Lipschitz Regularity for Elliptic Equations with Random Coefficients

Scott N Armstrong, Jean-Christophe Mourrat

▶ **To cite this version:**

Scott N Armstrong, Jean-Christophe Mourrat. Lipschitz Regularity for Elliptic Equations with Random Coefficients. Archive for Rational Mechanics and Analysis, 2016, 219, pp.255 - 348. 10.1007/s00205-015-0908-4 . ensl-01401892

HAL Id: ensl-01401892

https://ens-lyon.hal.science/ensl-01401892

Submitted on 23 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LIPSCHITZ REGULARITY FOR ELLIPTIC EQUATIONS WITH RANDOM COEFFICIENTS

SCOTT N. ARMSTRONG AND JEAN-CHRISTOPHE MOURRAT

Abstract. We develop a higher regularity theory for general quasilinear elliptic equations and systems in divergence form with random coefficients. The main result is a large-scale L^∞-type estimate for the gradient of a solution. The estimate is proved with optimal stochastic integrability under a one-parameter family of mixing assumptions, allowing for very weak mixing with non-integrable correlations to very strong mixing (e.g., finite range of dependence). We also prove a quenched L^2 estimate for the error in homogenization of Dirichlet problems. The approach is based on subadditive arguments which rely on a variational formulation of general quasilinear divergence-form equations.

1. Introduction

1.1. **Informal summary of the main results.** We are interested in the stochastic homogenization of divergence-form elliptic equations and systems which take the general quasilinear form

$$-\nabla\cdot(\mathbf{a}(\nabla u(x),x)) = 0 \quad \text{in } U\subseteq\mathbb{R}^d. \tag{1.1}$$

The precise assumptions are stated in detail below, but let us mention here that the map $\mathbf{a}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ is uniformly monotone and Lipschitz in its first argument. In addition, $x\mapsto\mathbf{a}(\cdot,x)$ is a stationary random field satisfying a quantitative ergodic assumption.

In this paper, we prove a "quenched" Lipschitz estimate for solutions of (1.1) which is optimal in terms of the stochastic integrability of the random variables appearing in the estimates. While the results are new for linear equations and even under strongly mixing conditions (see however the discussion of [3] and [24] below), they hold for the most general nonlinear divergence-form elliptic operators and under quite weak mixing conditions. We are primarily motivated by the fact that strong gradient bounds play a central role in the quantitative theory of stochastic homogenization for elliptic equations. This is because the magnitude of the gradient of a solution controls how sensitively it depends on the coefficients – and it turns out that good estimates of the latter, combined with appropriate concentration inequalities, yield optimal quantitative bounds in homogenization. This was shown for linear equations by Gloria and Otto [25, 26, 27] and Gloria, Neukamm and Otto [22, 23]. The work

Date: November 23, 2016.

2010 *Mathematics Subject Classification.* 35B27, 60H25, 35J20, 35J60.

Key words and phrases. Stochastic homogenization, Lipschitz regularity, error estimate, monotone operator.

in this paper is the crucial step to developing an optimal quantitative theory of stochastic homogenization for general quasilinear equations and systems in divergence form, as well as for improving the stochastic integrability of the current linear theory, as we will show in a future work.

To give a better idea of the gradient estimate we prove, we recall that a solution u of the linear equation

$$-\nabla \cdot (A(x)\nabla u) = \nabla \cdot \mathbf{f}(x) \quad \text{in } B_1,$$

under the assumption that the coefficient matrix $A(x)$ and vector field $\mathbf{f}(x)$ are Hölder continuous, satisfies the following pointwise bound:

$$|\nabla u(0)|^2 \le C\left(1 + \fint_{B_1} |\nabla u(x)|^2 \, dx\right), \tag{1.2}$$

where the constant C depends in particular on the regularity of the coefficients. Of course, the assumption of Hölder continuity of A is crucial, as the best regularity for general equations with rapidly oscillating coefficients is $W^{1,2+\varepsilon}$ (Meyers' estimate, for Sobolev regularity) and $C^{0,\varepsilon}$ (De Giorgi-Nash-Moser, for Hölder regularity) for some $\varepsilon > 0$ depending on the ellipticity. Therefore the estimate (1.2) is not scale invariant, and in particular at very large scales (i.e., with the ball B_1 replaced by B_R with radius $R \gg 1$) the estimate is false, in general, even if A is smooth.

The primary purpose of this paper is to show that, nevertheless, an equation *with random coefficients* has better regularity than a general equation, and we can in fact prove that, due to statistical effects, an estimate like (1.2) holds on large scales. The main result, stated precisely below in Theorem 1.1, states roughly that, under appropriate mixing conditions on the coefficients, for any $R \gg 1$, a solution u of (1.1) in a domain $U \supseteq B_R$ satisfies

$$\fint_{B_r} |\nabla u(x)|^2 \, dx \le C\left(1 + \fint_{B_R} |\nabla u(x)|^2 \, dx\right) \quad \text{for every } r \in \left[\mathcal{X}, \frac{1}{2}R\right]. \tag{1.3}$$

The main difference between (1.2) and (1.3) is that the latter estimate holds only for balls with radii larger than a random "minimal radius" $\mathcal{X}$, while the former holds in every ball (hence pointwise). The random variable $\mathcal{X}$ is almost surely finite – but not bounded – and thus the central task is to estimate the probability that $\mathcal{X}$ is very large. That is, we would like to specify which of the stochastic moments of $\mathcal{X}$ are bounded under various quantitative "mixing" assumptions on the coefficients. In this paper, we prove (1.3), with optimal stochastic integrability estimates on $\mathcal{X}$, under a continuum of mixing conditions on the coefficients, ranging from relatively weak mixing (allowing for non-integrable correlations) to very strong assumptions. For instance, if the coefficients satisfy a finite range of dependence condition, then we have the bound

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] < \infty \quad \text{for every } s < d, \tag{1.4}$$

which is optimal in that such an estimate is certainly false for $s > d$. On the other hand, if the coefficients satisfy a weaker "α–mixing" condition with an

algebraic rate given by an exponent $\beta > 0$ (see (P3) in the next subsection for the precise statement of this assumption), then we give the (optimal) bound

$$\mathbb{E}\left[\mathcal{X}^\theta\right] < \infty \quad \text{for every } \theta < \beta. \tag{1.5}$$

We call (1.3) a "quenched" estimate because the random variable $\mathcal{X}$ does not depend on the solution u, and thus in particular on what random dependence u may (or may not) possess. We call it a "Lipschitz" estimate because R is very large and in this scaling the microscopic scale is $O(1)$, and thus we think of the left side as approximating $|\nabla u(0)|$. Notice that (1.3) implies

$$\fint_{B_1} |\nabla u(x)|^2 \, dx \le C\mathcal{X}^d \left(1 + \fint_{B_R} |\nabla u(x)|^2 \, dx\right),$$

so we may also think of $\mathcal{X}$ as a random prefactor constant in an estimate with deterministic scales. Under the additional assumption that the coefficients are smooth on the microscopic scale, the left side of the latter inequality controls $|\nabla u(0)|$ (by the nonlinear version of the Schauder estimate (1.2), in fact) and, in that case, we then obtain a true pointwise Lipschitz estimate at the origin. However, we prefer to think of control over the very small scales as a separate issue. Indeed, questions about the behavior of solutions on scales smaller than the microscopic scale are outside of the realm of homogenization, as they cannot be due to statistical effects. An estimate like (1.3) is the strongest possible estimate for the control of microscopic-scale fluctuations in terms of large-scale fluctuations for solutions of (1.1); higher regularity such as $C^{1,\alpha}$ cannot hold at large scales without some reinterpretation of the meaning of such an estimate. In fact, for equations with periodic coefficients (a special case of what we consider here), there cannot exist uniform bounds on the modulus of continuity of the gradient unless the coefficients are constant.

While we use scalar notation throughout the paper, we only use arguments which apply to systems under the assumption that the constant-coefficient homogenized system admits $C^{1,\alpha}$ estimates (which is a necessary condition, satisfied both for linear systems as well as in the nonlinear scalar case, for example). Moreover, for the error estimate presented in Theorem 3.1, which is of independent interest, this assumption is not required. Therefore, the methods in this paper are quite flexible in their application to general elliptic systems.

Finally, we point out that the arguments leading to the Lipschitz estimate (Theorem 1.1) apply with only very minor modifications to yield Lipschitz estimates up to the boundary for Dirichlet problems (even with possibly oscillating boundary data).

1.2. **Motivation and previous works.** There has been much recent interest in quantitative results for the stochastic homogenization of uniformly elliptic equations, prompted by the striking results of Gloria and Otto [25, 26]. Partially inspired by the previous works of Naddaf and Spencer [33] and Conlon and Naddaf [10], they proved optimal estimates on the typical size of the

fluctuations of approximate correctors and their energy density for linear elliptic equations under strong independence assumptions on the coefficients. Expounding on this work and some ideas developed independently in [32], these authors and Neukamm [22, 23] then proved optimal estimates and a quantitative two-scale expansion for uniformly elliptic equations under quantitative assumptions built on spectral gap-type concentration inequalities.

At the analytic heart of all of these results are estimates on large stochastic moments of the gradient of the approximate correctors as well as the gradient of the Green's functions. Later, much stronger bounds on the Green's functions were obtained by Marahrens and Otto [31] for coefficient fields satisfying a logarithmic Sobolev inequality. They also proved the first "quenched" result: a $C^{0,\alpha}$ estimate for $0 < \alpha < 1$, with bounded (stochastic) moments (see also Gloria and Marahrens [21]). These estimates, and the latter result especially, suggested the possibility of developing a large scale Lipschitz regularity theory for random elliptic equations, which could potentially not only generalize these results but substantially improve the quantitative theory of stochastic homogenization.

The first quenched Lipschitz estimate for random elliptic equations was proved recently by the first author and Smart [3]. This result was obtained for solutions of variational quasilinear equations under the assumption that the coefficients satisfy a finite range of dependence, and gave optimal estimates on the stochastic integrability of the random variable $\mathcal{X}$ in (1.3), that is, they obtained (1.4).

