ANNEALED ESTIMATES ON THE GREEN FUNCTIONS AND UNCERTAINTY QUANTIFICATION

ANTOINE GLORIA AND DANIEL MARAHRENS

Abstract. We prove optimal annealed decay estimates on the derivative and mixed second derivative of the elliptic Green functions on $\mathbb{R}^d$ for random stationary measurable coefficients that satisfy a certain logarithmic Sobolev inequality and for periodic coefficients, extending to the continuum setting results by Otto and the second author for discrete elliptic equations. As a main application we obtain optimal estimates on the fluctuations of solutions of linear elliptic PDEs with “noisy” diffusion coefficients, an uncertainty quantification result. As a direct corollary of the decay estimates we also prove that for these classes of coefficients the H"older exponent of the celebrated De Giorgi-Nash-Moser theory can be taken arbitrarily close to 1 in the large (that is, away from the singularity).

Contents

1. Introduction and statement of the main results .................................................. 2
2. Structure of the proofs and auxiliary results ...................................................... 8
3. Proofs of the annealed estimates ........................................................................ 12
   3.1. Proof of Theorem 1.3 ................................................................................. 12
   3.2. Proof of Corollary 1.4 .............................................................................. 16
   3.3. Proof of Corollary 1.5 .............................................................................. 16
4. Proofs of the fluctuation estimates ...................................................................... 16
   4.1. Proof of Theorem 1.6 ................................................................................. 16
   4.2. Proof of Theorem 1.7 ................................................................................. 20
5. Proofs of the improved De Giorgi-Nash-Moser theory .................................... 22
   5.1. Proof of Theorem 1.10 ............................................................................... 22
   5.2. Proof of Corollary 1.11 ............................................................................. 27
6. Proofs of the auxiliary results ............................................................................. 30
   6.1. Proof of Lemma 2.4 .................................................................................. 30
   6.2. Proof of Lemma 2.6 .................................................................................. 32
   6.3. Proof of Lemma 2.5 .................................................................................. 35
   6.4. Proof of Lemma 2.10 for $p < \infty$ ............................................................ 39
   6.5. Proof of Lemma 2.9 .................................................................................. 41
Acknowledgments .................................................................................................... 43
References ............................................................................................................... 43
1. Introduction and Statement of the Main Results

We let $\lambda \in (0, 1]$ denote an ellipticity constant which is fixed throughout the paper, and set

$$
\Omega_0 := \left\{ A_0 \in \mathbb{R}^{d \times d} : A_0 \text{ is bounded, i.e. } |A_0\xi| \leq |\xi| \text{ for all } \xi \in \mathbb{R}^d, \right\}
$$

(1.1)

$$
A_0 \text{ is elliptic, i.e. } \lambda |\xi|^2 \leq \xi \cdot A_0 \xi \text{ for all } \xi \in \mathbb{R}^d \}
$$

We equip $\Omega_0$ with the usual topology of $\mathbb{R}^{d \times d}$. A coefficient field, denoted by $A$, is a Lebesgue-measurable function on $\mathbb{R}^d$ taking values in $\Omega_0$. We then define

$$
\Omega := \{ \text{measurable maps } A : \mathbb{R}^d \to \Omega_0 \},
$$

which we equip with the $\sigma$-algebra $F$ that makes the evaluations $A \mapsto \int_{\mathbb{R}^d} A_{i,j}(x)\chi(x)dx$ measurable for all $i, j \in \{1, \ldots, d\}$ and all smooth functions $\chi$ with compact support. This makes $F$ countably generated.

Following the convention in statistical mechanics, we describe a random coefficient field by equipping $(\Omega, F)$ with an ensemble $\langle \cdot \rangle$ (the expected value). Following [19], we shall assume that $\langle \cdot \rangle$ is stochastically continuous: For all $\delta > 0$ and $x \in \mathbb{R}^d$,

$$
\lim_{|h| \downarrow 0} \left\langle 1_{\{A : |A(x+h)-A(x)| > \delta\}} \right\rangle = 0
$$

We shall always assume that $\langle \cdot \rangle$ is stationary, i.e. for all translations $z \in \mathbb{R}^d$ the coefficient fields $\{\mathbb{R}^d \ni x \mapsto A(x)\}$ and $\{\mathbb{R}^d \ni x \mapsto A(x+z)\}$ have the same joint distribution under $\langle \cdot \rangle$. Let $\tau_z : \Omega \to \Omega, A(\cdot) \mapsto A(\cdot + z)$ denote the shift by $z$, then $\langle \cdot \rangle$ is stationary if and only if $\tau_z$ is $\langle \cdot \rangle$-preserving for all shifts $z \in \mathbb{R}^d$. The stochastic continuity assumption ensures that the map $\mathbb{R}^d \times \Omega \to \Omega, (x, \omega) \mapsto \tau_x \omega$ is measurable (where $\mathbb{R}^d$ is equipped with the $\sigma$-algebra of Lebesgue measurable sets).

A random variable is a measurable function on $(\Omega, F)$. A random field $\zeta$ is a measurable function on $\mathbb{R}^d \times \Omega$. In this article the random field under study is the Green function. We are interested in the behaviour of the (massive) Green function $G_\mu : \mathbb{R}^d \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, which is defined for all $\mu > 0$ and for all $y \in \mathbb{R}^d$ as the unique distributional solution in $W^{1,1}(\mathbb{R}^d)$ which is continuous away from the diagonal $x = y$ of the elliptic equation

$$
\mu G_\mu(x, y; A) - \nabla_x \cdot (A(x) \nabla_x G_\mu(x, y; A)) = \delta(x-y).
$$

(1.2)

For the existence, uniqueness and properties of $G_\mu$, see Definition 2.1. Note that by definition of the $\sigma$-algebra, $G_\mu$ is measurable.

We shall consider two classes of ensembles: periodic ensembles (in which case continuum stationarity is achieved by randomizing the origin) and ensembles satisfying a logarithmic Sobolev inequality.

**Definition 1.1** (Logarithmic Sobolev inequality (LSI)). We say that the ensemble $\langle \cdot \rangle$ satisfies a logarithmic Sobolev inequality if there exist constants $\rho, \ell > 0$, which we shall respectively call amplitude and correlation-length, such that

$$
\left\langle \zeta^2 \log \left(\frac{\zeta^2}{\langle \zeta^2 \rangle}\right) \right\rangle \leq \frac{2}{\rho} \left\langle \int_{\mathbb{R}^d} \left(\text{osc}_{A|_{B_\ell(y)}} \zeta \right)^2 dy \right\rangle
$$

(1.3)

for all measurable functions $\zeta : \Omega \to \mathbb{R}$, where the expectation in the RHS is an outer expectation (the oscillation is not necessarily measurable). Here the expression $\text{osc}_{A|_{B_\ell(y)}} \zeta$ denotes the oscillation of $\zeta$ with respect to all coefficient fields that coincide with $A$ outside of $B_\ell(y)$, where $B_\ell(y)$ is the
ball of radius \( \ell \) centered at \( y \in \mathbb{R}^d \), that is,

\[
\left( \text{osc}_{A|B_\ell} \zeta \right) (A) = \left( \sup_{A|B_\ell} \zeta \right) (A) - \left( \inf_{A|B_\ell} \zeta \right) (A)
\]

\[
= \sup \left\{ \zeta(\tilde{A}) | \tilde{A} \in \Omega, \tilde{A}|_{\mathbb{R}^d \setminus B_\ell} = A|_{\mathbb{R}^d \setminus B_\ell} \right\}
\]

\[
- \inf \left\{ \zeta(\tilde{A}) | \tilde{A} \in \Omega, \tilde{A}|_{\mathbb{R}^d \setminus B_\ell} = A|_{\mathbb{R}^d \setminus B_\ell} \right\}.
\]

\[
(1.4)
\]

An example of coefficient field which satisfies (LSI) is the Poisson inclusions process (and variants of it), see in particular [1], [8]. Without loss of generality, we assume in this article that \( \ell \geq 1 \).

**Remark 1.2.** The fact that outer expectations appear in the RHS of (1.3) is not a difficulty since in the rest of the article we shall always estimate the RHS of (1.3) by the expectation of measurable quantities (for which outer expectation and expectation coincide).

The first main result of this article are optimal annealed estimates on the elliptic Green function. In general, the only optimal decay result which holds without further smoothness assumption is the following consequence of the celebrated De Giorgi-Nash-Moser theory (in dimensions \( d > 2 \)) on the Green function itself: For all \( A \in \Omega \), and all \( \mu \geq 0 \),

\[
0 \leq G_\mu(x, y; A) \leq C_\mu e^{-c_\mu \sqrt{|x - y|}}
\]

for some constants \( c, C > 0 \) depending only on \( \lambda \) and \( d \), see Definition 2.1 below. For the constant-coefficient operator, i.e. the massive Laplacian, we also have the following optimal gradient estimate: For all \( \mu \geq 0 \),

\[
|\nabla G_\mu(x, y; \text{Id})| \leq C_\mu e^{-c_\mu \sqrt{|x - y|}}
\]

\[
(1.5)
\]

For variable-coefficients, the only generic bound which holds for the gradient of the elliptic Green function is another consequence of the De Giorgi-Nash-Moser theory: There exists \( 0 < \alpha \leq 1 \) depending only on \( \lambda \) and \( d \) (with \( \alpha \uparrow 1 \) as \( \lambda \uparrow 1 \)) such that for all \( x, y \in \mathbb{R}^d \)

\[
(1.6)
\]

if \( |x - y| \gtrsim 1 \), then

\[
\int_{B_1(x)} |\nabla_s G_\mu(x, y; A)| dx \leq C_\mu e^{-c_\mu \sqrt{|x - y|}}
\]

\[
|\nabla G_\mu(x, y; A)| dx \leq C_\mu e^{-c_\mu \sqrt{|x - y|}}
\]

\[
(1.5)
\]

see Lemma 2.6 below. As can be seen, there is a mismatch between the generic behavior and the heat kernel at the level of the gradient. The behavior at the singularity \( x = y \) only can be described for smooth coefficients (say, uniformly H"older-continuous). In that case, the optimal scaling of (1.5) holds for \( |x - y| \lesssim 1 \), cf. [13] Theorem 3.3 for \( \mu = 0 \). However, even for analytic coefficients, the heat-kernel bound (1.5) cannot hold generically in the large, that is in the regime \( |x - y| \uparrow \infty \), for this would contradict the counterexamples from quasiconformal mappings. These counterexamples are however very specific in the sense that they cannot be translation-invariant. Indeed, for the class of periodic uniformly H"older-continuous coefficients, Avellaneda and Lin already proved in [2] that (1.5) holds (for \( \mu = 0 \)), by exploiting the periodic structure through the introduction of harmonic coordinates. The result by Avellaneda and Lin contains again two regularity results: a regularity result in the small, that is (1.5) for \( |x - y| \lesssim 1 \), and a regularity result in the large, that is (1.5) for \( |x - y| \gtrsim 1 \). Although it does not appear explicitly in their arguments, the regularity in the large, in the form of (1.6) for \( \alpha = 1 \), is the sole consequence of periodicity. Indeed this is an elementary corollary of a result by Delmotte and Deuschel [7] on stationary coefficients, which is also at the basis of our results on random coefficients (under a slightly improved form obtained by Otto and the second author in [17], see also [12] for \( \mu > 0 \)). We shall show that for measurable
coefficients distributed according to an ensemble that satisfies (LSI), (1.10) holds almost surely for any \( \alpha > 0 \) (with however \( C(\alpha, d) \uparrow \infty \) as \( \alpha \uparrow 1 \)).

For all \( L > 0 \) and all \( |x - y| \geq 3L \) set

\[
(1.7) \quad (\nabla G_\mu)_L(x, y) := \left( \int_{B_L(x)} |\nabla G_\mu(x', y)|^2 \, dx' \right)^{\frac{1}{2}}
\]

\[
(1.8) \quad (\nabla \nabla G_\mu)_L(x, y) := \left( \int_{B_L(x)} \int_{B_L(y)} |\nabla \nabla G_\mu(x', y')|^2 \, dy' \, dx' \right)^{\frac{1}{2}},
\]

where (here and in the whole article) \( \nabla \nabla \) stands for the second mixed derivative \( \nabla x', \nabla y' \). The main result of this article is as follows.

**Theorem 1.3.** If the ensemble is stationary and satisfies (LSI) with correlation-length \( \ell \), then for all \( \mu \geq 0 \), the functions \( (\nabla G_\mu)_\ell \) and \( (\nabla \nabla G_\mu)_\ell \) decay optimally in all \( \langle \cdot \rangle \)-moments, i.e. we have

\[
(1.9) \quad \langle |(\nabla G_\mu)_\ell(x, 0)|^{2p} \rangle^{\frac{1}{p}} \leq C \frac{e^{-c \sqrt{\mu} |x|}}{|x|^{d-1}},
\]

\[
(1.10) \quad \langle |(\nabla \nabla G_\mu)_\ell(x, 0)|^{2p} \rangle^{\frac{1}{p}} \leq C \frac{e^{-c \sqrt{\mu} |x|}}{|x|^{d}},
\]

for all \( p < \infty \), and all \( x \in \mathbb{R}^d \) such that \( |x| \geq 3\ell \), where the constant \( C \) depends only on \( \lambda, p, \rho, \ell \).

If the ensemble is \( \ell \)-periodic, then

\[
(1.11) \quad |(\nabla G_\mu)_\ell(x, 0)| \leq C \frac{e^{-c \sqrt{\mu} |x|}}{|x|^{d-1}},
\]

\[
(1.12) \quad |(\nabla \nabla G_\mu)_\ell(x, 0)| \leq C \frac{e^{-c \sqrt{\mu} |x|}}{|x|^{d}},
\]

where the constant only depends on \( d \) and \( \lambda \).

Note that stationarity implies \( \langle |(\nabla G_\mu)_\ell(x, y)|^p \rangle = \langle |(\nabla G_\mu)_\ell(x - y, 0)|^p \rangle \) for all \( y \in \mathbb{R}^d \), so that the above result implies the annealed decay of \( (\nabla G_\mu)_\ell(x, y) \) for arbitrary \( x, y \in \mathbb{R}^d \).

This result is based on and extends the annealed estimates by Delmotte and Deuschel \([2]\), see Proposition 2.3 below. It is the extended continuum version of the result by Otto and the second author in \([13]\) for discrete elliptic equations. As an elementary corollary, in the random case these optimal annealed estimates yield almost optimal quenched estimates:

**Corollary 1.4.** If the ensemble is stationary and satisfies (LSI) with constant \( p \) and correlation-length \( \ell \), for all \( \beta > 0 \) there exists a random variable \( \mathcal{Y}_\beta \) such that for all \( |x - y| \geq 3\ell \) almost surely

\[
(\nabla G_\mu)_\ell(x, y) \leq \mathcal{Y}_\beta \frac{e^{-c \sqrt{\mu} |y - x|}}{|y - x|^{d-\gamma}}, \quad (\nabla \nabla G_\mu)_\ell(x, y) \leq \mathcal{Y}_\beta \frac{e^{-c \sqrt{\mu} |y - x|}}{|y - x|^{d-\gamma}},
\]

where the random variable \( \mathcal{Y}_\beta \) has bounded finite moments: for all \( p \geq 1 \), \( \langle \mathcal{Y}_\beta^p \rangle < \infty \).

Denote by \( C^\gamma(\mathbb{R}^d) \) the set of uniformly \( \gamma \)-Hölder continuous functions with Hölder semi-norm \( [.]_{C^\gamma} \). If in addition the coefficients are uniformly Hölder continuous (as considered by Avellaneda and Lin for the periodic case in \([2]\)), the averaged bounds of Theorem 1.3 hold pointwise.

**Corollary 1.5.** If the ensemble is stationary and satisfies (LSI) with constant \( p \) and correlation-length \( \ell \) and is supported on coefficient fields \( [A]_{C^\gamma(\mathbb{R}^d)} \leq C_\gamma \) for some \( \gamma > 0 \) and constant \( C_\gamma < \infty \),
then for all $\mu \geq 0$, the Green function satisfies
\begin{align}
(1.13) \quad & \langle |\nabla G_\mu(x,0)|^{2p} \rangle \leq C \frac{e^{-c\sqrt{|x|}}}{|x|^{d-1}}, \\
(1.14) \quad & \langle |\nabla \nabla G_\mu(x,0)|^{2p} \rangle \leq C \frac{e^{-c\sqrt{|x|}}}{|x|^d},
\end{align}
where the constant only depends on $d, \lambda, p, \rho, \ell, \gamma$.

If the ensemble is $\ell$-periodic and is supported on coefficient fields $[A]_{C^\gamma(\mathbb{R}^d)} \leq C_\gamma$ for some $\gamma > 0$ and constant $C_\gamma < \infty$, then for all $\mu \geq 0$, the Green function satisfies for all $x, y \in \mathbb{R}^d$
\begin{align}
(1.15) \quad & |\nabla G_\mu(x,y)| \leq C \frac{e^{-c\sqrt{|x-y|}}}{|y-x|^{d-1}}, \\
(1.16) \quad & |\nabla \nabla G_\mu(x,y)| \leq C \frac{e^{-c\sqrt{|x-y|}}}{|y-x|^d}.
\end{align}

As a first main application of the annealed estimates, we quantify the fluctuations of solutions of linear elliptic equations with “noisy” diffusion coefficients (a quantification of the propagation of uncertainty in elliptic PDEs). We consider diffusion coefficients $A_\varepsilon$ on $\mathbb{R}^d$ of the form
\begin{equation}
A_\varepsilon(x) := \text{Id} + B\left(\frac{x}{\varepsilon}\right)
\end{equation}
where $B$ is a random perturbation which has order 1, range of correlation unity (which we shall replace in the theorem by the (LSI) assumption), and vanishing expectation. Hence, $A_\varepsilon$ is a perturbation of the identity by some noise of range of correlation $\varepsilon$. Let $f$ be some RHS, and consider the random solution $u_\varepsilon$ of
\begin{equation}
u_\varepsilon - \nabla \cdot A_\varepsilon \nabla u_\varepsilon = f \quad \text{in} \, \mathbb{R}^d.
\end{equation}
The question we are interested in is the characterization of the fluctuations of $u_\varepsilon$ in function of $\varepsilon$ and of the statistics of $B$, first in terms of scaling and second in terms of law. In this contribution we address the question of the scaling wrt $\varepsilon$, and give optimal estimates of both weak and strong measures of the fluctuation, which generalize the bounds obtained for $B$ small (that is, in the regime of small ellipticity ratio) by the first author in [11].

The natural norms which control these fluctuations are mixed norms $L^p_{\lambda,\varepsilon}(\mathbb{R}^d)$ which measure local fluctuations at scale $\varepsilon$ in $L^p$ but large scale fluctuations in $L^q$. In particular, for all $q, \lambda \geq 1$, $\varepsilon > 0$ and $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ we set
\begin{equation}
||f||_{L^p_{\lambda,\varepsilon}(\mathbb{R}^d)} := \left( \int_{\mathbb{R}^d} \left( \int_{B_\varepsilon(z)} |f(y)|^q \, dy \right)^{\frac{p}{q}} \, dx \right)^{\frac{1}{p}}.
\end{equation}

In particular it is bounded by the $L^q(\mathbb{R}^d)$-norm for $q \geq \lambda$ by Jensen’s inequality. We start with the estimate of the fluctuations in a strong norm.

**Theorem 1.6.** Let $A_\varepsilon = A(\varepsilon)$ be the $\varepsilon$-rescaling of the coefficient field $A \in \Omega$ distributed according to a stationary ensemble $\langle \cdot \rangle$ that satisfies (LSI). Let $\mu \geq 0$. For all $\varepsilon > 0$, let $u_\varepsilon \in H^1(\mathbb{R}^d)$ be a distributional solution of
\begin{equation}
\mu u_\varepsilon - \nabla \cdot A_\varepsilon \nabla u_\varepsilon = f \quad \text{in} \, \mathbb{R}^d.
\end{equation}

\footnote{Note that the proof of [11] Lemma 2.1] is wrong under the general assumption of finite correlation-length. The assumption of [11] Theorem 3] should be replaced by “Assume that the stationary random field $B$ satisfies spectral gap”, as it is the case for Poisson inclusions for instance. The optimal form of [11] Theorem 3] is given by Theorems 1.9 and 1.10 below — the norms in [11] Theorem 3] have to be adapted accordingly.}
Then for all \( \lambda > \frac{d}{2} \), \( 1 \leq \theta < \infty \), \( 2 \leq p < \infty \), \( 1 \leq r \leq \frac{d}{\theta - 1} \), and \( q \) such that
\[
(1.19) \quad 1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{q},
\]
the fluctuations of \( u_\varepsilon \) satisfy
\[
\left\langle \left( \int_{\mathbb{R}^d} |u_\varepsilon - \langle u_\varepsilon \rangle|^p \, dx \right)^{\frac{1}{p}} \right\rangle \lesssim \left\{ \begin{array}{ll}
  d = 2 : & \ln(\mu \varepsilon^2)^{\frac{1}{d}} + 1 \\
  d > 2 : & 1
\end{array} \right\} \varepsilon \left( \mu - \frac{1}{\varepsilon^2} - \frac{1}{\varepsilon^d} + 1 \right) \|f\|_{L^\theta_{\lambda_1}(\mathbb{R}^d)}
\]
where \( \|f\|_{L^\theta_{\lambda_1}(\mathbb{R}^d)} \) is given by (1.17). In the border-line case \( r = \frac{d}{\theta - 1} \), we require in addition \( q > 1 \).

We then turn to the estimate of weak norm of the fluctuations.

