



Higher-Order Linearization and Regularity in Nonlinear Homogenization

SCOTT ARMSTRONG, SAMUEL J. FERGUSON & TUOMO KUUSI

Communicated by A. FIGALLI

Abstract

We prove large-scale C^∞ regularity for solutions of nonlinear elliptic equations with random coefficients, thereby obtaining a version of the statement of Hilbert’s 19th problem in the context of homogenization. The analysis proceeds by iteratively improving three statements together: (i) the regularity of the homogenized Lagrangian \bar{L} , (ii) the commutation of higher-order linearization and homogenization, and (iii) large-scale $C^{0,1}$ -type regularity for higher-order linearization errors. We consequently obtain a quantitative estimate on the scaling of linearization errors, a Liouville-type theorem describing the polynomially-growing solutions of the system of higher-order linearized equations, and an explicit (heterogenous analogue of the) Taylor series for an arbitrary solution of the nonlinear equations—with the remainder term optimally controlled. These results give a complete generalization to the nonlinear setting of the large-scale regularity theory in homogenization for linear elliptic equations.

Contents

1. Introduction	632
1.1. Motivation: Quantitative Homogenization for Nonlinear Equations	632
1.2. Background: Large-Scale Regularity Theory and Its Crucial Role in Quantitative Homogenization	634
1.3. Summary of the Results Proved in this Paper	637
1.4. Assumptions and Notation	641
1.5. Statement of the Main Results	642
2. Regularity Estimates for the Effective Lagrangian	651
2.1. The First-Order Correctors and Linearized Correctors	651
2.2. Smoothness of \bar{L}	663
2.3. Sublinearity of Correctors	665
3. Quantitative Homogenization of the Linearized Equations	669
3.1. Stationary, Finite Range Coefficient Fields	670

3.2. Setup of the Proof of Theorem 1.2	674
3.3. Estimates on Size of the w_m^ε and \bar{w}_m	676
3.4. The Mesoscopic Parameters k and l	678
3.5. The Minimal Scales	679
3.6. The Proof of the Local Stationary Approximation	679
3.7. Homogenization Estimates for the Approximating Equations	683
3.8. Homogenization of the Locally Stationary Equation	687
4. Large-Scale $C^{0,1}$ -Type Estimates for Linearized Equations	694
4.1. Approximation of w_{n+2} by Smooth Functions	694
4.2. Excess Decay Iteration for w_{n+2}	696
4.3. Improvement of Spatial Integrability	702
5. Large-Scale $C^{0,1}$ -Type Estimate for Linearization Errors	706
5.1. Excess Decay Iteration for ξ_{n+1}	706
5.2. Improvement of Spatial Integrability	713
6. Sharp Estimates of Linearization Errors	713
7. Liouville Theorems and Higher Regularity	715
Appendix A: Deterministic Regularity Estimates	724
References	740

1. Introduction

1.1. Motivation: Quantitative Homogenization for Nonlinear Equations

This article is concerned with nonlinear, divergence-form, uniformly elliptic equations of the form

$$-\nabla \cdot (D_p L(\nabla u(x), x)) = 0 \quad \text{in } U \subseteq \mathbb{R}^d, \quad d \geq 2. \quad (1.1)$$

The Lagrangian $L(p, x)$ is assumed to be uniformly convex and regular in p and possess some mild regularity in x . Furthermore, L is a stochastic object: it is sampled by a probability measure \mathbb{P} which is statistically stationary and satisfies a unit range of dependence. This essentially means that $x \mapsto L(\cdot, x)$ is a random field, valued in the space of uniformly convex functions, the law of which is independent of x (or, to be precise, periodic in x ; see Section 1.4 for the assumptions).

The objective is to describe the statistical behavior of the solutions of (1.1), with respect to the probability measure \mathbb{P} , on *large length scales*. In other words, we want to understand what the solution u looks like in the case that the “macroscopic” domain U is very large relative to “microscopic” scale, which is the correlation length scale of the coefficients (taken to the unit scale).

At a qualitative level, a satisfactory characterization of the solutions, in the regime in which the ratio of these two length scales is large, is provided by the principle of homogenization. First proved in this context by Dal Maso and Modica [11, 12], it asserts roughly that a solution of (1.1) is, with probability approaching one, close in L^2 (relative to its size in L^2) to a solution of a deterministic equation of the form

$$-\nabla \cdot (D_p \bar{L}(\nabla u_{\text{hom}})) = 0 \quad \text{in } U \quad (1.2)$$

for an *effective Lagrangian* \bar{L} which is also uniformly convex and $C^{1,1}$.

This result is of great importance, from both the theoretical and computation points of view, since the complexity of the homogenized equation (1.2) is significantly less than that of (1.1) as it is both deterministic and spatially homogeneous. It hints that the structure of (1.1) should, on large domains and with high probability, possess some of the structure of a constant coefficient equation and thus we may expect it to be more amenable to our analysis than the worst-possible heterogeneous equation of the form (1.1). In other words, since the Lagrangian L is sampled by a probability measure \mathbb{P} with nice ergodic properties, rather than given to us by the devil, can we expect its solutions to have a nicer structure? In order to answer this kind of question, we need to build a *quantitative* theory of homogenization.

To be of practical use, the principle of homogenization must be made quantitative. We need to have answers to questions such as:

- How large does the ratio of scale separation need to be before we can be reasonably sure that solutions of (1.1) are close to those of (1.2)? In other words, what is the size of a typical error in the homogenization approximation in terms of the size of U ?
- Can we estimate the probability of the unlikely event that the error is large?
- What is $D_p \bar{L}$ and how can we efficiently compute it? How regular can we expect it to be? Can we efficiently compute its derivatives?
- Can we describe the fluctuations of the solutions?

In this paper we show that (1.1) has a C^∞ structure. In particular, we will essentially answer the third question posed above by demonstrating that the effective Lagrangian \bar{L} is as regular as $L(\cdot, x)$ with estimates for its derivatives. We will identify the higher derivatives of \bar{L} as the homogenized coefficients of certain *linearized* equations and give quantitative homogenization estimates for these, implicitly indicating a computational method for approximating them and thus a Taylor approximation for \bar{L} . Finally, we will prove large-scale $C^{k,1}$ type estimates for solutions of (1.1), for $k \in \mathbb{N}$ as large as can be expected from the regularity assumptions on L , a result analogous to Hilbert's 19th problem, famously given for spatially homogeneous Lagrangians by De Giorgi and Nash. For this classical result, the main roadblock was in obtaining the continuity of solutions to linear equations, as it is completely straightforward to differentiate the equation apply the Schauder theory, which was previously known. In the case of homogenization, the situation is reversed, as it is less clear how one should “differentiate the equation” since literally doing so would produce negative powers of ε , even if the coefficients were smooth. Our analysis resolves this difficulty and reveals the interplay between these three seemingly different kinds of results: (i) the regularity of \bar{L} , (ii) the homogenization of linearized equations and the commutability of homogenization and linearization, and (iii) the large-scale regularity of the solutions. These three statements must be proved together, iteratively in the parameter $k \in \mathbb{N}$ which represents the degree of regularity of \bar{L} , the order of the linearized equation, and the order of the $C^{k,1}$ estimate.

1.2. Background: Large-Scale Regularity Theory and Its Crucial Role in Quantitative Homogenization

In the last decade, beginning with the work of Gloria and Otto [16, 18], a quantitative theory of homogenization has been developed to give precise answers to questions like the ones stated in the previous subsection. Until now, most of the progress has come in the case of *linear* equations

$$-\nabla \cdot \mathbf{a}(x)\nabla u = 0, \quad (1.3)$$

which corresponds to the special case $L(p, x) = \frac{1}{2}p \cdot \mathbf{a}(x)p$ of (1.1), where $\mathbf{a}(x)$ is a symmetric matrix. By now there is an essentially complete quantitative theory for linear equations, and we refer to the monograph [4] and the references therein for a comprehensive presentation of this theory. Quantitative homogenization for the nonlinear equation (1.1) has a comparatively sparse literature; in fact, the only such results of which we are aware are those of [5, 6] (see also [4, Chapter 11]), our previous paper [1] and a new paper of Fischer and Neukamm [14] which was posted to arXiv as we were finishing the present article.

Quantitative homogenization is inextractably linked to regularity estimates on the solutions of the heterogeneous equation. This is not surprising when we consider that the homogenized flux $D_p \bar{L}(\nabla u_{\text{hom}})$ should be related to the spatial average (say, on some mesoscopic scale) of the heterogeneous flux $D_p L(\nabla u(x), x)$. In order for spatial averages of the latter to converge nicely, we need to have bounds. It could be unfortunate and lead to a very slow rate of homogenization if, for instance, the flux was concentrated on sets of very small measure which percolate only on very large scales. To rule this out we need much better estimates: ideally, we would like to know that the size of the flux on small scales is the same as on large scales, which amounts to a $W^{1,\infty}$ estimate on solutions.

Unfortunately, solutions of equations with highly oscillating coefficients do not possess very strong regularity, in general. Indeed, the best deterministic elliptic regularity estimate for solutions of (1.1), which does not degenerate as the size of the domain U becomes large, is $C^{0,\delta}$ in terms of Hölder regularity (the De Giorgi–Nash estimate) and $W^{1,2+\delta}$ in terms of Sobolev regularity (the Meyers estimate). The tiny exponent $\delta > 0$ in each estimate becomes small as the ellipticity ratio becomes large (see [4, Example 3.1]) and thus both estimates are far short of the desired regularity class $W^{1,\infty} = C^{0,1}$.

One of the main insights in the quantitative theory of homogenization is that, compared to a generic (“worst-case”) L , solutions of the equation (1.1) have much better regularity if L is sampled by an ergodic probability measure \mathbb{P} . This is an effect of homogenization itself: on large scales, (1.1) should be a “small perturbation” of (1.2), and therefore better regularity estimates for the former can be inherited from the latter. This is the same idea used to prove the classical Schauder estimates. In the context of homogenization, the result states that there exists a random variable \mathcal{X} (sometimes called the minimal scale) which is finite almost surely such that, for every $\mathcal{X} < r < \frac{1}{2}R$ and solution $u \in H^1(B_R)$ of (1.1) with $U = B_R$, we have the estimate

$$\int_{B_r} |\nabla u|^2 \leq C \left(1 + \int_{B_R} |\nabla u|^2 \right). \quad (1.4)$$

Here C depends only on dimension and ellipticity and $f_U := \frac{1}{|U|} \int_U$ denotes the mean of a function in U . If we could send $r \rightarrow 0$ in (1.4), it would imply that

$$|\nabla u(0)|^2 \leq C \left(1 + \int_{B_R} |\nabla u|^2 \right),$$

which is a true Lipschitz estimate, same estimate in fact as holds for the homogenized equation (1.2). As (1.4) is valid only for $r > \mathcal{X}$, it is sometimes called a “large-scale $C^{0,1}$ estimate” or a “Lipschitz estimate down to the microscopic scale”. This estimate, first demonstrated in [6] in the stochastic setting, is a generalization of the celebrated result in the case of (non-random) periodic coefficients due to Avellaneda and Lin [7]. Of course, it then becomes very important to quantify the size of \mathcal{X} . The estimate proved in [6], which is essentially optimal, states that \mathcal{X} is bounded up to “almost volume-order large deviations”: for every $s < d$ and $r \geq 1$,

$$\mathbb{P}[\mathcal{X} > r] \leq C \exp(-cr^s). \quad (1.5)$$

Here the constant C depends only on s, d , and the ellipticity. A proof of this large-scale regularity estimate together with (1.5) can be found in [4, Chapter 3] in the linear case and in [4, Chapter 11] for the nonlinear case. The right side of (1.5) represents the probability of the unlikely event that the L sampled by \mathbb{P} will be a “worst-case” L in the ball of radius r . A proof of the optimality of (1.5) can be found in [4, Section 3.6].

This large-scale regularity theory introduced in [6] was further developed in the case of (1.3) in [2, 3, 5, 15, 17] and now plays an essential role in the quantitative theory of stochastic homogenization. Whether one employs functional inequalities [13, 17] or renormalization arguments [2, 3, 19], it is a crucial ingredient in the proof of the optimal error estimates in homogenization for (1.3): see the monograph [4] and the references therein for a complete presentation of these developments.

The large-scale $C^{0,1}$ estimate is, from one point of view, the best regularity one can expect solutions of (1.1) or (1.3) to satisfy: since the coefficients are rapidly oscillating, there is no hope for the gradient to exhibit continuity on the macroscopic scale. However, as previously shown in the periodic case in [7, 9], the solutions of the linear equation (1.3) still have a C^∞ structure. To state what we mean, let us first think of an (interior) $C^{k,1}$ estimate not as a pointwise bound on the $(k+1)$ th derivatives of a function, but as an estimate on how well a function may be approximated on small balls by a k th order polynomial. By Taylor’s theorem, these are of course equivalent in the following sense:

$$\left| \nabla^{k+1} u(0) \right| \simeq \limsup_{r \rightarrow 0} \frac{1}{r^{k+1}} \inf_{p \in \mathcal{P}_k} \|u - p\|_{\underline{L}^2(B_r)},$$

where \mathcal{P}_k denotes the set of polynomials of degree at most k and we use the notation $\|w\|_{\underline{L}^2(U)} := \left(\int_U |w|^2 \right)^{\frac{1}{2}}$ to denote the volume-normalized $L^2(U)$ norm.

Thus the interior $C^{k,1}$ estimate for a harmonic function can be stated in the form: for any harmonic function u in B_R and any $r \in (0, \frac{1}{2}R]$,

$$\inf_{p \in \mathcal{P}_k} \|u - p\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R}\right)^{k+1} \inf_{p \in \mathcal{P}_k} \|u - p\|_{\underline{L}^2(B_R)}. \tag{1.6}$$

Moreover, the infimum on the left side may be replaced by the set of *harmonic* polynomials of degree at most k .

As we cannot expect a solution of (1.1) to have regularity beyond $C^{0,1}$ in a classical (pointwise) sense, in order to make sense of a $C^{k,1}$ estimate for the heterogeneous equation (1.3) we need to replace the set of polynomials by a heterogeneous analogue. The classical Liouville theorem says that the set of harmonic functions which grow like $o(|x|^{k+1})$ is just the set of harmonic polynomials of degree at most k . This suggests that we should use the (random) vector space

$$\mathcal{A}_k := \left\{ u \in H^1_{\text{loc}}(\mathbb{R}^d) : -\nabla \cdot \mathbf{a} \nabla u = 0, \limsup_{r \rightarrow \infty} r^{k+1} \|u\|_{\underline{L}^2(B_r)} = 0 \right\}.$$

We think of these as “ $\mathbf{a}(x)$ -harmonic polynomials.” It turns out that one can prove that, \mathbb{P} -almost surely, this set is finite dimensional and has the same dimension as the set of at most k th order harmonic polynomials. In fact, one can match any $\mathbf{a}(x)$ -harmonic polynomial to an $\bar{\mathbf{a}}$ -harmonic polynomial in the highest degree, and vice versa. In close analogy to (1.6), the statement of large-scale $C^{k,1}$ regularity is then as follows: there exists a minimal scale \mathcal{X} satisfying (1.5) such that, for any $R > 2\mathcal{X}$ and any solution $u \in H^1(B_R)$ of $-\nabla \cdot \mathbf{a} \nabla u = 0$, we can find $\phi \in \mathcal{A}_k$ such that, for every $r \in [\mathcal{X}, \frac{1}{2}R]$,

$$\|u - \phi\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R}\right)^{k+1} \inf_{\phi \in \mathcal{A}_k} \|u - \phi\|_{\underline{L}^2(B_R)}; \tag{1.7}$$

see [4, Theorem 3.8] for the full statement, which was first proved in the periodic setting by Avellaneda and Lin [8]. Subsequent versions of this result, which are based on the ideas of [7, 8] in their more quantitative formulation given in [6], were proved in various works [2, 15, 17], with the full statement here given in [3, 10].

In all of its various forms, higher regularity in stochastic homogenizations is based on the simple idea that solutions of the heterogeneous equation should be close to those of the homogenized equation, which should have much better regularity. In the case of the linear equation (1.3), it does not a large leap, technically or philosophically, to go from (1.4) to (1.7). Indeed, to gain control over higher derivatives, one just needs to differentiate the equation (not in the microscopic parameters, of course, but in macroscopic ones) and, luckily, since the equation is linear, this does not change the equation. Roughly speaking, the idea is analogous to bootstrapping the regularity of a constant-coefficient, linear equation by differentiating it. Therefore the estimate (1.7) is perhaps not too surprising once one has the large-scale $C^{0,1}$ estimate in hand.

1.3. Summary of the Results Proved in this Paper

The situation is very different in the nonlinear case. When one differentiates the equation (again, in a macroscopic parameter) we get a new equation, namely the first-order linearized equation. If we want to apply a large-scale regularity result to this equation, we must first (quantitatively) homogenize it! Achieving higher-order regularity estimates requires repeatedly differentiating the equation, which lead to a hierarchy of linearized equations requiring homogenization estimates.

Let us be a bit more explicit. The gradient of the homogenized Lagrangian \bar{L} is given by the well-known formula

$$\begin{aligned} D_p \bar{L}(p) &= \mathbb{E} \left[\int_{[0,1]^d} D_p L(p + \nabla \phi_p(x), x) \, dx \right] \\ &= \lim_{r \rightarrow \infty} \int_{B_r} D_p L(p + \nabla \phi_p(x), x) \, dx, \end{aligned} \tag{1.8}$$

where ϕ_p is the first-order corrector with slope $p \in \mathbb{R}^d$, that is, it satisfies

$$\begin{cases} -\nabla \cdot D_p L(p + \nabla \phi_p(x), x) = 0 & \text{in } \mathbb{R}^d, \\ \nabla \phi_p \text{ is } \mathbb{Z}^d\text{-stationary, } \mathbb{E} \left[\int_{[0,1]^d} \nabla \phi_p(x) \, dx \right] = 0. \end{cases}$$

The limit in the second line of (1.8) is to be understood in a \mathbb{P} -almost sure sense, and it is a consequence of the ergodic theorem, which states that macroscopic averages of stationary fields must converge to their expectations. The formula (1.8) says that $D_p \bar{L}(p)$ is the flux per unit volume of the first-order corrector with slope $p \in \mathbb{R}^d$. It naturally arises when we homogenize the nonlinear equation. We can try to show that $\bar{L} \in C^2$ by formally differentiating (1.8), which leads to the expression

$$D_p \partial_{p_i} \bar{L}(p) = \mathbb{E} \left[\int_{[0,1]^d} D_p^2 L(p + \nabla \phi_p(x), x) (e_i + \nabla (\partial_{p_i} \phi_p(x))) \, dx \right].$$

If we define the linearized coefficients around $\ell_p + \phi_p$ by

$$\mathbf{a}_p(x) := D_p^2 L(p + \nabla \phi_p(x), x)$$

and put $\psi_{p,e_i}^{(1)} := \partial_{p_i} \phi_p(x)$, then we see that $\psi_{p,e_i}^{(1)}$ is the first-order corrector with slope e_i of the linear equation with coefficients \mathbf{a}_p :

$$\begin{cases} -\nabla \cdot \mathbf{a}_p(e_i + \nabla \psi_{p,e_i}^{(1)}) = 0 & \text{in } \mathbb{R}^d, \\ \nabla \psi_{p,e_i}^{(1)} \text{ is } \mathbb{Z}^d\text{-stationary, } \mathbb{E} \left[\int_{[0,1]^d} \nabla \psi_{p,e_i}^{(1)}(x) \, dx \right] = 0. \end{cases}$$

We call $\psi_{p,e}^{(1)}$ a *first-order linearized corrector*. Moreover, we have the formula

$$D_p \partial_{p_i} \bar{L}(p) = \mathbb{E} \left[\int_{[0,1]^d} \mathbf{a}_p(x) (e_i + \nabla \psi_{p,e_i}^{(1)}(x)) \, dx \right] = \bar{\mathbf{a}}_p e_i.$$

That is, “linearization and homogenization commute”: the Hessian of \bar{L} at p should be equal to the homogenized coefficients $\bar{\mathbf{a}}_p$ corresponding to the linear equation with coefficient field $\mathbf{a}_p = D_p^2 L(p + \nabla\phi_p(\cdot), \cdot)$. This reasoning is only formal, but since the right side of the above formula for $D_p^2 \bar{L}$ is well-defined (and only needs qualitative homogenization), we should expect that it should be rather easy to confirm it rigorously. Moreover, while quantitative homogenization of the original nonlinear equation gives us a $C^{0,1}$ estimate, we should expect that quantitative homogenization of the linearized equation should give us a $C^{0,1}$ estimate for *differences of solutions* and a large-scale $C^{1,1}$ estimate for solutions. This is indeed the case and was proved in our previous paper [1], where we were motivated by the goal of obtaining this regularity estimate for differences of solutions in anticipation of its important role in the proof of optimal quantitative homogenization estimates. Indeed, in the very recent preprint [14], Fischer and Neukamm showed that this estimate can be combined with spectral gap-type assumptions on the probability measure to obtain quantitative bounds on the first-order correctors which are optimal in the scaling of the error.

We may attempt to differentiate the formula for the homogenized Lagrangian a second time, with the ambition of obtaining a C^3 estimate for \bar{L} , a $C^{2,1}$ estimate for solutions and a higher-order improvement of our $C^{0,1}$ estimate for differences (which will be a $C^{0,1}$ estimate for *linearization errors*): we get

$$D_p \partial_{p_i} \partial_{p_j} \bar{L}(p) = \mathbb{E} \left[\int_{[0,1]^d} \mathbf{a}_p(x) \nabla \psi_{p,e_i,e_j}^{(2)}(x) \, dx \right] + \mathbb{E} \left[\int_{[0,1]^d} D^3 L(p + \nabla\phi_p(x), x) \left(e_i + \nabla\psi_{p,e_i}^{(1)} \right) \left(e_j + \nabla\psi_{p,e_j}^{(1)} \right) \, dx \right].$$

If we define the vector field

$$\mathbf{F}_{2,p,e_i,e_j}(x) := D^3 L(p + \nabla\phi_p(x), x) \left(e_i + \nabla\psi_{p,e_i}^{(1)} \right) \left(e_j + \nabla\psi_{p,e_j}^{(1)} \right),$$

then we see that $\psi_{p,e_i,e_j}^{(2)}$ is the first-order corrector with slope zero of the linear equation

$$-\nabla \cdot \mathbf{a}_p \nabla \psi_{p,e_i,e_j}^{(2)} = \nabla \cdot \mathbf{F}_{2,p,e_i,e_j} \quad \text{in } \mathbb{R}^d, \tag{1.9}$$

and the formula for the tensor $D_p^3 \bar{L}$ becomes

$$D_p \partial_{p_i} \partial_{p_j} \bar{L}(p) = \mathbb{E} \left[\int_{[0,1]^d} \mathbf{a}_p(x) \nabla \psi_{p,e_i,e_j}^{(2)}(x) + \mathbf{F}_{2,p,e_i,e_j}(x) \, dx \right] = \bar{\mathbf{F}}_{2,p,e_i,e_j}, \tag{1.10}$$

the corresponding homogenized coefficient. Unlike the case for the Hessian of \bar{L} , we should *not* expect this formula to be valid under qualitative ergodic assumptions! Indeed, the qualitative homogenization of (1.9) and thus the validity of (1.10) requires that \mathbf{F}_{2,p,e_i,e_j} belong to L^2 , in the sense that $\mathbb{E} \left[\int_{[0,1]^d} |\mathbf{F}_{2,p,e_i,e_j}(x)|^2 \, dx \right] < \infty$,

and due to the product of two first-order correctors we only have L^1 -type integrability for \mathbf{F}_{2,p,e_i,e_j} .¹ This is a serious problem which can only be fixed using the large-scale regularity theory for the first-order linearized equation, and a suitable bound on the minimal scale \mathcal{X} , thereby obtaining a bound in $L^\infty(\square_0)$ and hence $L^4(\square_0)$ for $\nabla\psi_{p,e_i}^{(1)}$, with at least a fourth moment in expectation. Note that this also requires some regularity of the Lagrangian L on the smallest scale.

If we differentiate the equation once more in an effort to prove that $\bar{L} \in C^4$, we will be faced with similar difficulties, this time with the more complicated vector field

$$\begin{aligned} \mathbf{F}_{p,e_i,e_j,e_k}^{(3)}(x) &:= D^3L(p + \nabla\phi_p(x), x) \left(e_i + \nabla\psi_{p,e_i}^{(1)} \right) \nabla\psi_{p,e_j,e_k}^{(2)} \\ &\quad + D^3L(p + \nabla\phi_p(x), x) \left(e_j + \nabla\psi_{p,e_j}^{(1)} \right) \nabla\psi_{p,e_i,e_k}^{(2)} \\ &\quad + D^3L(p + \nabla\phi_p(x), x) \left(e_k + \nabla\psi_{p,e_k}^{(1)} \right) \nabla\psi_{p,e_i,e_j}^{(2)} \\ &\quad + D^4L(p + \nabla\phi_p(x), x) \left(e_i + \nabla\psi_{p,e_i}^{(1)} \right) \\ &\quad \quad \left(e_j + \nabla\psi_{p,e_j}^{(1)} \right) \left(e_k + \nabla\psi_{p,e_k}^{(1)} \right). \end{aligned}$$

Notice that the last term has three factors of the first-order linearized correctors instead of two, and is thus “even further” from being obviously L^2 than was $\mathbf{F}_{p,e_i,e_j}^{(2)}$. Homogenizing the third-order linearized equation will therefore require large-scale regularity estimates for both the first-order and second-order linearized equations, and one can now see the situation will get worse as the order increases beyond three. Moreover, proving quantitative homogenization for these equations will also require some smoothness of the homogenized coefficients associated to the lower-order equations, due to the needs for the homogenized solutions to be smooth in quantitative two-scale expansion arguments.

This suggests a bootstrap argument for progressively and simultaneously obtaining (i) the smoothness of \bar{L} ; (ii) the higher-order large-scale regularity of solutions (and solutions of linearized equations); and (iii) the homogenization of the higher-order linearized equations and commutation of homogenization and higher-order linearization. The point of this paper is to formalize this idea and thereby give a proof of “Hilbert’s 19th problem for homogenization”. Here is a rough schematic of the argument, as discussed above, which comes in three distinct steps, discussed in more detail below:

- Homogenization & large-scale $C^{0,1}$ regularity for the linearized equations up to order $N \implies \bar{L} \in C^{2+N}$.
- $\bar{L} \in C^{2+N}$ and large-scale $C^{0,1}$ regularity for the linearized equations up to order $N \implies$ homogenization for the linearized equations up to order $N + 1$.
- $\bar{L} \in C^{2+N}$, large-scale $C^{0,1}$ regularity for the linearized equations up to order N and homogenization for the linearized equations up to order $N + 1 \implies$ large-scale $C^{0,1}$ regularity for the linearized equations up to order $N + 1$.

¹ To be pedantic, we actually have $L^{1+\delta}$ -type integrability for a tiny $\delta > 0$ by the Meyers estimate, but this does not help.

The three implications above are the focus of most of the paper and their proofs are given in Sections 2–5.

Once this induction argument is completed, we consequently obtain a full $C^{k,1}$ -type large scale regularity estimate for solutions of the original nonlinear equation, generalizing (1.7). The main question becomes what the replacement for \mathcal{A}_k should be, that is, what the “polynomials” should be. We show that these are certain solutions of the system of linearized equations (linearized around a first-order corrector) exhibiting polynomial-type growth, which we classify by providing a Liouville-type result which is part of the statement of the theorem (see the discussion between the statements of Theorems 1.5 and 1.6, below, for a definition of these spaces, which are denoted by W_n^p). The resulting theorem we obtain, which is a version of the statement of Hilbert’s 19th problem in the context of homogenization, provides a very precise description of the solutions of (1.1) in terms of the first-order correctors and the correctors of a hierarchy of linear equations.

We also obtain, as a corollary, the improvement of the scaling of linearization errors—which is closely related to the regularity of solutions. To motivate this result, suppose we are given two solutions $u, v \in H^1(B_R)$ of (1.1) in a large ball ($R \gg 1$) which are close to each other in the sense that

$$\|\nabla u - \nabla v\|_{L^2(B_R)} \ll \|\nabla u\|_{L^2(B_R)}. \tag{1.11}$$

Suppose that we attempt to approximate the difference $u - v$ by the solution $w \in H^1(B_{R/2})$ of the linearized problem

$$\begin{cases} -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w) = 0 & \text{in } B_R, \\ w = v - u & \text{on } \partial B_R. \end{cases}$$

Then we may ask the question of how small we should expect the first-order linearization error to be. The best answer that we have from deterministic elliptic regularity estimates is that there exists a small exponent $\delta(d, \Lambda) > 0$ such that

$$\frac{\|u - v - w\|_{L^2(B_R)}}{\|u\|_{L^2(B_R)}} \leq C \left(\frac{\|u - v\|_{L^2(B_R)}}{\|u\|_{L^2(B_R)}} \right)^{1+\delta}. \tag{1.12}$$

This can be easily proved for instance using the Meyers gradient $L^{2+\delta}$ estimate, and it is sharp in the sense that it is not possible to do better than the very small exponent δ . We can say roughly that the space of solutions of (1.1) is a $C^{1,\delta}$ manifold, but no better. Of course, if L does not depend on x , or if R is of order one and L is smooth in both variables (p, x) , then one expects to have the estimate above with $\delta = 1$ and to be able to prove more precise estimates using higher-order linearized equations. In fact, this is essentially a reformulation of the statement of Hilbert’s 19th problem (indeed—see Appendix E, where we give a proof of Hilbert’s 19th problem in its classical formulation by following this line of reasoning).

In this paper we also prove a large-scale version of the quadratic response to first-order linearization in the context of homogenization, which states that (1.12) holds with $\delta = 1$ whenever R is larger than a random minimal scale. Moreover,

we prove a full slate of higher-order versions of this result: see Corollary 1.4. These results roughly assert that, with probability one, the large-scale structure of solutions of (1.1) resembles that of a smooth manifold.

In the next two subsections, we state our assumptions and give the precise statements of the results discussed above.

1.4. Assumptions and Notation

In this subsection, we state the standing assumptions in force throughout the paper.

We fix following global parameters: the dimension $d \in \mathbb{N}$ with $d \geq 2$, a constant $\Lambda \in [1, \infty)$ measuring the ellipticity, an integer $N \in \mathbb{N}$ with $N \geq 1$ measuring the smoothness of the Lagrangian, and constants $M_0, K_0 \in [1, \infty)$. For short, we denote

$$\text{data} := (d, \Lambda, N, M_0, K_0).$$

This allows us to, for instance, denote constants C which depend on $(d, \Lambda, N, M_0, K_0)$ by simply $C(\text{data})$ instead of $C(d, \Lambda, N, M_0, K_0)$.

The probability space is the set Ω of all Lagrangians $L : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, written as a function of $(p, x) \in \mathbb{R}^d \times \mathbb{R}^d$, satisfying the following conditions:

(L1) L is $2+N$ times differentiable in the variable p and, for every $k \in \{2, \dots, 2+N\}$, the function $D_p^k L$ is uniformly Lipschitz in both variables and satisfies

$$\left[D_p^k L \right]_{C^{0,1}(\mathbb{R}^d \times \mathbb{R}^d)} \leq K_0. \tag{1.13}$$

For $k = 1$, we assume that, for $z \in \mathbb{R}^d$,

$$\left[D_p L(z, \cdot) \right]_{C^{0,1}(\mathbb{R}^d)} \leq K_0(1 + |z|). \tag{1.14}$$

(L2) L is uniformly convex in the variable p : for every $p \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$,

$$I_d \leq D_p^2 L(p, x) \leq \Lambda I_d.$$

(L3) $D_p L(0, \cdot)$ is uniformly bounded:

$$\| D_p L(0, \cdot) \|_{L^\infty(\mathbb{R}^d)} \leq M_0.$$

We define Ω to be the set of all such Lagrangians L :

$$\Omega := \{ L : L \text{ satisfies (L1), (L2) and (L3)} \}.$$

Note that Ω depends on the fixed parameters $(d, \Lambda, N, M_0, K_0)$. It is endowed with the following family of σ -algebras: for each Borel subset $U \subseteq \mathbb{R}^d$, define

$$\begin{aligned} \mathcal{F}(U) &:= \text{the } \sigma\text{-algebra generated by the family of random variables} \\ L &\mapsto L(p, x), \quad (p, x) \in \mathbb{R}^d \times U. \end{aligned}$$

The largest of these is denoted by $\mathcal{F} := \mathcal{F}(\mathbb{R}^d)$.

We assume that the law of the ‘‘canonical Lagrangian’’ L is a probability measure \mathbb{P} on (Ω, \mathcal{F}) satisfying the following two assumptions:

(P1) \mathbb{P} has a unit range of dependence: for all Borel subsets $U, V \subseteq \mathbb{R}^d$ such that $\text{dist}(U, V) \geq 1$,

$$\mathcal{F}(U) \text{ and } \mathcal{F}(V) \text{ are } \mathbb{P}\text{-independent.}$$

(P2) \mathbb{P} is stationary with respect to \mathbb{Z}^d -translations: for every $z \in \mathbb{Z}^d$ and $E \in \mathcal{F}$,

$$\mathbb{P}[E] = \mathbb{P}[T_z E],$$

where the translation group $\{T_z\}_{z \in \mathbb{Z}^d}$ acts on Ω by $(T_z L)(p, x) = L(p, x+z)$.

The expectation with respect to \mathbb{P} is denoted by \mathbb{E} .

Since we will be often concerned with measuring the stretched exponential moments of the random variables we encounter, the following notation is convenient: for every $\sigma \in (0, \infty)$, $\theta > 0$, and random variable X on Ω , we write

$$X \leq \mathcal{O}_\sigma(\theta) \iff \mathbb{E} \left[\exp \left(\left(\frac{X_+}{\theta} \right)^\sigma \right) \right] \leq 2.$$

This is essentially notation for an Orlicz norm on (Ω, \mathbb{P}) . Some basic properties of this notation is given in [4, Appendix A].

1.5. Statement of the Main Results

We begin by introducing the higher-order linearized equations. These can be computed by hand, as we did for the second and third linearized equations in Section 1.3, but it is convenient to work with more compact formulas. Observe that, by Taylor’s formula with remainder, we have, for every $n \in \{0, \dots, N + 1\}$,

$$\left| D_p L(p_0 + h, x) - \sum_{k=0}^n \frac{1}{k!} D_p^{k+1} L(p_0, x) h^{\otimes k} \right| \leq \frac{CK}{(n + 1)!} |h|^{n+1}.$$

Define, for $p, x, h_1, \dots, h_N \in \mathbb{R}^d$ and $t \in \mathbb{R}$,

$$\mathbf{G}(t, p, h_1, \dots, h_N, x) := \sum_{k=2}^{N+1} \frac{1}{k!} D_p^{k+1} L(p, x) \left(\sum_{j=1}^N \frac{t^j}{j!} h_j \right)^{\otimes k}.$$

Also define, for each $m \in \{1, \dots, N + 1\}$ and $p, x, h_1, \dots, h_{m-1} \in \mathbb{R}^d$,

$$\mathbf{F}_m(p, h_1, \dots, h_{m-1}, x) := (\partial_t^m \mathbf{G})(0, p, h_1, \dots, h_{m-1}, 0, \dots, 0, x). \tag{1.15}$$

Observe that $\mathbf{F}_1 \equiv 0$ by definition (that is, the right side of the first linearized equation is zero, as we have already seen).

Our first main result concerns the regularity of the effective Lagrangian \bar{L} and states that it has essentially the same regularity in p as we assumed for L .

Theorem 1.1. (Regularity of \bar{L}). *For every $\beta \in (0, 1)$, the effective Lagrangian \bar{L} belongs to $C_{\text{loc}}^{2+\mathbf{N},\beta}(\mathbb{R}^d)$ and, for every $\mathbf{M} \in [1, \infty)$, there exists $C(\beta, \mathbf{M}, \text{data}) < \infty$ such that*

$$\left\| D^2 \bar{L} \right\|_{C^{\mathbf{N},\beta}(B_{\mathbf{M}})} \leq C.$$

In view of Theorem 1.1, we may introduce homogenized versions of the above functions. We define, for every $p \in \mathbb{R}^d$, $\{h_i\}_{i=1}^{\mathbf{N}} \subseteq \mathbb{R}^d$ and $\delta \in \mathbb{R}$,

$$\bar{\mathbf{G}}(t, p, h_1, \dots, h_{\mathbf{N}}) := \sum_{k=2}^{\mathbf{N}+1} \frac{1}{k!} D_p^{k+1} \bar{L}(p) \left(\sum_{j=1}^{\mathbf{N}} \frac{t^j}{j!} h_j \right)^{\otimes k},$$

and then, for every $m \in \{1, \dots, \mathbf{N} + 1\}$ and $\{h_i\}_{i=1}^{m-1} \subseteq \mathbb{R}^d$,

$$\bar{\mathbf{F}}_m(p, h_1, \dots, h_{m-1}) := (\partial_t^m \bar{\mathbf{G}})(0, p, h_1, \dots, h_{m-1}, 0, \dots, 0). \tag{1.16}$$

As above, we have that $\bar{\mathbf{F}}_1 \equiv 0$ by definition.

In the next theorem, we present a statement concerning the commutability of homogenization and higher-order linearizations. It generalizes [1, Theorem 1.1], which proved the result in the case $\mathbf{N} = 1$.

Theorem 1.2. (Homogenization of higher-order linearized equations).

Let $n \in \{0, \dots, \mathbf{N}\}$, $\delta \in (0, \frac{1}{2}]$, $\mathbf{M} \in [1, \infty)$, and $U_1, \dots, U_{n+1} \subseteq \mathbb{R}^d$ be a sequence of bounded Lipschitz domains satisfying

$$\bar{U}_{m+1} \subseteq U_m, \quad \forall m \in \{1, \dots, n\}. \tag{1.17}$$

There exist $\sigma(\text{data}) > 0$, $\alpha(\{U_m\}, \delta, \text{data}) > 0$, $C(\{U_m\}, \mathbf{M}, \delta, \text{data}) < \infty$, and a random variable \mathcal{X} satisfying

$$\mathcal{X} = \mathcal{O}_{\sigma}(C) \tag{1.18}$$

such that the following statement is valid. Let $\varepsilon \in (0, 1]$, $f \in W^{1,2+\delta}(U_0)$ be such that

$$\|\nabla f\|_{L^{2+\delta}(U_1)} \leq \mathbf{M},$$

and, for each $m \in \{1, \dots, n + 1\}$, fix $g_m \in W^{1,2+\delta}(U_m)$ and let $u^\varepsilon \in H^1(U_1)$ and the functions $w_1^\varepsilon \in H^1(U_1)$, $w_2^\varepsilon \in H^1(U_2), \dots, w_{n+1}^\varepsilon \in H^1(U_{n+1})$ satisfy, for every $m \in \{1, \dots, n + 1\}$, the Dirichlet problems

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u^\varepsilon, \frac{x}{\varepsilon})) = 0 & \text{in } U_1, \\ u^\varepsilon = f, & \text{on } \partial U_1, \\ -\nabla \cdot (D_p^2 L(\nabla u^\varepsilon, \frac{x}{\varepsilon}) \nabla w_m^\varepsilon) = \nabla \cdot (\mathbf{F}_m(\nabla u^\varepsilon, \nabla w_1^\varepsilon, \dots, \nabla w_{m-1}^\varepsilon, \frac{x}{\varepsilon})) & \text{in } U_m, \\ w_m^\varepsilon = g_m & \text{on } \partial U_m. \end{cases} \tag{1.19}$$

Finally, let $\bar{u} \in H^1(U_1)$ and, for every $m \in \{1, \dots, n + 1\}$, the function $\bar{w}_m \in H^1(U_m)$ satisfy the homogenized problems

$$\begin{cases} -\nabla \cdot D\bar{L}(\nabla\bar{u}) = 0 & \text{in } U_1, \\ \bar{u} = f & \text{on } \partial U_1, \\ -\nabla \cdot (D^2\bar{L}(\nabla\bar{u}) \nabla\bar{w}_m) = \nabla \cdot (\bar{\mathbf{F}}_m(\nabla\bar{u}, \nabla\bar{w}_1, \dots, \nabla\bar{w}_{m-1})) & \text{in } U_m, \\ \bar{w}_m = g_m & \text{on } \partial U_m. \end{cases} \tag{1.20}$$

Then, for every $m \in \{1, \dots, n + 1\}$, we have the estimate

$$\|\nabla w_m^\varepsilon - \nabla\bar{w}_m\|_{H^{-1}(U_m)} \leq \mathcal{X}\varepsilon^\alpha \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \tag{1.21}$$

Observe that, due to the assumed regularity of L in the spatial variable and the Schauder estimates, the vector fields $\mathbf{F}_m(\nabla u^\varepsilon, \nabla w_1^\varepsilon, \dots, \nabla w_{m-1}^\varepsilon, \frac{x}{\varepsilon})$ on the right side of the equations for w_m in (1.19) belong to $L^\infty(U_m)$. In particular, they belong to $L^2(U_m)$ and therefore the Dirichlet problems in (1.19) are well-posed in the sense that the solutions w_m belong to $H^1(U_m)$. Of course, this regularity given by the application of the Schauder estimate depends on ε (and indeed the constants blow up like a large power of ε^{-1}) and therefore this remark is not very useful as a quantitative statement. To prove the homogenization result for the m th linearized equation, we will need to possess much better bounds on these vector fields, which amounts to better regularity on the solutions $\nabla u, \nabla w_1, \dots, \nabla w_{m-1}$.

This is the reason we must prove Theorems 1.1 and 1.2 at the same time (in an induction on the order m of the linearized equation and of the regularity of $D^2\bar{L}$) as the following result on the large-scale regularity of solutions of the linearized equations and of the linearization errors. Its statement is a generalization of the large-scale $C^{0,1}$ estimates for linearized equation and *differences* of solutions proved in [1].

As mentioned above, throughout the paper we use the following notation for volume-normalized L^p norms: for each $p \in [1, \infty)$, $U \subseteq \mathbb{R}^d$ with $|U| < \infty$ and $f \in L^p(U)$,

$$\|f\|_{\underline{L}^p(U)} := \left(\int_U |f|^p \, dx \right)^{\frac{1}{p}} = |U|^{-\frac{1}{p}} \|f\|_{L^p(U)}.$$

Theorem 1.3. (Large-scale $C^{0,1}$ estimates for the linearized equations).

Let $n \in \{0, \dots, \mathbf{N}\}$, $q \in [2, \infty)$, and $\mathbf{M} \in [1, \infty)$. Then there exist $\sigma(q, \text{data}) > 0$, a constant $C(q, \mathbf{M}, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying $\mathcal{X} \leq \mathcal{O}_\sigma(C)$ such that the following statement is valid. For $R \in [2\mathcal{X}, \infty)$ and $u, v, w_1, \dots, w_{n+1} \in H^1(B_R)$ satisfying, for every $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} \|\nabla u\|_{\underline{L}^2(B_R)} \vee \|\nabla v\|_{\underline{L}^2(B_R)} \leq \mathbf{M} \\ -\nabla \cdot (D_p L(\nabla u, x)) = 0 \text{ and } -\nabla \cdot (D_p L(\nabla v, x)) = 0 \text{ in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x)) \text{ in } B_R, \end{cases}$$

and defining, for $m \in \{0, \dots, n\}$, the m th order linearization error $\xi_m \in H^1(B_R)$ by

$$\xi_m := v - u - \sum_{k=1}^m \frac{1}{k!} w_k,$$

then we have, for every $r \in [\mathcal{X}, \frac{1}{2}R]$ and $m \in \{0, \dots, n\}$, the estimates

$$\|\nabla w_{m+1}\|_{\underline{L}^q(B_r)} \leq C \sum_{i=1}^{m+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i}} \tag{1.22}$$

and

$$\begin{aligned} & \|\nabla \xi_m\|_{\underline{L}^q(B_r)} \\ & \leq C \sum_{i=0}^m \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i+1}} + C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i}}. \end{aligned} \tag{1.23}$$

The main interest in the above theorem is the case of the exponent $q = 2$. However, we must consider arbitrarily large exponents $q \in [2, \infty)$ in order for the induction argument to work. In particular, in order to show that Theorem 1.3 for some n implies Theorem 1.2 for $n + 1$, we need to consider potentially very large q (depending on n).

As mentioned above, Theorems 1.1, 1.2 and 1.3 are proved together in an induction argument. Each of the theorems has already been proved in the case $N = 0$ and $q = 2$ in our previous paper [1]. The integrability in Theorem 1.3 is upgraded to $q \in (2, \infty)$ in Propositions 4.7 and 5.1 below for w_1 and ξ_0 , respectively. These serve as the base case of the induction. The main induction step is comprised of the following three implications:

- *Regularity of \bar{L}* (Section 2). We show that if Theorems 1.1, 1.2 and 1.3 are valid for some $n \in \{0, \dots, N - 1\}$, then Theorems 1.1 is valid for $n + 1$. The argument essentially consists of differentiating the corrector equation for the n th linearized equation in the parameter p . However, the reader should not be misled into expecting a simple argument based on the implicit function theorem. Due to the lack of sufficient spatial integrability of the vector fields \mathbf{F}_m , it is necessary to use the large-scale regularity theory (that is, the assumed validity of Theorem 1.3 for n) to complete the argument.
- *Homogenization of higher-order linearized equations* (Section 3). We argue, for $n \in \{0, \dots, N - 1\}$, that if Theorem 1.1 is valid for $n + 1$ and Theorems 1.2 and 1.3 are valid for n , then Theorems 1.2 is valid for $n + 1$. The regularity of \bar{L} allows us to write down the homogenized equation, while the homogenization and regularity estimates for the previous linearized equations allow us to *localize* the heterogeneous equation; that is, approximate it with another equation which has a finite range of dependence and bounded coefficients. This allows us to apply homogenization estimates from [4].

- *Large-scale $C^{0,1}$ regularity of linearized solutions and linearization errors* (Sections 4 and 5). We argue, for $n \in \{0, \dots, N - 1\}$, that if Theorems 1.1 and 1.2 are valid for $n + 1$ and Theorem 1.3 is valid for n , then we may conclude that Theorem 1.3 is also valid for $n + 1$. Here we use the method introduced in [6] of applying a quantitative excess decay iteration, based on the “harmonic” approximation provided by the quantitative homogenization statement. This estimate controls the regularity of the w_m ’s on “large” scales (that is, larger than a multiple of the microscopic scale). To obtain L^q -type integrability for ∇w_m , it is also necessary to control the small scales, and for this we apply deterministic Calderón–Zygmund-type estimates (this is our reason for assuming L possesses some small-scale spatial regularity). The estimates for the linearization errors ξ_{m-1} are then obtained as a consequence by comparing them to w_m .

From a high-level point of view, the induction argument summarized above resembles the resolution of Hilbert’s 19th problem on the regularity of minimizers of integral functionals with uniformly convex and smooth integrands. The previous three theorems allow us to prove the next two results, which can be considered as resolutions of Hilbert’s 19th problem in the context of homogenization.

