n,R} \nabla(\phi, \sigma), \frac{1}{C_0} \right\},$$

for which by design the unconditional bound for the Malliavin derivative

$$\left(\int |\frac{\partial \bar{F}_r}{\partial a}|^q dx \right)^{\frac{2}{q}} \lesssim r^{-\beta}$$

holds. Therefore, concentration of measure is applicable to $\bar{F}_r$. The remainder of the proof of the first part of our theorem below is dedicated to handling the (a priori unknown) expectation $\langle \bar{F}_r \rangle$.

The proof of the second assertion of our main theorem will mainly rely on the first assertion of the theorem as well as the quantitative improvement of the Malliavin derivative of averages of $(\nabla\phi, \nabla\sigma)$ on larger scales, as captured by the estimate (37).

3 Concentration of Measure and Estimates of the Malliavin Derivative

3.1 Concentration of Measure

Proof of Proposition 1. For the proof of the concentration of measure estimate (22), we refer the reader to [20, Proposition 2.18]. We now establish (23). By Chebychev's inequality, (22) implies $\langle I(F - \langle F \rangle \geq M) \rangle \leq \exp(-\frac{M^2}{2})$. In combination with the same estimate with F replaced by $-F$, we obtain (23). $\qquad\square$

Proof of Lemma 1. We need to verify that the condition (21) is implied by the assumption (25).

To do so, we first note that by Hölder's inequality we have for any exponent $1 < q < \infty$

$$\int \frac{\partial F}{\partial \tilde{a}} \mathrm{Cov} \frac{\partial F}{\partial \tilde{a}} dx \;\leq\; \Big(\int |\frac{\partial F}{\partial \tilde{a}}|^q dx \Big)^{\frac{1}{q}} \Big(\int |\mathrm{Cov}\frac{\partial F}{\partial \tilde{a}}|^{\frac{q}{q-1}} dx \Big)^{\frac{q-1}{q}}.$$

Since Cov is the convolution with $\langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle$ and since we have the bound $|\langle \tilde{a}(x) \otimes \tilde{a}(0) \rangle| \leq |x|^{-\beta}$, we have for the second factor

$$\Big(\int |\mathrm{Cov}\frac{\partial F}{\partial \tilde{a}}|^{\frac{q}{q-1}} dx \Big)^{\frac{q-1}{q}} \;\leq\; \Big(\int \big| \int \frac{1}{|x-y|^\beta} |\frac{\partial F}{\partial \tilde{a}}(y)| dy \big|^{\frac{q}{q-1}} dx \Big)^{\frac{q-1}{q}},$$

which allows us to use the Hardy-Littlewood-Sobolev inequality

$$\Big(\int \big| \int \frac{1}{|x-y|^\beta} |\frac{\partial F}{\partial \tilde{a}}(y)| dy \big|^{\frac{q}{q-1}} dx \Big)^{\frac{q-1}{q}} \;\lesssim\; \Big(\int |\frac{\partial F}{\partial \tilde{a}}|^q dx \Big)^{\frac{1}{q}},$$

provided the exponents q and β are related by (24). From this string of inequalities we learn that (21) also holds provided

$$\Big(\int |\frac{\partial F}{\partial \tilde{a}}|^q dx \Big)^{\frac{2}{q}} \ll 1 \quad \text{for almost every } \tilde{a}. \tag{40}$$

We now change variables according to $a(x) = \Phi(\tilde{a}(x))$; by the chain rule for $F(\tilde{a}) = F(\Phi(\tilde{a}))$ we have $\frac{\partial F}{\partial \tilde{a}}(\tilde{a}, x) = \Phi'(\tilde{a}(x))\frac{\partial F}{\partial a}(a, x)$, so that by the 1-Lipschitz continuity of Φ, our assumption (25) implies (40) and thus (21). $\quad\square$

3.2 Representation of the Malliavin derivative

Proof of Lemma 2. We first give the argument for the "vector potential" σ, fixing a component σ_{ijk}. Consider a functional of the form $F := F\nabla\sigma_{ijk}$ with Fh as in (26). We claim that the Fréchet derivative of F with respect to a is given by (29) where the functions $v = v(x)$ and $\tilde{v}_{jk} = \tilde{v}_{jk}(a, x)$ are determined as the decaying solutions of the elliptic equations (27) and (28). Computing the functional derivative of F as a function of a amounts to a linearization. We thus consider an arbitrary tensor field $\delta a = \delta a(x)$, which we think of as an infinitesimal perturbation of a, and which thus generates infinitesimal perturbations $\delta\phi$ and $\delta\sigma$ of ϕ and σ according to (4), (6), and (9), that is,

$$-\nabla \cdot (a\nabla\delta\phi_i + \delta a(\nabla\phi_i + e_i)) = 0 \tag{41}$$

and

$$-\triangle\delta\sigma_{ijk} = \partial_j\big(\delta a(\nabla\phi_i + e_i) + a\nabla\delta\phi_i\big)_k - \partial_k\big(\delta a(\nabla\phi_i + e_i) + a\nabla\delta\phi_i\big)_j. \tag{42}$$

In terms of the infinitesimal perturbation δF of F, this implies by integration by parts (or rather by directly appealing to the weak Lax-Milgram formulations of the elliptic equations)

$$
\begin{aligned}
\delta F &= \int g \cdot \nabla\delta\sigma_{ijk}dx \\
&\stackrel{(27)}{=} -\int \nabla v \cdot \nabla\delta\sigma_{ijk}dx \\
&\stackrel{(42)}{=} \int (\partial_j ve_k - \partial_k ve_j) \cdot \big(\delta a(\nabla\phi_i + e_i) + a\nabla\delta\phi_i\big)dx \\
&\stackrel{(28)}{=} \int (\partial_j ve_k - \partial_k ve_j) \cdot \delta a(\nabla\phi_i + e_i)dx - \int \nabla\tilde{v}_{jk} \cdot a\nabla\delta\phi_i dx \\
&\stackrel{(41)}{=} \int \big(\partial_j ve_k - \partial_k ve_j + \nabla\tilde{v}_{jk}\big) \cdot \delta a(\nabla\phi_i + e_i)dx,
\end{aligned}
$$

which is nothing else than (29).

Let us now establish the second part of our lemma. Consider a functional of the scalar potential of the form $F := F\nabla\phi_i$. To represent its Fréchet derivative, introduce the decaying solution $\overline{v}$ to the equation (30). We observe

16

that the variation of F with respect to a is given by

$$
\begin{aligned}
\delta F &= \int g \cdot \nabla \delta \phi_i dx \\
&\overset{(30)}{=} -\int a^* \nabla \overline{v} \cdot \nabla \delta \phi_i dx \\
&\overset{(41)}{=} \int \nabla \overline{v} \cdot \delta a (\nabla \phi_i + e_i) dx,
\end{aligned}
$$

which leads to the conclusion (31). $\qquad\square$

3.3 Sensitivity estimate

Proof of Lemma 3. We now argue that under certain boundedness assumptions on $F = Fh$ as a linear functional in vector fields $h = h(x)$, we control the size (25) of its Fréchet derivative $\frac{\partial F}{\partial a} = \frac{\partial F}{\partial a}(a, x)$ as a nonlinear functional $F\nabla \sigma_{ijk} = F(a)$ in coefficient fields $a = a(x)$ (and similarly in the case $F(a) = F\nabla\phi_i$; for this case, the (simpler) proof is sketched afterwards).

To this aim, let us first note that we have a Calderon-Zygmund estimate for $-\nabla \cdot a \nabla$ with the exponents p and its dual exponent $\frac{p}{p-1}$: For any decaying function w and vector field h on $\mathbb{R}^d$ related by

$$
-\nabla \cdot a \nabla w = \nabla \cdot h \tag{43}
$$

we have

$$
\int |\nabla w|^{\frac{p}{p-1}} dx \lesssim \int |h|^{\frac{p}{p-1}} dx \quad \text{and} \quad \int |\nabla w|^p dx \lesssim \int |h|^p dx. \tag{44}
$$

This assertion holds by Meyer's estimate (see e.g. [22]), which only requires the ellipticity and boundedness assumptions (10), (11) on a as well as the estimate $|p - 2| \ll 1$, which is ensured by our condition (33). Note that an analogous estimate would hold for the dual equation $-\nabla \cdot a^* \nabla w = \nabla \cdot h$ if our coefficient field were nonsymmetric.