**Theorem 1.7.** Let \( A_\varepsilon = A(\frac{\cdot}{\varepsilon}) \) be the \( \varepsilon \)-rescaling of the coefficient field \( A \in \Omega \) distributed according to a stationary ensemble \( \langle \cdot \rangle \) that satisfies (LSI). Let \( \mu > 0 \). For all \( \varepsilon > 0 \), let \( u_\varepsilon \in H^1(\mathbb{R}^d) \) be a distributional solution of (1.18). Then for all \( 1 \leq \theta < \infty \), \( 2 \leq p < \infty \), \( 1 \leq r, \tilde{r} \leq \frac{d}{\theta - 1} \), \( 1 \leq q \leq \frac{r}{\tilde{r} - 1} \), and \( 1 \leq \tilde{q} \leq \frac{1}{\theta - 1} \) such that
\[
(1.20) \quad 2 + \frac{1}{2} = \frac{1}{r} + \frac{1}{\tilde{r}} + \frac{1}{q} + \frac{1}{\tilde{q}}
\]
and for all \( \lambda_1, \lambda_2 \geq 1 \) such that
\[
(1.21) \quad \frac{1}{\lambda_1} + \frac{1}{\lambda_2} < \frac{d + 2}{d},
\]
the fluctuations of \( u_\varepsilon \) satisfy for all \( g \in L^1_{\text{loc}}(\mathbb{R}^d) \),
\[
\left\langle \left( \int_{\mathbb{R}^d} (u_\varepsilon - \langle u_\varepsilon \rangle) g \, dx \right)^{\frac{1}{d}} \right\rangle \lesssim \frac{\varepsilon}{d} \left( \mu^{-1} - \frac{1}{\varepsilon} - \frac{1}{\varepsilon^d} + 1 \right) \|g\|_{L^\theta_{\lambda_2}(\mathbb{R}^d)}
\]
In the border-line case \( r = \tilde{r} = \frac{d}{\theta - 1} \), we require in addition \( q, \tilde{q} > 1 \).

**Remark 1.8.** When the coefficients \( A \) in Theorems 1.6 and 1.7 are uniformly Hölder continuous, then we can replace the mixed norms \( L^\theta_{\lambda_2}(\mathbb{R}^d) \) by the usual norms \( L^q(\mathbb{R}^d) \) in the estimates. This shows that one can trade local integrability of \( f \) and \( g \) for regularity of \( A \). This is proved by replacing averaged bounds on the Green function by pointwise bounds, as in [16]. We leave the details to the reader.

**Remark 1.9.** Theorem [1.6] reveals the central limit scaling of the weak measure of the fluctuations. While the most natural norms for the RHS on \( \mathbb{R}^d \) are those which make the estimate independent of \( \mu \), the other estimates are valuable for \( \mu > 0 \) since the massive term essentially localizes the equation to a bounded domain of size \( \mu^{-\frac{1}{d}} \) (without boundary layers).

These results generalize both [11, Theorem 3] and [16, Corollaries 2 & 3] (cf. also [9] by Conlon and Naddaf in the case of discrete elliptic equations). Note that when the noise is in the zero-order term (that is, for \( \mu \) replaced by \( 1 + b_\varepsilon \) and \( A_\varepsilon \) by \( \text{Id} \) in (1.18)), the CLT scaling (and in addition the characterization of the limiting law) was established by Figari, Orlandi and Papanicolaou in [8] for \( d \geq 4 \) and by Bal in [3] for \( d \leq 3 \). The arguments involved in the proof of Theorems 1.6 and 1.7 have a different flavor since the randomness is in the derivative of highest order.

As a second main application of the annealed estimates we obtain an improved De Giorgi-Nash-Moser theory \textit{in the large} for the two classes of coefficients considered. One way to formulate the De Giorgi-Nash-Moser theory is as follows: There exists \( 0 < \alpha \leq 1 \) depending only on the ellipticity ratio \( \lambda \) such that for all \( p > \frac{d}{2} \), \( \kappa > 0 \), \( R > 0 \), and \( \mu \geq 0 \) with \( R^2 \mu \leq \kappa \), if \( u \) satisfies
\[
\mu u - \nabla \cdot A \nabla u = f \quad \text{in} \quad B_{2R},
\]
for some \( f \in L^p(B_{2R}) \), then

\[
R_α^μ \sup_{x, y \in B_R} \frac{|u(x) - u(y)|}{|x - y|^α} \lesssim \left( \int_{B_{2R}} u^2 \right)^{\frac{1}{2}} \left( \int_{B_{2R}} |R^2 f|^p \right)^{\frac{1}{p}},
\]

see for instance [10 Theorem 8.24]. (Note that this follows from the statement for \( R = 1 \) since by (1.23), \( f \) is replaced by \( R^2 f \) when performing a change of variables \( x \sim R^{-1} x_0 \).) In the supremum above, we have set by convention \( \frac{R}{R} := 0 \). This result has two aspects: a regularity in the small and a regularity in the large. In particular we may split the statement into two parts: in the small, that is for \( |x| \lesssim 1 \), (1.22) quantifies the high frequencies of \( u \) (local regularity),

\[
\sup_{B_1} \frac{|u(x) - u(0)|}{|x|^α} \lesssim \left( \int_{B_{2R}} u^2 \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} |f|^p \right)^{\frac{1}{p}},
\]

and in the large, (1.22) quantifies the low frequencies of \( u \) (growth at large scales),

\[
\sup_{B_R \setminus B_1} \frac{|u(x) - u(0)|}{|x|^α} \lesssim R^{-α} \left( \int_{B_{2R}} u^2 \right)^{\frac{1}{2}} + R^{-α} \left( \int_{B_{2R}} |R^2 f|^p \right)^{\frac{1}{p}}.
\]

If we assume that the coefficients \( A \) are uniformly Hölder-continuous, then we have an optimal regularity theory in the small, that is, (1.22) holds for the improved exponent \( α = 1 \) provided \( p > d \) (see for instance [15 Theorem 3.13]). However, the De Giorgi-Nash-Moser exponent cannot be improved in the large by increasing the regularity of the coefficients, as classical examples from quasiconformal mappings show. The improvement of the De Giorgi-Nash-Moser exponent in the large is the aim of the following result for stationary coefficients that satisfy \( (LSI) \) and for periodic coefficients.

**Theorem 1.10.** If the ensemble is stationary and satisfies \( (LSI) \) with constants \( ρ \) and \( ℓ \) (or if it is \( ℓ \)-periodic), then for all \( μ \geq 0 \) and \( d < p < \infty \), we have for all \( R \geq 2ℓ \) and all \( x \in B_R \setminus B_{2ℓ} \), and all \( 1 \leq q < \infty \),

\[
\left( \sup_{(u, f)} \left( \int_{B_1} \frac{|u(x) - u(x')|}{|x|} \frac{dx'}{|x|} \right)^q \right)^{\frac{1}{q}} \lesssim C,
\]

for some constant \( C < \infty \) depending only on \( d, λ, p, q, ρ, ℓ \), where the supremum is taken over all couples \( (u, f) \) satisfying \( f \in L^p(B_{2R}) \) and related via

\[
μu - ∇ · A∇u = f \text{ in } B_{2R}.
\]

In the random case, at the cost of a slightly smaller decay rate, this result holds quenched: For all \( 0 < β \leq 1 \) and \( d < p < \infty \), there exists a random variable \( Y_{β, p} \) with bounded finite moments such that for all \( μ \geq 0 \), if \( u \) and \( f \in L^p(B_{2R}) \) are related via (1.23), we have for all \( R \geq 2ℓ \) and all \( x \in B_R \setminus B_{2ℓ} \),

\[
R^{1-β} \int_{B_1} \frac{|u(x) - u(x')|}{|x|} dx' \leq Y_{β, p} \left( \left( \int_{B_{2R}} u^2 dx' \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} |R^2 f|^p dx' \right)^{\frac{1}{p}} \right).
\]

For uniformly Hölder continuous coefficients we also have:

**Corollary 1.11.** If the ensemble is stationary, satisfies \( (LSI) \) with constants \( ρ \) and \( ℓ \) (or if it is \( ℓ \)-periodic), and there exists a constant \( C_γ < \infty \) such that \( \langle \cdot \rangle \)-almost surely \( |A|_{C_γ} \leq C_γ \), then for
all $\mu \geq 0$, $d < p < \infty$, and $1 \leq q < \infty$, we have for all $R > 0$:

$$
(1.28) \quad \left\langle \left( \sup_{(u,f)} \frac{R |u(x) - u(0)|}{|x|} \right)^{\frac{1}{2}} + \left( \frac{\int_{B_{2R}} |R^2 f|^p dx}{|x|^\frac{p}{q}} \right)^{\frac{1}{p}} \right\rangle \leq C
$$

for some constant $C < \infty$ depending only on $d, \lambda, p, q, C_+, \rho, \ell$, where the supremum $\sup_{(u,f)}$ is taken over all couples $(u, f)$ satisfying $f \in L^p(B_{2R})$ and related via $(1.26)$.

In addition, for all $0 \leq \alpha < 1$, and $d < p < \infty$, there exists a random variable $Y_{\alpha,p}$ with bounded finite moments such that for all $\mu \geq 0$, if $u$ and $f \in L^p(B_{2R})$ are related via $(1.26)$, we have for all $R \geq 2\ell$ and all $x \in B_R$:

$$
(1.29) \quad R^\alpha \int_{B_{\ell}} \frac{|u(x + x') - u(x')|}{|x'|^\alpha} dx' \leq Y_{\alpha,p} \left( \left( \int_{B_{2R}} u^2 dx \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} |R^2 f|^p dx \right)^{\frac{1}{p}} \right).
$$

Let us conclude this introduction by mentioning an independent work by Armstrong and Smart. In [1], they obtain an improved version of $(1.27)$ in Theorem 1.10 for $\mu = 0$ which allows to treat the border case $\beta = 0$ (and better integrability of $Y$). Their analysis is based on convex duality theory and quantifies the proof by Dal Maso and Modica [6] at the level of the subadditive ergodic theorem. In particular, it covers Euler-Lagrange equations associated with the minimization of (non-necessarily quadratic) convex functionals under weaker assumptions on the probability measure. By rather standard arguments, this quenched H"{o}lder regularity estimate should allow to recover Theorem 1.3 for $\mu = 0$. However, the approach they develop in [1] is currently limited to variational problems and does not cover the case of nonsymmetric coefficients treated here. The estimates of the fluctuations in Theorems 1.6 and 1.7 are of independent interest and require the validity of a spectral gap estimate (which follows here from the associated logarithmic-Sobolev inequality).

We shall only prove the results in the random case. In the periodic case, the proofs are simpler since the results are quenched. Note that the method used in this article would also yield similar results for suitable subclasses of quasi-periodic coefficients.

## 2. Structure of the proofs and auxiliary results

We start with the definition and main properties of the elliptic Green function.

**Definition 2.1** (Green’s function). For all $A \in \Omega$ and every $\mu > 0$, there exists a unique function $G_\mu(x, y; A) \geq 0$ with the following properties

- Qualitative continuity off the diagonal, that is,

$$
(2.1) \quad \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d | x \neq y\} \ni (x, y) \mapsto G_\mu(x, y; A) \quad \text{is continuous.}
$$

- Upper pointwise bounds on $G_\mu$:

$$
(2.2) \quad G_\mu(x, y; A) \lesssim e^{-c \sqrt{|x - y|}} \begin{cases} 
\ln \left( 2 + \frac{1}{\sqrt{|x - y|}} \right) & \text{for } d = 2 \\
\ln \left( \frac{1}{|x - y|^{d-2}} \right) & \text{for } d > 2
\end{cases},
$$

where here and in the sequel the rate constant $c > 0$ in the exponential is generic and may change from term to term, but only depends on $d$ and $\lambda$. 

• Averaged bounds on $\nabla_x G_\mu$ and $\nabla_y G_\mu$:

\[
\left( R^{-d} \int_{R^d} |\nabla_x G_\mu(x, y; A)|^2 \, dx \right)^{\frac{1}{2}} \lesssim e^{-c\sqrt{R}} R^{1-d},
\]

(2.3)

\[
\left( R^{-d} \int_{R^d} |\nabla_y G_\mu(x, y; A)|^2 \, dy \right)^{\frac{1}{2}} \lesssim e^{-c\sqrt{R}} R^{1-d}.
\]

(2.4)

• Differential equation: We note that (2.2) and (2.3) & (2.4) imply that the maps $R^d \ni x \mapsto (G_\mu(x, y; A), \nabla_x G_\mu(x, y; A))$ and $R^d \ni y \mapsto (G_\mu(x, y; A), \nabla_y G_\mu(x, y; A))$ are (locally) integrable. Hence even for discontinuous $A$, we may formulate the requirement

\[
\mu G_\mu - \nabla_x \cdot A(x) \nabla_x G_\mu = \delta(x - y) \quad \text{distributionally in } R^d,
\]

(2.5)

\[
\mu G_\mu - \nabla_y \cdot A^*(y) \nabla_y G_\mu = \delta(y - x) \quad \text{distributionally in } R^d,
\]

(2.6)

where $A^*$ denotes the transpose of $A$.

We note that the uniqueness statement implies $G_\mu(x, y; A^*) = G_\mu(y, x; A)$ so that $G_\mu$ is symmetric when $A$ is symmetric.

These standard properties of the massive Green functions are proved in [12] (essentially following arguments of [14]).

**Remark 2.2.** All the main results of this article are stated for $\mu > 0$, whereas we shall only consider the case $\mu = 0$ in the proofs. Indeed, one can pass to the limit as $\mu \downarrow 0$ in all our estimates, and local averages of $\nabla G_\mu$ and $\nabla \nabla G_\mu$ converge to local averages of $\nabla G$ and $\nabla \nabla G$, where $G$ is the Green function for $\mu = 0$ (the existence of which is subtle for $d = 2$).

The improvement of the De Giorgi-Nash-Moser theory in the large is a consequence of the annealed bounds on the Green function of Theorem 1.3. As in the discrete case dealt with in [16] the strategy is to upgrade to any moment in probability the optimal bounds by Delmotte and Deuschel [7] on the first and second moments of $\nabla \nabla G_\mu$ and $\nabla G_\mu$, respectively. Yet, the bounds by Delmotte and Deuschel in [7, Theorem 1.2] are not enough at the level of the mixed second gradient, and we shall use the following result of [17] in its version with the massive term proved in [12] Lemma 2.11:

**Proposition 2.3.** If the ensemble is stationary, then the Green function satisfies for all $\mu > 0$, all $L \gtrsim 1$, and all $x \in R^d$ with $|x| \geq 2L$,

\[
\langle (\nabla_x G_\mu)_L(x, 0)^2 \rangle \lesssim C \frac{e^{c\sqrt{R}|x|}}{|x|^{d-1}},
\]

(2.7)

\[
\langle (\nabla \nabla G_\mu)_L(x, 0) \rangle \lesssim C \frac{e^{c\sqrt{R}|x|}}{|x|^{d}},
\]

(2.8)

for some constants $C$ and $c$ depending only on $\lambda$ and $d \geq 2$.

In the periodic case, Theorem 1.3 is a direct corollary of Proposition 2.3. In the random case, estimate (1.11) of Theorem 1.3 is a consequence of (2.8) and of the following reverse Hölder estimate valid for all $p > 1$ large enough:

\[
\sup_{x, y: |x - y| \geq \rho L} \left\{ |x - y|^d e^{c\sqrt{R}|x - y|} (\langle (\nabla \nabla G_\mu)_L(x, y) \rangle^{2p})^{\frac{1}{2p}} \right\} \lesssim C(d, \lambda, p, \rho, \ell) \sup_{x, y: |x - y| \geq \rho L} \left\{ |x - y|^d e^{c\sqrt{R}|x - y|} (\langle (\nabla \nabla G_\mu)_L(x, y) \rangle) \right\},
\]

and likewise for the first derivative. This gain of integrability is achieved by the following lemma in the spirit of [16] Lemma 4, where the assumption that $\langle \cdot \rangle$ satisfies (LSI) is crucial.
Lemma 2.4. Let \((\cdot)\) satisfy (LSI) with constants \(\rho, \ell > 0\). Then for arbitrary \(\delta > 0\) and \(1 \leq p < \infty\) and for any random variable \(\zeta\) we have

\[
\langle |\zeta|^{2p} \rangle^\frac{1}{2p} \leq C(d, \rho, p, \delta) \langle \langle |\zeta| \rangle \rangle^\frac{1}{2p} \left( \int_{\mathbb{R}^d} \left( \frac{\text{osc}}{A|_{B_L(z)}} G_{\mu} \right)^2 dz \right)^\frac{1}{2p}
\]

for some finite constant \(C(d, \rho, p, \delta)\), where we recall that the expectation in the RHS is an outer expectation.

Since \(G_{\mu}\) is measurable on \(\Omega\), one may apply this lemma to \(\zeta = \langle \nabla G_{\mu} \rangle_{\ell}(x, 0)\) and \(\zeta = \langle \nabla_x G_{\mu} \rangle_{\ell}(x, 0)\). In order to prove the reverse H"older inequality \((2.10)\), it suffices to absorb the second RHS term of \((2.10)\) in the RHS. This is the content of the following lemma, which is essentially based on deterministic arguments.

Lemma 2.5 (Absorption lemma). Let \(d \geq 2\). There exists \(p_0 > 1\) depending only on \(\lambda\) and \(d\) such that for all \(L \sim 1\) and \(p \geq p_0\), we have for the second derivative:

\[
\sup_{|x-y| \geq 6L} \left\{ |x-y|^{2p(d-1)} e^{2pc/\sqrt{\ell}|x-y|} \left( \int_{\mathbb{R}^d} \left( \frac{\text{osc}}{A|_{B_L(z)}} (\nabla^2 G_{\mu})_{L}(x,y) \right)^2 dz \right)^{\frac{1}{2p}} \right\} \leq \sup_{|x-y| \geq 6L} \left\{ |x-y|^{2p(d-1)} e^{2pc/\sqrt{\ell}|x-y|} \left( \left( \int_{\mathbb{R}^d} \left( \frac{\text{osc}}{A|_{B_L(z)}} (\nabla G_{\mu})_{L}(x,y) \right)^2 dz \right)^{\frac{1}{2}} \right)^{2p} \right\} + 1,
\]

and for the first derivative:

\[
\sup_{|x-y| \geq 6L} \left\{ |x-y|^{2p(d-1)} e^{2pc/\sqrt{\ell}|x-y|} \left( \int_{\mathbb{R}^d} \left( \frac{\text{osc}}{A|_{B_L(z)}} (\nabla_x G_{\mu})_{L}(x,y) \right)^2 dz \right)^{\frac{1}{2p}} \right\} \leq \sup_{|x-y| \geq 6L} \left\{ |x-y|^{2p(d-1)} e^{2pc/\sqrt{\ell}|x-y|} \left( \left( \int_{\mathbb{R}^d} \left( \frac{\text{osc}}{A|_{B_L(z)}} (\nabla_x G_{\mu})_{L}(x,y) \right)^2 dz \right)^{\frac{1}{2}} \right)^{2p} \right\} + 1,
\]

where \(\lesssim\) stands for \(\leq\) up to a multiplicative constant which depends on \(d, \lambda\), and \(p\).

A key ingredient for the proof of Lemma 2.5 are the following quenched estimates.

Lemma 2.6. Let \(d \geq 2\). There exist \(q_0 > 1\) and \(a_0 > 0\) depending only on \(d\) and \(\lambda > 0\) such that for all \(\mu > 0\), \(1 \leq q \leq q_0\), and all \(R \sim 4L \sim 1\),

\[
\int_{|x-y| < 2R} |\nabla_x G_{\mu}(x,y)|^{2q} \, dx \lesssim R^{d+(1-q)2e^{-c\sqrt{\lambda}R}},
\]

\[
\int_{|x-y| < 2R} |\nabla^2 G_{\mu}(x,y)|^{2q} \, dy \, dx \lesssim R^{-2q\alpha_0 e^{-c\sqrt{\lambda}R}},
\]

where the multiplicative constants depend only on \(d\) and \(\lambda\). In addition we have the following local boundedness estimate for all \(L \sim 1\)

\[
\sup_{x,y \in \mathbb{R}^d, |x-y| < 6L} \left\{ (\nabla^2 G_{\mu})_{L}(x,y) \right\} \lesssim 1.
\]

Remark 2.7. (i) Our results beg the question if we can upgrade \((2.9)\) and \((2.10)\) to a stronger version without space integrals as in \((2.9)\) and \((2.8)\). The answer is negative if \(p > 1\). Let us consider \((2.7)\) (in the parabolic setting, \((2.8)\) directly follows from \((2.7)\)). Using the De Giorgi-Nash-Moser
theory, we may upgrade \((1.9)\) to pointwise-estimates away from the singularity if \(p = 1\), but not otherwise. Indeed, the De Giorgi-Nash-Moser theory yields away from the singularity that
\[
\left\langle \int_{B_L(x)} \int_{B_L(y)} |\nabla_1 G(x', y')|^2 \, dy' \, dx' \right\rangle \gtrsim \left\langle \int_{B_L(x)} |\nabla_1 G(x, y)|^2 \, dx \right\rangle.
\]
Now by stationarity, the left hand side equals
\[
\left\langle \int_{B_L(0)} |\nabla_1 G(x + x', y')|^2 \, dx' \right\rangle = \left\langle \int_{B_L(0)} |\nabla_1 G(x - y, -x')|^2 \, dx' \right\rangle \gtrsim \left\langle |\nabla_x G(0, y - x)|^2 \right\rangle,
\]
where the last inequality again follows from de Giorgi-Nash-Moser theory. On the other hand, if \(p > 1\), pointwise bounds on \(\|\nabla G\|^{2p}\) cannot be expected since there is no local regularity to control \(\langle |\nabla G|^{2p} \rangle\). On the other hand, clearly energy methods allow to control locally the \(L^2\)-norms in \((1.9)\) and \((1.10)\) are necessary to smooth out local effects when the coefficients lack regularity if and only if \(p > 1\).

(ii) In a similar spirit, we observe that the restriction \(|x| \gtrsim L\) is not necessary in \((7)\), but cannot be avoided here. Indeed, assuming Proposition \((2.3)\) only for \(|x| \gtrsim 1\), we may remove this restriction by a simple scaling argument. The same is true if we (could) replace \((\nabla \nabla G)_L\) by \(\nabla \nabla G\) as discussed in (i). On the other hand, the presence of the averaging operation \((\cdot)_L\) breaks the scaling invariance by introducing a length scale \(L\). Therefore we cannot expect to obtain information on the blow-up of \((\nabla \nabla G)_L(x, y)\) as the singularity enters the integral, i.e. as \(|x - y| \downarrow 2L\). 

We turn now to the fluctuation estimates. By a scaling argument, it is enough to prove Theorems \((1.6)\) and \((1.7)\) for \(\varepsilon = 1\) and \(\ell = \frac{1}{2}\). We thus consider the solution \(u \in H^1(\mathbb{R}^d)\) of
\[
\mu u - \nabla \cdot A \nabla u = f, \quad \mu \geq 0.
\]
We shall only consider the case \(\mu > 0\) in the proofs. The results for \(\mu = 0\) are then obtained by letting \(\mu \downarrow 0\) in the estimates. The starting point is the following spectral gap estimate

**Lemma 2.8 \((q\text{-}\text{SG})\).** If \(\langle \cdot \rangle\) satisfies \((\text{LSI})\) with amplitude \(\rho > 0\) and correlation-length \(\ell < \infty\), then we have for all \(q \geq 1\) and all random variables \(\zeta\)
\[
\left\langle (\zeta - \langle \zeta \rangle)^2 \right\rangle^q \lesssim \left\langle \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A |_{B_L(z)}} \zeta \right)^2 \, dz \right)^q \right\rangle^{\frac{1}{q}},
\]
with \(\ell = 2\ell\), where the multiplicative constant depends on \(q\) and \(\rho\). 