The first is the following result on precision of the higher-order linearization approximations which matches the one we have in the constant-coefficient case, as discussed near the end of Section 1.3:

Corollary 1.4. (Large-scale estimates of linearization errors). *Fix $n \in \{0, \dots, N\}$, $M \in [1, \infty)$ and let $U_0, U_1, \dots, U_n \subseteq \mathbb{R}^d$ be a sequence of bounded Lipschitz domains satisfying*

$$\overline{U_{m+1}} \subseteq U_m, \quad \forall m \in \{1, \dots, n - 1\}. \tag{1.24}$$

There exist constants $\sigma(\text{data}) \in (0, \frac{1}{2}]$, $C(\{U_m\}, M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying

$$\mathcal{X} = \mathcal{O}_\sigma(C)$$

such that the following statement is valid. Let $r \in [\mathcal{X}, \infty)$, $n \in \{1, \dots, N\}$ and $u, v \in H^1(rU_0)$ satisfy

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 & \text{and} & -\nabla \cdot (D_p L(\nabla v, x)) = 0 & \text{in } rU_0, \\ \|\nabla u\|_{\underline{L}^2(rU_0)} \vee \|\nabla v\|_{\underline{L}^2(rU_0)} \leq M, \end{cases}$$

and recursively define $w_m \in H^1(rU_m)$, for every $m \in \{1, \dots, n\}$, to be the solution of the Dirichlet problem

$$\begin{cases} -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x)) & \text{in } rU_m, \\ w_m = v - u - \sum_{k=1}^{m-1} \frac{1}{k!} w_k & \text{on } r\partial U_m. \end{cases} \tag{1.25}$$

Then, for every $m \in \{1, \dots, n\}$,

$$\left\| \nabla v - \nabla \left(u + \sum_{k=1}^m \frac{1}{k!} w_k \right) \right\|_{\underline{L}^2(rU_m)} \leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{m+1}. \tag{1.26}$$

Corollary 1.4 is an easy consequence of the previous theorems stated above. Its proof is presented in Section 6.

The analysis of the linearized equations presented in the theorems above allow us to develop a higher regularity theory for solutions of the nonlinear equation on large scales, in analogy to the role of the Schauder theory in the resolution of Hilbert’s 19th problem on the regularity of solutions of nonlinear equations with smooth (or constant) coefficients. This result generalizes the large-scale $C^{1,1}$ -type estimate proved in our previous paper [1] to higher-order regularity as well as the result in the linear case [4, Theorem 3.6].

Before giving the statement of this result, we introduce some additional notation and provide some motivational discussion. Given a domain $U \subseteq \mathbb{R}^d$, we define

$$\mathcal{L}(U) := \left\{ u \in H_{\text{loc}}^1(U) : -\nabla \cdot D_p L(\nabla u, x) = 0 \text{ in } U \right\}.$$

This is the set of solutions of the nonlinear equation in the domain U , which we note is a stochastic object. We next define \mathcal{L}_1 to be the set of global solutions of the nonlinear equation which exhibit at most linear growth at infinity:

$$\mathcal{L}_1 := \left\{ u \in \mathcal{L}(\mathbb{R}^d) : \limsup_{r \rightarrow \infty} r^{-1} \|u\|_{\underline{L}^2(B_r)} < \infty \right\}.$$

For each $p \in \mathbb{R}^d$, we denote the affine function ℓ_p by $\ell_p(x) := p \cdot x$. Observe that if the difference of two elements of \mathcal{L}_1 has strictly sublinear growth at infinity, it must be constant, by the $C^{0,1}$ -type estimate for differences (the estimate (1.23) with $m = 0$). Therefore the following theorem, which was proved in [1], gives a complete classification of \mathcal{L}_1 .

Theorem 1.5. (Large-scale $C^{1,1}$ -type estimate [1, Theorem 1.3]).

Fix $\sigma \in (0, d)$ and $\mathbf{M} \in [1, \infty)$. There exist $\delta(\sigma, d, \Lambda) \in (0, \frac{1}{2}]$, $C(\mathbf{M}, \sigma, \text{data}) < \infty$ and a random variable \mathcal{X}_σ which satisfies the estimate

$$\mathcal{X}_\sigma \leq \mathcal{O}_\sigma(C) \tag{1.27}$$

such that the following statements are valid:

- (i) For every $u \in \mathcal{L}_1$ satisfying $\limsup_{r \rightarrow \infty} \frac{1}{r} \|u - (u)_{B_r}\|_{\underline{L}^2(B_r)} \leq \mathbf{M}$, there exist an affine function ℓ such that, for every $R \geq \mathcal{X}_\sigma$,

$$\|u - \ell\|_{\underline{L}^2(B_R)} \leq CR^{1-\delta}.$$

- (ii) For every $p \in B_{\mathbf{M}}$, there exists $u \in \mathcal{L}_1$ satisfying, for every $R \geq \mathcal{X}_\sigma$,

$$\|u - \ell_p\|_{\underline{L}^2(B_R)} \leq CR^{1-\delta}.$$

(iii) For every $R \geq \mathcal{X}_s$ and $u \in \mathcal{L}(B_R)$ satisfying $\frac{1}{R} \|u - (u)_{B_R}\|_{\underline{L}^2(B_R)} \leq \mathbf{M}$, there exists $\phi \in \mathcal{L}_1$ such that, for every $r \in [\mathcal{X}_\sigma, R]$,

$$\|u - \phi\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R}\right)^2 \inf_{\psi \in \mathcal{L}_1} \|u - \psi\|_{\underline{L}^2(B_R)}. \tag{1.28}$$

Statements (i) and (ii) of the above theorem, which give the characterization of \mathcal{L}_1 , can be considered as a first-order Liouville-type theorem. In the case of deterministic, periodic coefficient fields, this result was proved by Moser and Struwe [23], who generalized the result of Avellaneda and Lin [8] in the linear case. Part (iii) of the theorem is a quantitative version of this Liouville-type result, which we call a “large-scale $C^{1,1}$ estimate” since it states that, on large scales, an arbitrary solution of the nonlinear equation can be approximated by an elements of \mathcal{L}_1 with the same precision as harmonic functions can be approximated by affine functions. It can be compared to similar statements in the linear case (see for instance [4, 17]).

In this paper we prove a higher-order version of Theorem 1.5. We will show that, just as a harmonic function can be approximated locally by harmonic polynomials, we can approximate an arbitrary element of $\mathcal{L}(B_R)$ by elements of a random set of functions which are the natural analogue of harmonic polynomials. In order to state this result, we must first define this space of functions.

Let us first discuss the constant-coefficient case. If \bar{L} is a smooth Lagrangian, we know from the resolution of Hilbert’s 19th problem that solutions of $-\nabla \cdot D_p \bar{L}(\nabla \bar{u}) = 0$ are smooth and thus may be approximated by a Taylor expansion at each point. One may then ask, can we characterize the possible Taylor polynomials? In Appendix E we provide such a characterization in terms of the linearized equations. The quadratic part is an $\bar{\mathbf{a}}_p := D^2 \bar{L}(p)$ -harmonic polynomial and the higher-order polynomials satisfy the equations the linearized equations, involving the $\bar{\mathbf{F}}_m$ ’s as right hand sides. More precisely, for each $p \in \mathbb{R}^d$ and $n \in \mathbb{N}$, we set

$$\begin{aligned} \bar{\mathbf{W}}_n^{p, \text{hom}} := & \left\{ (\bar{w}_1, \dots, \bar{w}_n) \in H_{\text{loc}}^1(\mathbb{R}^d; \mathbb{R}^n) : \text{for } m \in \{1, \dots, n\} \text{ we have} \right. \\ & \lim_{r \rightarrow 0} r^{-m} \|\bar{w}_m\|_{\underline{L}^2(B_r)} = 0, \quad \lim_{R \rightarrow \infty} R^{-1-m} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_R)} = 0, \\ & \left. -\nabla \cdot (\bar{\mathbf{a}}_p \nabla \bar{w}_m) = \nabla \cdot \bar{\mathbf{F}}_m(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1}) \right\}. \end{aligned}$$

It is not too hard to show that $\bar{\mathbf{W}}_n^{p, \text{hom}} \subseteq \mathcal{P}_2^{\text{hom}} \times \dots \times \mathcal{P}_{n+1}^{\text{hom}}$, where $\mathcal{P}_j^{\text{hom}}$ stands for homogeneous polynomials of degree j . Indeed, we see, by Liouville’s theorem, that \bar{w}_1 is an $\bar{\mathbf{a}}_p$ -harmonic polynomial of degree two. More importantly, according to Appendix E, we have that if \bar{u} solves $-\nabla \cdot D_p \bar{L}(\nabla \bar{u}) = 0$ in the neighborhood of origin, and we set $p = \nabla \bar{u}(0)$ and $\bar{w}_m(x) := \frac{1}{m+1} \nabla^{m+1} \bar{u}(0) x^{\otimes(m+1)}$, then

$$(\bar{w}_1, \dots, \bar{w}_n) \in \bar{\mathbf{W}}_n^{p, \text{hom}}.$$

In particular, \bar{w}_m is a sum of a special solution of

$$\nabla \cdot (\bar{\mathbf{a}}_p \nabla \bar{w}_m + \bar{\mathbf{F}}_m(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1})) = 0$$

in $\mathcal{P}_{m+1}^{\text{hom}}$ and an $\bar{\mathbf{a}}_p$ -harmonic polynomial in $\mathcal{P}_{m+1}^{\text{hom}}$. For our purposes it is convenient to relax the growth condition at the origin and define simply

$$\bar{\mathbf{W}}_n^p := \left\{ (\bar{w}_1, \dots, \bar{w}_n) \in \mathcal{P}_2 \times \dots \times \mathcal{P}_{n+1} : \text{for } m \in \{1, \dots, n\} \text{ we have} \right. \\ \left. \begin{aligned} \lim_{R \rightarrow \infty} R^{-1-m} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_R)} &= 0, \\ -\nabla \cdot (\bar{\mathbf{a}}_p \nabla \bar{w}_m) &= \nabla \cdot \bar{\mathbf{F}}_m(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1}) \end{aligned} \right\}.$$

With this definition, we lose the homogeneity of polynomials. Following the approach in [4, Chapter 3], it is natural to define heterogeneous versions of these spaces by

$$\mathbf{W}_n^p := \left\{ (w_1, \dots, w_n) \in H_{\text{loc}}^1(\mathbb{R}^d; \mathbb{R}^n) : \text{for } m \in \{1, \dots, n\} \text{ we have} \right. \\ \left. \begin{aligned} \limsup_{R \rightarrow \infty} R^{-1-m} \|\nabla w_m\|_{\underline{L}^2(B_R)} &= 0, \\ -\nabla \cdot (D_p^2 L(p + \nabla \phi_p, \cdot) \nabla w_m) &= \nabla \cdot \mathbf{F}_m(p + \nabla \phi_p, \nabla w_1, \dots, \nabla w_{m-1}, \cdot) \end{aligned} \right\},$$

that is, the tuplets of heterogeneous solutions with prescribed growth. Here $\ell_p + \phi_p$ is the unique element of \mathcal{L}_1 (up to additive constants) satisfying

$$\lim_{r \rightarrow \infty} \frac{1}{r} \|\phi_p\|_{\underline{L}^2(B_r)} = 0.$$

In other words, ϕ_p is the first-order corrector with slope p : see Lemma 2.10.

The next theorem gives a higher-order Liouville-type result which classifies the spaces \mathbf{W}_n^p and states that they may be used to approximate any solution of the nonlinear equation with the precision of a $C^{n,1}$ estimate.

Theorem 1.6. (Large-scale regularity). *Fix $n \in \{1, \dots, \mathbf{N}\}$ and $\mathbf{M} \in [1, \infty)$. There exist constants $\sigma(n, \mathbf{M}, \text{data})$, $\delta(n, \text{data}) \in (0, \frac{1}{2}]$ and a random variable \mathcal{X} satisfying the estimate*

$$\mathcal{X} \leq \mathcal{O}_\sigma(C(n, \mathbf{M}, d, \Lambda)) \tag{1.29}$$

such that the following statements hold:

- (i)_n *There exists a constant $C(n, \mathbf{M}, \text{data}) < \infty$ such that, for every $p \in B_{\mathbf{M}}$ and $(w_1, \dots, w_n) \in \mathbf{W}_n^p$, there exists $(\bar{w}_1, \dots, \bar{w}_n) \in \bar{\mathbf{W}}_n^p$ such that, for every $R \geq \mathcal{X}$ and $m \in \{1, \dots, n\}$,*

$$\|w_m - \bar{w}_m\|_{\underline{L}^2(B_R)} \leq C R^{1-\delta} \left(\frac{R}{\mathcal{X}}\right)^m \sum_{i=1}^m \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{m}{i}}. \tag{1.30}$$

- (ii)_n *For every $p \in B_{\mathbf{M}}$ and $(\bar{w}_1, \dots, \bar{w}_n) \in \bar{\mathbf{W}}_n^p$, there exists $(w_1, \dots, w_n) \in \mathbf{W}_n^p$ satisfying (1.30) for every $R \geq \mathcal{X}$ and $m \in \{1, \dots, n\}$.*

(iii)_n *There exists $C(n, \mathbf{M}, \text{data}) < \infty$ such that, for every $R \geq \mathcal{X}$ and $v \in \mathcal{L}(B_R)$ satisfying $\|\nabla v\|_{\underline{L}^2(B_R)} \leq \mathbf{M}$, there exist $p \in B_C$ and $(w_1, \dots, w_n) \in W_n^p$ such that, defining*

$$\xi_k(x) := v(x) - p \cdot x - \phi_p(x) - \sum_{i=1}^k \frac{w_i}{i!},$$

we have, for every $r \in [\mathcal{X}, \frac{1}{2}R]$ and $k \in \{0, 1, \dots, n\}$, the following estimates:

$$\sum_{i=0}^k \left(\|\nabla \xi_i\|_{\underline{L}^2(B_r)} \right)^{\frac{k+1}{i+1}} \leq C \left(\frac{r}{R} \right)^{k+1} \sum_{i=0}^k \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+1}{i+1}} \tag{1.31}$$

and

$$\|\nabla \xi_k\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R} \right)^{k+1} \frac{1}{R} \inf_{\phi \in \mathcal{L}_1} \|v - \phi\|_{\underline{L}^2(B_R)}. \tag{1.32}$$

The proof of Theorem 1.6 is given in Section 7.

We conclude this introduction with some comments about the possibility of generalizing our results in two different directions. First, it is not possible that our results can be extended from the scalar case to general nonlinear elliptic systems. Indeed, for even for *constant-coefficient*, analytic Lagrangians exhibiting uniform ellipticity and quadratic growth, there are explicit counterexamples (in dimensions larger than two) demonstrating that solutions may not even be Lipschitz: see for instance the dark side survey [22, Sections 3.3 & 3.4]. However, our arguments readily adapt in certain situations special cases (as might be expected), such as the two dimensional case or under an assumption that L has a particular form such as an Uhlenbeck-type structure (see [22, Section 4.7]). Roughly speaking, if one can identify conditions under which constant-coefficient elliptic systems have C^1 regularity, and these conditions are preserved by homogenization, then our techniques should be expected to apply.

Second, one might wonder whether it is possible to develop an analogous theory (in the scalar case) for functionals exhibiting non-quadratic growth, for example heterogeneous versions of the p -Laplacian for $p \in (1, \infty) \setminus \{2\}$, under assumptions such as [22, (2.20)]. This is indeed an interesting problem which seems to be wide open. Of course, qualitative homogenization is well-known in such a setting [11, 12]. The main roadblock at the present time to establishing a quantitative theory of homogenization, as well as a large-scale regularity theory, is a better understanding of the homogenized Lagrangian. In particular, we would need to show that ellipticity and growth assumptions such as [22, (2.20)] are preserved by homogenization, without which there is no hope for higher regularity. To our knowledge this is open even in the periodic setting.

Appendices C–E of this paper contain estimates for constant-coefficient equations which are essentially known but not to our knowledge written anywhere. We also collect some auxiliary estimates and computations in Appendices A and B.

2. Regularity Estimates for the Effective Lagrangian

In this section, we suppose that $n \in \{0, \dots, N - 1\}$ is such that

$$\text{the statements of Theorems 1.1, 1.2 and 1.3 are valid for } n. \tag{2.1}$$

The goal is to prove Theorem 1.1 for $n + 1$.

We proceed by constructing the linearized correctors $\psi_{p,h}^{(m)}$ up to $m = n + 2$ and relate the correctors of different orders to each other via differentiation in the parameter p . We show that these results allow us to improve the regularity of $D^2\bar{L}$ up to $C^{n+1,\beta}$ and obtain the statement of Theorem 1.1 for $n + 1$. In particular, this allows us to define the effective coefficient $\bar{\mathbf{F}}_{n+1}$. We also give formulas for the derivatives of \bar{L} and for $\bar{\mathbf{F}}_m$ in terms of the correctors, which allow us to relate them to each other and show that (1.16) holds.

2.1. The First-Order Correctors and Linearized Correctors

In this subsection we construct the linearized correctors up to order $n + 2$.

For each $p \in \mathbb{R}^d$, we define ϕ_p to be the *first-order corrector* of the nonlinear equation, that is, the unique solution of

$$\begin{cases} -\nabla \cdot (D_p L(p + \nabla \phi_p(x), x)) = 0 & \text{in } \mathbb{R}^d, \\ \nabla \phi_p \text{ is } \mathbb{Z}^d\text{-stationary, and } \mathbb{E} \left[\int_{\square_0} \nabla \phi_p(x) \, dx \right] = 0. \end{cases} \tag{2.2}$$

The existence and uniqueness (up to additive constants) of the first-order corrector ϕ_p is classical: it can be obtained from a variational argument (applied to an appropriate function space of stationary functions). Alternatively, it can be shown (following the proof given in [4, Section 3.4]) that the elements of \mathcal{L}_1 , which was characterized in Theorem 1.5 above (which was proved already in [1]), have stationary gradients.

We define the coefficient field $\mathbf{a}_p(x)$ to be the coefficients for the linearized equation around the solution $x \mapsto p \cdot x + \phi_p(x)$:

$$\mathbf{a}_p(x) := D_p^2 L(p + \nabla \phi_p(x), x).$$

Given $p, h \in \mathbb{R}^d$, we define the *first linearized corrector* $\psi_{p,h}^{(1)}$ to satisfy

$$\begin{cases} -\nabla \cdot (\mathbf{a}_p(x) (h + \nabla \psi_{p,h}^{(1)})) = 0 & \text{in } \mathbb{R}^d, \\ \nabla \psi_{p,h}^{(1)} \text{ is } \mathbb{Z}^d\text{-stationary, and } \mathbb{E} \left[\int_{\square_0} \nabla \psi_{p,h}^{(1)}(x) \, dx \right] = 0. \end{cases}$$

In other words, $\psi_{p,h}^{(1)}$ is the first-order corrector with slope h for the equation which is the linearization around the first-order corrector $x \mapsto p \cdot x + \phi_p(x)$.

For $m \in \{2, \dots, N + 2\}$, we define the m th linearized corrector to be the unique (modulo additive constants) random field $\psi_{p,h}^{(m)}$ satisfying

$$\begin{cases} -\nabla \cdot (\mathbf{a}_p(x) \nabla \psi_{p,h}^{(m)}) \\ = \nabla \cdot \mathbf{F}_m(p + \nabla \phi_p(x), h + \nabla \psi_{p,h}^{(1)}, \nabla \psi_{p,h}^{(2)}(x), \dots, \nabla \psi_{p,h}^{(m-1)}(x), x) & \text{in } \mathbb{R}^d, \\ \nabla \psi_{p,h}^{(m)} \text{ is } \mathbb{Z}^d\text{-stationary, and } \mathbb{E} \left[\int_{\square_0} \nabla \psi_{p,h}^{(m)}(x) \, dx \right] = 0. \end{cases} \tag{2.3}$$

In other words, $\psi_{p,h}^{(m)}$ is the corrector with slope zero for the m th linearized equation around $x \mapsto p \cdot x + \phi_p(x)$ and $x \mapsto h \cdot x + \psi_{p,h}^{(1)}(x)$, $x \mapsto \psi_{p,h}^{(2)}(x)$, \dots , $x \mapsto \psi_{p,h}^{(m-1)}(x)$. Notice that this gives us the complete collection of correctors for the latter equation, since by linearity we observe that $\psi_{p,h}^{(m)} + \psi_{p,h'}^{(1)}$ is the corrector with slope h' . Furthermore, by the linearity of the map $h \mapsto h + \nabla \psi_{p,h}^{(1)}(x)$, it is easy to see from the structure of the equations of $\psi_{p,h}^{(m)}$ that, for $p, h \in \mathbb{R}^d$, and $t \in \mathbb{R}$,

$$\nabla \psi_{p,th}^{(m)} = t^m \nabla \psi_{p,h}^{(m)}. \tag{2.4}$$

For $p, h \in \mathbb{R}^d$, we define

$$\mathbf{f}_{p,h}^{(k)} := \mathbf{F}_k(p + \nabla \phi_p, h + \nabla \psi_{p,h}^{(1)}, \nabla \psi_{p,h}^{(2)}, \dots, \nabla \psi_{p,h}^{(k-1)}, \cdot). \tag{2.5}$$

By (2.4), we have that

$$\mathbf{f}_{p,h}^{(k)} = |h|^k \mathbf{f}_{p,h/|h|}^{(k)}$$

We first show that the the problem (2.3) for the m th linearized corrector is well-posed for $m \in \{2, \dots, n + 1\}$. This is accomplished by checking, inductively, using our hypothesis (2.1) (and in particular the validity of Theorem 1.3 for $m \leq n$), that we have appropriate estimates on the vector fields $\mathbf{F}_m(\dots)$ on the right side. We have to make this argument at the same time that we obtain estimates on the smoothness of the correctors $\psi_{p,h}^{(m)}$ as functions of p . In fact, we prove that $\nabla \psi_{p,h}^{(m+1)} = h \cdot D_p \nabla \psi_{p,h}^{(m)}$, which expressed in coordinates is

$$\nabla \psi_{p,h}^{(m+1)} = \sum_{i=1}^d h_i \partial_{p_i} \nabla \psi_{p,h}^{(m)}.$$

We will also obtain $C^{0,1}$ -type bounds on the linearized correctors, which together with the previous display yields good quantitative control on the smoothness of the correctors in p .

Before the main statements, let us collect a few preliminary elementary results needed in the proofs. The following lemma is well-known and can be proved by the Lax–Milgram lemma (see for instance [21, Chapter 7]).

Lemma 2.1. *Let $\mathbf{a}(\cdot)$ be a \mathbb{Z}^d -stationary random field valued in the symmetric matrices satisfying the ellipticity bound $I_d \leq \mathbf{a} \leq \Lambda I_d$. Suppose that \mathbf{f} is a \mathbb{Z}^d -stationary, \mathbb{R}^d -valued random field satisfying*

$$\mathbb{E} \left[\|\mathbf{f}\|_{L^2(\square_0)}^2 \right] < \infty.$$

Then there exists a unique random potential field ∇z satisfying

$$\begin{cases} -\nabla \cdot \mathbf{a} \nabla z = \nabla \cdot \mathbf{f} & \text{in } \mathbb{R}^d \\ \nabla z \text{ is } \mathbb{Z}^d\text{-stationary, and } \mathbb{E} \left[\int_{\square_0} \nabla z(x) \, dx \right] = 0. \end{cases}$$

and, for a constant $C(d, \Lambda) < \infty$, the estimate

$$\mathbb{E} \left[\|\nabla z\|_{L^2(\square_0)}^2 \right] \leq C \mathbb{E} \left[\|\mathbf{f}\|_{L^2(\square_0)}^2 \right].$$

Next, a central object in our analysis is the quantity \mathbf{F}_m defined in (1.15). Fix $m \in \{2, \dots, \mathbf{N} + 1\}$ and $x, p, h_1, \dots, h_{m-1} \in \mathbb{R}^d$. One can easily read from the definition that we have

$$\mathbf{F}_m(p, h_1, \dots, h_{m-1}, x) = m! \sum_{2 \leq j \leq m} \frac{1}{j!} D_p^{j+1} L(p, x) \sum_{\substack{i_1 + \dots + i_j = m \\ i_1, \dots, i_j \geq 1}} \prod_{k=1}^j \frac{h_{i_k}^{\otimes 1}}{i_k!}. \tag{2.6}$$

It then follows, by Young’s inequality, that there exists $C(m, \text{data}) < \infty$ such that

$$|\mathbf{F}_m(p, h_1, \dots, h_{m-1}, x)| \leq C \sum_{i=1}^{m-1} |h_i|^{\frac{m}{i}}. \tag{2.7}$$

We have, similarly, that

$$[\mathbf{F}_m(\cdot, h_1, \dots, h_{m-1}, x)]_{C^{0,1}(\mathbb{R}^n)} \leq C \sum_{i=1}^{m-1} |h_i|^{\frac{m}{i}} \tag{2.8}$$

and, for $k \in \{1, \dots, m - 1\}$,

$$|D_{h_k} \mathbf{F}_m(p, h_1, \dots, h_{m-1}, x)| \leq C \sum_{i=1}^{m-k} |h_i|^{\frac{m-k}{i}}. \tag{2.9}$$

Using these we get, for $x, p, \tilde{p}, h_1, \tilde{h}_1, \dots, h_{m-1}, \tilde{h}_{m-1} \in \mathbb{R}^d$,

$$\begin{aligned} & \left| \mathbf{F}_m(p, h_1, \dots, h_{m-1}, x) - \mathbf{F}_m(\tilde{p}, \tilde{h}_1, \dots, \tilde{h}_{m-1}, x) \right| \\ & \leq C |p - \tilde{p}| \sum_{i=1}^{m-1} \left(|h_i| + |\tilde{h}_i| \right)^{\frac{m}{i}} + C \sum_{i=1}^{m-1} |h_i - \tilde{h}_i| \sum_{j=1}^{m-i} \left(|h_j| + |\tilde{h}_j| \right)^{\frac{m-j}{j}}. \end{aligned} \tag{2.10}$$

Therefore, by Young’s inequality, we get, for all $\delta > 0$,

$$\begin{aligned} & \left| \mathbf{F}_m(p, h_1, \dots, h_{m-1}, x) - \mathbf{F}_m(\tilde{p}, \tilde{h}_1, \dots, \tilde{h}_{m-1}, x) \right| \\ & \leq C (|p - \tilde{p}| + \delta) \sum_{i=1}^{m-1} \left(|h_i| + |\tilde{h}_i| \right)^{\frac{m}{i}} + C\delta \sum_{i=1}^{m-1} \left(\delta^{-1} |h_i - \tilde{h}_i| \right)^{\frac{m}{i}}. \end{aligned} \tag{2.11}$$

Furthermore, as we will be employing an induction argument in m , it is useful to notice that the leading term in \mathbf{F}_m has a simple form, and we have

$$\mathbf{F}_m(p, h_1, \dots, h_{m-1}, x) = mD_p^3 L(p, x) h_{m-1}^{\otimes 1} h_1^{\otimes 1} + \tilde{\mathbf{F}}_m(p, h_1, \dots, h_{m-2}, x), \tag{2.12}$$

with $\tilde{\mathbf{F}}_2 = 0$. Moreover, as is shown in Appendix B, if, for $p, h \in \mathbb{R}^d, t \mapsto \mathbf{g}(p+th)$ is m times differentiable at $t = 0$ and noticing that since $m \leq \mathbf{N} + 1, p \mapsto L(p, x)$ is in C^{m+2} , then

$$\begin{aligned} & \mathbf{F}_{m+1}(\mathbf{g}(p), D_p \mathbf{g}(p) h^{\otimes 1}, \dots, D_p^m \mathbf{g}(p) h^{\otimes m}, x) \\ & = D_p \left(\mathbf{F}_m(\mathbf{g}(p), D_p \mathbf{g}(p) h^{\otimes 1}, \dots, D_p^{m-1} \mathbf{g}(p) h^{\otimes (m-1)}, x) \right) \cdot h \\ & \quad + D_p \left(D_p^2 L(\mathbf{g}(p), x) \right) h^{\otimes 1} \left(D_p^m \mathbf{g}(p) h^{\otimes m} \right)^{\otimes 1}. \end{aligned} \tag{2.13}$$

Our first result in this section gives direct consequences of (2.1) for estimates on the first-order correctors and linearized correctors.

Theorem 2.2. (Quantitative estimates on linearized correctors). *Assume (2.1) is valid. Fix $\mathbf{M} \in [1, \infty)$. For every $m \in \{2, \dots, n + 1\}$ and $p, h \in \mathbb{R}^d$, there exists a function $\psi_{p,h}^{(m)}$ satisfying (2.3). Moreover, there exist constants $C(\mathbf{M}, \text{data}) < \infty$ and $\sigma(n, d, \Lambda) \in (0, \frac{1}{2}]$ and a random variable \mathcal{X} satisfying $\bar{\mathcal{X}} \leq \mathcal{O}_\sigma(C)$ such that the following statement is valid. For every $p \in B_{\mathbf{M}}, h \in \bar{B}_1, m \in \{1, \dots, n\}$, and $r \geq \mathcal{X}$,*

$$\left\| \nabla \phi_{p+h} - \left(\nabla \phi_p + h + \sum_{k=1}^m \frac{1}{k!} \nabla \psi_{p,h}^{(k)} \right) \right\|_{L^2(B_r)} \leq C|h|^{m+1} \tag{2.14}$$

and, for every $p \in B_{\mathbf{M}}, h \in \bar{B}_1, m \in \{1, \dots, n + 1\}$ and $r \geq \mathcal{X}$,

$$\left\| \nabla \psi_{p,h}^{(m)} \right\|_{L^2(B_r)} \leq C|h|^m. \tag{2.15}$$

Finally, for $q \in [2, \infty)$ and $m \in \{1, \dots, n + 1\}$, there exist constants $\delta(m, d, \Lambda) \in (0, \frac{1}{2}]$ and $C(q, m, \mathbf{M}, \text{data}) < \infty$ such that, for every $p \in B_{\mathbf{M}}$ and $h \in \bar{B}_1$,

$$\left\| \nabla \psi_{p,h}^{(m)} \right\|_{L^q(\square_0)} \leq \mathcal{O}_\delta(C|h|^m). \tag{2.16}$$

Proof. Set, for $m \in \{1, \dots, n + 1\}$,

$$\xi_0 := (p + h) \cdot x + \phi_{p+h}(x) - (p \cdot x + \phi_p(x)) \quad \text{and} \quad \xi_m := \xi_0 - \sum_{k=1}^m \frac{1}{k!} \psi_{p,h}^{(k)}.$$

We first collect two consequences of Theorem 1.3 assumed for n . Fix $q \in [2, \infty)$. Theorem 1.3 implies that there is a minimal scale \mathcal{X} such that (1.22) and (1.23) are valid with $q(n + 1)$ instead of q and for every $r \in [\mathcal{X}, \frac{1}{2}R]$. Hence, for every $r \in [\mathcal{X}, \frac{1}{2}R]$ and $m \in \{0, \dots, n\}$, we get the estimates

$$\left\| \nabla \psi_{p,h}^{(m+1)} \right\|_{\underline{L}^{q(n+1)}(B_r)} \leq C \sum_{i=1}^{m+1} \left(\left\| \nabla \psi_{p,h}^{(i)} \right\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i}} \tag{2.17}$$

and

$$\left\| \nabla \xi_m \right\|_{\underline{L}^{q(n+1)}(B_r)} \leq C \sum_{i=0}^{m-1} \left(\left\| \nabla \xi_i \right\|_{\underline{L}^2(B_R)} + \left\| \nabla \psi_{p,h}^{(i+1)} \right\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i+1}}. \tag{2.18}$$

On the other hand, using (2.7) we obtain

$$\left| \mathbf{f}_{p,h}^{(m)} \right| \leq C \sum_{i=1}^{m-1} \left| \nabla \psi_{p,h}^{(i)} \right|^{\frac{m}{i}}.$$

In particular, it follows by (2.17) that, for $m \in \{1, \dots, n + 2\}$, $R > 2\mathcal{X}$ and $r \in [\mathcal{X}, \frac{1}{2}R]$,

$$\left\| \mathbf{f}_{p,h}^{(m)} \right\|_{\underline{L}^q(B_r)} \leq C \sum_{i=1}^{m-1} \left\| \nabla \psi_{p,h}^{(i)} \right\|_{\underline{L}^2(B_R)}^{\frac{m}{i}}.$$

Since $\nabla \psi_{p,h}^{(i)}$ is \mathbb{Z}^d -stationary random field, we have by the ergodic theorem, after sending $R \rightarrow \infty$, that a.s.

$$\left\| \mathbf{f}_{p,h}^{(m)} \right\|_{\underline{L}^q(\mathcal{X}\square_0)} \leq C \sum_{i=1}^{m-1} \mathbb{E} \left[\left\| \nabla \psi_{p,h}^{(i)} \right\|_{L^2(\square_0)}^2 \right]^{\frac{m}{2i}}.$$

Furthermore, by Lemma 2.1 and the previous display we get, for $m \in \{1, \dots, n + 2\}$, that

$$\begin{aligned} \mathbb{E} \left[\left\| \nabla \psi_{p,h}^{(m)} \right\|_{L^2(\square_0)}^2 \right] &\leq C \mathbb{E} \left[\left\| \mathbf{f}_{p,h}^{(m)} \right\|_{L^2(\square_0)}^2 \right] \\ &\leq C \mathbb{E} \left[\mathcal{X}^d \left\| \mathbf{f}_{p,h}^{(m)} \right\|_{\underline{L}^2(\mathcal{X}_m \square_0)}^2 \right] \leq C \sum_{i=1}^{m-1} \mathbb{E} \left[\left\| \nabla \psi_{p,h}^{(i)} \right\|_{L^2(\square_0)}^2 \right]^{\frac{m}{i}}. \end{aligned}$$

Observe that the limiting behavior of ξ_0 and $\psi_{p,h}^{(1)}$ can be identified via their equations

$$-\nabla \cdot (\tilde{\mathbf{a}}_p (h + \nabla(\phi_{p+h} - \phi_p))) = 0 \quad \text{and} \quad -\nabla \cdot (\mathbf{a}_p (h + \nabla \psi_{p,h}^{(1)})) = 0$$

respectively, where

$$\tilde{\mathbf{a}}_p := \int_0^1 D_p^2 L(p + th + t\nabla\phi_{p+h} + (1-t)\nabla\phi_p, \cdot) dt.$$

By \mathbb{Z}^d -stationary of $\nabla\phi_p$ and $\nabla\phi_{p+h}$, implying \mathbb{Z}^d -stationarity of \mathbf{a}_p and $\tilde{\mathbf{a}}_p$, we may apply Lemma 2.1 to obtain

$$\mathbb{E} \left[\|\nabla\xi_0\|_{L^2(\square_0)}^2 + \|\nabla\psi_{p,h}^{(1)}\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C|h|.$$

It then follows inductively that, for $m \in \{1, \dots, n + 2\}$,

$$\mathbb{E} \left[\|\nabla\psi_{p,h}^{(m)}\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C|h|^m.$$

Using this together with the ergodic theorem and (2.17), (2.18), we obtain inductively, for $q \in [2, \infty)$, $m \in \{0, \dots, n + 1\}$ and $r \geq \mathcal{X}$,

$$\|\nabla\psi_{p,h}^{(m)}\|_{\underline{L}^q(B_r)} \leq C_q|h|^m \quad \text{and} \quad \|\nabla\xi_{m-1}\|_{\underline{L}^q(B_r)} \leq C|h|^m.$$

Now (2.16) follows by giving up a volume factor. The proof is complete. \square

We next show, again using (2.1), that the corrector $\psi_{p,h}^{(n+2)}$ satisfies an L^2 -type gradient estimate.

Lemma 2.3. *Assume (2.1) is valid. Let $\mathbf{M} \in [1, \infty)$. Suppose that $p \in B_{\mathbf{M}}$ and $h \in \overline{B}_1$. There exists $\psi_{p,h}^{(n+2)}$ satisfying (2.3) for $m = n + 2$ and a constant $C(n, \mathbf{M}, \text{data})$ such that*

$$\mathbb{E} \left[\|\nabla\psi_{p,h}^{(n+2)}\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C|h|^{n+2}.$$

Proof. The result follows directly by Lemma 2.1 and (2.16) using (2.7). \square

Lemma 2.4. *Assume (2.1) is valid. Suppose that $p \in B_{\mathbf{M}}$ and $h \in \overline{B}_1$. Then, for $m \in \{1, \dots, n + 1\}$, we have, a.s. and almost everywhere,*

$$\nabla\psi_{p,h}^{(m+1)} = \sum_{i=1}^d h_i \partial_{p_i} \nabla\psi_{p,h}^{(m)}. \tag{2.19}$$

Moreover, for $\beta \in (0, 1)$ and $m \in \{1, \dots, n + 1\}$, there exists $C(\beta, m, \mathbf{M}, \text{data}) < \infty$ such that, for $t \in (-1, 1)$,

$$\mathbb{E} \left[\|\nabla\psi_{p+th,h}^{(m)} - \nabla\psi_{p,h}^{(m)} - t\nabla\psi_{p,h}^{(m+1)}\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C|h|^{m+1}|t|^{1+\beta}. \tag{2.20}$$

For $m \in \{1, \dots, n\}$, we can take $\beta = 1$ in (2.20). Finally, we have, for $m \in \{1, \dots, n + 2\}$, that, a.s. and almost everywhere,

$$\mathbf{f}_{p,h}^{m+1} = (D_p \mathbf{a}_p \cdot h) \nabla\psi_{p,h}^{(m)} + D_p \mathbf{f}_{p,h}^m \cdot h. \tag{2.21}$$

Proof. Fix $p \in \mathbb{R}^d$ and, without loss of generality, $h \in \partial B_1$. By (2.14) we have that $p \mapsto p + \nabla \phi_p(x)$ is C^n and $D_p^j \nabla \phi_p(x) h^{\otimes j} = \psi_{p,h}^{(j)}$ for every $j \in \{2, \dots, n\}$ almost surely for almost every x . Thus (2.20) is valid with $\beta = 1$ for $m \in \{1, \dots, n-1\}$ and, by (2.13),

$$\mathbf{f}_{p,h}^{n+1} = \left(D_p \mathbf{f}_{p,h}^n + D_p \mathbf{a}_p \left(\nabla \psi_{p,h}^{(n)} \right)^{\otimes 1} \right) \cdot h. \quad (2.22)$$

We denote, in short, for $t \neq 0$,

$$\zeta_{p,h,t}^{(n)} := \frac{1}{t} \left(\psi_{p+th,h}^{(n)} - \psi_{p,h}^{(n)} - t \psi_{p,h}^{(n+1)} \right).$$

Observe that, by (2.16), $\mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] < \infty$ for $t \neq 0$.

Step 1. We prove that, for $t \in (-1, 1)$, $t \neq 0$,

$$\begin{aligned} \mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] &\leq C \mathbb{E} \left[\left\| D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p,h}^{(n)} - \nabla \psi_{p+th,h}^{(n)} \right) \right\|_{L^2(\square_0)}^2 \right] \\ &\quad + C \mathbb{E} \left[\left\| \frac{1}{t} \left(\mathbf{a}_{p+th} - \mathbf{a}_p - D_p \mathbf{a}_p \cdot h \right) \nabla \psi_{p+th,h}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \\ &\quad + C \mathbb{E} \left[\left\| \frac{1}{t} \left(\mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h \right) \right\|_{L^2(\square_0)}^2 \right]. \end{aligned} \quad (2.23)$$

To show (2.23), we first claim that the difference quotient solves the equation

$$\begin{aligned} \nabla \cdot \left(\mathbf{a}_p \nabla \zeta_{p,h,t}^{(n)} \right) &= t \nabla \cdot \left(D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p,h}^{(n)} - \nabla \psi_{p+th,h}^{(n)} \right) \right) \\ &\quad - \nabla \cdot \left(\left(\mathbf{a}_{p+th} - \mathbf{a}_p - D_p \mathbf{a}_p \cdot h \right) \nabla \psi_{p+th,h}^{(n)} \right) \\ &\quad - \left(\mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h \right). \end{aligned} \quad (2.24)$$

Indeed, then (2.23) follows by Lemma 2.1. Rewriting

$$\begin{aligned} \mathbf{a}_p \nabla \zeta_{p,h,t}^{(n)} &= \left(\mathbf{a}_{p+th} \nabla \psi_{p+th,h}^{(n)} + \mathbf{f}_{p+th,h}^n \right) \\ &\quad - \left(\mathbf{a}_p \nabla \psi_{p,h}^{(n)} + \mathbf{f}_{p,h}^n \right) - t \left(\mathbf{a}_p \nabla \psi_{p,h}^{(n+1)} + \mathbf{f}_{p,h}^{n+1} \right) \\ &\quad + t \left(\mathbf{f}_{p,h}^{n+1} - D_p \mathbf{a}_p \cdot h \psi_{p+th,h}^{(n)} - D_p \mathbf{f}_{p,h}^n \cdot h \right) \\ &\quad + t D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p,h}^{(n)} - \nabla \psi_{p+th,h}^{(n)} \right) \\ &\quad - \left(\mathbf{a}_{p+th} - \mathbf{a}_p - D_p \mathbf{a}_p \cdot h \right) \nabla \psi_{p+th,h}^{(n)} - \left(\mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h \right), \end{aligned}$$

we observe that the first three terms on the right are solenoidal by the equations of $\psi_{p+th,h}^{(n)}$, $\psi_{p,h}^{(n)}$ and $\psi_{p,h}^{(n+1)}$, respectively, and the fourth term on the right is zero by (2.22). We thus obtain (2.24).

Step 2. We will estimate the terms on the right in (2.23) separately, and in this step we first show that, for $t \in (-1, 1)$, $t \neq 0$,

$$\begin{aligned} & \mathbb{E} \left[\left\| D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p,h}^{(n)} - \nabla \psi_{p+th,h}^{(n)} \right) \right\|_{L^2(\square_0)}^2 \right] \\ & \leq C |t|^{2\beta} \left(1 + \mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \right)^\beta. \end{aligned} \tag{2.25}$$

By the triangle inequality we have

$$\begin{aligned} & \left| \nabla \psi_{p+th,h}^{(n)} - \nabla \psi_{p,h}^{(n+1)} \right| \\ & \leq |t|^\beta \left(\left| \nabla \psi_{p+th,h}^{(n)} \right| + \left| \nabla \psi_{p,h}^{(n)} \right| \right)^{1-\beta} \left(\left| \nabla \psi_{p,h}^{(n+1)} \right| + \left| \nabla \zeta_{p,h,t}^{(n)} \right| \right)^\beta. \end{aligned}$$

Therefore, by Hölder’s inequality and (2.16),

$$\begin{aligned} & \mathbb{E} \left[\left\| D_p \mathbf{a}_p h^{\otimes 1} \left(\nabla \psi_{p+th,h}^{(n)} - \nabla \psi_{p,h}^{(n)} \right) \right\|_{L^2(\square_0)}^2 \right] \\ & \leq C |t|^{2\beta} \mathbb{E} \left[\left\| D_p \mathbf{a}_p \right\|_{L^{\frac{4}{1-\beta}}(\square_0)}^2 \left(\left\| \nabla \psi_{p+th,h}^{(n)} \right\|_{L^4(\square_0)} + \left\| \nabla \psi_{p,h}^{(n)} \right\|_{L^4(\square_0)} \right)^2 \right]^{1-\beta} \\ & \quad \times \left(\mathbb{E} \left[\left\| \nabla \psi_{p,h}^{(n+1)} \right\|_{L^2(\square_0)}^2 \right] + \mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \right)^\beta \\ & \leq C |t|^{2\beta} \left(1 + \mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \right)^\beta, \end{aligned}$$

which is (2.25).

Step 3. We show that

$$\mathbb{E} \left[\left\| \left(\mathbf{a}_{p+th} - \mathbf{a}_p - D_p \mathbf{a}_p \cdot h \right) \nabla \psi_{p+th,h}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \leq C t^4. \tag{2.26}$$

We have that

$$\mathbf{a}_{p+th} - \mathbf{a}_p - D_p \mathbf{a}_p \cdot h = t^2 \int_0^1 s_1 \int_0^1 D_p^2 \mathbf{a}_{p+s_1 s_2 t h} h^{\otimes 2} ds_1 ds_2,$$

and since

$$D_p^2 \mathbf{a}_p \Big|_{p=z} h^{\otimes 2} = D_p^3 L(z + \phi_z, \cdot) \left(\nabla \psi_{z,h}^{(2)} \right)^{\otimes 1} + D_p^4 L(p + \phi_p, \cdot) \left(h + \nabla \psi_{p,h}^{(1)} \right)^{\otimes 2},$$

we obtain (2.26) by (2.16).

Step 4. We then prove that

$$\begin{aligned} & \mathbb{E} \left[\left\| \mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h \right\|_{L^2(\square_0)}^2 \right] \\ & \leq C |t|^{2(1+\beta)} \left(1 + \mathbb{E} \left[\left\| \nabla \zeta_{p,h,t}^{(n)} \right\|_{L^2(\square_0)}^2 \right] \right)^\beta. \end{aligned} \tag{2.27}$$

Using decomposition (2.12), we have that

$$\begin{aligned} D_p \mathbf{f}_{p,h}^n \cdot h &= D_p \left(n D_p \mathbf{a}_p \cdot h \nabla \psi_{p,h}^{(n-1)} + \tilde{\mathbf{f}}_{p,h}^n \right) \cdot h \\ &= n D_p \mathbf{a}_p \cdot h \nabla \psi_{p,h}^{(n)} + n D_p^2 \mathbf{a}_p h^{\otimes 2} \nabla \psi_{p,h}^{(n-1)} + D_p \tilde{\mathbf{f}}_{p,h}^n \\ &= n D_p \mathbf{a}_p \cdot h \nabla \psi_{p,h}^{(n)} + \mathbf{g}_{p,h}^{n-1}, \end{aligned}$$

where

$$\mathbf{g}_{p,h}^{n-1} := n D_p^2 \mathbf{a}_p h^{\otimes 2} \nabla \psi_{p,h}^{(n-1)} + D_p \tilde{\mathbf{f}}_{p,h}^n$$

is a function of $\nabla \phi_p, \nabla \psi_{p,h}^{(1)}, \dots, \nabla \psi_{p,h}^{(n-1)}$ for $n \geq 2$, and thus differentiable in p . In particular, by (2.16), for every $q \in [2, \infty)$ there is $\delta(q, n, \text{data}) > 0$ and $C(q, n, M, \text{data}) < \infty$ such that

$$\left\| \mathbf{g}_{p,h}^{n-1} \right\|_{L^q(\square_0)} + \left\| D_p \mathbf{g}_{p,h}^{n-1} \right\|_{L^q(\square_0)} \leq \mathcal{O}_\delta(C).$$

Using the above decomposition for $D_p \mathbf{f}_{p,h}^n \cdot h$, we compute

$$\begin{aligned} \mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h &= t \int_0^1 \left(D_p \mathbf{f}_{p+s_1th,h}^n - D_p \mathbf{f}_{p,h}^n \right) \cdot h \, ds_1 \\ &= nt \int_0^1 D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p+s_1th,h}^{(n)} - \nabla \psi_{p,h}^{(n)} \right) \, ds \\ &\quad + t^2 \int_0^1 s_1 \int_0^1 \left(n D_p^2 \mathbf{a}_{p+s_1s_2th} h^{\otimes 2} \nabla \psi_{p+s_1th,h}^{(n)} + \left(D_p \mathbf{g}_{p+s_1s_2th,h}^{n-1} \right) \cdot h \right) \, ds_1 \, ds_2. \end{aligned}$$

We have by (2.16) that

$$\left\| D_p^2 \mathbf{a}_{p+s_1s_2th} h^{\otimes 2} \nabla \psi_{p+s_1th,h}^{(n)} + \left(D_p \mathbf{g}_{p+s_1s_2th,h}^{n-1} \right) \cdot h \right\|_{L^q(\square_0)} \leq \mathcal{O}_\delta(C),$$

and therefore

$$\mathbf{f}_{p+th,h}^n - \mathbf{f}_{p,h}^n - D_p \mathbf{f}_{p,h}^n \cdot h = t \int_0^1 D_p \mathbf{a}_p \cdot h \left(\nabla \psi_{p+s_1th,h}^{(n)} - \nabla \psi_{p,h}^{(n)} \right) + \mathcal{O}_\delta(Ct^2).$$

The right hand side can be estimated with the aid of (2.25) to obtain (2.27).