There were two separate ideas central to the arguments of [3], which differed substantially from those of the previous works mentioned above: the first was, following the philosophy of Avellaneda and Lin [4] in the context of periodic homogenization, to "borrow" the higher regularity of the constant-coefficient homogenized equation. The technique is to prove a Lipschitz estimate by implementing a Campanato-type $C^{1,\alpha}$ iteration and comparing, at each dyadic scale, the heterogeneous solution to the homogenized one with the same boundary values. Since the latter is more regular, the "flatness" of the solution (the normalized L^2 difference between the solution and the nearest affine function) improves at each step of the iteration by a universal factor less than one. It is necessary to keep track of the error made by substituting the solution for that of the homogenized equation at each iteration. It was shown in [3, Lemma 5.1] that one obtains a Lipschitz bound on scales larger than the minimal scale above which this error in homogenization is controlled by a fixed *algebraic* (or at least Dini) modulus. This quantitative variation of the compactness argument of [4] does not depend on the precise form of the equation (in particular, it does not require the equation to be variational) and thus in certain cases it may give more information than the compactness arguments of [4] (for example, in a recent preprint, the first author and Shen [2] used the idea to generalize the results of [4, 5] as well as those of Kenig, Lin and Shen [28, 29] to certain linear systems with almost periodic coefficients).

The key step in quantifying the stochastic moments of $\mathcal{X}$ in the Lipschitz estimate, if we follow this method, is to obtain a quantitative estimate on the

error in homogenization for the Dirichlet problem which is strong in stochastic integrability: here the estimate may be suboptimal in the typical size of the error (provided it is algebraic or Dini) in order to bound the probability of deviations as strongly as possible.

This brings us to the second central theme of [3], which is that the method of subadditivity, introduced in stochastic homogenization by Dal Maso and Modica [11, 12], is very well-suited for obtaining estimates which (although suboptimal in the typical size of the error) are strong in stochastic integrability. It is here that variational techniques play a central role, as the natural subadditive quantities for analyzing the problem arise from the variational characterization of the equation. Therefore, the results in [3] were stated for local minimizers of uniformly convex integral functionals, which is equivalent to the special case of (1.1) in which $\mathbf{a}(\cdot, x)$ is the gradient of a uniformly convex function. If the equation is linear (i.e., $\mathbf{a}(p, x) = A(x)p$ for a square matrix A with positive eigenvalues), then this reduces to the requirement that A be symmetric.

More recently, Gloria, Neukamm and Otto [24], partially inspired by [3], obtained a quenched Lipschitz estimate for linear equations and systems. In particular, they give the first such estimate for linear equations and systems with nonsymmetric coefficients, although their techniques do not seem to generalize in a straightforward way to nonlinear equations. Their arguments use a similar scheme based on a Campanato iteration and approximation of the Dirichlet problem for the heterogeneous equation by the homogenized problem. They introduce a random variable r_* which is related to the quantity we denote by $\mathcal{X}$ (and the analogous random variable in [3]) but is defined using harmonic coordinates, and is therefore perhaps more intuitive and geometric. It also allows to establish qualitative results (without imposing any mixing conditions) and gives control of higher regularity of the solutions after subtracting the corrector (i.e., the regularity of the error in the two-scale expansion).

The arguments in [24] and [3] are very different in how the integrability of $\mathcal{X}$ or r_* is obtained (that is, in how the error in homogenization of the Dirichlet problem is quantified in each step of the Campanato iteration), and those of [24] so far yield suboptimal estimates of the stochastic integrability of $\mathcal{X}$ under strong mixing conditions (for example, under independence assumptions, they get (1.4) for some $s > 0$ rather than every $s < d$).

We conclude this subsection by mentioning two more recent related works: both Ben-Artzi, Marahrens and Neukamm [8] and Bella and Otto [7] obtained moment bounds on the gradient of the approximate correctors for linear elliptic systems using concentration inequalities, and Lamacz, Neukamm and Otto [30] obtained such estimates for a percolation model using similar methods.

1.3. **An overview of the approach.** Before stating the main results, it is necessary to give a brief overview of the ideas leading to the proofs (since the latter affects the former).

First, the Campanato scheme for proving a Lipschitz estimate from [3] reduces our task to that of obtaining an estimate for the homogenization error

for the Dirichlet problem which is optimal in stochastic integrability. This estimate, stated below in Theorem 3.1, is of independent interest as it is the first quantitative result in stochastic homogenization for equations taking the general quasilinear form (1.1) and, except for the result in [3] that it generalizes, the only quantitative result in stochastic homogenization for nonlinear elliptic equations in divergence form. The proof of the error estimate is an adaption of the variational, subadditive approach of [3], although here we do not have the same variational structure. We instead start from the fact that solutions of

$$-\nabla\cdot(\mathbf{a}(\nabla u,x)) = f \quad \text{in } U \tag{1.6}$$

may be characterized as the null minimizers in $H^1(U)$ of a nonnegative functional taking the form

$$w \mapsto \mathcal{J}[w,f] := \inf\left\{ \int_U \left(F(\nabla w(x),\mathbf{g}(x),x) - \nabla w(x)\cdot\mathbf{g}(x)\right)dx : \mathbf{g}\in L^2(U;\mathbb{R}^d),\ -\nabla\cdot\mathbf{g} = f\right\}, \tag{1.7}$$

where the function $F(p,q,x)$ in the integrand is uniformly convex in (p,q). Moreover, if u is a null minimizer of $\mathcal{J}[\,\cdot\,,f]$, then the corresponding vector field $\mathbf{g}$ which attains the infimum is the flux, that is, $\mathbf{g} = \mathbf{a}(\nabla u,x)$, and the integrand vanishes identically. This variational formulation does not coincide with the classical one in the case that $\mathbf{a}$ is the gradient of a Lagrangian.

This more general variational principle first appeared in special cases in the work of Brezis and Ekeland [9] and Neyroles [35]. Later, major progress was made by Fitzpatrick [14], who showed that general maximal monotone maps admit a variational characterization. Variational formulations of monotone maps were subsequently explored in some details by Ghoussoub and his collaborators (see [16] and the references therein). In Section 2, we review the connection between variational problems of this form and general quasilinear, divergence-form elliptic equations.

We consider the functional $\mathcal{J}[\cdot,\cdot]$ to be the fundamental object of study in this paper and we seldom refer to the PDE or the coefficients $\mathbf{a}$. Thus, rather than looking to identify effective coefficients $\overline{\mathbf{a}}$ directly, we search for an effective variational problem characterized by an effective integrand $\overline{F}$. The latter, as we will see, may be defined by the ($\mathbb{P}$–almost sure) limit

$$\overline{F}(p,q) = \lim_{n\to\infty} \inf_{u\in H^1_0(\square_n),\,\mathbf{g}\in L^2_{\mathrm{sol},0}(\square_n)} \fint_{\square_n} F\left(p+\nabla u(x), q+\mathbf{g}(x), x\right)dx.$$

Here $\square_n$ is the cube centered at the origin with side length 3^n, and $L^2_{\mathrm{sol},0}(U)$ denotes the set of solenoidal (i.e., divegence-free) L^2 vector fields in U with normal component vanishing on the boundary of U.

In the context of periodic homogenization, such an approach to homogenization has already been explored by Ghoussoub, Moameni and Sáiz [17], who

gave an analogous formula for $\overline{F}$ and a qualitative proof of homogenization, as well as by Visintin [38, 39]. Since the (normalized) energy

$$\mu_0(U,p,q) := \inf\left\{ \fint_U F\left(p+\nabla u(x), q+\mathbf{g}(x), x\right)\, dx \,:\, u \in H^1_0(U),\, \mathbf{g} \in L^2_{\mathrm{sol},0}(U)\right\}$$

of the minimizing pair with given affine boundary data is a naturally monotone/subadditive quantity, the argument of Dal Maso and Modica [12] naturally extends to this setting to give a qualitative proof of homogenization in the general (stationary-ergodic) random setting as well.

Here however, our interest is in quantitative results, and for this μ_0 is only half of the picture. We also need the naturally superadditive quantity, modeled on the analogous one from [3], which we denote by

$$\mu(U,q^*,p^*) := \inf\left\{ \fint_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - q^*\cdot\nabla u(x) - p^*\cdot\mathbf{g}(x)\right)\, dx \,:\, u \in H^1(U),\, \mathbf{g} \in L^2_{\mathrm{sol}}(U)\right\}.$$

Notice that μ imposes no boundary conditions on u and $\mathbf{g}$, which is what gives it the property of superadditivity. As we will see, the large scale limit of $-\mu$ is the convex dual (Legendre-Fenchel transform) of that of μ_0, and the variables (p,q) are dual to (q^*,p^*). The fact that μ_0 and μ are monotone from different sides will allow us to "trap" the limiting behavior between these two extremes, in effect revealing the additive structure of the problem which renders a more precise quantitative analysis possible.

1.4. **Statement of hypotheses and main results.** As mentioned in the previous section, we consider the variational problem associated to (1.1) to be the fundamental object of study, and therefore we state our results in terms of null minimizers of the functional $\mathcal{J}[\,\cdot\,, f]$ rather than solutions of (1.6). The link between the variational problem and the PDE is studied in detail in Section 2, where it becomes clear that our assumptions cover the case of (1.1) in full generality.