This is a standard result. It is indeed enough to assume that \(\langle \zeta \rangle = 0\) and \(\langle \zeta^2 \rangle = 1\). To prove estimate \((2.17)\) for \(q = 1\) it suffices to apply \((\text{LSI})\) to the random variable \(\chi = \sqrt{1 - \alpha^2} + \alpha \zeta\) and make a Taylor expansion as \(\alpha \downarrow 0\), this yields the result for the correlation-length \(\ell\). The estimate for \(q > 1\) is a consequence of the estimate for \(q = 1\) (up to increasing \(\ell\) to \(\ell = 2\ell\), see for instance [12] Corollary 2.3]. Since we have assumed that \(\ell = \frac{1}{2}\), \((2.3)\) holds for \(\ell = 1\).

The following lemma is a sensitivity estimate which quantifies how much the solution \(u\) of \((2.10)\) depends on the coefficients \(A\).

**Lemma 2.9.** Let \(\lambda_1, \lambda_2 \in [1, +\infty]\) satisfy
\[
\frac{1}{\lambda_1} + \frac{1}{\lambda_2} < \frac{d + 2}{d} \quad \iff \quad \frac{1}{\lambda_1} + \frac{1}{\lambda_2} > \frac{d - 2}{d}.
\]
Lemma 2.10. Let $K_{\mu}(x, y) = \sup_{\Omega} (|\nabla G_{\mu}(x, y)|^2 + (\nabla G_{\mu})^2_{L^2}(x, y))^{1/2}$, where $K_{\mu}(x, y)$ follows from the Nash-Aronson bounds (if $\mu \geq 0$, Young’s inequality and the well-known estimate
\[ \|u\|_{L^p(B_\ell)} \lesssim \|u\|_{L^2(B_\ell)} + \|f\|_{L^p(B_\ell)}, \]
where the multiplicative constant depends on $\mu$, $d$, and $q$, but not on $\mu \geq 0$.

This result is usually stated for $p = \infty$ only, cf. [10, Theorem 8.17]. Although we think it should follow from the Nash-Aronson bounds (if $d > 2$), Young’s inequality and the well-known estimate with $p = +\infty$, we display a direct proof for $p < \infty$ using a (simplified) Moser-type iteration that works for $d = 2$ and uses less machinery.

3. Proofs of the annealed estimates

3.1. Proof of Theorem 1.13

The proof is a simple combination of Proposition 2.3, Lemma 2.4 and Lemma 2.5.

Step 1. Proof of 1.10.

We apply 2.10 of Lemma 2.4 to $\zeta(A) = (\nabla G_{\mu})_{L^p}(x, y)$ for some $x, y \in \mathbb{R}^d$ such that $|x - y| \geq \ell$ to the effect of
\[ \langle (\nabla G_{\mu})_{L^p}(x, y) \rangle \]
\[ \lesssim C(d, \ell, p, \ell, \rho, \delta) \left\langle \left( \int_{\mathbb{R}^d} (\text{osc} A|_{B_{\ell}(x, y)})^2 dz \right) \right\rangle \]
Combined with 2.3 in Proposition 2.3 this yields
\[ |x - y|^d e^{c\sqrt{d|x - y|}} \left( \langle (\nabla G_{\mu})_{L^p}(x, y) \rangle \right) \lesssim C(d, \lambda, \ell, p, \delta) \]
\[ + \delta |x - y|^d e^{c\sqrt{d|x - y|}} \left( \int_{\mathbb{R}^d} (\text{osc} A|_{B_{\ell}(x, y)})^2 dz \right) \]

We take the supremum over all $x$ and $y$ such that $|x - y| \geq 6\ell$ and insert (2.11) in Lemma 2.5 to obtain that
\[
\sup_{|x - y| \geq 6\ell} \left\{ |x - y|^{d-1} e^{c\sqrt{|x - y|}} \langle (\nabla \nabla G_\mu)\ell(x, y)\rangle^{2p} \right\}^{\frac{1}{2p}} \leq C(d, \lambda, \rho, \ell, p, \delta) + C(d, \lambda, p, \ell, \delta)\sup_{|x - y| \geq 6\ell} \left\{ |x - y|^{d-1} e^{c\sqrt{|x - y|}} \langle (\nabla \nabla G_\mu)\ell(x, y)\rangle^{2p} \right\}^{\frac{1}{2p}} + 1.
\]

Choosing $\delta$ small enough, we may absorb the last RHS term in the LHS. This yields (1.10).

Step 2. Proof of (1.9).

We proceed as in Step 1: Take $\zeta(A) = (\nabla \nabla G_\mu)\ell(A; x, y)$ in Lemma 2.4 to deduce
\[
\langle (\nabla \nabla G_\mu)\ell(x, y)\rangle^{2p} \leq C(d, \rho, \ell, p, \delta) + \delta \left\{ \left( \int_{\mathbb{R}^d} \left( \text{osc} (\nabla \nabla G_\mu)\ell(x, y) \right)^2 \, dz \right) \right\}^{\frac{1}{p}}.
\]

Combined with (2.7) in Proposition 2.2, this turns into
\[
|x - y|^{d-1} e^{c\sqrt{|x - y|}} \langle (\nabla \nabla G_\mu)\ell(x, y)\rangle^{2p} \leq C(d, \lambda, \rho, \ell, p, \delta) + \delta |x - y|^{d-1} e^{c\sqrt{|x - y|}} \left\{ \left( \int_{\mathbb{R}^d} \left( \text{osc} (\nabla \nabla G_\mu)\ell(x, y) \right)^2 \, dz \right) \right\}^{\frac{1}{p}}.
\]

After taking the supremum over all $x, y$ such that $|x - y| \geq 6\ell$, the estimate (2.12) from Lemma 2.5 yields
\[
\sup_{|x - y| \geq 6\ell} \left\{ |x - y|^{d-1} e^{c\sqrt{|x - y|}} \left( \int_{\mathbb{R}^d} \left( \text{osc} (\nabla \nabla G_\mu)\ell(x, y) \right)^2 \, dz \right)^\frac{1}{p} \right\}^{\frac{1}{2p}} \leq C(d, \lambda, \rho, \ell) \left( 1 + \delta \right) \sup_{|x - y| \geq 6\ell} \left\{ |x - y|^{2p(d-1)} e^{2pc\sqrt{|x - y|}} \left( \left| (\nabla \nabla G_\mu)\ell(x, y) \right|^{2p} \right) \right\} + C(d, \lambda, p, \ell) \delta \sup_{|x - y| \geq 6\ell} \left\{ |x - y|^{2pd} e^{2pc\sqrt{|x - y|}} \left( \left| (\nabla \nabla G_\mu)\ell(x, y) \right|^{2p} \right) \right\} + C(d, \lambda, p, \ell, \gamma, \beta).
\]

By (1.10) (proved in Step 1), the last term is bounded by a constant $C(d, \lambda, \rho, \ell, p, \delta)$. We then conclude by taking $\delta$ small enough so that we can absorb the remaining supremum on the LHS.

3.2. Proof of Corollary 1.4

For every $x \in \mathbb{R}^d$, there exists some $x' \in \frac{\ell}{2^k} \mathbb{Z}^d$ such that the difference $x - x'$ has max-norm $|x - x'|_\infty \leq \frac{\ell}{2^k}$. Hence its Euclidean norm satisfies $|x - x'| \leq \frac{\ell}{2^k}$. Consequently, we have that $\left| (\nabla \nabla G_\mu)\ell(x', 0) \right| \geq \left| (\nabla \nabla G_\mu)\ell(x, 0) \right|$ and it holds that
\[
\left( \sup_{x \in \mathbb{R}^d \setminus B_{4\ell}} \left\{ |x|^{d-\beta} \left| (\nabla \nabla G_\mu)\ell(x, 0) \right| \right\} \right)^p \leq \left( \sup_{x' \in \left( \frac{\ell}{2^k} \mathbb{Z}^d \right)^d, |x'| \geq 4\ell} \left\{ |x'|^{-\beta} \left| (\nabla \nabla G_\mu)\ell(x', 0) \right| \right\} \right)^p \leq \left( \sup_{x' \in \left( \frac{\ell}{2^k} \mathbb{Z}^d \right)^d, |x'| \geq 4\ell} \left\{ |x'|^{-\beta} \left| (\nabla \nabla G_\mu)\ell(x', 0) \right| \right\} \right)^p \leq C(d, \lambda, \rho, \ell, \gamma, \beta)
\]
as long as $\beta p > d$, which we may assume without loss of generality since by Jensen’s inequality we may always increase $p$. The same remark applies to $(\nabla G_\mu)\ell$. The choice
\[
\mathcal{Y}_\beta := \max \left\{ \sup_{x \in \mathbb{R}^d \setminus B_{4\ell}} \left\{ |x|^{d-\beta} \left| (\nabla \nabla G_\mu)\ell(x, 0) \right| \right\}, \sup_{x \in \mathbb{R}^d \setminus B_{4\ell}} \left\{ |x|^{d-1-\beta} \left| (\nabla \nabla G_\mu)\ell(x, 0) \right| \right\} \right\}
\]
concludes the proof.
3.3. Proof of Corollary 1.5

We split the proof into two steps.

Step 1. Near-field estimates.

The results of [14, Theorem 3.3] yield

\[ |\nabla G_\mu(x,0)| \lesssim |x|^{1-d} \quad \text{and} \quad |\nabla \nabla G_\mu(x,0)| \lesssim |x|^{-d} \]

for all $|x| \leq 3\ell$. (The fact that $G_\mu$ does not vanish on $\partial B_{3\ell}$ can be dealt with by substracting the corresponding boundary value problem, which is clearly bounded by the classical Schauder estimates and the Nash-Aronson $L^\infty$-estimate on $G_\mu$ away from the origin. The arguments are uniform wrt $\mu \geq 0$. The estimate for $d=2$ can be deduced from the corresponding estimate for $d=3$ by using the elegant argument by Avellaneda and Lin [2], see for instance Step 2 of the proof of Lemma 2.6 below.)

Step 2. Far-field estimates.

It remains to treat the $|x| \geq 3\ell$. Let $u$ be a $(\mu - \nabla \cdot A\nabla)$-harmonic function in $\mathbb{R}^d \setminus B_\ell$. Our goal is to prove the following reverse Hölder inequality

\[ |\nabla u(x)|^2 \lesssim \int_{B_\ell(x)} |\nabla u(x')|^2 \, dx' + \mu \int_{B_\ell(x)} |u(x)| \, dx, \]

with a constant depending on $\ell, d, \lambda$, and $\gamma$ only. Without the derivative, this is a consequence of the De Giorgi-Nash-Moser theory. Since we are interested in $\nabla u$, we require the Hölder-continuity of the coefficient field. In the following, we will nonetheless pursue a strategy similar to Moser iteration to achieve the desired bound in (3.1). Since $A$ is Hölder-continuous, the function $u$ satisfies $u \in C^{2,\gamma}(\mathbb{R}^d \setminus \overline{B}_\ell)$ by interior Schauder theory. Now consider some length $0 < L \leq \frac{\ell}{2}$, and denote by $\overline{\nabla u}$ the average of $u$ on $B_L(x)$. Let $\eta \in C^\infty_0(B_L(x))$. By assumption, we have that $\eta(\mu u - \nabla \cdot A\nabla u) = 0$ in $\mathbb{R}^d$. Fix some $y' \in B_L(x)$. The product rule yields

\[ \mu(\eta(u - \overline{\nabla u}))(y) - \nabla \cdot A(y')\nabla (\eta(u - \overline{\nabla u}))(y) = -\mu\overline{\nabla u}\eta(y) + \nabla \cdot ((A(y) - A(y'))\eta(y)\nabla (u - \overline{\nabla u}))(y) \]

\[ - \nabla \cdot ((u - \overline{\nabla u})(y)A(y')\nabla \eta(y)) - \nabla \eta(y) \cdot A(y)\nabla (u - \overline{\nabla u})(y), \]

for all $y \in \mathbb{R}^d$. This is a constant-coefficient elliptic equation in $y$ with a right hand side in $H^{-1}(\mathbb{R}^d)$ and associated Green function $G_0(\cdot) \equiv G_\mu(\cdot,0;A(y'))$. The Green function representation yields for all $x' \in B_{2\ell}(x)$

\[ (\eta(u - \overline{\nabla u}))(x') = \int_{\mathbb{R}^d} \left( \nabla G_0(x' - y) \cdot (A(y') - A(y))\nabla (u - \overline{\nabla u})(y) 
\right. \]

\[ + (u - \overline{\nabla u})(y)\nabla G_0(x' - y) \cdot A(y)\nabla \eta(y) 
\left. + G_0(x' - y)(\nabla \eta(y) \cdot A(y)\nabla (u - \overline{\nabla u})(y) - \mu\overline{\nabla u}\eta(y)) \right) \, dy. \]

This can be made rigorous by mollification of the RHS of (3.2). Indeed, since $u \in C^{2,\gamma}(\overline{B_\ell(x)})$, the limit exists and is given by (3.3). Assume now that $\eta$ is a cutoff function for $B_{2\ell}(x)$ in $B_L(x)$ such that $|\nabla \eta| \lesssim \frac{1}{\ell^2}$. We may also take the gradient in (3.3) w. r. t. $x'$ at the point $y' \in B_{2\ell}(x)$ to obtain

\[ \nabla u(y') = \int_{\mathbb{R}^d} \left( \nabla \nabla G_0(x' - y) \cdot (A(y') - A(y))\nabla (u - \overline{\nabla u})(y) 
\right. \]

\[ + (u - \overline{\nabla u})(y)\nabla G_0(x' - y) \cdot A(y)\nabla \eta(y) 
\left. + \nabla G_0(x' - y)(\nabla \eta(y) \cdot A(y)\nabla (u - \overline{\nabla u})(y) - \mu\overline{\nabla u}\eta(y)) \right) \, dy. \]
As above, this can be justified by mollification of the RHS of (3.2). Indeed, the limit is well-defined since the constant-coefficient Green function $G_0$ classically satisfies
\begin{align*}
|\nabla G_0(y)| &= |\nabla G_{\mu}(y, 0; A(y'))(y)| 
&\leq C(d, \lambda)|y|^{1-d}
|\nabla \nabla G_0(y)| &= |\nabla \nabla G_{\mu}(y, 0; A(y'))| 
&\leq C(d, \lambda)|y|^{-d}
\end{align*}
uniformly in $y, y' \in \mathbb{R}^d$, while by assumption, the coefficient field satisfies $|A(y') - A(y)| \leq C_\gamma |y' - y|$. It then follows
\begin{equation}
|\nabla u(y')| \lesssim \rho|\varpi_L| + \int_{B_L(x)} |\nabla G_0(y' - y)||A(y') - A(y)||\nabla u(y)| \, dy
+ L^{-1} \int_{A_{2L}(x)} \left( |u(y) - \varpi_L||\nabla G_0(y' - y)| + |\nabla G_0(y' - y)||\nabla u(y)| \right) \, dy
\end{equation}
for all $y' \in B_L(x)$, where $A_{L'}(x) := \{ y : L' \leq |y - x| \leq L'' \}$ denotes the annulus centered at $x$ and of radii $L'$ and $L''$. Since $L \sim 1$, we allow the constant in $\lesssim$ to depend on $L$. The constant-coefficient bounds yield
\begin{equation}
\int_{A_{2L}(x)} |u(y) - \varpi_L||\nabla G_0(y' - y)| \, dy \lesssim \int_{B_L(x)} |u(y) - \varpi_L| \, dy.
\end{equation}
Combined with Jensen’s and Poincaré’s inequalities, this turns into
\begin{equation}
\int_{A_{2L}(x)} |u(y) - \varpi_L||\nabla G_0(y' - y)| \, dy \lesssim \left( \int_{B_L(x)} |\nabla u(y)|^2 \, dy \right)^{\frac{1}{2}}.
\end{equation}
Likewise we obtain
\begin{equation}
\int_{A_{2L}(x)} |\nabla G_0(y' - y)||\nabla u(y)| \, dy \lesssim \left( \int_{B_L(x)} |\nabla u(y)|^2 \, dy \right)^{\frac{1}{2}}.
\end{equation}
We are left with the second RHS term of (3.4), which we bound, by the decay of $\nabla \nabla G_0$ and the Hölder continuity of $A$, by
\begin{equation}
\int_{B_L(x)} |\nabla \nabla G_0(y' - y)||A(y') - A(y)||\nabla u(y)| \, dy \lesssim \int_{B_L(x)} |x' - y|^{-d-r} |\nabla u(y)| \, dy.
\end{equation}
Let $p \geq 2$. We then take the $p$-th power of (3.4), use (3.5)–(3.7), and integrate over $y'$ in $B_L(x)$. This yields
\begin{equation}
\int_{B_L(x)} |\nabla u(y')|^p \, dy' \lesssim (\rho|\varpi_L|)^p + \int_{B_L(x)} \left( \int_{B_L(x)} |y' - y|^{-d-r}|\nabla u(y)| \, dy \right)^p \, dy' + \left( \int_{B_L(x)} |\nabla u(y)|^2 \, dy \right)^{\frac{p}{2}}.
\end{equation}
We are almost in position to apply Young’s convolution inequality. Mimicking its proof, we let $r$ and $p'$ be such that
\begin{equation}
p \geq p' \geq 1, \ p > r \geq 1, \ 1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{p'},
\end{equation}
and use Hölder’s inequality with exponents $(\frac{p'}{p-r}, \frac{r}{p-r})$ on the integrand
\begin{equation}
|y' - y|^{-d-r}|\nabla u(y)| = \left( |y' - y|^{(\gamma-d)\frac{p'}{2}} |\nabla u(y)| \frac{r}{r'} \right) \left( |\nabla u(y)| \frac{p'}{p} \right) \left( |y' - y|^{(\gamma-d)\frac{p'}{2}} \right).
\end{equation}
In addition, we consider exponents and ball size: 

\[ p_\theta \]

We start from 

\[ L(x) \]

As long as we choose \( 1 \leq r < \frac{d}{2} \gamma < 2 \) (since \( \gamma < 1 \) and \( d \geq 2 \)), the last RHS term is bounded (depending on \( L \)). Let us fix such an \( 1 < r < 2 \leq p \), in which case the exponents \( (p, r, p', r') = \frac{pr}{r+(r-1)p} \) satisfy (3.10). Integrating (3.10) over \( y' \in B_\frac{r}{2} \) then yields

\[
\int_{B_{\frac{r}{2}}(x)} \left( \int_{B_{\frac{r}{2}}(x)} |y' - y|^{\gamma - d} |\nabla u(y')| \, dy' \right)^p \, dy 
\]

Combined with (3.8), this gives for all \( p \geq 2 \), \( 1 < r < 2 \), and \( p' = \frac{pr}{r+(r-1)p} \),

\[
\|u\|_{L^p(B_{\frac{r}{2}}(x))} \lesssim \sqrt{\mu} \|u\|_{L^\infty(B_{\frac{r}{2}}(x))} + \|\nabla u\|_{L^p(B_{\frac{r}{2}}(x))} + \|\nabla u\|_{L^2(B_{\frac{r}{2}}(x))}.
\]

We start from \( p_0' = 2 \) (that is, with \( p_0 = \frac{2r}{r+(r-1)} > 2 \)) and iterate using the following exponents and ball size:

\[
p_{n+1}' := p_n, \quad p_{n+1} := \frac{p_n r}{r - (r-1)p_n}, \quad L_{n+1} = \frac{L_n}{2}.
\]

So defined, \( p_n \) is a monotonically increasing sequence, so that \( (p_n, r, p_n') \) satisfies (3.9) for all \( n \in \mathbb{N}_0 \) such that \( p_n < \frac{r}{2} \). In particular, (3.11) then yields

\[
\|\nabla u\|_{L^{p_n}(B_{\frac{r}{2}}(x))} \lesssim \sqrt{\mu} \|u\|_{L^\infty(B_{\frac{r}{2}}(x))} + \|\nabla u\|_{L^{p_n}(B_{\frac{r}{2}}(x))}.
\]

In addition, \( p_n \) satisfies \( p_n \geq \left( \frac{r}{r - (r-1)p_n} \right)^{n+1} \), so that after finitely many steps, \( p_n \) is such that \( \frac{p_n r}{r - (r-1)p_n} > \frac{r}{2} \), at which point we may choose \( p_{n+1} = \infty \). This proves (3.11), and Corollary 1.5 now follows directly from Theorem 1.3 with \( L = \ell \), noting that the quenched estimate on the Green function itself yields

\[
0 \leq \mu G_\mu(x, y) \lesssim \mu \ln(2 + \frac{1}{\sqrt{|x - y|}} \frac{e^{-c\sqrt{\mu}|x - y|}}{|x - y|^{d-1}} \lesssim \frac{e^{-c\sqrt{\mu}|x - y|}}{|x - y|^{d-1}}),
\]

for any \( 0 < c' < c \), as desired.