In the following, we will use the abbreviation $\|\cdot\|_{p,B}$ for the spatial L^p-norm on the set B; we write $\|\cdot\|_p$ when $B = \mathbb{R}^d$. We start by arguing that because $\frac{p}{p-1} \in (1, 2)$, (35) also entails

$$
\fint_{|x| \leq \rho} |\nabla u|^{\frac{p}{p-1}} dx \lesssim \fint_{|x| \leq R} |\nabla u|^{\frac{p}{p-1}} dx. \tag{45}
$$

It is obviously enough to establish (45) only for $R \geq 2\rho$; hence by Jensen's inequality, (45) follows from (35) once we establish the reverse Hölder inequality

$$\left(\fint_{|x| \leq \frac{R}{2}} |\nabla u|^2 dx\right)^{\frac{1}{2}} \lesssim \fint_{|x| \leq R} |\nabla u| dx. \tag{46}$$

To this purpose, we test $-\nabla \cdot a \nabla u = 0$ with $\eta^{2\gamma}(u - m)$, where η is a smooth cut-off of $\chi_{\{|x| \leq \frac{R}{2}\}}$ in $\{|x| \leq R\}$ (with the property $|\nabla \eta| \lesssim \frac{1}{R}$) and where the exponent $\gamma \geq 1$ and the constant $m \in \mathbb{R}$ will be chosen later. By the ellipticity and boundedness assumptions (10), (11) and Young's inequality we obtain

$$\int (\eta^\gamma |\nabla u|)^2 dx \lesssim \int ((u - m)|\nabla \eta^\gamma|)^2 dx,$$

and thus

$$\int |\nabla(\eta^\gamma (u - m))|^2 dx \lesssim \int ((u - m)|\nabla \eta^\gamma|)^2 dx,$$

which by the estimate on $\nabla \eta$ gives

$$\|\nabla(\eta^\gamma (u - m))\|_2 \lesssim \frac{1}{R}\|\eta^{\gamma - 1}(u - m)\|_2. \tag{47}$$

On the r. h. s. of (47) we use first Hölder's inequality, then the isoperimetric inequality on $\{|x| \leq R\}$ and finally Sobolev's inequality on the whole space (for simplicity, we assume $d > 2$ here)

$$\begin{aligned}
\|\eta^{\gamma-1}(u - m)\|_2 &\leq \|\eta^\gamma(u - m)\|_{\frac{2d}{d-2}}^{\frac{\gamma-1}{\gamma}} \|u - m\|_{\frac{d}{d-1}, |x| \leq R}^{\frac{1}{\gamma}} \\
&\lesssim \|\nabla(\eta^\gamma(u - m))\|_2^{\frac{\gamma-1}{\gamma}} \|\nabla u\|_{1, |x| \leq R}^{\frac{1}{\gamma}},
\end{aligned}$$

provided the exponent $\gamma \in (1, \infty)$ is chosen such that $\frac{1}{2} = (1 - \frac{1}{\gamma})\frac{d-2}{2d} + \frac{1}{\gamma}\frac{d-1}{d}$ (which – as a simple computation shows – is satisfied precisely for $\gamma = \frac{d}{2}$) and the constant m is the spatial average of u on $\{|x| \leq R\}$. The combination of the last two estimates yields

$$\|\nabla(\eta^\gamma(u - m))\|_2 \lesssim \frac{1}{R}\|\nabla(\eta^\gamma(u - m))\|_2^{\frac{\gamma-1}{\gamma}} \|\nabla u\|_{1, |x| \leq R}^{\frac{1}{\gamma}},$$

which (by $\gamma = \frac{d}{2}$) entails $\|\nabla u\|_{2, |x| \leq \frac{R}{2}} \leq \|\nabla(\eta^\gamma(u - m))\|_2 \lesssim R^{-\frac{d}{2}}\|\nabla u\|_{1, |x| \leq R}$ and thus (46).

We now give the argument for (37) in case of a functional of the form $F\nabla\sigma_{ijk}$ (the case $F\nabla\phi_i$ will be treated below). Clearly (32) implies that there exists a (deterministic) vector field $g = g(x)$ with

$$\operatorname{supp} g \subset \{|x| \leq r\} \quad \text{and} \quad \|g\|_p \leq r^{-\frac{p-1}{p}d} \tag{48}$$

18

such that we have the representation for $F = Fh$ as a linear functional on vector fields $h = h(x)$

$$Fh = \int g \cdot h\, dx. \tag{49}$$

This gives us access to the representation (29) of its Fréchet derivative $\frac{\partial F}{\partial a}$ considered as a nonlinear functional $F\nabla\sigma_{ijk} = F(a)$ of a. Using this representation, a partition into dyadic annuli, and Hölder's estimate (recall (34)) we obtain

$$
\begin{aligned}
\|\frac{\partial F}{\partial a}\|_q \\
\lesssim \|(|\nabla v| + |\nabla\tilde{v}_{jk}|)|\nabla\phi_i + e_i|\|_{q,|x|\leq 2r} \\
+ \sum_{n=1}^{\infty} \|(|\nabla v| + |\nabla\tilde{v}_{jk}|)|\nabla\phi_i + e_i|\|_{q,2^n r \leq |x| \leq 2^{n+1}r} \\
\lesssim (\|\nabla v\|_{p,|x|\leq 2r} + \|\nabla\tilde{v}_{jk}\|_{p,|x|\leq 2r})\|\nabla\phi_i + e_i\|_{2,|x|\leq 2r} \\
+ \sum_{n=1}^{\infty}(\|\nabla v\|_{p,2^n r \leq |x| \leq 2^{n+1}r} + \|\nabla\tilde{v}_{jk}\|_{p,2^n r \leq |x| \leq 2^{n+1}r})\|\nabla\phi_i + e_i\|_{2,2^n r \leq |x| \leq 2^{n+1}r}.
\end{aligned}
\tag{50}
$$

In view of (35) applied to the a-harmonic function $u(x) = x_i + \phi_i(x)$, cf. (4), we obtain for all radii $\rho \geq r$ using Caccioppoli's inequality and (36)

$$
\begin{aligned}
\fint_{|x|\leq\rho} |\nabla\phi_i + e_i|^2 dx &\lesssim \rho^d \liminf_{R\uparrow\infty} \fint_{|x|\leq R} |\nabla\phi_i + e_i|^2 dx \\
&\lesssim \rho^d \liminf_{R\uparrow\infty} \frac{1}{R^2}\fint_{|x|\leq 2R} \left|\phi_i + x_i - \fint_{|x|\leq 2R}(\phi_i + x_i)\right|^2 dx \lesssim \rho^d.
\end{aligned}
$$

Hence (50) turns into

$$
\|\frac{\partial F}{\partial a}\|_q \lesssim r^{\frac{d}{2}}(\|\nabla v\|_p + \|\nabla\tilde{v}_{jk}\|_p) \tag{51}
$$

$$
+ \sum_{n=1}^{\infty}(2^n r)^{\frac{d}{2}}(\|\nabla v\|_{p,|x|\geq 2^n r} + \|\nabla\tilde{v}_{jk}\|_{p,|x|\geq 2^n r}). \tag{52}
$$

It thus remains to estimate the auxiliary functions v and $\tilde{v}_{jk}$. The estimate of the terms in line (51) is easy: By (48) and Calderon-Zygmund for (27) we obtain $\|\nabla v\|_p \lesssim \|g\|_p \leq r^{-\frac{p-1}{p}d}$. By (44) we have Calderon-Zygmund with exponent p for the equation (28), so that $\|\nabla\tilde{v}_{jk}\|_p \lesssim \|\nabla v\|_p \lesssim r^{-\frac{p-1}{p}d}$.

In order to control the terms in line (52), we shall establish the following estimates for $n \in \mathbb{N}$

$$\|\nabla v\|_{p,|x|\geq 2^n r} \lesssim (2^n)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d}, \tag{53}$$

$$\|\nabla \tilde{v}_{jk}\|_{p,|x|\geq 2^n r} \lesssim n(2^n)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d}. \tag{54}$$

We note that since $p > 2$, these estimates imply that the sum over n in (52) converges and gives (37).