Step 5. Conclusion. Combining (2.23) with (2.25), (2.26) and (2.27) yields (2.20) by Young’s inequality. Now, (2.20) implies (2.19) for $m = n$. Therefore, we may replace n by $n + 1$ in Steps 1-4 above, and conclude that (2.20) is valid for $m = n + 1$ as well, which then gives (2.19) for $m = n + 1$. Using obtained formula, it is straightforward to show that (2.20) is valid for $m = n$ with $\beta = 1$. Indeed, we notice that

$$\mathbb{E} \left[\left\| \nabla \psi_{p+th,h}^{(m)} - \nabla \psi_{p,h}^{(m)} - t \nabla \psi_{p,h}^{(m+1)} - \frac{t^2}{2} \nabla \psi_{p,h}^{(m+2)} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C |t|^{2+\beta},$$

from which we get (2.20) for $m = n$ with $\beta = 1$. Finally, (2.20) implies that $t \mapsto \nabla \phi_{p+th}$ is in $C^{n+2,\beta}$ close to $t = 0$, and thus we have that (2.21) is valid by (2.13). The proof is complete. \square

Lemma 2.5. *Assume (2.1) is valid. Let $M \in [1, \infty)$ and $p \in B_M$. Then $p \mapsto p + \nabla\phi_p$ is $(n + 2)$ times differentiable with respect to p and, for $q \in [2, \infty)$, there are constants $\delta(q, n, \text{data}) \in (0, \frac{1}{2}]$ and $C(q, n, M, \text{data}) < \infty$ such that, for $m \in \{1, \dots, n + 1\}$,*

$$\left\| D_p^m \phi_p \right\|_{L^q(\square_0)} \leq \mathcal{O}_\delta(C) \quad \text{and} \quad \mathbb{E} \left[\left\| D_p^{n+2} \phi_p \right\|_{L^2(\square_0)}^2 \right] \leq C.$$

The proof of Lemma 2.5 relies on a general principle in polarization based on multilinear analysis, and it is formalized in the following lemma:

Lemma 2.6. *Let V be a real, finite-dimensional vector space, and let $\Phi : V^n \rightarrow \mathbb{R}$ be a multilinear, symmetric form, that is, for all $v_1, \dots, v_n \in V$ and any permutation σ of $\{1, \dots, n\}$, we have*

$$\Phi(v_1, \dots, v_n) = \Phi(v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

For $v \in V$, define $\phi(v) := \Phi(v, v, \dots, v)$. Then, for $v_1, \dots, v_n \in V$, the polarization formula

$$\Phi(v_1, \dots, v_n) = \frac{1}{n!} \sum_{A \subseteq \{1, \dots, n\}} (-1)^{n-|A|} \phi \left(\sum_{j \in A} v_j \right)$$

holds, where the leftmost summation is over all non-empty subsets $A \subseteq \{1, 2, \dots, n\}$ and $|A|$ is the number of elements in A .

Proof. We show the equivalent statement that

$$\sum_{A \subseteq \{1, \dots, n\}} (-1)^{n-|A|} \phi \left(\sum_{j \in A} v_j \right) = n! \Phi(v_1, \dots, v_n),$$

where the sum on the left is over all non-empty subsets A of $\{1, 2, \dots, n\}$. For this, we begin by expanding each summand $\phi \left(\sum_{j \in A} v_j \right) = \Phi \left(\sum_{j \in A} v_j, \dots, \sum_{j \in A} v_j \right)$ fully, as a sum of terms of the form $\Phi(v_{j_1}, \dots, v_{j_n})$ with $j_1, \dots, j_n \in A$. Using the symmetry of Φ , each such term can be written as $\Phi(v_{f(1)}, \dots, v_{f(n)})$, with non-decreasing indices $f(1) \leq f(2) \leq \dots \leq f(n)$ in A . Denote

$$\mathcal{M} := \{f : \{1, \dots, n\} \rightarrow \{1, \dots, n\} : f(1) \leq f(2) \leq \dots \leq f(n)\}$$

and, for $f \in \mathcal{M}$,

$$\text{im } f := \bigcup_{j \in \{1, \dots, n\}} \{f(j)\}.$$

Letting $c_A(f)$ denote the number of ordered n -tuples (j_1, \dots, j_n) of elements of A which can be reordered to form $(f(1), \dots, f(n))$, it follows that the expression $\sum_{A \subseteq \{1, \dots, n\}} (-1)^{n-|A|} \phi \left(\sum_{j \in A} v_j \right)$ can be expanded to give

$$\sum_{A \subseteq \{1, \dots, n\}} (-1)^{n-|A|} \sum_{f \in \mathcal{M}, \text{ im } f \subseteq A} c_A(f) \Phi(v_{f(1)}, \dots, v_{f(n)}).$$

Changing the order of summation gives

$$\sum_{f \in \mathcal{M}} c_{\{1, \dots, n\}}(f) \Phi(v_{f(1)}, \dots, v_{f(n)}) \sum_{A \supseteq \text{im } f} (-1)^{n-|A|},$$

where the sum on the right is over all subsets A of $\{1, 2, \dots, n\}$ which contain $\text{im } f$. Each such subset A can be written as $\{f(1), \dots, f(n)\} \cup B$, for some set B , possibly empty, satisfying $B \subseteq \{1, 2, \dots, n\} \setminus \text{im } f$. Hence, as $|A| = |\text{im } f| + |B|$ for B defined in this way, we can write our expression as

$$\sum_{f \in \mathcal{M}} c_{\{1, \dots, n\}}(f) \Phi(v_{f(1)}, \dots, v_{f(n)}) s_f,$$

where

$$s_f = (-1)^{n-|\text{im } f|} \sum_{B \subseteq (\{1, \dots, n\} \setminus \text{im } f)} (-1)^{|B|}.$$

Finally, it is a well-known combinatorial fact that for every nonempty finite set S , we have

$$\sum_{B \subseteq S} (-1)^{|B|} = \sum_{j=0}^{|S|} (-1)^j \binom{|S|}{j} = 0, \tag{2.28}$$

where the sum on the left is over all subsets B of S . Thus, above, we have

$$s_f = 0$$

unless $f(1) < \dots < f(n)$. Therefore $f(1) = 1, f(2) = 2, \dots, f(n) = n$, and in this case $s_f = 1$ and $c_{\{1, \dots, n\}}(f) = n!$. It follows that

$$\sum_{A \subseteq \{1, \dots, n\}} (-1)^{n-|A|} \phi \left(\sum_{j \in A} v_j \right) = n! \Phi(v_1, \dots, v_n),$$

as was to be shown; this proves the polarization formula. \square

Proof of Lemma 2.5. The lemma follows from Lemmas 2.3, 2.4, 2.6, and (2.16). \square

We next prove Hölder continuity of $p \mapsto D_p^{n+2} \phi_p$ and $p \mapsto \mathbf{f}_{p,h}^{n+2}$.

Lemma 2.7. *Assume (2.1) is valid. Let $M \in [1, \infty)$ and $\beta \in (0, 1)$. Then there is a constant $C(\beta, n, M, \text{data}) < \infty$ such that, for all $p, p' \in B_M$ and $h \in \overline{B}_1 \setminus \{0\}$,*

$$\begin{aligned} & \mathbb{E} \left[\left\| \nabla D_p^{n+2} \phi_p - \nabla D_{p'}^{n+2} \phi_{p'} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \\ & + |h|^{-n-2} \mathbb{E} \left[\left\| \mathbf{f}_{p',h}^{n+2} - \mathbf{f}_{p,h}^{n+2} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C |p - p'|^\beta. \end{aligned} \tag{2.29}$$

Proof. Fix $M \in [1, \infty)$ and $\beta \in (0, 1)$, and fix $p, p' \in B_M$ and $h \in \overline{B}_1 \setminus \{0\}$. By Lemma 2.1 and equations of $\psi_{p,h}^{(n+2)}$ and $\psi_{p',h}^{(n+2)}$ we have that

$$\mathbb{E} \left[\left\| \nabla \psi_{p,h}^{(n+2)} - \nabla \psi_{p',h}^{(n+2)} \right\|_{L^2(\square_0)}^2 \right] \leq C \mathbb{E} \left[\left\| \mathbf{f}_{p',h}^{n+2} - \mathbf{f}_{p,h}^{n+2} \right\|_{L^2(\square_0)}^2 \right].$$

Therefore, in view of Lemma 2.6, it is enough to show that

$$\mathbb{E} \left[\left\| \mathbf{f}_{p',h}^{n+2} - \mathbf{f}_{p,h}^{n+2} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \leq C |h|^{n+2} |p - p'|^\beta. \quad (2.30)$$

By (2.12) we may decompose $\mathbf{f}_{p,h}^{n+2}$ as

$$\begin{aligned} \mathbf{f}_{p,h}^{n+2} &= (n+2) D_p^3 L(p + \nabla \phi_p, \cdot) \left(h + \nabla \left(D_p \phi_p h^{\otimes 1} \right) \right)^{\otimes 1} \left(\nabla \psi_{p,h}^{(n+1)} \right)^{\otimes 1} \\ &\quad + \tilde{\mathbf{F}}_{n+2} \left(p + \nabla \phi_p, h + \nabla \left(D_p \phi_p h^{\otimes 1} \right), \dots, \nabla \left(D_p^n \phi_p h^{\otimes n} \right), \cdot \right). \end{aligned} \quad (2.31)$$

Since $\tilde{\mathbf{F}}_{n+2}$ is $C^{0,1}$ in its first argument and polynomial in its last n arguments, we obtain by homogeneity, the chain rule and Lemma 2.5 that

$$\begin{aligned} &\mathbb{E} \left[\left\| \tilde{\mathbf{F}}_{n+2} \left(p + \nabla \phi_p, h + \nabla \left(D_p \phi_p h^{\otimes 1} \right), \dots, \nabla \left(D_p^n \phi_p h^{\otimes n} \right), \cdot \right) \right. \right. \\ &\quad \left. \left. - \tilde{\mathbf{F}}_{n+2} \left(p' + \nabla \phi_{p'}, h + \nabla \left(D_p \phi_{p'} h^{\otimes 1} \right), \dots, \nabla \left(D_p^n \phi_{p'} h^{\otimes n} \right), \cdot \right) \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \\ &\leq C |h|^{n+2} |p - p'|. \end{aligned}$$

Therefore, the leading term is the first one on the right in (2.31). Indeed, we observe by the above display that

$$\begin{aligned} &\mathbb{E} \left[\left\| \mathbf{f}_{p',h}^{n+2} - \mathbf{f}_{p,h}^{n+2} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \\ &\leq C \mathbb{E} \left[\left\| |h + \nabla(D_p \phi_p h^{\otimes 1})| \left| \nabla \psi_{p',h}^{(n+1)} - \nabla \psi_{p,h}^{(n+1)} \right| \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} + C |h|^{n+2} |p - p'| \end{aligned}$$

To conclude, we need to estimate the first term on the right. To this end, we use the triangle inequality to get, for any $\beta \in [0, 1]$,

$$\left| \nabla \psi_{p',h}^{(n+1)} - \nabla \psi_{p,h}^{(n+1)} \right| \leq \left(\left| \nabla \psi_{p',h}^{(n+1)} \right| + \left| \nabla \psi_{p,h}^{(n+1)} \right| \right)^{1-\beta} \left(\left| \nabla \psi_{p',h}^{(n+1)} - \nabla \psi_{p,h}^{(n+1)} \right| \right)^\beta.$$

It follows, by Hölder's inequality, that

$$\begin{aligned} &\mathbb{E} \left[\left\| \mathbf{f}_{p',h}^{n+2} - \mathbf{f}_{p,h}^{n+2} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} \\ &\leq C \mathbb{E} \left[\left\| |h + \nabla(D_p \phi_p h^{\otimes 1})| \right\|_{L^{\frac{4}{1-\beta}}(\square_0)}^{\frac{4}{1-\beta}} \right]^{\frac{1-\beta}{4}} \end{aligned}$$

$$\begin{aligned} & \mathbb{E} \left[\left\| \left| \nabla \psi_{p',h}^{(n+1)} \right| + \left| \nabla \psi_{p,h}^{(n+1)} \right| \right\|_{L^4(\square_0)}^4 \right]^{\frac{1-\beta}{4}} \\ & \times \mathbb{E} \left[\left\| \nabla \psi_{p',h}^{(n+1)} - \nabla \psi_{p,h}^{(n+1)} \right\|_{L^2(\square_0)}^2 \right]^{\frac{\beta}{2}}. \end{aligned}$$

By Lemma 2.5 and Hölder’s inequality, we get

$$\begin{aligned} & \mathbb{E} \left[\left\| \nabla \psi_{p',h}^{(n+1)} - \nabla \psi_{p,h}^{(n+1)} \right\|_{L^2(\square_0)}^2 \right]^{\frac{\beta}{2}} \\ & \leq C \mathbb{E} \left[\left\| \nabla D_p^{n+1} \phi_{p'} - \nabla D_p^{n+1} \phi_p \right\|_{L^2(\square_0)}^2 \right]^{\frac{\beta}{2}} |h|^{\beta(n+1)} \\ & \leq C \mathbb{E} \left[\left\| \nabla \int_0^1 D_p^{n+2} \phi_{tp'+(1-t)p} dt \right\|_{L^2(\square_0)}^2 \right]^{\frac{\beta}{2}} |h|^{\beta(n+1)} |p - p'|^\beta \\ & \leq C \left(\int_0^1 \mathbb{E} \left[\left\| \nabla D_p^{n+2} \phi_{tp'+(1-t)p} \right\|_{L^2(\square_0)}^2 \right]^{\frac{1}{2}} dt \right)^\beta |h|^{\beta(n+1)} |p - p'|^\beta \\ & \leq C |h|^{\beta(n+1)} |p - p'|^\beta, \end{aligned}$$

together with

$$\mathbb{E} \left[\left\| h + \nabla(D_p \phi_p h^{\otimes 1}) \right\|_{L^{\frac{4}{1-\beta}}(\square_0)}^{\frac{4}{1-\beta}} \right]^{\frac{1-\beta}{4}} \leq C |h|$$

and

$$\mathbb{E} \left[\left\| \left| \nabla \psi_{p',h}^{(n+1)} \right| + \left| \nabla \psi_{p,h}^{(n+1)} \right| \right\|_{L^2(\square_0)}^4 \right]^{\frac{1-\beta}{4}} \leq C |h|^{(1-\beta)(n+1)}.$$

Consequently, combining above displays yields (2.30), finishing the proof. \square

2.2. Smoothness of \bar{L}

It is a well-known and easy consequence of qualitative homogenization that $D\bar{L}$ is given by the formula

$$D\bar{L}(p) = \mathbb{E} \left[\int_{\square_0} D_p L(p + \nabla \phi_p(x), x) dx \right].$$

In [1], we proved that $\bar{L} \in C^2$ and $D^2\bar{L}$ is given by the formula

$$D^2\bar{L}(p)h = \mathbb{E} \left[\int_{\square_0} \mathbf{a}_p \left(h + \nabla \psi_{p,h}^{(1)} \right) dx \right]. \tag{2.32}$$

We next generalize this result to higher derivatives of \bar{L} by essentially differentiating the previous formula many times and applying the results of the previous subsection. Moreover, we validate the statement of Theorem 1.1 for $n + 1$.

Proposition 2.8. *Assume that (2.1) is valid. Then Theorem 1.1 is valid for $n + 1$. Moreover, for every $m \in \{1, \dots, n + 1\}$ and $h \in \overline{B}_1$, we have the formula*

$$D_p^{m+2}\overline{L}(p)h^{\otimes(m+1)} = \mathbb{E} \left[\int_{\square_0} \left(\mathbf{a}_p \nabla \psi_{p,h}^{(m+1)} + \mathbf{f}_{p,h}^{(m+1)} \right) \right]. \tag{2.33}$$

Proof. Fix $M \in [1, \infty)$ and $p \in B_M$. We begin by showing (2.33). Starting from (2.32), and observing that since

$$\mathbf{f}_{p,h}^{(2)} = D_p^3 L(p + \nabla \phi_p, \cdot) \left(h + \nabla \psi_{p,h}^{(1)} \right)^{\otimes 2},$$

we have that

$$D_p \left(\mathbf{a}_p \left(h + \nabla \psi_{p,h}^{(1)} \right) \right) \cdot h = \mathbf{a}_p \nabla \psi_{p,h}^{(2)} + \mathbf{f}_{p,h}^{(2)}.$$

This implies that

$$D_p^3 \overline{L}(p)h^{\otimes 2} = \mathbb{E} \left[\int_{\square_0} \left(\mathbf{a}_p \nabla \psi_{p,h}^{(2)} + \mathbf{f}_{p,h}^{(2)} \right) \right].$$

Assume then, inductively, that for some $m \in \{3, \dots, n + 1\}$ we have that, for all $k \in \{2, \dots, m\}$,

$$D_p^{k+1} \overline{L}(p)h^{\otimes k} = \mathbb{E} \left[\int_{\square_0} \left(\mathbf{a}_p \nabla \psi_{p,h}^{(k)} + \mathbf{f}_{p,h}^{(k)} \right) \right]. \tag{2.34}$$

We prove that (2.34) is valid for $k = m + 1$ as well. Differentiating with respect to p yields, using (2.19) and (2.21), that

$$\begin{aligned} D_p^{m+2} \overline{L}(p)h^{\otimes(m+1)} &= \mathbb{E} \left[\int_{\square_0} \left(D_p \left(\mathbf{a}_p \nabla \psi_{p,h}^{(m)} + \mathbf{f}_{p,h}^{(m)} \right) \cdot h \right) \right] \\ &= \mathbb{E} \left[\int_{\square_0} \left(\mathbf{a}_p \nabla \psi_{p,h}^{(m+1)} + \left(D_p \mathbf{a}_p \cdot h \nabla \psi_{p,h}^{(m)} + D_p \mathbf{f}_{p,h}^{(m)} \cdot h \right) \right) \right] \\ &= \mathbb{E} \left[\int_{\square_0} \left(\mathbf{a}_p \nabla \psi_{p,h}^{(m+1)} + \mathbf{f}_{p,h}^{(m+1)} \right) \right], \end{aligned}$$

proving the induction step. This validates (2.33).

To show the regularity of \overline{L} , we first observe that Theorem 2.2 and Lemma 2.3, together with (2.34) and Lemma 2.6, yield that

$$\max_{k \in \{2, \dots, n+3\}} \left\| D_p^k \overline{L}(p) \right\|_{L^\infty(B_M)} \leq C. \tag{2.35}$$

Fix then $p' \in B_M$. Since

$$\nabla \psi_{p,h}^{(n+2)} = \nabla \left(D_p^{n+2} \phi_p h^{\otimes(n+2)} \right),$$

decomposing

$$\begin{aligned} & \mathbf{a}_p \nabla \psi_{p,h}^{(m+1)} + \mathbf{f}_{p,h}^{(m+1)} - (\mathbf{a}_{p'} \nabla \psi_{p',h}^{(m+1)} + \mathbf{f}_{p',h}^{(m+1)}) \\ &= \mathbf{a}_p \nabla \left((D_p^{n+2} \phi_p - D_p^{n+2} \phi_{p'}) h^{\otimes(n+2)} \right) \\ & \quad + (\mathbf{a}_p - \mathbf{a}_{p'}) \nabla \left(D_p^{n+2} \phi_{p'} h^{\otimes(n+2)} \right) + \left(\mathbf{f}_{p,h}^{(m+1)} - \mathbf{f}_{p',h}^{(m+1)} \right), \end{aligned}$$

and noticing that

$$\begin{aligned} |\mathbf{a}_p(x) - \mathbf{a}_{p'}(x)| &\leq \left[D_p^2 L(\cdot, x) \right]_{C^{0,1}} (|p - p'| + |\nabla \phi_p(x) - \nabla \phi_{p'}(x)|) \\ &\leq C \int_0^1 (1 + |\nabla D_p \phi_{tp+(1-t)p'}(x)|) dt |p - p'|, \end{aligned}$$

we obtain by Lemmas 2.5 and 2.7, together with Hölder’s inequality, that

$$\left| D_p^{n+3} \bar{L}(p) h^{\otimes(n+3)} - D_p^{n+3} \bar{L}(p') h^{\otimes(n+3)} \right| \leq C |h|^{n+3} |p - p'|^\beta.$$

In view of Lemma 2.6 this yields

$$\left| D_p^{n+3} \bar{L}(p) - D_p^{n+3} \bar{L}(p') \right| \leq C |p - p'|^\beta,$$

proving that $[D_p^{n+3} \bar{L}]_{C^{0,\beta}(B_M)} \leq C$. Together with (2.35) we thus get

$$\left\| D_p^2 \bar{L} \right\|_{C^{n+1,\beta}(B_M)} \leq C,$$

which is the statement of Theorem 1.1 for $n + 1$. The proof is complete. \square

Remark 2.9. Since we now have that Theorem 1.1 is valid for $n + 1$, we also have that $p \mapsto \bar{F}_{n+2}(p, \cdot, \dots, \cdot)$ belongs to $C^{0,\beta}$ for all $\beta \in (0, 1)$. In particular, for $\beta \in (0, 1)$ and $M \in [1, \infty)$, there exists $C(n, \beta, M, \text{data}) < \infty$ such that, for every tuple $(h_1, \dots, h_{n+1}) \in \mathbb{R}^d \times \dots \times \mathbb{R}^d$, we have that

$$[F_{n+2}(\cdot, h_1, \dots, h_{n+1})]_{C^{0,\beta}(B_M)} \leq C \sum_{i=1}^{n+1} |h_i|^{\frac{n+2}{i}}.$$

2.3. Sublinearity of Correctors

By the ergodic theorem, we have that, for every $p, h \in \mathbb{R}^d$ and $m \in \{1, \dots, n + 1\}$, the correctors and linearized correctors are (qualitatively) sublinear at infinity:

$$\begin{cases} \limsup_{r \rightarrow \infty} \frac{1}{r} \left\| \phi_p - (\phi_p)_{B_r} \right\|_{\underline{L}^2(B_r)} = 0, \\ \limsup_{r \rightarrow \infty} \frac{1}{r} \left\| \psi_{p,h}^{(m)} - (\psi_{p,h}^{(m)})_{B_r} \right\|_{\underline{L}^2(B_r)} = 0, \end{cases} \quad \mathbb{P}\text{-a.s.}$$

The assumption (2.1) allows us to give a quantitative estimate of this sublinearity.

Lemma 2.10. (Sublinearity of correctors). *Assume (2.1) is valid. Let $M \in [1, \infty)$. There exist $\alpha(\text{data}), \delta(\text{data}) > 0, C(M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying $\mathcal{X} \leq \mathcal{O}_\delta(C)$ such that, for every $r \geq \mathcal{X}, p \in B_M, h \in \bar{B}_1$ and $m \in \{1, \dots, n + 1\}$,*

$$\left\| \phi_p - (\phi_p)_{B_r} \right\|_{\underline{L}^2(B_r)} + \left\| \psi_{p,h}^{(m)} - (\psi_{p,h}^{(m)})_{B_r} \right\|_{\underline{L}^2(B_r)} \leq Cr^{1-\alpha}. \tag{2.36}$$

Proof. Fix $p \in B_M$ and $h \in \bar{B}_1$. Clearly $x \mapsto p \cdot x + \phi_p(x)$ belongs to \mathcal{L}_1 , and thus the result follows from [1, Theorem 1.3] as in [4, Section 3.4]. Hence we are left to show that $\psi_{p,h}^{(m)}$ satisfies the estimate in the statement. For $m = 1$ the result follows by [1, Theorem 5.2] and [4, Section 3.4]. We thus proceed inductively. Assume that (2.36) is valid for $m \in \{1, \dots, k\}$ for some $k \in \{1, \dots, n\}$. We then show that it continues to hold for $m = k + 1$. Since (2.1) is valid, by taking α and \mathcal{X} as in Theorem 1.2, and setting $\mathcal{Y} := \mathcal{X}^{2/\alpha}$, we have that if $R := \varepsilon^{-1} \geq \mathcal{Y}$, then $\varepsilon^\alpha \mathcal{X} \leq R^{-\frac{\alpha}{2}}$. We relabel \mathcal{Y} to \mathcal{X} and $\frac{\alpha}{2}$ to α . We further take \mathcal{X} larger, if necessary, so that [1, Proposition 4.3] is at our disposal. Suppressing both p and h from the notation, we let ϕ_r and $\psi_r^{(m)}$, for $m \in \{1, \dots, n + 1\}$, solve

$$\begin{cases} -\nabla \cdot (D_p L(p + \nabla \phi_r, \cdot)) = 0 & \text{in } B_{2^{n+1}r}, \\ -\nabla \cdot (D_p^2 L(p + \nabla \phi) \nabla \psi_r^{(m)}) = \nabla \cdot \mathbf{f}^{(k+1)} & \text{in } B_{2^{n+1-m}r}, \\ -\nabla \cdot (D_p^2 L(p + \nabla \phi_r) \nabla \tilde{\psi}_r^{(m)}) = \nabla \cdot \mathbf{f}_r^{(k+1)} & \text{in } B_{2^{n+1-m}r}, \\ \phi_r = 0 & \text{on } \partial B_{2^{n+1}r}, \\ \psi_r^{(m)} = 0 & \text{on } \partial B_{2^{n+1-m}r}, \end{cases} \tag{2.37}$$

where

$$\begin{aligned} \mathbf{f}^{(m)} &:= \mathbf{F}_m \left(p + \nabla \phi, h + \nabla \psi^{(1)}, \nabla \psi^{(2)}, \dots, \nabla \psi^{(m-1)}, \cdot \right), \\ \mathbf{f}_r^{(m)} &:= \mathbf{F}_m \left(p + \nabla \phi, h + \nabla \psi_r^{(1)}, \nabla \psi_r^{(2)}, \dots, \nabla \psi_r^{(m-1)}, \cdot \right). \end{aligned}$$

Now, $\phi_r, \psi_r^{(1)}, \dots, \psi_r^{(m-1)}$ all homogenize to zero and we get, by Theorem 1.2, that, for $r \geq \mathcal{X}$,

$$\|\phi_r\|_{\underline{L}^2(B_r)} + \|\tilde{\psi}_r^{(m)}\|_{\underline{L}^2(B_r)} \leq Cr^{1-\alpha}.$$

This and the induction assumption, that is that (2.36) is valid for $m \in \{1, \dots, k\}$, together with Lemma 2.11 below, imply that

$$\frac{1}{r} \left\| \psi_r^{(k+1)} - \tilde{\psi}_r^{(k+1)} \right\|_{\underline{L}^2(B_r)} + \left\| \mathbf{f}^{(k+1)} - \mathbf{f}_r^{(k+1)} \right\|_{\underline{L}^2(B_r)} \leq Cr^{-\alpha}.$$

Combining the previous two displays yields

$$\|\phi_r\|_{\underline{L}^2(B_r)} + \|\psi_r^{(k+1)}\|_{\underline{L}^2(B_r)} \leq Cr^{1-\alpha}.$$

Now, since $\psi_{2r}^{(k+1)} - \psi_r^{(k+1)}$ is \mathbf{a}_p -harmonic in B_r , we have by the Lipschitz estimate [1, Proposition 4.3] that, for $r \geq \mathcal{X}$ and $t \in [\mathcal{X}, r]$,

$$\left\| \nabla \psi_{2r}^{(k+1)} - \nabla \psi_r^{(k+1)} \right\|_{\underline{L}^2(B_t)} \leq \frac{C}{r} \left\| \psi_{2r}^{(k+1)} - \psi_r^{(k+1)} \right\|_{\underline{L}^2(B_r)} \leq Cr^{-\alpha}.$$

Therefore, by compactness, there exists $\hat{\psi}^{(k+1)}$ such that, for $t \in [\mathcal{X}, r]$,

$$\frac{1}{r} \left\| \hat{\psi}^{(k+1)} \right\|_{\underline{L}^2(B_r)} + \left\| \nabla \psi_r^{(k+1)} - \nabla \hat{\psi}^{(k+1)} \right\|_{\underline{L}^2(B_t)} \leq Cr^{-\alpha}.$$

Proceeding now as in [4, Section 3.4] proves that $\nabla \hat{\psi}^{(k+1)}$ is \mathbb{Z}^d -stationary. Finally, by integration by parts we also obtain that $\mathbb{E} \left[\int_{\square_0} \nabla \hat{\psi}^{(k+1)} \right] = 0$ and, therefore, since $\hat{\psi}^{(k+1)}$ solves the same equation as $\psi^{(k+1)}$, by the uniqueness we have that $\hat{\psi}^{(k+1)} = \psi^{(k+1)}$ up to a constant. The proof is hence complete by the previous display. \square

Above we made use of a lemma, which roughly states that if two solutions of the systems of linearized equations are close in L^2 then their gradients are also close. This lemma will also be applied repeatedly in the following sections.

Lemma 2.11. *Suppose that (3.1) holds. Let $q \in [2, \infty)$ and $\mathbf{M} \in [1, \infty)$. There exist $\delta(q, \text{data}) > 0$, $C(q, \mathbf{M}, \text{data}) < \infty$ and a random variable $\mathcal{X} \leq \mathcal{O}_\delta(C)$ such that the following holds. Let $R \geq \mathcal{X}$, $N \leq n + 1$ and (u, w_1, \dots, w_N) , $(\tilde{u}, \tilde{w}_1, \dots, \tilde{w}_N) \in (H^1(B_R))^{N+1}$ each be a solution of the system of the linearized equations, that is, for every $m \in \{1, \dots, N\}$, we have*

$$\left\{ \begin{array}{l} \|\nabla u\|_{\underline{L}^2(B_R)} \vee \|\nabla v\|_{\underline{L}^2(B_R)} \leq \mathbf{M} \\ -\nabla \cdot (D_p L(\nabla u, x)) = 0 \text{ and } -\nabla \cdot (D_p L(\nabla v, x)) = 0 \text{ in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x)) \text{ in } B_R, \end{array} \right.$$

and the same holds with $(\tilde{u}, \tilde{w}_1, \dots, \tilde{w}_N)$ in place of (u, w_1, \dots, w_N) . Then

$$\|\nabla u - \nabla \tilde{u}\|_{\underline{L}^q(B_{R/2})} \leq \frac{C}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \tag{2.38}$$

and, denoting

$$h_i := \frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} + \frac{1}{R} \|\tilde{w}_i - (\tilde{w}_i)_{B_R}\|_{\underline{L}^2(B_R)},$$

we have, for every $m \in \{1, \dots, N\}$,

$$\begin{aligned} & \|\nabla w_m - \nabla \tilde{w}_m\|_{\underline{L}^2(B_{R/2})} \\ & \leq C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} + C \sum_{i=1}^m \frac{1}{R} \|w_i - \tilde{w}_i\|_{\underline{L}^2(B_R)} \sum_{j=1}^{m-i} h_j^{\frac{m-j}{j}}. \end{aligned} \tag{2.39}$$

and

$$\begin{aligned} & \|\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, \cdot) - \mathbf{F}_m(\nabla \tilde{u}, \nabla \tilde{w}_1, \dots, \nabla \tilde{w}_{m-1}, \cdot)\|_{\underline{L}^2(B_{R/2})} \\ & \leq C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} + C \sum_{i=1}^{m-1} \frac{1}{R} \|w_i - \tilde{w}_i\|_{\underline{L}^2(B_R)} \sum_{j=1}^{m-i} h_j^{\frac{m-i}{j}}. \end{aligned} \tag{2.40}$$

Proof. Let us define $\mathbf{a}(x) := D_p^2 L(\nabla u, x)$ and $\mathbf{f}_m(x) := \mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x)$ and analogously define $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{f}}_m$. We assume that $R \geq 2^{m+2} \mathcal{X}$, where \mathcal{X} is as in Theorem 1.3 for n , valid by the assumption of (3.1).

The estimate (2.38) is just the estimate for ξ_0 in Theorem 1.3. It also implies

$$\|\mathbf{a} - \tilde{\mathbf{a}}\|_{\underline{L}^q(B_{R/2})} \leq \frac{C}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)}. \tag{2.41}$$

By (2.10), we have that

$$\begin{aligned} \|\mathbf{f}_m - \tilde{\mathbf{f}}_m\| & \leq C |\nabla u - \nabla \tilde{u}| \sum_{i=1}^{m-1} (|\nabla w_i| \vee |\nabla \tilde{w}_i|)^{\frac{m}{i}} \\ & \quad + C \sum_{i=1}^{m-1} |\nabla w_i - \nabla \tilde{w}_i| \sum_{j=1}^{m-i} (|\nabla w_j| \vee |\nabla \tilde{w}_j|)^{\frac{m-i}{j}}. \end{aligned}$$

Using Hölder’s inequality and applying Theorem 1.3 for n , we obtain, for any $p \in [2, \infty)$ and $\delta > 0$,

$$\begin{aligned} & \|\mathbf{f}_m - \tilde{\mathbf{f}}_m\|_{L^p(B_{2^{-m}R})} \\ & \leq C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} + C \sum_{i=1}^{m-1} \|\nabla w_i - \nabla \tilde{w}_i\|_{\underline{L}^{p+\delta}(B_{2^{-m}R})} \sum_{j=1}^{m-i} h_j^{\frac{m-i}{j}}. \end{aligned} \tag{2.42}$$

We observe that $\zeta := w_m - \tilde{w}_m$ satisfies the equation

$$-\nabla \cdot \mathbf{a} \nabla \zeta = \nabla \cdot (\mathbf{f}_m - \tilde{\mathbf{f}}_m) + \nabla \cdot ((\mathbf{a} - \tilde{\mathbf{a}}) \nabla \tilde{w}_m) \quad \text{in } B_R. \tag{2.43}$$

By Meyer’s estimate, if $\delta > 0$ is small enough, then

$$\begin{aligned} \|\nabla w_m - \nabla \tilde{w}_m\|_{\underline{L}^{2+\delta}(B_{2^{-m-1}R})} & \leq C \|\mathbf{f}_m - \tilde{\mathbf{f}}_m\|_{\underline{L}^{2+\delta}(B_{2^{-m}R})} + \frac{C}{R} \|w_m - \tilde{w}_m\|_{\underline{L}^2(B_R)} \\ & \quad + \|\mathbf{a} - \tilde{\mathbf{a}}\|_{\underline{L}^5(B_{R/2})} \|\nabla \tilde{w}_m\|_{\underline{L}^5(B_{R/2})}. \end{aligned}$$

Combining these and using (2.41) and the validity of Theorem 1.3, we get

$$\begin{aligned} \|\nabla w_m - \nabla \tilde{w}_m\|_{\underline{L}^{2+\delta}(B_{2^{-m-1}R})} &\leq \frac{C}{R} \|w_m - \tilde{w}_m\|_{\underline{L}^2(B_R)} \\ &\quad + C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} \\ &\quad + C \sum_{i=1}^{m-1} \|\nabla w_i - \nabla \tilde{w}_i\|_{\underline{L}^{2+2\delta}(B_{2^{-m}R})} \sum_{j=1}^{m-i} h_j^{\frac{m-i}{j}}. \end{aligned}$$

Taking δ_0 sufficiently small and putting $\delta_m := 2^{-m} \delta_0$, we get by induction (using Young’s inequality and rearranging several sums) that, for every $m \in \{1, \dots, N\}$,

$$\begin{aligned} \|\nabla w_m - \nabla \tilde{w}_m\|_{\underline{L}^{2+\delta_m}(B_{2^{-m-1}R})} \\ \leq C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} + C \sum_{i=1}^m \frac{1}{R} \|w_i - \tilde{w}_i\|_{\underline{L}^2(B_R)} \sum_{j=1}^{m-i} h_j^{\frac{m-i}{j}}. \end{aligned}$$

Combining this with (2.42), we get

$$\begin{aligned} \|\mathbf{f}_m - \tilde{\mathbf{f}}_m\|_{\underline{L}^2(B_{2^{-m-1}R})} \\ \leq C \left(\frac{1}{R} \|u - \tilde{u}\|_{\underline{L}^2(B_R)} \right) \sum_{i=1}^{m-1} h_i^{\frac{m}{i}} + C \sum_{i=1}^{m-1} \frac{1}{R} \|w_i - \tilde{w}_i\|_{\underline{L}^2(B_R)} \sum_{j=1}^{m-i} h_j^{\frac{m-i}{j}}. \end{aligned}$$

These imply (2.39) and (2.40) after a covering argument. \square

3. Quantitative Homogenization of the Linearized Equations

In this section, we suppose that $n \in \{0, \dots, N - 1\}$ is such that

$$\begin{cases} \text{Theorem 1.1 is valid with } n + 1 \text{ in place of } N, \\ \text{Theorems 1.2 and 1.3 are valid for } n. \end{cases} \tag{3.1}$$

The goal is to prove that Theorem 1.2 is also valid for $n + 1$ in place of n . That is, we need to homogenize the $(n + 2)$ th linearized equation.

In order to prove homogenization for the $(n + 2)$ th linearized equation, we follow the procedure used in [1] for homogenizing the first linearized equation. We first show, using the induction hypothesis (3.1), that the coefficients $D_p^2 L(\nabla u^\varepsilon, \frac{x}{\varepsilon})$ and $\mathbf{F}_{n+1}(\frac{x}{\varepsilon}, \nabla u^\varepsilon, \nabla w_1^\varepsilon, \dots, \nabla w_n^\varepsilon)$ can be approximated by random fields which are *local* (they satisfy a finite range of dependence condition) and *locally stationary* (they are stationary up to dependence on a slowly varying macroscopic variable). This then allows us to apply known quantitative homogenization results for linear elliptic equations which can be found for instance in [4].

3.1. Stationary, Finite Range Coefficient Fields

We proceed in a similar fashion as in [1, Section 3.4] by introducing approximating equations which are stationary and localized.

We fix an integer $k \in \mathbb{N}$ which represents the scale on which we localize. Let $v_{p,z}^{(k)}$ denote the solution of the Dirichlet problem

$$\begin{cases} -\nabla \cdot \left(D_p L(\nabla v_{p,z}^{(k)}, x) \right) = 0 & \text{in } z + \square_{k+1}, \\ v_{p,z}^{(k)} = \ell_p & \text{on } \partial(z + \square_{k+1}), \end{cases}$$

where ℓ_p is the affine function $\ell_p(x) := p \cdot x$. We then define, for each $z \in 3^k \mathbb{Z}^d$, a coefficient field $\tilde{\mathbf{a}}_{p,z}^{(k)}$ in $z + \square_{k+1}$ by

$$\tilde{\mathbf{a}}_{p,z}^{(k)}(x) := D_p^2 L(\nabla v_{p,z}^{(k)}(x), x), \quad x \in z + \square_{k+1}$$

and then recursively define, for each $\Theta = (p, h_1, \dots, h_{n+1}) \in (\mathbb{R}^d)^{n+2}$, $m \in \{1, \dots, n+1\}$ and $z \in 3^k \mathbb{Z}^d$, the functions $w_{m,\Theta,z}^{(k)} \in H^1(z + (1 + 2^{-m})\square_k)$ to be the solutions of the sequence Dirichlet problems

$$\begin{cases} -\nabla \cdot \left(\tilde{\mathbf{a}}_{p,z}^{(k)} \nabla w_{m,\Theta,z}^{(k)} \right) = \nabla \cdot \tilde{\mathbf{F}}_{m,\Theta,z}^{(k)} & \text{in } z + (1 + 2^{-m})\square_k, \\ w_{m,\Theta,z}^{(k)} = \ell_{h_m} & \text{on } \partial(z + (1 + 2^{-m})\square_k), \end{cases}$$

where $\tilde{\mathbf{F}}_{m,\Theta,z}^{(k)} \in L^2(z + (1 + 2^{-(m-1)})\square_k)$ is defined for $m \in \{1, \dots, n+2\}$ by

$$\tilde{\mathbf{F}}_{m,\Theta,z}^{(k)}(x) := \mathbf{F}_m(\nabla v_{p,z}^{(k)}(x, z + \square_{k+1}, p), \nabla w_{1,\Theta,z}^{(k)}(x), \dots, \nabla w_{m-1,\Theta,z}^{(k)}(x), x). \tag{3.2}$$

Finally, we create $3^k \mathbb{Z}^d$ -stationary fields by gluing the above functions together: for each $x \in \mathbb{R}^d$, we define

$$\begin{cases} v_p^{(k)}(x) := v_{p,z}^{(k)}(x) \\ w_{m,\Theta}^{(k)}(x) := w_{m,\Theta,z}^{(k)}(x) \\ \mathbf{a}_p^{(k)}(x) := \tilde{\mathbf{a}}_{p,z}^{(k)}(x), \\ \mathbf{F}_{m,\Theta}^{(k)}(x) := \tilde{\mathbf{F}}_{m,\Theta,z}^{(k)}(x), \end{cases} \quad z \in 3^k \mathbb{Z}^d \text{ is such that } x \in z + \square_k.$$

Notice that $v_p^{(k)}$ and $w_{m,\Theta}^{(k)}$ might not be H^1 functions globally, but we can nevertheless define their gradients locally in $z + \square_k$. The $\mathbb{R}^{d(n+3)}$ -valued random field $(\nabla v_p^{(k)}, \nabla w_{1,\Theta}^{(k)}, \dots, \nabla w_{n+2,\Theta}^{(k)})$ is $3^k \mathbb{Z}^d$ -stationary and has a range of dependence of at most $3^k \sqrt{15 + d}$, by construction. The same is also true of the corresponding coefficient fields, since these are local functions of this random field:

$$\begin{cases} (\mathbf{a}_p^{(k)}, \mathbf{F}_{2,\Theta}^{(k)}, \dots, \mathbf{F}_{n+2,\Theta}^{(k)}) \text{ is } 3^k \mathbb{Z}^d\text{-stationary and} \\ \text{has a range of dependence of at most } 3^k \sqrt{15 + d}. \end{cases} \tag{3.3}$$

In the next subsection, we will apply some known quantitative homogenization estimates to the linear equation

$$-\nabla \cdot \left(\mathbf{a}_p^{(k)} \nabla v \right) = \nabla \cdot \mathbf{F}_{n+2, \Theta}^{(k)}. \tag{3.4}$$

In order to anchor these estimates, we require some deterministic bounds on the vector field $\mathbf{F}_{n+2, \Theta}^{(k)}$ and thus on the gradients of the functions $w_{m, \Theta}^{(k)}$ defined above. These bounds are exploding as a large power of 3^k , but we will eventually choose 3^k to be relatively small compared with the macroscopic scale (so these large powers of 3^k will be absorbed).

Lemma 3.1. *Fix $M \in [1, \infty)$. There exist exponents $\beta(\text{data}) > 0$, $Q(\text{data}) < \infty$ and a constant $C(M, \text{data}) < \infty$ such that, for every $k \in \mathbb{N}$, $m \in \{1, \dots, n + 2\}$ and $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p| \leq M$, we have*

$$\left\| \mathbf{F}_{m, \Theta}^{(k)} \right\|_{C^{0, \beta}(\square_k)} \leq C 3^{Qk} \sum_{j=1}^{m-1} |h_j|^{\frac{m}{j}}. \tag{3.5}$$

Proof. By [1, Proposition A.3], there exist $\beta(d, \Lambda) \in (0, 1)$ and $C(M, \text{data}) < \infty$ such that, for every $z \in 3^k \mathbb{Z}^d$ and $p \in B_M$,

$$\begin{aligned} \left[\nabla v_{p, z}^{(k)} \right]_{C^{0, \beta}(z+2\square_k)} &\leq C + C \sup_{x \in z+2\square_k} \left\| \nabla v_{p, z}^{(k)} \right\|_{L^2(B_1(x))} \\ &\leq C + C \left\| \nabla v_{p, z}^{(k)} \right\|_{L^2(z+\square_{k+1})} \\ &\leq C + C(1 + |p|) 3^{\frac{kd}{2}}. \end{aligned}$$

We deduce the existence of $C(M, \text{data}) < \infty$ such that

$$\left[\tilde{\mathbf{a}}_{p, z}^{(k)} \right]_{C^{0, \beta}(z+2\square_k)} \leq C 3^{\frac{kd}{2}}. \tag{3.6}$$

We will argue by induction in $m \in \{1, \dots, n + 2\}$ that there exist $Q(\text{data}) < \infty$ and $C(M, \text{data}) < \infty$ such that

$$\left\| \nabla w_{m, \Theta, z}^{(k)} \right\|_{C^{0, \beta}(z+(1+\frac{2}{3})2^{-m})\square_k)} \leq C 3^{Qk} \sum_{j=1}^m |h_j|^{\frac{m}{j}}. \tag{3.7}$$

For $m = 1$ the claim follows from Proposition A.2 and (3.6). Suppose now that there exists $M \in \{2, \dots, n + 2\}$ such that the bound (3.7) holds for each $m \in \{1, \dots, M - 1\}$. Then we obtain that, for some $Q(M, \text{data}) < \infty$,

$$\left[\tilde{\mathbf{F}}_{M, \Theta, z}^{(k)} \right]_{C^{0, \beta}(z+(1+2^{-M})\square_k)} \leq C 3^{Qk} \sum_{j=1}^M |h_j|^{\frac{M}{j}}$$

By the Caccioppoli inequality, we get

$$\left\| \nabla w_{M, \Theta, z}^{(k)} \right\|_{L^2(z+(1+\frac{3}{4})2^{-M})\square_k)} \leq C 3^{Qk} \sum_{j=1}^M |h_j|^{\frac{M}{j}}.$$

In view of (3.6) and the previous two displays, another application of Proposition A.2 yields, after enlarging Q and C , the bound (3.7) for $m = M$. This completes the induction argument and the proof of (3.7). The bound (3.5) immediately follows. \square

By the assumed validity of Theorem 1.3 for n , we also have that, for each $M, q \in [2, \infty)$ there exist $\delta(q, \text{data}) > 0$ and $C(M, q, \text{data}) < \infty$ and a random variable $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, for every $k \in \mathbb{N}$ with $3^k \geq \mathcal{X}$ and every $m \in \{1, \dots, n + 1\}$ and $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p| \leq M$, we have

$$\left\| \nabla w_{m, \Theta, 0}^{(k)} \right\|_{\underline{L}^q((1+2^{-m-1})\square_k)} \leq C \sum_{j=1}^{m-1} |h_j|^{\frac{m}{j}} \tag{3.8}$$

and hence, for such k and Θ and $m \in \{1, \dots, n + 2\}$,

$$\left\| \tilde{\mathbf{F}}_{m, \Theta, 0}^{(k)} \right\|_{\underline{L}^q((1+2^{-m-1})\square_k)} \leq C \sum_{j=1}^{m-1} |h_j|^{\frac{m}{j}}. \tag{3.9}$$

Observe that (3.5) and (3.9) together imply that, for δ and C as above and every $m \in \{1, \dots, n + 2\}$,

$$\left\| \tilde{\mathbf{F}}_{m, \Theta, 0}^{(k)} \right\|_{\underline{L}^q(\square_k)} \leq \mathcal{O}_\delta \left(C \sum_{j=1}^{m-1} |h_j|^{\frac{m}{j}} \right). \tag{3.10}$$

We next study the continuity of $\mathbf{a}_p^{(k)}$ and $\mathbf{F}_{m, \Theta}^{(k)}$ in the parameter Θ .