4. PROOFS OF THE FLUCTUATION ESTIMATES

4.1. Proof of Theorem 1.6 We first assume that the coefficient field \( A \) and the right-hand side \( f \) are smooth. Since the estimates do not depend on the smoothness of the parameters, we may at the end lift this restriction by approximation. The triangle inequality yields

\[
\left( \int_{\mathbb{R}^d} |u(x) - \langle u(x) \rangle|^p \, dx \right)^\frac{1}{p} \leq \left( \int_{\mathbb{R}^d} \langle |u(x) - \langle u(x) \rangle|^p \rangle \, dx \right)^\frac{1}{p}.
\]

Appealing to the spectral gap estimate of Lemma 2.8 with exponent \( \frac{d}{2} \geq 1 \) yields

\[
\left( \int_{\mathbb{R}^d} \langle |u(x) - \langle u(x) \rangle|^p \rangle \frac{1}{d} \, dx \right)^\frac{1}{p} \lesssim \left( \int_{\mathbb{R}^d} \left\langle \left( \int_{\mathbb{R}^d} \langle \text{osc} u(x) \rangle^2 dz \right)^{\frac{d}{2}} \right\rangle \, dx \right)^\frac{1}{p},
\]

This yields

\[
(3.10) \left( \int_{B_L(x)} |y' - y|^{\gamma - d} |\nabla u(y')| \, dy' \right)^p \leq \int_{B_L(x)} |y' - y|^{(\gamma - d)r} |\nabla u(y')|^{p'} \, dy'

\times \left( \int_{B_L(x)} |\nabla u(y')|^{p'} \, dy' \right)^{\frac{p}{p'-1}} \left( \int_{B_L(x)} |y' - y|^{(\gamma - d)r} \, dy' \right)^{\frac{1}{p'-1}}.
\]
and by the triangle inequality
\[
\left( \int_{\mathbb{R}^d} \left| u(x) - \langle u(x) \rangle \right|^p \, dx \right)^{\frac{1}{p}} \leq \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |\text{osc}_{A_1 B_\lambda} u(x)|^p \, dz \right)^{\frac{1}{p}} \, dx \right)^{\frac{1}{p}}.
\]

By the oscillation estimate of Lemma 2.3, this turns into
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |u(x) - \langle u(x) \rangle| \, dx \right)^p \, dx \right)^{\frac{1}{p}} \leq \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \left( \mathcal{K}_{G_{\mu}} u(x, z) \right)^{\frac{1}{p}} \, dz \right)^p \, dx \right)^{\frac{1}{p}}.
\]

We now estimate the RHS. By the Cauchy-Schwarz inequality and Theorem 1.3, we have
\[
\langle K(x-z)^2 \rangle \leq K(x-z)^2 \langle (\nabla u)_t(z) \rangle \chi_{B_\lambda}(x-z) \| f \|_{L^p(B_\lambda)}^2,
\]
where again $\chi_D$ denotes the characteristic function of the set $D \subseteq \mathbb{R}^d$ and $K$ is the kernel $K(x-z) = \frac{e^{-c\sqrt{\|x-z\|}}}{1 + |x-z|^{d-1}}$.

In the following, the constant $c > 0$ in $K$ may change from line to line (and only depends on $\lambda$ and $d$). In order to correctly capture the decay of $(\nabla u)_t(z)$, we write $u$ in terms of its Green function representation and split the sum into two contributions:
\[
u(z) = \int_{\mathbb{R}^d} G_{\mu}(z, y) f(y) \, dy = \int_{\mathbb{R}^d \setminus B_{11\ell}(z)} G_{\mu}(z, y) f(y) \, dy + \int_{B_{11\ell}(z)} G_{\mu}(z, y) f(y) \, dy.
\]

Thus
\[
\langle |(\nabla u)_t(z)|^{2p} \rangle \leq \left( \int_{B_{11\ell}(z)} \left( \int_{\mathbb{R}^d} \nabla z' G_{\mu}(z', y) f(y) \, dy \right)^2 \, dz' \right)^{\frac{1}{p}} \leq \left( \int_{B_{11\ell}(z)} \left( \int_{\mathbb{R}^d \setminus B_{11\ell}(z)} \nabla z' G_{\mu}(z', y) f(y) \, dy \right)^2 \, dz' \right)^{\frac{1}{p}} + \left( \int_{B_{11\ell}(z)} \left( \int_{B_{11\ell}(z)} \nabla z' G_{\mu}(z', y) f(y) \, dy \right)^2 \, dz' \right)^{\frac{1}{p}}.
\]

We start by estimating the second RHS term, and consider the function
\[
u : z' \mapsto \int_{B_{11\ell}(z)} G_{\mu}(z', y) f(y) \, dy,
\]
which solves on $\mathbb{R}^d$
\[
\mu v - \nabla \cdot A \nabla v = f \chi_{B_{11\ell}(z)}.
\]

Set $\bar{v} := \int_{B_{11\ell}(z)} v \, dy$. An energy estimate combined with the Sobolev embedding on $B_{11\ell}(z)$ yields for $\lambda > \frac{d}{2} \geq \frac{2d}{d+2}$.
\[
\begin{align*}
\|\nabla v\|^2_{L^2(\mathbb{R}^d)} & \leq \int_{B_{11\ell}(z)} f \, v \, dy \leq \int_{B_{11\ell}(z)} f (v - \bar{v}) \, dy + \bar{v} \int_{B_{11\ell}(z)} f \, dy \\
& \leq \|f\|_{L^p(B_{11\ell}(z))} \|\nabla v\|_{L^2(\mathbb{R}^d)} + |\bar{v}| \|f\|_{L^1(B_{11\ell}(z))}.
\end{align*}
\]
It remains to estimate $\bar{v}$. By the triangle inequality and H"{o}lder’s inequality with exponents $(\lambda', \lambda)$, we have using the pointwise bounds (2.2) on $G_\mu$ in Definition 2.1

\begin{align}
(4.5) \quad |\bar{v}| & \leq \int_{B_{11}(z)} \int_{B_{11}(z)} |G_\mu(z', y)||f(y)| \, dy 
\lesssim \int_{B_{11}(z)} \left( \int_{B_{11}(z)} G_\mu(z', y)^{\lambda'} \, dy \right)^{\frac{1}{\lambda'}} \left( \int_{B_{11}(z)} |f|^\lambda \, dy \right)^{\frac{1}{\lambda'}}
\lesssim \kappa_d(\mu) \|f\|_{L^\lambda(B_{11}(z))},
\end{align}

where $\kappa_d(\mu) = 1$ if $d > 2$ and $\kappa_d(\mu) = |\mu| + 1$ if $d = 2$, since $1 \leq \lambda' < \frac{d}{d-2}$. By (4.4), (4.5), and Young’s inequality, we may thus bound the second RHS of (4.3) by

\begin{align}
(4.6) \quad \int_{B_{\ell}(z)} \left( \int_{B_{11}(z)} \nabla z' G_\mu(z', y)f(y) \, dy \right)^2 \, dz' & = \|\nabla v\|_{L^2(B_{\ell}(z))}^2 \lesssim \kappa_d(\mu) \|f\|_{L^2(B_{11}(z))}^2.
\end{align}

We then turn to the first RHS term of (4.3), and take local averages using H"{o}lder’s inequality with exponents $(\lambda', \lambda)$ (with respect to $dy$):

\begin{align*}
&\left\langle \left( \int_{B_{\ell}(z)} \left( \int_{\mathbb{R}^d \setminus B_{11}(z)} \nabla z' G(z', y)f(y) \, dy \right)^2 \, dz' \right)^{\rho_0} \right\rangle^\frac{1}{\rho_0} \\
&\quad \lesssim \left\langle \left( \int_{B_{11}(z)} \left( \int_{\mathbb{R}^d \setminus B_{11}(z)} \|\nabla z' G(z', y')\|_{L^\lambda(B(y))} \|f\|_{L^\lambda(B(y))} \, dy \right)^2 \, dz' \right)^{\rho_0} \right\rangle^\frac{1}{\rho_0}.
\end{align*}

Combined with the triangle inequality in $L^2_{\rho}(B_{\ell}(z))$, this yields

\begin{align*}
&\left\langle \left( \int_{B_{\ell}(z)} \left( \int_{\mathbb{R}^d \setminus B_{11}(z)} \nabla z' G(z', y)f(y) \, dy \right)^2 \, dz' \right)^{\rho_0} \right\rangle^\frac{1}{\rho_0} \\
&\quad \lesssim \left\langle \left( \int_{\mathbb{R}^d \setminus B_{11}(z)} \|\nabla z' G(z', y')\|_{L^\lambda(B(y)), L^2_{\rho}(B_{\ell}(z))} \|f\|_{L^\lambda(B(y))} \, dy \right)^{2\rho_0} \right\rangle^\frac{1}{\rho_0}.
\end{align*}

From the De Giorgi-Nash-Moser theory in the form of Lemma 2.10 (with RHS zero), we then have

$$
\|\nabla z' G(z', y')\|_{L^\lambda(B(y), L^2_{\rho}(B_{\ell}(z)))} \lesssim (\nabla G)_{\mu}(z, y).
$$

We then finally appeal to Theorem 1.3 and the triangle inequality with respect to $L^2_{(\rho)}$ to obtain the following estimate of the first RHS term of (4.3):

\begin{align}
&\left\langle \left( \int_{B_{\ell}(z)} \left( \int_{\mathbb{R}^d \setminus B_{11}(z)} \nabla z' G(z', y)f(y) \, dy \right)^2 \, dz' \right)^{\rho_0} \right\rangle^\frac{1}{\rho_0} \\
&\quad \lesssim \int_{\mathbb{R}^d \setminus B_{11}(z)} K(z - y) \|f\|_{L^\lambda(B_{\ell}(y))} \, dy.
\end{align}

Since $K(z - y) \sim 1$ for $y \in B_{11}(z)$, the combination of (4.3), (4.6), and (4.7) yields

\begin{align}
(4.8) \quad \langle (\nabla u)_{\mu}(z)^{2\rho_0} \rangle^\frac{1}{2\rho_0} \lesssim \int_{\mathbb{R}^d} K(z - y) \|f\|_{L^\lambda(B_{11}(y))} \, dy.
\end{align}
In total, collecting (4.1), (4.2) and (4.3), we then have
\[
\left( \int_{\mathbb{R}^d} |u(x) - \langle u(x) \rangle|^p \, dx \right)^{\frac{1}{p}} \lesssim \kappa_d(\mu)^{\frac{1}{p}} \left( \int_{\mathbb{R}^d} \|f\|_{L^q(B_{2r}(z))}^p \, dz \right)^{\frac{1}{p}}
+ \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 \left( \int_{\mathbb{R}^d} K(z-y) \|f\|_{L^q(B_{3r}(y))} \, dy \right)^2 \, dz \right)^{\frac{1}{p}} \, dx \right)^{\frac{1}{p}}.
\]

Since \( q \leq p \) and the integral of the RHS term is equivalent to a discrete sum over an appropriate lattice of size \( \ell \), we have that
\[
\left( \int_{\mathbb{R}^d} \|f\|_{L^q(B_{2r}(z))}^p \, dz \right)^{\frac{1}{p}} \lesssim \left( \int_{\mathbb{R}^d} \|f\|_{L^q(B_{3r}(y))}^p \, dy \right)^{\frac{1}{p}} \lesssim \left( \int_{\mathbb{R}^d} \|f\|_{L^q(B_{4r}(y))}^p \, dy \right)^{\frac{1}{p}}.
\]

The most important term is the last one. By the triangle inequality in \( L^2(\mathbb{R}^d) \),
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 \left( \int_{\mathbb{R}^d} K(z-y) \|f\|_{L^q(B_{4r}(y))} \, dy \right)^2 \, dz \right)^{\frac{1}{p}} \, dx \right)^{\frac{1}{p}} \leq \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 K(z-y)^2 \|f\|_{L^q(B_{4r}(y))}^2 \, dz \right)^{\frac{1}{p}} \, dy \right)^{\frac{1}{p}} \, dx \right)^{\frac{1}{p}}.
\]

We bound the integral over \( z \) as follows:
\[
\int_{\mathbb{R}^d} e^{-c\sqrt{\sqrt{|x-z|^2}}} e^{-c\sqrt{\sqrt{|x-y|^2}}} \, dz \lesssim \begin{cases} e^{-c\sqrt{|x-y|^2}} & \text{if } d > 2, \\ (|\ln \mu| + 1) e^{-c\sqrt{|x-y|^2}} & \text{if } d = 2. \end{cases}
\]

In other words,
\[
\int_{\mathbb{R}^d} K(x-z)^2 K(z-y)^2 \, dz \lesssim K(x-y)^2 \kappa_d(\mu),
\]
where we recall that \( \kappa_d(\mu) = 1 \) for \( d > 2 \) and \( \mu_d(\mu) = |\ln \mu| + 1 \) for \( d = 2 \). We thus have
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 K(z-y)^2 \|f\|_{L^q(B_{4r}(y))}^2 \, dz \right)^{\frac{1}{p}} \, dy \right)^{\frac{1}{p}} \lesssim \kappa_d(\mu)^\frac{1}{p} \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 K(z-y)^2 \|f\|_{L^q(B_{4r}(y))} \, dy \right)^p \, dx \right)^{\frac{1}{p}}.
\]

Let us pick \( 1 \leq r \leq \frac{d}{d-1} \) and \( 1 \leq q < +\infty \) such that (1.19) holds. If \( r < \frac{d}{d-1} \), Young’s inequality yields
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z)^2 K(z-y)^2 \|f\|_{L^q(B_{4r}(y))} \, dy \right)^p \, dx \right)^{\frac{1}{p}} \lesssim \|K\|_{L^r(\mathbb{R}^d)} \left( \int_{\mathbb{R}^d} \|f\|_{L^l(B_{4r}(y))}^p \, dx \right)^{\frac{1}{p}}.
\]

We easily check that
\[
\|K\|_{L^r(\mathbb{R}^d)} = \left( \int_{\mathbb{R}^d} K(x)^r \, dx \right)^{\frac{1}{r}} \lesssim 1 + \mu^{\left(\frac{1-d}{d}x + \epsilon\right)}.
\]
In the border-line case \( r = \frac{d}{d-1} \), the Hardy-Littlewood-Sobolev inequality immediately yields provided \( q > 1 \)

\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |x-y|^{1-d} \| f \|_{L^\lambda(B_{1r}(y))} \, dy \right)^p \, dx \right)^{\frac{1}{p}} \lesssim \left( \int_{\mathbb{R}^d} \| f \|_{L^\lambda(B_{1r}(y))} \, dx \right)^{\frac{1}{q}},
\]

where we have also used the elementary fact that \( \frac{1}{1+|x-y|^r} \lesssim \frac{1}{|x-y|^r} \). Collecting \((4.11), (4.11)\), and \((4.12)\) yields

\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-y) \| f \|_{L^\lambda(B_{1r}(y))} \, dy \right)^p \, dx \right)^{\frac{1}{p}} \lesssim (1 + \mu)^{\frac{4-d}{2p}} \kappa_d(\mu)^{\frac{1}{2}} \left( \int_{\mathbb{R}^d} \| f \|_{L^\lambda(B_{2r}(x))} \, dx \right)^{\frac{1}{q}}
\]

\[
\lesssim (1 + \mu)^{\frac{4-d}{2p}} \kappa_d(\mu)^{\frac{1}{2}} \left( \int_{\mathbb{R}^d} \| f \|_{L^\lambda(B_{2r}(x))} \, dx \right)^{\frac{1}{q}},
\]

where \( p, q \) and \( r \) are related by \((1.19)\). This concludes the proof of the theorem.

### 4.2. Proof of Theorem 1.7

Since transposition is a linear local operator, if \( A \) satisfies the assumptions of Theorem 1.7, then \( A^* \) does as well, so that the statement of Theorem 1.7 is symmetric with respect to interchanging \( f \) and \( g \) provided \( A \) is replaced by \( A^* \). Hence we may without loss of generality assume that \( \lambda_1 \leq \lambda_2 \). By \((2.18)\), this implies that

\[
\lambda_2 \geq \frac{2d}{d+2}.
\]

By Jensen’s inequality in probability we may assume w. l. o. g. that \( \theta \geq 2 \). The spectral gap estimate of Lemma 2.8 for \( q = \frac{\theta}{2} \geq 1 \) yields

\[
\left\langle \left| \int_{\mathbb{R}^d} (u(x) - \langle u(x) \rangle) g(x) \, dx \right|^\theta \right\rangle \lesssim \left\langle \left( \int_{\mathbb{R}^d} \left( \text{osc}_{B_{1r}(x)} \int_{\mathbb{R}^d} u(x) g(x) \, dx \right)^2 \, dx \right)^{\frac{\theta}{2}} \right\rangle^{\frac{2}{\theta}}.
\]

By the triangle inequality, we may insert the unperturbed solution \( u \) and estimate

\[
\left\langle \left( \int_{\mathbb{R}^d} \left( \text{osc}_{B_{1r}(x)} \int_{\mathbb{R}^d} u(x) g(x) \, dx \right)^2 \, dx \right)^{\frac{\theta}{2}} \right\rangle^{\frac{2}{\theta}} \lesssim 2 \left\langle \left( \int_{\mathbb{R}^d} \sup_{B_{1r}(x)} \int_{\mathbb{R}^d} |u(x) - \tilde{u}(x)| g(x) \, dx \right)^2 \, dx \right)^{\frac{\theta}{2}} \right\rangle^{\frac{2}{\theta}}.
\]

Taking local averages combined with Hölder’s inequality with exponents \((\lambda_1', \lambda_1)\) yields

\[
\left\langle \left( \int_{\mathbb{R}^d} (u(x) - \langle u(x) \rangle) g(x) \, dx \right|^\theta \right\rangle \lesssim \left\langle \left( \int_{\mathbb{R}^d} \sup_{B_{1r}(x)} \int_{\mathbb{R}^d} \|u - \tilde{u}\|_{L^{\lambda_1}(B_{1r}(x))} \|g\|_{L^{\lambda_1}(B_{1r}(x))} \, dx \right)^2 \, dx \right)^{\frac{\theta}{2}} \right\rangle^{\frac{2}{\theta}}.
\]
We then put the supremum inside the inner integral and appeal to the sensitivity estimate of Lemma 2.10 to obtain

\[
\left\langle \left( \int_{\mathbb{R}^d} \left( \sup_{A \in P(\mathbb{R}^d)} \int_{\mathbb{R}^d} \|u - \tilde{u}\|_{L^{q_1}(B_1(x))} \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^2 \, dz \right)^{\frac{q}{2}} \right\rangle
\]

\[
\lesssim \left\langle \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \mathcal{K}_{G,u}(x,z) \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^2 \, dz \right)^{\frac{q}{2}} \right\rangle.
\]

It remains to estimate the RHS. By the triangle inequality in \( L^1_{s,1} \), first with \( s = \frac{\theta}{2} \geq 1 \) and then \( s = 2 \), we have

\[
\left\langle \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \mathcal{K}_{G,u}(x,z) \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^2 \, dz \right)^{\frac{q}{2}} \right\rangle
\]

\[
\lesssim \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \mathcal{K}_{G,u}(x,z) \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^2 \, dz \right)^{\frac{1}{2}}
\]

\[
+ \left( \int_{\mathbb{R}^d} \|f\|_{L^{q_1}(B_2(z))}^2 \|g\|_{L^{q_1}(B_{r_1}(z))}^2 \, dz \right)^{\frac{1}{2}}.
\]

By Hölder’s inequality with \( \frac{1}{2} = \frac{1}{q_1} + \frac{1}{q_1} \), we bound the second RHS term by

\[
\left( \int_{\mathbb{R}^d} \|f\|_{L^{q_1}(B_2(z))}^2 \|g\|_{L^{q_1}(B_{r_1}(z))}^2 \, dz \right)^{\frac{1}{2}} \lesssim \|f\|_{L^{q_1}_{2,1}(\mathbb{R}^d)} \|g\|_{L^{q_1}_{1,1}(\mathbb{R}^d)}.
\]

By \( 1.20 \), since \( r, \tilde{r} \geq 1 \), we may choose \( q_1 \geq q \) and \( \tilde{q}_1 \geq \tilde{q} \) so that

\[
\|f\|_{L^{q_1}_{2,1}(\mathbb{R}^d)} \|g\|_{L^{q_1}_{1,1}(\mathbb{R}^d)} \lesssim \|f\|_{L^{\tilde{q}_1}_{3,1}(\mathbb{R}^d)} \|g\|_{L^{\tilde{q}_1}_{3,1}(\mathbb{R}^d)}.
\]

From \( 4.8 \) (with \( p = 1 \)) in the proof of Theorem 1.6 we learn that

\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z) \langle (\nabla u)_2(z) \rangle_B^2 \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^2 \, dz \right)^{\frac{1}{2}}
\]

\[
\lesssim \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x-z) K(z-y) \|f\|_{L^{q_1}(B_1(x))} \|g\|_{L^{q_1}(B_1(x))} \, dx \, dy \right)^2 \, dz \right)^{\frac{1}{2}},
\]

which holds by our choice \( \lambda_2 \geq \lambda_1 \) which implies \( \lambda_2 > \frac{2d}{\lambda_1} \) by \( 1.24 \). Let \( p, \tilde{p} \geq 1 \) be two exponents to be specified later such that \( \frac{1}{2} = \frac{1}{p} + \frac{1}{\tilde{p}} \). We then have that

\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x-z) K(z-y) \|f\|_{L^{q_1}(B_1(x))} \|g\|_{L^{q_1}(B_1(x))} \, dx \, dy \right)^2 \, dz \right)^{\frac{1}{2}}
\]

\[
\lesssim \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(z-y) \|f\|_{L^{q_1}(B_1(x))} \, dy \right)^p \, dz \right)^{\frac{1}{p}} \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x-z) \|g\|_{L^{q_1}(B_1(x))} \, dx \right)^q \, dz \right)^{\frac{1}{q}}.
\]
We treat the two factors of the RHS the same way. First we consider the non-borderline case \( r < \frac{d}{d + 1}, \) in which case Young’s convolution inequality with \( 1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{q} \) yields
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(z - y) f \|L^p(B_{11r}(y)) \|g \|L^q(B_{11r}(x)) \|dx \right)^p dy \right)^\frac{1}{p} \lesssim \|K\|_{L^r(\mathbb{R}^d)} \|f\|_{L^p_{2,1}(\mathbb{R}^d)} \lesssim \mu^{-1 - \frac{d}{d + 1}} \|f\|_{L^p_{2,1}(\mathbb{R}^d)}.
\]
In the borderline case \( r = \frac{d}{d + 1}, \) the result follows from the Hardy-Littlewood-Sobolev inequality provided \( q > 1. \) An identical estimate holds for the second factor with exponents \( 1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{q} \) (provided \( \tilde{q} > 1 \) in the borderline case). Gathering these two estimates yields
\[
\left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} K(x - z) K(z - y) f \|L^p_{2,1}(B_{11r}(y)) \|g \|L^q_1(B_{11r}(x)) \|dx \right)^p dy \right)^\frac{1}{p} \lesssim \mu^{-(1 - d) - \frac{d}{2(d + 1)}} \|f\|_{L^p_{2,1}(\mathbb{R}^d)} \|g\|_{L^q_{1,1}(\mathbb{R}^d)},
\]
with
\[
2 + \frac{1}{2} = 1 + \frac{1}{p} + 1 + \frac{1}{q} = \frac{1}{r} + \frac{1}{r} + \frac{1}{q} + \frac{1}{q}
\]
This completes the proof.