The estimate (53) for the solution v of the constant coefficient equation (27) is classical: We already argued that $\|\nabla v\|_p \lesssim r^{-\frac{p-1}{p}d}$; by the estimate on the support of g in (48) we have that v is harmonic in $\{|x| \geq r\}$ and that it has vanishing flux $\int_{|x|=r} x \cdot \nabla v = 0$. It thus decays as $|\nabla v(x)| \lesssim |x|^{-d} r^{d-\frac{d}{p}} \|\nabla v\|_p$ for $|x| \geq 2r$, which in particular yields (53). We now turn to (54) and to this purpose rewrite the equation (28) for $\tilde{v}_{jk}$ as

$$-\nabla \cdot a^* \nabla \tilde{v}_{jk} = \nabla \cdot \tilde{g}$$

with the r. h. s. $\tilde{g} := -a^*(\partial_j v e_k - \partial_k v e_j)$. We already argued that $\|\tilde{g}\|_p \lesssim r^{-\frac{p-1}{p}d}$ and (53) translates into

$$\|\tilde{g}\|_{p,|x|\geq 2^n r} \lesssim (2^n)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d}. \tag{55}$$

In order to proceed, we split $\tilde{g}$ into $\{\tilde{g}_m\}_{m=0,1,\cdots}$ where $\tilde{g}_0$ is supported in $\{|x| \leq 2r\}$ and for $m \geq 1$ $\tilde{g}_m$ is supported in $\{2^m r \leq |x| \leq 2^{m+1} r\}$, so that (55) translates into

$$\|\tilde{g}_m\|_p \lesssim (2^m)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d}. \tag{56}$$

This entails a splitting of $\tilde{v}_{jk}$ into $\{\tilde{v}_m\}_{m=0,1,\cdots}$, where $\tilde{v}_m$ is the Lax-Milgram solution of

$$-\nabla \cdot a^* \nabla \tilde{v}_m = \nabla \cdot \tilde{g}_m. \tag{57}$$

We will now argue that

$$\|\nabla \tilde{v}_m\|_{p,|x|\geq 2^n r} \lesssim \min\{(2^n)^{-d+\frac{d}{p}}, (2^m)^{-d+\frac{d}{p}}\} r^{-\frac{p-1}{p}d}, \tag{58}$$

which implies the estimate (54) by the triangle inequality $\|\nabla \tilde{v}_{jk}\|_{p,|x|\geq 2^n r} \leq \sum_{m=0}^{\infty} \|\nabla \tilde{v}_m\|_{p,|x|\geq 2^n r}$. We note that (56) together with our Calderon-Zygmund estimate (44) applied to (57) yields $\|\nabla \tilde{v}_m\|_p \lesssim (2^m)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d}$. In order to establish (58), it thus remains to show

$$\|\nabla \tilde{v}_m\|_{p,|x|\geq 2^n r} \lesssim (2^n)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d} \quad \text{for } m < n. \tag{59}$$

We argue in favor of (59) by duality and thus consider an arbitrary $h \in L^{\frac{p}{p-1}}$ supported in $\{|x| \geq 2^n r\}$ and denote by w the corresponding Lax-Milgram solution of (43). By integration by parts, we deduce from (43) and (57) that $\int h \cdot \nabla \tilde{v}_m dx = \int \tilde{g}_m \cdot \nabla w \, dx$. By the support condition on $\tilde{g}_m$ this yields

$$\left| \int h \cdot \nabla \tilde{v}_m dx \right| \leq \|\tilde{g}_m\|_p \|\nabla w\|_{\frac{p}{p-1}, |x| \leq 2^{m+1} r}.$$

By the support assumption on h we have that w is a-harmonic in $\{|x| \leq 2^n r\}$. Since $m < n$, we may use (45) applied to w in form of

$$\|\nabla w\|_{\frac{p}{p-1}, |x| \leq 2^{m+1} r} \lesssim (2^{n-m})^{-d+\frac{d}{p}} \|\nabla w\|_{\frac{p}{p-1}, |x| \leq 2^n r}.$$

We combine this with (44) in form of $\|\nabla w\|_{\frac{p}{p-1}} \lesssim \|h\|_{\frac{p}{p-1}}$, and with (56), to obtain

$$\left| \int h \cdot \nabla \tilde{v}_m dx \right| \lesssim (2^n)^{-d+\frac{d}{p}} r^{-\frac{p-1}{p}d} \|h\|_{\frac{p}{p-1}},$$

which gives (59).

In the case of a functional of the scalar potential of the form $F(a) = F \nabla \phi_i$, we claim that the Fréchet derivative of F is again controlled in the sense of (37). The proof is mostly analogous to the previous one; we again rewrite F as in (49) with some g satisfying (48). Starting from the representation (31), one derives an analogue of estimate (50) reading

$$
\begin{aligned}
\|\frac{\partial F}{\partial a}\|_q \\
\lesssim \quad &\|\nabla \overline{v}\|_{p, |x| \leq 2r} \|\nabla \phi_i + e_i\|_{2, |x| \leq 2r} \\
&+ \sum_{n=1}^{\infty} \|\nabla \overline{v}\|_{p, 2^n r \leq |x| \leq 2^{n+1} r} \|\nabla \phi_i + e_i\|_{2, 2^n r \leq |x| \leq 2^{n+1} r}.
\end{aligned}
$$

The second factors on the right in this estimate coincide with the ones in the case $F(a) = F \nabla \sigma_{ijk}$; therefore, we get the following analogue to estimate (52):

$$\|\frac{\partial F}{\partial a}\|_q \quad \lesssim \quad r^{\frac{d}{2}} \|\nabla \overline{v}\|_p + \sum_{n=1}^{\infty} (2^n r)^{\frac{d}{2}} \|\nabla \overline{v}\|_{p, |x| \geq 2^n r}.$$

The equation (30) for $\overline{v}$ has the structure of the equation (57) with $m = 0$ (including the estimate $\||g|\|_p \lesssim r^{-\frac{p-1}{p}d}$ and the inclusion supp $g \subset \{|x| \leq 2r\}$, cf. (48)); therefore, the decay property (59) carries over to our $\overline{v}$. This establishes the estimate

$$\|\frac{\partial F}{\partial a}\|_q \lesssim r^{-\frac{p-2}{2p}d}$$

also in the case $F(a) = F \nabla \phi_i$. $\qquad \square$

3.4 Sufficient conditions for the mean value property in terms of linear functionals of the corrector

Proof of Lemma 4. In order to show that (39) and (36) imply (35), we only need to show the existence of functionals $F_1, \cdots, F_N$ such that (39) and (36) imply

$$\frac{1}{R}\Big(\fint_{|x|\le R}|(\phi,\sigma)-\fint_{|x|\le R}(\phi,\sigma)|^2 dx\Big)^{\frac{1}{2}} \ll 1 \quad \text{for all } R \ge r. \qquad (60)$$

By Proposition 2, the estimate (35) follows from (60).

Let us now give the argument for (60). First, it is clearly enough to show that for any $0 < \delta \ll 1$, there exists $N \lesssim \delta^{-d}$ functionals $F_1, \cdots, F_N$ on vector fields which are bounded in the sense of

$$|F_n h| \lesssim \delta^{-d}\Big(\int_{|x|\le 1} |h|^{\frac{p}{p-1}}dx\Big)^{\frac{p-1}{p}} \qquad (61)$$

and such that for any dyadic $\rho \ge 1$

$$\frac{1}{\rho}\Big(\fint_{|x|\le \rho}|(\phi,\sigma)-\fint_{|x|\le \rho}(\phi,\sigma)|^2 dx\Big)^{\frac{1}{2}} \lesssim \delta + \sup_{R\ge\rho \text{ dyadic}} \max_{n=1,\cdots,N} |F_{n,R}\nabla(\phi,\sigma)|. \qquad (62)$$

By dyadic iteration, it is enough to show for any dyadic $\rho \ge 1$

$$\frac{1}{\rho}\Big(\fint_{|x|\le \rho} |(\phi,\sigma)-\fint_{|x|\le \rho}(\phi,\sigma)|^2 dx\Big)^{\frac{1}{2}}$$
$$\lesssim \max_{n=1,\cdots,N} |F_{n,\rho}\nabla(\phi,\sigma)|$$
$$+ \delta\Big(\frac{1}{2\rho}\Big(\fint_{|x|\le 2\rho} |(\phi,\sigma)-\fint_{|x|\le 2\rho}(\phi,\sigma)|^2 dx\Big)^{\frac{1}{2}} + 1\Big). \qquad (63)$$

Indeed, abbreviating $D_m := \frac{1}{2^m}\Big(\fint_{|x|\le 2^m}|(\phi,\sigma)-\fint_{|x|\le 2^m}(\phi,\sigma)|^2 dx\Big)^{\frac{1}{2}}$, the estimate (63) may be rewritten as (using a slight readjustment of δ)

$$D_m \le C_0 \max_{n=1,\cdots,N} |F_{n,2^m}\nabla(\phi,\sigma)| + \delta\Big(D_{m+1} + 1\Big),$$

which may be iterated to

$$D_m \le \frac{1}{1-\delta}C_0 \max_{M=m,m+1,\cdots,m+m_0} \max_{n=1,\cdots,N} |F_{n,2^M}\nabla(\phi,\sigma)|$$
$$+ \frac{\delta}{1-\delta} + \delta^{m_0+1}D_{m+m_0+1}.$$

By our sublinearity assumption on the corrector (36) (which may be rewritten as $\lim_{m_0 \uparrow \infty} D_{m_0} = 0$), this yields (62).