Lemma 3.2. (Continuity of $\mathbf{a}_p^{(k)}$ and $\mathbf{F}_{m, \Theta}^{(k)}$ in Θ). *Fix $q \in [2, \infty)$ and $M \in [1, \infty)$. There exist constants $\delta(q, \text{data}) > 0$ and $Q(q, \text{data}), C(q, M, \text{data}) < \infty$ and a random variable $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, for every $k \in \mathbb{N}$ with $3^k \geq \mathcal{X}$, $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ and $\Theta' = (p', h'_1, \dots, h'_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p|, |p'| \leq M$,*

$$\left\| \mathbf{a}_p^{(k)} - \mathbf{a}_{p'}^{(k)} \right\|_{\underline{L}^q(\square_k)} \leq C |p - p'| \tag{3.11}$$

and, for every $m \in \{1, \dots, n + 2\}$,

$$\begin{aligned} & \left\| \mathbf{F}_{m, \Theta}^{(k)} - \mathbf{F}_{m, \Theta'}^{(k)} \right\|_{\underline{L}^2(\square_k)} \\ & \leq C |p - p'| \sum_{i=1}^{m-1} (|h_i| \vee |h'_i|)^{\frac{m}{i}} + C \sum_{i=1}^{m-1} |h_i - h'_i| \sum_{j=1}^{m-i} \left(|h_j| \vee |h'_j| \right)^{\frac{m-j}{j}}. \end{aligned} \tag{3.12}$$

Proof. We take \mathcal{X} to be larger than the random variables in the statements of Lemma 2.11 and Theorems 1.2 and 1.3 for n . The bound (3.11) is then an immediate consequence of (2.38) and the obvious fact that

$$\left\| \nabla v_{p, 0}^{(k)} - \nabla v_{p', 0}^{(k)} \right\|_{\underline{L}^2(z+\square_{k+1})} \leq C |p - p'|.$$

We then use the equation for the difference $w_{m,\Theta}^{(k)} - w_{m,\Theta'}^{(k)}$ [see (2.43)] and then apply the result of Lemma 2.11 to obtain, for every $m \in \{1, \dots, n + 1\}$,

$$\begin{aligned} & \left\| \nabla w_{m,\Theta,0}^{(k)} - \nabla w_{m,\Theta',0}^{(k)} \right\|_{\underline{L}^2((1+2^{-m})\square_k)} \\ & \leq C |h_m - h'_m| + C \left\| \tilde{\mathbf{F}}_{m,\Theta,0}^{(k)} - \tilde{\mathbf{F}}_{m,\Theta',0}^{(k)} \right\|_{\underline{L}^2((1+2^{-m})\square_k)} + C |p - p'| \sum_{i=1}^m |h_i|^{\frac{m}{i}} \\ & \leq C |h_m - h'_m| + C |p - p'| \sum_{i=1}^{m-1} (|h_i| \vee |h'_i|)^{\frac{m}{i}} \\ & \quad + C \sum_{i=1}^{m-1} \frac{1}{R} \left\| w_{i,\Theta,0}^{(k)} - w_{i,\Theta',0}^{(k)} \right\|_{\underline{L}^2((1+2^{-i})\square_k)} \sum_{j=1}^{m-i} (|h_i| \vee |h'_i|)^{\frac{m-i}{j}}. \end{aligned}$$

By induction we now obtain, for every $m \in \{1, \dots, n + 1\}$,

$$\begin{aligned} & \left\| \nabla w_{m,\Theta,0}^{(k)} - \nabla w_{m,\Theta',0}^{(k)} \right\|_{\underline{L}^2((1+2^{-m})\square_k)} \\ & \leq C |p - p'| \sum_{i=1}^{m-1} (|h_i| \vee |h'_i|)^{\frac{m}{i}} + C \sum_{i=1}^{m-1} |h_i - h'_i| \sum_{j=1}^{m-i} (|h_j| \vee |h'_j|)^{\frac{m-i}{j}}. \end{aligned}$$

This implies (3.12). \square

By combining (3.5) and (3.12) and using interpolation, we obtain the existence of $\alpha(\text{data}) > 0$, $Q(\text{data}) < \infty$ and $C(\text{data}) > 0$ such that, with \mathcal{X} as in the statement of Lemma 3.2, then for every $k \in \mathbb{N}$ with $3^k \geq \mathcal{X}$, $m \in \{1, \dots, n + 2\}$, and $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ and $\Theta' = (p', h'_1, \dots, h'_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p|, |p'| \leq M$, we have

$$\begin{aligned} & \left\| \mathbf{F}_{m,\Theta}^{(k)} - \mathbf{F}_{m,\Theta'}^{(k)} \right\|_{L^\infty(\square_k)} \\ & \leq C 3^{Qk} |p - p'|^\alpha \sum_{i=1}^{m-1} (|h_i| \vee |h'_i|)^{\frac{m}{i}} + C 3^{Qk} \sum_{i=1}^{m-1} |h_i - h'_i|^\alpha \sum_{j=1}^{m-i} (|h_j| \vee |h'_j|)^{\frac{m-i}{j}}. \end{aligned} \tag{3.13}$$

Likewise, we can use (3.6) and (3.11) to obtain

$$\left\| \mathbf{a}_p^{(k)} - \mathbf{a}_{p'}^{(k)} \right\|_{\underline{L}^q(\square_k)} \leq C 3^{Qk} |p - p'|^\alpha. \tag{3.14}$$

This variation of Lemma 3.2 will be needed below.

3.2. Setup of the Proof of Theorem 1.2

We are now ready to begin the proof of the implication

$$(3.1) \implies \text{the statement of Theorem 1.2 with } n + 1 \text{ in place of } n.$$

We fix parameters $M \in [1, \infty)$, $\delta > 0$ and $\varepsilon \in (0, 1)$, a sequence of Lipschitz domains $U_1, U_2, \dots, U_{n+2} \subseteq \square_0$ satisfying

$$\bar{U}_{m+1} \subseteq U_m, \quad \forall m \in \{1, \dots, n + 1\}. \tag{3.15}$$

a function $f \in W^{1,2+\delta}(U_1)$ satisfying

$$\|\nabla f\|_{L^{2+\delta}(U_1)} \leq M, \tag{3.16}$$

and a sequence of boundary conditions $g_1 \in W^{1,2+\delta}(U_1), \dots, g_{n+2} \in W^{1,2+\delta}(U_{n+2})$. We let $u^\varepsilon \in f + H_0^1(U_1)$ and $w_1^\varepsilon \in g_1 + H_0^1(U_1), \dots, w_{n+2}^\varepsilon \in g_{n+2} + H_0^1(U_{n+2})$ as well as $\bar{u} \in f + H_0^1(U_1)$ and $\bar{w}_1 \in g_1 + H_0^1(U_1), \dots, \bar{w}_{n+2} \in g_{n+2} + H_0^1(U_{n+2})$ be as in the statement of Theorem 1.2 for $n + 1$ in place of n .

We denote

$$\begin{cases} \mathbf{a}^\varepsilon(x) := D_p^2 L(\nabla u^\varepsilon(x), \frac{x}{\varepsilon}), \\ \bar{\mathbf{a}}(x) := D_p^2 \bar{L}(\nabla \bar{u}(x)). \end{cases} \tag{3.17}$$

and

$$\begin{cases} \mathbf{F}_m^\varepsilon(x) := \mathbf{F}_m(\nabla u^\varepsilon(x), \nabla w_1^\varepsilon(x), \dots, \nabla w_{m-1}^\varepsilon(x), \frac{x}{\varepsilon}), \\ \bar{\mathbf{F}}_m(x) := \bar{\mathbf{F}}_m(\nabla \bar{u}(x), \nabla \bar{w}_1(x), \dots, \nabla \bar{w}_{m-1}(x)). \end{cases} \tag{3.18}$$

We also choose $K \in \mathbb{N}$ to be the unique positive integer satisfying $3^{-K-1} < \varepsilon \leq 3^{-K}$. We write

$$\bar{\Theta}(x) := (\nabla \bar{u}(x), \nabla \bar{w}_1(x), \dots, \nabla \bar{w}_{n+1}(x)). \tag{3.19}$$

By the assumption (3.1), we only need to homogenize the linearized equation for $m = n + 2$. As we have already proved a homogenization result in [1, Theorem 1.1] for the linearized equation with zero right-hand side, it suffices to prove (1.21) for $m = n + 2$ under the assumption that the boundary condition vanishes:

$$g_{n+2} = 0 \quad \text{in } U_{n+2}. \tag{3.20}$$

We can write the equations for w_{n+2}^ε and \bar{w}_{n+2} respectively as

$$\begin{cases} -\nabla \cdot (\mathbf{a}^\varepsilon \nabla w_{n+2}^\varepsilon) = \nabla \cdot \mathbf{F}_{n+2}^\varepsilon & \text{in } U_{n+2}, \\ w_{n+2}^\varepsilon = 0, & \text{on } \partial U_{n+2}, \end{cases} \tag{3.21}$$

and

$$\begin{cases} -\nabla \cdot (\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}^\varepsilon) = \nabla \cdot \bar{\mathbf{F}}_{n+2} & \text{in } U_{n+2}, \\ \bar{w}_{n+2}^\varepsilon = 0, & \text{on } \partial U_{n+2}. \end{cases} \tag{3.22}$$

Our goal is to prove the following estimate: there exists a constant $C(\mathbf{M}, \text{data}) < \infty$, exponents $\sigma(\text{data})$ and $\alpha(\text{data}) > 0$ and random variable \mathcal{X} satisfying

$$\mathcal{X} = \mathcal{O}_\sigma(C), \tag{3.23}$$

as in the statement of the theorem:

$$\|\nabla w_{n+2}^\varepsilon - \nabla \bar{w}_{n+2}\|_{H^{-1}(U_m)} \leq \mathcal{X} \varepsilon^\alpha \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \tag{3.24}$$

To prove (3.24), we first compare the solution w_{n+2} to the solution $\tilde{w}_{n+2}^\varepsilon$ of a second heterogeneous problem, namely

$$\begin{cases} -\nabla \cdot (\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{w}_{n+2}^\varepsilon) = \nabla \cdot \tilde{\mathbf{F}}_{n+2}^\varepsilon & \text{in } U_{n+2}, \\ \tilde{w}_{n+2}^\varepsilon = 0, & \text{on } \partial U_{n+2}, \end{cases} \tag{3.25}$$

where the coefficient fields $\tilde{\mathbf{a}}^\varepsilon$ and $\tilde{\mathbf{F}}_{n+2}^\varepsilon$ are defined in terms of the localized, stationary approximating coefficients (introduced above in Section 3.1) by

$$\begin{cases} \tilde{\mathbf{a}}^\varepsilon(x) := \mathbf{a}_{\nabla \bar{u}(x)}^{(k)}\left(\frac{x}{\varepsilon}\right), \\ \tilde{\mathbf{F}}_{n+2}^\varepsilon := \mathbf{F}_{n+2, \bar{\Theta}(x)}^{(k)}\left(\frac{x}{\varepsilon}\right). \end{cases} \tag{3.26}$$

The parameter $k \in \mathbb{N}$ will be chosen below in such a way that $1 \ll 3^k \ll \varepsilon^{-1}$. We also need to declare a second mesoscopic scale by taking $l \in \mathbb{N}$ such that $1 \ll 3^k \ll 3^l \ll \varepsilon^{-1}$ and, for every $m \in \{1, \dots, n+1\}$,

$$U_{m+1} + \varepsilon \square_l \subseteq U_m. \tag{3.27}$$

Like k , the parameter l will be declared later in this section. For convenience we will also take a slightly smaller domain U_{n+3} than U_{n+2} , which also depends on ε and l and is defined by

$$U_{n+3} := \{x \in U_{n+2} : x + \varepsilon \square_l \subseteq U_{n+2}\}. \tag{3.28}$$

Thus we have (3.27) for every $m \in \{1, \dots, n+2\}$. See Section 3.4 below for more on the choices of the parameters k and l .

The estimate (3.24) follows from the following two estimates, which are proved separately below (here \mathcal{X} denotes a random variable as in the statement of the theorem):

$$\|\nabla w_{n+2}^\varepsilon - \nabla \tilde{w}_{n+2}^\varepsilon\|_{L^2(U_m)} \leq \mathcal{X} \varepsilon^\alpha \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}, \tag{3.29}$$

and

$$\|\nabla \tilde{w}_{n+2}^\varepsilon - \nabla \bar{w}_{n+2}\|_{H^{-1}(U_m)} \leq \mathcal{X} \varepsilon^\alpha \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \tag{3.30}$$

3.3. Estimates on Size of the w_m^ε and \bar{w}_m

To prepare for the proofs of (3.29) and (3.30), we present some preliminary bounds on the size of the w_m^ε 's. The first is a set of deterministic bounds representing the “worst-case scenario” in which nothing has homogenized up to the current scale.

Lemma 3.3. *There exist exponents $\beta(d, \Lambda) \in (0, 1)$ and $Q(\text{data}) < \infty$ and a constant $C(\{U_m\}, M, K_0, \text{data}) < \infty$ such that, for every $m \in \{1, \dots, n + 2\}$,*

$$\|\nabla w_m^\varepsilon\|_{C^{0,\beta}(U_{m+1})} \leq C\varepsilon^{-Q} \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \tag{3.31}$$

Proof. By [1, Proposition A.3], there exist $\beta(d, \Lambda) \in (0, 1)$, $Q(\text{data}) < \infty$ and $C(\{U_m\}, M, K_0, \text{data}) < \infty$ such that

$$[\nabla u^\varepsilon]_{C^{0,\beta}(U_1)} \leq C\varepsilon^{-Q}$$

and hence

$$[\mathbf{a}^\varepsilon]_{C^{0,\beta}(U_1)} \leq C\varepsilon^{-Q}. \tag{3.32}$$

We will argue by induction in $m \in \{1, \dots, n + 2\}$ that there exists $Q(m, \text{data}) < \infty$ and $C(m, \{U_k\}, M, K_0, \text{data}) < \infty$ such that

$$\|\nabla w_m^\varepsilon\|_{C^{0,\beta}(U_{m+1})} \leq C\varepsilon^{-Q} \sum_{j=1}^m |h_j|^{\frac{m}{j}}. \tag{3.33}$$

For $m = 1$ the claim follows from Proposition A.2 and (3.32). Suppose now that there exists $M \in \{2, \dots, n + 2\}$ such that the bound (3.33) holds for each $m \in \{1, \dots, M - 1\}$. Then we obtain that, for some $Q(M, \text{data}) < \infty$,

$$\|\mathbf{F}_M^\varepsilon\|_{L^2(U_M)} \leq C\varepsilon^{-Q} \sum_{j=1}^M |h_j|^{\frac{M}{j}}$$

Then by the basic energy estimate, we get

$$\|\nabla w_M^\varepsilon\|_{L^2(U_M)} \leq C\varepsilon^{-Q} \sum_{j=1}^M |h_j|^{\frac{M}{j}}.$$

In view of (3.6) and the previous two displays, another application of Proposition A.2 yields, after enlarging Q and C , the bound (3.33) for $m = M$. This completes the induction argument and the proof of the lemma. \square

The typical size of $|\nabla w_m^\varepsilon|$ is much better than (3.31) gives.

Lemma 3.4. *For each $q \in [2, \infty)$, there exist $\sigma(q, \text{data}) > 0$, $C(q, \mathbf{M}, \mathbf{K}_0, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying*

$$\mathcal{X} = \mathcal{O}_\sigma(C)$$

such that, for each $m \in \{1, \dots, n+1\}$,

$$\|\nabla w_m^\varepsilon\|_{L^q(U_{m+1})} \leq \mathcal{X} \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \quad (3.34)$$

Proof. We argue by induction in m . Assume that (3.34) holds for $m \in \{1, \dots, M-1\}$ for some $M \leq n+1$. By (3.1), in particular the assumed validity of Theorem 1.3 for $m \leq n+1$, it suffices to show that

$$\|\nabla w_M^\varepsilon\|_{L^2(U_M)} \leq C\mathcal{X} \sum_{j=1}^M \|\nabla g_j\|_{L^2(U_j)}^{\frac{M}{j}}. \quad (3.35)$$

The induction hypothesis yields that

$$\|\mathbf{F}_M^\varepsilon\|_{L^2(U_M)} \leq C\mathcal{X} \sum_{j=1}^{M-1} \|\nabla g_j\|_{L^2(U_j)}^{\frac{M}{j}} \quad (3.36)$$

and then the basic energy estimate yields (3.35). \square

We also require bounds on the homogenized solutions \bar{u} and $\bar{w}_1, \dots, \bar{w}_{n+2}$. These are consequences of elliptic regularity estimates presented in Appendix D, namely Lemma D.1, which is applicable here because of the assumption (3.1) which ensures that $\bar{L} \in C^{3+\mathbf{N}, \beta}$ for every $\beta \in (0, 1)$. We obtain the existence of $C(\mathbf{M}, \text{data}) < \infty$ such that, for every $m \in \{1, \dots, n+1\}$,

$$\begin{cases} \|\nabla \bar{u}\|_{C^{0,1}(\bar{U}_2)} \leq C, \\ \|\nabla \bar{w}_m\|_{C^{0,1}(\bar{U}_{n+2})} \leq C \sum_{i=1}^m \|g_i\|_{L^2(U_i)}^{\frac{m}{i}}. \end{cases} \quad (3.37)$$

In particular, the function $\bar{\Theta}$ defined in (3.19) is Lipschitz continuous. By the global Meyers estimate, we also have, for some $\delta(d, \Lambda) > 0$ and $C(\mathbf{M}, \text{data}) < \infty$, the bound

$$\|\nabla \bar{w}_{n+2}\|_{L^{2+\delta}(\bar{U}_{n+2})} \leq C \|\bar{\mathbf{F}}_{n+2}\|_{L^{2+\delta}(U_{n+2})} \leq C \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \quad (3.38)$$

We may also apply Lemma D.1 to get a bound on w_{n+2} . In view of the merely mesoscopic gap between ∂U_{n+2} and U_{n+3} , which is much smaller than the macroscopic gaps between ∂U_m and U_{m+1} for $m \in \{1, \dots, n+1\}$, cf. (3.15) and (3.28), the estimate we obtain is, for every $\beta \in (0, 1)$,

$$\|\nabla \bar{w}_{n+2}\|_{C^{0,\beta}(\bar{U}_{n+3})} \leq C \left(3^l \varepsilon\right)^{-Q} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}, \quad (3.39)$$

for an exponent $Q(\beta, \text{data}) < \infty$ which can be explicitly computed rather easily (but for our purposes is not worth the bother) and a constant $C(\beta, \mathbf{M}, \text{data}) < \infty$.

3.4. The Mesoscopic Parameters k and l

Here we enter into a discussion regarding the choice of the parameters k and l which specify the mesoscopic scales. Recall that $1 \ll 3^k \ll 3^l \ll \varepsilon^{-1}$. As we will eventually see later in this section, we will estimate the left sides of (3.29) and (3.30) by expressions of the form

$$C \left(\varepsilon^\alpha (\varepsilon 3^l)^{-Q} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}, \tag{3.40}$$

where $Q(\text{data}) < \infty$ is a large exponent and $\alpha(\text{data}) > 0$ is a small exponent. We then need to choose k and l so that the expression in parentheses is a positive power of ε . We may do this as follows. First, we may assume throughout for convenience that $Q \geq 1 \geq 4\alpha$. Then we may take care of the first two terms by choosing l so that $\varepsilon 3^l$ is a ‘‘very large’’ mesoscale: we let l be defined so that

$$\varepsilon 3^l \leq \varepsilon^{\frac{\alpha}{4Q}} < \varepsilon 3^{l+1}.$$

Then we see that, for some $\beta > 0$,

$$\varepsilon^\alpha (\varepsilon 3^l)^{-Q} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} \leq C\varepsilon^\beta.$$

Next, we take care of the last term: since, for some $\beta > 0$,

$$(\varepsilon 3^l)^\alpha 3^{kQ} \leq C\varepsilon^\beta 3^{kQ},$$

we can make this smaller than $\varepsilon^{\frac{\beta}{2}}$ by taking $\varepsilon 3^k$ to be a ‘‘very small’’ mesoscale. We take k so that, for β and Q as in the previous display,

$$3^k \leq \varepsilon^{-\frac{\beta}{2Q}} < 3^{k+1}.$$

From this we deduce that $(\varepsilon 3^l)^\alpha 3^{kQ} \leq C\varepsilon^{\frac{\beta}{2}}$. We see that this choice of k also makes the third term inside the parentheses on the right side of (3.40) smaller than a positive power of ε . With these choices, we obtain that, for some $\beta > 0$,

$$C \left(\varepsilon^\alpha (\varepsilon 3^l)^{-Q} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}} \leq C\varepsilon^\beta. \tag{3.41}$$

Throughout the rest of this section, we will allow the exponents $\alpha \in (0, \frac{1}{4}]$ and $Q \in [1, \infty)$ to vary in each occurrence, but will always depend only on data. Similarly, we let c and C denote positive constants which may vary in each occurrence and whose dependence will be clear from the context.

3.5. The Minimal Scales

Many of the estimates we will use in the proof of Theorem 1.2 are deterministic estimates (that is, the constants C in the estimate are not random) but which are valid only above a minimal scale \mathcal{X} which satisfies $\mathcal{X} = \mathcal{O}_\delta(C)$ for some $\delta(\text{data}) > 0$ and $C(\mathbf{M}, \text{data}) < \infty$. This includes, for instance, the estimates of Theorems 1.1, 1.2 and 1.3 assumed to be valid by our assumption (3.1), as well as Lemmas 2.11, 3.1, 3.2 and 3.4, some estimates like (3.13) which appear in the text, and future estimates such as those of Lemmas 3.5, 3.6 and 3.7. We stress that this list is not exhaustive.

In our proofs of (3.29) and (3.30), we may always suppose that $3^k \geq \mathcal{X}$, where \mathcal{X} is the maximum of all of these minimal scales. In fact, we may suppose that 3^k is larger than the stationary translation $T_z \mathcal{X}$ of \mathcal{X} by any element of $z \in \mathbb{Z}^d \cap \varepsilon^{-1}U_1$. To see why, first we remark that if \mathcal{X} is any random variable satisfying $\mathcal{X} = \mathcal{O}_\delta(C)$, then a union bounds gives that, for any $Q < \infty$ the random variable

$$\tilde{\mathcal{X}} := \sup \left\{ 3^{j+1} : \sup_{z \in \mathbb{Z}^d \cap 3^Q U_1} T_z \mathcal{X} \geq 3^j \right\} \tag{3.42}$$

satisfies, for a possibly smaller δ and larger $C < \infty$, depending also on Q , the estimate $\tilde{\mathcal{X}} = \mathcal{O}_{\delta/2}(C)$. Since, as explained in the previous subsection, 3^k is a small but positive power of ε^{-1} , we see that by choosing Q suitably we have that $3^k \geq \tilde{\mathcal{X}}$ implies the validity of all of our estimates. Finally, in the event that $3^k < \tilde{\mathcal{X}}$, we can use the deterministic bounds obtained in Lemma 3.3 and (3.38) very crudely as follows:

$$\begin{aligned} \|\nabla w_{n+2}^\varepsilon - \nabla \bar{w}_{n+2}\|_{H^{-1}(U_{n+2})} &\leq C \|\nabla w_{n+2}^\varepsilon - \nabla \bar{w}_{n+2}\|_{L^2(U_{n+2})} \\ &\leq C \|\nabla w_{n+2}^\varepsilon\|_{L^2(U_{n+2})} + C \|\nabla \bar{w}_{n+2}\|_{L^2(U_{n+2})} \\ &\leq C \left(1 + \varepsilon^{-Q}\right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned}$$

Then we use that

$$(1 + \varepsilon^{-Q}) \mathbb{1}_{\{3^k < \tilde{\mathcal{X}}\}} \leq \varepsilon \tilde{\mathcal{X}}^Q = \mathcal{O}_{\delta/Q}(C\varepsilon), \tag{3.43}$$

where in the previous display the exponent Q is larger in the second instance than in the first. This yields (3.29) and (3.30) in the event $\{3^k > \tilde{\mathcal{X}}\}$, so we do not have to worry about this event.

Therefore, throughout the rest of this section, we let \mathcal{X} denote a minimal scale satisfying $\mathcal{X} = \mathcal{O}_\delta(C)$ which, in addition to both δ and C , may vary in each occurrence.

3.6. The Proof of the Local Stationary Approximation

We now turn to the proof of (3.29), which amounts to showing that the difference between the coefficient fields $(\mathbf{a}^\varepsilon, \mathbf{F}_{n+2}^\varepsilon)$ and $(\tilde{\mathbf{a}}^\varepsilon, \tilde{\mathbf{F}}_{n+2}^\varepsilon)$ is small. This is accomplished by an application of Lemma 2.11.

Lemma 3.5. *There exist $\alpha(\text{data}) > 0$, $Q(\text{data}) < \infty$ and $C(\mathbf{M}, \text{data}) < \infty$ and a minimal scale $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, if $3^k \geq \mathcal{X}$, then*

$$\begin{aligned} & \|\nabla w_{n+2}^\varepsilon - \nabla \tilde{w}_{n+2}^\varepsilon\|_{L^2(U_m)} \\ & \leq C \left(\varepsilon^\alpha (\varepsilon 3^l)^{-\frac{d}{2}-1} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned} \tag{3.44}$$

Proof. Throughout \mathcal{X} denotes a random scale which may change from line to line and satisfies $\mathcal{X} = \mathcal{O}_\delta(C)$.

For each $z \in \varepsilon 3^l \mathbb{Z}^d$ with $z + \varepsilon \square_{l+1} \subseteq U_{n+1}$, we compare ∇u and ∇w_m^ε to the functions $\nabla v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right)$ and $\nabla w_{m, \Theta(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right)$, using Lemma 2.11 and the assumed validity of Theorems 1.2 and 1.3 for n . The latter yields that, if $\varepsilon^{-1} \geq \mathcal{X}$, then for every such z , and every $m \in \{1, \dots, n + 1\}$

$$\begin{cases} \|u^\varepsilon - \bar{u}\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \leq C \varepsilon^\alpha (\varepsilon 3^l)^{-\frac{d}{2}}, \\ \|w_m^\varepsilon - \bar{w}_m\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \leq C \varepsilon^\alpha (\varepsilon 3^l)^{-\frac{d}{2}} \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{cases}$$

Likewise, if $3^l \geq \mathcal{X}$, then for every z as above, and every $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} \left\| v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{u}(z)} \right\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \leq C 3^{-l\alpha}, \\ \left\| w_{m, \Theta(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_m(z)} \right\|_{\underline{L}^2(z+\varepsilon(1+2^{-m}) \square_l)} \leq C 3^{-l\alpha} \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{cases}$$

Here we used (3.37) in order that the C not depend implicitly on $|\nabla u(z)|$ and in order that the prefactor on the right side of the second line be as it is, rather than $\sum_{j=1}^m |\nabla \bar{w}_j(z)|^{\frac{m}{j}}$. Using (3.37) again, we see that

$$\begin{cases} \|\bar{u} - (\bar{u}(z) + \ell_{\nabla \bar{u}(z)})\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \leq C \varepsilon^2 3^{2l}, \\ \|\bar{w}_m - (\bar{w}_m(z) + \ell_{\nabla \bar{w}_m(z)})\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \leq C \varepsilon^2 3^{2l} \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{cases}$$

Combining the three previous displays with the triangle inequality, we get, for all such m and z and provided that $3^l \geq \mathcal{X}$,

$$\left\{ \begin{aligned} & \left(\varepsilon 3^l \right)^{-1} \left\| u^\varepsilon - v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \bar{u}(z) \right\|_{\underline{L}^2(z+\varepsilon \square_{l+1})} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right), \\ & \left(\varepsilon 3^l \right)^{-1} \left\| w_m^\varepsilon - w_{m, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \bar{w}_m(z) \right\|_{\underline{L}^2(z+\varepsilon(1+2^{-m}) \square_l)} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right) \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{aligned} \right.$$

An application of Lemma 2.11 then yields

$$\left\{ \begin{aligned} & \left\| \nabla u^\varepsilon - \nabla v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^2(z+\varepsilon \square_l)} \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right), \\ & \left\| \nabla w_m^\varepsilon - \nabla w_{m, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^2(z+\varepsilon \square_l)} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right) \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{aligned} \right.$$

By L^p interpolation and the bounds (3.8) and (3.34), we obtain, for every $q \in [2, \infty)$, an \mathcal{X} such that $3^l \geq \mathcal{X}$ implies that, for every m and z as above,

$$\left\{ \begin{aligned} & \left\| \nabla u^\varepsilon - \nabla v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right), \\ & \left\| \nabla w_m^\varepsilon - \nabla w_{m, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right) \sum_{j=1}^m \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{m}{j}}. \end{aligned} \right.$$

This estimate implies

$$\left\{ \begin{aligned} & \left\| \mathbf{a}^\varepsilon - \mathbf{a}_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right), \\ & \left\| \mathbf{F}_{n+2}^\varepsilon - \mathbf{F}_{n+2, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + \varepsilon 3^l \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned} \right.$$

By a very similar argument, comparing the functions $\nabla v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right)$ and $\nabla w_{m, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(l)} \left(\frac{\cdot}{\varepsilon} \right)$ to $\nabla v_{\nabla \bar{u}(z), \frac{z}{\varepsilon}}^{(k)} \left(\frac{\cdot}{\varepsilon} \right)$ and $\nabla w_{m, \bar{\Theta}(z), \frac{z}{\varepsilon}}^{(k)} \left(\frac{\cdot}{\varepsilon} \right)$, also using Lemma 2.11 and the assumed

validity of Theorems 1.2 and 1.3 for n , we obtain, for all m and z as above,

$$\left\{ \begin{aligned} & \left\| \mathbf{a}_{\nabla \bar{u}(z)}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \mathbf{a}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C 3^{-k\alpha}, \\ & \left\| \mathbf{F}_{n+2, \bar{\Theta}(z)}^{(l)} \left(\frac{\cdot}{\varepsilon} \right) - \mathbf{F}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C 3^{-k\alpha} \sum_{j=1}^{n+2} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+1}{j}}. \end{aligned} \right.$$

Using (3.13) and (3.14), in view of (3.37), we find an exponent $Q(\text{data}) < \infty$ such that, for every m and z as above,

$$\left\{ \begin{aligned} & \left\| \tilde{\mathbf{a}}^\varepsilon - \mathbf{a}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C \left(\varepsilon 3^l \right) 3^{Qk}, \\ & \left\| \tilde{\mathbf{F}}_{n+2}^\varepsilon - \mathbf{F}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C \left(\varepsilon 3^l \right) 3^{Qk} \sum_{j=1}^{n+2} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+1}{j}}. \end{aligned} \right.$$

Combining the above estimates, we finally obtain that, for every z and m as above,

$$\left\{ \begin{aligned} & \left\| \mathbf{a}^\varepsilon - \tilde{\mathbf{a}}^\varepsilon \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + 3^{-k\alpha} + \varepsilon 3^{l+kQ} \right), \\ & \left\| \mathbf{F}_{n+2}^\varepsilon - \tilde{\mathbf{F}}_{n+2}^\varepsilon \right\|_{\underline{L}^q(z+\varepsilon \square_l)} \\ & \leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + 3^{-k\alpha} + \varepsilon 3^{l+kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}, \end{aligned} \right.$$

Using that these coefficient fields are bounded and that the boundary layer (that is, the set of points which lie in U_{n+2} but not in any cube of the form $z + \varepsilon \square_l$ with z as above) has thickness $O(\varepsilon 3^l)$, we get from the previous estimate that

$$\left\{ \begin{aligned} & \left\| \mathbf{a}^\varepsilon - \tilde{\mathbf{a}}^\varepsilon \right\|_{\underline{L}^q(U_{n+2})} \leq C \mathbf{A}, \\ & \left\| \mathbf{F}_{n+2}^\varepsilon - \tilde{\mathbf{F}}_{n+2}^\varepsilon \right\|_{\underline{L}^q(U_{n+2})} \leq C \mathbf{A} \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}, \end{aligned} \right.$$

where (in order to shorten the expressions) we denote

$$\mathbf{A} := \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-\frac{d}{2}-1} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + 3^{-k\alpha} + \left(\varepsilon 3^l \right)^\alpha 3^{kQ} \right).$$

To complete the proof of the lemma, we observe that $\zeta := \tilde{w}_{n+2}^\varepsilon - w_{n+2}^\varepsilon \in H_0^1(U_{n+2})$ satisfies the equation

$$-\nabla \cdot \tilde{\mathbf{a}}^\varepsilon \nabla \zeta = \nabla \cdot \left(\tilde{\mathbf{F}}_{n+2}^\varepsilon - \mathbf{F}_{n+2}^\varepsilon \right) + \nabla \cdot \left((\tilde{\mathbf{a}}^\varepsilon - \mathbf{a}^\varepsilon) \nabla w_{n+2}^\varepsilon \right). \tag{3.45}$$

By the basic energy estimate (test the equation for ζ with ζ), we obtain

$$\|\nabla \zeta\|_{L^2(U_{n+2})} \leq \left(\left\| \tilde{\mathbf{F}}_{n+2}^\varepsilon - \mathbf{F}_{n+2}^\varepsilon \right\|_{L^2(U_{n+2})} + \left\| (\tilde{\mathbf{a}}^\varepsilon - \mathbf{a}^\varepsilon) \nabla w_{n+2}^\varepsilon \right\|_{L^2(U_{n+2})} \right).$$

The first term on the right side is already estimated above. For the second term, we use the Meyers estimate and (3.34), without forgetting (3.20), to find $\delta(d, \Lambda) > 0$ such that

$$\|\nabla w_{n+2}^\varepsilon\|_{\underline{L}^{2+\delta}(U_{n+2})} \leq C \|\mathbf{F}_{n+2}^\varepsilon\|_{L^{2+\delta}(U_{n+2})} \leq C \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}.$$

Therefore, by Hölder’s inequality and the above estimate on $\|\mathbf{a}^\varepsilon - \tilde{\mathbf{a}}^\varepsilon\|_{\underline{L}^q(z+\varepsilon\Box_l)}$ with exponent $q := \frac{4+2\delta}{\delta}$, we obtain

$$\begin{aligned} & \|(\tilde{\mathbf{a}}^\varepsilon - \mathbf{a}^\varepsilon) \nabla w_{n+2}^\varepsilon\|_{L^2(U_{n+2})} \\ & \leq \|\tilde{\mathbf{a}}^\varepsilon - \mathbf{a}^\varepsilon\|_{L^{\frac{4+2\delta}{\delta}}(U_1)} \|\nabla w_{n+2}^\varepsilon\|_{L^{2+\delta}(U_{n+2})} \leq CA \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned}$$

This completes the proof of (3.44). \square

In view of the discussion in Sections 3.4 and 3.5, Lemma 3.5 implies (3.29).

3.7. Homogenization Estimates for the Approximating Equations

To prepare for the proof of (3.29), we apply the quantitative homogenization estimates proved in [4] for the linear equation (3.4). We denote by $\tilde{\mathbf{a}}_p^{(k)}$ and $\tilde{\mathbf{F}}_{n+2,\Theta}^{(k)}$ the homogenized coefficients.

By an application of the results of [4, Chapter 11], the solutions $v(\cdot, U, e, \Theta)$ of the family of Dirichlet problems

$$\begin{cases} -\nabla \cdot (\mathbf{a}_p^{(k)} \nabla v(\cdot, U, e, \Theta)) = \nabla \cdot \mathbf{F}_{n+2,\Theta}^{(k)} & \text{in } U, \\ v(\cdot, U, e, \Theta) = \ell_e & \text{on } \partial U, \end{cases} \quad (3.46)$$

indexed over bounded Lipschitz domains $U \subseteq \mathbb{R}^d$ and $e \in B_1$ and $\Theta \in \mathbb{R}^{d(n+2)}$, satisfy the following quantitative homogenization estimate: there exist $\alpha(s, d, \Lambda) > 0$, $\beta(\text{data}) > 0$, $\delta(\text{data}) > 0$ and $Q(\text{data}) < \infty$ and $C(s, \mathbf{M}, \text{data}) < \infty$ and a random variable $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, for every $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p| \leq \mathbf{M}$, $e \in \mathbb{R}^d$ and every $l \in \mathbb{N}$ with $3^{l-k} \geq \mathcal{X}$,

$$\begin{aligned} & 3^{-l} \|v(\cdot, \Box_{l+1}, e, \Theta) - \ell_e\|_{\underline{L}^2(\Box_l)} + 3^{-l} \|\nabla v(\cdot, \Box_{l+1}, e, \Theta) - e\|_{\underline{H}^{-1}(\Box_l)} \\ & + 3^{-l} \left\| \mathbf{a}_p^{(k)} \nabla v(\cdot, \Box_{l+1}, e, \Theta) + \mathbf{F}_{n+2,\Theta}^{(k)} - (\tilde{\mathbf{a}}_p^{(k)} e + \tilde{\mathbf{F}}_{n+2,\Theta}^{(k)}) \right\|_{\underline{H}^{-1}(\Box_l)} \\ & \leq C 3^{-\alpha(l-k)} 3^{Qk} \left(|e| + \sum_{j=1}^{n+1} |h_j|^{\frac{n+2}{j}} \right) \end{aligned} \quad (3.47)$$

as well as

$$\|\nabla v(\cdot, \Box_{l+1}, e, \Theta)\|_{C^{0,\beta}(\Box_l)} \leq Cl^Q 3^{Qk} \left(|e| + \sum_{j=1}^{n+1} |h_j|^{\frac{n+2}{j}} \right). \quad (3.48)$$

To obtain (3.47), we change the scale by performing a dilation $x \mapsto 3^k \left[(15 + d)^{\frac{1}{2}} \right] x$ so that the resulting coefficient fields have a unit range of dependence and are still \mathbb{Z}^d -stationary, cf. (3.3). We then apply [4, Lemma 11.11] to obtain (3.47), using Lemma 3.1, namely (3.5), to give a bound on the parameter K in the application of the former. The reason we quote results for general nonlinear equations, despite the fact that (3.46) is linear, is because (3.46) has a nonzero right-hand side and the results for linear equations have not been formalized in that generality.

The bound (3.48) is a consequence the large-scale $C^{0,1}$ -type estimate for the equation (3.4) (see [4, Theorem 11.13]) which, together with a routine union bound, yields (3.48) with the left side replaced by $\sup_{z \in \mathbb{Z}^d \cap \square_l} \|\nabla v\|_{\underline{L}^2(z+(l-k)\square_k)}$. Indeed, if we denote by \mathcal{X}_z the minimal scale in the large-scale $C^{0,1}$ estimate, then a union bound yields (for any $s \in (0, d)$, so in particular we can take $s > 1$),

$$\max \left\{ k' \in \mathbb{N} : \max_{z \in \mathbb{Z}^d \cap \square_{k'}} \mathcal{X}_z > \left(\log 3^{k'-k} \right)^{\frac{1}{s}} \right\} \leq \mathcal{O}_\delta(C).$$

We obtain (3.48) by combining this bound for $\sup_{z \in \mathbb{Z}^d \cap \square_l} \|\nabla v\|_{\underline{L}^2(z+(l-k)\square_k)}$ with the deterministic bounds (3.5), (3.6) and an application of the Schauder estimates (see Proposition A.2).

We also need the following estimate for the difference of v 's on overlapping cubes, which can be inferred from (3.47) and the Caccioppoli inequality: there exist $\alpha(s, d, \Lambda) > 0, \beta(\text{data}) > 0, \delta(\text{data}) > 0$ and $Q(\text{data}) < \infty$ and $C(s, \mathbf{M}, \text{data}) < \infty$ and a random variable $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, for every $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p| \leq \mathbf{M}, e \in \mathbb{R}^d$ and every $l \in \mathbb{N}$ with $3^{l-k} \geq \mathcal{X}$,

$$\begin{aligned} & \sum_{z' \in 3^l \mathbb{Z}^d \cap \square_{l+1}} \left\| \nabla v(\cdot, z + \square_{l+1}, e, \Theta) - \nabla v(\cdot, z' + \square_l, e, \Theta) \right\|_{\underline{L}^2(z' + \square_l)} \\ & \leq C 3^{-\alpha(l-k)} \left(|e| + 3^{Qk} \sum_{j=1}^{n+1} |h_j|^{\frac{n+2}{j}} \right). \end{aligned} \tag{3.49}$$

Finally, we mention that we have the following deterministic estimate on the functions $v(\cdot, z + \square_{l+1}, e, \Theta)$ which states that

$$\begin{aligned} \left\| \nabla v(\cdot, z + \square_{l+1}, e, \Theta) \right\|_{\underline{L}^2(z + \square_{l+1})} & \leq C \left(|e| + \left\| \mathbf{F}_{n+2, \Theta}^{(k)} \right\|_{\underline{L}^2(z + \square_{l+1})} \right) \\ & \leq C \left(|e| + \sum_{j=1}^{n+1} |h_j|^{\frac{n+2}{j}} \right). \end{aligned} \tag{3.50}$$

This follows from testing the problem (3.46) with $U = z + \square_{l+1}$ with the test function $v(\cdot, z + \square_{l+1}, e, \Theta) - \ell_e \in H_0^1(z + \square_{l+1})$ and then using (3.9) with $m = n + 2$.

Our next goal is to compare the homogenized coefficients $\overline{\mathbf{F}}_{m, \Theta}^{(k)}$ for the approximating equations to the functions $\overline{\mathbf{F}}_m(\Theta)$ defined in (1.16). In view of the results of Section 2, it is natural to proceed by comparing the vector fields $\mathbf{F}_{m, \Theta}^{(k)}$ to $\mathbf{f}_{p, h}^{(m)}$ defined in (2.5). This is accomplished by invoking Lemma 2.11 again.

Lemma 3.6. Fix $q \in [2, \infty)$ and $M \in [1, \infty)$. There exist $\delta(q, \text{data}), \alpha(q, \text{data}) \in (0, \frac{1}{2}]$, $C(q, M, \text{data}) < \infty$ such that, for every $k \in \mathbb{N}$ with $3^k \geq \mathcal{X}$, $m \in \{1, \dots, n + 2\}$ and $p, h \in \mathbb{R}^d$ with $|p| \leq M$, we have

$$\left\| \mathbf{a}_p - \mathbf{a}_p^{(k)} \right\|_{\underline{L}^q(\square_k)} \leq \mathcal{O}_\delta \left(C 3^{-k\alpha} \right) \tag{3.51}$$

and, for $\Theta := (p, h, 0, \dots, 0) \in \mathbb{R}^{d(n+2)}$,

$$\left\| \mathbf{F}_{m, \Theta}^{(k)} - \mathbf{f}_{p, h}^{(m)} \right\|_{\underline{L}^q(\square_k)} \leq \mathcal{O}_\delta \left(C 3^{-k\alpha} |h|^m \right). \tag{3.52}$$

Proof. We take \mathcal{X} to be the maximum of the random scales in Theorem 2.2, Lemmas 2.10 and 2.11 and Theorems 1.2 and 1.3 for n , the latter two being valid by assumption (3.1). Assume $3^k \geq \mathcal{X}$.

By Lemma 2.10 and the assumed validity of Theorem 1.2 for n , we have

$$\begin{aligned} & 3^{-k} \left\| v_{p, 0}^{(k)} - \left(\ell_p + \phi_p - (\phi_p)_{\square_{k+1}} \right) \right\|_{\underline{L}^2(\square_{k+1})} \leq C 3^{-k\alpha}, \\ & 3^{-k} \left\| w_{1, \Theta, 0}^{(k)} - \left(\ell_h + \psi_{p, h}^{(1)} - \left(\psi_{p, h}^{(1)} \right)_{\frac{3}{2}\square_k} \right) \right\|_{\underline{L}^2(\frac{3}{2}\square_k)} \leq C 3^{-k\alpha} |h| \end{aligned}$$

and, for $m \in \{2, \dots, n + 1\}$,

$$3^{-k} \left\| w_{m, \Theta, 0}^{(k)} - \left(\psi_{p, h}^{(m)} - \left(\psi_{p, h}^{(m)} \right)_{\frac{3}{2}\square_k} \right) \right\|_{\underline{L}^2((1+2^{-m})\square_k)} \leq C 3^{-k\alpha} |h|^m.$$

By Theorem 2.2 and the assumed validity of Theorem 1.3 for n , we also have that, for every $m \in \{1, \dots, n + 1\}$,

$$\left\| \nabla w_{m, \Theta, 0}^{(k)} \right\|_{\underline{L}^2((1+2^{-m})\square_k)} + \left\| \nabla \psi_{p, h}^{(m)} \right\|_{\underline{L}^2((1+2^{-m})\square_k)} \leq C |h|^m. \tag{3.53}$$

Using these estimates and Lemma 2.11, we obtain, for $m \in \{2, \dots, n + 1\}$,

$$\left\| \nabla w_{m, \Theta, 0}^{(k)} - \nabla \psi_{p, h}^{(m)} \right\|_{\underline{L}^2(\square_k)} \leq C 3^{-k\alpha} |h|^m. \tag{3.54}$$

The previous display holds for $3^k \geq \mathcal{X}$. Combining this estimate with (2.16) and (3.10) and using L^p interpolation, we obtain

$$\left\| \nabla w_{m, \Theta, 0}^{(k)} - \nabla \psi_{p, h}^{(m)} \right\|_{\underline{L}^2(\square_k)} \leq \mathcal{O}_\delta \left(C 3^{-k\alpha} |h|^m \right). \tag{3.55}$$

The conclusion of the lemma now follows. \square

We next observe that the homogenized coefficients $\bar{\mathbf{a}}_p^{(k)}$ and $\bar{\mathbf{F}}_{n+2, \Theta}^{(k)}$ agree, up to a small error, with $\bar{\mathbf{a}}_p := D^2 \bar{L}(p)$ and $\bar{\mathbf{F}}_{n+2}(\Theta)$. This will help us eventually to prove that the homogenized coefficients for the linearized equations are the linearized coefficients of the homogenized equation.