5. Proofs of the improved De Giorgi-Nash-Moser theory

5.1. Proof of Theorem \ref{thm1}. As opposed to the corresponding proof in the discrete case, cf. \cite{16} Corollary 4), we have to take care of the singularity of the Green function. This prevents us to make use of Morrey’s inequality when the coefficients are only measurable, and we propose a more direct approach which partly mimics the proof of Morrey’s inequality. We assume w. l. o. g. that \( R > 9L. \) In the first five steps we assume that \( d > 2, \) and indicate the changes for \( d = 2 \) in Step 6.

Step 1. Representation formula for \( u(x + x') - u(x'), x \in B_R \setminus B_{2L}, x' \in B_L. \) In order to make use of the annealed estimates of Theorem \ref{thm1} we rewrite equation \ref{eq1} on \( \mathbb{R}^d \) as follows. Let \( \eta : \mathbb{R}^d \to [0, 1] \) be a cutoff-function for \( B^*_{2L} \) in \( B^*_{2R} \) such that \( |\nabla \eta| \lesssim R^{-1}. \) A direct calculation shows that \( \eta u \in H^1(\mathbb{R}^d) \) satisfies
\[
\mu \eta u - \nabla \cdot A \nabla (\eta u) = \mu \eta u - \eta \nabla \cdot A \nabla u - \nabla \eta \cdot A \nabla u - \nabla \cdot (u A \nabla \eta).
\]
The sum of the first two RHS terms equals \( \eta f \) while the other two terms belong to \( H^{-1}(\mathbb{R}^d) \) and have compact support. The Green representation formula yields
\[
(\eta u)(x) = \int_{B_{2R}} \left( G_\mu(x, y)(\eta(y) f(y) - \nabla \eta(y) \cdot A(y) \nabla u(y)) + u(y) \nabla_y G_\mu(x, y) \cdot A(y) \nabla \eta(y) \right) dy.
\]
Assume first that \( f \) and \( A \) are smooth (so that \( u \) is smooth and the formula holds classically). We argue by density. Since \( 0 \leq G_\mu(x, y) \leq |y - x|^{d - d}, \eta f \in L^p(\mathbb{R}^d) \) with \( p > \frac{d}{d}, \nabla \eta = 0 \) on \( B_{3R}/2 \) (and in particular at the singularity of \( G_\mu(x, \cdot) \)), the first term of the integral is well-defined at the limit. Recalling that \( y \mapsto \nabla_y G_\mu(x, y) \) is locally square-integrable away from \( y = x, \) the second term of the integral is well-defined as well since \( \nabla \eta \) vanishes in a neighborhood of the singularity of \( y \mapsto \nabla_y G_\mu(x, y) \) and \( u \in L^2(B_{2R}). \) Since \( u \) is uniformly Hölder continuous, one can also take the limit of the LHS, so that the Green representation formula holds by a density and regularization argument.
We thus have for all \( x \in B_R \setminus B_{2L} \) and \( x' \in B_L \), using in addition that \( \nabla \eta \) vanishes on \( B_{4R} \),

\[
(5.2) \quad u(x + x') - u(x') = \int_{B_{2R}} (G_\mu(x + x', y) - G_\mu(x', y))\eta(y)f(y)dy \\
- \int_{B_{2R} \setminus B_{4R}} \left( (G_\mu(x + x', y) - G_\mu(x', y))\nabla \eta(y) \cdot A(y)\nabla u(y) \\
- u(y)\nabla \eta(G_\mu(x + x', y) - G_\mu(x', y)) \cdot A(y)\nabla \eta(y) \right) dy.
\]

**Step 2.** Estimate of the integral on \( B_{3L}(x) \cup B_{3L} \subset B_{2L} \).

Since \( R > 9L, x \in B_R \) and \( x' \in B_L, B_{3L}(x) \cup B_{3L} \subset B_{4R} \), only the first integral term of the RHS of (5.2) has a contribution. We shall argue that

\[
(5.3) \quad \left| \int_{B_{3L}(x) \cup B_{3L}} (G_\mu(x + x', y) - G_\mu(x', y))\eta(y)f(y)dy \right| \lesssim \left( \int_{B_{2R}} |f|^q dy \right)^{\frac{1}{q}}.
\]

Indeed, the quenched pointwise estimates on \( G_\mu \) for \( d > 2 \) combined with Hölder’s inequality with exponents \( \frac{2}{q} \) yield

\[
\left| \int_{B_{3L}(x) \cup B_{3L}} (G_\mu(x + x', y) - G_\mu(x', y))\eta(y)f(y)dy \right| \lesssim \left( \int_{B_{3L}(x) \cup B_{3L}} (|x + x' - y|^{\frac{2(d-2)}{d-1}} + |x' - y|^{\frac{2(d-2)}{d-1}})dy \right)^{\frac{1}{2}} \times \left( \int_{B_{2R}} |f|^q dy \right)^{\frac{1}{q}}.
\]

Since \( q > \frac{d}{2} \) implies \( \frac{2(d-2)}{d-1} < d \), the first factor is of order 1, and (5.3) follows.

**Step 3.** Representation formulas for \( G_\mu(x + x', y) - G_{\mu}(x', y) \) and \( \nabla_y G_\mu(x + x', y) - \nabla_y G_{\mu}(x', y) \), \( x \in B_R \setminus B_{2L}, x' \in B_L, y \notin B_{3L}(x) \cup B_{3L} \).

When \( y \) is not at the singularity of the Green function, we may write the difference of Green functions as the directional integral of its gradient: for all \( y \notin [x', x'] \),

\[
(5.4) \quad G_\mu(x + x', y) - G_{\mu}(x', y) = \int_0^1 \nabla_x G_\mu(tx + x', y) \cdot xdt,
\]

and for all \( i \in \{1, \ldots, d\} \),

\[
(5.5) \quad \nabla_{y_i} G_\mu(x + x', y) - \nabla_{y_i} G_{\mu}(x', y) = \int_0^1 \nabla_x \nabla_{y_i} G_\mu(tx + x', y) \cdot xdt.
\]

When \( y \) is close to \([x', x'] \), we have to refine this decomposition. To this end, we define two points \( x^+ \) and \( x^- \) and two sets \( B^+ \) and \( B^- \) as follows:

\[
x^+ := \frac{x}{2} + \left( \frac{|x|}{2} + L \right)e_1, \quad x^- := \frac{x}{2} - \left( \frac{|x|}{2} + L \right)e_1,
\]

where \( e_1 \) is the first unit vector of the canonical basis of \( \mathbb{R}^d \), and

\[
B^+ := \{ y \in B_{2R} \setminus (B_{3L}(x) \cup B_{3L}), (y - x) \cdot e_1 \leq 0 \}, \\
B^- := \{ y \in B_{2R} \setminus (B_{3L}(x) \cup B_{3L}), (y - x) \cdot e_1 > 0 \}.
\]

Note that \( B^+ \cup B^- = B_{2R} \setminus (B_{3L}(x) \cup B_{3L}) \). For \( x \in B^+ \) we write \( G_\mu(x + x', y) - G_{\mu}(x', y) = G_\mu(x + x', y) - G_\mu(x' + x', y) + G_\mu(x' + x', y) - G_{\mu}(x', y) \), so that

\[
(5.6) \quad G_\mu(x + x', y) - G_{\mu}(x', y) = \int_0^1 \nabla_x G_\mu(x^+ + t(x - x^+) + x', y) \cdot (x - x^+) dt + \int_0^1 \nabla_x G_\mu(tx^+ + x', y) \cdot x^+ dt.
\]
We proceed correspondingly for \( y \in B^- \).

In the following step we estimate the RHS of (5.2). In view of Step 2, it only remains to estimate the integrals on \( \tilde{B} := B_{2R} \setminus (B_{3L}(x) \cup B_{3L}) \).

**Step 4.** Estimates of the integrals on \( \tilde{B} \).

We shall prove three estimates. First,

\[
\int_{\tilde{B}} \left| (G_\mu(x + x', y) - G_\mu(x', y)) \eta(y) f(y) dy \right| 
\lesssim |x| \left( \int_{B_{2R}} |f|^q dy \right)^{\frac{1}{q}} \int_0^1 \left( \int_{B_{2R}} \left| \nabla G_\mu(x^\pm + t(x - x^\pm) + x', y) \right|^{\frac{q}{q-1}} \right. 
\left. + \left| \nabla G_\mu(tx^\pm + x', y) \right|^{\frac{q}{q-1}} \right) dy \right)^{\frac{q-1}{2}} dt, 
\]

where \( B^\pm \) is a shorthand notation we use when the inequality holds both on \( B^+ \) and \( B^- \). We only prove the claim for \( B^+ \). Since \(|x| > L\), by construction \(|x - x^\pm| \lesssim |x|\) and \(|x^\pm| \lesssim |x|\), so that (5.7) follows from (5.6) and Hölder’s inequality with exponents \( (\frac{q}{q-1}, q) \).

The second estimate is:

\[
\int_{\tilde{B} \setminus B_{\frac{3L}}}(G_\mu(x + x', y) - G_\mu(x', y))\nabla \eta(y) \cdot A(y) \nabla u(y) dy
\lesssim |x| \left( R^{-1} \left( \int_{B_{2R}} u^2 dy \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} f^2 dy \right)^{\frac{1}{2}} \right)
\times \int_0^1 \left( \int_{\tilde{B} \setminus B_{\frac{3L}}} \left| \nabla G_\mu(tx^\pm + x', y) \right|^2 dy \right)^{\frac{1}{2}} dt.
\]

We proceed as for the proof of (5.7) and use in addition the following consequence of the definition of \( \eta \) and Caccioppoli’s inequality:

\[
\left( \int_{B_{2R}} |\nabla \eta|^2 |\nabla u|^2 dy \right)^{\frac{1}{2}} \lesssim R^{-1} \left( \int_{B_{2R}} u^2 dy \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} f^2 dy \right)^{\frac{1}{2}}.
\]

Indeed, since \( \nabla \eta \) has support in \( B_{\frac{3L}2} \setminus B_{\frac{3L}4} \),

\[
\int_{B_{2R}} |\nabla \eta|^2 |\nabla u|^2 dy \lesssim R^{-2} \int_{B_{\frac{3L}2} \setminus B_{\frac{3L}4}} |\nabla u|^2 dy.
\]

Testing equation (1.26) with test-function \( \tilde{\eta}^2 u \in H^1_0(B_{2R}) \), where \( \tilde{\eta} \) has support in \( B_{2R} \) and is such that \( \tilde{\eta}|_{B_{\frac{3L}2}} \equiv 1 \) and \( |\nabla \tilde{\eta}| \lesssim \frac{1}{R} \), yields the Caccioppoli estimate

\[
\int_{B_{2R}} \nabla (\tilde{\eta} u) \cdot A \nabla (\tilde{\eta} u) dy \lesssim \int_{B_{2R}} u^2 \nabla \tilde{\eta} \cdot A \nabla \tilde{\eta} dy + \int_{B_{2R}} \tilde{\eta}^2 f dy,
\]

which, by definition of \( \tilde{\eta} \) and Young’s inequality on the last term, we may use in the form

\[
\int_{B_{\frac{3L}2}} |\nabla u|^2 dy \lesssim R^{-2} \int_{B_{2R}} u^2 dy + R^2 \int_{B_{2R}} f^2 dy.
\]
Finally, we prove

\begin{equation}
\left| \int_{B \setminus B_{4R}} u(y) \nabla_y (G_\mu(x + x', y) - G_\mu(x', y)) \cdot A(y) \nabla \eta(y) dy \right| \lesssim |x|R^{-1} \left( \int_{B_{2R}} u^2 dy \right)^{\frac{1}{2}} \int_0^1 \left( \int_{B \setminus B_{4R}} |\nabla \nabla G_\mu(tx + x', y)|^2 dy \right)^{\frac{1}{2}} dt.
\end{equation}

This estimate follows from (5.9), the bound \(|\eta| \lesssim R^{-1}\), and Cauchy-Schwarz’ inequality.

**Step 5.** Conclusion for \(d > 2\).

The combination of (5.2), (5.3), (5.7), (5.8), and (5.10) yields, using that \(|x| > L \sim 1\) and that \(B \setminus B_{4R} = B_{2R} \setminus B_{4R}\),

\[
|u(x + x') - u(x')| \lesssim R^{-1} \left( \left( \int_{B_{2R}} |R f|^q dy \right)^{\frac{1}{q}} + \left( \int_{B_{2R}} (R f)^2 dy \right)^{\frac{1}{2}} \right) \times \left\{ \begin{array}{l}
R^{-1+\frac{d}{q}} \int_0^1 \left( \int_{B^{-}} \left( |\nabla_x G_\mu(x^+ + t(x - x^+) + x', y)|^{\frac{q}{q-1}} + |\nabla_x G_\mu(tx^+ + x', y)|^{\frac{q}{q-1}} \right) dy \right) \frac{dx'}{t} dt \\
+ R^{-1+\frac{d}{q}} \int_0^1 \left( \int_{B^{-}} \left( |\nabla_x G_\mu(x^- + t(x - x^-) + x', y)|^{\frac{q}{q-1}} + |\nabla_x G_\mu(tx^- + x', y)|^{\frac{q}{q-1}} \right) dy \right) \frac{dx'}{t} dt \\
+ R^{-1+\frac{d}{q}} \int_0^1 \left( \int_{B_{2R} \setminus B_{4R}} |\nabla \nabla G_\mu(tx + x', y)|^2 dy \right)^{\frac{1}{2}} dt \\
+ R^{\frac{d}{2}} \int_0^1 \left( \int_{B_{2R} \setminus B_{4R}} |\nabla \nabla G_\mu(tx + x', y)|^2 dy \right)^{\frac{1}{2}} dt + 1.
\end{array} \right.
\]

Dividing both sides of the inequality by the first RHS term and averaging over \(x' \in B_L\) yield using Jensen’s inequality and that \(q > d\) (so that \(R^{-1+\frac{d}{q}} \lesssim 1\)):

\[
R \int_{B_L} \frac{|u(x + x') - u(x')|}{|x|} dx' \left( \int_{B_{2R}} u^2 dy \right)^{\frac{1}{2}} + \left( \int_{B_{2R}} |R f|^q dy \right)^{\frac{1}{q}} \lesssim \int_0^1 \left( \int_{B_L} \int_{B^+} \left( |\nabla_x G_\mu(x^+ + t(x - x^+) + x', y)|^{\frac{q}{q-1}} + |\nabla_x G_\mu(tx^+ + x', y)|^{\frac{q}{q-1}} \right) dy dx' \right) \frac{dx'}{t} dt \\
+ \int_0^1 \left( \int_{B_L} \int_{B^{-}} \left( |\nabla_x G_\mu(x^- + t(x - x^-) + x', y)|^{\frac{q}{q-1}} + |\nabla_x G_\mu(tx^- + x', y)|^{\frac{q}{q-1}} \right) dy dx' \right) \frac{dx'}{t} dt \\
+ R^{-1+\frac{d}{q}} \int_0^1 \left( \int_{B_{2R} \setminus B_{4R}} |\nabla \nabla G_\mu(tx + x', y)|^2 dy dx' \right)^{\frac{1}{2}} dt \\
+ R^{\frac{d}{2}} \int_0^1 \left( \int_{B_{2R} \setminus B_{4R}} |\nabla \nabla G_\mu(tx + x', y)|^2 dy dx' \right)^{\frac{1}{2}} dt + 1.
\]

Estimate (5.25) formally follows from taking the expectation of the \(p\)-th power of this inequality and bounding \(|\nabla G_\mu(x, y)|\) by \(|x - y|^{-d}\) and \(|\nabla \nabla G_\mu(x, y)|\) by \(|x - y|^{-d}\). It remains to show that it is enough to use bounds on large moments of local square averages of \(|\nabla G_\mu(x, y)|\) and \(|\nabla \nabla G_\mu(x, y)|\) instead, which we control optimally by Theorem 1.3. We only treat the first term in detail (the other terms are treated similarly). By bounding the integral on \(B^+\) by the sum of integrals on
Recall that by construction of $G$ we follow the elegant argument by Avellaneda and Lin [2] and add a third dimension. We denote $d > 2$, the estimate fails optimality by a logarithm of $\mu$ due to the bound on the Green function $G_{\mu}$ in dimension 2 close to the singularity. Recall that $p > \frac{d}{2} = 1$. To avoid this logarithmic correction, we follow the elegant argument by Avellaneda and Lin [2] and add a third dimension. We denote by $G_{\mu}^{(2)}$ and $A^{(2)}$ the fields in dimension 2 and consider the following extensions to dimension 3: $A^{(3)}(x_{1}, x_{2}, x_{3}) := \text{diag} \left[ A^{(2)}(x_{1}, x_{2}), 1 \right]$ and $G_{\mu}^{(3)}$ the Green function associated with $A^{(3)}$. It is elementary to check using Definition 2.1 that for all $x \neq y \in \mathbb{R}^{2}$,

$$G_{\mu}^{(2)}(x, y) = \int_{\mathbb{R}} G_{\mu}^{(3)}((x, 0), (y, t)) dt,$$

balls of radius $L$ and by Hölder’s inequality, we have

$$\int_{B_{L}} \int_{B^{+}} \left( |\nabla_{x} G_{\mu}(x^{+} + t(x - x^{+}) + x', y)|^{\frac{p}{q}} + |\nabla_{x} G_{\mu}(tx^{+} + x', y)|^{\frac{p}{q}} \right) dy dx' \leq \sum_{i \in B^{+} \cap \frac{\delta}{2} \mathbb{Z}^{d}} \left( \langle (\nabla_{x} G_{\mu}) L(x^{+} + t(x - x^{+}), i) \rangle \right)^{\frac{q-1}{q}} + \sum_{i \in B^{+} \cap \frac{\delta}{2} \mathbb{Z}^{d}} \left( \langle (\nabla_{x} G_{\mu}) L(tx^{+}, i) \rangle \right)^{\frac{q-1}{q}}.$$

We only treat the first RHS term. By Jensen’s inequality in probability it is enough to prove the claim for $p$ large enough, which we take such that $p \geq \frac{d}{q - 1}$. By Jensen’s inequality on $\int_{0}^{1} dt$ and by the triangle inequality for $\mu$

$$\left\langle \left( \int_{0}^{1} \left( \sum_{i \in B^{+} \cap \frac{\delta}{2} \mathbb{Z}^{d}} \left( \langle (\nabla_{x} G_{\mu}) L(x^{+} + t(x - x^{+}), i) \rangle \right)^{\frac{q-1}{q}} dt \right)^{p} \right) \right\rangle \leq \left( \int_{0}^{1} \left( \sum_{i \in B^{+} \cap \frac{\delta}{2} \mathbb{Z}^{d}} \left( \langle (\nabla_{x} G_{\mu}) L(x^{+} + t(x - x^{+}), i) \rangle \right)^{\frac{q-1}{q}} dt \right)^{\frac{p}{q}} \right)^{q}. $$

Recall that by construction of $x^{+}$ and $B^{+}$, $|x^{+} + t(x - x^{+}) - i| \sim |x - i|$ for all $t \in [0, 1]$, so that by Theorem 1.3

$$\left\langle \int_{0}^{1} ((\nabla_{x} G_{\mu}) L(x^{+} + t(x - x^{+}), i)) dt \right\rangle^{\frac{q}{p}} \leq e^{-c_{\mu} |x - i| d}.$$ 

Giving up the exponential cut-off, this yields

$$\left\langle \left( \int_{0}^{1} \left( \sum_{i \in B^{+} \cap \frac{\delta}{2} \mathbb{Z}^{d}} \left( |x - i|^{(1-d)\frac{q-1}{q}} \right)^{\frac{p}{q}} dt \right)^{\frac{q}{p}} \right) \right\rangle \leq \left( \int_{B_{2R} \setminus B_{L}(x)} |x - y|^{(1-d)\frac{q-1}{q}} dy \right)^{\frac{q}{p}} \leq 1$$

since $q > d$. This completes the proof of (1.25). To prove (1.27), we have to take the supremum over $x$ inside the expectation. We then appeal to the quenched estimates of Corollary 1.3.

**Step 6.** Proof for $d = 2$.

The proof for $d = 2$ is identical as for $d > 2$ except for Step 2. Indeed, if we proceed there as for $d > 2$, the estimate fails optimality by a logarithm of $\mu$ due to the bound on the Green function $G_{\mu}$ in dimension 2 close to the singularity. Recall that $p > \frac{d}{2} = 1$. To avoid this logarithmic correction, we follow the elegant argument by Avellaneda and Lin [2] and add a third dimension. We denote by $G_{\mu}^{(3)}$ and $A^{(3)}$ the fields in dimension 2 and consider the following extensions to dimension 3: $A^{(3)}(x_{1}, x_{2}, x_{3}) := \text{diag} \left[ A^{(3)}(x_{1}, x_{2}), 1 \right]$ and $G_{\mu}^{(3)}$ the Green function associated with $A^{(3)}$. It is elementary to check using Definition 2.1 that for all $x \neq y \in \mathbb{R}^{2}$,

$$G_{\mu}^{(2)}(x, y) = \int_{\mathbb{R}} G_{\mu}^{(3)}((x, 0), (y, t)) dt,$$
and we rewrite the LHS of (5.3) as

\[
\begin{aligned}
(5.11) \quad & \int_{B_3(x) \cup B_3} (G^2_\mu(x + x', y) - G^2_\mu(x', y)) \eta(y) f(y) dy \\
& = \int_{B_3(x) \cup B_3} (G^3_\mu((x + x', 0), (y, t)) - G^3_\mu((x', 0), (y, t))) \eta(y) f(y) dt dy.
\end{aligned}
\]

We then split the integral over \( t \) into two parts: \(|t| \leq 1\) and \(|t| > 1\). We start by estimating the first part, and appeal to the quenched pointwise estimate on \( G^3_\mu \). By the triangle inequality,

\[
\begin{aligned}
& \left| \int_{B_3(x) \cup B_3} \int_{|t| \leq 1} (G^3_\mu((x + x', 0), (y, t)) - G^3_\mu((x', 0), (y, t))) \eta(y) f(y) dt dy \right| \\
& \leq \int_{B_3(x) \cup B_3} \int_{|t| \leq 1} \left( |x + x' - y|^2 + t^2 \right)^{-\frac{1}{2}} \left( (|x' - y|^2 + t^2)^{-\frac{1}{2}} \right) |f(y)| dt dy.
\end{aligned}
\]

We first integrate in \( y \) and use Hölder’s inequality with exponents \( \left( \frac{q}{q-1}, q \right) \) for some \( 1 < q \leq p \) small enough so that \( \frac{q}{q-1} > 2 \). This yields

\[
(5.12) \quad \left| \int_{B_3(x) \cup B_3} \int_{|t| \leq 1} (G^3_\mu((x + x', 0), (y, t)) - G^3_\mu((x', 0), (y, t))) \eta(y) f(y) dt dy \right| \\
\leq \int_{|t| \leq 1} |t|^{\frac{q-1}{q}} \left( \int_{B_3(x) \cup B_3} |f(y)|^q dy \right)^{\frac{1}{q}} \leq \left( \int_{B_3(x) \cup B_3} |f(y)|^q dy \right)^{\frac{1}{q}},
\]

by Jensen’s inequality since \( L \sim 1 \).