For the rest of the paper, we assume that $d \geq 2$ and fix the parameters $K_0 \geq 0$ and $\Lambda \geq 3$. We consider coefficient fields $F = F(p,q,x)$ satisfying

$$F \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d) \tag{1.8}$$

such that $(p,q) \mapsto F(p,q,x)$ is uniformly convex, that is,

$$(p,q) \mapsto F(p,q,x) - \frac{1}{2\Lambda}\left(|p|^2+|q|^2\right) \quad \text{is convex} \tag{1.9}$$

as well as uniformly $C^{1,1}$, uniformly in x; in view of (1.9), it suffices to assume

$$(p,q) \mapsto F(p,q,x) - \frac{\Lambda}{2}\left(|p|^2+|q|^2\right) \quad \text{is concave.} \tag{1.10}$$

We require also that, for every $p, x \in \mathbb{R}^d$,

$$\inf_{q\in\mathbb{R}^d}\left(F(p,q,x) - p\cdot q\right) = 0. \tag{1.11}$$

We define the set of all coefficient fields by

$$\Omega := \{F : F \text{ satisfies (1.8), (1.9), (1.10) and (1.11)}\}. \tag{1.12}$$

We endow Ω with the translation group $\{T_y\}_{y\in\mathbb{R}^d}$, which acts on Ω via

$$(T_z F)(p,q,x) := F(p,q,x+z),$$

and the family $\{\mathcal{F}_U\}$ of σ–algebras, with $\mathcal{F}_U$ defined for each Borel $U \subseteq \mathbb{R}^d$ by

$$\mathcal{F}_U := \sigma\text{–algebra on } \Omega \text{ generated by the family of maps } F \mapsto \int_U F(p,q,x)\varphi(x)\,dx, \quad p,q\in\mathbb{R}^d,\ \varphi\in C_c^\infty(\mathbb{R}^d).$$

The largest of these σ–algebras is denoted by $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. The translation group also acts naturally on $\mathcal{F}$ via the definition

$$T_z A := \{T_z F : F \in A\}, \quad A \in \mathcal{F}.$$

Associated to each $F\in\Omega$ and bounded open $U\subseteq\mathbb{R}^d$ is the functional

$$\mathcal{J} : H^1(U)\times H^{-1}(U) \to [0,\infty),$$

defined for each $(u,u^*)\in H^1(U)\times H^{-1}(U)$ by

$$\mathcal{J}[u,u^*] := \inf\left\{ \int_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - \nabla u(x)\cdot\mathbf{g}(x)\right) dx \,:\, \mathbf{g}\in L^2(U;\mathbb{R}^d),\ -\nabla\cdot\mathbf{g} = u^* \right\}, \tag{1.13}$$

where the condition $-\nabla\cdot\mathbf{g} = u^*$ is to be understood in the sense that

$$\int_U \mathbf{g}(x)\cdot\nabla\phi(x)\,dx = \langle \phi, u^*\rangle \quad \text{for every } \phi\in H_0^1(U)$$

and $\langle\cdot,\cdot\rangle$ denotes the canonical pairing between $H_0^1(U)$ and $H^{-1}(U)$. Although $\mathcal{J}$ depends on both F and U, we keep this implicit in our notation as the identities of F and U are always clear from context.

We consider a probability measure $\mathbb{P}$ on $(\Omega,\mathcal{F})$ which is assumed to satisfy the following three conditions:

(P1) The random field F is uniformly bounded below on the support of $\mathbb{P}$, in the sense that

$$\mathbb{P}\left[\operatorname*{ess\,inf}_{x\in\mathbb{R}^d} \inf_{p,q\in\mathbb{R}^d} F(p,q,x) \ge -K_0^2\right] = 1.$$

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

$$\mathbb{P}[A] = \mathbb{P}[T_z A].$$

(P3) $\mathbb{P}$ is "α–mixing" with an algebraic decorrelation rate: there exist an exponent $\beta>0$ and a constant $C_3>0$ such that, for all Borel subsets $U,V\subseteq\mathbb{R}^d$, $A\in\mathcal{F}_U$ and $B\in\mathcal{F}_V$, we have

$$|\mathbb{P}[A\cap B] - \mathbb{P}[A]\,\mathbb{P}[B]| \le C_3\left(1+\operatorname{dist}(U,V)\right)^{-\beta}.$$

The assumption (P3) is a relatively weak mixing condition. The conclusions of the main result can be considerably strengthened, with minor modifications to the arguments, if one is willing to assume stronger mixing conditions. To make this point explicitly, we also state results under an α–mixing condition with a faster rate of decorrelation:

(P4) $\mathbb{P}$ is "α–mixing" with a stretched exponential decorrelation rate: there exist an exponent $\gamma > 0$ and a constant $C_4 > 0$ such that, for all Borel subsets $U, V \subseteq \mathbb{R}^d$, $A \in \mathcal{F}_U$ and $B \in \mathcal{F}_V$, we have

$$|\mathbb{P}[A \cap B] - \mathbb{P}[A]\,\mathbb{P}[B]| \leq C_4 \exp\left(-\operatorname{dist}(U,V)^\gamma\right).$$

Notice that a finite range of dependence hypothesis on the probability measure $\mathbb{P}$ (as assumed for instance in [3]) implies (P4) for every $\gamma < \infty$.

We emphasize that assumptions (P1)-(P3) are assumed to be in force throughout the paper. This is not the case for (P4), which is only assumed to be in force where explicitly stated.

We now present the statement of the main result.

Theorem 1.1 (Lipschitz regularity)**.** *Assume that $\mathbb{P}$ satisfies (P1), (P2), (P3). Fix $p > d$ and $\theta \in (0, \beta)$. There exist $C(d, \Lambda, \beta, C_3, p, \theta) \geq 1$ and a random variable $\mathcal{X} \geq 1$ satisfying*

$$\mathbb{E}\left[\mathcal{X}^\theta\right] \leq C \tag{1.14}$$

such that the following holds: if $R \geq 1$, $f \in L^2(B_R)$ and $u \in H^1(B_R)$ satisfy

$$\mathcal{J}[u, f] = 0, \tag{1.15}$$

and if we define

$$M := K_0 + \frac{1}{R} \inf_{a \in \mathbb{R}} \left(\fint_{B_R} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} + R \left(\fint_{B_R} |f(x)|^p \, dx \right)^{\frac{1}{p}}, \tag{1.16}$$

then we have the estimate

$$\fint_{B_r} |\nabla u(x)|^2 \, dx \leq CM^2 \quad \text{for every} \quad \mathcal{X}(1+M)^{\frac{2d}{\theta}} \leq r \leq \frac{1}{2}R. \tag{1.17}$$

Moreover, if $\mathbb{P}$ satisfies (P4), then (1.14) may be improved to the statement that, for every $s \in (0, d\gamma/(d+\gamma))$,

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] \leq C, \tag{1.18}$$

where the constant C depends additionally on (γ, C_4, s), and (1.17) to the statement that

$$\fint_{B_r} |\nabla u(x)|^2 \, dx \leq CM^2 \quad \text{for every} \quad \mathcal{X}\log(1+M) \leq r \leq \frac{1}{2}R. \tag{1.19}$$

Remark 1.2. The condition (1.15) says that u is a null minimizer (on B_R) of the functional $\mathcal{J}[\cdot, f]$. According to Proposition 2.15, this is equivalent to the statement that u is a solution of

$$-\nabla \cdot (\mathbf{a}(\nabla u(x), x)) = f(x) \quad \text{in } B_R, \tag{1.20}$$

where $\mathbf{a}$ is the uniformly monotone vector field represented by F in the sense defined in Section 2. Moreover, according to Theorem 2.9, any uniformly monotone vector field $\mathbf{a}$ admits a variational representation satisfying our assumptions, and thus there is no loss of generality in the variational formulation compared to the PDE itself. In particular, the conclusion of Theorem 1.1 applies to solutions of equations of the form (1.20) with random coefficients which are uniformly monotone, Lipschitz and bounded. (It is obvious how to translate hypotheses (P2) and (P3), since these are essentially unchanged; the proper translation of (P1) into a statement about $\mathbf{a}(p,x)$ is indicated in Lemma 2.14, below.)

Remark 1.3. The integrability of $\mathcal{X}$ given in Theorem 1.1 is optimal under each of assumptions (P3) and (P4) and for every value of the exponents β and γ. This can be verified by the construction of explicit examples. Observe that if $\mathbb{P}$ satisfies a finite range of dependence, then we may take any $\gamma < \infty$ in (P4) and we thus get (1.18) for every exponent $s < d$, which matches the result in [3]. This is certainly optimal, because the probability of obtaining any *particular* coefficient field (for the random checkerboard, say) in a ball of radius R is of order $\exp(-cR^d)$.

While we do not give further details here, we also obtain appropriate bounds on $\mathcal{X}$ under mixing conditions other than α–mixing. For instance, if $\mathbb{P}$ is assumed to satisfy a spectral gap inequality, then we obtain (1.18) for every $s < d/2$ and, under a logarithmic Sobolev inequality, for every $s < d$ (we also expect both of these to be optimal).

Remark 1.4. The parameter M in the lower bound for r in (1.17) and (1.19) may be removed in the case that F is positively homogeneous of order two (e.g., in the case that the corresponding PDE is linear). The reason it arises in the general setting has to do with the fact that one may not obtain full information regarding the random fields $\{F(p,q,\cdot)\}_{p,q\in\mathbb{R}^d}$ by looking at (p,q) lying in a bounded set, and thus the random scale above which the Lipschitz estimate holds necessarily exhibits some dependence on the size of the solution. See also Remark 3.2 below.

1.5. **Outline of the paper.** In Section 2, we review the variational characterization of general divergence-form elliptic equations. In Section 3, we explain the general scheme for obtaining the Lipschitz estimate and reduce the main result to an estimate on the error in homogenization for the Dirichlet problem. The rest of the paper is focused on obtaining the latter result, stated in Theorem 3.1. We introduce the subadditive and superadditive quantities in Section 4, and review some of their basic properties. In Section 5, we reduce Theorem 3.1 to two ingredients: one is probabilistic and gives a quantitative estimate for the convergence of the subadditive and superadditive quantities, while the other is a deterministic statement that asserts that these quantities control Dirichlet problems with arbitrary boundary data. The proofs of these results are given in Sections 6 and 7, respectively, which are the analytic core of the paper. We also discuss how to construct $\overline{\mathbf{a}}$ from $\overline{F}$ in Section 6.5. In Appendix A, we put the mixing conditions into a convenient form for their use

in Section 6. In Appendix B, we give the proofs of some regularity estimates (such as the Caccioppoli and Meyers estimates) needed at several points in the paper, which although well-known, we nevertheless could not find in the literature in the generality required here.