Lemma 3.7. Fix $M \in [1, \infty)$. There exist $\alpha(\text{data}) > 0$ and $C(n, M, \text{data}) < \infty$ such that, for every $\Theta = (p, h_1, \dots, h_{n+1}) \in B_M \times \mathbb{R}^{d(n+1)}$,

$$\left| \bar{\mathbf{a}}_p - \bar{\mathbf{a}}_p^{(k)} \right| \leq C 3^{-k\alpha} \tag{3.56}$$

and

$$\left| \bar{\mathbf{F}}_{n+2, \Theta}^{(k)} - \bar{\mathbf{F}}_{n+2}(\Theta) \right| \leq C 3^{-k\alpha} \left(\sum_{j=1}^{n+1} |h_j| \frac{n+1}{j} \right). \tag{3.57}$$

Proof. We first give the argument in the case that $\Theta = (p, h, 0, \dots, 0)$. From the definition (1.16) and Proposition 2.8, we see that $(\bar{\mathbf{a}}_p, \bar{\mathbf{F}}_{n+2}(\Theta))$ are the homogenized coefficients for the stationary coefficient fields $(\mathbf{a}_p, \mathbf{f}_{p,h}^{(n+2)})$. By Lemma 3.6, and the bounds (3.5) and (2.16), we find that, for every $q \in [2, \infty)$,

$$\left\| \mathbf{a}_p - \mathbf{a}_p^{(k)} \right\|_{\underline{L}^q(\square_k)} \leq \mathcal{O}_\delta \left(3^{-k\alpha} \right) \tag{3.58}$$

and

$$\left\| \mathbf{F}_{m, \Theta}^{(k)} - \mathbf{f}_{p,h}^{(m)} \right\|_{\underline{L}^q(\square_k)} \leq \mathcal{O}_\delta \left(C 3^{-k\alpha} |h|^m \right), \tag{3.59}$$

with the constants C depending additionally on q . The result now follows easily by the Meyers estimate in the case $\Theta = (p, h, 0, \dots, 0)$. Indeed, by subtracting the equations and applying Hölder’s inequality and the Meyers estimate, we can show that the first-order correctors for the two linear problems are close in the $\underline{L}^2(\square_k)$ norm (with a second moment in expectation). Therefore their fluxes are also close, and taking the expectations of the fluxes gives the homogenized coefficients. This argument can also be performed in a finite volume box \square_M with the result obtained after sending $M \rightarrow \infty$. See [1, Lemma 3.2] for a similar argument.

For general Θ the result follows by polarization (see Lemma 2.6) and the multilinear structure shared by the functions $\Theta \mapsto \bar{\mathbf{F}}_{n+2, \Theta}^{(k)}$ and $\Theta \mapsto \bar{\mathbf{F}}_{n+2}(\Theta)$. \square

We also require the following continuity estimate for the functions $w(\cdot, \square_l, e, \Theta)$ in (e, Θ) . Following the proof of Lemma 3.2 for L^2 dependence in Θ (we need need to do one more step of the iteration described there), using (3.47) for L^2 dependence in e , and then interpolating this result with (3.48), we obtain exponents $\alpha(\text{data}) > 0$ and $Q(\text{data}) < \infty$ and a random variable $\mathcal{X} = \mathcal{O}_\delta(C)$ such that, for every $e, e' \in \mathbb{R}^d$ and $\Theta = (p, h_1, \dots, h_{n+1}) \in \mathbb{R}^{d(n+2)}$ and $\Theta' = (p', h'_1, \dots, h'_{n+1}) \in \mathbb{R}^{d(n+2)}$ with $|p|, |p'| \leq M$, if $3^k \geq \mathcal{X}$, then

$$\begin{aligned} & \left\| \nabla v(\cdot, \square_{l+1}, e, \Theta) - \nabla v(\cdot, \square_{l+1}, e', \Theta') \right\|_{L^\infty(\square_l)} \\ & \leq C l^Q 3^{kQ} |p - p'|^\alpha \sum_{i=1}^{m-1} (|h_i| \vee |h'_i|)^{\frac{m}{i}} \\ & \quad + C l^Q 3^{kQ} \sum_{i=1}^{m-1} |h_i - h'_i|^\alpha \sum_{j=1}^{m-i} \left(|h_j| \vee |h'_j| \right)^{\frac{m-i}{j}}. \end{aligned} \tag{3.60}$$

3.8. Homogenization of the Locally Stationary Equation

We next present the proof of (3.30), which is the final step in the proof of Theorem 1.2. The argument follows a fairly routine (albeit technical) two-scale expansion argument and requires the homogenization estimates given in Section 3.7.

The proof of (3.30). In view of the discussion in Section 3.5, we may assume throughout that $3^k \geq \mathcal{X}$ where \mathcal{X} is any minimal scale described above. We begin by building an approximation of the solution $\tilde{w}_{n+2}^\varepsilon$ of (3.25) built out of the solutions of (3.46). We select a smooth function $\zeta \in C_c^\infty(\mathbb{R}^d)$ such that, for some $C(d) < \infty$ and each $i \in \{1, 2\}$,

$$\mathbb{1}_{\varepsilon \square_l} \leq \zeta \leq \mathbb{1}_{\varepsilon \square_{l+1}} \quad \left\| \nabla^i \zeta \right\|_{L^\infty} \leq C \left(\varepsilon 3^l \right)^{-i}, \quad \text{and} \quad \sum_{z \in \varepsilon 3^l \mathbb{Z}^d} \zeta(\cdot - z) = 1 \quad \text{in } \mathbb{R}^d. \quad (3.61)$$

We can construct such a ζ by a suitable mollification of $\mathbb{1}_{\varepsilon \square_l}$. We write $\zeta_z := \zeta(\cdot - z)$ for short. We next define a function $\tilde{v}_{n+2}^\varepsilon \in H_0^1(U_{n+2})$ by

$$\begin{aligned} \tilde{v}_{n+2}^\varepsilon(x) &:= \bar{w}_{n+2}(x) \\ &+ \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\varepsilon v \left(\frac{x}{\varepsilon}, \frac{z}{\varepsilon} + \square_{l+1}, \nabla \bar{w}_{n+2}(z), \bar{\Theta}(z) \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right). \end{aligned}$$

Since $\tilde{v}_{n+2}^\varepsilon - \tilde{w}_{n+2}^\varepsilon \in H_0^1(U_{n+2})$, We have that

$$\begin{aligned} \left\| \nabla \tilde{v}_{n+2}^\varepsilon - \nabla \tilde{w}_{n+2}^\varepsilon \right\|_{L^2(U_{n+2})} &\leq C \left\| \nabla \cdot \tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon - \nabla \cdot \tilde{\mathbf{a}}^\varepsilon \nabla \tilde{w}_{n+2}^\varepsilon \right\|_{H^{-1}(U_{n+2})} \\ &= C \left\| \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon + \tilde{\mathbf{F}}_{n+2}^\varepsilon \right) \right\|_{H^{-1}(U_{n+2})}. \end{aligned} \quad (3.62)$$

A preliminary goal is therefore to prove the estimate

$$\begin{aligned} &\left\| \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon + \tilde{\mathbf{F}}_{n+2}^\varepsilon \right) \right\|_{H^{-1}(U_{n+2})} \\ &\leq C \left(\varepsilon^\alpha \left(\varepsilon 3^l \right)^{-Q} + 3^{-l\alpha} \left(\varepsilon 3^l \right)^{-1} + 3^{-k\alpha} + \left(\varepsilon 3^l \right)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \left\| \nabla g_j \right\|_{L^{2+\delta}(U_j)}. \end{aligned} \quad (3.63)$$

The proof of (3.63) is essentially completed in Steps 1–5 below: see (3.72) and (3.73). After obtaining this estimate, we will prove that

$$\left\| \nabla \tilde{v}_{n+2}^\varepsilon - \nabla \bar{w}_{n+2} \right\|_{H^{-1}(U_{n+2})} \leq C 3^{-\alpha(l-k)} 3^{Qk} \sum_{i=1}^{n+1} \left\| g_i \right\|_{L^2(U_i)}^{\frac{n+2}{i}}. \quad (3.64)$$

Together these inequalities imply (3.30).

Denote $v_z := v \left(\cdot, \frac{\cdot}{\varepsilon} + \square_{l+1}, \nabla \bar{w}_{n+2}(z), \bar{\Theta}(z) \right)$ for short and compute

$$\begin{aligned} \nabla \tilde{v}_{n+2}^\varepsilon(x) &= \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \left(\nabla \zeta_z(x) \left(\varepsilon v_z \left(\frac{x}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) + \zeta_z(x) \nabla v_z \left(\frac{x}{\varepsilon} \right) \right) \\ &+ \left(\nabla \bar{w}_{n+2}(x) - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \nabla \bar{w}_{n+2}(z) \right). \end{aligned} \tag{3.65}$$

Thus

$$\begin{aligned} \tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon &= \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \\ &+ \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \tilde{\mathbf{a}}^\varepsilon \nabla \zeta_z \left(\varepsilon v_z \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \\ &+ \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \\ &+ \tilde{\mathbf{a}}^\varepsilon \left(\nabla \bar{w}_{n+2}(x) - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \nabla \bar{w}_{n+2}(z) \right). \end{aligned} \tag{3.66}$$

Using the equation for v_z , we find that

$$\begin{aligned} \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon + \tilde{\mathbf{F}}_{n+2}^\varepsilon \right) &= \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \cdot \left(\tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \\ &+ \nabla \cdot \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \\ &+ \nabla \cdot \left(\tilde{\mathbf{F}}_{n+2}^\varepsilon - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \\ &+ \nabla \cdot \left(\sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \tilde{\mathbf{a}}^\varepsilon \nabla \zeta_z \left(\varepsilon v_z \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right) \\ &+ \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \left(\nabla \bar{w}_{n+2}(x) - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \nabla \bar{w}_{n+2}(z) \right) \right). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} &\left\| \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon + \tilde{\mathbf{F}}_{n+2}^\varepsilon \right) \right\|_{H^{-1}(U_{n+2})} \\ &\leq C \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \cdot \left(\tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \right\|_{H^{-1}(U_{n+2})} \end{aligned}$$

$$\begin{aligned}
 &+ C \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(U_{n+2})} \\
 &+ C \left\| \tilde{\mathbf{F}}_{n+2}^\varepsilon - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(U_{n+2})} \\
 &+ C \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(\varepsilon v_z \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right\|_{L^2(U_{n+2})} \\
 &+ C \left\| \nabla \bar{w}_{n+2}(x) - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \nabla \bar{w}_{n+2}(z) \right\|_{L^2(U_{n+2})}. \tag{3.67}
 \end{aligned}$$

We next estimate each of the five terms on the right side of the previous display.

Step 1. The estimate of the first term on the right side of (3.67). This is where we use the homogenized equation for \bar{w}_{n+2} . Fix $h \in H_0^1(U_{n+2})$ with $\|h\|_{H^1(U_{n+2})} = 1$. By the equation for \bar{w}_{n+2} and integration by parts that

$$\begin{aligned}
 0 &= \int_{U_{n+2}} \nabla h(x) \cdot \left(\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}(x) + \bar{\mathbf{F}}_{n+2}(x) \right) dx \\
 &= - \int_{U_{n+2}} h(x) \sum_{z \in 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z(x) \cdot \left(\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}(x) + \bar{\mathbf{F}}_{n+2}(x) \right) dx \\
 &\quad + \int_{U_{n+2}} \nabla h(x) \cdot \sum_{z \in 3^l \mathbb{Z}^d \setminus U_{n+3}} \zeta_z(x) \left(\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}(x) + \bar{\mathbf{F}}_{n+2}(x) \right) dx.
 \end{aligned}$$

Therefore we may have

$$\begin{aligned}
 &\left| \int_{U_{n+2}} h(x) \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z(x) \cdot \left(\tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{x}{\varepsilon} \right) \nabla v_z \left(\frac{x}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{x}{\varepsilon} \right) \right) dx \right| \\
 &\leq \left| \int_{U_{n+2}} h \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \cdot \left(\tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) - \bar{\mathbf{a}} \nabla \bar{w}_{n+2} + \bar{\mathbf{F}}_{n+2} \right) dx \right| \\
 &\quad + \int_{U_{n+2}} |\nabla h(x)| \left| \sum_{z \in 3^l \mathbb{Z}^d \setminus U_{n+3}} \zeta_z(x) \left(\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}(x) + \bar{\mathbf{F}}_{n+2}(x) \right) \right| dx.
 \end{aligned}$$

For the first term on the right, we use (3.37), (3.47), (3.56) and (3.57) to obtain

$$\begin{aligned}
 &\left| \int_{U_{n+2}} h \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \cdot \left(\tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) - \bar{\mathbf{a}} \nabla \bar{w}_{n+2} + \bar{\mathbf{F}}_{n+2} \right) dx \right| \\
 &\leq C \|h\|_{H^1(U_{n+2})} \left(3^{-\alpha(l-k)} 3^{Qk} + 3^{-k\alpha} \right) \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{l}}.
 \end{aligned}$$

For the second term, we denote

$$U_{n+4} := \{x \in U_{n+3} : x + \varepsilon \square_{l+1} \subseteq U_{n+3}\}, \tag{3.68}$$

observe that $|U_{n+2} \setminus U_{n+4}| \leq C\varepsilon 3^l$, and compute, using the Hölder inequality, to obtain, using also (3.38),

$$\begin{aligned} & \int_{U_{n+2}} |\nabla h(x)| \left| \sum_{z \in 3^l \mathbb{Z}^d \setminus U_{n+3}} \zeta_z(x) (\bar{\mathbf{a}}(x) \nabla \bar{w}_{n+2}(x) + \bar{\mathbf{F}}_{n+2}(x)) \right| dx \\ & \leq C \|\nabla h\|_{L^2(U_{n+2})} \|\bar{\mathbf{a}} \nabla \bar{w}_{n+2} + \bar{\mathbf{F}}_{n+2}\|_{L^2(U_{n+2} \setminus U_{n+4})} \\ & \leq C \|\nabla h\|_{L^2(U_{n+2})} \left((\varepsilon 3^l)^{\frac{\delta}{4+2\delta}} \|\nabla \bar{w}_{n+2}\|_{L^{2+\delta}(U_{n+2})} + \varepsilon 3^l \|\bar{\mathbf{F}}_{n+2}\|_{L^\infty(U_{n+2})} \right) \\ & \leq C \|\nabla h\|_{L^2(U_{n+2})} (\varepsilon 3^l)^\alpha \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Combining the above estimates and taking the supremum over $h \in H_0^1(U_{n+2})$ with $\|h\|_{H^1(U_{n+2})} = 1$ yields

$$\begin{aligned} & \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \cdot \left(\bar{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) + \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \right\|_{H^{-1}(U_{n+2})} \\ & \leq C \left(3^{-\alpha(l-k)} 3^{Qk} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha \right) \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Step 2. The estimate of the second term on the right side of (3.67). We use (3.14), (3.37) and (3.50) to find

$$\begin{aligned} & \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(U_{n+2})} \\ & \leq \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \left\| \left(\tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right) \nabla v_z \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(z + \varepsilon \square_{l+1})} \\ & \leq C \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \left\| \tilde{\mathbf{a}}^\varepsilon - \tilde{\mathbf{a}}_{\nabla \bar{u}(z)}^{(k)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^\infty(z + \varepsilon \square_{l+1})} \|\nabla v_z \left(\frac{\cdot}{\varepsilon} \right)\|_{L^2(z + \varepsilon \square_{l+1})} \\ & \leq C 3^{Qk} (\varepsilon 3^l)^\alpha \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Step 3. The estimate of the third term on the right side of (3.67). We have by (3.13) and (3.37) that

$$\begin{aligned} & \left\| \tilde{\mathbf{F}}_{n+2}^\varepsilon - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(U_{n+2})} \\ & \leq C \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \left(\tilde{\mathbf{F}}_{n+2}^\varepsilon - \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)} \left(\frac{\cdot}{\varepsilon} \right) \right) \right\|_{L^2(U_{n+2})} \\ & \leq C \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \left\| \tilde{\mathbf{F}}_{n+2}^\varepsilon - \tilde{\mathbf{F}}_{n+2, \bar{\Theta}(z)} \left(\frac{\cdot}{\varepsilon} \right) \right\|_{L^2(z + \varepsilon \square_{l+1})} \\ & \leq C 3^k Q^k \left(\varepsilon 3^l \right)^\alpha \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Step 4. The estimate of the fourth term on the right side of (3.67). For convenience, extend v_z so that it is equal to the affine function $\ell_{\nabla \bar{w}_{n+2}(z)}$ outside of the cube $z + \varepsilon \square_{l+1}$. By the triangle inequality, we have

$$\begin{aligned} & \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(\varepsilon v_z \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right\|_{L^2(U_{n+2})} \\ & \leq \left\| \sum_{y, z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(\varepsilon v_y \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(y)} \right) \right\|_{L^2(U_{n+2})} \\ & \quad + \left\| \sum_{y, z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(v_z \left(\frac{\cdot}{\varepsilon} \right) - v_y \left(\frac{\cdot}{\varepsilon} \right) \right) \right\|_{L^2(U_{n+2})} \\ & \quad + \left\| \sum_{y, z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}, y \sim z} \nabla \zeta_z \varepsilon \left| \nabla \bar{w}_{n+2}(z) - \nabla \bar{w}_{n+2}(y) \right| \right\|_{L^2(U_{n+2})}, \quad (3.69) \end{aligned}$$

where here and below we use the notation $y \sim z$ to mean that $y + \square_{l+1}$ and $z + \square_{l+1}$ have nonempty intersection. Using (3.61) and the fact that $\sum_{z \in \varepsilon 3^l \mathbb{Z}^d} \nabla \zeta_z = 0$, we have that

$$\left| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \right| \leq C \left(\varepsilon 3^l \right)^{-1} \mathbf{1}_V. \quad (3.70)$$

where

$$V := \left\{ x \in \mathbb{R}^d : (x + \varepsilon \square_{l+1}) \cap \partial U_{n+3} \neq \emptyset \right\} \quad (3.71)$$

is a boundary layer of U_{n+3} of thickness $\varepsilon 3^l$. Observe that $|V| \leq C\varepsilon 3^l$. Thus by (3.47) we find that

$$\begin{aligned} & \left\| \sum_{y,z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(\varepsilon v_y \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(y)} \right) \right\|_{L^2(U_{n+2})} \\ & \leq C \left(\varepsilon 3^l \right)^{-1} \varepsilon \left\| \mathbb{1}_V \sum_{y \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \left(\varepsilon v_y \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(y)} \right) \right\|_{L^2(U_{n+2})} \\ & \leq C 3^{-\alpha(l-k)} 3^{Qk} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Combining (3.61) with (3.49), (3.37) and the triangle inequality to compare the v_z 's on overlapping cubes, we find that

$$\begin{aligned} & \left\| \varepsilon \sum_{y,z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(v_z \left(\frac{\cdot}{\varepsilon} \right) - v_y \left(\frac{\cdot}{\varepsilon} \right) \right) \right\|_{L^2(U_{n+2})} \\ & \leq \sum_{y,z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}, y \sim z} \|\nabla \zeta_z\|_{L^\infty(\mathbb{R}^d)} \cdot \varepsilon \|v_y - v_z\|_{L^2((y+\square_{l+1}) \cap (z+\square_{l+1}))} \\ & \leq C 3^{-\alpha(l-k)} 3^{Qk} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

For the last term, we compute, using (3.39),

$$\begin{aligned} & \left\| \sum_{y,z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}, y \sim z} \nabla \zeta_z \varepsilon |\nabla \bar{w}_{n+2}(z) - \nabla \bar{w}_{n+2}(y)| \right\|_{L^2(U_{n+2})} \\ & \leq C \varepsilon^{1+\beta} 3^l \|\nabla \bar{w}_{n+2}\|_{C^{0,\beta}(U_{n+3})} \left\| \sum_{y,z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}, y \sim z} |\nabla \zeta_z| \right\|_{L^2(U_{n+2})} \\ & \leq C \varepsilon^\alpha \left(3^l \varepsilon \right)^{-Q} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Putting the above inequalities together, we find that

$$\begin{aligned} & \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \nabla \zeta_z \left(\varepsilon v_z \left(\frac{\cdot}{\varepsilon} \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right\|_{L^2(U_{n+2})} \\ & \leq C \left(3^{-\alpha(l-k)} 3^{Qk} + \varepsilon^\alpha \left(3^l \varepsilon \right)^{-Q} \right) \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Step 5. The estimate of the fifth term on the right side of (3.67). Recall that U_{n+4} is defined in (3.68). We have that, using (3.38) and (3.39),

$$\begin{aligned} & \left\| \nabla \bar{w}_{n+2} - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \nabla \bar{w}_{n+2}(z) \right\|_{L^2(U_{n+2})} \\ & \leq \left\| \nabla \bar{w}_{n+2} \left(1 - \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z \right) \right\|_{L^2(U_{n+2})} \\ & \quad + \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z (\nabla \bar{w}_{n+2} - \nabla \bar{w}_{n+2}(z)) \right\|_{L^2(U_{n+2})} \\ & \leq \|\nabla \bar{w}_{n+2}\|_{L^2(U_{n+2} \setminus U_{n+4})} + C \varepsilon 3^l \|\nabla \bar{w}_{n+2}\|_{C^{0,\beta}(U_{n+3})} \\ & \leq C |U_{n+2} \setminus U_{n+4}|^{\frac{\delta}{4+2\delta}} \|\nabla \bar{w}_{n+2}\|_{L^{2+\delta}(U_{n+2})} + C \varepsilon^\alpha (\varepsilon 3^l)^{-Q} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}} \\ & \leq C \left((\varepsilon 3^l)^\alpha + \varepsilon^\alpha (\varepsilon 3^l)^{-Q} \right) \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

Step 6. Let us summarize what we have shown so far. Substituting the results of the five previous steps into (3.67), we obtain that

$$\begin{aligned} & \left\| \nabla \cdot \left(\tilde{\mathbf{a}}^\varepsilon \nabla \tilde{v}_{n+2}^\varepsilon + \tilde{\mathbf{F}}_{n+2}^\varepsilon \right) \right\|_{H^{-1}(U_{n+2})} \\ & \leq C \left(\varepsilon^\alpha (\varepsilon 3^l)^{-Q} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned} \tag{3.72}$$

This implies by (3.62) that

$$\begin{aligned} & \left\| \nabla \tilde{v}_{n+2}^\varepsilon - \nabla \tilde{w}_{n+2}^\varepsilon \right\|_{L^2(U_{n+2})} \\ & \leq C \left(\varepsilon^\alpha (\varepsilon 3^l)^{-Q} + 3^{-l\alpha} (\varepsilon 3^l)^{-1} + 3^{-k\alpha} + (\varepsilon 3^l)^\alpha 3^{kQ} \right) \sum_{j=1}^{n+1} \|\nabla g_j\|_{L^{2+\delta}(U_j)}^{\frac{n+2}{j}}. \end{aligned} \tag{3.73}$$

Therefore to obtain the lemma it suffices to prove (3.64). To prove this bound, we use the identity

$$\begin{aligned} & \nabla \tilde{v}_{n+2}^\varepsilon(x) - \nabla \bar{w}_{n+2}(x) \\ & = \nabla \left(\sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\varepsilon v \left(\frac{x}{\varepsilon}, \frac{z}{\varepsilon} + \square_{l+1}, \nabla \bar{w}_{n+2}(z), \bar{\Theta}(z) \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right), \end{aligned}$$

which, combined with (3.47), and (3.37), yields

$$\begin{aligned} & \left\| \nabla \tilde{v}_{n+2}^\varepsilon(x) - \nabla \bar{w}_{n+2}(x) \right\|_{H^{-1}(U_{n+2})} \\ & \leq \left\| \sum_{z \in \varepsilon 3^l \mathbb{Z}^d \cap U_{n+3}} \zeta_z(x) \left(\varepsilon v \left(\frac{x}{\varepsilon}, \frac{z}{\varepsilon} + \square_{l+1}, \nabla \bar{w}_{n+2}(z), \bar{\Theta}(z) \right) - \ell_{\nabla \bar{w}_{n+2}(z)}(x) \right) \right\|_{L^2(U_{n+2})} \\ & \leq C \varepsilon 3^l 3^{-\alpha(l-k)} 3^{Qk} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}} \leq C 3^{-\alpha(l-k)} 3^{Qk} \sum_{i=1}^{n+1} \|g_i\|_{L^2(U_i)}^{\frac{n+2}{i}}. \end{aligned}$$

This is (3.64).

In view of the selection of the mesoscopic parameters k and l in Section 3.4, which gives us (3.41), the proof of (3.30) is now complete. \square

4. Large-Scale $C^{0,1}$ -Type Estimates for Linearized Equations

In the next two sections, we suppose that $n \in \{0, \dots, N - 1\}$ is such that

$$\begin{cases} \text{Theorems 1.1 and 1.2 are valid for } n + 1 \text{ in place of } n, \\ \text{and Theorem 1.3 is valid for } n. \end{cases} \tag{4.1}$$

The goal is to prove that Theorem 1.3 is also valid for $n + 1$. Combined with the results of the previous two sections and an induction argument, this completes the proof of Theorems 1.1, 1.2 and 1.3.

The goal of this section is to prove the half of Theorem 1.3, namely estimate (1.22) for w_{n+2} . The estimate (1.23) will be the focus of Section 5.

Both of the estimate in the conditional proof of Theorem 1.3 are obtained by “harmonic” approximation: homogenization says that on large scales the heterogeneous equations behave like the homogenized equations, and therefore we may expect the former to inherit some of the better regularity estimates of the latter. The quantitative version of the homogenization statement provided by Theorem 1.2 allows us to prove a $C^{0,1}$ -type estimate, following a well-known excess decay argument originally introduced in [6].

4.1. Approximation of w_{n+2} by Smooth Functions

The large-scale regularity estimates are obtained by approximating the solutions of the linearized equations for the heterogeneous Lagrangian L , as well as the linearization errors, by the solutions of the linearized equations for the homogenized Lagrangian \bar{L} . Since the latter possess better smoothness properties, this allows us to implement an excess decay iteration for the former.

We begin by giving a quantitative statement concerning the smoothness of solutions to the linearized equations for the homogenized operator \bar{L} . This is essentially well-known, but we need a more precise statement than we could find in the literature.

We next present the statement concerning the approximating of the solutions w_m of the linearized equations for L , as well as the linearized errors ξ_m , by solutions of the linearized equations for \bar{L} . For the w_m 's, this is essentially a rephrasing of the assumed validity of Theorem 1.2 for $n + 1$ in place of N , see (4.1). For the ξ_m 's, this can be thought of as a homogenization result.

Lemma 4.1. (Smooth approximation of w_{n+2}). *Assume that (4.1) holds. Fix $M \in [1, \infty)$. There exist $\delta(n, d, \Lambda) \in (0, d)$, $\alpha(d, \Lambda) \in (0, \frac{1}{2}]$, $C(n, M, d, \Lambda) < \infty$ and a random variable \mathcal{X} satisfying*

$$\mathcal{X} = \mathcal{O}_\delta(C)$$

such that the following statement is valid. For every $R \geq \mathcal{X}$ and $u, v, w_1, \dots, w_{n+2} \in H^1(B_R)$ satisfying, for each $m \in \{1, \dots, n + 2\}$,

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 \quad \text{and} \quad -\nabla \cdot (D_p L(\nabla v, x)) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(x, \nabla u, \nabla w_1, \dots, \nabla w_{m-1})) & \text{in } B_R, \\ \|\nabla u\|_{\underline{L}^2(B_R)} \vee \|\nabla v\|_{\underline{L}^2(B_R)} \leq M, \end{cases} \tag{4.2}$$

if we let $\bar{u}, \bar{v}, \bar{w}_1, \dots, \bar{w}_{n+2} \in H^1(B_{R/2})$ be the solutions of the Dirichlet problems

$$\begin{cases} -\nabla \cdot (D_p \bar{L}(\nabla \bar{u})) = 0 \quad \text{and} \quad -\nabla \cdot (D_p \bar{L}(\nabla \bar{v})) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 \bar{L}(\nabla \bar{u}) \nabla \bar{w}_m) = \nabla \cdot (\bar{\mathbf{F}}_m(\nabla \bar{u}, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1})) & \text{in } B_{\frac{1}{2}(1+2^{-m})R}, \\ \bar{u} = u, \quad \bar{v} = v, \quad \bar{w}_m = w_m & \text{on } \partial B_{\frac{1}{2}(1+2^{-m})R}, \end{cases} \tag{4.3}$$

then we have, for each $m \in \{1, \dots, n + 2\}$, the estimate

$$\|w_m - \bar{w}_m\|_{\underline{L}^2(B_{\frac{1}{2}(1+2^{-m})R})} \leq C R^{-\alpha} \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \tag{4.4}$$

Proof. Denote, in short, $R_m := \frac{1}{2}(1 + 2^{-m})R$ and $r_m := \frac{1}{4}(R_m - R_{m-1})$. Since we assume (4.1), we have that Theorem 1.2 applied for $n + 1$ instead of N and Theorem 1.3, assumed now for n in place of N , are both valid. In particular, there is $\sigma(n, \text{data}) \in (0, 1)$ and $C(n, M, \text{data}) < \infty$, a random variable $\tilde{\mathcal{X}} \leq \mathcal{O}_\sigma(C)$ such that, for $R \geq \tilde{\mathcal{X}}$ and $m \in \{1, \dots, n + 2\}$, Theorem 1.3 gives

$$\sum_{i=1}^m \|\nabla w_i\|_{\underline{L}^{\frac{m}{i}(2+\delta)}(B_{R_i})}^{\frac{m}{i}} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \tag{4.5}$$

and Theorem 1.2 yields

$$\|w_m - \bar{w}_m\|_{\underline{L}^2(B_{R_m})} \leq \tilde{\mathcal{X}} R^{-\alpha} \sum_{i=1}^m \|\nabla w_j\|_{\underline{L}^{2+\delta}(B_{R_i})}^{\frac{m}{i}}. \tag{4.6}$$

We set

$$\mathcal{X} := \left(1 \vee \tilde{\mathcal{X}} \right)^{\frac{2}{\alpha}}.$$

Clearly $\mathcal{X} = \mathcal{O}_{\frac{1}{2}\alpha\sigma}(C)$ and $\mathcal{X} \geq \tilde{\mathcal{X}}$, and if $R \geq \mathcal{X}$, then $\tilde{\mathcal{X}}R^{-\alpha} \leq R^{-\frac{1}{2}\alpha}$. In conclusion, by taking α smaller if necessary, we obtain by (4.6) that, for $m \in \{1, \dots, n + 2\}$ and $R \geq \mathcal{X}$,

$$\|w_m - \bar{w}_m\|_{\underline{L}^2(B_{R_m})} \leq CR^{-\alpha} \sum_{i=1}^m \|\nabla w_i\|_{\underline{L}^{2+\delta}(B_{R_i})}^{\frac{m}{i}}. \tag{4.7}$$

Furthermore, we notice that (2.7) yields

$$|\mathbf{F}_m(x, \nabla u, \nabla w_1, \dots, \nabla w_{n+1})| \leq C \sum_{i=1}^{n+1} |\nabla w_i|^{\frac{m}{i}},$$

and thus we get, by (4.5) and (4.7), that

$$\|\mathbf{F}_m(x, \nabla u, \nabla w_1, \dots, \nabla w_{m-1})\|_{\underline{L}^{2+\delta}(B_{R_{m-1}})} \leq C \sum_{i=1}^{m-1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}.$$

It then follows by the Meyers estimate and equation of w_m that

$$\|\nabla w_m\|_{\underline{L}^{2+\delta}(B_{R_m})} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}},$$

finishing the proof of (4.4) by (4.7). \square

4.2. Excess Decay Iteration for w_{n+2}

In this subsection we conditionally prove the statement of Theorem 1.3 for $n + 1$ and for $q = 2$. The restriction on q will be removed in the next subsection. The proof is by a decay of excess iteration, following along similar lines as the argument from [6], using ‘‘harmonic’’ approximation. The statement we prove is summarized in the following proposition:

Proposition 4.2. *Assume that (4.1) holds. Fix $M \in [1, \infty)$. Then there exist constants $\sigma(\text{data}), \alpha(d, \Lambda) \in (0, \frac{1}{2}]$, $C(M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying*

$$\mathcal{X} \leq \mathcal{O}_\sigma(C)$$

such that the following statement is valid: let $R \in [\mathcal{X}, \infty)$ and u, w_1, \dots, w_n satisfy, for each $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1})) & \text{in } B_R, \end{cases} \tag{4.8}$$

where u satisfy the normalization

$$\frac{1}{R} \|u - (u)_{B_R}\|_{\underline{L}^2(B_R)} \leq M.$$

Then, for $m \in \{1, \dots, n + 2\}$ and for every $r \in [\mathcal{X}, \frac{1}{2}R]$, we have

$$\begin{aligned} & \left(\frac{r}{R} + \frac{1}{r}\right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|w_m - \ell\|_{\underline{L}^2(B_r)} + \|\nabla w_m\|_{\underline{L}^2(B_r)} \\ & \leq C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}\right)^{\frac{m}{i}}. \end{aligned} \tag{4.9}$$

Remark 4.3. Proposition 4.2 holds for $n = 0$ by [1, Proposition 4.3] without assuming (4.1).

Proof. The proof is based on combination of harmonic approximation and decay estimates for homogenized solutions presented in Appendix D. The necessary estimate is (D.7). We take the minimal scale \mathcal{X} as the maximum of the minimal scale in Lemma 4.1 and in Theorem 1.3, which is valid with n in place of \mathbb{N} , corresponding $q = 2$, and a constant $\mathbf{R}(\mathbf{M}, \text{data}) \in [1, \infty)$ to be fixed in the course of the proof. This choice, in particular, implies that there exist constants $C(\mathbf{M}, \text{data}) < \infty$ and $\sigma(\text{data}) \in (0, \frac{1}{2}]$ such that $\mathcal{X} \leq \mathcal{O}_\sigma(C)$ and $\mathcal{X} \geq \mathbf{R}$.

We will prove the statement using induction in shrinking radii. Indeed, we set, for $j \in \mathbb{N}$ and $\theta \in (0, \frac{1}{2}]$, $r_j := \theta^j r_0$, where $r_0 \in [\mathcal{X}, \delta R]$ and $\delta \in (0, \frac{1}{2}]$. Parameters θ, δ and \mathbf{R} will all be fixed in the course of the proof. Having fixed θ, δ and \mathbf{R} , we assume that there is $J \in \mathbb{N}$ such that $r_{J+1} < \mathcal{X} \leq r_J$ for some $J \in \mathbb{N}$. If there is no such J or $\mathcal{X} \geq \delta R$, the result will follow simply by giving up a volume factor. Furthermore, we device the notation of this proof in such a way that it will also allow us to prove the result of the next lemma, Lemma 4.4.

We denote, in short, for $m \in \{1, \dots, n + 2\}$ and $\gamma \in [0, 1)$ to be fixed,

$$\mathbf{D}_m := \frac{1}{r_0} \|w_m - (w_m)_{B_{r_0}}\|_{\underline{L}^2(B_{r_0})} + \sum_{i=1}^{m-1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}\right)^{\frac{m}{i}} \tag{4.10}$$

and

$$\mathbf{E}_m := \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_0} \|w_m - \ell\|_{\underline{L}^2(B_{r_0})} + \left(\frac{r_0}{R} + \frac{1}{r_0}\right)^\gamma \mathbf{D}_m. \tag{4.11}$$

Theorem 1.3 implies that, for $r \in [\mathcal{X}, \frac{1}{2}R]$ and $m \in \{1, \dots, n + 1\}$, we have that

$$\|\nabla w_m\|_{\underline{L}^2(B_r)} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}\right)^{\frac{m}{i}}. \tag{4.12}$$

Notice then that, by (4.12) and Poincaré’s inequality, we have, for $r \in [\mathcal{X}, \frac{1}{2}R]$ and $m \in \{1, \dots, n + 2\}$, that

$$\sum_{i=1}^{m-1} \left(\frac{1}{r} \|w_i - (w_i)_{B_r}\|_{\underline{L}^2(B_r)}\right)^{\frac{m}{i}} \leq C \mathbf{D}_m. \tag{4.13}$$

Step 1. Letting \bar{u}_j solve

$$\begin{cases} -\nabla \cdot (D_p \bar{L}(\nabla \bar{u}_j)) = 0 & \text{in } B_{2r_j}, \\ \bar{u}_j = u & \text{on } \partial B_{2r_j}, \end{cases}$$

we show that there exist, for $\eta \in (0, 1]$, constants $\alpha_1(d, \Lambda) \in (0, \frac{1}{2})$, $\delta(\eta, \mathbf{M}, \text{data}) < \infty$ and $\mathbf{R}(\eta, \mathbf{M}, \text{data}) < \infty$ such that, for $j \in \{0, \dots, J\}$,

$$\frac{1}{2r_j} \inf_{\ell \in \mathcal{P}_1} \|u - \ell\|_{\underline{L}^2(B_{2r_j})} + \frac{1}{2r_j} \inf_{\ell \in \mathcal{P}_1} \|\bar{u}_j - \ell\|_{\underline{L}^2(B_{2r_j})} \leq \eta \left(\frac{r_j}{\mathbf{R}} + \frac{1}{r_j} \right)^{\alpha_1}. \tag{4.14}$$

The parameter η shall be fixed later in (4.19). On the one hand, we have from [1, Theorem 2.3] that there exist constants $\beta_1(d, \Lambda) \in (0, 1)$ and $C(\mathbf{M}, d, \Lambda) < \infty$ such that

$$\frac{1}{2r_0} \inf_{\ell \in \mathcal{P}_1} \|u - \ell\|_{\underline{L}^2(B_{2r_0})} \leq C \left(\frac{r_0}{\mathbf{R}} + \frac{1}{r_0} \right)^{\beta_1}.$$

On the other hand, by harmonic approximation ([1, Corollary 2.2]) and Lipschitz estimate for u ([1, Theorem 2.3]) we get that there exist constants $\beta_2(d, \Lambda) \in (0, 1)$ and $C(\mathbf{M}, d, \Lambda) < \infty$ such that

$$\|u - \bar{u}_j\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{-\beta_2}.$$

Thus (4.14) follows by the triangle inequality by taking $\alpha_1 := \frac{1}{2}(\beta_1 \wedge \beta_2)$, and choosing δ small enough and \mathbf{R} large enough so that

$$C \left(\delta + \frac{1}{\mathbf{R}} \right)^{\frac{1}{2}\beta_1} + C\mathbf{R}^{-\frac{1}{2}\beta_2} \leq \eta. \tag{4.15}$$

We assume, from now on, that δ and \mathbf{R} are such that (4.15) is valid.

Step 2. Letting $j \in \{0, \dots, J\}$ and $m \in \{1, \dots, n + 2\}$, and $\bar{w}_{1,j}, \dots, \bar{w}_{m,j}$ to solve, with \bar{u}_j as in Step 2, equations

$$\begin{cases} -\nabla \cdot (D_p^2 \bar{L}(\nabla \bar{u}_j) \nabla \bar{w}_{m,j}) = \nabla \cdot \bar{\mathbf{F}}_m(\nabla \bar{u}_j, \bar{w}_{1,j}, \dots, \bar{w}_{m-1,j}) & \text{in } B_{\frac{1}{2}(1+2^{-m})r_j}, \\ \bar{w}_{m,j} = w_m & \text{on } \partial B_{\frac{1}{2}(1+2^{-m})r_j}, \end{cases}$$

we show that then there is $\alpha_2(\text{data}) \in (0, \frac{1}{2}]$ such that

$$\begin{aligned} \frac{1}{r_j} \|w_m - \bar{w}_{m,j}\|_{\underline{L}^2(B_{r_j/2})} &\leq Cr_j^{-\alpha_2} \frac{1}{r_j} \|w_m - (w_m)_{B_{r_j}}\|_{\underline{L}^2(B_{r_j})} \\ &\quad + Cr_j^{-\alpha_2} \sum_{i=1}^{m-1} \left(\frac{1}{\mathbf{R}} \|w_i - (w_i)_{B_{\mathbf{R}}}\|_{\underline{L}^2(B_{\mathbf{R}})} \right)^{\frac{m}{i}}. \end{aligned} \tag{4.16}$$

This, however, is a direct consequence of (4.12) and Lemma 4.1.

Step 3. Induction assumption. Set $\alpha := \frac{\beta}{2}(\alpha_1 \wedge \alpha_2)$, where α_1 and α_2 are as in Steps 1 and 2, respectively, and β comes from the $C^{n+2,\beta}$ regularity of \bar{L} . Let δ_j be defined as

$$\delta_j := \left(\frac{r_j}{r_0} + \frac{1}{r_j^{d-\sigma}} \right)^\alpha. \tag{4.17}$$

We assume inductively that, for $j^* \in \{1, \dots, J\}$, $j \in \{0, \dots, j^*\}$, and $m \in \{1, \dots, n + 2\}$, we have, for a constant $\mathbf{C} \in [1, \infty)$ to be fixed in Step 5, that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|w_m - \ell\|_{\underline{L}^2(B_{r_j})} \leq \delta_j \mathbf{E}_m \quad \text{and} \quad \frac{1}{r_j} \|w_m - (w_m)_{B_{r_j}}\|_{\underline{L}^2(B_{r_j})} \leq \mathbf{C} \mathbf{D}_m. \tag{4.18}$$

Here \mathbf{D}_m and \mathbf{E}_m are defined in (4.10) and (4.11), respectively. Obviously (4.18) is valid for $j = 0$ by the definitions of \mathbf{D}_m and \mathbf{E}_m . Fixing

$$\eta := (1 \vee \mathbf{C})^{1/\beta}, \tag{4.19}$$

we have that (4.14) implies

$$\mathbf{C} \left(\frac{1}{r_j} \inf_{\ell \in \mathcal{P}_1} \|\bar{u}_j - \ell\|_{\underline{L}^2(B_{r_j})} \right)^\beta \leq \delta_j \left(\frac{r_0}{R} + \frac{1}{r_0} \right)^\alpha. \tag{4.20}$$

Using (4.16) and (4.18) as well, we obtain, for $m \in \{1, \dots, n + 1\}$,

$$\frac{1}{r_j} \|w_m - \bar{w}_{m,j}\|_{L^2(B_{r_j/2})} \leq \mathbf{C} \mathbf{C} r_j^{-\alpha_2} \mathbf{D}_m \leq \frac{1}{2} \theta^{\frac{d}{2}+1} \delta_{j+1} \mathbf{E}_m \tag{4.21}$$

provided that

$$\mathbf{C} \mathbf{C} \theta^{-\frac{d}{2}-2} \mathbf{R}^{-\frac{1}{2}\alpha_2} \leq \frac{1}{2}. \tag{4.22}$$

We assume, from now on, that \mathbf{R} is such that both (4.15) and (4.22) are valid.

Step 4. We show that the first inequality in (4.18) continues to hold for $j = j^* + 1$. First, by the triangle inequality, (4.21) and (4.18), we have that

$$\begin{aligned} & \frac{1}{\delta_j r_j} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_{m,j} - \ell\|_{L^2(B_{r_j/2})} \\ & \leq \frac{1}{\delta_j r_j} \inf_{\ell \in \mathcal{P}_1} \|w_m - \ell\|_{L^2(B_{r_j/2})} + \frac{1}{\delta_j r_j} \|w_m - \bar{w}_{m,j}\|_{L^2(B_{r_j/2})} \\ & \leq 2\mathbf{E}_m \end{aligned}$$

and

$$\frac{1}{\delta_{j+1} r_{j+1}} \inf_{\ell \in \mathcal{P}_1} \|w_m - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \leq \frac{1}{\delta_{j+1} r_{j+1}} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_{m,j} - \ell\|_{\underline{L}^2(B_{r_{j+1}})} + \frac{1}{2} \mathbf{E}_m.$$

By a similar computation, using also the induction assumption (4.18),

$$\sum_{i=1}^m \left(\frac{1}{r_j} \|\bar{w}_i - (\bar{w}_i)_{B_{r_j/2}}\|_{\underline{L}^2(B_{r_j/2})} \right)^{\frac{m}{i}} \leq \mathbf{C} \mathbf{C} \mathbf{D}_m.$$

Now, applying (D.7) we obtain by the previous three displays, (4.20) and (4.13), for each $m \in \{1, \dots, n + 2\}$, that

$$\begin{aligned}
 & \left(\frac{r_j}{r_{j+1}}\right)^\beta \frac{1}{r_{j+1}} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_{m,j} - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \\
 & \leq C\delta_j \sum_{i=1}^m \left(\frac{1}{\delta_j r_j} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_{i,j} - \ell\|_{\underline{L}^2(B_{r_j/2})}\right)^{\frac{m}{i}} \\
 & \quad + C\delta_j \sum_{i=1}^{m-1} \left(\frac{1}{r_j} \|\bar{w}_i - (\bar{w}_i)_{B_{r_j/2}}\|_{\underline{L}^2(B_{r_j/2})}\right)^{\frac{m}{i}} \\
 & \quad + C \left(\frac{1}{r_j} \inf_{\ell \in \mathcal{P}_1} \|\bar{u}_j - \ell\|_{\underline{L}^2(B_{r_j/2})}\right)^\beta \sum_{i=1}^m \left(\frac{1}{r_j} \|\bar{w}_i - (\bar{w}_i)_{B_{r_j/2}}\|_{\underline{L}^2(B_{r_j/2})}\right)^{\frac{m}{i}} \\
 & \leq C\delta_j \mathbf{E}_m. \tag{4.23}
 \end{aligned}$$

and, consequently,

$$\frac{1}{\delta_{j+1} r_{j+1}} \inf_{\ell \in \mathcal{P}_1} \|w_m - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \leq \frac{1}{2} \mathbf{E}_m + C \left(\frac{\delta_j}{\delta_{j+1}}\right) \left(\frac{r_{j+1}}{r_j}\right)^\beta \mathbf{E}_m \leq \left(\frac{1}{2} + C\theta^{\frac{\beta}{2}}\right) \mathbf{E}_m.$$

Thus, choosing θ small enough so that $C\theta^{\frac{\beta}{2}} \leq \frac{1}{2}$, we obtain that the first inequality in (4.18) is valid for $j = j^* + 1$.

Step 5. The last step in the proof is to show the second inequality in (4.18) for $j = j^* + 1$. Let ℓ_j be the minimizing affine function in $\inf_{\ell \in \mathcal{P}_1} \|w_m - \ell\|_{\underline{L}^2(B_{r_j})}$. Then, by the triangle inequality and the first inequality in (4.18) valid for $j \in \{0, \dots, j^*\}$,

$$|\nabla \ell_{j+1} - \nabla \ell_j| \leq C(\delta_{j+1} + \delta_j) \mathbf{E}_m$$

Thus, summation gives that

$$|\nabla \ell_{j^*+1} - \nabla \ell_0| \leq C \mathbf{E}_m \sum_{j=0}^{j^*+1} \delta_j.$$

Therefore,

$$\begin{aligned}
 \frac{1}{r_{j^*+1}} \|w_m - (w_m)_{B_{r_{j^*+1}}}\|_{\underline{L}^2(B_{r_{j^*+1}})} & \leq \frac{1}{r_{j^*+1}} \|w_m - \ell_{j^*+1}\|_{\underline{L}^2(B_{r_{j^*+1}})} + |\nabla \ell_{j^*+1}| \\
 & \leq C \mathbf{E}_m \sum_{j=0}^{j^*+1} \delta_j + |\nabla \ell_0|.
 \end{aligned}$$

By the triangle inequality we have that

$$|\nabla \ell_0| \leq \frac{8}{r_0} \|w_m - \ell_0\|_{\underline{L}^2(B_{r_0})} + \frac{8}{r_0} \|w_m - (w_m)_{B_{r_0}}\|_{\underline{L}^2(B_{r_0})} \leq 8\mathbf{E}_m + 8\mathbf{D}_m$$

and hence

$$\frac{1}{r^{j^*+1}} \left\| w_m - (w_m)_{B_{r^{j^*+1}}} \right\|_{\underline{L}^2(B_{r^{j^*+1}})} \leq C D_m \sum_{j=0}^{j^*+1} \delta_j.$$

Choosing $\mathbf{C} = C$, where C is as in the above inequality, verifies the second inequality in (4.18) for $j = j^* + 1$. This finishes the proof of the induction step, and thus completes the proof. \square

To show Lipschitz estimates for the linearization errors in the next section, we need a small variant of the previous proposition.