We now turn to the argument for (63). By Caccioppoli's estimate on (4) we have

$$\left(\int_{|x| \leq \frac{3}{2}\rho} |\nabla \phi_i|^2 dx\right)^{\frac{1}{2}} \lesssim \frac{1}{\rho}\left(\int_{|x| \leq 2\rho} \left(\phi_i - \fint_{|x| \leq 2\rho} \phi_i\right)^2 dx\right)^{\frac{1}{2}} + \rho^{d/2},$$

and thus in particular for the flux $q_i = a(\nabla \phi_i + e_i)$

$$\left(\int_{|x| \leq \frac{3}{2}\rho} |q_i|^2 dx\right)^{\frac{1}{2}} \lesssim \frac{1}{\rho}\left(\int_{|x| \leq 2\rho} \left(\phi_i - \fint_{|x| \leq 2\rho} \phi_i\right)^2 dx\right)^{\frac{1}{2}} + \rho^{d/2}.$$

Caccioppoli's estimate on (9) gives

$$\left(\int_{|x| \leq \rho} |\nabla \sigma_{ijk}|^2 dx\right)^{\frac{1}{2}}$$

$$\lesssim \frac{1}{\rho}\left(\int_{|x| \leq \frac{3}{2}\rho} \left(\sigma_{ijk} - \fint_{|x| \leq \frac{3}{2}\rho} \sigma_{ijk}\right)^2 dx\right)^{\frac{1}{2}} + \left(\int_{|x| \leq \frac{3}{2}\rho} |q_i|^2 dx\right)^{\frac{1}{2}}$$

$$\leq \frac{1}{\rho}\left(\int_{|x| \leq 2\rho} \left(\sigma_{ijk} - \fint_{|x| \leq 2\rho} \sigma_{ijk}\right)^2 dx\right)^{\frac{1}{2}} + \left(\int_{|x| \leq \frac{3}{2}\rho} |q_i|^2 dx\right)^{\frac{1}{2}}$$

The last three estimates combine to

$$\left(\int_{|x| \leq \rho} |\nabla(\phi, \sigma)|^2 dx\right)^{\frac{1}{2}}$$

$$\lesssim \frac{1}{\rho}\left(\int_{|x| \leq 2\rho} \left|(\phi, \sigma) - \fint_{|x| \leq 2\rho} (\phi, \sigma)\right|^2 dx\right)^{\frac{1}{2}} + \rho^{d/2}. \qquad (64)$$

Hence, to establish (63) it is enough to show

$$\frac{1}{\rho}\left(\fint_{|x| \leq \rho} \left|(\phi, \sigma) - \fint_{|x| \leq \rho} (\phi, \sigma)\right|^2 dx\right)^{\frac{1}{2}}$$

$$\leq \max_{n=1,\cdots,N} |F_{n,\rho} \nabla(\phi, \sigma)| + \delta\left(\fint_{|x| \leq \rho} |\nabla(\phi, \sigma)|^2 dx\right)^{\frac{1}{2}}.$$

This statement is not just true for $(\phi, \sigma) - \fint_{|x| \leq \rho} (\phi, \sigma)$, but for any function ζ of vanishing spatial average on $\{|x| \leq \rho\}$: By rescaling, it is sufficient to show the estimate on the unit ball $\{|x| \leq 1\}$. It is more convenient to see it when the unit ball $\{|x| \leq 1\}$ is replaced by the unit square $(0,1)^d$:

$$\fint_{(0,1)^d} \zeta^2 dx \leq \max_{n=1,\cdots,N} |F_n \nabla \zeta|^2 + \delta^2 \int_{(0,1)^d} |\nabla \zeta|^2 dx. \qquad (65)$$

23

Indeed, dividing $(0,1)^d$ into $N = \delta^{-d}$ (suppose that δ^{-1} is an integer) sub-cubes $\{Q_n\}_{n=1,\cdots,N}$ of side length δ and setting $F_n \nabla \zeta := \fint_{Q_n} \zeta dx$ (recall that $\int_{(0,1)^d} \zeta dx = 0$ so that F_n is indeed a function of $\nabla \zeta$), (65) follows from using Poincaré's estimate on each Q_n in form of $\int_{Q_n} \zeta^2 dx - |Q_n|(\fint_{Q_n} \zeta dx)^2 \lesssim \delta^2 \int_{Q_n} |\nabla \zeta|^2 dx$ and then summing up. We note that by Poincaré's estimate on $(0,1)^d$, the F_n have the desired boundedness property (61), at first on gradient fields $\nabla \zeta$

$$|F_n \nabla \zeta| \leq \delta^{-d} \Big(\int_{(0,1)^d} |\zeta|^{\frac{p}{p-1}} dx \Big)^{\frac{p-1}{p}} \lesssim \delta^{-d} \Big(\int_{(0,1)^d} |\nabla \zeta|^{\frac{p}{p-1}} dx \Big)^{\frac{p-1}{p}},$$

and then on any vector field h by extension à la Hahn-Banach. $\qquad\square$

4   Proof of Main Result

Proof of Theorem 1.
Proof of Assertion i).
Consider the functionals $\{F_n\}_{n=1,\cdots,N}$ and their rescalings $\{F_{n,R}\}_{n=1,\cdots,N;R \text{ dyadic}}$ constructed in Lemma 4. Let us abbreviate $F_{n,R} \nabla(\phi, \sigma)$ as $F_{n,R}$. We would like to apply concentration of measure to these functionals.
The main difficulty that we need to overcome is that our sensitivity estimate (37) in Lemma 3 for the quantity $F_{n,r}$ is based on the assumption that the mean-value property (35) holds for a-harmonic functions down to scale r. By Lemma 4 this assumption may be reduced to the smallness assumption (39) for our functionals $F_{m,R}$ on scales $R \geq r$, so that Lemma 3 becomes applicable under the assumption (39): Let q be related to β through (24) and let p be related to q through (34). By the smallness assumption on β in our theorem (cf. (13)), we deduce that (33) holds. By scaling, our functionals $F_{n,r}$ satisfy the estimate (32) up to a universal constant factor. Furthermore, by ergodicity the property (36) holds for $\langle \cdot \rangle$-almost every coefficient field a (regarding σ, this result has been shown in [11, Lemma 1]; for ϕ, it is classical but may also be found in [11]). Thus, the estimate (37) holds for $F_{n,r}$ under the assumption (39), i.e. there exists a constant C_0 only depending on d, λ, and β, such that for any $n = 1, \cdots, N$ and any radius r the implication

$$\sup_{m,R \geq r \text{ dyadic}} F_{m,R} \leq \frac{1}{C_0} \quad \Longrightarrow \quad \|\frac{\partial F_{n,r}}{\partial a}\|_q^2 \lesssim \frac{1}{r^\beta} \tag{66}$$

holds for $\langle \cdot \rangle$-a.e. coefficient field a.

To apply concentration of measure in the form of Proposition 1 to some functional F, we however need an unconditional bound on the Malliavin derivative (cf. (21) respectively (25)).