2. Variational formulation of divergence-form elliptic equations

At the core of the proof of Theorem 1.1 is a subadditivity argument which generalizes the one used in [3]. The purpose of this section is to set up this argument by first formulating the PDE as a variational problem. This section is thus somewhat independent of the rest of the paper, and we also prove a new result (Theorem 2.9 below), which is of independent interest. Apart from this result, the material covered here is well-known and can be found for example in [14, 16, 36, 39] and is included for the convenience of the reader.

2.1. Variational characterization of monotone maps.

The variational formulation of maximal monotone operators, which was pioneered by Fitzpatrick [14], makes it possible to interpret the quasilinear PDE (1.1) as a minimization problem for an integral functional. We now review the basics of this (relatively recently developed) theory.

In this subsection, we drop x-dependence and consider a Lipschitz, uniformly monotone vector field $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d$. That is, we assume that $\mathbf{a}$ satisfies, for some constant $\lambda \geq 1$,

$$\begin{cases} |\mathbf{a}(p_1) - \mathbf{a}(p_2)| \leq \lambda |p_1 - p_2|, \\ (\mathbf{a}(p_1) - \mathbf{a}(p_2)) \cdot (p_1 - p_2) \geq \dfrac{1}{\lambda} |p_1 - p_2|^2. \end{cases} \tag{2.1}$$

We next recall that (2.1) implies that $\mathbf{a}$ is *maximal monotone*.

Lemma 2.1. *The map $\mathbf{a}$ is maximal monotone, that is, for every $p, q \in \mathbb{R}^d$,*

$$\inf_{\xi \in \mathbb{R}^d} (\mathbf{a}(\xi) - q) \cdot (\xi - p) \geq 0 \iff q = \mathbf{a}(p). \tag{2.2}$$

Proof. Suppose the first statement in (2.2) holds. Fix $h \in \mathbb{R}^d$ and $\delta > 0$ and take $\xi = p + \delta h$ to obtain

$$\begin{aligned} 0 \leq (\mathbf{a}(\xi) - q) \cdot (\xi - p) &= (\mathbf{a}(p) - q) \cdot (\xi - p) + (\mathbf{a}(\xi) - a(p)) \cdot (\xi - p) \\ &\leq \delta (\mathbf{a}(p) - q) \cdot h + \lambda \delta^2 |h|^2. \end{aligned}$$

Dividing by δ and passing to the limit $\delta \to 0$, we obtain

$$(\mathbf{a}(p) - q) \cdot h \geq 0.$$

This holds for arbitrary $h \in \mathbb{R}^d$, thus $\mathbf{a}(p) = q$. This gives one implication of (2.2), and the other is obvious. □

Remark 2.2. An inspection of the proof of Lemma 2.1 reveals that, for $\mathbf{a}$ to be maximal monotone, it suffices that $\mathbf{a}$ be merely monotone and continuous.

According to [37, Proposition 12.54] (or the Browder-Minty theorem), $\mathbf{a}$ is surjective. It then follows that the inverse $\mathbf{a}^{-1}$ of $\mathbf{a}$ is also Lipschitz and uniformly monotone (with the same constants), that is, for every $q_1, q_2 \in \mathbb{R}^d$,

$$\begin{cases} \left|\mathbf{a}^{-1}(q_1) - \mathbf{a}^{-1}(q_2)\right| \le \lambda \left|q_1 - q_2\right|, \\ \left(\mathbf{a}^{-1}(q_1) - \mathbf{a}^{-1}(q_2)\right) \cdot (q_1 - q_2) \ge \dfrac{1}{\lambda} \left|q_1 - q_2\right|^2. \end{cases} \tag{2.3}$$

We look for a variational representation of the map $\mathbf{a}$ in the following sense.

Definition 2.3. Let $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex function. We say that F *(variationally) represents* the monotone vector field $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d$ if the following two conditions hold for every $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$:

(i) $F(p, q) \ge p \cdot q$,
(ii) $F(p, q) = p \cdot q \quad \Leftrightarrow \quad q = \mathbf{a}(p)$.

Representatives of maximal monotone vector fields are not unique. It is sometimes desirable to find a representative that has the following additional property.

Definition 2.4. Let $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex function, and let

$$F^*(q^*, p^*) = \sup_{p,q \in \mathbb{R}^d} \left\{(p, q) \cdot (q^*, p^*) - F(p, q)\right\} \tag{2.4}$$

be the Legendre-Fenchel transform of F (where we write $(p, q) \cdot (q^*, p^*)$ for $p \cdot q^* + q \cdot p^*$). We say that F is *self-dual* if $F^*(q, p) = F(p, q)$ for every $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$.

We continue with some examples of representatives of monotone vector fields.

Example 2.5 (Gradients of convex functions)**.** Consider a uniformly convex function $L : \mathbb{R}^d \to \mathbb{R}$ satisfying

$$\frac{1}{2\Lambda}|p_1 - p_2|^2 \le \frac{1}{2}L(p_1) + \frac{1}{2}L(p_2) - L\left(\frac{1}{2}p_1 + \frac{1}{2}p_2\right) \le \frac{\Lambda}{2}|p_1 - p_2|^2. \tag{2.5}$$

Then the uniformly monotone map given by

$$\mathbf{a}(p) = \nabla L(p)$$

is represented by the self-dual function

$$F(p, q) = L(p) + L^*(q),$$

where L^* is the Legendre-Fenchel transform of L. Moreover, property (2.5) ensures that F is uniformly convex and $C^{1,1}$.

Example 2.6 (Linear modifications [9, 34])**.** Suppose $\mathbf{a}_0$ is a monotone map represented by F_0 and let M be a $d \times d$ matrix such that $p \cdot Mp \ge 0$ for every $p \in \mathbb{R}^d$. Then the monotone map

$$\mathbf{a}(p) = \mathbf{a}_0(p) + Mp$$

is represented by the function

$$F(p, q) = F_0(p, q - Mp) + p \cdot Mp.$$

Indeed, the fact that F is convex is clear. We also have

$$F(p,q) \geq p \cdot (q - Mp) + p \cdot Mp = p \cdot q,$$

with equality if and only if $q - Mp = \mathbf{a}_0(p)$, that is, $q = \mathbf{a}(p)$. Besides, if F_0 is uniformly convex and $C^{1,1}$, then so is F. In particular, if

$$\mathbf{a}(p) = \nabla L(p) + Mp$$

with M skew-symmetric, then $\mathbf{a}$ can be represented by the self-dual function

$$F(p,q) = L(p) + L^*(q - Mp),$$

and moreover, F is uniformly convex and $C^{1,1}$ if L satisfies (2.5).

A fortiori, if A is a symmetric positive definite matrix and M is skew-symmetric, then the monotone map

$$\mathbf{a}(p) = (A + M)p$$

is represented by

$$F(p,q) = \frac{1}{2}\, p \cdot Ap + \frac{1}{2}\, (q - Mp) \cdot A^{-1}(q - Mp). \tag{2.6}$$

Those readers interested only in the case that (1.1) is linear may skip the rest of this subsection and take (2.6) to be the definition of F in the rest of the paper.

Remark 2.7. For $G : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, denote $G^{\mathsf{T}}(q,p) := G(p,q)$. The function G represents $\mathbf{a}$ if and only if G^{T} represents $\mathbf{a}^{-1}$.

Example 2.8 (Representing general maximal monotone maps)**.** Let $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d$ be a maximal monotone map. The *Fitzpatrick function* [14] of $\mathbf{a}$, defined by

$$F(p,q) := \sup_{\xi \in \mathbb{R}^d} \big(q \cdot \xi - \mathbf{a}(\xi) \cdot (\xi - p)\big), \tag{2.7}$$

represents $\mathbf{a}$. Indeed, the facts that F is convex and that $F(p,q) \geq p \cdot q$ for every (p,q) are clear; the condition $F(p,q) \leq p \cdot q$ is equivalent to

$$\inf_{\xi \in \mathbb{R}^d} (\mathbf{a}(\xi) - q) \cdot (\xi - p) \geq 0,$$

which is equivalent to having $q = \mathbf{a}(p)$ since $\mathbf{a}$ is maximal monotone, see (2.2). As will be clear from Step 3 of the proof of Proposition 2.10 below, the Fitzpatrick function is the smallest representative of $\mathbf{a}$.

The previous example demonstrates that an arbitrary maximal monotone map is variationally representable. However, for Lipschitz, uniformly monotone maps, it is desirable (and necessary for our purposes) to find a representative that is uniformly convex and $C^{1,1}$. Unfortunately, the Fitzpatrick function does not satisfy this property in general – in particular, it is never uniformly convex if the map is linear. This issue is resolved by the following theorem.

Theorem 2.9 (Uniformly convex and $C^{1,1}$ representative)**.** *Assume* $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d$ *satisfies* (2.1) *for some* $\lambda \geq 1$*. Then there exists a self-dual* $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *that represents* $\mathbf{a}$ *and satisfies*

$$(p,q) \mapsto F(p,q) - \frac{1}{2\Lambda}(|p|^2 + |q|^2) \quad \text{is convex,} \tag{2.8}$$

$$(p,q) \mapsto F(p,q) - \frac{\Lambda}{2}(|p|^2 + |q|^2) \quad \text{is concave,} \tag{2.9}$$

where we have set

$$\Lambda := 2\lambda + 1. \tag{2.10}$$

The proof of Theorem 2.9 uses the following result.

Proposition 2.10. *Let* $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d$ *be a Lipschitz, uniformly monotone vector field, and let* $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ *represent* $\mathbf{a}$*. The dual function* F^* *represents* $\mathbf{a}^{-1}$*.*

Proof. We decompose the proof into four steps. In the first step, we introduce a function F_∞ on $\mathbb{R}^d \times \mathbb{R}^d$ that can be thought of as an indicator function of the graph of $\mathbf{a}$ [36]. It is designed to be larger than any representative of $\mathbf{a}$. We then note that the Fitzpatrick function introduced in (2.7) is $F_\infty^{*\mathsf{T}}$. In the second step, we show that F_∞^{**} represents $\mathbf{a}$. In the third step, we show that $F_\infty^{*\mathsf{T}} \leq F \leq F_\infty$. In the last step, we deduce that $F_\infty^* \leq F^* \leq F_\infty^{**\mathsf{T}}$ and complete the proof.