Lemma 4.4. *Assume that (4.1) holds. Fix $M \in [1, \infty)$. Then there exist constants $\alpha(\text{data})$, $\sigma(n, M, \text{data})$, $\theta(n, M, \text{data}) \in (0, \frac{1}{2}]$ and $C(\sigma, M, \text{data}) < \infty$, and a random variable \mathcal{X} satisfying*

$$\mathcal{X} \leq \mathcal{O}_\sigma(C)$$

such that the following statement is valid. Let $R \in [\mathcal{X}, \infty)$ and u, w_1, \dots, w_n satisfy, for each $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1})) & \text{in } B_R, \end{cases} \tag{4.24}$$

and, for $r \in [\mathcal{X}, \frac{1}{2}R]$,

$$-\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_{n+2}) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{n+1})) \text{ in } B_r,$$

where u satisfy the normalization

$$\frac{1}{R} \|u - (u)_{B_R}\|_{\underline{L}^2(B_R)} \leq M.$$

Then

$$\begin{aligned} & \inf_{\ell \in \mathcal{P}_1} \frac{1}{\theta r} \|w_{n+2} - \ell\|_{\underline{L}^2(B_{\theta r})} \\ & \leq \frac{1}{2} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|w_{n+2} - \ell\|_{\underline{L}^2(B_r)} + C \left(\frac{r}{R} + \frac{1}{r}\right)^\alpha \frac{1}{r} \|w_{n+2} - (w_{n+2})_{B_r}\|_{\underline{L}^2(B_r)} \\ & \quad + C \left(\frac{r}{R} + \frac{1}{r}\right)^\alpha \sum_{i=1}^{n+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}\right)^{\frac{n+2}{i}}. \end{aligned} \tag{4.25}$$

Proof. The proof is a rearrangement of the argument in the proof of Proposition 4.2. Indeed, we take $r_0 = r$ and combine the first inequality of (4.23) with (4.16) and (4.9), which is valid for $m \in \{1, \dots, n + 1\}$. \square

4.3. Improvement of Spatial Integrability

We next complete the conditional proof of (1.22) by improving the spatial integrability of (4.25) from L^2 to L^q for every $q \in [2, \infty)$. To do this, we use the estimate (4.25) to pass from the large scale R to the microscopic, random scale \mathcal{X} . We then use deterministic estimates from classical elliptic regularity theory to obtain local L^q estimates in balls of radius one, picking up a volume factor—which is power of \mathcal{X} —as a price to pay. The first formalize the latter step in the following lemma:

Lemma 4.5. *Assume (4.1) and the hypotheses of Theorem 1.3. Let $\beta \in (0, 1)$ and $q \in (2, \infty)$. Then there exist $\sigma(q, \text{data}) \in (0, d)$, $C(q, \mathbf{M}, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying $\mathcal{X} = \mathcal{O}_\sigma(C)$ such that, for every $r \geq \mathcal{X}$ and $m \in \{1, \dots, n + 2\}$,*

$$\|\nabla w_m\|_{\underline{L}^q(B_{r/2})} \leq C \left(1 + r^{\frac{d^2(q-2)}{4q\beta} + \frac{d(q-2)}{2q}}\right) \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}\right)^{\frac{m}{i}}. \tag{4.26}$$

Proof. In view of the assumption of (4.1) and thus the validity of Theorem 1.3 for n , we only need to prove (4.26) for $m = n + 2$. Fix $q \in (2, \infty)$, $r \in [2, \infty)$, $\beta \in (0, 1)$ to be selected below. The $C^{1,\beta}$ -estimate in Proposition A.2, together with a covering argument, yields

$$\|\nabla u\|_{L^\infty(B_r)} + [\nabla u]_{C^{0,\beta}(B_r)} \leq Cr^{\frac{d}{2}} \|\nabla u\|_{\underline{L}^2(B_{2r})}.$$

Setting $\mathbf{a} := D_p^2 L(\nabla u, x)$, we deduce by the assumption of (L1) that

$$[\mathbf{a}]_{C^{0,\beta}(B_r)} \leq C (1 + [\nabla u]_{C^{0,\beta}(B_r)}) \leq C \left(1 + r^{\frac{d}{2}} \|\nabla u\|_{\underline{L}^2(B_{2r})}\right).$$

Applying Proposition A.1 we find that, for each $x \in B_{r/2}$ and $q \in (2, \infty)$,

$$\begin{aligned} &\|\nabla w_{n+2}\|_{L^q(B_1(x))} \\ &\leq C [\mathbf{a}]_{C^{0,\beta}(B_r)}^{\frac{d(q-2)}{2q\beta}} \|\nabla w_{n+2}\|_{L^2(B_2(x))} + C \|\mathbf{f}_m(\nabla u, \nabla w_1, \dots, \nabla w_{n+1})\|_{L^q(B_2(x))} \\ &\leq C \left(1 + r^{\frac{d^2(q-2)}{4q\beta}} \|\nabla u\|_{\underline{L}^2(B_{2r})}\right) \|\nabla w_{n+2}\|_{L^2(B_2(x))} \\ &\quad + C \|\mathbf{f}_{n+2}(\nabla u, \nabla w_1, \dots, \nabla w_{n+1})\|_{L^q(B_2(x))}. \end{aligned}$$

By a covering argument, we therefore obtain

$$\begin{aligned} \|\nabla w_{n+2}\|_{L^q(B_{r/2})} &\leq C \left(1 + r^{\frac{d^2(q-2)}{4q\beta}} \|\nabla u\|_{\underline{L}^2(B_{2r})}\right) \|\nabla w_{n+2}\|_{L^2(B_r)} \\ &\quad + C \|\mathbf{f}_{n+2}(\nabla u, \nabla w_1, \dots, \nabla w_{n+1})\|_{L^q(B_r)}. \end{aligned} \tag{4.27}$$

If we now take \mathcal{X} to be the maximum of the minimal scales in the statements of:

- (1) [4, Theorem 11.13];

- (2) Theorem 1.3 for n in place of \mathbf{N} and with a sufficiently large exponent of spatial integrability q' in place of q (which can be computed explicitly in terms of our q using the Hölder inequality, although we omit this computation)—the validity of which is given by assumption (4.1);
- (3) Proposition 4.2;

then we have that $\mathcal{X} = \mathcal{O}_\sigma(C)$ as stated in the lemma and that $r \geq \mathcal{X}$ implies the following estimates:

$$\|\nabla u\|_{\underline{L}^2(B_{2r})} \leq C,$$

$$\|\mathbf{f}_{n+2}(\nabla u, \nabla w_1, \dots, \nabla w_{n+1})\|_{\underline{L}^q(B_r)} \leq C \sum_{i=1}^{n+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}$$

and

$$\|\nabla w_{n+2}\|_{\underline{L}^2(B_r)} \leq C \sum_{i=1}^{n+2} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}.$$

Combining these with (4.27), we obtain

$$\|\nabla w_{n+2}\|_{\underline{L}^q(B_{r/2})} \leq C \left(1 + r^{\frac{d^2(q-2)}{4q\beta} + \frac{d(q-2)}{2q}} \right) \sum_{i=1}^{n+2} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}.$$

This completes the proof of the lemma. \square

In the next lemma we finally achieve the goal of this section, which is to show that (4.1) implies (1.22) for $m = n + 1$.

Lemma 4.6. *Assume that (4.1) holds. Fix $M \in [1, \infty)$ and $q \in (2, \infty)$. Then there exist constants $\sigma(q, \text{data}) \in (0, d)$, $C(q, M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying*

$$\mathcal{X} \leq \mathcal{O}_\sigma(C)$$

such that the following is valid. Suppose that $R \in [\mathcal{X}, \infty)$ and $u, w_1, \dots, w_n \in H^1(B_R)$ such that, for every $m \in \{1, \dots, n + 2\}$,

$$\begin{cases} \frac{1}{R} \|u - (u)_{B_R}\|_{\underline{L}^2(B_R)} \leq M, \\ -\nabla \cdot (D_p L(\nabla u, x)) = 0 \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1})) \end{cases} \quad \begin{array}{l} \text{in } B_R, \\ \\ \text{in } B_R, \end{array} \tag{4.28}$$

Then, for every $r \in [\mathcal{X}, \frac{1}{2}R]$, we have

$$\|\nabla w_{n+2}\|_{\underline{L}^q(B_r)} \leq C \sum_{i=1}^{n+2} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}. \tag{4.29}$$

Proof. Fix $q \in (2, \infty)$. Select a parameter $\theta \in (0, 1)$ which will denote a mesoscopic scale. For each $z \in \mathbb{R}^d$, we take \mathcal{X}_z to be the random variable \mathcal{X} in the statement of Proposition 4.2, centered at the point z . Define another random variable (a minimal scale) by

$$\mathcal{Y} := \sup \left\{ 3^k : k \in \mathbb{N}, \sup_{z \in \mathbb{Z}^d \cap B_{3^k}} \mathcal{X}_z \geq 3^{k\theta} \right\}.$$

It is clear from Proposition 4.2 and an routine union bound argument that

$$\mathcal{Y} \leq \mathcal{O}_\sigma(C).$$

Next, for every $k \in \mathbb{N}$ and $z \in \mathbb{R}^d$ we let $\mathcal{Z}_{k,z}$ denote the random variable

$$\mathcal{Z}_{k,z} := \sup_{(u, w_1, \dots, w_{n+2})} \sup_{m \in \{1, \dots, n+2\}} \frac{\|\nabla w_m\|_{\underline{L}^q(z + \square_k)}}{\sum_{i=1}^m \left(3^{-k} \|w_i - (w_i)_{z + \square_{k+1}}\|_{\underline{L}^2(z + \square_{k+1})} \right)^{\frac{m}{i}}},$$

where the supremum is taken over all $(u, w_1, \dots, w_{n+2}) \in (H^1(z + \square_{k+1}))^{n+3}$ satisfying, for every $m \in \{1, \dots, n + 2\}$,

$$\begin{cases} \|\nabla u\|_{\underline{L}^2(z + \square_{k+1})} \leq M, \\ -\nabla \cdot D_p L(\nabla u, x) = 0 & \text{in } z + \square_{k+1}, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x)) & \text{in } z + \square_{k+1}. \end{cases} \tag{4.30}$$

Observe that $\mathcal{Z}_{k,z}$ is $\mathcal{F}(z + \square_{k+1})$ -measurable and, by Lemma 4.5 and an easy covering argument, it satisfies the estimate

$$\mathcal{Z}_{k,z} \leq \mathcal{O}_\sigma(C). \tag{4.31}$$

Fix $\mathbf{A} \in [1, \infty)$ and define another random variable (a minimal scale) \mathcal{Z} by

$$\mathcal{Z} := \sup \left\{ 3^k : \left(\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1} \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}} \mathcal{Z}_{\lceil \theta k \rceil, z}^q \right)^{\frac{1}{q}} \geq \mathbf{A} \right\}.$$

We will show below that, if \mathbf{A} is chosen sufficiently large (depending of course on the appropriate parameters) then

$$\mathcal{Z} \leq \mathcal{O}_\sigma(C). \tag{4.32}$$

Assuming that (4.32) holds for the moment, let us finish the proof of the lemma. Suppose now that $k \in \mathbb{N}$ satisfies $\mathcal{Y} \vee \mathcal{Z} \leq 3^k \leq 3^{k+1} \leq R$. Let $(u, w_1, \dots, w_{n+2}) \in (H^1(B_R))^{n+3}$ satisfy (4.28). Then we have that

$$\begin{aligned}
 & \|\nabla w_{n+2}\|_{\underline{L}^q(\square_k)} \\
 &= \left(\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k} \|\nabla w_{n+2}\|_{\underline{L}^q(z + \square_{\lceil \theta k \rceil})}^q \right)^{\frac{1}{q}} \\
 &\leq \left(\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k} \sum_{i=1}^{n+2} \left(3^{-\theta k} \|w_i - (w_i)_{z + \square_{\lceil \theta k \rceil + 1}}\|_{\underline{L}^2(z + \square_{\lceil \theta k \rceil + 1})} \right)^{\frac{(n+2)q}{i}} \mathcal{Z}_{\lceil \theta k \rceil, z}^q \right)^{\frac{1}{q}} \\
 &\leq C \sum_{i=1}^{n+2} \left(3^{-k} \|w_i - (w_i)_{z + \square_{k+1}}\|_{\underline{L}^2(\square_{k+1})} \right)^{\frac{n+2}{i}} \left(\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_k} \mathcal{Z}_{\lceil \theta k \rceil, z}^q \right)^{\frac{1}{q}} \\
 &\leq CA \sum_{i=1}^{n+2} \left(3^{-k} \|w_i - (w_i)_{z + \square_{k+1}}\|_{\underline{L}^2(\square_{k+1})} \right)^{\frac{n+2}{i}} \\
 &\leq CA \sum_{i=1}^{n+2} \left(\frac{1}{R} \|w_i - (w_i)_{z + \square_{k+1}}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}.
 \end{aligned}$$

Note that in the third and final lines we used that $3^k \geq \mathcal{Y}$, that is, we used the result of Proposition 4.2. This is the desired estimate for $\mathcal{X} = \mathcal{Y} \vee \mathcal{Z}$, and so the proof of the lemma is complete subject to the verification of (4.32).

Turning to the proof of (4.32), we notice first that it suffices to show, for A sufficiently large, the existence of $\sigma(q, \text{data}) > 0$ and $C(q, M, \text{data}) < \infty$ such that, for every $k \in \mathbb{N}$,

$$\mathbb{P} \left[\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1} \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}} \mathcal{Z}_{\lceil \theta k \rceil, z}^q \geq A^q \right] \leq C \exp(-c3^{k\sigma}). \tag{4.33}$$

Indeed, we can see that (4.33) implies (4.32) using a simple union bound. Fix then a parameter $\lambda \in [1, \infty)$ and compute, using (4.31) and the Chebyshev inequality,

$$\mathbb{P} \left[\sup_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}} \mathcal{Z}_{\lceil \theta k \rceil, z} > \lambda \right] \leq \exp(-c\lambda^\sigma). \tag{4.34}$$

Using the simple large deviations bound for sums of bounded, independent random variables, we have

$$\begin{aligned}
 & \mathbb{P} \left[\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1} \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}} \left(\mathcal{Z}_{\lceil \theta k \rceil, z}^q \wedge \lambda \right) \geq \mathbb{E} \left[\mathcal{Z}_{\lceil \theta k \rceil, z}^q \wedge \lambda \right] + 1 \right] \\
 & \leq 3^d \exp \left(-c\lambda^{-2} \left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1} \right| \right) \leq 3^d \exp \left(-c3^{d(1-\theta)k} \lambda^{-2} \right). \tag{4.35}
 \end{aligned}$$

Here we are careful to notice that, while the collection $\{\mathcal{Z}_{\lceil \theta k \rceil, z}^q \wedge \lambda : z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}\}$ of random variables is not independent (since adjacent cubes are touching and thus not separated by a unit distance), we can partition this collection into 3^d many subcollections which have an equal number of elements and each of which is independent. The large deviations estimate can then be applied to each

subcollection, and then a union bound yields (4.35). Combining (4.34), (4.35) and the observation that $\mathbb{E} \left[\mathcal{Z}_{[\theta k],z}^q \wedge \lambda \right] \leq \mathbb{E} \left[\mathcal{Z}_{[\theta k],z}^q \right] \leq C$ by (4.31), we obtain

$$\begin{aligned} \mathbb{P} \left[\left| 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1} \right|^{-1} \sum_{z \in 3^{\lceil \theta k \rceil} \mathbb{Z}^d \cap \square_{k+1}} \mathcal{Z}_{[\theta k],z}^q \geq C + 1 \right] \\ \leq C \exp \left(-c \left(\lambda^\sigma \wedge 3^{d(1-\theta)k} \lambda^{-2} \right) \right). \end{aligned}$$

Taking $\lambda := 3^{\frac{d}{4}(1-\theta)k}$ and $A^q := C + 1$ yields (4.33). \square

The above proof, in view of Remark 4.3, gives the following result without assuming (4.1). This, together with (5.1) below, serves as the base case for the induction.

Proposition 4.7. *Let $q \in [2, \infty)$, $M \in [1, \infty)$. Then there exist $\sigma(q, \text{data}) > 0$ and $C(q, M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying $\mathcal{X} \leq \mathcal{O}_\sigma(C)$ such that the following statement is valid. For $R \in [2\mathcal{X}, \infty)$ and $u, w_1 \in H^1(B_R)$ satisfying $\|\nabla u\|_{\underline{L}^2(B_R)} \leq M$ and*

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_1) = 0 & \text{in } B_R, \end{cases}$$

we have, for all $r \in [\mathcal{X}, \frac{1}{2}R]$,

$$\|\nabla w_1\|_{\underline{L}^q(B_r)} \leq \frac{C}{R} \|w_1 - (w_1)_{B_r}\|_{\underline{L}^2(B_r)}. \tag{4.36}$$

5. Large-Scale $C^{0,1}$ -Type Estimate for Linearization Errors

In this section we continue to suppose that $n \in \{0, \dots, N-1\}$ is such that (4.1) holds. The goal is to complete the proof that Theorem 1.3 is also valid for $n + 1$. Combined with the results of the previous three sections and an induction argument, this completes the proof of Theorems 1.1, 1.2 and 1.3.

5.1. Excess Decay Iteration for ξ_{n+1}

We start by proving higher integrability for a difference of two solutions. This, together with Proposition 4.7, yields the base case for the induction.

Proposition 5.1. *Fix $M \in [1, \infty)$ and $q \in [2, \infty)$. There exist $\alpha(\text{data})$, $\sigma(q, \text{data}) \in (0, \frac{1}{2}]$, $C(q, M, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying*

$$\mathcal{X} = \mathcal{O}_\sigma(C)$$

such that the following statement is valid: for every $R \geq \mathcal{X}$ and $u, v \in H^1(B_R)$ satisfying, for each $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 \quad \text{and} \quad -\nabla \cdot (D_p L(\nabla v, x)) = 0 & \text{in } B_R, \\ \|\nabla u\|_{\underline{L}^2(B_R)} \vee \|\nabla v\|_{\underline{L}^2(B_R)} \leq \mathbf{M}, \end{cases} \tag{5.1}$$

Then, for $m \in \{1, \dots, n + 1\}$, $\xi_0 = u - v$ and $r \in [\mathcal{X}, \frac{1}{2}R]$, we have

$$\|\nabla \xi_0\|_{\underline{L}^q(B_r)} + \left(\frac{r}{R} + \frac{1}{r}\right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\xi_0 - \ell\|_{\underline{L}^2(B_r)} \leq C \frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)}. \tag{5.2}$$

Proof. On the one hand, the estimate

$$\|\nabla \xi_0\|_{\underline{L}^2(B_r)} + \left(\frac{r}{R} + \frac{1}{r}\right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\xi_0 - \ell\|_{\underline{L}^2(B_r)} \leq C \frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} \tag{5.3}$$

follows by [1, Proposition 4.2]. On the other hand, the proof of

$$\|\nabla \xi_0\|_{\underline{L}^q(B_r)} \leq C \frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} \tag{5.4}$$

is similar to the proof of L^q -integrability of w_1 presented in Section 4.3. Indeed, noticing that ξ_0 satisfies the equation

$$-\nabla \cdot \tilde{\mathbf{a}} \nabla \xi_0 = 0, \quad \tilde{\mathbf{a}}(x) := \int_0^1 D_p^2 F(t \nabla u(x) + (1-t) \nabla v(x), x) dt,$$

we have by the normalization in (5.1) and $C^{1,\alpha}$ regularity of u and v that we may replace w_1 with ξ_0 in Lemma 4.5 applied with $m = 1$. Using this together with the Lipschitz estimate (5.3) for ξ_0 as in the proof of Lemma 4.6, concludes the proof of (5.4). We omit further details. \square

Proposition 5.2. Assume that (4.1) holds. Fix $\mathbf{M} \in [1, \infty)$. There exist constants $\alpha(n, \text{data}), \sigma(n, \text{data}) \in (0, \frac{1}{2}]$, $C(n, \mathbf{M}, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying

$$\mathcal{X} = \mathcal{O}_\sigma(C)$$

such that the following statement is valid. For every $R \geq \mathcal{X}$ and $u, v, w_1, \dots, w_{n+1} \in H^1(B_R)$ solving, for each $m \in \{1, \dots, n + 1\}$,

$$\begin{cases} -\nabla \cdot (D_p L(\nabla u, x)) = 0 \quad \text{and} \quad -\nabla \cdot (D_p L(\nabla v, x)) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot (\mathbf{F}_m(x, \nabla u, \nabla w_1, \dots, \nabla w_{m-1})) & \text{in } B_R, \\ \|\nabla u\|_{\underline{L}^2(B_R)} \vee \|\nabla v\|_{\underline{L}^2(B_R)} \leq \mathbf{M}, \end{cases} \tag{5.5}$$

we have, for $m \in \{1, \dots, n + 1\}$ and $r \in [\mathcal{X}, \frac{1}{2}R]$, the estimate

$$\begin{aligned} & \left(\frac{r}{R} + \frac{1}{r}\right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\nabla \xi_m - \ell\|_{\underline{L}^2(B_r)} + \|\nabla \xi_m\|_{\underline{L}^2(B_r)} \\ & \leq C \sum_{i=0}^m \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)}^{\frac{1}{i+1}} + \frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)}^{\frac{1}{i}} \right)^{m+1}. \end{aligned} \tag{5.6}$$

Proof. We start by fixing some notation. Let $q(n, d, \Lambda)$ be as in Lemma C.1, applied for $n + 1$ instead of n . Corresponding this q , choose \mathcal{X} to be the maximum of minimum scales appearing in Proposition 5.1, Theorem 1.3 and Lemma 4.4, of which last two are valid for $n + 1$ in place of N by (4.1). We also assume that $\mathcal{X} \geq R$, by taking $\mathcal{X} \vee R$ instead of \mathcal{X} , where R will be fixed in the course of the proof to depend on parameters (n, M, data) . Chosen this way, \mathcal{X} satisfies $\mathcal{X} \leq \mathcal{O}_\sigma(C)$ for some constants $\sigma(n, \text{data}) > 0$ and $C(n, M, \text{data}) < \infty$.

Let $r_j = \theta^j \eta R$, where θ is as in Lemma 4.4 and $\eta \in (0, \frac{1}{2}]$. The constant η , as well as R , will be fixed in the course of the proof, so that η is small and R is large. We track down the dependencies on η and R carefully and, in particular, constants denoted by C below do not depend on them. We may assume, without loss of generality, that $\eta R \geq \mathcal{X}$.

Set, for $k \in \{0, 1, \dots, n\}$,

$$\begin{aligned} \mathbf{E}_k & := \frac{1}{2r_0} \left\| \xi_{k+1} - (\xi_{k+1})_{B_{2r_0}} \right\|_{\underline{L}^2(B_{2r_0})} \\ & \quad + \sum_{i=0}^k \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} + \frac{1}{R} \|w_{i+1} - (w_{i+1})_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+2}{i+1}}. \end{aligned} \tag{5.7}$$

We denote

$$R_k := \frac{1}{2} \left(1 + 2^{-k} \right) R.$$

Step 1. Induction on degree. We assume that, for fixed $k \in \{0, \dots, n\}$, we have, for every $m \in \{0, \dots, k\}$ and every $r \in [\mathcal{X}, R_m]$,

$$\begin{aligned} & \left(\frac{r}{R} + \frac{1}{r}\right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\nabla \xi_m - \ell\|_{\underline{L}^2(B_r)} + \|\nabla \xi_m\|_{\underline{L}^2(B_r)} \\ & \leq C \sum_{i=0}^m \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i+1}} + \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i}}. \end{aligned} \tag{5.8}$$

Notice that if $k = 0$, then (5.8) follows by Proposition 5.1.

Step 2. Caccioppoli estimate for ξ_{k+1} . We show that, for all $r \in [\mathcal{X}, R_k]$,

$$\|\nabla \xi_{k+1}\|_{\underline{L}^2(B_{(1-2^{-k-2})r})} \leq \frac{C}{r} \|\xi_{k+1} - (\xi_{k+1})_{B_r}\|_{\underline{L}^2(B_r)} + C \mathbf{E}_k. \tag{5.9}$$

We first have by (C.2) that

$$\begin{aligned} & \|\nabla \xi_{k+1}\|_{\underline{L}^2(B_{(1-2^{-k-3})r})} \\ & \leq C \sum_{i=1}^k \left(\frac{1}{r} \|\xi_i - (\xi_i)_{B_r}\|_{\underline{L}^2(B_r)} \right)^{\frac{k+2}{i+1}} + C \|\nabla \xi_0\|_{\underline{L}^q(B_r)}^{k+2} + C \sum_{i=1}^{k+1} \left\| \nabla \frac{w_i}{i!} \right\|_{\underline{L}^q(B_r)}^{\frac{k+2}{i}}. \end{aligned}$$

By Theorem 1.3 and the choice of q in the beginning of the proof, we obtain, for every $r \in [\mathcal{X}, \frac{1}{2}R]$ and $m \in \{1, \dots, k+1\}$, the estimates

$$\|\nabla w_m\|_{\underline{L}^q(B_r)}^{\frac{1}{m}} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{1}{i}} \tag{5.10}$$

and, by Proposition 5.1,

$$\|\nabla \xi_0\|_{\underline{L}^q(B_r)} \leq \frac{C}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)}.$$

Combining above displays yields (5.9) in view of the induction assumption, that is, that (5.8) holds for $m \in \{0, \dots, k\}$.

Step 3. We prove that there is a constant $C < \infty$ independent of η and R such that, for $j \in \mathbb{N}_0$ such that $r_j \geq 2(\mathcal{X} \vee R)$, we have

$$\begin{aligned} & \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_{j+1}} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \\ & \leq \frac{1}{2} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})} + C \frac{\varepsilon_j}{r_j} \|\xi_{k+1} - (\xi_{k+1})_{B_{r_j}}\|_{\underline{L}^2(B_{r_j})} + C \varepsilon_j^{\frac{1}{k+1}} \mathbf{E}_k, \end{aligned} \tag{5.11}$$

where

$$\varepsilon_j := \frac{1}{2} \left(\frac{r_j}{R} + \frac{1}{r_j} \right)^{\frac{\alpha}{2}}, \tag{5.12}$$

and $\alpha(\text{data})$ is the minimum of parameter α appearing in statements of Proposition 5.1 and Lemma 4.4.

Let us fix $j \in \mathbb{N}_0$ such that $r_j \geq 2(\mathcal{X} \vee R)$. We argue in two cases, namely, either (5.13) or (5.15) below is valid. We prove that in both cases (5.11) follows.

Step 3.1. We first assume that

$$\mathbf{E}_k > \varepsilon_j \left(\frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} + \sum_{i=1}^{k+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+1}{i}} \right), \tag{5.13}$$

where ε_j has been fixed in (5.12). We show that this implies

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_r)} \leq C \varepsilon_j \mathbf{E}_k. \tag{5.14}$$

Notice that this gives (5.11). To show (5.14), we have by the triangle inequality that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})} \leq \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_0 - \ell\|_{\underline{L}^2(B_{r_j})} + \sum_{i=1}^{k+1} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|w_i - \ell\|_{\underline{L}^2(B_{r_j})}.$$

By the choice of α , we get by Proposition 5.1 that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_0 - \ell\|_{\underline{L}^2(B_{r_j})} \leq C \left(\frac{r_j}{R} + \frac{1}{r_j} \right)^\alpha \frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)}$$

and, by Proposition 4.2,

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|w_i - \ell\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r_j}{R} + \frac{1}{r_j} \right)^\alpha \sum_{h=1}^i \left(\frac{1}{R} \|w_h - (w_h)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{i}{h}}.$$

Combining the estimates and using (5.13), we have that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r_j}{R} + \frac{1}{r_j} \right)^\alpha \varepsilon_j^{-1} \mathbf{E}_k.$$

We then obtain (5.14) by the choice of ε_j in (5.12), provided that (5.13) is valid.

Step 3.2. We then assume that (5.13) is not the case, that is,

$$\mathbf{E}_k \leq \varepsilon_j \left(\frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} + \sum_{i=1}^{k+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+1}{i}} \right) \tag{5.15}$$

holds with ε_j defined in (5.12). We validate (5.11) also in this case. To this end, let us first observe an immediate consequence of (5.15). By Young’s inequality, we have

$$\begin{aligned} & \varepsilon_j \sum_{i=1}^{k+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+1}{i}} \\ & \leq C \varepsilon_j^{\frac{k+2}{k+1}} + \frac{1}{4} \sum_{i=0}^k \left(\frac{1}{R} \|w_{i+1} - (w_{i+1})_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+2}{i+1}} \end{aligned}$$

and

$$\varepsilon_j \frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} \leq 4\varepsilon_j^{\frac{k+2}{k+1}} + \frac{1}{4} \left(\frac{1}{R} \|\xi_0 - (\xi_0)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{k+2}.$$

Hence we get, by the definition of \mathbf{E}_k in (5.7) and reabsorption, that

$$\mathbf{E}_k \leq C \varepsilon_j^{\frac{k+2}{k+1}}. \tag{5.16}$$

Let then $w_{m+2,j}$ solve

$$\begin{cases} -\nabla \cdot \left(D_p^2 L(\nabla u, x) \nabla w_{m+2,j} \right) = \nabla \cdot \left(\mathbf{F}_{m+2}(\nabla u, \nabla w_1, \dots, \nabla w_{m+1}, \cdot) \right) & \text{in } B_{r_j}, \\ w_{m+2,j} = \xi_{m+1} & \text{on } \partial B_{r_j}. \end{cases}$$

It follows from Lemma C.1, together with (5.2), (5.10) and (5.8), assumed inductively for $m \in \{1, \dots, k\}$, that

$$\begin{aligned} & \left\| \nabla \cdot \left(D^2 L(\nabla u, \cdot) \nabla (\xi_{k+1} - w_{k+2,j}) \right) \right\|_{\underline{H}^{-1}(B_{r_j})} \\ & \leq C \sum_{i=1}^{k+1} \left(\|\nabla \xi_{i-1}\|_{\underline{L}^{2+\delta}(B_{r_j})}^{\frac{1}{7}} + \|\nabla \xi_0\|_{\underline{L}^{\frac{8}{\delta}(k+3)}(B_{r_j})} + \|\nabla w_i\|_{\underline{L}^{\frac{8}{\delta}(k+3)}(B_{r_j})}^{\frac{1}{7}} \right)^{k+3} \\ & \leq C \mathbf{E}_k^{\frac{k+3}{k+2}}. \end{aligned}$$

Consequently, since $\xi_{k+1} - w_{k+2,j} \in H_0^1(B_{r_j})$, (5.16) yields

$$\frac{1}{r_j} \|\xi_{k+1} - w_{k+2,j}\|_{\underline{L}^2(B_{r_j})} \leq C \varepsilon_j^{\frac{1}{k+1}} \mathbf{E}_k. \quad (5.17)$$

By Lemma 4.4 we have

$$\begin{aligned} & \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_{j+1}} \|w_{k+2,r} - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \\ & \leq \frac{1}{2} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|w_{k+2,j} - \ell\|_{\underline{L}^2(B_{r_j})} + C \varepsilon_j \frac{1}{r_j} \|w_{k+2,j} - (w_{k+2,j})_{B_{r_j}}\|_{\underline{L}^2(B_{r_j})} \\ & \quad + C \varepsilon_j \sum_{i=1}^{k+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{k+2}{i}}. \end{aligned} \quad (5.18)$$

Combining this with (5.17) and the triangle inequality yields (5.11).

Step 4. We show that, for $r \in [2(\mathcal{X}_\sigma \vee \mathbf{R}), r_0]$ we have

$$\frac{1}{r} \|\xi_{k+1} - (\xi_{k+1})_{B_r}\|_{\underline{L}^2(B_r)} \leq C \mathbf{E}_k. \quad (5.19)$$

We proceed inductively, and assume $J \in \mathbb{N}_0$ is such that $r_j \geq 2(\mathcal{X} \vee \mathbf{R})$ and for all $j \in \{0, \dots, J\}$ we assume that there exists a constant $\mathbf{C}(d, \theta) \in [1, \infty)$, independent of η and \mathbf{R} , such that

$$\frac{1}{r_j} \|\xi_{k+1} - (\xi_{k+1})_{B_{r_j}}\|_{\underline{L}^2(B_{r_j})} \leq C \mathbf{E}_k.$$

This is true for $J = 0$ by the definition of \mathbf{E}_k . We claim that it continues to hold for $j = J + 1$ as well. Combining (5.14) and (5.11) with the induction assumption we have, for $j \in \{1, \dots, J\}$, that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_{j+1}} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_{j+1}})} \leq \frac{1}{2} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})} + C \left(C \varepsilon_j + \varepsilon_j^{\frac{1}{k+1}} \right) \mathbf{E}_k.$$

Since $r_j \geq \mathbf{R}$, we may take \mathbf{R} large and η small enough so that

$$C \sum_{j=0}^J \left(C \varepsilon_j + \varepsilon_j^{\frac{1}{k+1}} \right) \leq C \mathbf{C} \left(1 - \theta^{\frac{\alpha}{n+1}} \right)^{-1} \left(\eta + \frac{1}{\mathbf{R}} \right)^{\frac{1}{k+1}} \leq \frac{1}{2}.$$

Thus, by summing and reabsorbing,

$$\sum_{j=0}^{J+1} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})} \leq \inf_{\ell \in \mathcal{P}_1} \frac{1}{r_0} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_0})} + \mathbf{E}_k \leq 2\mathbf{E}_k. \tag{5.20}$$

Letting ℓ_j being the minimizing affine function in $\inf_{\ell \in \mathcal{P}_1} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})}$, we obtain by the above display and the triangle inequality that

$$|\nabla \ell_{j+1} - \nabla \ell_0| \leq C(d)\theta^{-\frac{d}{2}-1}\mathbf{E}_k.$$

By the triangle inequality again, we obtain that

$$|\nabla \ell_0| \leq C(d) \left(\frac{1}{r_0} \|\xi_{k+1} - \ell_0\|_{\underline{L}^2(B_{r_0})} + \frac{1}{r_0} \|\xi_{k+1} - (\xi_{k+1})_{B_{r_0}}\|_{\underline{L}^2(B_{r_0})} \right) \leq 2\mathbf{E}_k$$

and, consequently, for $C(d, \theta) < \infty$,

$$|\nabla \ell_{J+1}| \leq C\mathbf{E}_k.$$

We thus obtain by the triangle inequality, together with (5.20) and the above display, that, there exists $C(d, \theta) < \infty$ such that

$$\frac{1}{r_{J+1}} \|\xi_{k+1} - (\xi_{k+1})_{B_{r_{J+1}}}\|_{\underline{L}^2(B_{r_{J+1}})} \leq C\mathbf{E}_k.$$

Hence we can take $C = C$, proving the induction step. Letting then J be such that $r \in (r_{J+1}, r_J]$, we obtain (5.19) by giving up a volume factor.

Step 5. Conclusion. To conclude, we obtain from (5.11) and (5.19) by iterating that

$$\inf_{\ell \in \mathcal{P}_1} \frac{1}{r_j} \|\xi_{k+1} - \ell\|_{\underline{L}^2(B_{r_j})} \leq C \left(2^{-j} + \sum_{i=0}^j 2^{i-j} \varepsilon_j \right) \mathbf{E}_k$$

We hence find α such that, for all $r \in [2(\mathcal{X} \wedge \mathbf{R}), r_0]$,

$$\left(\frac{r}{\mathbf{R}} + \frac{1}{r} \right)^{-\alpha} \inf_{\ell \in \mathcal{P}_1} \frac{1}{r} \|\nabla \xi_{k+1} - \ell\|_{\underline{L}^2(B_r)} \leq C\mathbf{E}_k.$$

Moreover, by (5.19) and (5.9) we deduce that, for all $r \in [2(\mathcal{X} \wedge \mathbf{R}), R_{k+1}]$,

$$\|\nabla \xi_{k+1}\|_{\underline{L}^2(B_r)} \leq C \sum_{i=0}^{k+1} \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} + \frac{1}{R} \|w_{i+1} - (w_{i+1})_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m+1}{i+1}}.$$

Therefore, (5.8) is valid for $m = k + 1$, by giving up a volume factor. This proves the induction step and completes the proof. \square

5.2. Improvement of Spatial Integrability

Following the strategy in Section 4.3, we improve the spatial integrability of (5.6) from L^2 to L^q for every $q \in [2, \infty)$. Fix $q \in [2, \infty)$. Now, ξ_{n+1} solves

$$-\nabla \cdot \left(D_p^2 L(\nabla u, \cdot) \nabla \xi_{n+1} \right) = \nabla \cdot \mathbf{E}_{n+1},$$

where \mathbf{E}_{n+1} satisfies the estimate (C.13) for $\delta = q$ and $n + 1$ instead of n . Recall that both (1.22) and (1.23) are valid for $m \in \{1, \dots, n\}$ with $2nq$ instead of q . These, together with Proposition 5.1, yields by (C.13) that, for $R \geq \mathcal{X}$,

$$\begin{aligned} \|\mathbf{E}_{n+1}\|_{\underline{L}^q(B_{\mathcal{X}/2})} &\leq C \sum_{i=0}^n \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i+1}} \\ &\quad + C \sum_{i=1}^{n+1} \left(\frac{1}{R} \|w_i - (w_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+2}{i}}. \end{aligned}$$

Having this at our disposal, we may repeat the proof of Section 4.3 to conclude (5.6) for $q \in [2, \infty)$.

6. Sharp Estimates of Linearization Errors

Here we show that Corollary 1.4 is a consequence of Theorems 1.1, 1.2 and 1.3.

Proof of Corollary 1.4. For each $k \in \{0, \dots, n - 1\}$, let $\{B_j^{(k)}\}$ be a finite family of balls such that

$$U_k \subseteq \bigcup_j B_j^{(k)} \quad \text{and} \quad \bigcup_j 4B_j^{(k)} \subseteq U_{k+1}. \tag{6.1}$$

By the Vitali covering lemma, we may further assume that $\frac{1}{3}B_j^{(k)} \cap \frac{1}{3}B_i^{(k)} = \emptyset$ whenever $i \neq j$. Let \mathcal{Z} be the finite set consisting of the centers of these balls. The size of \mathcal{Z} depends only on the geometry of the sequence of domains U_0, U_1, \dots, U_n . Let \mathcal{X} be the maximum of the random variables \mathcal{X} given in Theorem 1.3, centered at elements of \mathcal{Z} , divided by the radius of the smallest ball. We assume that $r \geq \mathcal{X}$. This ensures the validity of Theorem 1.3 in each of the balls $B_j^{(k)}$: that is, for every $q \in [2, \infty)$ and $m \in \{1, \dots, n\}$, we have the estimate

$$\|\nabla w_m\|_{\underline{L}^q(B_j^{(m)})} \leq C \sum_{i=1}^m \|\nabla w_i\|_{\underline{L}^2(2B_j^{(m)})}^{\frac{m}{i}} \tag{6.2}$$

and hence, by the covering,

$$\|\nabla w_m\|_{\underline{L}^q(U_m)} \leq C \sum_{i=1}^m \|\nabla w_i\|_{\underline{L}^2(U_i)}^{\frac{m}{i}}. \tag{6.3}$$

Proceeding with the proof of the corollary, we define, as usual,

$$\xi_m := v - u - \sum_{k=1}^m \frac{1}{k!} w_k.$$

Arguing by induction, let us suppose that $m \in \mathbb{N}$ with $m \geq 1$ and $\theta > 0$ are such that, for every $j \in \{0, \dots, m - 1\}$ and $q \in [2, \infty)$, there exists $C(q, \text{data}) < \infty$ such that

$$\|\nabla \xi_j\|_{\underline{L}^{2+\theta}(U_j)} + \|\nabla w_{j+1}\|_{\underline{L}^q(U_{j+1})} \leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{j+1}. \tag{6.4}$$

This obviously holds for $m = 1$ and some $\theta(d, \Lambda) > 0$ by the Meyers estimate and Theorem 1.3. We will show that it must also hold for $m + 1$ and some other (possibly smaller) exponent $\theta > 0$.

Step 1. We show that

$$\|\nabla w_{m+1}\|_{\underline{L}^2(U_{m+1})} \leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{m+1}. \tag{6.5}$$

By the basic energy estimate,

$$\|\nabla w_{m+1}\|_{\underline{L}^2(U_{m+1})} \leq C \cdot \begin{cases} \|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} & \text{if } m = 0, \\ \|\mathbf{F}_{m+1}(\cdot, \nabla u, \nabla w_1, \dots, \nabla w_m)\|_{L^2(U_{m+1})} & \text{if } m \geq 1, \end{cases}$$

where

$$|\mathbf{F}_{m+1}(\cdot, \nabla u, \nabla w_1, \dots, \nabla w_m)| \leq C \sum_{k=1}^m |\nabla w_k|^{\frac{m+1}{k}}.$$

By Theorem 1.3 (using our definition of \mathcal{X} and the fact that $r \geq \mathcal{X}$) and the induction hypothesis, we have, for every $k \in \{1, \dots, m\}$,

$$\|\nabla w_k\|_{\underline{L}^{\frac{2(m+1)}{k}}(U_m)}^{\frac{m+1}{k}} \leq C \sum_{i=1}^{k-1} \left(\frac{1}{R} \|w_i\|_{\underline{L}^2(U_{m-1})} \right)^{\frac{m+1}{i}} \leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{m+1}.$$

This completes the proof of (6.5).

Step 2. We show that

$$\|\nabla \xi_m\|_{\underline{L}^{2+\theta/2}(U_m)} \leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{m+1}. \tag{6.6}$$

Observe that, since $m \geq 1$, $\xi_m \in H_0^1(U_m)$, that is, ξ_m vanishes on ∂U_m . Therefore, by Meyers estimate and Lemma C.1, in particular (C.13) with $q = 2 + \frac{1}{2}\theta$ and $\delta = \frac{1}{2}\theta$, we get

$$\begin{aligned} \|\nabla \xi_m\|_{\underline{L}^{2+\theta/2}(U_m)} &\leq C \left(\sum_{i=1}^{m-1} \|\nabla \xi_i\|_{\underline{L}^{2+\theta}(B_R)}^{\frac{m+1}{i+1}} + \|\nabla \xi_0\|_{\underline{L}^{\frac{2m(4+\theta)}{\theta}}(B_R)}^{m+1} + \sum_{i=1}^{m-1} \|\nabla w_i\|_{\underline{L}^{\frac{2m(4+\theta)}{\theta}}(B_R)}^{\frac{m+1}{i}} \right) \\ &\leq C \left(\|\nabla u - \nabla v\|_{\underline{L}^2(rU_0)} \right)^{m+1}. \end{aligned}$$

This completes the proof of (6.6).

The corollary now follows by induction. \square

7. Liouville Theorems and Higher Regularity

In this section we prove Theorem 1.6 by an induction in the degree n . The initial step $n = 1$ has been already established in [1]. Indeed, (i)₁ and (ii)₁ are consequences of [1, Theorem 5.2], and (iii)₁ follows from Theorem 1.5 which is [1, Theorem 1.3]. Moreover, these estimates hold with optimal stochastic integrability, namely we may take any $\sigma \in (0, d)$ for $n = 1$ (with the constant C then depending additionally on σ).

Throughout the section we will use the following notation: given $p \in \mathbb{R}^d$ and $k \in \mathbb{N}$, we denote

$$\mathcal{A}_k^p := \left\{ u \in H_{\text{loc}}^1(\mathbb{R}^d) : -\nabla \cdot (\mathbf{a}_p \nabla u) = 0, \lim_{r \rightarrow \infty} r^{-k-1} \|u\|_{\underline{L}^2(B_r)} = 0 \right\}$$

and

$$\overline{\mathcal{A}}_k^p := \left\{ \bar{u} \in H_{\text{loc}}^1(\mathbb{R}^d) : -\nabla \cdot (\bar{\mathbf{a}}_p \nabla \bar{u}) = 0, \lim_{r \rightarrow \infty} r^{-k-1} \|\bar{u}\|_{\underline{L}^2(B_r)} = 0 \right\}.$$

Remark 7.1. In proving Theorem 1.6(ii)_n by induction in n , it will be necessary to prove a stronger statement, namely that, for every $p \in B_M$, $(\bar{w}_1, \dots, \bar{w}_n) \in \overline{\mathcal{W}}_n^p$ and $(w_1, \dots, w_{n-1}) \in \mathcal{W}_{n-1}^p$ satisfying (1.30) for $m \in \{1, \dots, n-1\}$, there exists w_n satisfying (1.30) for $m = n$ such that $(w_1, \dots, w_n) \in \mathcal{W}_n^p$.

Proof of Theorem 1.6 (ii)_n For fixed n , we will take \mathcal{X} as the maximum of random variables \mathcal{X} appearing in Theorem 1.2, Theorem 1.3 corresponding $q = 2 + \delta$, where δ is as in Theorem 1.2, and a deterministic constant $\mathbf{R}(n, M, d, \Lambda) < \infty$. Clearly (1.29) holds then.

Given $p \in B_M$ and $(\bar{w}_1, \dots, \bar{w}_n) \in \overline{\mathcal{W}}_n^p$, our goal is to prove that there exists a tuple $(w_1, \dots, w_n) \in \mathcal{W}_n^p$ such that (1.30) holds for every $R \geq \mathcal{X}$ and $k \in \{1, \dots, n\}$.