Therefore we first introduce a new random variable whose derivative vanishes whenever the smallness condition in (66) is violated: Consider the auxiliary random variable

$$\bar{F}_r := \min\{\sup_{m,R \geq r} |F_{m,R}|, \frac{1}{C_0}\},\tag{67}$$

where the sup runs over all dyadic radii $R = 2^k r$, $k \in \mathbb{N}_0$. By the usual differentiation rules applied to the Fréchet derivative $\frac{\partial}{\partial a}$ in the norm $\|\cdot\|_q$, we obtain

$$\|\frac{\partial \bar{F}_r}{\partial a}\|_q \leq I(\sup_{m,R \geq r} F_{m,R} \leq \frac{1}{C_0}) \sup_{m,R \geq r} \|\frac{\partial F_{m,R}}{\partial a}\|_q$$

and thus by (66)

$$\|\frac{\partial \bar{F}_r}{\partial a}\|_q^2 \lesssim \frac{1}{r^\beta}.$$

By Lemma 1, we may apply concentration of measure in form of (23) to the random variable $cr^{\beta/2}\bar{F}_r$ (where c is some small universal constant). This yields

$$\langle I(|\bar{F}_r - \langle \bar{F}_r \rangle| \geq M)\rangle \lesssim \exp(-\frac{1}{C}r^\beta M^2) \quad \text{for all } M \geq 0,\tag{68}$$

so that it remains to control the expectation $\langle \bar{F}_r \rangle$.

Because of (8) and the definition of $F_{m,R}$, it follows from qualitative ergodicity of $\langle \cdot \rangle$ and Birkhoff's ergodic theorem that $\lim_{R \uparrow \infty} F_{m,R} = 0$ almost surely, so that by dominated convergence $\lim_{r \uparrow \infty} \langle \bar{F}_r \rangle = 0$. Hence there exists a *finite* radius r_0 which is minimal with the property

$$\langle \bar{F}_r \rangle \leq \frac{1}{4C_0} \quad \text{for all } r \geq r_0.\tag{69}$$

Hence using (68) for $M = \frac{1}{4C_0}$ we get in view of the definition (67) of $\bar{F}_r$

$$\langle I(\sup_{m,R \geq r} |F_{m,R}| \geq \frac{1}{2C_0})\rangle \lesssim \exp(-\frac{1}{C}r^\beta) \quad \text{for all } r \geq r_0.\tag{70}$$

On the basis of (70), we now get a *quantitative* estimate on r_0. To this purpose we now consider the auxiliary variable

$$\bar{F}_{n,r} := \eta(\sup_{m,R \geq r} F_{m,R})F_{n,r},\tag{71}$$

where again the sup runs over all dyadic radii $R = 2^k r$, $k \in \mathbb{N}_0$, and where the cut-off function $\eta = \eta(F)$ is given by

$$\eta(F) = \max\{\min\{2C_0(\frac{1}{C_0} - F), 1\}, 0\}.\tag{72}$$

The advantage of the auxiliary variable (71) over (67) is that we control its expectation: Since the stationary $\nabla(\phi, \sigma)$ has vanishing expectation, cf. (8), and by the linearity of $F_{n,r}$ in $\nabla(\phi, \sigma)$ we have $\langle F_{n,r} \rangle = 0$ and thus $\langle \bar{F}_{n,r} \rangle = \langle (\eta - 1) F_{n,r} \rangle$ so that by construction of η

$$|\langle \bar{F}_{n,r} \rangle| \leq \langle I(\sup_{m,R \geq r} |F_{m,R}| \geq \frac{1}{2C_0}) |F_{n,r}| \rangle.$$

Since the stationary $\nabla(\phi, \sigma)$ has bounded second moments, cf. (8), and by the boundedness property of $F_{n,r}$ in $\nabla(\phi, \sigma)$ we obtain from the Cauchy-Schwarz inequality

$$\langle \bar{F}_{n,r} \rangle^2 \lesssim \langle I(\sup_{m,R \geq r} |F_{m,R}| \geq \frac{1}{2C_0}) \rangle,$$

which in view of (70) improves to

$$|\langle \bar{F}_{n,r} \rangle| \lesssim \exp(-\frac{1}{C} r^\beta) \quad \text{for any } r \geq r_0. \tag{73}$$

By differentiation rules for the Fréchet derivative $\frac{\partial}{\partial a}$ in the norm $\| \cdot \|_q$ we obtain for the auxiliary random variable $\bar{F}_{n,r}$

$$
\begin{aligned}
\|\frac{\partial \bar{F}_{n,r}}{\partial a}\|_q &\leq I(\sup_{m,R \geq r} |F_{m,R}| \leq \frac{1}{C_0})\big(2C_0 |F_{n,r}| \sup_{m,R \geq r} \|\frac{\partial F_{m,R}}{\partial a}\|_q + \|\frac{\partial F_{n,r}}{\partial a}\|_q\big) \\
&\leq 3 I(\sup_{m,R \geq r} |F_{m,R}| \leq \frac{1}{C_0}) \sup_{m,R \geq r} \|\frac{\partial F_{m,R}}{\partial a}\|_q,
\end{aligned}
$$

and thus by (66)

$$\|\frac{\partial \bar{F}_{n,r}}{\partial a}\|_q^2 \lesssim \frac{1}{r^\beta},$$

and hence by concentration of measure in form of (23) (applied to $cr^{\beta/2} \bar{F}_{n,r}$ by means of Lemma 1, c being a small universal constant)

$$\langle I(|\bar{F}_{n,r} - \langle \bar{F}_{n,r} \rangle| \geq M) \rangle \lesssim \exp(-\frac{1}{C} r^\beta M^2) \quad \text{for all } M \geq 0.$$

Together with (73) this yields

$$\langle I(|\bar{F}_{n,r}| \geq M) \rangle \lesssim \exp(-\frac{1}{C} r^\beta M^2) \quad \text{for } M \gg \exp(-\frac{1}{C} r^\beta) \text{ and } r \geq r_0.$$

By definition (71) we have $\langle I(|F_{n,r}| \geq M) \rangle \leq \langle I(\sup_{m,R \geq r} |F_{m,R}| \geq \frac{1}{2C_0}) \rangle + \langle I(|\bar{F}_{n,r}| \geq M) \rangle$ so that by (70) the above upgrades to

$$\langle I(|F_{n,r}| \geq M) \rangle \lesssim \exp(-\frac{1}{C} r^\beta M^2) \quad \text{for } 1 \geq M \gg \exp(-\frac{1}{C} r^\beta) \text{ and } r \geq r_0.$$

Since $r^\beta \exp(-\frac{1}{C}r^\beta) \lesssim 1$ for all r, the above holds without the lower restriction on M:

$$\langle I(|F_{n,r}| \geq M)\rangle \lesssim \exp(-\frac{1}{C}r^\beta M^2) \quad \text{for } M \leq 1 \text{ and } r \geq r_0. \qquad (74)$$

Using this estimate with r replaced by R and summing over the finite index set $n = 1, \cdots, N$ and all dyadic $R \geq r$ we obtain

$$\langle I(\sup_{m,R\geq r} |F_{m,R}| \geq M)\rangle \lesssim \exp(-\frac{1}{C}r^\beta M^2)$$
$$\text{for } M \leq 1 \text{ and } r \geq r_0 \qquad (75)$$

and thus in particular for the auxiliary random variable (67)

$$\langle I(\bar{F}_r \geq M)\rangle \lesssim \exp(-\frac{1}{C}r^\beta M^2) \quad \text{for all } M \geq 0 \text{ and } r \geq r_0,$$

where the upper bound on M is immaterial since $\bar{F}_r \leq \frac{1}{C_0} \leq 1$. Using $\langle \bar{F}_r \rangle = \int_0^\infty \langle I(\bar{F}_r \geq M)\rangle dM$, this yields the following quantification of $\lim_{r\uparrow\infty} \langle \bar{F}_r \rangle = 0$:

$$\langle \bar{F}_r \rangle \lesssim \int_0^\infty \exp(-r^\beta M^2) dM \lesssim r^{-\frac{\beta}{2}} \quad \text{for all } r \geq r_0.$$

Since r_0 was minimal in (69) and since $\langle \bar{F}_r \rangle$ depends continuously on r, this yields the desired

$$r_0 \lesssim 1. \qquad (76)$$

It remains to argue why (74), which together with (76) may be rephrased as

$$\langle I(|F_{n,r}| \geq M)\rangle \lesssim \exp(-\frac{1}{C}r^\beta M^2) \quad \text{for } M \leq 1 \text{ and } r \gg 1,$$

yields (15). It just suffices to include the given functional F from (14) into the list of finitely many functionals $F_1, \cdots, F_N$, say, as the last functional $F_N = F$, and then to specify the above to $n = N$. We note that for q related to β through (24) and p related to q through (34) one has $\frac{p}{p-1} = \frac{2d}{d+\beta}$, i.e. (14) entails (32). Note that by adjusting the constants, (15) is trivial for $r \lesssim 1$, so that we obtain (15) over the whole range $r \geq 0$.