Step 1. Let $F_\infty : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be the (possibly non-convex) function defined by

$$F_\infty(p,q) := \begin{cases} p \cdot q & \text{if } q = \mathbf{a}(p), \\ +\infty & \text{otherwise.} \end{cases}$$

It is clear that any F that represents $\mathbf{a}$ must satisfy $F \leq F_\infty$. Moreover,

$$F_\infty^*(q,p) = \sup_{\xi \in \mathbb{R}^d} \left(q \cdot \xi + a(\xi) \cdot p - a(\xi) \cdot \xi \right).$$

In other words, $F_\infty^{*\mathsf{T}}$ is the Fitzpatrick function introduced in (2.7). Since the Fitzpatrick function represents $\mathbf{a}$, Remark 2.7 implies that F_∞^* represents $\mathbf{a}^{-1}$.

Step 2. We now show that F_∞^{**} represents $\mathbf{a}$. The function is convex. Since $F_\infty^{**} \leq F_\infty$, it is also clear that if $q = \mathbf{a}(p)$, then $F_\infty^{**}(p,q) \leq F_\infty(p,q) = p \cdot q$. It thus suffices to demonstrate the following, for every $p, q \in \mathbb{R}^d$:

$$F_\infty^{**}(p,q) \geq p \cdot q \quad \text{and} \quad q \neq \mathbf{a}(p) \Rightarrow F_\infty^{**}(p,q) > p \cdot q. \tag{2.11}$$

Since the Fitzpatrick function $F_\infty^{*\mathsf{T}}$ represents $\mathbf{a}$, we have $F_\infty^{*\mathsf{T}} \leq F_\infty$. By duality, $F_\infty^{*\mathsf{T}} \leq F_\infty^{**}$. Using again that $F_\infty^{*\mathsf{T}}$ represents $\mathbf{a}$, we obtain (2.11).

Step 3. We now show that

$$F_\infty^{*\mathsf{T}} \leq F \leq F_\infty. \tag{2.12}$$

The second inequality is clear since F represents $\mathbf{a}$. For the first one, the convexity of F ensures that for every $p, q, \xi \in \mathbb{R}^d$ and every $\varepsilon > 0$,

$$F(p,q) - F(\xi, \mathbf{a}(\xi)) \geq \varepsilon^{-1} \left[F\big(\xi + \varepsilon(p - \xi), \mathbf{a}(\xi) + \varepsilon(q - \mathbf{a}(\xi))\big) - F(\xi, \mathbf{a}(\xi)) \right].$$

Since $F(p,q) \geq p \cdot q$ for every (p,q), the latter quantity is larger than

$$\mathbf{a}(\xi) \cdot (p - \xi) + (q - \mathbf{a}(\xi)) \cdot \xi + O(\varepsilon) \qquad (\varepsilon \to 0).$$

Sending $\varepsilon \to 0$ and recalling that $F(\xi, \mathbf{a}(\xi)) = \mathbf{a}(\xi) \cdot \xi$, we arrive at

$$F(p,q) \geq q \cdot \xi - \mathbf{a}(\xi) \cdot (\xi - p).$$

Taking the supremum over $\xi \in \mathbb{R}^d$ gives $F \geq F_\infty^{*\mathsf{T}}$.

Step 4. By duality, we deduce from (2.12) that $F_\infty^* \leq F^* \leq F_\infty^{**\mathsf{T}}$. We have seen in Steps 1 and 2 respectively that F_∞^* and $F_\infty^{**\mathsf{T}}$ represent $\mathbf{a}^{-1}$. So F^* represents $\mathbf{a}^{-1}$ as well, and the proof is complete. □

Remark 2.11. The assumption that $\mathbf{a}$ be Lipschitz and uniformly monotone is not necessary for Proposition 2.10. The result can be proved along the same lines assuming only that $\mathbf{a}$ is a (possibly set-valued) maximal monotone map.

Proof of Theorem 2.9. We decompose the proof into four steps. In the first step, we construct a preliminary representative $\widetilde{F}$ of $\mathbf{a}$ which is uniformly convex. In the second step, we use $\widetilde{F}$ to construct a representative F of $\mathbf{a}$ that is also self-dual, following an approach due to [6]. In the third and fourth steps, we check that F is uniformly convex and $C^{1,1}$.

Step 1. We construct a uniformly convex representative of $\mathbf{a}$. For convenience, denote $\tau := 1/\lambda$. In view of (2.1) and (2.3), we can write

$$\mathbf{a}(p) = \tau p + \mathbf{a}_0(p) \quad \text{and} \quad \mathbf{a}^{-1}(q) = \tau q + \mathbf{a}_1(q), \tag{2.13}$$

where $\mathbf{a}_0$ and $\mathbf{a}_1$ are maximal monotone operators (see Lemma 2.1 and Remark 2.2). Let F_0 and F_1 be the Fitzpatrick functions representing $\mathbf{a}_0$ and $\mathbf{a}_1$, respectively:

$$F_0(p,q) := \sup_{\xi \in \mathbb{R}^d} \big(q \cdot \xi - \mathbf{a}_0(\xi) \cdot (\xi - p)\big) \quad \text{and} \quad F_1(q,p) := \sup_{\xi \in \mathbb{R}^d} \big(p \cdot \xi - \mathbf{a}_1(\xi) \cdot (\xi - q)\big).$$

By Example 2.6, the functions

$$(p,q) \mapsto F_0(p, q - \tau p) + \tau |p|^2 \quad \text{and} \quad (q,p) \mapsto F_1(q, p - \tau q) + \tau |q|^2$$

represent $\mathbf{a}$ and $\mathbf{a}^{-1}$ respectively. It then follows that $\mathbf{a}$ is represented by

$$\widetilde{F}(p,q) := \frac{1}{2}\left(F_0(p, q - \tau p) + \tau |p|^2 + F_1(q, p - \tau q) + \tau |q|^2\right). \tag{2.14}$$

Observe that $\widetilde{F}$ can be written in the form

$$\widetilde{F}(p,q) = \frac{\tau}{2}(|p|^2 + |q|^2) + G(p,q), \tag{2.15}$$

with G convex.

Step 2. We define F to be the proximal average of $\widetilde{F}$ and $\widetilde{F}^*$:

$$F(p,q) = \inf \left\{ \frac{1}{2}\widetilde{F}(p_1,q_1) + \frac{1}{2}\widetilde{F}^*(q_2,p_2) + \frac{1}{8}|p_1 - p_2|^2 + \frac{1}{8}|q_1 - q_2|^2 \right\}, \tag{2.16}$$

where the infimum is taken over all $p_1, q_1, p_2, q_2 \in \mathbb{R}^d$ such that

$$(p,q) = \frac{1}{2}(p_1,q_1) + \frac{1}{2}(p_2,q_2) \tag{2.17}$$

and the Legendre-Fenchel transform $\widetilde{F}^*$ of $\widetilde{F}$ is defined in (2.4). According to [6, Theorem 2.2], the function $(p,q) \mapsto F(p,q)$ is self-dual. Using the notation of Remark 2.7, we let $H = (\widetilde{F} + \widetilde{F}^{*\mathsf{T}})/2$. It is clear from (2.16) that $F \le H$. By duality, we also have $H^{*\mathsf{T}} \le F^{*\mathsf{T}} = F$. Moreover, we learn from Proposition 2.10 (and Remark 2.7) that both H and $H^{*\mathsf{T}}$ represent $\mathbf{a}$. Since $H^{*\mathsf{T}} \le F \le H$, it thus follows that F represents $\mathbf{a}$.

Step 3. We now show that F is uniformly convex using the representation (2.15), that is,

$$\widetilde{F}(p_1,q_1) = \frac{\tau}{2}(|p_1|^2 + |q_1|^2) + G(p_1,q_1)$$

with G convex. In order to lighten notation, let us denote $\xi := (p_1 - p_2)/2$ and $\eta := (q_1 - q_2)/2$. For $p_1, p_2, q_1, q_2 \in \mathbb{R}^d$ satisfying (2.17), since $p_1 = p + \xi$ and $q_1 = q + \eta$, we have

$$\begin{aligned}
&\tau(|p_1|^2 + |q_1|^2) + \frac{1}{2}|p_1 - p_2|^2 + \frac{1}{2}|q_1 - q_2|^2 \\
&\quad = \tau\left(|p|^2 + |q|^2 + 2p\cdot\xi + 2q\cdot\eta + |\xi|^2 + |\eta|^2\right) + 2|\xi|^2 + 2|\eta|^2 \\
&\quad = \frac{2\tau}{2+\tau}\left(|p|^2 + |q|^2\right) + \underbrace{(2+\tau)\left|\xi + \frac{\tau}{2+\tau}p\right|^2 + (2+\tau)\left|\eta + \frac{\tau}{2+\tau}q\right|^2}_{=:2\widetilde{G}(p,q,\xi,\eta)}.
\end{aligned}$$

Defining $\Lambda := (2+\tau)/\tau$ (which matches (2.10) as $\tau = 1/\Lambda$), we thus obtain

$$\begin{aligned}
&F(p,q) - \frac{1}{2\Lambda}\left(|p|^2 + |q|^2\right) \\
&\quad = \inf_{\xi,\eta\in\mathbb{R}^d}\left\{\frac{1}{2}\left(G(p+\xi,q+\eta) + \widetilde{G}(p,q,\xi,\eta) + \widetilde{F}^*(q+\eta,p+\xi)\right) + \frac{1}{2}|\eta|^2\right\}.
\end{aligned} \tag{2.18}$$

The expression inside the infimum on the right side above is a convex function of (p,q,ξ,η). It follows that the infimum itself is a convex function of (p,q); see for example [37, Proposition 2.22]. This yields the uniform convexity of F or, more precisely, (2.8).