Step 1. Induction assumption. We proceed inductively and assume that there is a tuple $(w_1, \dots, w_{n-1}) \in \mathcal{W}_{n-1}^p$ such that (1.30) is true for every $R \geq \mathcal{X}$ and $m \in \{1, \dots, n-1\}$. The base case for the induction is valid by the results of [1], as mentioned at the beginning of this section. Our goal is therefore to construct w_n such that $(w_1, \dots, w_n) \in \mathcal{W}_n^p$ and (1.30) holds for every $R \geq \mathcal{X}$ and $k \in \{1, \dots, n\}$. Recall that since $(\bar{w}_1, \dots, \bar{w}_n) \in \overline{\mathcal{W}}_n^p$, we have, for every $R \geq \mathcal{X}$, that

$$\sum_{i=1}^n \left(\frac{1}{R} \|\bar{w}_i\|_{\underline{L}^2(B_R)} \right)^{\frac{n}{i}} \leq C(d) \left(\frac{R}{\mathcal{X}} \right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})} \right)^{\frac{n}{i}}. \tag{7.1}$$

Moreover, by the induction assumption, if the lower bound \mathbf{R} for \mathcal{X} is large enough, we deduce by (1.30) that, for $R \geq \mathcal{X}$ and $m \in \{1, \dots, n-1\}$,

$$\sum_{i=1}^n \left(\frac{1}{R} \|w_i\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \leq 2 \left(\frac{R}{\mathcal{X}} \right)^m \sum_{i=1}^m \left(\frac{1}{\mathcal{X}} \|w_i\|_{\underline{L}^2(B_{\mathcal{X}})} \right)^{\frac{m}{i}}. \tag{7.2}$$

Step 1. Let $r_j := 2^j \mathcal{X}$ and let $w_{n,j}, j \in \mathbb{N}$, solve

$$\begin{cases} -\nabla \cdot (\mathbf{a}_p \nabla w_{n,j}) = \nabla \cdot \mathbf{F}_n(p + \nabla \phi_p, \nabla w_1, \dots, \nabla w_{n-1}, \cdot) & \text{in } B_{r_j}, \\ w_{n,j} = \bar{w}_n & \text{on } \partial B_{r_j}. \end{cases} \quad (7.3)$$

We show that

$$\|w_{n,j} - \bar{w}_n\|_{\underline{L}^2(B_{r_{j-1}})} \leq Cr_j^{1-\alpha} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \quad (7.4)$$

Theorem 1.2 yields that, for $m \in \{1, \dots, n\}$, solutions of

$$\begin{cases} -\nabla \cdot (D_p \bar{L}(\nabla \bar{u}_j)) = 0 & \text{in } B_{r_{j+n}}, \\ -\nabla \cdot (D_p^2 \bar{L}(\nabla \bar{u}_j) \nabla \bar{w}_{m,j}) = \nabla \cdot \bar{\mathbf{F}}_m(\nabla \bar{u}_j, \nabla \bar{w}_{1,j}, \dots, \nabla \bar{w}_{m-1,j}) & \text{in } B_{r_{j+n-m}}, \\ \bar{u}_j(x) = p \cdot x + \phi_p(x) & \text{for } x \in \partial B_{r_{j+n}}, \\ \bar{w}_{m,j} = w_m, \quad m \in \{1, \dots, n-1\}, & \text{on } \partial B_{r_{j+n-m}}, \\ \bar{w}_{n,j} = \bar{w}_n & \text{on } \partial B_{r_j} \end{cases}$$

satisfy the estimate

$$\|w_{n,j} - \bar{w}_{n,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\alpha} \left(\|\nabla \bar{w}_n\|_{\underline{L}^{2+\delta}(B_{r_j})} + \sum_{i=1}^{n-1} \left(\|\nabla w_i\|_{\underline{L}^{2+\delta}(B_{r_{j+n-m}})} \right)^{\frac{n}{i}} \right).$$

By Theorem 1.3, together with (7.1) and (7.2), assumed for $m \in \{1, \dots, n-1\}$, we obtain

$$\|w_{n,j} - \bar{w}_{n,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\alpha} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \quad (7.5)$$

Similarly, by Theorems 1.2 and 1.3, together with (7.2) valid for $m \in \{1, \dots, n-1\}$, we get, for $m \in \{1, \dots, n-1\}$, that

$$\|w_m - \bar{w}_{m,j}\|_{\underline{L}^2(B_{r_{j+n-m}})} \leq Cr_j^{1-\alpha} \left(\frac{r_j}{\mathcal{X}}\right)^m \sum_{i=1}^m \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{m}{i}}.$$

Consequently, again by (1.30) assumed for $m \in \{1, \dots, n-1\}$, that

$$\|\bar{w}_m - \bar{w}_{m,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\delta} \left(\frac{r_j}{\mathcal{X}}\right)^m \sum_{i=1}^m \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{m}{i}}. \quad (7.6)$$

We denote, in short,

$$\bar{\mathbf{f}}_{m,j} := \bar{\mathbf{F}}_m(\nabla \bar{u}_j, \nabla \bar{w}_{1,j}, \dots, \nabla \bar{w}_{m-1,j}),$$

together with

$$\bar{\mathbf{a}}_p := D_p^2 \bar{L}(p) \quad \text{and} \quad \bar{\mathbf{f}}_m := \bar{\mathbf{F}}_m(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1}).$$

Now, it is easy to see (cf. the proof of [1, Theorem 5.2]) that

$$\|\nabla \bar{u}_j - p\|_{L^\infty(B_{r_{j+n-1}})} \leq Cr_j^{-\alpha},$$

so that, by Theorem 1.1,

$$\|D_p^2 \bar{L}(\nabla \bar{u}_j) - \bar{\mathbf{a}}_p\|_{L^\infty(B_{r_{j+n-1}})} \leq \|D_p^3 \bar{L}\|_{L^\infty(B_{2M})} \|\nabla \bar{u}_j - p\|_{L^\infty(B_{r_{j+n-1}})} \leq Cr_j^{-\alpha}.$$

Therefore, by an analogous computation to the proof of Lemma 2.11, using (7.6), we get

$$\|\bar{\mathbf{f}}_{n,j} - \bar{\mathbf{f}}_n\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{-\alpha/2} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}$$

as well as

$$\|D_p^2 \bar{L}(\nabla \bar{u}_j) - \bar{\mathbf{a}}_p\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{-\alpha/2-1} \|\bar{w}_n - (\bar{w}_n)_{B_{r_n}}\|_{\underline{L}^2(B_{r_j})}.$$

By testing the equation

$$\begin{cases} -\nabla \cdot (\bar{\mathbf{a}}_p \nabla (\bar{w}_{n,j} - \bar{w}_n)) = \nabla \cdot \left((D_p^2 \bar{L}(\nabla \bar{u}_j) - \bar{\mathbf{a}}_p) \nabla \bar{w}_{n,j} + \bar{\mathbf{f}}_{n,j} - \bar{\mathbf{f}}_n \right) & \text{in } B_{r_j}, \\ \bar{w}_{n,j} - \bar{w}_n = 0 & \text{on } \partial B_{r_j} \end{cases}$$

with $\bar{w}_{n,j} - \bar{w}_n$, we then obtain

$$\|\bar{w}_{n,j} - \bar{w}_n\|_{\underline{L}^2(B_{r_j})} \leq C \left(r_j^{1-\delta} + r_j^{1-\alpha/2}\right) \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}.$$

Therefore, (7.4) follows by (7.5) and the above display by taking $\delta = \alpha/2$.

Step 2. We show that there is w_n such that $(w_1, \dots, w_n) \in \mathbf{W}_n^p$ and w_n satisfies (1.30). Setting $z_{n,j} := w_{n,j} - w_{n,j+1}$ we have by the triangle inequality that

$$\|z_{n,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\delta} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.7}$$

Notice that $z_{n,j}$ is \mathbf{a}_p -harmonic in B_{r_j} . Thus, by [1, Theorem 5.2], we find $\phi_{n,j} \in \mathcal{A}_n^p$ such that, for every $r \in [\mathcal{X}, r_j]$,

$$\|z_{n,j} - \phi_{n,j}\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{r_j}\right)^{n+1} \|z_{n,j}\|_{\underline{L}^2(B_{r_j})}. \tag{7.8}$$

Consequently, for every $r \in [\mathcal{X}, r_j]$,

$$\|z_{n,j} - \phi_{n,j}\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{r_j}\right)^\delta r^{1-\delta} \left(\frac{r}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}.$$

Setting then $\tilde{w}_{n,j} := w_{n,j} - \sum_{h=1}^{j-1} \phi_{n,h}$, we have that $\tilde{w}_{n,k} - \tilde{w}_{n,j} = \sum_{i=j}^{k-1} (z_{n,i} - \phi_{n,i})$ and it follows that, for all $j, k \in \mathbb{N}$, $j < k$,

$$\|\tilde{w}_{n,k} - \tilde{w}_{n,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\delta} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.9}$$

Therefore, $\{\tilde{w}_{n,k}\}_{k=j}^\infty$ is a Cauchy sequence and, by the Caccioppoli estimate and the diagonal argument, we find w_n such that

$$-\nabla \cdot (\mathbf{a}_p \nabla w_n) = \nabla \cdot \mathbf{F}_n(p + \nabla \phi_p, \nabla w_1, \dots, \nabla w_{n-1}, \cdot) \quad \text{in } \mathbb{R}^d \tag{7.10}$$

and, for all $j \in \mathbb{N}$,

$$\|w_n - \tilde{w}_{n,j}\|_{\underline{L}^2(B_{r_j})} \leq Cr_j^{1-\delta} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.11}$$

We now use the facts that $(\bar{w}_1, \dots, \bar{w}_n) \in \overline{\mathbf{W}}_n^p$ and $\phi_{n,j} \in \mathcal{A}_n^p$, together with (7.7) and (7.8), to deduce that

$$\begin{aligned} \sum_{h=1}^{j-1} \|\phi_{n,h}\|_{\underline{L}^2(B_{r_j})} &\leq \sum_{h=1}^{j-1} \left(\frac{r_j}{r_h}\right)^n \|\phi_{n,h}\|_{\underline{L}^2(B_{r_h})} \\ &\leq \sum_{h=1}^{j-1} \left(\frac{r_j}{r_h}\right)^n \left(\|\phi_{n,j}\|_{\underline{L}^2(B_{r_h})} + \|z_{n,j} - \phi_{n,j}\|_{\underline{L}^2(B_{r_h})}\right) \\ &\leq C \sum_{h=1}^{j-1} \left(\frac{r_j}{r_h}\right)^n r_h^{1-\delta} \left(\frac{r_h}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}} \\ &\leq Cr_j^{1-\delta} \left(\frac{r_j}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \end{aligned}$$

Combining this with (7.11) yields that w_n satisfies (1.30). Moreover, obviously $(w_1, \dots, w_n) \in \mathbf{W}_n^p$. The proof is complete. \square

Proof of Theorem 1.6 (i)_n Let \mathcal{X} be as in the beginning of the proof of (ii)_n. Fix $R \geq \mathcal{X}$. We proceed via induction. Assume that (w_1, \dots, w_{n-1}) satisfy (1.30), that is, we find $(\bar{w}_1, \dots, \bar{w}_{n-1}) \in \overline{\mathbf{W}}_{n-1}^p$ such that, for $k \in \{1, \dots, n-1\}$ and $t \geq \mathcal{X}$,

$$\|w_k - \bar{w}_k\|_{\underline{L}^2(B_t)} \leq Ct^{1-\delta} \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.12}$$

Since $(\bar{w}_1, \dots, \bar{w}_{n-1}) \in \mathcal{P}_2 \times \dots \times \mathcal{P}_n$, we have by the homogeneity that

$$\overline{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1}) \in \mathcal{P}_n.$$

We find, by Remark 7.2 below, a solution $\bar{w} \in \mathcal{P}_{n+1}$ such that $(\bar{w}_1, \dots, \bar{w}_{n-1}, \bar{w}) \in \overline{\mathbf{W}}_n^p$ and that there exists a constant $\mathbf{C}(n, d, \Lambda) < \infty$ such that, for every $t \geq \mathcal{X}$,

$$\|\bar{w}\|_{\underline{L}^2(B_t)} \leq \mathbf{C}t \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.13}$$

Consequently, Remark 7.1 provides us \tilde{w} such that $(w_1, \dots, w_{n-1}, \tilde{w}) \in \mathbf{W}_n^p$ and, for $t \geq \mathcal{X}$,

$$\|\tilde{w} - \bar{w}\|_{\underline{L}^2(B_t)} \leq C C t^{1-\alpha} \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.14}$$

Moreover, by the equations and the growth condition at infinity, $w_n - \tilde{w} \in \mathcal{A}_{n+1}^p$. Therefore, by [1, Theorem 5.2], there is $q \in \bar{\mathcal{A}}_{n+1}^p$ such that, for all $t \geq \mathcal{X}$,

$$\|w_n - \tilde{w} - q\|_{\underline{L}^2(B_t)} \leq C t^{-\delta} \|q\|_{\underline{L}^2(B_t)}.$$

We set $\bar{w}_n := \bar{w} + q$. Obviously, $(\bar{w}_1, \dots, \bar{w}_{n-1}, \bar{w}_n) \in \bar{\mathbf{W}}_n^p$ since q is $\bar{\mathbf{a}}_p$ -harmonic. By (7.13) and the triangle inequality we have, for $t \geq \mathcal{X}$, that

$$\begin{aligned} t \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}} &\leq \frac{1}{2C} \|q\|_{\underline{L}^2(B_t)} \\ \implies \|q\|_{\underline{L}^2(B_t)} + \|\bar{w}\|_{\underline{L}^2(B_t)} &\leq 3 \|\bar{w}_n\|_{\underline{L}^2(B_t)} \leq C \left(\frac{t}{\mathcal{X}}\right)^{n+1} \|\bar{w}_n\|_{\underline{L}^2(B_{\mathcal{X}})} \end{aligned}$$

and

$$\begin{aligned} \|q\|_{\underline{L}^2(B_t)} &\leq 2Ct \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}} \\ \implies \|q\|_{\underline{L}^2(B_t)} + \|\bar{w}\|_{\underline{L}^2(B_t)} &\leq 3Ct \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \end{aligned}$$

We thus have by the triangle inequality that

$$\begin{aligned} \|w_n - \bar{w}_n\|_{\underline{L}^2(B_t)} &\leq \frac{1}{t} \|w_n - \tilde{w} - q\|_{\underline{L}^2(B_t)} + \|\tilde{w} - \bar{w}\|_{\underline{L}^2(B_t)} \\ &\leq C t^{-\delta \wedge \alpha} \left(\|q\|_{\underline{L}^2(B_t)} + \|\bar{w}\|_{\underline{L}^2(B_t)} + t \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^{n-1} \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}} \right) \\ &\leq C t^{1-\delta \wedge \alpha} \left(\frac{t}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}, \end{aligned}$$

proving the induction step and finishing the proof of (i)_n. \square

Remark 7.2. We show that there is \bar{w} such that (7.13) is valid. Indeed, by letting \bar{v} solve

$$\begin{cases} -\nabla \cdot \bar{\mathbf{a}}_p \nabla \bar{v} = \nabla \cdot \bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1}) & \text{in } B_R, \\ \bar{v} = 0 & \text{on } \partial B_R, \end{cases} \tag{7.15}$$

using the fact that $\bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1})$ is a polynomial of degree n , it is straightforward to show by homogeneity that

$$\bar{w}(x) = \sum_{m=2}^{n+1} \frac{1}{m!} \nabla^m \bar{v}(0) x^{\otimes m}$$

solves

$$-\nabla \cdot (\bar{\mathbf{a}} \nabla \bar{w}) = \nabla \cdot \bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1}) \quad \text{in } \mathbb{R}^d.$$

We have the estimate

$$\left| \nabla^{m+1} \bar{v}(0) \right| \leq C \sum_{k=0}^{m+1} R^{k-m} \left\| \nabla^k \bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1}) \right\|_{\underline{L}^2(B_R)}$$

for all $m \in \{1, \dots, n\}$, and by the equations of $\bar{w}_1, \dots, \bar{w}_{n-1}$ we see that

$$\left\| \nabla^k \bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1}) \right\|_{\underline{L}^2(B_R)} \leq C R^{-k} \left(\sum_{i=1}^{n-1} \frac{1}{R} \|\bar{w}_i\|_{\underline{L}^2(B_R)} \right)^{\frac{n}{i}}.$$

Therefore, for all $R > 0$ and $m \in \{1, \dots, n\}$,

$$\left| \nabla^{m+1} \bar{v}(0) \right| \leq C R^{-m} \left(\sum_{i=1}^{n-1} \frac{1}{R} \|\bar{w}_i\|_{\underline{L}^2(B_R)} \right)^{\frac{n}{i}}.$$

Thus we have that, for $t \geq R$,

$$\|\nabla \bar{w}\|_{\underline{L}^2(B_t)} \leq C \left(\frac{t}{R} \right)^n \left(\sum_{i=1}^{n-1} \frac{1}{R} \|\bar{w}_i\|_{\underline{L}^2(B_R)} \right)^{\frac{n}{i}},$$

which yields (7.13).

We now turn to the proof of the large-scale $C^{n,1}$ estimate.

Proof of Theorem 1.6 (iii)_n Fix $\mathbf{M} \in [1, \infty)$. By Theorem 1.3 there exist constants $\sigma(n, \mathbf{M}, \text{data}) \in (0, 1)$ and $C(n, \mathbf{M}, \text{data}) < \infty$ and a random variable \mathcal{X} satisfying $\mathcal{X} \leq \mathcal{O}_\sigma(C)$ so that the statement of Theorem 1.3 is valid with $q = 2(n + 2)$. We now divide the proof in two steps.

Step 1. Induction assumption. Assume (iii) _{$n-1$} . Consequently there is $p \in B_C$ and a tuple (w_1, \dots, w_{n-1}) such that, for $k \in \{0, \dots, n - 1\}$ and

$$\xi_k(x) = v(x) - p \cdot x - \phi_p(x) - \sum_{i=1}^k \frac{w_i(x)}{i!}, \quad \xi_0(x) := v(x) - p \cdot x - \phi_p(x),$$

we have that there exists $C(k, \mathbf{M}, \text{data})$ such that, for every $m \in \{0, \dots, n - 1\}$ and for every $r \in [\mathcal{X}, \frac{1}{2}(1 + 2^{-k-2})R]$,

$$\|\nabla \xi_m\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R} \right)^{m+1} \left(\mathbf{H}_k^{m+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right), \quad (7.16)$$

where we denote, in short,

$$\mathbf{H}_m := \sum_{i=0}^m \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{1}{i+1}}.$$

Our goal is to show that (7.16) continues to hold with $m = n$ and for every $r \in [\mathcal{X}, \frac{1}{2}(1 + 2^{-n-2})R]$. The base case $n = 1$ is valid since, by [1, Theorem 1.3], we have that there is $p \in B_C$ such that, for all $r \in [\mathcal{X}, \frac{3}{4}R]$,

$$\|\nabla \xi_0\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R} \right) \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)}. \quad (7.17)$$

Step 2. Construction of a special solution. We construct a solution \tilde{w}_n of

$$-\nabla \cdot (\mathbf{a}_p \nabla \tilde{w}_n) = \nabla \cdot \mathbf{F}_n(p + \nabla \phi_p, \nabla w_1, \dots, \nabla w_{n-1}, \cdot) \quad \text{in } \mathbb{R}^d \quad (7.18)$$

satisfying, for $r \geq \mathcal{X}$,

$$\|\nabla \tilde{w}_n\|_{\underline{L}^{2(n+1)}(B_r)} \leq C \left(\frac{r}{R} \right)^n \left(\mathbf{H}_{n-1}^n \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \quad (7.19)$$

To show this, it first follows by (7.16) and the triangle inequality that, for $m \in \{1, \dots, n-1\}$,

$$\|\nabla w_m\|_{\underline{L}^2(B_r)} \leq C \left(\inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right)^{\frac{1}{m+1}} \|\nabla \xi_m\|_{\underline{L}^2(B_r)}^{\frac{m}{m+1}} + \|\nabla \xi_{m-1}\|_{\underline{L}^2(B_r)}.$$

Since we have Theorem 1.3 at our disposal with $q = 2(n+1)$, we can increase the integrability and obtain by (1.22) and (7.16) that, for $r \in [\mathcal{X}, \frac{1}{2}(1 + \frac{3}{8}2^{-n})R]$ and $m \in \{1, \dots, n-1\}$,

$$\|\nabla w_m\|_{\underline{L}^{2(n+1)}(B_r)} \leq C \left(\frac{r}{R} \right)^m \left(\mathbf{H}_m^m \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \quad (7.20)$$

Consequently, by (1.23) and (7.20), we also get, for $m \in \{0, \dots, n-1\}$ and $r \in [\mathcal{X}, \frac{1}{2}(1 + \frac{3}{16}2^{-m})R]$, that

$$\|\nabla \xi_m\|_{\underline{L}^{2(n+1)}(B_r)} \leq C \left(\frac{r}{R} \right)^{m+1} \left(\mathbf{H}_m^{m+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \quad (7.21)$$

Next, Theorem 1.6(i) $_{n-1}$ yields that we find $(\bar{w}_1, \dots, \bar{w}_{n-1}) \in \overline{\mathbf{W}}_{n-1}^p$ such that, for $m \in \{1, \dots, n-1\}$,

$$\frac{1}{\mathcal{X}} \|w_m - \bar{w}_m\|_{\underline{L}^2(B_{\mathcal{X}})} \leq C \mathcal{X}^{-\delta} \sum_{i=1}^m \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})} \right)^{\frac{m}{i}}. \quad (7.22)$$

In particular, applying this inductively, assuming that the lower bound \mathbf{R} for \mathcal{X} is such that $C\mathbf{R}^{-\delta} \leq \frac{1}{2}$, we deduce by (7.20) that, for $k \in \{1, \dots, n-1\}$,

$$\|\nabla \bar{w}_k\|_{\underline{L}^2(B_{\mathcal{X}})} \leq C \left(\frac{\mathcal{X}}{R} \right)^k \left(\mathbf{H}_k^k \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \quad (7.23)$$

By Remark 7.2 we find a solution \bar{w}_n of

$$-\nabla \cdot (\mathbf{a}_p \nabla \bar{w}_n) = \nabla \cdot \bar{\mathbf{F}}(p, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{n-1})$$

satisfying, for $r \geq \mathcal{X}$,

$$\|\nabla \bar{w}_n\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{\mathcal{X}}\right)^n \left(\sum_{i=1}^{n-1} \|\nabla \bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}.$$

In view of (7.23) this yields, for $r \geq \mathcal{X}$, that

$$\|\nabla \bar{w}_n\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R}\right)^n \left(\mathbf{H}_{n-1}^n \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)}\right). \tag{7.24}$$

By Remark 7.1 we then find \tilde{w}_n solving (7.18) such that, for $r \geq \mathcal{X}$,

$$\|\tilde{w}_n - \bar{w}_n\|_{\underline{L}^2(B_r)} \leq Cr^{1-\delta} \left(\frac{r}{\mathcal{X}}\right)^n \sum_{i=1}^n \left(\frac{1}{\mathcal{X}} \|\bar{w}_i\|_{\underline{L}^2(B_{\mathcal{X}})}\right)^{\frac{n}{i}}. \tag{7.25}$$

Now (7.19) follows by (7.23), (7.24) and (7.25) together with Theorem 1.3.

Step 3. We show the induction step, that is, we validate (7.16) for $k = n$ and $r \in [\mathcal{X}, \frac{1}{2}(1 + 2^{-n-2})R]$. Denote $\tilde{\xi}_n := \xi_{n-1} - \frac{1}{n!}\tilde{w}_n$. We begin by deducing an estimate for $\tilde{\xi}_n$. Appendix C tells us that $\tilde{\xi}_n$ solves the equation

$$-\nabla \cdot (\mathbf{a}_p \nabla \tilde{\xi}_n) = \nabla \cdot \mathbf{E}_n \quad \text{in } B_R$$

and there exist constants $C(n, \mathbf{M}, \text{data}) < \infty$ such that

$$\|\mathbf{E}_n\|_{\underline{L}^2(B_r)} \leq C \sum_{i=0}^{n-1} \|\nabla \xi_i\|_{\underline{L}^{\frac{n+1}{i+1}}(B_r)}^{\frac{n+1}{i+1}} + C \sum_{i=1}^{n-1} \|\nabla w_i\|_{\underline{L}^{\frac{n+1}{i}}(B_r)}^{\frac{n+1}{i}}. \tag{7.26}$$

By (7.21) and (7.20) we then obtain, for $r \in [\mathcal{X}, \frac{1}{2}(1 + \frac{3}{8}2^{-n})R]$, that

$$\|\mathbf{E}_n\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R}\right)^{n+1} \left(\mathbf{H}_{n-1}^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)}\right). \tag{7.27}$$

Next, set, for $j \in \mathbb{N}_0$, $r_j := \frac{1}{2}\theta^j(1 + 2^{-n-2})R$, where $\theta(n, \mathbf{M}, \text{data}) \in (0, \frac{1}{2}]$ will be fixed shortly. Let $\phi_0 \in \mathcal{A}_{n+2}^p$ and, for given $\phi_j \in \mathcal{A}_{n+2}^p$, let h_j solve

$$\begin{cases} -\nabla \cdot (\mathbf{a}_p \nabla h_j) = 0 & \text{in } B_{r_j}, \\ h_j = \tilde{\xi}_n - \phi_j & \text{on } \partial B_{r_j}. \end{cases} \tag{7.28}$$

By testing and (7.27) we get

$$\|\nabla \tilde{\xi}_n - \nabla \phi_j - \nabla h_j\|_{\underline{L}^2(B_{r_j})} \leq C \left(\frac{r}{R}\right)^{n+1} \left(\mathbf{H}_{n-1}^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)}\right). \tag{7.29}$$

Furthermore, by [1, Theorem 5.2], there is $\tilde{\phi}_{j+1} \in \mathcal{A}_{n+2}^p$ such that

$$\left\| \nabla h_j - \nabla \tilde{\phi}_{j+1} \right\|_{\underline{L}^2(B_{r_{j+1}})} \leq C\theta^{n+2} \left\| \nabla h_j - \nabla \phi_j \right\|_{\underline{L}^2(B_{r_j})}.$$

Combining, we have by the triangle inequality, for $\phi_{j+1} := \tilde{\phi}_{j+1} + \phi_j \in \mathcal{A}_{n+2}^p$, that

$$\begin{aligned} \left\| \nabla \tilde{\xi}_n - \nabla \phi_{j+1} \right\|_{\underline{L}^2(B_{r_{j+1}})} &\leq C\theta^{n+2} \left\| \nabla \tilde{\xi}_n - \nabla \phi_j \right\|_{\underline{L}^2(B_{r_j})} \\ &\quad + C\theta^{-\frac{d}{2}} \left(\frac{r}{R} \right)^{n+1} \left(H_{n-1}^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \end{aligned}$$

Choosing $C\theta^{1/2} = 1$, we thus arrive at

$$\begin{aligned} &\frac{1}{r_{j+1}^{n+1}} \left\| \nabla \tilde{\xi}_n - \nabla \phi_{j+1} \right\|_{\underline{L}^2(B_{r_{j+1}})} \\ &\leq \frac{\theta^{1/2}}{r_j^{n+1}} \left\| \nabla \tilde{\xi}_n - \nabla \phi_j \right\|_{\underline{L}^{2+\theta}(B_{r_j})} + \frac{C}{R^{n+1}} \left(H_{n-1}^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \end{aligned}$$

It follows by iteration that, for $r \in [\mathcal{X}, \frac{1}{2}(1 + 2^{-n-2})R]$,

$$\begin{aligned} \inf_{\phi \in \mathcal{A}_{n+2}^p} \left\| \nabla \tilde{\xi}_n - \nabla \phi \right\|_{\underline{L}^2(B_r)} &\leq C \left(\frac{r}{R} \right)^{n+3/2} \inf_{\phi \in \mathcal{A}_{n+2}^p} \left\| \nabla \tilde{\xi}_n - \nabla \phi \right\|_{\underline{L}^2(B_{r_0})} \\ &\quad + C \left(\frac{r}{R} \right)^{n+3/2} \left(H_{n-1}^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right). \end{aligned}$$

We can now revisit [4, Proof of (3.49)] and obtain that there exists $\phi \in \mathcal{A}_{n+1}^p$ such that, taking $w_n = \tilde{w}_n + \phi$ and $\xi_n := \xi_{n-1} - \frac{1}{n!}w_n$, we get, for all $r \in [\mathcal{X}, \frac{1}{2}(1 + 2^{-n-2})R]$,

$$\left\| \nabla \xi_n \right\|_{\underline{L}^2(B_r)} \leq C \left(\frac{r}{R} \right)^{n+1} \left(H_n^{n+1} \wedge \inf_{\phi \in \mathcal{L}_1} \frac{1}{R} \|v - \phi\|_{\underline{L}^2(B_R)} \right) \tag{7.30}$$

Since $\phi \in \mathcal{A}_{n+1}^p$, we see that $(w_1, \dots, w_n) \in W_n^p$. Now (7.16) follows for $k = n$ by the previous inequality together with the Caccioppoli estimate and (7.27). The proof is complete. \square

Acknowledgements. Open access funding provided by University of Helsinki including Helsinki University Central Hospital. SCOTT ARMSTRONG was partially supported by the NSF Grant DMS-1700329. SAMUEL J. FERGUSON was partially supported by NSF Grants DMS-1700329 and DMS-1311833. TUOMO KUUSI was supported by the Academy of Finland and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No 818437).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: Deterministic Regularity Estimates

In this first appendix, we record some deterministic regularity estimates of Schauder and Calderón–Zygmund type for linear equations with Hölder continuous coefficients. These estimates, while well-known, are not typically written with explicit dependence on the regularity of the coefficients, which is needed for our purposes in this paper.

Proposition A.1. (Calderón–Zygmund gradient L^q estimates). *Let $\beta \in (0, 1]$, $q \in [2, \infty)$ and $\mathbf{a} \in \mathbb{R}^{d \times d}$ be a symmetric matrix with entries in $C^{0,\beta}(B_2)$ satisfying*

$$I_d \leq \mathbf{a}(x) \leq \Lambda I_d, \quad \forall x \in B_2.$$

Suppose $\mathbf{f} \in L^q(B_2; \mathbb{R}^d)$ and $u \in H^1(B_2)$ is a solution of

$$-\nabla \cdot (\mathbf{a} \nabla u) = \nabla \cdot \mathbf{f} \quad \text{in } B_2.$$

Then $u \in W_{\text{loc}}^{1,q}(B_2)$ and there exists $C(q, d, \Lambda) < \infty$ such that

$$\|\nabla u\|_{L^q(B_1)} \leq C \exp\left(\frac{C}{\beta} \left(1 - \frac{2}{q}\right)\right) \left(1 + [\mathbf{a}]_{C^{0,\beta}(B_2)}^{\frac{d}{2\beta} \left(1 - \frac{2}{q}\right)}\right) \|\nabla u\|_{L^2(B_2)} + C \|\mathbf{f}\|_{L^q(B_2)}. \quad (\text{A.1})$$

Proof. We will explain how to extract the statement of the proposition from that of [4, Proposition 7.3]. The latter asserts the existence of $\delta_0(q, d, \Lambda) > 0$ such that, for every ball $x \in B_1$ and $r \in \left(0, \frac{1}{2}\right]$ satisfying

$$\text{osc}_{B_{2r}(x)} \mathbf{a} \leq \delta_0,$$

we have, for a constant $C(q, d, \Lambda) < \infty$, the estimate

$$\|\nabla u\|_{\underline{L}^q(B_r(x))} \leq C \left(\|\nabla u\|_{\underline{L}^2(B_{2r}(x))} + \|\mathbf{f}\|_{\underline{L}^q(B_{2r}(x))} \right).$$

Since $\text{osc}_{B_{2r}} \mathbf{a} \leq (2r)^\beta [\mathbf{a}]_{C^{0,\beta}(B_2)}$, we have the above estimate for every $x \in B_1$ and

$$r := \frac{1}{2} \wedge \frac{1}{2} \left(\delta_0 [\mathbf{a}]_{C^{0,\beta}(B_2)}^{-1} \right)^{\frac{1}{\beta}}.$$

From this, Fubini's theorem and Young's inequality for convolutions, we obtain

$$\|\nabla u\|_{L^q(B_1)}^q \leq C \int_{B_1} \left(|\nabla u|^q * \left(\frac{1}{|B_r|} \mathbf{1}_{B_r} \right) \right) (x) dx$$

$$\begin{aligned}
 &= C \int_{B_1} \int_{B_r(x)} |\nabla u(y)|^q \, dy \, dx \\
 &\leq C \int_{B_1} \left(\int_{B_{2r}(x)} |\nabla u(y)|^2 \, dy \right)^{\frac{q}{2}} \, dx + C \int_{B_1} \int_{B_{2r}(x)} |\mathbf{f}(y)|^q \, dy \, dx \\
 &\leq C \int_{B_1} \left| \left(|\nabla u|^2 * \left(\frac{1}{|B_{2r}|} \mathbb{1}_{B_{2r}} \right) \right) (x) \right|^{\frac{q}{2}} \, dx + C \|\mathbf{f}\|_{L^q(B_2)}^q \\
 &\leq C \|\nabla u\|_{L^2(B_2)}^q \left\| \frac{1}{|B_{2r}|} \mathbb{1}_{B_{2r}} \right\|_{L^{q/2}(B_2)}^{\frac{q}{2}} + C \|\mathbf{f}\|_{L^q(B_2)}^q \\
 &= Cr^{-d(\frac{q}{2}-1)} \|\nabla u\|_{L^2(B_2)}^q + C \|\mathbf{f}\|_{L^q(B_2)}^q.
 \end{aligned}$$

This completes the proof. \square

Proposition A.2. Let $\beta \in (0, 1)$ and $\mathbf{a} \in \mathbb{R}^{d \times d}$ be a symmetric matrix with entries in $C^{0,\beta}(B_2)$ satisfying

$$I_d \leq \mathbf{a}(x) \leq \Lambda I_d, \quad \forall x \in B_2.$$

Suppose $\mathbf{f} \in C^{0,\beta}(B_2; \mathbb{R}^d)$ and $u \in H^1(B_2)$ is a solution of

$$-\nabla \cdot (\mathbf{a} \nabla u) = \nabla \cdot \mathbf{f} \text{ in } B_2.$$

Then $u \in C_{\text{loc}}^{1,\beta}(B_2)$ and there exists $C(\beta, d, \Lambda) < \infty$ such that

$$\|\nabla u\|_{L^\infty(B_1)} \leq C \left(1 + [\mathbf{a}]_{C^{0,\beta}(B_2)}^{\frac{d}{2\beta}} \right) \|\nabla u\|_{L^2(B_2)} + C \|\mathbf{f}\|_{C^{0,\beta}(B_2)} \tag{A.2}$$

and

$$[\nabla u]_{C^{0,\beta}(B_1)} \leq C \left(1 + [\mathbf{a}]_{C^{0,\beta}(B_2)}^{1+\frac{d}{2\beta}} \right) \|\nabla u\|_{L^2(B_2)} + C \|\mathbf{f}\|_{C^{0,\beta}(B_2)}. \tag{A.3}$$

Proof. We will explain how to extract the statement of the proposition from the gradient Hölder estimate found in [20, Theorem 3.13]. The latter states that, under the assumption that

$$[\mathbf{a}]_{C^{0,\beta}(B_2)} \leq 1,$$

there exists $C(\beta, d, \Lambda) < \infty$ such that

$$\|\nabla u\|_{C^{0,\beta}(B_1)} \leq C \left(\|\nabla u\|_{L^2(B_2)} + [\mathbf{f}]_{C^{0,\beta}(B_2)} \right).$$

After changing the scale, we obtain the corresponding statement in B_r , which asserts that, under the assumption that

$$r^\beta [\mathbf{a}]_{C^{0,\beta}(B_{2r})} \leq 1,$$

there exists $C(\beta, d, \Lambda) < \infty$ such that

$$\|\nabla u\|_{L^\infty(B_r)} + r^\beta [\nabla u]_{C^{0,\beta}(B_r)} \leq C \left(\|\nabla u\|_{L^2(B_{2r})} + r^\beta [\mathbf{f}]_{C^{0,\beta}(B_{2r})} \right).$$

Therefore we take $r := \frac{1}{2} \wedge [\mathbf{a}]_{C^{0,\beta}(B_2)}^{-\frac{1}{\beta}}$ and apply the previous statement in every ball $B_r(x)$ with $x \in B_1$ to obtain

$$\begin{aligned}
 \|\nabla u\|_{L^\infty(B_1)} + \sup_{x \in B_1} r^\beta [\nabla u]_{C^{0,\beta}(B_r(x))} &\leq C \sup_{x \in B_1} \left(\|\nabla u\|_{L^2(B_{2r}(x))} + r^\beta [\mathbf{f}]_{C^{0,\beta}(B_{2r}(x))} \right) \\
 &\leq C \left(r^{-\frac{d}{2}} \|\nabla u\|_{L^2(B_2)} + r^\beta [\mathbf{f}]_{C^{0,\beta}(B_2)} \right).
 \end{aligned}$$

After a covering argument, we obtain

$$\|\nabla u\|_{L^\infty(B_1)} + r^\beta [\nabla u]_{C^{0,\beta}(B_1)} \leq C \left(r^{-\frac{d}{2}} \|\nabla u\|_{L^2(B_2)} + r^\beta [\mathbf{f}]_{C^{0,\beta}(B_2)} \right),$$

which yields the proposition. \square

Appendix B: Differentiation of \mathbf{F}_m

In this appendix, we show that (2.13) holds.

Lemma B.1. *Fix $m \in \mathbb{N}$ and $h, p \in \mathbb{R}^d$. Suppose that $z \mapsto L(z, x)$ is C^{m+2} and $t \mapsto \mathbf{g}(p + th)$ is m times differentiable at 0. Then*

$$\begin{aligned} & \mathbf{F}_{m+1}(\mathbf{g}(p), D_p \mathbf{g}(p)h^{\otimes 1}, \dots, D_p^m \mathbf{g}(p)h^{\otimes m}, x) \\ &= D_p \left(\mathbf{F}_m(\mathbf{g}(p), D_p \mathbf{g}(p)h^{\otimes 1}, \dots, D_p^{m-1} \mathbf{g}(p)h^{\otimes(m-1)}, x) \right) \cdot h \\ &+ D_p \left(D_p^2 L(\mathbf{g}(p), x) \right) h^{\otimes 1} \left(D_p^m \mathbf{g}(p)h^{\otimes m} \right)^{\otimes 1}. \end{aligned} \quad (\text{B.1})$$

Proof. We first observe that the terms $D_p \mathbf{g}(p)h^{\otimes 1}, \dots, D_p^m \mathbf{g}(p)h^{\otimes m}$ in (B.1) are precisely the directional derivatives of \mathbf{g} in the h direction, up to m th degree. The terms in (B.1) involve derivatives of $z \mapsto L(z, x)$ up to the $(m+2)$ th degree. Hence, we can assume by approximation, without loss of generality, that $z \mapsto L(z, x)$ and $p \mapsto \mathbf{g}(p)$ are polynomials, of degrees at most $m+2$ and m , respectively. Fix $h \in \mathbb{R}^d$ and let $t \in \mathbb{R}$. We write

$$\mathbf{g}(p + th) = \mathbf{g}(p) + \sum_{j=1}^m \frac{t^j}{j!} D_p^j \mathbf{g}(p)h^{\otimes j}.$$

Denote

$$\mathbf{z}_j(p) := D_p^j \mathbf{g}(p)h^{\otimes j} \quad \text{and} \quad \mathbf{Z}(p, t) := \sum_{j=1}^m \frac{t^j}{j!} \mathbf{z}_j(p).$$

Examining the relation between the p and t derivatives of $\mathbf{Z}(p, t)$, we find that

$$D_p \mathbf{Z}(p, t) \cdot h = \sum_{j=1}^m \frac{t^j}{j!} \mathbf{z}_{j+1}(p) = \partial_t \mathbf{Z}(p, t) - \mathbf{z}_1(p). \quad (\text{B.2})$$

Set now, for fixed $h, x \in \mathbb{R}^d$,

$$\mathbf{G}_{h,x}(t, p) := \sum_{k=2}^{m+1} \frac{1}{k!} D_p^{k+1} L(\mathbf{g}(p), x) (\mathbf{Z}(p, t))^{\otimes k}$$

and, by the definition of \mathbf{F}_m in (1.15),

$$\partial_t^m \mathbf{G}_{h,x}(0, p) = \mathbf{F}_m(\mathbf{g}(p), \mathbf{z}_1(p), \dots, \mathbf{z}_{m-1}(p), x),$$

and similarly for \mathbf{F}_{m+1} . Computing the directional derivative, recalling that we assume that $z \mapsto L(z, x)$ is a $(m+2)$ th degree polynomial,

$$D_p \mathbf{G}_{h,x}(t, p) \cdot h = \sum_{k=2}^{m+1} \frac{1}{(k-1)!} D_p^{k+1} L(\mathbf{g}(p), x) (\mathbf{Z}(p, t))^{\otimes(k-1)} (D_p \mathbf{Z}(p, t) \cdot h)^{\otimes 1}$$

$$+ \sum_{k=3}^{m+1} \frac{1}{(k-1)!} D_p^{k+1} L(\mathbf{g}(p), x) (\mathbf{Z}(p, t))^{\otimes(k-1)} (\mathbf{z}_1(p))^{\otimes 1},$$

we get by (B.2) that

$$\partial_t \mathbf{G}_{h,x}(t, p) = D_p \mathbf{G}_{h,x}(t, p) \cdot h + D_p^3 L(\mathbf{g}(p), x) (\mathbf{Z}(p, t))^{\otimes 1} (\mathbf{z}_1(p))^{\otimes 1}.$$

Consequently, we have

$$\partial_t^{m+1} \mathbf{G}_{h,x}(0, p) = D_p \partial_t^m \mathbf{G}_{h,x}(0, p) \cdot h + D_p^3 L(\mathbf{g}(p), x) (\mathbf{z}_1(p))^{\otimes 1} (\mathbf{z}_m(p))^{\otimes 1},$$

which is (B.1), concluding the proof. \square

Appendix C: Linearization Errors

In this appendix we compute the equation satisfied by a higher-order linearization error and thereby obtain gradient estimates.

Lemma C.1. (Equation for the linearization error). Fix $M, R \in (0, \infty)$ and $n \in \mathbb{N}$ with $n \geq 2$. Assume that $p \mapsto L(p, x)$ is $C^{n+1,1}$ for every $x \in \mathbb{R}^d$ and

$$\sum_{k=1}^n \frac{1}{k!} \|D_p^{k+1} L\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq M.$$

Suppose that $u, v, w_1, \dots, w_n \in H^1(B_R)$ satisfy

$$\nabla \cdot (D_p L(\nabla u, x) - D_p L(\nabla v, x)) = 0 \quad \text{in } B_R$$

and, for each $m \in \{1, \dots, n\}$,

$$-\nabla \cdot (D_p^2 L(\nabla u, x) \nabla w_m) = \nabla \cdot \mathbf{F}_m(\nabla u, \nabla w_1, \dots, \nabla w_{m-1}, x) \quad \text{in } B_R,$$

where \mathbf{F}_m is defined in (1.15). Denote $\xi_0 = v - u$ and, for each $m \in \{1, \dots, n\}$,

$$\xi_m := v - u - \sum_{k=1}^m \frac{w_k}{k!}.$$

Then there is vector field \mathbf{E}_n such that ξ_n solves

$$-\nabla \cdot (D^2 L(\nabla u, \cdot) \nabla \xi_n) = \nabla \cdot \mathbf{E}_n \quad \text{in } B_R$$

and there exists a constant $C(n, M, d) < \infty$ such that

$$|\mathbf{E}_n| \leq C \sum_{h=0}^{n-1} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{i=1}^{n-1} \left| \nabla \frac{w_i}{i!} \right|^{\frac{1}{i}} \right)^{n-h}. \quad (\text{C.1})$$

Furthermore, there exist constants $C(n, M, d) < \infty$, $q(n, d) \in (2, \infty)$ and $\delta(d, \Lambda) \in (0, \frac{1}{2}]$ such that

$$\begin{aligned} \|\nabla \xi_n\|_{\underline{L}^{2+\delta} \left(B_{\frac{1}{2}(1+2^{-n}R)} \right)} &\leq C \sum_{i=1}^n \left(\frac{1}{R} \|\xi_i - (\xi_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{n+1}{i+1}} \\ &\quad + C \|\nabla \xi_0\|_{\underline{L}^q(B_R)}^{n+1} + C \sum_{i=1}^{n-1} \left\| \nabla \frac{w_i}{i!} \right\|_{\underline{L}^q(B_R)}^{\frac{n+1}{i}}. \end{aligned} \quad (\text{C.2})$$

Proof. Throughout the proof we use the notation $s_k = \sum_{j=1}^k \frac{w_j}{j!}$ and $\xi_0 = v - u$, so that $\xi_k = \xi_0 - s_k$.

Step 1. Recalling that $\mathbf{F}_1 = 0$, we may rewrite

$$\begin{aligned} & D_p L(\nabla v, x) - D_p L(\nabla u, x) - D^2 L(\nabla u, x) \nabla \xi_n \\ &= \sum_{k=1}^n \frac{1}{k!} \left(D^2 L(\nabla u, x) \nabla w_k + \mathbf{F}_k(\nabla u, \nabla w_1, \dots, \nabla w_{k-1}, x) \right) + \mathbf{E}_n \end{aligned}$$

where we define

$$\begin{aligned} \mathbf{E}_n := & \sum_{k=2}^n \frac{1}{k!} \left(D_p^{k+1} L(\nabla u, x) (\nabla \xi_0)^{\otimes k} - \mathbf{F}_k(\nabla u, \nabla w_1, \dots, \nabla w_{k-1}, x) \right) \\ & + D_p L(\nabla v, x) - \sum_{k=0}^n \frac{1}{k!} D_p^{k+1} L(\nabla u, x) (\nabla \xi_0)^{\otimes k}. \end{aligned}$$

By the equations of u, v and w_k , we have that

$$-\nabla \cdot \left(D^2 L(\nabla u, x) \nabla \xi_n \right) = \nabla \cdot \mathbf{E}_n.$$

It thus remains to estimate \mathbf{E}_n .

Step 2. We show that, for $k \in \{2, \dots, n\}$ and $m \in \{k, \dots, n\}$,

$$(\nabla \xi_0)^{\otimes k} = \mathbf{S}_m^{(k)} + \mathbf{E}_m^{(k)}, \tag{C.3}$$

where $\mathbf{S}_m^{(j)}$ and $\mathbf{E}_m^{(j)}$ are defined, for $j \in \{2, \dots, k\}$, recursively as

$$\mathbf{S}_m^{(j)} := \sum_{i=1}^{m+1-j} \mathbf{S}_{m-i}^{(j-1)} \otimes \nabla \frac{w_i}{i!} \tag{C.4}$$

and

$$\mathbf{E}_m^{(j)} := \sum_{i=1}^{m+1-j} \mathbf{E}_{m-i}^{(j-1)} \otimes \nabla \frac{w_i}{i!} + (\nabla \xi_0)^{\otimes (j-1)} \otimes \nabla \xi_{m-(j-1)}, \tag{C.5}$$

with

$$\mathbf{S}_i^{(1)} := \nabla s_i \quad \text{and} \quad \mathbf{E}_i^{(1)} := \nabla \xi_i. \tag{C.6}$$

Indeed, suppose that we have, for $j \in \{1, \dots, k-1\}$ and $m \in \{j, \dots, n\}$, that

$$(\nabla \xi_0)^{\otimes j} = \mathbf{S}_m^{(j)} + \mathbf{E}_m^{(j)}.$$

This is obviously true for $j = 1$. We compute, for $m \in \{k, \dots, n\}$,

$$\begin{aligned} (\nabla \xi_0)^{\otimes k} &= \sum_{i=1}^{m+1-k} (\nabla \xi_0)^{\otimes (k-1)} \otimes \nabla \frac{w_i}{i!} + (\nabla \xi_0)^{\otimes (k-1)} \otimes \nabla \xi_{m-(k-1)} \\ &= \sum_{i=1}^{m+1-k} \mathbf{S}_{m-i}^{(k-1)} \otimes \nabla \frac{w_i}{i!} + \sum_{i=1}^{m+1-k} \mathbf{E}_{m-i}^{(k-1)} \otimes \nabla \frac{w_i}{i!} + (\nabla \xi_0)^{\otimes (k-1)} \otimes \nabla \xi_{m-(k-1)} \\ &= \mathbf{S}_m^{(k)} + \mathbf{E}_m^{(k)}, \end{aligned}$$

which proves the recursive formula (C.3).