Proof of Assertion ii).
The arguments in this section require $\beta < 2$, which in view of our assumption $\beta \ll 1$ is no restriction. Let r_* denote the minimal dyadic radius with the property (16); note that the proof to follow does *not* assume $r_* < \infty$. In order to establish (17), it is enough to show for a given dyadic $r_0 \geq 1$ that

$$\langle I(r_* > r_0)\rangle \lesssim \exp(-\frac{1}{C}r_0^\beta). \qquad (77)$$

It will be convenient to replace balls by cubes. Moreover, all radii or rather side length are dyadic. By definition of r_* as the smallest radius with (16), the event $r_* > r_0$ means that there exists a radius $R \ge r_0$ with

$$\frac{1}{R^2} \fint_{(-R,R)^d} |(\phi,\sigma) - \fint_{(-R,R)^d} (\phi,\sigma)|^2 dx > (\frac{r_0}{R})^\beta f(\frac{R}{r_0}), \qquad (78)$$

where

$$f(z) := \log(e + \log z).$$

In the sequel, the intermediate (dyadic) radius $r_1 \in [r_0, R]$ with

$$r_1 \sim r_0^{\frac{\beta}{2}} R^{1-\frac{\beta}{2}} f^{\frac{1}{2}}(\frac{R}{r_0}) \quad \text{so that} \quad (\frac{r_1}{R})^2 \sim (\frac{r_0}{R})^\beta f(\frac{R}{r_0}) \qquad (79)$$

will play a role. Note that we use here $\beta > 0$ and that $f(z)$ grows sub-algebraically. For the l. h. s. of (78) we note

$$\fint_{(-R,R)^d} |(\phi,\sigma) - \fint_{(-R,R)^d} (\phi,\sigma)|^2 dx$$
$$= \fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx + \sum_{\substack{r \in [2r_1, R] \\ r \text{ dyadic}}} \fint_{(-R,R)^d} |(\phi,\sigma)_{\frac{r}{2}} - (\phi,\sigma)_r|^2 dx,$$

$$(80)$$

where $(\phi,\sigma)_r$ denotes the $L^2((-R,R)^d)$-orthogonal projection of (ϕ,σ) onto the space of functions that are piecewise constant on the $(\frac{R}{r})^d$ dyadic sub-cubes Q of "level r" (that is, of side length $2r$) of the cube $(-R,R)^d$. In other words, on such a sub-cube Q, $(\phi,\sigma)_r = \fint_Q (\phi,\sigma) dx$. With this language, we may rewrite the first r. h. s. term of (80) as

$$\fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx = (\frac{r_1}{R})^d \sum_{Q \text{ level } r_1} \fint_Q |(\phi,\sigma) - \fint_Q (\phi,\sigma)|^2 dx,$$

so that by Poincaré's estimate on each of the cubes Q we obtain

$$\fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx \lesssim r_1^2 \fint_{(-R,R)^d} |\nabla(\phi,\sigma)|^2 dx,$$

and then by Caccioppoli's estimate based on (4) & (9), cf. (64),

$$\fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx$$
$$\lesssim (\frac{r_1}{R})^2 \fint_{(-2R,2R)^d} |(\phi,\sigma) - \fint_{(-2R,2R)^d} (\phi,\sigma)|^2 dx + r_1^2.$$

As a consequence of Lemma 4, there exist $N \sim 1$ linear functionals $\{F_n\}_{n=1,\cdots,N}$ whose rescaled versions $F_{n,r}$ satisfy the boundedness property (14) such that for any $r \geq 2R$ we have the implication

$$\sup_{r \geq 2R \text{ dyadic}} \max_{n=1,\cdots,N} (F_{n,r}\nabla(\phi,\sigma))^2 \ll 1$$

$$\implies \frac{1}{R^2} \fint_{(-2R,2R)^d} |(\phi,\sigma) - \fint_{(-2R,2R)^d}(\phi,\sigma)|^2 dx \lesssim 1.$$

From the two last statements we gather

$$\sup_{r \geq 2R \text{ dyadic}} \max_{n=1,\cdots,N} (F_{n,r}\nabla(\phi,\sigma))^2 \ll 1$$

$$\implies \frac{1}{R^2} \fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx \lesssim (\frac{r_1}{R})^2.$$

In view of (79) this can be rewritten as

$$\forall\, r \geq 2R \text{ dyadic, } n = 1, \cdots, N \quad (F_{n,r}\nabla(\phi,\sigma))^2 \ll 1$$

$$\implies \frac{1}{R^2} \fint_{(-R,R)^d} |(\phi,\sigma) - (\phi,\sigma)_{r_1}|^2 dx \leq \frac{1}{2}(\frac{r_0}{R})^\beta f(\frac{R}{r_0}), \qquad (81)$$

provided that we adjust the definition (79) of r_1 appropriately (to obtain the estimate $\leq \frac{1}{2}\cdot$ in (81) in place of just $\lesssim$). We now turn to the second r. h. s. term in (80), which in view of the definition of $(\phi,\sigma)_r$ we may estimate as follows

$$\sum_{\substack{r \in [2r_1,R] \\ r \text{ dyadic}}} \fint_{(-R,R)^d} |(\phi,\sigma)_{\frac{r}{2}} - (\phi,\sigma)_r|^2 dx$$

$$\leq \sum_{\substack{r \in [2r_1,R] \\ r \text{ dyadic}}} \max_{Q \text{ level } r} \max_{Q' \subset Q \text{ level } \frac{r}{2}} \left| \fint_{Q'}(\phi,\sigma) - \fint_Q(\phi,\sigma) \right|^2.$$

Hence if for any of the $(\frac{R}{r})^d$ dyadic sub-cubes Q of $(-R,R)^d$ of level r we introduce the $N = 2^d$ linear functionals $F_{Q,n}$ as an extension of

$$F_{Q,n}\nabla\zeta := \frac{1}{r}(\fint_{Q'_n} \zeta dx - \fint_Q \zeta dx),$$

where $\{Q'_n\}_{n=1,\cdots,2^d}$ is an enumeration of the sub-cubes of level $\frac{r}{2}$ of Q, and which satisfy the desired boundedness property (14) restricted to gradient

29

fields (which is no issue because of Hahn-Banach extension) and translated (which will be no issue because of stationarity), that is,

$$|F_{Q,n}\nabla\zeta| \lesssim \Big(\fint_Q |\nabla\zeta|^{\frac{2d}{d+\beta}} dx \Big)^{\frac{d+\beta}{2d}}, \qquad (82)$$

we have

$$\frac{1}{R^2} \sum_{\substack{r \in [2r_1, R] \\ r \text{ dyadic}}} \fint_{(-R,R)^d} |(\phi,\sigma)_{\frac{r}{2}} - (\phi,\sigma)_r|^2 dx$$

$$\leq \sum_{\substack{r \in [2r_1, R] \\ r \text{ dyadic}}} (\frac{r}{R})^2 \max_{Q \text{ level } r} \max_{n=1,\cdots,2^d} (F_{Q,n}\nabla(\phi,\sigma))^2.$$

From this we learn, since for the auxiliary function $g(z) := \log^{-2}(z+e)$, the dyadic sum $\sum_{r \in [2r_1, R]} g(\frac{R}{r})$ is universally bounded,

$$\forall \, r \in [2r_1, R] \text{ dyadic}, \; Q \text{ level } r, \; n = 1, \cdots, 2^d$$

$$(F_{Q,n}\nabla(\phi,\sigma))^2 \ll (\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})$$

$$\implies \frac{1}{R^2} \sum_{\substack{r \in [2r_1, R] \\ r \text{ dyadic}}} \fint_{(-R,R)^d} |(\phi,\sigma)_{\frac{r}{2}} - (\phi,\sigma)_r|^2 dx \leq \frac{1}{2}(\frac{r_0}{R})^\beta f(\frac{R}{r_0}).$$

In view of (80), the combination of this with (81) yields

$$\forall \, r \in [2r_1, R] \text{ dyadic}, \; Q \text{ level } r, \; n = 1, \cdots, 2^d$$

$$(F_{Q,n}\nabla(\phi,\sigma))^2 \ll (\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})$$

$$\text{and} \quad \forall \, r \geq 2R \text{ dyadic}, \; n = 1, \cdots, N \quad (F_{n,r}\nabla(\phi,\sigma))^2 \ll 1$$

$$\implies \frac{1}{R^2} \fint_{(-R,R)^d} |(\phi,\sigma) - \fint_{(-R,R)^d} (\phi,\sigma)|^2 dx \leq (\frac{r_0}{R})^\beta f(\frac{R}{r_0}). \qquad (83)$$

Equipped with this deterministic argument, we now may proceed to the stochastic part: In the event of $r_* > r_0$, there exists a dyadic $R \geq r_0$ such that (78) holds, so that we learn from (83) that there exists

- a (dyadic) $r \in [2r_1, R]$, a sub-cube Q of $(-R,R)^d$ of level r, and an index $n = 1, \cdots, 2^d$ such that $(F_{Q,n}\nabla(\phi,\sigma))^2 \gtrsim (\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})$. In view of the boundedness condition (82) and stationarity, we may apply (15) with F replaced by $F_{Q,n}$ and M^2 replaced by $(\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})$.