Step 4. We show that F is $C^{1,1}$ as a consequence of its self-duality and uniform convexity. In fact, this follows from the general fact that the Legendre-Fenchel transform of a uniformly convex function is $C^{1,1}$ from above. Since $F^*(q,p) = F(p,q)$, this will complete the proof. Letting $z = (p,q)$, we have shown that there exists a convex $\hat{G}$ such that

$$F(z) - \frac{1}{2\Lambda}|z|^2 = \hat{G}(z).$$

We now observe that for every $z^* \in \mathbb{R}^{2d}$,

$$\begin{aligned}
F^*(z^*) &= \sup_{z\in\mathbb{R}^{2d}}\left(z\cdot z^* - F(z)\right) \\
&= \sup_{z\in\mathbb{R}^{2d}}\left(z\cdot z^* - \frac{1}{2\Lambda}|z|^2 - \hat{G}(z)\right) \\
&= \frac{\Lambda}{2}|z^*|^2 - \inf_{z\in\mathbb{R}^{2d}}\left(\frac{1}{2\Lambda}\left|z - \Lambda z^*\right|^2 + \hat{G}(z)\right).
\end{aligned}$$

Since the function $(z, z^*) \mapsto |z - \Lambda z^*|^2 / 2\Lambda + \hat{G}(z)$ is convex, the infimum on the right side of the previous line is also convex, and we obtain (2.9). □

Remark 2.12. An alternative proof of Theorem 2.9 that does not rely on a proximal average (and thus gives a somewhat more explicit formula), but does not give a self-dual representative, is as follows. Choosing $\tau = 1/(2\Lambda)$ in (2.13), we see that $\mathbf{a}_0$ and $\mathbf{a}_1$ are uniformly convex and $C^{1,1}$. Hence, each has a uniformly convex representative by Step 1 of the above proof. We can then construct representatives F_0 and F_1 of $\mathbf{a}_0$ and $\mathbf{a}_1$ respectively that are $C^{1,1}$, by considering the dual functions. The function $\widetilde{F}$ defined in (2.14) is then uniformly convex and $C^{1,1}$, and it represents $\mathbf{a}$.

We also need a converse of Theorem 2.9, which is easier to prove.

Lemma 2.13. *Suppose that $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies (2.8) for some $\Lambda \geq 1$ and has the property that, for every $p \in \mathbb{R}^d$,*

$$\inf_{q \in \mathbb{R}^d} \left(F(p, q) - p \cdot q \right) = 0. \tag{2.19}$$

Let $\mathbf{a}(p) \in \mathbb{R}^d$ be the point at which this infimum is attained, which is necessarily unique since $q \mapsto F(p, q)$ is uniformly convex. Then $\mathbf{a}$ is variationally represented by F and satisfies (2.1) with $\lambda = 4\Lambda$.

Proof. By the definition of $\mathbf{a}$, we have that, for every $p, q \in \mathbb{R}^d$,

$$F(p, q) = p \cdot q \quad \iff \quad \mathbf{a}(p) = q.$$

We have left to show that $\mathbf{a}$ satisfies (2.1) with $\lambda = 4\Lambda$. Let $p_1, p_2 \in \mathbb{R}^d$. By (2.8), (2.19) and the definition of $\mathbf{a}$,

$$\begin{aligned}
&(\mathbf{a}(p_1) - \mathbf{a}(p_2)) \cdot (p_1 - p_2) \\
&\quad = F(p_1, \mathbf{a}(p_1)) + F(p_2, \mathbf{a}(p_2)) - p_1 \cdot \mathbf{a}(p_2) - p_2 \cdot \mathbf{a}(p_1) \\
&\quad \geq 2F\left(\frac{1}{2}(p_1 + p_2), \frac{1}{2}(\mathbf{a}(p_1) + \mathbf{a}(p_2)) \right) - p_1 \cdot \mathbf{a}(p_2) - p_2 \cdot \mathbf{a}(p_1) \\
&\qquad + \frac{1}{8\Lambda} \left(|p_1 - p_2|^2 + |\mathbf{a}(p_1) - \mathbf{a}(p_2)| \right) \\
&\quad \geq \frac{1}{2}(p_1 + p_2) \cdot (\mathbf{a}(p_1) + \mathbf{a}(p_2)) - p_1 \cdot \mathbf{a}(p_2) - p_2 \cdot \mathbf{a}(p_1) \\
&\qquad + \frac{1}{8\Lambda} \left(|p_1 - p_2|^2 + |\mathbf{a}(p_1) - \mathbf{a}(p_2)| \right) \\
&\quad = \frac{1}{2}(\mathbf{a}(p_1) - \mathbf{a}(p_2)) \cdot (p_1 - p_2) + \frac{1}{8\Lambda} \left(|p_1 - p_2|^2 + |\mathbf{a}(p_1) - \mathbf{a}(p_2)| \right).
\end{aligned}$$

Rearranging this gives

$$\frac{1}{4\Lambda} \left(|p_1 - p_2|^2 + |\mathbf{a}(p_1) - \mathbf{a}(p_2)|^2 \right) \leq (\mathbf{a}(p_1) - \mathbf{a}(p_2)) \cdot (p_1 - p_2).$$

Discarding the second term on the left side gives the first inequality of (2.1) with $\lambda = 4\Lambda$; discarding the first term on the left side and using Young's inequality gives the second inequality of (2.1) with $\lambda = 2\Lambda$. □

We conclude this subsection by stating a connection between the minimum of F and $\mathbf{a}(0)$, for F and $\mathbf{a}$ as in the statement of Theorem 2.9.

Lemma 2.14. *Suppose that F represents $\mathbf{a}$ and satisfies (2.8) and (2.9). Then there exist $C, c > 0$, depending only on Λ, such that*

$$-C\left|\mathbf{a}(0)\right|^2 \le \inf_{p,q\in\mathbb{R}^d} F(p,q) \le -c\left|\mathbf{a}(0)\right|^2. \tag{2.20}$$

Proof. Since F is uniformly elliptic, ∇F exists in the classical sense and is itself a uniformly monotone and Lipschitz map on $\mathbb{R}^d\times\mathbb{R}^d$ with constant Λ, as is its inverse. Moreover, there exists a unique point $(p_0,q_0)\in\mathbb{R}^d\times\mathbb{R}^d$ at which F attains its minimum. We will show that these facts imply the lemma.

First, using that the inverse of ∇F is Lipschitz with constant Λ, we get

$$\begin{aligned} |p_0|^2 + |\mathbf{a}(0)-q_0|^2 &= \left|(0,\mathbf{a}(0)) - (p_0,q_0)\right|^2 \\ &\le \Lambda\left|\nabla F(0,\mathbf{a}(0)) - \nabla F(p_0,q_0)\right|^2 = \Lambda\left|\mathbf{a}(0)\right|^2. \end{aligned} \tag{2.21}$$

In particular, by Young's inequality,

$$|p_0|^2 + |q_0|^2 \le C\left|\mathbf{a}(0)\right|^2,$$

and therefore

$$F(p_0,q_0) \ge p_0\cdot q_0 \ge -\frac{1}{2}\left(|p_0|^2+|q_0|^2\right) \ge -C\left|\mathbf{a}(0)\right|^2.$$

This yields the first inequality of (2.20).

Similarly to (2.21), we next use that ∇F is Lipschitz with constant Λ to get

$$\left|\mathbf{a}(0)\right|^2 \le \Lambda\left(|p_0|^2+|\mathbf{a}(0)-q_0|^2\right),$$

and thus, by the uniform convexity of F,

$$\begin{aligned} 0 = F(0,\mathbf{a}(0)) &\ge F(p_0,q_0) + \frac{1}{2\Lambda}\left(|p_0|^2+|\mathbf{a}(0)-q_0|^2\right) \\ &\ge \inf_{p,q\in\mathbb{R}^d} F(p,q) + \frac{1}{2\Lambda^2}\left|\mathbf{a}(0)\right|^2. \end{aligned}$$

Rearranging gives the second inequality of (2.20) and completes the proof of the lemma. □

2.2. **Variational formulation of quasilinear elliptic PDEs.** A uniformly elliptic operator on a bounded domain U can be thought of as a uniformly monotone mapping from $H^1(U)$ into $H^{-1}(U)$. From this point of view, it is natural to seek a variational representative in terms of a given representative of the underlying vector field $\mathbf{a}$. This is provided below in Proposition 2.15 which, together with Theorem 2.9, completes the link between general equations of the form (1.1) and null minimizers of functions of the form (1.7).

We now reintroduce x-dependence, consider $K_0\ge 0$ and $\Lambda\ge 3$ as in the hypotheses in the introduction and consider a Lebesgue measurable

$$\mathbf{a}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$$

such that for each $x\in\mathbb{R}$, $\mathbf{a}(\cdot,x)$ satisfies (2.1) for $\lambda := \frac{1}{2}(\Lambda-1)$ and

$$|\mathbf{a}(0,x)| \le K_0.$$

By Theorem 2.9 and Lemma 2.14, there exists $F \in \Omega$ such that, for each $x \in \mathbb{R}^d$,

$$F(\cdot,\cdot,x) \quad \text{variationally represents the monotone map} \quad \mathbf{a}(\cdot,x), \tag{2.22}$$

and such that for every $p, q, x \in \mathbb{R}^d$,

$$\frac{1}{2\Lambda}\left(|p|^2+|q|^2\right) - CK_0^2 \le F(p,q,x) \le \frac{\Lambda}{2}\left(|p|^2+|q|^2\right) + CK_0^2, \tag{2.23}$$

where the constant C depends only on Λ.

The next proposition asserts that the functional $\mathcal{J} + \langle\cdot,\cdot\rangle$ defined in (1.13) variationally represents the quasilinear elliptic operator $u \mapsto -\nabla\cdot\mathbf{a}(\nabla u,\cdot)$ as a monotone function from $H^1(U)$ to $H^{-1}(U)$ (compare with Definition 2.3). This result, like many of those in this section, is well-known (cf. [16]).