Step 3. We show that, for $k \in \{2, \dots, n\}$ there exists a constant $C(n, k, d) < \infty$ such that

$$\left| \mathbf{E}_n^{(k)} \right| \leq C \sum_{h=1}^{n+1-k} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{i=1}^{n+1-k} \left| \nabla \frac{w_i}{i!} \right|^{\frac{1}{i}} \right)^{n-h}. \quad (\text{C.7})$$

The statement is easy to verify by induction. Indeed, for $m = j = 2$ we have by (C.5) that

$$\left| \mathbf{E}_2^{(2)} \right| \leq |\nabla \xi_1 \otimes \nabla w_1 + \nabla \xi_0 \otimes \nabla \xi_1| \leq C \sum_{h=1}^1 |\nabla \xi_h| (|\nabla \xi_0| + |\nabla w_1|)^{2-h}.$$

Assume then that, for $m \in \{2, \dots, n-1\}$ and $j \in \{2, \dots, m\}$, we have

$$\left| \mathbf{E}_m^{(j)} \right| \leq C \sum_{h=0}^{m+1-j} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{i=1}^{m+1-j} \left| \nabla \frac{w_i}{i!} \right|^{\frac{1}{i}} \right)^{m-h}. \quad (\text{C.8})$$

By the definition of $\mathbf{E}_n^{(j)}$, we have, for $j \in \{2, \dots, n\}$, that

$$\left| \mathbf{E}_n^{(j)} \right| \leq C \sum_{i=1}^{n+1-j} \left| \mathbf{E}_{n-i}^{(j-1)} \right| \left| \nabla \frac{w_i}{i!} \right| + C |\nabla \xi_0|^{j-1} |\nabla \xi_{n-(j-1)}| \quad (\text{C.9})$$

By (C.8), using Fubini for sums, we obtain,

$$\begin{aligned} \sum_{i=1}^{n+1-j} \left| \mathbf{E}_{n-i}^{(j-1)} \right| \left| \nabla \frac{w_i}{i!} \right| &\leq C \sum_{i=1}^{n+1-j} \sum_{h=0}^{n+2-i-j} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{\ell=1}^{n+2-i-j} \left| \nabla \frac{w_\ell}{\ell!} \right|^{\frac{1}{\ell}} \right)^{n-i-h} \left| \nabla \frac{w_i}{i!} \right| \\ &\leq C \sum_{h=0}^{n+1-j} |\nabla \xi_h| \sum_{i=1}^{n+2-j-h} \left(|\nabla \xi_0| + \sum_{\ell=1}^{n+1-j} \left| \nabla w_\ell \right|^{\frac{1}{\ell}} \right)^{n-i-h} \left| \nabla \frac{w_i}{i!} \right| \\ &\leq C \sum_{h=0}^{n+1-j} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{\ell=1}^{n+1-j} \left| \nabla \frac{w_\ell}{\ell!} \right|^{\frac{1}{\ell}} \right)^{n-h}. \end{aligned}$$

This, together with (C.9), proves the induction step, and gives also (C.7).

Step 4. We show that

$$\left| \sum_{k=2}^n \left(\frac{1}{k!} D_p^{k+1} L(\nabla u, x) \left((\nabla \xi_0)^{\otimes k} - \mathbf{S}_n^{(k)} \right) \right) \right| \leq C \sum_{h=0}^{n-1} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{i=1}^{n-1} \left| \nabla \frac{w_i}{i!} \right|^{\frac{1}{i}} \right)^{n-h}. \quad (\text{C.10})$$

By the recursive formula we have, for $k \in \{2, \dots, n\}$, that

$$(\nabla \xi_0)^{\otimes k} = \mathbf{S}_n^{(k)} + \mathbf{E}_n^{(k)},$$

and thus (C.10) follows by (C.7).

Step 5. We show that

$$\sum_{k=2}^n \frac{1}{k!} \left(D_p^{k+1} L(\nabla u, x) \mathbf{S}_n^{(k)} - \mathbf{F}_k(\nabla u, \nabla w_1, \dots, \nabla w_{k-1}, x) \right) = 0. \quad (\text{C.11})$$

For this, we first abbreviate

$$\mathbf{F}_k = \mathbf{F}_k(\nabla u, \nabla w_1, \dots, \nabla w_{k-1}, x)$$

and observe that, by definition,

$$\frac{1}{k!} \mathbf{F}_k = \sum_{j \geq 2} \frac{1}{j!} D_p^{j+1} L(\nabla u, x) \left(\sum_{i_1 + \dots + i_j = k : i_1, \dots, i_j \geq 1} \nabla \frac{w_{i_1}}{i_1!} \otimes \dots \otimes \nabla \frac{w_{i_j}}{i_j!} \right).$$

Second, we observe that, by induction on $j \geq 2$, we have

$$\mathbf{S}_n^{(j)} = \sum_{m \leq n} \left(\sum_{i_1 + \dots + i_j = m : i_1, \dots, i_j \geq 1} \nabla \frac{w_{i_1}}{i_1!} \otimes \dots \otimes \nabla \frac{w_{i_j}}{i_j!} \right)$$

for all $n \geq j$. Third, by commutativity of addition, we observe that

$$\begin{aligned} & \sum_{m \leq n} \left(\sum_{j \geq 2} \frac{1}{j!} D_p^{j+1} L(\nabla u, x) \left(\sum_{i_1 + \dots + i_j = m : i_1, \dots, i_j \geq 1} \nabla \frac{w_{i_1}}{i_1!} \otimes \dots \otimes \nabla \frac{w_{i_j}}{i_j!} \right) \right) \\ &= \sum_{j \geq 2} \frac{1}{j!} D_p^{j+1} L(\nabla u, x) \left(\sum_{m \leq n} \left(\sum_{i_1 + \dots + i_j = m : i_1, \dots, i_j \geq 1} \nabla \frac{w_{i_1}}{i_1!} \otimes \dots \otimes \nabla \frac{w_{i_j}}{i_j!} \right) \right). \end{aligned}$$

Finally, letting $\mathbf{F}_m = 0$ for $m < 2$ and $\mathbf{S}_n^{(j)} = 0$ for $j > n$ for notational convenience, we note that the above equation may be rewritten as

$$\sum_{m \leq n} \frac{1}{m!} \mathbf{F}_m = \sum_{j \geq 2} \frac{1}{j!} D_p^{j+1} L(\nabla u, x) \mathbf{S}_n^{(j)},$$

which is (C.11).

Step 6. Conclusion. We have that

$$\left| D_p L(\nabla v, x) - \sum_{k=0}^n \frac{1}{k!} D_p^{k+1} L(\nabla u, x) (\nabla v - \nabla u)^{\otimes k} \right| \leq C |\nabla v - \nabla u|^{n+1}. \quad (\text{C.12})$$

Indeed, by a Taylor expansion, we see that

$$\left| D_p L(z_0 + z, x) - \sum_{k=0}^n \frac{1}{k!} D_p^{k+1} L(z_0, x) z^{\otimes k} \right| \leq \left[D_p^{n+1} L(\cdot, x) \right]_{C^{0,1}(B_{|z|}(z_0))} \frac{|z|^{n+1}}{(k+1)!}.$$

Applying this with $z_0 = \nabla u$ and $z = \nabla v - \nabla u$ gives (C.12). Combining this with the previous steps yields the desired estimate (C.1) for \mathbf{E}_n . Finally, by the Hölder and Young inequalities, we get, for all $q \in [2, \infty)$ and $r \in (0, R]$,

$$\begin{aligned} \|\mathbf{E}_n\|_{\underline{L}^q(B_r)} &\leq \sum_{h=1}^{n-1} \|\nabla \xi_h\|_{\underline{L}^{q+\delta}(B_r)} \left\| |\nabla \xi_0| + \sum_{i=1}^{n-1} \left| \nabla \frac{w_i}{i!} \right|^{\frac{1}{i}} \right\|_{\underline{L}^{\frac{nq(q+\delta)}{\delta}}(B_r)}^{n-h} \\ &\leq C \left(\sum_{h=1}^{n-1} \|\nabla \xi_h\|_{\underline{L}^{q+\delta}(B_R)}^{\frac{n+1}{h+1}} + \|\nabla \xi_0\|_{\underline{L}^{\frac{nq(q+\delta)}{\delta}}(B_R)}^{n+1} + \sum_{i=1}^{n-1} \left\| \nabla \frac{w_i}{i!} \right\|_{\underline{L}^{\frac{nq(q+\delta)}{\delta}}(B_R)}^{\frac{n+1}{i}} \right), \end{aligned} \tag{C.13}$$

Let δ_0 be the Meyers exponent corresponding Λ . Let

$$q_h := q + \frac{1}{2}\delta_0 + \frac{1}{2} \frac{n+1-h}{n+1} \delta_0 \quad \text{and} \quad q := \frac{16}{\delta_0} n(n+1).$$

Set also $R_h := \frac{1}{2}(1 + 2^{-h})R$. With this notation the previous display yields, by Hölder’s inequality, that

$$\|\mathbf{E}_m\|_{\underline{L}^{qm}(B_{R_m})} \leq C \left(\sum_{h=1}^{m-1} \|\nabla \xi_h\|_{\underline{L}^{q_h}(B_{R_h})}^{\frac{n+1}{h+1}} + \|\nabla \xi_0\|_{\underline{L}^q(B_R)}^{n+1} + \sum_{i=1}^{n-1} \left\| \nabla \frac{w_i}{i!} \right\|_{\underline{L}^q(B_R)}^{\frac{n+1}{i}} \right),$$

Now (C.2) follows by the Caccioppoli estimate, concluding the proof. \square

Appendix D: Regularity for Constant Coefficient Linearized Equations

In this appendix we prove a lemma tracking down the regularity of a solution $(\bar{w}_1, \dots, \bar{w}_n)$ of the linearized system in the case that \bar{L} is a smooth, constant-coefficient Lagrangian. Throughout we fix $n \in \mathbb{N}_0$, $\Lambda \in [1, \infty)$, $\beta \in (0, 1)$, and assume that \bar{L} satisfies

$$I_d \leq D_p^2 \bar{L} \leq \Lambda I_d \tag{D.1}$$

and for all $M_0 \in [1, \infty)$ there is $C(M_0, \beta, d) < \infty$ such that

$$\|D^2 \bar{L}\|_{C^{n,\beta}(B_{M_0})} \leq C. \tag{D.2}$$

Lemma D.1. (Regularity of \bar{w}_m). *Let $\eta \in [\frac{1}{2}, 1)$, $M \in [1, \infty)$ and $R \in (0, \infty)$. Assume that \bar{L} satisfies (D.1) and (D.2). Let $\bar{u}, \bar{w}_1, \dots, \bar{w}_n$ solve the equations, for $m \in \{1, \dots, n+1\}$,*

$$\begin{cases} -\nabla \cdot (D_p \bar{L}(\nabla \bar{u})) = 0 & \text{in } B_R, \\ -\nabla \cdot (D_p^2 \bar{L}(\nabla \bar{u}) \nabla \bar{w}_m) = \nabla \cdot (\bar{\mathbf{F}}_m(\nabla \bar{u}, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1})) & \text{in } B_R, \end{cases} \tag{D.3}$$

where $\bar{\mathbf{F}}_m$ has been defined in (1.16) and \bar{u} satisfies

$$\frac{1}{R} \|\bar{u} - (\bar{u})_{B_R}\|_{\underline{L}^2(B_R)} \leq M. \tag{D.4}$$

Then, for $m \in \{1, \dots, n+1\}$, there exists a constant $C(m, \eta, M, \beta, d, \Lambda) < \infty$ such that

$$\|\nabla \bar{w}_m\|_{L^\infty(B_{\eta R})} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \tag{D.5}$$

Moreover, letting

$$\delta \in (0, \infty) \cap \left[\left(\frac{1}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{\underline{L}^2(B_R)} \right)^\beta, \infty \right), \tag{D.6}$$

we have, for $m \in \{1, \dots, n + 1\}$, that

$$\begin{aligned} & R^\beta [\nabla \bar{w}_m]_{C^{0,\beta}(B_{\eta R})} \\ & \leq C\delta \sum_{i=1}^m \left(\frac{1}{\delta R} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_i - \ell\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} + C\delta \sum_{i=1}^{m-1} \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \\ & \quad + C \left(\frac{1}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{\underline{L}^2(B_R)} \right)^\beta \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \end{aligned} \tag{D.7}$$

and, for $m \in \{1, \dots, n\}$ and $k \in \{1, \dots, n + 1 - m\}$,

$$\begin{aligned} & R^k \|\nabla^{k+1} \bar{w}_m\|_{L^\infty(B_{\eta R})} + R^{k+\beta} [\nabla^{k+1} \bar{w}_m]_{C^{0,\beta}(B_{\eta R})} \\ & \leq C\delta \sum_{i=1}^m \left(\frac{1}{\delta R} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_i - \ell\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} + C\delta \sum_{i=1}^{m-1} \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \\ & \quad + C \left(\frac{1}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{\underline{L}^2(B_R)} \right)^\beta \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \end{aligned} \tag{D.8}$$

Notice that by (D.4) we may always take $\delta = M^\beta$ in (D.6). When applying the result in practice, we typically take δ to be very small.

Proof. Fix $m \in \{1, \dots, n + 1\}$, $\eta \in [\frac{1}{2}, 1)$, $M \in [1, \infty)$. Let $\bar{u}, \bar{w}_1, \dots, \bar{w}_n$ solve (D.3) and assume (D.4). Fix also δ as in (D.6).

Throughout the proof we denote, for $\theta \in (0, \infty)$,

$$\begin{aligned} \mathbb{E}_m^{(\theta)} & := \theta \sum_{i=1}^m \left(\frac{1}{\theta R} \inf_{\ell \in \mathcal{P}_1} \|\bar{w}_i - \ell\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} + \theta \sum_{i=1}^{m-1} \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \\ & \quad + \left(\frac{1}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{\underline{L}^2(B_R)} \right)^\beta \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \end{aligned}$$

We also denote

$$\bar{\mathbf{f}}_m := \bar{\mathbf{F}}_m(\nabla \bar{u}, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1})$$

and, for $j \in \mathbb{N}$,

$$R_j := (\eta + 2^{-j}(1 - \eta))R \quad \text{and} \quad r_j := 2^{-j-8}(1 - \eta)R.$$

Below we denote by C a constant depending only on parameters $(m, \eta, M, \beta, d, \Lambda)$. It may change from line to line.

Step 1. Basic properties of \bar{u} . In view of [1, Proposition A.1], assumption (D.1) and normalization (D.4) imply that there exists a constant $M_0(\eta, M, d, \Lambda) < \infty$ such that

$$\|\nabla \bar{u}\|_{L^\infty(B_{\frac{1}{3}(2+\eta)R})} \leq M_0. \tag{D.9}$$

Therefore, (D.2) is applicable in $B_{\frac{1}{3}(2+\eta)R}$, and we obtain by [1, Proposition A.1] that

$$R^2 \left\| \nabla^2 \bar{u} \right\|_{L^\infty \left(B_{\frac{1}{2}(1+\eta)R} \right)} \leq C \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{L^2(B_R)}. \tag{D.10}$$

We also define

$$\mathbf{b}(x) := D_p^2 \bar{L}(\nabla \bar{u}). \tag{D.11}$$

We have by (D.10) and (D.4) that

$$I_d \leq \mathbf{b}(x) \leq \Lambda I_d \quad \text{and} \quad R \|\nabla \mathbf{b}\|_{L^\infty(B_{\frac{1}{2}(1+\eta)R})} \leq \frac{C}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{L^2(B_R)} \leq C\delta^{1/\beta}. \tag{D.12}$$

Notice also that, by (D.12), we have

$$\mathbf{E}_m^{(1)} \leq C \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}. \tag{D.13}$$

Step 2. Induction assumption on degree m . We assume inductively that, for $j \in \{1, \dots, m-1\}$ there exists a constant $\mathbf{K}_j(\eta, \mathbf{M}, \mathbf{K}_0, d, \Lambda) < \infty$ such that

$$R_j^\beta [\nabla \bar{w}_j]_{C^{0,\beta}(B_{R_j})} \leq \mathbf{K}_j \mathbf{E}_j^{(\delta)} \tag{D.14}$$

and

$$\|\nabla \bar{w}_j\|_{\underline{L}^\infty(B_{R_j})} \leq \mathbf{K}_j \sum_{i=1}^j \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{j}{i}}. \tag{D.15}$$

Notice that the case $m = 2$ has been established in [1, Proposition A.1].

Throughout the next steps of the proof, we let constants C depend on parameters $(\{\mathbf{K}_i\}_{i=1}^{m-1}, m, \eta, \mathbf{M}, \beta, \mathbf{K}_0, d, \Lambda)$ and they may change from line to line.

Step 3. Bounds on \mathbf{f} . We show that under induction assumptions (D.14) and (D.15), we have that

$$\|\bar{\mathbf{f}}_m\|_{L^\infty(B_{R_{m-1}})} \leq C \sum_{i=1}^{m-1} \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}} \tag{D.16}$$

and, for $r \in (0, r_m]$ and $y \in B_{R_m}$, defining

$$\bar{\mathbf{f}}_{m,y,r} := \bar{\mathbf{F}}_m \left((\nabla \bar{u})_{B_r(y)}, (\nabla \bar{w}_1)_{B_r(y)}, \dots, (\nabla \bar{w}_{m-1})_{B_r(y)} \right), \tag{D.17}$$

we have that

$$\|\bar{\mathbf{f}}_m - \bar{\mathbf{f}}_{m,y,r}\|_{L^\infty(B_r(y))} \leq C \left(\frac{r}{R} \right)^\beta \mathbf{E}_m^{(\delta)}. \tag{D.18}$$

To show (D.16), we have by (D.15) that

$$\begin{aligned} \sum_{i=1}^{m-1} \|\nabla \bar{w}_i\|_{\underline{L}^\infty(B_{R_{m-1}})}^{\frac{m}{i}} &\leq C \sum_{i=1}^{m-1} \mathbf{K}_i^m \sum_{j=1}^i \left(\frac{1}{R} \|\bar{w}_j - (\bar{w}_j)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{j}} \\ &\leq C \sum_{i=1}^{m-1} \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}, \end{aligned} \tag{D.19}$$

which yields (D.16) by (2.7).

To show (D.18), using Hölder regularity of $\bar{\mathbf{f}}_m$ with respect to $\nabla \bar{u}$ variable, similarly to (2.11), gives us

$$|\bar{\mathbf{f}}_m - \bar{\mathbf{f}}_{m,r,y}| \leq C \left(r^\beta \|\nabla^2 \bar{u}\|_{L^\infty(B_r(y))}^\beta + \delta \left(\frac{r}{R}\right)^\beta \right) \sum_{i=1}^{m-1} (|\nabla \bar{w}_j| + |(\nabla \bar{w}_j)_{B_r(y)}|)^\frac{m}{i} + C \delta \left(\frac{r}{R}\right)^\beta \sum_{i=1}^{m-1} \left(\delta^{-1} \left(\frac{r}{R}\right)^{-\beta} |\nabla \bar{w}_i - (\nabla \bar{w}_i)_{B_r(y)}| \right)^\frac{m}{i}.$$

Applying (D.14) and (D.6) yields that

$$\delta \left(\delta^{-1} \left(\frac{r}{R}\right)^{-\beta} |\nabla \bar{w}_i - (\nabla \bar{w}_i)_{B_r(y)}| \right)^\frac{m}{i} \leq C \delta \left(\delta^{-1} \mathbf{E}_i^{(\delta)} \right)^\frac{m}{i} \leq C \mathbf{E}_m^{(\delta)}.$$

Thus (D.18) follows by (D.12) and (D.19).

Step 4. Caccioppoli estimate. We show that under induction assumptions (D.14) and (D.15), we have that, for all $y \in B_{R_m+r_m}$ and $r \in (0, r_m]$,

$$\|\nabla \bar{w}_m - \nabla \ell\|_{\underline{L}^2(B_{r/2}(y))} \leq \frac{C}{r} \|\bar{w}_m - \ell\|_{\underline{L}^2(B_r(y))} + Cr \|\nabla^2 \bar{u}\|_{L^\infty(B_r(y))} (|\nabla \ell| \wedge \|\nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))}) + C \left(\frac{r}{R}\right)^\beta \mathbf{E}_m^{(\delta)}. \tag{D.20}$$

Since \bar{w}_m solves the equations, for $\mathbf{f}_{m,y,r}$ defined in (D.17) and any affine function ℓ ,

$$-\nabla \cdot (\mathbf{b}(\nabla(\bar{w}_m - \ell))) = \nabla \cdot ((\mathbf{b} - \mathbf{b}(y))\nabla \ell + \mathbf{f}_m - \mathbf{f}_{m,y,r}),$$

and

$$-\nabla \cdot (\mathbf{b}(y)\nabla(\bar{w}_m - \ell)) = \nabla \cdot ((\mathbf{b} - \mathbf{b}(y))\nabla \bar{w}_m + \mathbf{f}_m - \mathbf{f}_{m,y,r}),$$

we obtain (D.20) simply by testing and (D.18).

Step 5. Induction assumption on the scale. We now assume that we find $\varepsilon \in (0, 1]$, a constant C_ε , and $r^* \in (0, \varepsilon r_m]$ such that

$$\sup_{y \in B_{R_m}} \sup_{t \in [r^*, \varepsilon r_m]} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_t(y))} \leq C_\varepsilon \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^\frac{m}{i}. \tag{D.21}$$

Notice that, by the Caccioppoli estimate (D.20) and (D.13), we have, for any $\varepsilon \in (0, 1]$, that

$$\|\nabla \bar{w}_m\|_{\underline{L}^2(B_{\varepsilon r_m}(y))} \leq C_\varepsilon \sum_{i=1}^j \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^\frac{j}{i}, \tag{D.22}$$

implying that (D.21) is valid for $r^* = \varepsilon r_m$ provided that $C_\varepsilon \geq C_\varepsilon$.

Step 6. We verify that (D.15) is true for $j = m$. This gives us also (D.5). Actually, we prove that if (D.21) is valid for some $r^* \in (0, \varepsilon r_m]$, then it remains valid for $\frac{1}{2}r^*$ instead. This proves, by induction, that we may take any $r^* \in (0, \varepsilon r_m]$ in (D.21). In particular, we obtain (D.15) for $j = m$.

Fix $y \in B_{R_m}$. Rewriting the equation of \bar{w}_m as before, for $r \in (0, r_m]$,

$$-\nabla \cdot (\mathbf{b}(y)\nabla \bar{w}_m) = \nabla \cdot ((\mathbf{b} - \mathbf{b}(y))\nabla \bar{w}_m + (\bar{\mathbf{f}}_m - \bar{\mathbf{f}}_{m,y,r})),$$

where $\mathbf{f}_{m,y,r}$ defined in (D.17), and consequently solving

$$\begin{cases} -\nabla \cdot (\mathbf{b}(y)\nabla \bar{w}_{m,y,r}) = 0 & \text{in } B_r(y), \\ \bar{w}_{m,y,r} = \bar{w}_m & \text{on } \partial B_r(y), \end{cases}$$

we obtain by testing and (D.18) that

$$\|\nabla \bar{w}_{m,y,r} - \nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))} \leq Cr \|\nabla^2 \bar{u}\|_{L^\infty(B_r(y))} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))} + C \left(\frac{r}{R}\right)^\beta \mathbf{E}_m^{(\delta)}.$$

In particular, we get by (D.21) for $r \in [r^*, \varepsilon r_m]$ that

$$\|\nabla \bar{w}_{m,y,r} - \nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))} \leq C \left(\frac{r}{R}\right)^\beta \left(\mathbf{C}_\varepsilon \varepsilon^{1-\beta} + 1\right) \mathbf{E}_m^{(\delta)}.$$

By decay estimate for harmonic functions we have for small enough $\theta(\beta, d, \Lambda) \in \left(0, \frac{1}{2}\right]$ that

$$\|\nabla \bar{w}_{m,y,r} - (\nabla \bar{w}_{m,y,r})_{B_{\theta r}(y)}\|_{\underline{L}^2(B_{\theta r}(y))} \leq \frac{1}{2} \theta^\beta \|\nabla \bar{w}_{m,y,r} - (\nabla \bar{w}_{m,y,r})_{B_r(y)}\|_{\underline{L}^2(B_r(y))}.$$

Therefore, by the triangle inequality, we get

$$\begin{aligned} & \|\nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{\theta r}(y)}\|_{\underline{L}^2(B_{\theta r}(y))} \\ & \leq \frac{1}{2} \theta^\beta \|\nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_r(y)}\|_{\underline{L}^2(B_r(y))} + C \left(\frac{r}{R}\right)^\beta \left(\mathbf{C}_\varepsilon \varepsilon^{1-\beta} + 1\right) \mathbf{E}_m^{(\delta)}. \end{aligned}$$

By an iteration argument we thus obtain that, for $r \in [\theta r^*, \varepsilon r_m]$

$$\begin{aligned} & \left(\frac{r}{\varepsilon r_m}\right)^{-\beta} \|\nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_r(y)}\|_{\underline{L}^2(B_r(y))} \\ & \leq C \|\nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{\varepsilon r_m}(y)}\|_{\underline{L}^2(B_{\varepsilon r_m}(y))} + C(\varepsilon \mathbf{C}_\varepsilon + 1) \mathbf{E}_m^{(\delta)}. \end{aligned} \tag{D.23}$$

Letting $r \in [\theta r^*, \varepsilon r_m]$ and $n \in \mathbb{N}_0$ be such that $r \in (\theta^{n+1} \varepsilon r_m, \theta^n \varepsilon r_m]$, we obtain by the triangle inequality that

$$\begin{aligned} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))} & \leq C \|\nabla \bar{w}_m\|_{\underline{L}^2(B_{\theta^n \varepsilon r_m}(y))} \\ & \leq C \left\| \nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{\theta^n \varepsilon r_m}(y)} \right\|_{\underline{L}^2(B_{\theta^n \varepsilon r_m}(y))} \\ & \quad + C |(\nabla \bar{w}_m)_{B_{\varepsilon r_m}(y)}| + C \sum_{i=1}^n \left| (\nabla \bar{w}_m)_{B_{\theta^i \varepsilon r_m}(y)} - (\nabla \bar{w}_m)_{B_{\theta^{i-1} \varepsilon r_m}(y)} \right|. \end{aligned}$$

Thus, by the previous two displays and (D.22), we obtain, for $r \in [\theta r^*, \varepsilon r_m]$, that

$$\|\nabla \bar{w}_m\|_{\underline{L}^2(B_r(y))} \leq (C_\varepsilon + C_\varepsilon \mathbf{C}_\varepsilon) \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}}.$$

We first take ε so small that $C_\varepsilon = \frac{1}{2}$ and then choose $\mathbf{C}_\varepsilon \geq 2C_\varepsilon$. All in all, we have proved that

$$\sup_{t \in [\theta r^*, \varepsilon r_m]} \|\nabla \bar{w}_m\|_{\underline{L}^2(B_t(y))} \leq C_\varepsilon \sum_{i=1}^m \left(\frac{1}{R} \|\bar{w}_i - (\bar{w}_i)_{B_R}\|_{\underline{L}^2(B_R)} \right)^{\frac{m}{i}},$$

which implies that (D.21) is valid for $\frac{1}{2}r^*$ instead of r^* , which was to be shown.

Step 7. We now prove that (D.14) is valid for $j = m$, giving also (D.7). An application of the Caccioppoli estimate (D.20), together with (D.5), which was proved in Step 6 above, we have that, by giving up volume factors,

$$\left\| \nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{\varepsilon r_m}(y)} \right\|_{\underline{L}^2(B_{\varepsilon r_m}(y))} \leq C E_m^{(\delta)}.$$

Therefore, (D.23) yields, for all $y \in B_{\eta R}$, that

$$\sup_{r \in (0, r_m)} \left(\frac{r}{R} \right)^{-\beta} \left\| \nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_r(y)} \right\|_{\underline{L}^2(B_r(y))} \leq C E_m^{(\delta)}.$$

This yields, via telescoping summation as in Step 6, that

$$\left| \nabla \bar{w}_m(y) - (\nabla \bar{w}_m)_{B_r(y)} \right| \leq C \left(\frac{r}{R} \right)^\beta E_m^{(\delta)}.$$

Thus, if, on the one hand, $r = |x - y| \in (0, r_m]$, we get by the above two displays that

$$\begin{aligned} \left| \nabla \bar{w}_m(y) - \nabla \bar{w}_m(x) \right| &\leq \left| \nabla \bar{w}_m(y) - (\nabla \bar{w}_m)_{B_r(y)} \right| + \left| \nabla \bar{w}_m(x) - (\nabla \bar{w}_m)_{B_r(x)} \right| \\ &\quad + C \left\| \nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{2r}(y)} \right\|_{\underline{L}^2(B_{4r}(y))} \\ &\leq C \left(\frac{r}{R} \right)^\beta E_m^{(\delta)}. \end{aligned}$$

If, on the other hand $|x - y| > r_m$, we get

$$\left| \nabla \bar{w}_m(y) - \nabla \bar{w}_m(x) \right| \leq C E_m^{(\delta)}$$

by noticing that

$$\left| (\nabla \bar{w}_m)_{B_{r_m}(y)} - (\nabla \bar{w}_m)_{B_{r_m}(x)} \right| \leq C \left\| \nabla \bar{w}_m - (\nabla \bar{w}_m)_{B_{R_m+r_m}(y)} \right\|_{\underline{L}^2(B_{R_m+r_m}(y))},$$

and applying once more (D.20). Thus we have proved (D.7).

Step 8. We finally sketch the proof of (D.8). Since it is very similar to the above reasoning, we will omit most of the details. We prove the statement by using induction in m and in k . First, we observe that by differentiation we see that $\partial_{x_j}^k \bar{u}$ satisfies the equation

$$-\nabla \cdot (\mathbf{b} \nabla \partial_{x_j}^k u) = \nabla \cdot \mathbf{F}_k(\nabla \bar{u}, \nabla \partial_{x_j} \bar{u}, \dots, \nabla \partial_{x_j}^{k-1} \bar{u}).$$

Thus we can apply (D.5) and (D.8) for $\bar{w}_k = \partial_{x_j}^k u$ recursively and obtain, by polarization as in Lemma 2.6, that, for every $k \in \{1, \dots, n + 1\}$ and $\eta \in \left[\frac{1}{2}, 1 \right)$, there is a constant $C(k, \eta, M, \beta, d, \Lambda)$ such that

$$R^k \left\| \nabla^{k+1} \bar{u} \right\|_{L^\infty(B_{(2+\eta)R/3})} + R^{k+\beta} \left[\nabla^{k+1} \bar{u} \right]_{C^{0,\beta}(B_{(2+\eta)R/3})} \leq C \frac{1}{R} \inf_{\ell \in \mathcal{P}_1} \|\bar{u} - \ell\|_{L^2(B_R)}. \tag{D.24}$$

Next, \bar{w}_1 solves $-\nabla \cdot (\mathbf{b} \nabla \bar{w}_1) = 0$ and $z \mapsto \mathbf{b}(z) = D_p^2 \bar{L}(\nabla u(z))$ is in $C^{n,\beta}$ by (D.24), we may differentiate the equation at most n times and obtain that $\bar{w}_1 \in C^{n+1,\beta}$ and it is also straightforward to show that \bar{w}_1 satisfies (D.8) for $m = 1$; this is just classical Schauder theory.

We then assume that (D.8) is valid for every $m \in \{1, \dots, M\}$ and $k \in \{1, \dots, n + 1 - m\}$ with some $M \in \{1, \dots, n - 1\}$. We then show that it continues to hold for $m = M + 1$ and $k \in \{1, \dots, n + M\}$ as well. Now

$$-\nabla \cdot (\mathbf{b} \nabla \bar{w}_{M+1}) = -\nabla \cdot \bar{F}_m(\nabla \bar{u}, \bar{w}_1, \dots, \bar{w}_M).$$

Recalling that $(h_1, \dots, h_{m-1}) \mapsto \overline{F}_m(\nabla \bar{u}, h_1, \dots, h_{m-1})$ is a polynomial, using (D.24) and (D.8) for $m \in \{1, \dots, M\}$ and $k \in \{1, \dots, n + 1 - m\}$, we can actually deduce that

$$R^{k+\beta} \left[\nabla^k \overline{F}_m(\nabla \bar{u}, \bar{w}_1, \dots, \bar{w}_M) \right]_{C^{0,\beta}(B_{(1+\eta)R/2})} \leq C E_m^{(\delta)}.$$

Therefore, using (D.24) once more, we can differentiate the equation of \bar{w}_{M+1} k times and then show that \bar{w}_{M+1} satisfies (D.8). The proof is complete. \square

Appendix E: C^∞ Regularity for Smooth Constant-Coefficient Lagrangians

In this section we give an alternative proof of the statement that $C^{1,1}$ regularity implies C^∞ regularity for smooth, constant-coefficient Lagrangians. Our argument is similar to the classical argument by Schauder theory, but we keep track of the linearized equations to obtain a Taylor series with an explicit representation of the Taylor polynomials in terms of the linearized equations. We note that it is relatively simple to obtain real analyticity for solutions using this argument.

Proposition E.1. Fix $\varepsilon \in (0, \frac{1}{2}]$, $M \in [0, \infty)$, $N \in \mathbb{N}$, $\sigma \in [2, \infty)$, and $R \in (0, \infty]$. Suppose that $\bar{L} \in C^{N+2,1}(\mathbb{R}^d)$ is uniformly convex, that is, for all $\zeta, \xi \in \mathbb{R}^d$,

$$|\zeta|^2 \leq D^2 \bar{L}(\xi) \zeta \cdot \zeta \leq \Lambda |\zeta|^2$$

Let $u \in H^1(B_R)$ solve $\nabla \cdot D \bar{L}(\nabla u) = 0$ such that $\|\nabla u\|_{L^2(B_R)} \leq M$. Then there exist constants $C(\bar{L}, M, N, \varepsilon, \text{data})$ and polynomials q_1, \dots, q_{N+2} such that q_{m+1} is homogeneous polynomial of degree $m + 1$ solving

$$-\nabla \cdot \left(D_p^2 \bar{L}(\nabla q_1) \nabla \frac{q_{m+1}}{m+1} \right) = \nabla \cdot \bar{F}_m \left(\nabla q_1, \nabla \frac{q_2}{2}, \dots, \nabla \frac{q_m}{m} \right) \text{ in } \mathbb{R}^d,$$

and, for all $r \in (0, \frac{R}{2}]$,

$$\left\| \nabla u - \nabla \sum_{j=1}^{N+2} \frac{q_j}{j!} \right\|_{L^p(B_r)} \leq C \left(\frac{r}{R} \right)^{N+2-\varepsilon}.$$

Proof. Without loss of generality, we may take $R = 1$. By $C^{1,1}$ -estimates, see for example [1, Proposition A.1], we have that

$$|\nabla u(0)| \leq CM \text{ and } \sup_{r \in (0, 3/4)} r^{-1} \|\nabla u - \nabla u(0)\|_{L^\infty(B_r)} \leq C. \tag{E.1}$$

We set

$$q_0 = u(0) \text{ and } q_1(x) = \nabla u(0) \cdot x.$$

Assume then inductively that, for $m \in \{1, \dots, n\}$, there exists homogeneous polynomials q_m of degree m such that, for every $\sigma \in [2, \infty)$ and $\varepsilon \in (0, \frac{1}{2}]$, there exists a constant $N_{m,\sigma}(\varepsilon, d, \Lambda)$ such that

$$|\nabla^m q_m| \leq N_{m,2}, \quad \sup_{r \in (0, \frac{2m+1}{4m})} r^{-m+\varepsilon} \left\| \nabla u - \nabla \sum_{j=1}^m \frac{q_j}{j!} \right\|_{L^\sigma(B_r)} \leq N_{m,\sigma}.$$

and that, for $m \in \{1, \dots, n - 1\}$, q_{m+1} satisfies the equation

$$-\nabla \cdot \left(D_p^2 \bar{L}(\nabla q_1) \nabla \frac{q_{m+1}}{m+1} \right) = \nabla \cdot \bar{\mathbf{F}}_m \left(\nabla q_1, \nabla \frac{q_2}{2}, \dots, \nabla \frac{q_m}{m} \right) \quad \text{in } \mathbb{R}^d,$$

Let us denote $\bar{\mathbf{a}} := D_p^2 \bar{L}(\nabla q_1)$ and, for $m \in \{1, \dots, n - 1\}$,

$$\bar{w}_m := \frac{q_{m+1}}{m+1}.$$

By homogeneity, we find a homogeneous polynomial q_{n+1} of degree $n + 1$ solving the equation

$$-\nabla \cdot \left(D_p^2 \bar{L}(\nabla q_1) \nabla \frac{q_{n+1}}{n+1} \right) = \nabla \cdot \bar{\mathbf{F}}_n \left(\nabla q_1, \nabla \frac{q_2}{2}, \dots, \nabla \frac{q_n}{n} \right) \quad \text{in } \mathbb{R}^d,$$

Notice that there is a degree of freedom in the choice of q_{n+1} . Namely, the solution is unique up to an $\bar{\mathbf{a}}$ -harmonic polynomial of degree $n + 1$. We will fix this shortly. To draw parallels between this appendix and Appendix C, we set

$$\bar{w}_n := \frac{q_{n+1}}{n+1} \quad \text{and} \quad \xi_m := u - q_1 - \sum_{k=1}^m \frac{\bar{w}_k}{k!} = u - \sum_{k=1}^{m+1} \frac{q_k}{k!}.$$

Rewrite

$$\begin{aligned} & D_p \bar{L}(\nabla u) - D_p \bar{L}(\nabla q_1) - D^2 \bar{L}(\nabla q_1) \nabla \xi_n \\ &= \sum_{m=1}^n \frac{1}{m!} \left(D^2 \bar{L}(\nabla q_1) \nabla \bar{w}_m + \bar{\mathbf{F}}_m(\nabla q_1, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1}) \right) + \bar{\mathbf{E}}_n \end{aligned}$$

where

$$\begin{aligned} \bar{\mathbf{E}}_n &:= \sum_{m=2}^n \frac{1}{m!} \left(D_p^{m+1} \bar{L}(\nabla q_1) (\nabla u - \nabla q_1)^{\otimes m} - \bar{\mathbf{F}}_m(\nabla q_1, \nabla \bar{w}_1, \dots, \nabla \bar{w}_{m-1}) \right) \\ &+ D_p \bar{L}(\nabla u) - \sum_{k=0}^{n+1} \frac{1}{k!} D_p^{k+1} \bar{L}(\nabla q_1) (\nabla u - \nabla q_1)^{\otimes k}. \end{aligned}$$

By the estimate in Appendix C, we have that

$$|\bar{\mathbf{E}}_n| \leq C \sum_{h=0}^{n-1} |\nabla \xi_h| \left(|\nabla \xi_0| + \sum_{i=1}^{n-1} \left| \nabla \frac{\bar{w}_i}{i!} \right|^{\frac{1}{i}} \right)^{n-h}.$$

Taking divergence gives us, by the equations of $u, \bar{w}_1, \dots, \bar{w}_n$, that

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \xi_n = \nabla \cdot \bar{\mathbf{E}}_n.$$

Using the induction assumption we get that

$$\|\bar{\mathbf{E}}_n\|_{\underline{L}^\sigma(B_r)} \leq Cr^{n+1-\varepsilon}.$$

Now Lemma E.2 below allows us to identify the homogeneous $\bar{\mathbf{a}}$ -harmonic polynomial part of q_{n+1} of degree $n + 1$ such that

$$\sup_{r \in \left(0, \frac{2(n+1)+1}{4(n+1)}\right)} r^{-(n+1)+\varepsilon} \left\| \nabla u - \nabla \sum_{j=1}^{n+1} \frac{q_j}{j!} \right\|_{\underline{L}^\sigma(B_r)} \leq C$$

and

$$\left| \nabla^{n+1} q_{n+1} \right| \leq C.$$

This proves the induction step, and finishes the proof. \square

Lemma E.2. *Let $n \in \mathbb{N}$ and let $\alpha \in (n, n + 1)$. Let $M \in [0, \infty)$ and $\varepsilon \in (0, 1)$. Suppose that \mathbf{A} is a constant symmetric matrix having eigenvalues on the interval $[1, \Lambda]$. There is a constant $C(n, \alpha, \varepsilon, d, \Lambda)$ such that if $\mathbf{F} \in L^p(B_1)$ and $u \in H^1(B_1)$ solve*

$$\nabla \cdot \mathbf{A} \nabla u = \nabla \cdot \mathbf{F},$$

and that $\mathbf{F} \in L^p(B_1)$ satisfies, for $r \in (0, 1)$,

$$\|\mathbf{F}\|_{\underline{L}^p(B_r)} \leq Mr^\alpha, \tag{E.2}$$

then there is \mathbf{A} -harmonic $q \in \mathcal{P}_{n+1}$ such that

$$\sup_{r \in (0, 1-\varepsilon)} r^{-\alpha} \|\nabla u - \nabla q\|_{L^p(B_r)} \leq C \left(\|\nabla u\|_{L^2(B_1)} + M \right).$$

Proof. We proceed via harmonic approximation. Let $v_r \in u + H_0^1(B_r)$ be A -harmonic. Denote

$$\overline{\mathcal{A}}_{n+1} := \{p \in \mathcal{P}_{n+1} : p \text{ is } A\text{-harmonic}\}$$

By Calderón–Zygmund estimates and (E.2),

$$\|\nabla v_r - \nabla u\|_{\underline{L}^2(B_r)} \leq CMr^\alpha.$$

Using the oscillation decay estimate

$$\inf_{\tilde{q} \in \overline{\mathcal{A}}_{n+1}} \|\nabla v_r - \nabla \tilde{q}\|_{L^\infty(B_r)} \leq C\theta^{n+1} \inf_{\tilde{q} \in \overline{\mathcal{A}}_{n+1}} \|\nabla v_r - \nabla \tilde{q}\|_{\underline{L}^2(B_r)}$$

and defining

$$D(r) := r^{-\alpha} \inf_{\tilde{q} \in \overline{\mathcal{A}}_{n+1}} \|\nabla u - \nabla \tilde{q}\|_{\underline{L}^p(B_r)},$$

we obtain by the triangle inequality, for $\theta > 0$ such that $C\theta^{n+1-\alpha} = \frac{1}{2}$, that

$$D(\theta r) \leq \frac{1}{2} D(r) + CM.$$

It follows by reabsorption that, for $r \in (0, 1)$,

$$\sup_{t \in (0, r)} D(t) \leq C(D(r) + M).$$

In particular, letting \tilde{q}_t be the minimizing element of $\overline{\mathcal{A}}_{n+1}$ in the definition of $D(t)$, we get by the triangle inequality that

$$t^{-n} \|\nabla \tilde{q}_{t/2} - \nabla \tilde{q}_t\|_{\underline{L}^p(B_r)} \leq Ct^{\alpha-n} (D(t/2) + D(t)) \leq Ct^{\alpha-n} (D(r) + M).$$

This allows us to identify $q \in \overline{\mathcal{A}}_{n+1}$ such that, for $t \in (0, r)$,

$$\sum_{j=1}^{n+1} t^{n+1-j} \left| \nabla^j \tilde{q}_t(0) - \nabla^j q(0) \right| \leq Ct^{\alpha-n} (D(1) + M),$$

and it follows that

$$\|\nabla u - \nabla q\|_{\underline{L}^p(B_r)} \leq Cr^\alpha (D(1 - \varepsilon) + M).$$

The proof is complete by an easy estimate $D(1 - \varepsilon) \leq C_\varepsilon (\|\nabla u\|_{\underline{L}^2(B_1)} + M)$. \square

References

1. Armstrong, S., Ferguson, S., Kuusi, T.: Homogenization, linearization and large-scale regularity for nonlinear elliptic equations. *Comm. Pure Appl. Math.* [arXiv:1805.00467](https://arxiv.org/abs/1805.00467) (to appear)
2. ARMSTRONG, S., KUUSI, T., MOURRAT, J.-C.: Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Commun. Math. Phys.* **347**(2), 315–361, 2016
3. ARMSTRONG, S., KUUSI, T., MOURRAT, J.-C.: The additive structure of elliptic homogenization. *Invent. Math.* **208**(3), 999–1154, 2017
4. ARMSTRONG, S., KUUSI, T., MOURRAT, J.-C.: *Quantitative Stochastic Homogenization and Large-Scale Regularity*, Vol. 352. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2019
5. ARMSTRONG, S.N., MOURRAT, J.-C.: Lipschitz regularity for elliptic equations with random coefficients. *Arch. Rational Mech. Anal.* **219**(1), 255–348, 2016
6. ARMSTRONG, S.N., SMART, C.K.: Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)* **49**(2), 423–481, 2016
7. AVELLANEDA, M., LIN, F.-H.: Compactness methods in the theory of homogenization. *Commun. Pure Appl. Math.* **40**(6), 803–847, 1987
8. AVELLANEDA, M., LIN, F.-H.: Un théorème de Liouville pour des équations elliptiques à coefficients périodiques. *C. R. Acad. Sci. Paris Sér. I Math.* **309**(5), 245–250, 1989
9. AVELLANEDA, M., LIN, F.-H.: L^p bounds on singular integrals in homogenization. *Commun. Pure Appl. Math.* **44**(8–9), 897–910, 1991
10. BELLA, P., GIUNTI, A., OTTO, F.: Effective multipoles in random media. [arXiv:1708.07672](https://arxiv.org/abs/1708.07672)
11. Dal MASO, G., MODICA, L.: Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl.* **4**(144), 347–389, 1986
12. Dal MASO, G., MODICA, L.: Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.* **368**, 28–42, 1986
13. DUERINCKX, M., GLORIA, A., OTTO, F.: The structure of fluctuations in stochastic homogenization. [arXiv:1602.01717](https://arxiv.org/abs/1602.01717)
14. FISCHER, J., NEUKAMM, S.: Optimal homogenization rates in stochastic homogenization of nonlinear uniformly elliptic equations and systems, August 2019. [arXiv:1908.02273](https://arxiv.org/abs/1908.02273)
15. FISCHER, J., OTTO, F.: A higher-order large-scale regularity theory for random elliptic operators. *Commun. Partial Differ. Equ.* **41**(7), 1108–1148, 2016
16. GLORIA, A., NEUKAMM, S., OTTO, F.: Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199**(2), 455–515, 2015
17. GLORIA, A., NEUKAMM, S., OTTO, F.: A regularity theory for random elliptic operators. [arXiv:1409.2678](https://arxiv.org/abs/1409.2678)
18. GLORIA, A., OTTO, F.: An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.* **39**(3), 779–856, 2011
19. GLORIA, A., OTTO, F.: The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations, p. 114, May 2016. [arXiv:1510.08290v3](https://arxiv.org/abs/1510.08290v3)
20. HAN, Q., LIN, F.: *Elliptic Partial Differential Equations*, Vol. 1. Courant Lecture Notes in Mathematics. American Mathematical Society, Providence, 1997
21. JIKOV, V.V., KOZLOV, S.M., OLENIK, O.A.: *Homogenization of Differential Operators and Integral Functionals*. Springer, Berlin, 1994. Translated from the Russian by G.A. Yosifian
22. MINGIONE, G.: Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.* **51**(4), 355–426, 2006
23. MOSER, J., STRUWE, M.: On a Liouville-type theorem for linear and nonlinear elliptic differential equations on a torus. *Bol. Soc. Brasil. Mat. (N.S.)* **23**(1–2), 1–20, 1992

SCOTT ARMSTRONG & SAMUEL J. FERGUSON
Courant Institute of Mathematical Sciences,
New York University,
251 Mercer St.,
New York
NY
10012 USA.
e-mail: scotta@cims.nyu.edu

SAMUEL J. FERGUSON
e-mail: sjf370@nyu.edu

and

TUOMO KUUSI
Department of Mathematics and Statistics,
University of Helsinki,
P.O. Box 68, (Gustaf Hällströmin katu 2),
00014 Helsinki
Finland.
e-mail: tuomo.kuusi@helsinki.fi

(Received October 10, 2019 / Accepted March 23, 2020)
Published online April 9, 2020
© The Author(s) (2020)