30

This M is admissible in the sense of $M \lesssim 1$ because by (79) we have $(\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0}) \le (\frac{R}{r_1})^2 (\frac{r_0}{R})^\beta f(\frac{R}{r_0}) \sim 1$. Hence the probability of each single of this events is estimated as follows

$$\left\langle I\left((F_{Q,n}\nabla(\phi,\sigma))^2 \ge \frac{1}{C}(\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})\right)\right\rangle$$
$$\lesssim \exp\left(-\frac{1}{C}(\frac{R}{r})^{2-\beta} g(\frac{R}{r}) f(\frac{R}{r_0}) r_0^\beta\right).$$

Since $g(z)$ decays sub-algebraically in z and since $\beta < 2$, this yields the simpler form

$$\left\langle I\left((F_{Q,n}\nabla(\phi,\sigma))^2 \ge \frac{1}{C}(\frac{R}{r})^2 g(\frac{R}{r})(\frac{r_0}{R})^\beta f(\frac{R}{r_0})\right)\right\rangle$$
$$\lesssim \exp\left(-\frac{1}{C}(\frac{R}{r})^{1-\frac{\beta}{2}} f(\frac{R}{r_0}) r_0^\beta\right).$$

- *or* a (dyadic) $r \ge 2R$ and an index $n = 1, \cdots, N$ for which the estimate $(F_{n,r}\nabla(\phi,\sigma))^2 \gtrsim 1$ holds. By the boundedness property of $F_{n,r}$, each single of these events is estimated as

$$\left\langle I\left((F_{n,r}\nabla(\phi,\sigma))^2 \ge \frac{1}{C}\right)\right\rangle \lesssim \exp(-\frac{1}{C}r^\beta).$$

Taking the number $(\frac{R}{r})^d$ of sub-cubes Q into account and recalling $N \lesssim 1$, this implies

$$\langle I(r_* > r_0)\rangle$$
$$\lesssim \sum_{\substack{R \ge r_0 \\ R \text{ dyadic}}} \left(\sum_{\substack{r \in [2r_0, R] \\ r \text{ dyadic}}} (\frac{R}{r})^d \exp(-\frac{1}{C}(\frac{R}{r})^{1-\frac{\beta}{2}} f(\frac{R}{r_0}) r_0^\beta) + \sum_{\substack{r \ge 2R \\ r \text{ dyadic}}} \exp(-\frac{1}{C}r^\beta)\right).$$

(84)

Again, since $1 - \frac{\beta}{2} > 0$, we have the calculus estimate

$$\sum_{\substack{r \in [2r_0, R] \\ r \text{ dyadic}}} (\frac{R}{r})^d \exp\left(-A(\frac{R}{r})^{1-\frac{\beta}{2}}\right)$$
$$\lesssim \exp(-A) \sum_{\substack{r \in [2r_0, R] \\ r \text{ dyadic}}} (\frac{R}{r})^d \exp\left(-A\log(\frac{R}{r})\right)$$
$$\lesssim \exp(-A) \quad \text{for } A \gg 1.$$

31

Applying this to the first sum over r in (84) and $A = \frac{1}{C} f(\frac{R}{r_0}) r_0^{\beta}$, which satisfies $A \gg 1$ for $r_0 \gg 1$, and using the estimate $\sum_{r \geq 2R; r \text{ dyadic}} \exp(-\frac{1}{C} r^{\beta}) \lesssim \exp(-\frac{1}{C} R^{\beta})$ (which holds provided that $R \geq r_0 \geq 1$) for the second sum over r, we obtain

$$\langle I(r_* > r_0) \rangle \; \lesssim \; \sum_{\substack{R \geq r_0 \\ R \text{ dyadic}}} \left(\exp(-\frac{1}{C} f(\frac{R}{r_0}) r_0^{\beta}) + \exp(-\frac{1}{C} R^{\beta}) \right) \quad \text{for } r_0 \gg 1.$$

Thanks to $\beta > 0$, we have $\exp(-\frac{1}{C} R^{\beta}) \lesssim \exp(-\frac{1}{C} f(\frac{R}{r_0}) r_0^{\beta})$, so that the second summand is dominated by the first one:

$$\langle I(r_* > r_0) \rangle \; \lesssim \; \sum_{\substack{R \geq r_0 \\ R \text{ dyadic}}} \exp(-\frac{1}{C} f(\frac{R}{r_0}) r_0^{\beta}) = \sum_{m=0}^{\infty} \exp(-\frac{1}{C} f(2^m) r_0^{\beta}).$$

Now we see the reason for the choice of $f(z) = \log(e + \log z)$ for which $f(2^m) \geq \frac{1}{C}(1 + \log(m+1))$ and thus

$$\sum_{m=0}^{\infty} \exp(-A f(2^m)) \leq \sum_{m=0}^{\infty} \exp(-\frac{A}{C}(1 + \log(m+1))) \lesssim \exp(-\frac{A}{C}) \quad \text{for } A \gg 1.$$

With $\frac{1}{C} r_0^{\beta}$ playing the role of A this yields (77). Note that the condition $r_0 \gg 1$ is immaterial after adjusting the constants, as the l. h. s. of (77) is bounded by 1. $\qquad\square$

A   Law of the iterated logarithm

We would now like to demonstrate that the scaling $r^{-\beta} \log \log r$ in the bound (16) – which is reminiscent of the law of the iterated logarithm for Brownian motion $\limsup_{t \to \infty} B_t / \sqrt{2t \log \log t} = 1$ almost surely – is optimal. To see this, let us consider in the perturbative setting of (18), (19) for some fixed $m \in \mathbb{N}$ the sequence of random variables

$$X_i := \frac{1}{(2^{2mi})^{1-\beta/2}} \left(\fint_{|x| \leq 2 \cdot 2^{2mi}} \phi_k \, dx - \fint_{|x| \leq 2^{2mi}} \phi_k \, dx \right).$$

Note that these random variables may be written in the form

$$X_i = 2^{2mi(\beta/2 - d)} \int \theta(2^{-2mi} |x|) \frac{x}{|x|} \cdot \nabla \phi_k \, dx$$

32

with $\theta = \theta(|x|)$ being independent of i and m.

By (20) and the discussion preceding (20), the X_i are therefore identically distributed centered Gaussian random variables with variance $\sim \delta^2$. One observes that their covariance satisfies

$$\langle X_i X_j \rangle \lesssim \delta^2 2^{-\beta m |i-j|}. \tag{85}$$

By a purely linear algebra argument (namely, the Gram-Schmidt algorithm, see below), one may construct a sequence of independent centered Gaussian random variables from the X_i:

Lemma 5. *If m is chosen large enough, then there exists a sequence of independent centered Gaussian random variables Y_i with the properties $\langle Y_i \rangle \sim \delta^2$ and*

$$Y_i = X_i + \sum_{j=1}^{i-1} a_{ij} X_j, \tag{86}$$

where $|a_{ij}| \lesssim 2^{-\beta m |i-j|/2}$.