Proposition 2.15 (Variational formulation of (1.1))**.** *Fix $U \subseteq \mathbb{R}^d$ and $F \in \Omega$ satisfying (2.22) and (2.23), and let $\mathcal{J}$ be given by (1.13). For every $(u,u^*) \in H^1(U)\times H^{-1}(U)$,*

$$\mathcal{J}[u,u^*] \ge 0.$$

Moreover, the following statements are equivalent:

(i) *(u,u^*) is a solution of the quasilinear equation*

$$-\nabla\cdot(\mathbf{a}(\nabla u,\cdot)) = u^* \quad \text{in } U;$$

(ii) *$\mathcal{J}[u,u^*] = 0$;*

(iii) *u is the unique minimizer of the functional*

$$w \mapsto \mathcal{J}[w,u^*], \quad w \in u + H_0^1(U);$$

(iv) *u^* is the unique minimizer of the functional*

$$w^* \mapsto \mathcal{J}[u,w^*], \quad w^* \in H^{-1}(U).$$

Proof. By $F(p,q,\cdot) \ge p\cdot q$, we have, for any $u \in H^1(U)$ and $\mathbf{g} \in L^2(U;\mathbb{R}^d)$,

$$\int_U \left(F\left(\nabla u(x),\mathbf{g}(x),x\right) - \nabla u(x)\cdot\mathbf{g}(x)\right)dx \ge 0. \tag{2.24}$$

Moreover, equality holds in the previous inequality if and only if

$$F\left(\nabla u(x),\mathbf{g}(x),x\right) = \nabla u(x)\cdot\mathbf{g}(x) \quad \text{a.e. in } U,$$

which by (2.22) is equivalent to

$$\mathbf{g}(x) = \mathbf{a}\left(\nabla u(x)\right) \quad \text{a.e. in } U. \tag{2.25}$$

This yields the equivalence of statements (i) and (ii) in the proposition.

The equivalence with (iv) also follows from these facts and the observation that the minimum of the functional appearing in (iv) is zero, that is, for every $u \in H^1(U)$,

$$\inf_{u^*\in H^{-1}(U)} \mathcal{J}\left[u,u^*\right] = 0. \tag{2.26}$$

This can be seen for a given $u \in H^1(U)$ by simply taking $u^* := -\nabla\cdot(\mathbf{a}(\nabla u(x),x))$. That this u^* is unique minimizer follows from the equivalence between equality holding in (2.24) and (2.25).

It remains to establish the equivalence between (iii) and the other statements. Fix $f \in H^1(U)$ and $u^* \in H^{-1}$ and notice that, for $u \in f + H_0^1(U)$, we have

$$\mathcal{J}[u,u^*] = \inf\left\{ \int_U \left(F(\nabla u(x), \mathbf{g}(x), x) - \nabla f(x) \cdot \mathbf{g}(x)\right) dx - \langle u - f, u^* \rangle \right. \\ \left. : \mathbf{g} \in L^2(U;\mathbb{R}^d),\ -\nabla \cdot \mathbf{g} = u^* \right\}.$$

It is clear that, for fixed $f \in H^1(U)$ and $u^* \in H^{-1}(U)$, the quantity in the infimum on the right side is bounded from below (by 0), uniformly convex as well as lower semicontinuous on the linear space $(f + H_0^1(U)) \times \{\mathbf{g} \in L^2(U;\mathbb{R}^d) : -\nabla \cdot \mathbf{g} = u^*\}$. It therefore has a unique minimizer. It suffices then to show that this minimization problem is null, that is, for every $f \in H^1(U)$ and $u^* \in H^{-1}(U)$,

$$\inf_{u_0 \in H_0^1(U)} \mathcal{J}[u_0 + f, u^*] = 0. \tag{2.27}$$

This follows from the solvability of the Dirichlet problem for the PDE and the equivalence of (i) and (ii). Alternatively, we give a variational proof of (2.27) by considering the dual convex optimization problem (following, e.g., the arguments in [13, Proposition III.2.1] or [16, Proposition 6.1]).

Fix $f \in H^1(U)$ and $u^* \in H^{-1}(U)$ and define $G : H^{-1}(U) \to \mathbb{R}$ by

$$G(v^*) := \inf_{u_0 \in H_0^1(U)} \left(\mathcal{J}[u_0 + f, v^* + u^*] + \langle u_0, v^* \rangle \right).$$

We want to argue that $G(0) = 0$. It is clear that $G(0) \geq 0$, so it suffices to argue that $G(0) \leq 0$. It is easy to check that G is bounded below, convex and lower semicontinuous. The Legendre-Fenchel transform of G satisfies, for every $v \in H_0^1(U)$,

$$\begin{aligned} G^*(v) &:= \sup_{v^* \in H^{-1}(U)} \left(\langle v, v^* \rangle - G(v^*) \right) \\ &= \sup_{v^* \in H^{-1}(U)} \sup_{u_0 \in H_0^1(U)} \left(\langle v - u_0, v^* \rangle - \mathcal{J}[u_0 + f, v^* + u^*] \right) \\ &\geq \sup_{v^* \in H^{-1}(U)} \left(-\mathcal{J}[v + f, v^* + u^*] \right) \\ &= 0, \end{aligned}$$

where we used (2.26) in the last step. By duality, using the fact that G is convex and lower semicontinuous, we have $G^{**} = G$. Thus

$$G(0) = G^{**}(0) = \sup_{v \in H_0^1(U)} \left(-G^*(v) \right) \leq 0. \qquad \square$$

3. Lipschitz regularity: structure of the proof of Theorem 1.1

In this section, we present the scheme for obtaining a Lipschitz estimate from Campanato iterations. We also state the error estimate in homogenization of the Dirichlet problem and argue that these two elements imply Theorem 1.1.

3.1. **Two ingredients in the proof of Theorem 1.1.** The primary ingredient in the proof of Theorem 1.1 is the following estimate for the error in homogenization of the Dirichlet problem. Under strong mixing conditions, it is necessarily *suboptimal* in its estimate of the typical size of the error because it is *optimal* in terms of stochastic integrability. The functional $\mathcal{J}$ is defined in (1.13) and the homogenized functional $\overline{\mathcal{J}}$ referred to in the statement of the theorem is defined below in Section 4.8 (see (4.41)).

Theorem 3.1 (Quenched L^2 error estimate in homogenization)**.** *Fix a bounded Lipschitz domain $U_0 \subseteq \mathbb{R}^d$ and exponents $\delta \in (0,1]$ and $\theta \in (0,\beta)$. There exist $s_0(d,\Lambda,\beta,\delta,\theta) > 0$, $C(d,\Lambda,\beta,C_3,\delta,U_0,\theta) \geq 1$ and, for every $s \in [s_0,\infty)$, a nonnegative random variable $\mathcal{X}_s$ satisfying*

$$\mathbb{E}\left[\mathcal{X}_s^s\right] \leq C \tag{3.1}$$

such that the following holds. For $R \geq 1$, $U := RU_0$, $f \in W^{1,2+\delta}(U)$ and setting

$$M := K_0 + \left(\fint_U |\nabla f(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}}, \tag{3.2}$$

the unique functions $u, \overline{u} \in f + H_0^1(U)$ such that

$$\mathcal{J}\left[u,0\right] = \overline{\mathcal{J}}\left[\overline{u},0\right] = 0, \tag{3.3}$$

satisfy the estimate

$$R^{-2} \fint_U |u(x) - \overline{u}(x)|^2 \, dx \leq CM^2(1+M)^{2d/s} \mathcal{X}_s R^{-\theta/s}. \tag{3.4}$$

Moreover, under assumption (P4), for every $s \in (0, d\gamma/(d+\gamma))$, there exist $\alpha(d,\Lambda,\gamma,\delta,s) > 0$, $C(d,\Lambda,\delta,U_0,\gamma,C_4,s) \geq 1$ and a random variable $\mathcal{Y}_s$ satisfying

$$\mathbb{E}\left[\exp\left(\mathcal{Y}_s\right)\right] \leq C \tag{3.5}$$

such that, for u and $\overline{u}$ as above,

$$R^{-2} \fint_U |u(x) - \overline{u}(x)|^2 \, dx \leq CM^2 \left(1 + \mathcal{Y}_s R^{-s} \log(1+M)\right) R^{-\alpha}. \tag{3.6}$$

Remark 3.2. The non-quadratic dependence of the estimates (3.4) and (3.6) on the parameter M is a nonlinear phenomenon. Indeed, in the linear case (as usual we mean the PDE is linear, i.e., F is quadratic) it may be removed by homogeneity. However, in the nonlinear case, this is not an artifact of our method: thinking in terms of the PDE, it arises because it is not necessarily the case that the span of the random fields $\{\mathbf{a}(p,\cdot)\}_{p\in\mathbb{R}^d}$ lie in a finite dimensional space. Thus different p's may exhibit different (i.e., independent) randomness, and since the estimates are quenched, we are required to control them all at once, up to the fixed size M. Thus the larger M is, the larger we may expect the random part of the right sides in the estimates to be.

The following proposition encapsulates the general scheme for proving Lipschitz estimates introduced in [3]. It reduces Theorem 1.1 to Theorem 3.1 by a Campanato-type iteration, which allows us to focus most of the effort in this paper on the proof of the latter. It is almost the same as [3, Lemma 5.1], but

we have included a slight variation to allow us to handle equations with non-zero right hand sides. We also formulate the result using the L^2 norm rather than the L^∞ norm, but this makes no difference in the argument.

Before giving the statement, we briefly introduce some notation. We take $\mathcal{L}$ to be the set of affine functions on $\mathbb{R}^d$ and define, for each $\sigma \in (0, \frac{1}{2}]$ and $r > 0$, the set

$$\mathcal{A}(r,\sigma) := \left\{ v \in L^2(B_r) \,:\, \frac{1}{\sigma r} \inf_{l \in \mathcal{L}} \left(\fint_{B_{\sigma r}} |v(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right. \\ \left. \leq \frac{1}{2} \left(\frac{1}{r} \inf_{l \in \mathcal{L}} \left(\fint_{B_r} |v(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right) \right\}.$$

In words, the set $\mathcal{A}(r,\sigma)$ consists of those $L^2(B_r)$ functions u which satisfy one step of a $C^{1,\alpha}$ Campanato iteration with dilation factor σ: the flatness of u in $B_{\sigma r}$ is improved from its flatness in B_r by a factor of two.