We turn to the second part of the integral. We bound the difference of the Green functions by the oscillation, and appeal to the De Giorgi-Nash-Moser theory in the form of the quenched estimate: For all \(|t| > 1\), and all \( z, y \in \mathbb{R}^d \),

\[
\text{osc}_{z \in B_2R} G^3_\mu((z, 0), (y, t)) \lesssim |t|^{-1+\alpha_0},
\]

for some \( \alpha_0 > 0 \) depending only on \( \lambda \) (see (6.10) in Step 2 of the proof of Lemma 4.6 for details). Since \( x, x' \in B_2R \), this yields

\[
(5.13) \quad \left| \int_{B_3(x) \cup B_3} \int_{|t| > 1} (G^3_\mu((x + x', 0), (y, t)) - G^3_\mu((x', 0), (y, t))) \eta(y) f(y) dt dy \right| \\
\leq \int_{B_3(x) \cup B_3} \int_{|t| > 1} \left( \text{osc}_{z \in B_2R} G^3_\mu((z, 0), (y, t)) \right) |f(y)| dt dy \\
\leq \int_{B_3(x) \cup B_3} \int_{|t| > 1} |t|^{-1+\alpha_0} |f(y)| dt dy \lesssim \int_{B_3(x) \cup B_3} |f(y)| dy.
\]

The desired estimate (5.3) for \( d = 2 \) and \( p > 1 \) follows from (5.11), (5.12), and (5.13).

5.2. Proof of Corollary 1.11. Estimate (1.28) is a straightforward combination of (1.25) in Theorem 1.10 and of Schauder interior estimates. We focus on the proof of (1.29), which closely follows its discrete counterpart, cf. [16 Corollary 4].

Step 1. Representation formula for solutions \( u \in H^1(B_R) \) of

\[
\mu u - \nabla \cdot A \nabla u = f \in L^p(B_{2R})
\]

in \( B_{2R} \) for some \( p > d \). Let \( \eta \) be a smooth cutoff function for \( B_{3R} \) in \( B_{2R} \) such that \( |\nabla \eta| \lesssim R^{-1} \).

We claim that for all \( x \in B_{\frac{R}{2}} \),

\[
(5.14) \quad \nabla u(x) = \int_{B_{2R}} \left( \nabla_x G(x, y) (\eta(y) f(y) - \nabla \eta(y) \cdot A(y) \nabla u(y)) + u(y) \nabla \nabla G(x, y) \cdot A(y) \nabla \eta(y) \right) dy
\]
Indeed the Leibniz rule yields

\begin{equation}
\mu\eta u - \nabla \cdot A\nabla (\eta u) = \mu\eta u - \eta \nabla \cdot A\nabla u - \nabla \eta \cdot A\nabla u - \nabla \cdot (uA\nabla \eta).
\end{equation}

The sum of the first two RHS terms equals \(\eta f\) while the other two terms are in \(H^{-1}(\mathbb{R}^d)\) and have compact support. Hence testing (5.15) with \(G\) yields

\begin{equation}
(\eta u)(x) = \int_{B_2R} \left( G_\mu(x,y)(\eta(y)f(y) - \nabla \eta(y) \cdot A(y)\nabla u(y)) + u(y)\nabla_y G_\mu(x,y) \cdot A(y)\nabla \eta(y) \right) dy
\end{equation}

and (5.14) follows by taking the derivative w. r. t. \(G\). Hence testing (5.15) with the terms involving \(\eta\) is well-defined for all \(x \in B_{\frac{2}{3}}\) (so that the Green representation formula follows from mollifying the RHS).

On the one hand, \(G\) \(\in L^\infty(\mathbb{R}^d)\) and \(\nabla G \in L^\infty(\mathbb{R}^d)\) for all \(\varepsilon > 0\) and \(f \in L^p(B_R)\) for some \(p > d\), so that the terms involving \(f\) are well-defined. On the other hand, \(\nabla G_\mu, G_\mu\), and \(\nabla \nabla G_\mu\) are locally square-integrable away from the singularity, and \(\nabla \eta\) vanishes in \(B_{\frac{2}{3}}\) so that the terms involving \(\nabla G_\mu\) or \(\nabla \nabla G_\mu\) and \(u\) or \(\nabla u\) in (5.16) and (5.14) are not singular and are integrable.

**Step 2.** Proof that for \(\alpha = 1 - \frac{d}{p}\),

\begin{equation}
\left( R^\alpha [u]_{C^\alpha(B_R)} \right)^p \lesssim \int_{B_{\frac{2}{3}}} |\nabla G_\mu(x,y)||f(y)| dy
\end{equation}

Indeed, in view of the definition of \(\eta\), (5.14) in Step 1 yields for all \(x \in B_{\frac{2}{3}}\)

\[
|\nabla u(x)| \lesssim \int_{B_{\frac{2}{3}}} |\nabla G_\mu(x,y)||f(y)| dy
\]

where \(A_{\frac{2}{3}, \frac{4}{3}} = \left\{ \frac{4R}{3} < |y| < \frac{5R}{3} \right\}\). The desired estimate (5.17) then follows from Morrey’s inequality

\[
[u]_{C^\alpha(B_R)} = \sup_{x, y \in B_R, x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\alpha} \lesssim \left( \int_{B_R} |\nabla u|^p dy \right)^{\frac{1}{p}}
\]

and the triangle inequality.

**Step 3.** Proof of

\begin{equation}
\left( \sup_{u,f} \left( \frac{R^\alpha \sup_{x,y \in B_R} \frac{|u(x) - w(y)|}{|x - y|^\alpha}}{\sup_{B_{\frac{2}{3}}} |u| + (\frac{1}{B_{2R}} |\nabla^2 f|^p)^{\frac{1}{p}}^2} \right)^p \right)^{\frac{1}{p}}
\end{equation}

\[
\lesssim \left( \int_{B_{\frac{2}{3}}} |\nabla^2 \nabla G_\mu(x,y)|^p dy dx + \int_{B_R} \int_{A_{\frac{2}{3}, \frac{4}{3}}} |\nabla \nabla G_\mu(x,y)|^p dy dx \right)^{\frac{1}{p}}
\]

\[
+ R^{-2p} \int_{B_{2R}} \left( \int_{B_{2R}} |\nabla \nabla G_\mu(x,y)|^p \frac{dy}{dy} \right)^{p-1} dx.
\]
The starting point is (5.17) in Step 2, and we treat each of the three RHS terms separately. For the first term, we use Hölder’s inequality with exponents \((\frac{p}{p-1}, p)\):

\[
\int_{B_R} \left( \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)||f(y)| \, dy \right)^{p-1} \, dx 
\leq R^{-2p} \int_{B_R} \left( \int_{B_{2R}} |\nabla_x G_\mu(x, y)||f(y)| \, dy \right)^p \, dx \int_{B_{2R}} |R^2 f|^p \, dy.
\]

For the second term, we also use Hölder’s inequality with exponents \((p, \frac{p}{p-1})\):

\[
\int_{B_R} \left( \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)||u(y)| \, dy \right)^p \, dx 
\leq \int_{B_R} \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)|^p \, dy \, dx \left( \int_{A_{\frac{4p}{p-1}}} |u|^{\frac{p}{p-1}} \, dy \right)^{p-1}
\leq R^{d(p-2)} \int_{B_R} \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)|^p \, dy \, dx R^d \left( \sup_{B_{2R}} |u| \right)^p.
\]

Likewise, for the third term we have

\[
\int_{B_R} \left( \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)||u(y)| \, dy \right)^p \, dx \leq \int_{B_R} \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)|^p \, dy \, dx 
\times \left( \int_{B_R} |u|^{\frac{p}{p-1}} \, dy \right)^{p-1}.
\]

Since \(p > d \geq 2\), we have \(\frac{p}{p-1} < 2\), so that Jensen’s inequality yields

\[
\left( \int_{B_R} |u|^{\frac{p}{p-1}} \, dy \right)^{p-1} \leq R^{d(\frac{p}{2}-1)} \left( \int_{B_R} |u|^2 \, dy \right)^{\frac{p}{2}}.
\]

By Caccioppoli’s inequality (cf. (5.9)),

\[
\int_{B_R} |\nabla u|^2 \, dy \leq R^{-2} \int_{B_R} |u|^2 \, dy + R^2 \int_{B_{2R}} f^2 \, dy \leq R^{d-2} \sup_{B_{2R}} |u|^2 + R^2 \int_{B_{2R}} f^2 \, dy;
\]

and consequently, by Jensen’s inequality on \(f\) (using that \(p > d \geq 2\)),

\[
\left( \int_{B_R} |\nabla u|^{\frac{p}{p-1}} \, dy \right)^{p-1} \leq R^{d(p-2)-p} \left( R^d \sup_{B_{2R}} |u|^p + \int_{B_{2R}} |R^2 f|^p \, dy \right).
\]

Hence we have proved the following bound for the third RHS term of (5.17):

\[
\int_{B_R} \left( \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)||u(y)| \, dy \right)^p \, dx 
\leq R^{d(p-2)-p} \left( R^d \sup_{B_{2R}} |u|^p + \int_{B_{2R}} |R^2 f|^p \, dy \right) \int_{B_R} \int_{A_{\frac{4p}{p-1}}} |\nabla_x G_\mu(x, y)|^p \, dy \, dx.
\]

This concludes the proof of (5.18) recalling that \(R^d = R^d\).

**Step 4. Conclusion.**
We bound each term of the r. h. s. of (5.18) separately. The first term is bounded by
\[ R^{d(p-2)-p} \int_{B_R} \int_{A_{\frac{1}{2} R} \frac{3}{4} R} |(\nabla_x G(x, y)|^p| dydx \lesssim R^{d(p-2)-p} \int_{B_R} \int_{A_{\frac{1}{2} R} \frac{3}{4} R} |x-y|^{(1-d)p} dydx. \]

For \( x \in B_R \) and \( y \in A_{\frac{1}{2} R} \frac{3}{4} R \), we have that \(|x-y| \geq |y| - |x| \geq \frac{R}{4} \), so that
\[ R^{d(p-2)-p} \int_{B_R} \int_{A_{\frac{1}{2} R} \frac{3}{4} R} |x-y|^{(1-d)p} dydx \lesssim R^{d(p-2)-p+2d+(1-d)p} = 1. \]

Likewise, the second term is bounded by
\[ R^{d(p-2)-p} \int_{B_R} \int_{A_{\frac{1}{2} R} \frac{3}{4} R} |(\nabla_x G_\mu(x, y)|^p| dydx \lesssim R^{d(p-2)-p} \int_{B_R} \int_{A_{\frac{1}{2} R} \frac{3}{4} R} |x-y|^{-dp} dydx \lesssim R^{d(p-2)+2d-dp} = 1. \]

For the third term, we use the triangle inequality in form of
\[ \left\langle \left( \int_{B_{2R}} |\nabla_x G_\mu(x, y)|^{\frac{p}{p-1}} dy \right)^{p-1} \right\rangle \leq \left( \int_{B_{2R}} |(\nabla_x G_\mu(x, y)|^p| \right)^{\frac{2}{p}} dy \right)^{\frac{1}{2}} \]
\[ \lesssim \left( \int_{B_{2R}} |x-y|^{1-d} dy \right)^{p-1} \lesssim R^{d(p-1)+(1-d)p} = R^{d-d}. \]

Hence,
\[ \left\langle R^{-2p} \int_{B_R} \left( \int_{B_{2R}} |\nabla_x G(x, y)|^{\frac{p}{p-1}} dy \right)^{p-1} dx \right\rangle \lesssim R^{-2p+d-p} = R^{-p} \lesssim 1. \]

This completes the proof.

6. PROOFS OF THE AUXILIARY RESULTS

6.1. Proof of Lemma 2.4. The proof is essentially identical to the proof in the discrete case. The only difference lies in the different form of the (LSI). We reproduce the proof for completeness.

**Step 1.** Result for \( p = 1 \).
We claim that for any \( \delta > 0 \) and all \( \zeta(a) \):
\[ (\zeta^2)^{\frac{1}{p}} \leq \left( \exp \left( \frac{2}{\rho^2} \right) + \frac{\rho^2}{2e} \right) (|\zeta|) + \delta \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_a}^{(z)}} \zeta \right)^2 dz \right)^{\frac{1}{p}}, \]
where \( \rho \) denotes the constant in the (LSI), see Definition 1.1. By homogeneity, we may assume \( (\zeta^2) = 1 \). For all real-valued \( \zeta \) we have that
\[ \zeta^2 \leq \begin{cases} \exp \left( \frac{2}{\rho^2} \right) |\zeta| & \text{if } |\zeta| \leq \exp \left( \frac{2}{\rho^2} \right) \\ \rho^2 - \frac{1}{2} \zeta^2 \log \zeta^2 & \text{if } |\zeta| \geq \exp \left( \frac{2}{\rho^2} \right) \end{cases}. \]

Since \( x \log x \) is bounded from below by \( \frac{1}{e} \), we have that \( \frac{2}{e} |\zeta| + \zeta^2 \log \zeta^2 \geq 0 \) for all \( \zeta \). It follows that
\[ \zeta^2 \leq \left( \exp \left( \frac{2}{\rho^2} \right) + \frac{\rho^2}{2e} \right) |\zeta| + \frac{\rho^2}{4} \zeta^2 \log \zeta^2. \]

Hence taking the expectation \( \langle \cdot \rangle \) yields
\[ \langle \zeta^2 \rangle \leq \left( \exp \left( \frac{2}{\rho^2} \right) + \frac{\rho^2}{2e} \right) |\zeta| + \frac{\rho^2}{4} \langle \zeta^2 \log \zeta^2 \rangle. \]
Since $\langle \zeta^2 \rangle = 1$, Young’s inequality yields

$$\langle|\zeta|\rangle \leq \frac{1}{2} \left( \exp\left( \frac{2}{\rho \delta^2} \right) + \frac{\rho \delta^2}{2e} \right) \langle|\zeta|\rangle^2 + \frac{1}{2} \left( \exp\left( \frac{2}{\rho \delta^2} \right) + \frac{\rho \delta^2}{2e} \right)^{−1} \langle|\zeta|\rangle^2,$$

Combining the last two estimates, we deduce

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

Hence (LSI) yields

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

and estimate (6.1) follows from taking the square root and applying the inequality $\sqrt{\zeta} + \zeta \leq e^{\zeta}$ for all numbers $\zeta, \xi \geq 0$.

**Step 2.** We finish the proof of (2.10), i.e. we show that

$$\langle |\zeta|^{2p} \rangle^{\frac{1}{2p}} \leq C(p, \ell, p, \delta) \langle|\zeta|\rangle + \delta \left( \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \, dz \right)^p \right)^{\frac{1}{p}}$$

for general $p \geq 1$. To that end, we apply (6.1) to $\zeta$ replaced by $|\zeta|^p$:

$$\langle |\zeta|^{2p} \rangle \leq C(p, \rho, \delta) \langle|\zeta|^p\rangle^2 + \delta \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} |\zeta|^p \right)^2 \, dz \right),$$

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

$$\langle |\zeta|^{2p} \rangle \leq C(p, \rho, \delta) \langle|\zeta|^p\rangle^2 + 2\delta \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} |\zeta|^p \right)^2 \, dz \right).$$

Now we use that

$$\text{osc}_{A_{B_{\ell}(z)}} |\zeta|^p \leq C(p) \left( \langle|\zeta|^{p-1}\rangle \text{osc}_{A_{B_{\ell}(z)}} \zeta + \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^p \right),$$

which follows from the elementary inequality $|\zeta|^p - \zeta^p \leq C(p)(\zeta^{p-1}|\zeta| - \zeta + |\zeta| - \zeta^p)$ for all numbers $\zeta, \xi > 0$ and the triangle inequality in form of $\text{osc}_{A_{\zeta}(\xi)} |\zeta| \leq \text{osc}_{A_{\zeta}(\xi)} \zeta$. Hence (6.2) yields

$$\langle |\zeta|^{2p} \rangle \leq C(p, \rho, \delta) \langle|\zeta|^p\rangle^2 + 2C(p)\delta \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \, dz \right)^2 \right) + 2C(p)\delta \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \, dz \right).$$

The last term on the right-hand side may be estimated by discreteness, using the argument developed in [12] Proof of Lemma 2.3]. Since every ball $B_{\ell}(z)$, $z \in \mathbb{R}^d$ is contained in the collection $(B_{2\ell}(z'))_{z' \in \frac{1}{2\ell} \mathbb{Z}^d}$, we have that

$$\left\langle \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \, dz \right\rangle \leq C\left( \sum_{z \in \frac{1}{2\ell} \mathbb{Z}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \right)^p.$$

Hence, by discreteness, we find have

$$\left\langle \int_{\mathbb{R}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \, dz \right\rangle \leq C\left( \sum_{z \in \frac{1}{2\ell} \mathbb{Z}^d} \left( \text{osc}_{A_{B_{\ell}(z)}} \zeta \right)^2 \right)^p.$$
Furthermore, Hölder’s inequality followed by Young’s inequality yields
\[ \left\langle |\zeta|^{2p-2} \int_{\mathbb{R}^d} \left( \text{osc}_{A|B_{\ell}(z)} \zeta \right)^{2p} \, dx \right\rangle \leq \left( \left\langle \left( \int_{\mathbb{R}^d} \left( \text{osc}_{A|B_{\ell}(z)} \zeta \right)^{2p} \, dx \right)^p \right\rangle \right)^{\frac{1}{p}} \]
(6.5)
\[ \leq \frac{1}{4C(p)\delta} (\langle |\zeta|^{2p} \rangle + (4C(p)\delta)^{p-1} \left\langle \left( \int_{A|B_{\ell}(z)} \left( \text{osc}_{A|B_{\ell}(z)} \zeta \right)^{2p} \, dx \right)^p \right\rangle) \]
Hence collecting (6.3), (6.4) and (6.5) yields
\[ \langle |\zeta|^{2p} \rangle \leq C(p, \delta) (\langle |\zeta|^{2p} \rangle + 2(2C(p)\delta + (4C(p)\delta)^p) \langle \sum_{z \in \frac{d}{2}Z^d} \left( \text{osc}_{A|B_{\ell}(z)} \zeta \right)^2 \rangle \]
where we have absorbed the second term of (6.5) in the LHS. Since every ball $B_{2\ell}(z')$, $z' \in \frac{2d}{\sqrt{d}} Z^d$ is contained in the collection $(B_{2\ell}(z))_{z \in \frac{d}{2}Z^d}$, we also deduce
\[ \sum_{z \in \frac{d}{2}Z^d} \left( \text{osc}_{A|B_{2\ell}(z)} \zeta \right)^2 \leq \frac{1}{\ell} \int_{\mathbb{R}^d} \left( \text{osc}_{A|B_{2\ell}(z)} \zeta \right)^2 \, dx \]
By redefining $\delta$, we obtain (2.10).

6.2. Proof of Lemma 2.6
Estimate (2.13) is a Meyers’ type estimate, for which we refer the reader to [12, Lemma 2.9]. We split the rest of the proof into four steps. For $d > 2$, (2.13) is a consequence of (2.13) and of Meyers’ estimate, see Step 1. For $d = 2$, however, we need sharper quenched estimates on the decay of local averages of the gradient of the Green function. These are consequence of (2.13) and of Meyers’ estimate, see Step 1. For $d = 2$, however, we need sharper quenched estimates on the decay of local averages of the gradient of the Green function. These are obtained using the De Giorgi-Nash-Moser theory and pointwise bounds on the Green function in Step 2. We then prove (2.14) for all $d \geq 2$ in Step 3. We prove (2.15) in the fourth and last step.

Step 1. Proof of
\[ \int_{R \leq |x-y| < 2R} \int_{|y| < L} |\nabla \nabla G_{\mu}(x,y)|^{2q_1} \, dy \, dx \lesssim R^{-2q_1 \alpha_1} e^{-c \sqrt{R}}, \]
for some $q_1 > 1$ and $\alpha_1 > 0$ and all $R \geq 4L - 1$.
This follows from Meyers’ estimate in the form of: There exists some $q_0 > 1$ depending only on $\lambda$ and $d$ such that for all $1 \leq q \leq q_0$ and all functions $u \in H^1(\mathbb{R}^d)$, $g \in L^2(\mathbb{R}^d, \mathbb{R}^d)$, $f \in L^2(\mathbb{R}^d)$ supported in $B_{2\ell}$ with $L \sim 1$ and related through
\[ -\nabla \cdot A \nabla u = \nabla \cdot g + f, \]
we have
\[ \left( \int_{\mathbb{R}^d} |\nabla u|^{2q_0} \, dy \right)^{\frac{1}{q_0}} \lesssim \left( \int_{\mathbb{R}^d} |g|^{2q_0} \, dy \right)^{\frac{1}{q_0}} + \left( \int_{\mathbb{R}^d} f^{2q_0} \, dy \right)^{\frac{1}{q_0}}. \]
For this estimate we refer the reader to the original article by Meyers [13] or to [13] (4.31) in Proof of Lemma 2.9 (the proof of which is first presented in the continuum setting dealt with here).