Would the law of the iterated logarithm be non-optimal for the X_i, the same would hold for the Y_i: If we had with nonzero probability

$$\limsup_{i \to \infty} \frac{|X_i|}{\sqrt{\log \log 2^{2mi}}} = 0,$$

by the bound on the a_{ij} we would have

$$\limsup_{i \to \infty} \frac{|Y_i|}{\sqrt{\log \log 2^{2mi}}}$$

$$\leq \limsup_{i \to \infty} \frac{|X_i|}{\sqrt{\log \log 2^{2mi}}} + C \limsup_{i \to \infty} \frac{\sum_{k=1}^{i-1} 2^{-\beta mk/2} |X_{i-k}|}{\sqrt{\log \log 2^{2mi}}}$$

$$\leq 0 + C \sum_{k=1}^{N} 2^{-\beta mk/2} \limsup_{i \to \infty} \frac{|X_{i-k}|}{\sqrt{\log \log 2^{2mi}}} + C \sum_{k=N+1}^{\infty} 2^{-\beta mk/2} \sup_{j \in \mathbb{N}} \frac{|X_j|}{\sqrt{\log \log 2^{2m(j+k)}}}$$

$$\leq 0 + C \sum_{k=1}^{N} 2^{-\beta mk/2} \limsup_{i \to \infty} \frac{|X_i|}{\sqrt{\log \log 2^{2mi}}} + C \sum_{k=N+1}^{\infty} 2^{-\beta mk/2} \sup_{j \in \mathbb{N}} \frac{|X_j|}{\sqrt{\log \log 2^{2mj}}}$$

$$\leq 0 + 0 + C 2^{-\beta mN/2} \sup_{i \in \mathbb{N}} \frac{|X_i|}{\sqrt{\log \log 2^{2mi}}}$$

for all $N \in \mathbb{N}$ and therefore with the same nonzero probability

$$\limsup_{i\to\infty} \frac{|Y_i|}{\sqrt{\log\log 2^{2mi}}} = 0.$$

However, for a sequence of independent centered Gaussian random variables Y_i with $\sigma_i^2 := \langle Y_i^2 \rangle \sim \delta^2$, the law of the iterated logarithm in the simplified variant

$$\limsup_{i\to\infty} \frac{|Y_i|}{\sigma_i\sqrt{\log i}} \geq c \qquad \text{almost surely}$$

for some $c > 0$ is easily established: For each i, we have the lower bound

$$\mathbb{P}\big[|Y_i| > c\sigma_i\sqrt{\log i}\big] = 2\int_{c\sigma_i\sqrt{\log i}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{s^2}{2\sigma_i^2})\, ds$$

$$\gtrsim \frac{1}{c\sqrt{\log i}} \exp(-\frac{1}{2}c^2\log i).$$

As the sum of the right-hand sides gives ∞ for c small enough, the second part of the Borel-Cantelli lemma provides the conclusion.

Proof of Lemma 5. One may use the Gram-Schmidt algorithm to obtain the sequence of independent centered Gaussian random variables Y_i: One sets $Y_1 := X_1$ and

$$Y_i := X_i - \sum_{j=1}^{i-1} \frac{\langle X_i Y_j \rangle}{\langle Y_j^2 \rangle} Y_j, \tag{87}$$

which enforces $\langle Y_i Y_j \rangle = 0$ for $i \neq j$, provided that the algorithm does not terminate (which would happen in case $\langle Y_i^2 \rangle = 0$). To obtain a lower bound for $\langle Y_i^2 \rangle$, let us first estimate for $k > i$

$$|\langle Y_i X_k \rangle| \leq |\langle X_i X_k \rangle| - \sum_{j=1}^{i-1} \frac{|\langle X_i Y_j \rangle \langle X_k Y_j \rangle|}{\langle Y_j^2 \rangle},$$

which gives using $|j-i| + |j-k| = |i-k| + 2|j-i|$ (note that $k > i > j$) and the bound (85)

$$|\langle Y_i X_k \rangle| \lesssim \delta^2 2^{-\beta m|i-k|} + \delta^2 2^{-\beta m|i-k|} \max\left\{ \frac{\delta^2}{\langle Y_j^2 \rangle} : j \leq i-1 \right\}$$

$$\times \sum_{j=1}^{i-1} 2^{-2\beta m|i-j|} \frac{|\langle Y_j X_i \rangle|}{\delta^2 2^{-\beta m|j-i|}} \frac{|\langle Y_j X_k \rangle|}{\delta^2 2^{-\beta m|j-k|}}.$$

34

Let us denote

$$B_i := \sup_{k > i} \frac{|\langle Y_i X_k \rangle|}{\delta^2 2^{-\beta m |i-k|}}.$$

The previous estimate then entails

$$B_i \lesssim 1 + \max\left\{ \frac{\delta^2}{\langle Y_j^2 \rangle} : j \leq i - 1 \right\} \sum_{j=1}^{i-1} 2^{-2\beta m |i-j|} |B_j|^2.$$

If we can derive by induction a lower bound for $\langle Y_j^2 \rangle$ of the form $\langle Y_j^2 \rangle \gtrsim \delta^2$, for m large enough the previous bound entails by induction

$$\sup_i B_i \lesssim 1.$$

Note that we only require the lower bound on $\langle Y_j^2 \rangle$ for all $j \leq i - 1$ to obtain such an estimate for B_i.
To obtain the desired lower bound, let us compute

$$\langle Y_i^2 \rangle = \langle X_i^2 \rangle - \sum_{j=1}^{i-1} \frac{|\langle X_i Y_j \rangle|^2}{\langle Y_j^2 \rangle}.$$

As we have $\langle X_i^2 \rangle \sim \delta^2$, an estimate of the form

$$\sum_{j=1}^{i-1} \frac{|\langle X_i Y_j \rangle|^2}{\langle Y_j^2 \rangle} \ll \delta^2$$

would imply the desired lower bound for $\langle Y_i^2 \rangle$. Suppose that we have shown the lower bound $\langle Y_j^2 \rangle \gtrsim \delta^2$ for all $j \leq i-1$. To show the desired lower bound for $\langle Y_i^2 \rangle$, it is then sufficient to have

$$\sum_{j=1}^{i-1} 2^{-2\beta m |i-j|} |B_j|^2 \ll 1.$$

By our estimate on the B_j for $j \leq i - 1$, this is true for m large enough.
Recursively inserting the definition (87) into itself, we obtain the representation (86). To see the bound on the a_{ij}, one proceeds by induction in this iterative insertion; the desired bound is then obtained for m large enough, provided that we have for all i and all $j < i$ the bound $\frac{|\langle X_i Y_j \rangle|}{\langle Y_j^2 \rangle} \leq 2^{-2\beta m |i-j|/3}$
(a bound that we have already shown if m is large enough). $\qquad \square$

References

[1] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Preprint*, 2015. arXiv:1507.06935.

[2] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Preprint*, 2016. arXiv:1602.00512.

[3] S. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.

[4] S. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér*, 48:423–481, 2016.

[5] M. Avellaneda and F.H. Lin. Une théorème de liouville pour des équations elliptiques à coefficients périodiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 309:245–250, 1989.

[6] P. Bella and F. Otto. Corrector estimates for elliptic systems with random periodic coefficients. *Preprint*, 2014. arXiv:1409.5271.

[7] I. Benjamini, H. Duminil-Copin, G. Kozma, and A. Yadin. Disorder, entropy and harmonic functions. *Ann. Probab.*, 43(5):2332–2373, 2015.

[8] J. G. Conlon and A. Naddaf. Greens functions for elliptic and parabolic equations with random coefficients. *New York J. Math.*, 6:153–225, 2000.

[9] J. G. Conlon and T. Spencer. Strong convergence to the homogenized limit of elliptic equations with random coefficients. *Trans. Amer. Math. Soc.*, 366:1257–1288, 2014.

[10] J. Fischer and F. Otto. A higher-order large-scale regularity theory for random elliptic operators. *Comm. Partial Differential Equations*, 41(7):1108–1148, 2016.

[11] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. *Preprint*, 2014. arXiv:1409.2678.

[12] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.

[13] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 05 2011.

[14] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 02 2012.

[15] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *to appear in J. Eur. Math. Soc.*, 2014. arXiv:1409.0801.

[16] A. Gloria and F. Otto. Quantitative estimates on the periodic approximation of the corrector in stochastic homogenization. *ESAIM: Proc.*, 48:80–97, 2015.

[17] A. Gloria and F. Otto. The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. *Preprint*, 2015. arXiv:1510.08290.

[18] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, Berlin, 1994.

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

[20] M. Ledoux. *The Concentration of Measure Phenomenon*, volume 89 of *Mathematical Surveys and Monographs.* American Mathematical Society, 2001.

[21] D. Marahrens and F. Otto. Annealed estimates on the green function. *Probab. Theory Related Fields*, 163(3):527–573, 2014.

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

[23] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. *Unpublished preprint*, 1998.

[24] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random Fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloquia Mathematica Societatis Jnos Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[25] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27:167–180, 215, 1986.