Let $\eta : \mathbb{R}^d \to \mathbb{R}$ be such that $\eta \equiv 1$ on $B_L$, $\eta \equiv 0$ on $\mathbb{R}^d \setminus B_{2L}$ and $|\nabla \eta| \lesssim 1$. Assume momentarily that $A$ is smooth, so that $(x, y) \mapsto G_{\mu}(x, y)$ is smooth away from the diagonal $x = y$. Let $i \in \{1, \ldots, d\}$, we apply Meyers’ estimate to the smooth function $u(y) = \eta(y)\nabla_{x_i} G_{\mu}(y, x)$ for $|x| \geq 4L$. Indeed, the defining equation for $G_{\mu}$ yields
\[ -\nabla \cdot A(y) \nabla u(y) = -\mu \eta(y) \nabla_{x_i} G_{\mu}(y, x) - \nabla_{x_i} \eta(y) \cdot A(y) \nabla_{x_i} G_{\mu}(y, x) \]
so that Meyers’ estimate with exponent $q_0 > 1$ takes the form
\[ \left( \int_{B_L} |\nabla_{x_i} G_{\mu}(y, x)|^{2q_0} \, dy \right)^{\frac{1}{q_0}} \lesssim \left( \int_{B_{2L}} |\nabla_{x_i} G_{\mu}(y, x)|^{2q_0} \, dy \right)^{\frac{1}{q_0}} + \left( \int_{B_{2L}} \mu^2 |\nabla_{x_i} G_{\mu}(y, x)|^2 \, dy \right)^{\frac{1}{2}}. \]
By Caccioppoli’s inequality (cf. \[5.3\]) in the proof of Theorem 1.10, since \(L \sim 1\),
\[
\int_{B_{2L}} |\nabla_y \nabla x_i G_\mu(y, x)|^2 \, dy \lesssim \int_{B_{2L}} (\nabla_x G_\mu(y, x))^2 \, dy,
\]
so that by Hölder’s inequality,
\[
\left( \int_{B_L} |\nabla_y \nabla x_i G_\mu(y, x)|^{2q_0} \, dy \right)^{\frac{1}{2q_0}} \lesssim (1 + \mu) \left( \int_{B_L} |\nabla_x G_\mu(y, x)|^{2q_0} \, dy \right)^{\frac{1}{q_0}}.
\]
Taking the \((2q_0)\)th power of this inequality, summing over \(i = 1, \ldots, d\), and integrating over \(\{ R \leq |x| < 2R \} \) yield combined with (2.13) and \(L \sim 1\)
\[
\int_{R \leq |x| < 2R} \int_{|y| < L} |\nabla \nabla G_\mu(y, x)|^{2q_0} \, dy \, dx \lesssim (1 + \mu)^{2q_0} R^d R^{2q_0(1 - d)} \exp \left( - c\sqrt{\mu} R \right).
\]
Since \(q_0 > 1\) and \(d \geq 2\),
\[
d + 2q_0(1 - d) = d(1 - q_0) - q_0(d - 2) \leq -2q_0 \frac{q_0 - 1}{q_0},
\]
\[(6.7)\] implies \[(6.6)\] for \(q_1 = q_0 > 1\) and \(\alpha_1 = \frac{q_0 - 1}{q_0} > 0\). This result carries over to general measurable coefficients \(A\) by density. (Note that for \(d > 2\), this already yields the desired result \[(2.14)\] for all \(1 \leq q \leq q_0\) and \(\alpha_0 = \frac{1}{2}\). The following two steps are forced upon us to deal with \(d = 2\).)

**Step 2.** Quenched estimates on the gradient of the Green function.
In this step we show that there exists a Hölder exponent \(\alpha_2 > 0\) such that for all \(L \sim 1\) and \(|x| \geq R \geq 4L \sim 1\),
\[
\left( \int_{B_L} |\nabla_y G_\mu(x, y)|^2 \, dy \right)^{\frac{1}{2}} \lesssim \frac{e^{-c\sqrt{\mu}|x|}}{|x|^{d - 2 + \alpha_2}}.
\]
Since \(G_\mu(x, y; A) = G_\mu(y, x; A^*)\) (where \(A^*\) is the transpose of \(A\)) and the bounds are uniform wrt \(A \in \Omega\), it is enough to prove \[(6.8)\] with \(\nabla_y G_\mu(x, y)\) replaced by \(\nabla_y G_\mu(y, x)\). We shall first prove \[(6.8)\] for \(d > 2\) and then deduce it for \(d = 2\) from the result for \(d = 3\) following the argument by Avellaneda and Lin already used in Step 6 of the proof of Theorem 1.10. By Caccioppoli’s inequality, for all \(K \in \mathbb{R}\), since \(L \sim 1\),
\[
\int_{B_L} |\nabla_y G_\mu(y, x)|^2 \, dy \lesssim \int_{B_{2L}} (G_\mu(y, x) - K)^2 \, dy + \mu K^2,
\]
so that
\[
\int_{B_L} |\nabla_y G_\mu(y, x)|^2 \, dy \lesssim \left( \operatorname{osc}_{y \in B_{2L}} G_\mu(y, x) \right)^2 + \left( \sqrt{\mu} \int_{B_{2L}} G_\mu(y, x) \, dy \right)^2.
\]
From [10] Theorem 8.22, since \(\{ y : |y| \leq 2L \} \subset \{ y : |y| \leq \| x \|_2 \} \) and
\[
\mu G_\mu(y, x) - \nabla_y \cdot A(y) \nabla_y G_\mu(y, x) = 0 \text{ in } \{ y : |y| \leq \| x \|_2 \},
\]
we learn that there exists \(\alpha_2 > 0\) such that
\[
\operatorname{osc}_{y \in B_{2L}} G_\mu(x, y) \lesssim L_\alpha \| x \|_2^{-\alpha_2} (1 + |x|^2 \mu) \sup_{|y| \leq \| x \|_2} G_\mu(x, y).
\]
Appealing to the pointwise estimate \[(2.2)\] for \(d > 2\) to bound the supremum and using that \(|x - y| \geq \| x \|_2\), this turns into
\[
\operatorname{osc}_{y \in B_{2L}} G_\mu(x, y) \lesssim \frac{|x|^2}{2} (2 - \alpha_2) e^{-c\sqrt{\mu} \| x \|_2}.
\]

Likewise the pointwise estimate (2.2) for $d > 2$ allows one to bound the average in the RHS of (6.8) by

$$
\sqrt{d} \int_{B_2} G_\mu(y,x) dy \lesssim \sqrt{d} \int_{|x|^{2}/2} e^{\alpha \sqrt{|x|^2} + |z|^2} dx \lesssim \frac{|x|^{2}}{2} e^{\alpha \sqrt{|x|^2} + |z|^2}
$$

for some slightly smaller $c > 0$ in the RHS. Hence, (6.8) follows from (6.9) for $d > 2$.

We now turn to $d = 2$, which is the aim of this step, and prove the result by integrating the three-dimensional Green function. Denote by $A^{(2)}$ the coefficients in $\mathbb{R}^{2 \times 2}$, and let $A^{(3)}$ be the block diagonal matrix of $\mathbb{R}^{3 \times 3}$ given by $\text{diag} \left[A^{(2)}, 1\right]$. We denote by $G^{(3)}_\mu$ the Green function associated with $A^{(3)}$ and define a function $G^{(2)}_\mu : \mathbb{R}^2 \times \mathbb{R}^2 \setminus \{x = y\} \to \mathbb{R}^+$, $(x,y) \mapsto G^{(2)}_\mu(x,y)$ as follows:

$$
G^{(2)}_\mu(x,y) = \int_{R} G^{(3)}_\mu(x,z,y,0) dz.
$$

Then, $G^{(2)}_\mu = G_\mu(\cdot, \cdot; A^{(2)})$. By the triangle inequality,

$$
\int_{B_2} |\nabla_y G^{(2)}_\mu(y,x)|^2 dy = \int_{B_2} |\nabla_y G^{(3)}_\mu(y,z,x,0)|^2 dz \lesssim \left( \int_{R} \left( \int_{B_2} |\nabla_y G^{(3)}_\mu(y,z,x,0)|^2 dy \right)^{\frac{1}{2}} dz \right)^2.
$$

Using Cauchy-Schwarz’ inequality locally, this yields

$$
\int_{B_2} |\nabla G^{(2)}_\mu(y,x)|^2 dy \lesssim \left( \int_{R} \left( \int_{|y-z| \leq |x|} |\nabla_y G^{(3)}_\mu(y,z,x,0)|^2 dy dz \right)^{\frac{1}{2}} dz \right)^2.
$$

We then appeal to (6.8) for $d = 3$, which yields

$$
\left( \int_{|y-z| \leq |x|} |\nabla_y G^{(3)}_\mu(y,z',x,0)|^2 dy dz \right)^{1/2} \lesssim \frac{e^{-c\sqrt{|x|^2 + |z|^2}}}{(|x|^2 + |z|^2)^{1/2}}.
$$

Estimating the $z$-integral as follows,

$$
\int_{R} \frac{e^{-c\sqrt{|x|^2 + |z|^2}}}{|x|^2 + |z|^2} dz \lesssim \frac{1}{|x|^2 + |z|^2} dz \lesssim \frac{1}{|x|^2 + |z|^2},
$$

completes the proof of (6.8) for $d = 2$.

Step 3. Proof of (2.11) for all $1 \leq q \leq q_0$.

We first prove that (2.11) holds for $q = 1$ using Caccioppoli’s inequality combined with (6.8), and then conclude by interpolation using Step 1. Assume that $A$ is smooth, so that $\nabla_y G_\mu(y,x)$ is smooth for $x \neq y$. Since for all $i \in \{1, \ldots, d\}$

$$
\mu \nabla_y G_\mu(y,x) = \nabla_x \cdot A(x) \nabla_x \nabla_y G_\mu(y,x) = 0 \text{ in } \{ R/2 \leq |x| < 4R \},
$$

Caccioppoli’s inequality yields

$$
\int_{|y| \leq L} \int_{R \leq |x| < 2R} |\nabla_x \nabla_y G_\mu(y,x)|^2 dx dy \lesssim R^{-2} \int_{|y| \leq L} \int_{R/2 \leq |x| < 4R} (\nabla_y G_\mu(y,x))^2 dx dy.
$$

Combined with (6.8) this turns into

$$
\int_{|y| \leq L} \int_{R \leq |x| < 2R} |\nabla \nabla G_\mu(y,x)|^2 dx dy \lesssim R^{-2} R^d R^{2(2-d)-2\alpha_2} = R^{2-d} R^{-2\alpha_2} e^{-c\sqrt{R}},
$$

that is (2.11) for $q = 1$ and exponent $\alpha_2$. The case of measurable coefficients $A$ follows by density.
Set \( \alpha_0 = \min\{\alpha_1, \alpha_2\} \). An elementary interpolation argument between (6.12) and (6.6) then shows that for all \( 1 \leq q \leq q_0 \),
\[
\int_{|y| \leq L} \int_{R \leq |x| < 2R} |\nabla \nabla G_\mu(y, x)|^{2q} \, dx \, dy \lesssim R^{-2q_0 \alpha_0} e^{-c \sqrt{q} R},
\]
as desired.

**Step 4. Proof of (6.15).**
This is a consequence of Caccioppoli’s inequality and (6.13). Indeed, for all \( 3L \leq |x - y| < 6L \) with \( L \sim 1 \),
\[
(\nabla \nabla G_\mu)_L(x, y) = \int_{B_L(y)} \int_{B_L(x)} |\nabla_{x'} \nabla_{y'} G_\mu(x', y')|^2 \, dx' \, dy' \leq L \sim 1.
\]

### 6.3. Proof of Lemma 2.5
We only prove (2.11), the proof of (2.12) is similar and left to the reader. We split the proof of (2.11) into three steps. In the first step we estimate the oscillation of the mixed second derivative of the Green function. In the second step we control the RHS of this estimate using Lemma 2.7, and we conclude in the third step.

We let \( \hat{A} \) be a coefficient field which coincides with \( A \) outside of \( B_L(z) \), for \( z \in \mathbb{R}^d \), and denote by \( G_\mu \) and \( \tilde{G}_\mu \) the Green functions associated with \( A \) and \( \hat{A} \), respectively, for some \( \mu > 0 \). Set \( \delta G_\mu := \tilde{G}_\mu - G_\mu \).

**Step 1. Proof of (6.13)**

\[
(\nabla \nabla G_\mu)_L(x, y) \lesssim (\nabla \nabla G_\mu)_L(z, y) \begin{cases} \frac{1}{(\nabla \nabla G_\mu)_L(x, z)} & \text{if } |z - x| \leq 6L, \\ (\nabla \nabla G_\mu)_L(x, z) & \text{if } |z - x| > 6L. \end{cases}
\]

for all \( x, y \) with \( |z - y| > 3L \) and \( |x - y| > 3L \).

By density it is enough to take \( A \) and \( \hat{A} \) smooth. Estimate (6.13) follows from the combination of a Green representation formula and an a priori estimate. We start with the former and proceed by regularization. Let \( (\rho_r)_{r>0} \) be a family of smooth non-negative approximations of the Dirac mass with total mass unity and support in \( B_r \). For all \( r > 0 \) and \( y' \in \mathbb{R}^d \), let \( G_{\mu, r}(\cdot, y') \) be the unique weak solution in \( H^1(\mathbb{R}^d) \) of
\[
\mu G_{\mu, r}(x', y') - \nabla_{x'} \cdot A(x') \nabla_{x'} G_{\mu, r}(x', y') = \rho_r(y' - x').
\]

By standard elliptic regularity theory, \( G_{\mu, r} \) is smooth on \( \mathbb{R}^d \times \mathbb{R}^d \). In addition, from the existence/uniqueness theory for the Green function, we learn that for all \( y' \in \mathbb{R}^d \),
\[
G_{\mu, r}(\cdot, y') \overset{r \downarrow 0}{\rightarrow} G_{\mu}(\cdot, y') \quad \text{in } W^{1,1}(\mathbb{R}^d).
\]

Hence, for all \( y' \in \mathbb{R}^d \),
\[
(6.14) \quad G_{\mu, r}(\cdot, y') \overset{r \downarrow 0}{\rightarrow} G_{\mu}(\cdot, y') \quad \text{in } W^{1,1}(\mathbb{R}^d).
\]

For all \( y' \in \mathbb{R}^d \), \( \delta G_{\mu, r}(\cdot, y') \) is a classical solution of
\[
\mu \delta G_{\mu, r}(x', y') - \nabla_{x'} \cdot \hat{A}(x') \nabla_{x'} \delta G_{\mu, r}(x', y') = \nabla_{x'} \cdot (\hat{A} - A)(x') \nabla_{x'} G_{\mu, r}(x', y').
\]

Since the RHS has compact support, \( \delta G_{\mu, r}(\cdot, y') \) satisfies the Green representation formula for all \( x', y' \in \mathbb{R}^d \),
\[
(6.16) \quad \delta G_{\mu, r}(x', y') = \int_{\mathbb{R}^d} \nabla_{x'} \tilde{G}_\mu(x', z') \cdot (\hat{A} - A)(z') \nabla_{z'} G_{\mu, r}(z', y') \, dz'.
\]
Provided $|z - x'| > 2L$ and $|z - y'| > 2L$, standard quenched estimates on the gradient of the Green function yield:

$$
\sup_{z' \in B_L(z)} |\nabla_{x'} \delta G_\mu(x', z')| \lesssim \sup_{z' \in B_L(z)} |x' - z'|^{2-d} \lesssim L^{2-d} \sim 1.
$$

Hence, using (6.14) and (6.15), as $r \downarrow 0$, the Green representation formula (6.16) turns into

$$
\delta G_\mu(x', y') = \int_{\mathbb{R}^d} \nabla_{x'} \delta G_\mu(x', z') \cdot (\hat{A} - A)(z') \nabla_{z'} G_\mu(z', y') dz'.
$$

for all $|z - x'| > 2L$ and $|z - y'| > 2L$. Since $G_\mu$ and $\hat{G}_\mu$ are smooth away from the diagonal, we may differentiate twice (6.17), which yields for all $|z - x'| > 2L$ and $|z - y'| > 2L$,

$$
\nabla_{x'} \delta G_\mu(x', y') = \int_{\mathbb{R}^d} \nabla_{x'} \delta G_\mu(x', z') \cdot (\hat{A} - A)(z') \nabla_{z'} G_\mu(z', y') dz'.
$$

Recall that $|z - x| > 3L$ and $|z - y| > 3L$. Integrating (6.13) over $x' \in B_L(x)$ and $y' \in B_L(y)$, we obtain by Cauchy-Schwarz’ inequality

$$
(\nabla_{x'} \delta G_\mu)_L(x, y) \lesssim (\nabla_{x'} \hat{G}_\mu)_L(x, z) (\nabla_{x'} G_\mu)_L(z, y).
$$

We turn now to the a priori estimate. Let $|y' - z| > 2L$. Then, $\delta G_\mu(\cdot, y')$ is the unique distributional solution in $W^{1,1}(\mathbb{R}^d)$ of

$$
\mu \delta G_\mu(x', y') - \nabla_{x'} \hat{A}(x') \nabla_{x'} \delta G_\mu(x', y') = \nabla_{x'} \cdot (\hat{A} - A)(x') \nabla_{x'} G_\mu(x', y').
$$

Since $G_\mu$ is smooth away from the diagonal, the RHS is smooth with compact support, so that $\delta G_\mu(\cdot, y')$ is a classical solution. We then differentiate the equation with respect to $y_i'$ for $i \in \{1, \ldots, d\}$:

$$
\mu \nabla_{y_i'} \delta G_\mu(x', y') - \nabla_{x'} \hat{A}(x') \nabla_{x'} \nabla_{y_i'} \delta G_\mu(x', y') = \nabla_{x'} \cdot (\hat{A} - A)(x') \nabla_{x'} \nabla_{y_i'} G_\mu(x', y').
$$

Since the RHS is smooth and has compact support, $\nabla_{y_i'} \delta G_\mu(\cdot, y') \in H^1(\mathbb{R}^d)$, and we may test the weak formulation of the equation with the solution itself. This yields

$$
\int_{\mathbb{R}^d} |\nabla_{x'} \delta G_\mu(x', y')|^2 dx' \lesssim \int_{B_L(z)} |\nabla_{x'} \delta G_\mu(x', y')||\nabla_{x'} G_\mu(x', y')| dx',
$$

which, by Young’s inequality, turns into

$$
\int_{\mathbb{R}^d} |\nabla_{x'} \delta G_\mu(x', y')|^2 dx' \lesssim \int_{B_L(z)} |\nabla_{x'} G_\mu(x', y')|^2 dx'.
$$

We are in position to conclude. On the one hand, integrating (6.20) over $y' \in B_L(y)$ yields

$$
(\nabla_{x'} \delta G_\mu)_L(x, y) \lesssim (\nabla_{x'} G_\mu)_L(z, y).
$$

On the other hand, assume that $|z - x| > 3L$. Denote by $G^*_\mu$, $\hat{G}^*_\mu$ and $\delta G^*_\mu$ the Green functions associated with $A^*$, $\hat{A}^*$, and their difference. Estimate (6.20) takes the form

$$
\int_{\mathbb{R}^d} |\nabla_{x'} \delta G^*_\mu(y', x')|^2 dx' \lesssim \int_{B_L(z)} |\nabla_{x'} G^*_\mu(y', x')|^2 dx',
$$

so that by integration over $y' \in B_L(x)$ and by the symmetry properties of the Green function,

$$
(\nabla_{x'} \delta G^*_\mu)_L(x, z) = (\nabla_{x'} \delta G^*_\mu)_L(z, x) \lesssim (\nabla_{x'} G^*_\mu)_L(z, x) = (\nabla_{x'} G^*_\mu)_L(x, z).
$$

Hence by the triangle inequality, the estimate (6.19) for $|z - x| > 3L$ turns into

$$
(\nabla_{x'} \delta G^*_\mu)_L(x, y) \lesssim (\nabla_{x'} G^*_\mu)_L(x, z)(\nabla_{x'} G^*_\mu)_L(z, y).
$$

The claim (6.13) follows from the combination of (6.21) and (6.22).
We first show that for all 

\[
\sum_{|z-x'| \leq 6L} \quad \text{uniformly wrt } (6.25)
\]

On each dyadic annulus, 

\[
\text{For larger } |z-x'|, \text{ we decompose } \{ z : |z-x'| > 6L \} \text{ into dyadic annuli:}
\]

\[
\int_{|z-x'| > 6L} (|z-x'| + 1)^{2\alpha} \left( \frac{1}{|\nabla \nabla G_{\mu}} \right)^{2\alpha} (x', z) \, dz \leq \int_{|z-x'| > 6L} (|z-x'| + 1)^{2\alpha} (x', z) \, dz.
\]

On each dyadic annulus, 

\[
\int_{2^6L < |z-x'| < 2^{n+1}6L} (|z-x'| + 1)^{2\alpha} (x', z) \, dz
\]

which we bound using Jensen’s inequality and (2.13) as 

\[
\sum_{n \in \mathbb{N}} 2^{-q \alpha n} = 2^{-q \alpha n},
\]

uniformly wrt \( x' \in \mathbb{R}^d \). The combination of (6.24), (6.25), and (6.26) yields the claim (6.23) since \( \sum_{n \in \mathbb{N}} 2^{-q \alpha n} \lesssim 1 \).

Step 3. Conclusion.

We first show that for all \( |x-y| > 6L \) and all \( p \) large enough, we have

\[
\left( \int_{|z-y| > 3L} \frac{\text{osc}_{B_L(z)} (\nabla \nabla G_{\mu})}{B_L(z)} (x, y) \right)^2 |x-y|^{2d} e^{2c \sqrt{|z-y|}} \, dz \right)^p 
\]

We claim that it is enough to prove that

\[
\left( \int_{|z-y| > 3L} \frac{\text{osc}_{B_L(z)} (\nabla \nabla G_{\mu})}{B_L(z)} (x, y) \right)^2 |x-y|^{2d} dz \right)^p
\]

\[
\left( \int_{|z-y| > 3L} \frac{\text{osc}_{B_L(z)} (\nabla \nabla G_{\mu})}{B_L(z)} (x, y) \right)^2 |x-y|^{2d} dz \right)^p
\]
To this aim we have to prove that the corresponding integral on the LHS of (6.28), this time over \(|z - y| \leq |z - x|\), is bounded by the RHS of (6.28). Indeed, (6.28) for \(G^*_\mu\) with \(x\) and \(y\) switched takes the form after using the symmetry properties of the Green function

\[
\left\langle \left( \int_{|z-x|>|z-y|} \left( \frac{\osc_{B_L(z)} (\nabla \nabla G_{\mu})_L(x,y)}{B_L(z)} \right)^2 |y-x|^2dz \right)^p \right\rangle \leq \sup_{z,x:|z-x|>3L} \left\{ |z-x|^{2pd} \left\langle \left( \nabla \nabla G_{\mu} \right)_L^2(x,z) \right\rangle \right\}.
\]

The conclusion follows by stationarity since

\[
\sup_{z,x:|z-x|>3L} \left\{ |z-x|^{2pd} \left\langle \left( \nabla \nabla G_{\mu} \right)_L^2(x,z) \right\rangle \right\} = \sup_{z:|z|>3L} \left\{ |z|^{2pd} \left\langle \left( \nabla \nabla G_{\mu} \right)_L^2(z,0) \right\rangle \right\} = \sup_{z,y:|z-y|>3L} \left\{ |z-y|^{2pd} \left\langle \left( \nabla \nabla G_{\mu} \right)_L^2(z,y) \right\rangle \right\}.
\]

It is therefore enough to prove (6.28).

For \(|z - y| \geq |z - x|\), we have \(|z - y| \geq \frac{|z-x|}{16} \geq 3L\), so that taking the supremum over \(\tilde{A}\) (by a density argument the supremum can be taken on smooth fields \(A\)) in the estimate (6.13) of Step 1 yields

\[
\int_{|z-y|>|z-x|} \left( \frac{\osc_{B_L(z)} (\nabla \nabla G_{\mu})_L(x,y)}{B_L(z)} \right)^2 dz \lesssim \int_{|z-y|>|z-x|} (\nabla \nabla G_{\mu})_L^2(z,y) \begin{cases} 1 & \text{if } |z-x| \leq 6L \\ \mu & \text{if } |z-x| > 6L \end{cases} dz.
\]

We smuggle in the weight \((|z - x| + 1)^\alpha\) and apply Hölder’s inequality with exponents \((p,q)\) for some \(p > 1\) to be fixed below:

\[
\left\langle \left( \int_{|z-y|>|z-x|} (\nabla \nabla G_{\mu})_L^2(z,y) \begin{cases} 1 & \text{if } |z-x| \leq 6L \\ \mu & \text{if } |z-x| > 6L \end{cases} dz \right)^p \right\rangle \lesssim \left\langle \left( \int_{|z-y|>3L} (|z-x| + 1)^{2\alpha q} (\nabla \nabla G_{\mu})_L^2(x,z) \begin{cases} 1 & \text{if } |z-x| \leq 6L \\ \mu & \text{if } |z-x| > 6L \end{cases} dz \right)^q \right\rangle^{\frac{p}{q}} \times \left\langle \int_{|z-y|>|z-x|} (|z-x| + 1)^{-2\alpha p} (\nabla \nabla G_{\mu})_L^2(z,y) dz \right\rangle^{\frac{q}{p}}.
\]

By (6.23) in Step 2, the first term on the r. h. s. is bounded uniformly wrt \(A\) as long as \(1 \leq q \leq q_0\), i.e. \(p = \frac{q}{q-1} \geq \frac{q_0}{q_0-1} =: p_0\). Hence, using that \(|z - y| \geq |x - y|/2\), this yields

\[
\left\langle \left( |x-y|^{2d} e^{2\sqrt{|m|}x|y|} \int_{|z-y|>|z-x|} (\nabla \nabla G_{\mu})_L^2(z,y) \begin{cases} 1 & \text{if } |z-x| \leq 6L \\ \mu & \text{if } |z-x| > 6L \end{cases} dz \right)^p \right\rangle^{\frac{1}{p}} \lesssim \left\langle \int_{|z-y|>|z-x|} (|z-x| + 1)^{-2\alpha_0 p} |z-y|^{2pd} e^{2\sqrt{|m|}x|y|} (\nabla \nabla G_{\mu})_L^2(z,y) dz \right\rangle^{\frac{1}{p}}.
\]
In the following, we require that
\[ k := v \]

Without loss of generality we may assume that \( q < \lambda \) Moser iteration. We follow the proof of [10, Theorem 8.17] and mainly focus on the differences.

Since by construction \( v \) the derivative of \( u \) is enough to prove the claim for the positive part

\[ (6.32) \lambda \beta \]

so that

\[ \text{Estimate (6.28), which implies (6.27), is now a consequence of (6.29) and (6.30).} \]

6.4. Proof of Lemma 2.10 for \( p < \infty \). The proof consists in a minor modification of the usual Moser iteration. We follow the proof of [10] Theorem 8.17 and mainly focus on the differences. Without loss of generality we may assume that \( q < \frac{d}{2} \). Up to multiplying the equation by \( -1 \) it is enough to prove the claim for the positive part \( \lambda^{\frac{d}{2}} = \max(0, u) \) of \( u \). Set \( \bar{u} = u^+ + k \), where \( k := \|f\|_{L^p(B_2)} \) with \( p \) given in the statement. We test the equation (2.19) with the test function \( v = \eta^2(\bar{u}^3 - k^3) \geq 0 \), where \( \beta > 0 \) and \( \eta \) is a smooth cut-off function for \( B_1 \) in \( B_2 \) with \( 0 \leq \eta \leq 1 \). In the following, we require that

\[ 0 < \beta < \frac{(q - 1)d}{d - 2q}. \]

The derivative of \( v \) is given by

\[ \nabla v = 2\eta(\bar{u}^3 - k^3)\nabla \eta + \eta^2\beta\bar{u}^{\beta-1}\nabla \bar{u}. \]

Since by construction \( \mu\bar{u}(\bar{u}^3 - k^3) \geq 0 \) and either \( \nabla \bar{u} \) and \( \bar{u}^3 - k^3 \) vanish or \( \nabla \bar{u} \) equals \( \nabla u \), equation (2.19) with test function \( v \) yields

\[ 0 = \int_{\mathbb{R}^d} (\mu \bar{u}^3 + \nabla v \cdot A \nabla u - vf) \, dx \]

\[ = \int_{\mathbb{R}^d} \left( \mu\eta^2(\bar{u}^3 - k^3) + \beta\eta^2\bar{u}^{\beta-1}\nabla \bar{u} \cdot A \nabla \bar{u} + (2\eta(\bar{u}^3 - k^3))\nabla \eta \cdot A \nabla \bar{u} - vf \right) \, dx \]

\[ \geq \int_{\mathbb{R}^d} \left( \lambda\beta\eta^2\bar{u}^{\beta-1}\nabla \bar{u}^2 - 2|\nabla \eta|(\bar{u}^3 - k^3)|\nabla \bar{u}| - \eta^2(\bar{u}^3 - k^3)|f| \right) \, dx \]

\[ \geq \int_{\mathbb{R}^d} \left( \beta\eta^2\bar{u}^{\beta-1}\nabla \bar{u}^2 - 2|\nabla \eta|\bar{u}^3|\nabla \bar{u}| - \eta^2\bar{u}^3|f| \right) \, dx. \]

By Young’s inequality,

\[ \int_{\mathbb{R}^d} 2|\nabla \eta|\bar{u}^2|\nabla \bar{u}| \, dx \leq \frac{\lambda\beta}{2} \int_{\mathbb{R}^d} \eta^2\bar{u}^{\beta-1}|\nabla \bar{u}|^2 \, dx + \frac{2}{\lambda\beta} \int_{\mathbb{R}^d} |\nabla \eta|^2\bar{u}^\beta+1 \, dx, \]

so that

\[ \frac{\lambda\beta}{2} \int_{\mathbb{R}^d} \eta^2\bar{u}^{\beta-1}|\nabla \bar{u}|^2 \leq \frac{2}{\lambda\beta} \int_{\mathbb{R}^d} |\nabla \eta|^2\bar{u}^\beta+1 \, dx + \frac{1}{\lambda\beta} \int_{\mathbb{R}^d} \eta^2\bar{u}^3|f| \, dx. \]
So far, the computations are identical to the usual Moser iteration. Here comes the difference: Let \( \chi = \frac{d}{q-2} \) if \( d > 2 \) (or fix any \( 1 < \chi < +\infty \) if \( d = 2 \)) and let \( s > 1 \) be such that \( 1 = \frac{1}{s} + \frac{\beta}{(\beta+1)\chi} \).

Then, the choice (6.31) implies that \( 1 \leq s < q \). Indeed,

\[
\frac{1}{s} - \frac{1}{q} = 1 - \frac{\beta}{(\beta+1)\chi} - \frac{1}{q} > \frac{q - 1}{q} - \frac{\beta}{\beta + 1} \cdot \frac{d - 2}{d} = \frac{-\beta(d - 2q) + d(q - 1)}{q(\beta + 1)d} > 0.
\]

We now treat the second RHS term of (6.32). H"older’s inequality on \( k \)-s

\[
\int_{\mathbb{R}^d} \eta^2 \tilde{u}^\beta \, |f| \, dx \leq \left( \int_{\mathbb{R}^d} \eta^2 \tilde{u}^{(\beta+1)\chi} \, dx \right)^{\frac{\beta}{(\beta+1)\chi}} \left( \int_{\mathbb{R}^d} \eta^2 \tilde{u}^\chi |f|^\chi \, dx \right)^{\frac{1}{\chi}},
\]

where we recall that \( k = \|f\|_{L^\chi(B_d)} \). By Young’s inequality we thus have for all \( \varepsilon > 0 \):

\[
\int_{\mathbb{R}^d} \eta^2 \tilde{u}^\beta \, |f| \, dx \leq \varepsilon \left( \int_{\mathbb{R}^d} \eta^2 \tilde{u}^{(\beta+1)\chi} \, dx \right)^{\frac{\beta}{(\beta+1)\chi}} + C k^{\beta+1},
\]

where \( C \) depends now in addition on \( \varepsilon \). Combined with (6.32) this yields

\[
\frac{\lambda \beta}{2} \int_{\mathbb{R}^d} \eta^2 \tilde{u}^{\beta - 1} |\nabla \tilde{u}|^2 \, dx \leq \frac{2}{\lambda \beta} \int_{\mathbb{R}^d} |\nabla \tilde{u}|^2 \tilde{u}^{\beta - 1} \, dx + C k^{\beta+1} + \varepsilon \left( \int_{\mathbb{R}^d} \eta^2 \tilde{u}^{(\beta+1)\chi} \, dx \right)^{\frac{1}{\beta}}.
\]

Next we introduce another function \( w := \tilde{u}^{\frac{\beta + 1}{\beta}} \) and rewrite this inequality as

\[
(6.33) \quad \lambda \int_{\mathbb{R}^d} \eta^2 |\nabla w|^2 \, dx \leq \frac{2}{\lambda \beta} \int_{\mathbb{R}^d} |\nabla \eta|^2 \tilde{u}^{(\beta+1)\chi} \, dx + C k^{\beta+1} + \varepsilon \left( \int_{\mathbb{R}^d} |\eta w|^{2\chi} \, dx \right)^{\frac{1}{\chi}}.
\]

This yields

\[
\int_{\mathbb{R}^d} |\nabla (\eta w)|^2 \, dx \leq \frac{2}{\lambda \beta} \int_{\mathbb{R}^d} |\nabla \eta|^2 \tilde{u}^{(\beta+1)\chi} \, dx + C k^{\beta+1} + \varepsilon \left( \int_{\mathbb{R}^d} |\eta w|^{2\chi} \, dx \right)^{\frac{1}{\chi}}.
\]

By the Sobolev embedding, this turns into

\[
C_{\text{Sob}} \left( \int_{\mathbb{R}^d} |\eta w|^{2\chi} \, dx \right)^{\frac{1}{\chi}} \leq \frac{2}{\lambda \beta} \int_{\mathbb{R}^d} |\nabla \eta|^2 \tilde{u}^{(\beta+1)\chi} \, dx + C k^{\beta+1} + \varepsilon \left( \int_{\mathbb{R}^d} |\eta w|^{2\chi} \, dx \right)^{\frac{1}{\chi}},
\]

so that for \( \varepsilon \) small enough (and only depending on \( d, \lambda \) and \( q \)), we have

\[
\left( \int_{\mathbb{R}^d} |\eta w|^{2\chi} \, dx \right)^{\frac{1}{\chi}} \leq C \int_{\mathbb{R}^d} |\nabla \eta|^2 \tilde{u}^{(\beta+1)\chi} \, dx.
\]

This corresponds to the usual Moser iteration (albeit the dependence of the constants on \( \beta \) is worse), and yields the desired result for \( p = \chi (\beta + 1) \). We can then iterate by increasing \( \beta \) to yield bounds as long as \( \beta < \frac{(q-1)d}{d-2q} \). In this case any exponent of the form \( p = (\beta + 1)\chi \) can be attained, which yields

\[
\frac{1}{p} > \frac{d - 2q}{((q-1)d + d - 2q)\chi} = \frac{d - 2q}{dq} = \frac{1}{q} - \frac{2}{d}.
\]

as claimed. Note that (unlike the usual Moser iteration) the dependence of the constants on \( \beta \) does not matter since we only need to iterate finitely many times in order to reach \( p < +\infty \).
6.5. Proof of Lemma 2.9. Let \( u \) and \( \tilde{u} \) be solutions of \((2.16)\) with coefficient fields \( A \) and \( \tilde{A} \), respectively, where the coefficients coincide outside the ball \( B_\ell(z) \). Their difference \( \delta u \) solves
\[
\begin{align*}
\mu \delta u - \nabla \cdot \tilde{A} \nabla \delta u &= -\nabla \cdot (\tilde{A} - A) \nabla u, \\
\mu \delta u - \nabla \cdot A \nabla \delta u &= -\nabla \cdot (A - \tilde{A}) \nabla \tilde{u}.
\end{align*}
\]

Step 1. Preliminary result and proof of
\[
\sup_{x, x' \in \mathbb{R}^d} \sup_{A \in \Omega} \left\| \int_{B_\ell(x)} \nabla y' \tilde{G}_\mu(y, y') \, dy \right\|_{L^2(B_{2\ell}(x'))} \lesssim 1.
\]
To see this, we note that
\[
v : y' \mapsto \int_{B_\ell(x)} \tilde{G}_\mu(y, y') \, dy \text{ solves } \mu v - \nabla \cdot \tilde{A} \nabla v = \frac{1}{\ell^d(x)} \chi_{B_{2\ell}(x)},
\]
where \( \chi_D \) denotes the characteristic function of the set \( D \subseteq \mathbb{R}^d \), that is, a regularized version of the defining equation for the Green function without singularity. The proof that \( \int_{B_{2\ell}(x')} |\nabla v|^2 \, dy' \) is bounded and only depends on \( \ell \) and \( \lambda \) is similar to the corresponding proof of [13, Corollary 2.3] in the discrete case (since there is no singularity to be taken care of).

Step 2. Proof of \((2.19)\) for \(|x - z| > 6\ell\).
The Green function representation formula associated with \((6.35)\) yields
\[
u(x) - \tilde{u}(x) = \int_{B_\ell(z)} \nabla z G_\mu(x, z') \cdot (\tilde{A}(z') - A(z')) \nabla \tilde{u}(z') \, dz'.
\]
Hence, by the triangle inequality and Hölder’s inequality,
\[
\|\nu - \tilde{u}\|_{L^{\lambda_1}(B_\ell(z))} \lesssim \|\nabla_2 G_\mu\|_{L^{\lambda_1}(B_\ell(z), L^2(B_\ell(z)))} \|\nabla \tilde{u}\|_{L^{\lambda_1}(B_\ell(z))},
\]
where we recall that \( (\nabla \tilde{u})(z) = \|\nabla \tilde{u}\|_{L^2(B_\ell(z))} \). Since \(|x - z| > 6\ell\), for all \( i \in \{1, \ldots, d\} \) the function \( x \mapsto \nabla_{x_i} G_\mu(x, z) \) is in the kernel of \( (\mu - \nabla \cdot A \nabla) \) in \( B_{2\ell}(x) \) for all \( z \in B_\ell(z) \) and Lemma 2.10 implies that
\[
\|\nabla_2 G_\mu\|_{L^{\lambda_1}(B_\ell(z), L^2(B_\ell(z)))} \lesssim \|\nabla_2 G_\mu\|_{L^2(B_{2\ell}(x) \times B_\ell(z))} \lesssim (\nabla G_\mu)_{2\ell}(x, z).
\]
On the other hand, an energy estimate based on \((6.34)\) yields
\[
(\nabla \tilde{u})_{2\ell}(z) \lesssim (\nabla u)_{2\ell}(z).
\]
Estimate \((2.19)\) for \(|x - z| > 6\ell\) is proved.

Step 3. Proof of \((2.19)\) for \(|x - z| \leqslant 6\ell\).
Let \( x \) be fixed such that \(|x - z| \leqslant 6\ell\). We shall consider a third coefficient field \( A_0 \in \Omega \) such that \( A_0|_{\mathbb{R}^d \setminus B_{6\ell}(z)} = A|_{\mathbb{R}^d \setminus B_{6\ell}(z)} \), \( A_0|_{B_{6\ell}(z)} = 1d \), and denote by \( u_0 \) the associated solution of \((2.16)\) with coefficient fields \( A_0 \). We denote the local averages of \( u \), \( \tilde{u} \), and \( u_0 \) around \( x \) by
\[
\bar{u} = \int_{B_\ell(x)} u \, dy, \quad \bar{\tilde{u}} = \int_{B_\ell(x)} \tilde{u} \, dy, \quad \text{and} \quad \bar{u}_0 = \int_{B_\ell(x)} u_0 \, dy.
\]
The triangle inequality yields
\[
\|u - \bar{u}\|_{L^{\lambda_1}(B_\ell(z))} \lesssim \|u - \bar{u}_0\|_{L^{\lambda_1}(B_\ell(z))} + \|\bar{u} - \bar{u}_0\|_{L^{\lambda_1}(B_\ell(z))}.
\]
By the De Giorgi-Nash-Moser estimate of Lemma 2.10 with \( p = \lambda_1 \) and \( q = \lambda_2 \) (note \( \frac{1}{\lambda_1} < \frac{2}{d} + \frac{1}{\lambda_2} \) and \( u - \bar{u}_0 \) solves the same equation as \( u \) with the addition of \(-\mu \bar{u}_0 \) on the RHS), the triangle
inequality, and Poincaré’s inequality on $B_M^+(x)$, the first term yields

$$(6.37) \| u - \tilde{u}_0 \|_{L^1(B_r(x))} \lesssim \left( \int_{B_M^+(x)} |u(y) - \tilde{u}_0|^2 \, dy \right)^{\frac{1}{2}} + \| \mu \tilde{u}_0 + f \|_{L^2(B_M^+(x))}$$

$$\lesssim \left( \int_{B_M^+(x)} |u(y) - \bar{u}|^2 \, dy \right)^{\frac{1}{2}} + \| \bar{u}_0 - \bar{u} \|_{L^2(B_M^+(x))} + \| f \|_{L^2(B_M^+(x))}$$

Likewise,

$$(6.38) \| \tilde{u} - \tilde{u}_0 \|_{L^1(B_r(x))} \lesssim \left( \int_{B_M^+(x)} |\nabla \tilde{u}|^2 \, dy \right)^{\frac{1}{2}} + \| \bar{u}_0 - \bar{u} \|_{L^2(B_M^+(x))} + \| f \|_{L^2(B_M^+(x))}.$$
and therefore by definition of \(u_\ell\) and \(f_\ell\) and Cauchy-Schwarz’ inequality,
\[
\mu^2 \bar{u}_0^2 = \mu^2 \left( \int_{B_{2\ell}(x)} u_0 dy \right)^2 \lesssim \left( \int_{B_{2\ell}(x)} |f| dy \right)^2 + \int_{B_{2\ell}(x)} |\nabla u_0|^2 dy.
\]

The combination of (6.36), (6.37), (6.38), (6.39), (6.40), and (6.41) then yields
\[
\|u - \bar{u}\|_{L^2(B_{\ell}(x))} \lesssim (\nabla u)_{3\ell}(z) + \|f\|_{L^2(B_{2\ell}(x))},
\]
which proves (2.19) for \(|x-z| \leq 6\ell\).

Acknowledgments

The first author acknowledges financial support from the European Research Council under the European Community Seventh Framework Programme (FP7/2014-2019 Grant Agreement QUANTHOM 335410).

References

[1] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. ArXiv e-prints, June 2014.
[2] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. Comm. Pure and Applied Math., 40(6):803–847, 1987.
[3] G. Bal. Central limits and homogenization in random media. Multiscale Model. Simul., 7:677–702, 2008.
[4] P. Bella, D. Marahrens, and F. Otto. Corrector estimates for systems in stochastic homogenization. In preparation.
[5] J.C. Conlon and A. Naddaf. On homogenization of elliptic equations with random coefficients. Elec. Journal of Probability, 5:1–58, 2000. Paper no. 9.
[6] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. J. Reine Angew. Math., 368:28–42, 1986.
[7] T. Delmotte and J.-D. Deuschel. On estimating the derivatives of symmetric diffusions in stationary random environments, with applications to the \(\nabla \phi\) interface model. Probab. Theory Rel. Fields, 133:358–390, 2005.
[8] M. Duerinckx and A. Gloria. Weighted nonlocal Poincaré and logarithmic-Sobolev inequalities. In Preparation.
[9] R. Figari, E. Orlandi, and G. Papanicolaou. Mean field and Gaussian approximation for partial differential equations with random coefficients. SIAM J. Appl. Math., 42:1069–1077, 1982.
[10] D. Gilbarg and N.S. Trudinger. Elliptic partial differential equations of second order. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
[11] A. Gloria. Fluctuation of solutions to linear elliptic equations with noisy diffusion coefficients. Comm. Partial Differential Equations, 38(2):304–338, 2013.
[12] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. Preprint.
[13] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. Ann. Probab., 39(3):779–856, 2011.
[14] M. Grüber and K.-O. Widman. The Green function for uniformly elliptic equations. Manuscripta Math., 37:303–342, 1982.
[15] Q. Han and F. Lin. Elliptic partial differential equations. Courant Institute of Mathematical Sciences, New York, 1997.
[16] D. Marahrens and F. Otto. Annealed estimates on the Green’s function. MPI Preprint 69/2012.
[17] D. Marahrens and F. Otto. On annealed elliptic Green function estimates. arXiv:1401.2859v2.
[18] A. Meyers. An \(L^p\)-estimate for the gradient of solutions of second order elliptic divergence equations. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3), 17(3):189–206, 1963.
[19] G.C. Papanicolaou and S.R.S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In Random fields, Vol. I, II (Esztergom, 1979), volume 27 of Colloq. Math. Soc. János Bolyai, pages 835–873. North-Holland, Amsterdam, 1981.

(A. Gloria) Université Libre de Bruxelles (ULB), Brussels, Belgium, and Project-team SIMPAF, INRIA Lille - Nord Europe, Villeneuve d’Ascq, France
E-mail address: antoine.gloria@ulb.ac.be

(D. Marahrens) Max-Planck-Institute for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig, Germany
E-mail address: daniel.marahrens@mis.mpg.de