Proposition 3.3 (Campanato scheme)**.** *For $\sigma \in (0, \frac{1}{2}]$ and $\gamma \in (0,1]$ such that $\sigma^\gamma \geq \frac{2}{3}$ and $\alpha > 0$, there exists $C(\sigma,\gamma,\alpha) \geq 1$ such that the following holds. Suppose that $R \geq 1$, $K, L \geq 0$, $r_0 \in [1, R/8]$ and $u \in L^2(B_R)$ have the property that for every $r \in [r_0, R/8]$,*

$$\inf_{v \in \mathcal{A}(r,\sigma)} \frac{1}{r} \left(\fint_{B_r} |u(x) - v(x)|^2 \, dx \right)^{\frac{1}{2}} \\ \leq r^{-\alpha} \left(K + \frac{1}{2r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{2r}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right) + L r^\gamma. \tag{3.7}$$

Then for every $r \in [r_0, R/2]$,

$$\frac{1}{r} \inf_{a \in \mathbb{R}} \left(\fint_{B_r} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \\ \leq C \left(\frac{1}{R} \inf_{a \in \mathbb{R}^d} \left(\fint_{B_R} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} + K \left(\frac{r}{R} \right)^\alpha + L R^\gamma \right). \tag{3.8}$$

Proposition 3.3 asserts that if a function $u \in L^2(B_R)$ (we are thinking of R very large) has the property that, in every ball B_r with radius r between $R/8$ and a "minimal radius" r_0, it can be well-approximated by a function in $\mathcal{A}(r,\sigma)$, then in fact u does not oscillate too much on scales larger than the minimal radius. Its proof appears in Section 3.3 below.

In order to make use of Proposition 3.3, we need to check that null minimizers of the homogenized functional belong to $\mathcal{A}(r,\sigma)$. This is handled by the following simple lemma, which is a reflection of the well-known fact that a family of scale-invariant functions satisfies a $C^{1,\alpha}$ estimate if and only if it satisfies an improvement of flatness property.

Lemma 3.4. *Suppose that $\alpha \in (0,1]$, $K \geq 0$, and $u \in C^{1,\alpha}(B_R)$ has the property that, for every $r \in (0, R/2]$,*

$$[\nabla u]_{C^{0,\alpha}(B_r)} \leq K r^{-\alpha} \left(\frac{1}{2r} \inf_{l \in \mathcal{L}} \left(\fint_{B_{2r}} |u(x) - l(x)|^2 \, dx \right)^{\frac{1}{2}} \right).$$

Then there exists $\sigma(\alpha, K) \in (0, \frac{1}{2}]$ such that $u \in \mathcal{A}(r, \sigma)$ for every $r \in (0, R/2]$.

The proof of Lemma 3.4 is also given in Section 3.3 below.

3.2. **Proof of Theorem 1.1.** In this subsection, we prove the main result of the paper by showing that the combination of Theorem 3.1 and Proposition 3.3 yields a pointwise Lipschitz estimate at large scales.

Proof of Theorem 1.1. Throughout we fix $p > d$, $\theta \in (0, \beta)$. We take $\delta(d, \Lambda) > 0$ to be the exponent δ_0 in the statement of the interior Meyers estimate given in Proposition B.4 in the appendix. For each $r \in [1, R/8]$, let $v_r \in u + H_0^1(B_{4r})$ denote the unique minimizer of $\mathcal{J}[\cdot, 0]$ belonging to $u + H_0^1(B_{4r})$, and $\overline{u}_r \in v_r + H_0^1(B_r)$ the unique minimizer of $\overline{\mathcal{J}}[\cdot, 0]$ in $v_r + H_0^1(B_r)$.

By the definition of $\overline{\mathcal{J}}$ and $\overline{\mathbf{a}}$ in Sections 4.8 and 6.5, respectively, and using Lemma 2.13 and Proposition 2.15, we see that $\overline{u}_r$ is a solution of the constant-coefficient equation

$$-\nabla \cdot \overline{\mathbf{a}} (\nabla \overline{u}_r) = 0 \quad \text{in } B_r.$$

Note that $\overline{\mathbf{a}}$ is uniformly monotone and Lipschitz with constants depending only on Λ. Applying Lemma 3.4 and [20, Theorem 8.9], we have that $\overline{u}_r \in \mathcal{A}(r, \sigma)$ for some $\sigma(d, \Lambda) \in (0, \frac{1}{2}]$. Note that while the assumptions in [20, Theorem 8.9] require that the coefficients $\overline{\mathbf{a}}$ be differentiable, this assumption is not quantitative and therefore the theorem holds for Lipschitz coefficients by approximation.

By making p smaller if necessary, we may assume without loss of generality that $\sigma^{1-d/p} \geq \frac{2}{3}$, which is convenient in view of the hypothesis of Proposition 3.3.

Throughout the argument, we let C denote a positive constant which may vary in each occurrence and depends only on $(d, \Lambda, \beta, C_3, p, \theta)$ or, in the case that (P4) holds, depends only on $(d, \Lambda, \beta, C_3, p, \gamma, C_4, s)$.

Step 1. We record some preliminary estimates involving v_r and $u - v_r$. The interior Meyers estimate (see Proposition B.4) and the Caccioppoli inequality (see Proposition B.2) imply that

$$\begin{aligned} \left(\fint_{B_r} |\nabla v_r(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}} &\leq C \left(K_0 + \left(\fint_{B_{2r}} |\nabla v_r(x)|^2 \, dx \right)^{\frac{1}{2}} \right) \\ &\leq C \left(K_0 + \frac{1}{4r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{4r}} |v_r(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right). \end{aligned}$$

We next apply Proposition B.1 in the appendix and then use Hölder's inequality to obtain

$$\fint_{B_r} |\nabla u(x) - \nabla v_r(x)|^2 \, dx \leq Cr^2 \fint_{B_r} |f(x)|^2 \, dx \leq CR^{\frac{2d}{p}} r^{2-\frac{2d}{p}} \left(\fint_{B_R} |f(x)|^p \, dx \right)^{\frac{2}{p}} \leq CM^2. \tag{3.9}$$

Then by the Poincaré inequality, we get, for every $r \leq R/8$,

$$\frac{1}{r^2} \fint_{B_r} |u(x) - v_r(x)|^2 \, dx \leq CR^{\frac{2d}{p}} r^{2-\frac{2d}{p}} \left(\fint_{B_R} |f(x)|^p \, dx \right)^{\frac{2}{p}} \leq CM^2. \tag{3.10}$$

Combining these with the triangle inequality yields that

$$\left(\fint_{B_r} |\nabla v_r(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}} \leq C_0 \left(M + \frac{1}{4r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{4r}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \right), \tag{3.11}$$

where we write $C = C_0$ for future reference.

Step 2. We plan to appeal to Theorem 3.1 in a sequence of dyadic balls. In order to prepare the ground for this, we fix $C' \geq 1$ to be selected below and observe that letting $c = 1/(2C_0) > 0$, for any $r \leq R/8$, the condition

$$\frac{1}{4r} \inf_{a \in \mathbb{R}} \left(\fint_{B_{4r}} |u(x) - a|^2 \, dx \right)^{\frac{1}{2}} \leq cC'M \tag{3.12}$$

implies by (3.11) that

$$M_r := K_0 + \left(\fint_{B_r} |\nabla v_r(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}} \leq K_0 + C_0 M (1 + cC') \leq C'M \tag{3.13}$$

provided that $C' \geq 2 + 2C_0$.

Step 3. We define the random minimal radius $\mathcal{X} \geq 1$ in the general case that only (P3) holds. Set $\theta' = \theta + \frac{1}{2}(\beta - \theta)$. With $s = s_0(d, \Lambda, \beta, \delta, \theta') \geq 1$ and $\mathcal{X}_s$ as in the statement of Theorem 3.1 with θ' in place of θ, we set

$$\mathcal{X} := \left[(C')^{2+2d/s} \mathcal{X}_s \right]^{\frac{s}{\theta}}.$$

According to the conclusions of Theorem 3.1, we then have that

$$\mathbb{E}\left[\mathcal{X}^\theta \right] \leq C (C')^{2s+2d}$$

and, for every $r \in \left[\mathcal{X}(1+M)^{\frac{2d}{\theta}}, \frac{1}{8} R \right]$,

$$\frac{1}{r^2} \fint_{B_r} |\overline{u}_r(x) - v_r(x)|^2 \, dx \leq \frac{C}{(C')^2} M_r^2 \, r^{-(\theta' - \theta)/s} \leq CM^2 r^{-\alpha}, \tag{3.14}$$

provided that r satisfies (3.12), since (3.12) implies (3.13). Note that the constant C in (3.14) does not depend on our choice of C', and that the exponent $\alpha > 0$ depends only on $(d, \Lambda, \beta, \theta)$.

Step 4. In the case that assumption (P4) holds, we define $\mathcal{X}$ differently. Here we fix an exponent $s \in (0, d\gamma/(d+\gamma))$ and allow the constant C to depend

additionally on (γ, C_4, s). Letting $\mathcal{Y}_s$ and $\alpha > 0$ be as in the second statement of Theorem 3.1 with M_0 replaced by $C'M_0$, we define

$$\mathcal{X} := \max\left\{\mathcal{Y}_s^{\frac{1}{s}}, (C')^{\frac{4}{\alpha}}\right\},$$

so that

$$\mathbb{E}\left[\exp\left(\mathcal{X}^s\right)\right] \le Ce^{(C')^{\frac{4s}{\alpha}}}(C'M_0)^{2d},$$

and for every $r \in [\mathcal{X}, \frac{1}{8}R]$,

$$\frac{1}{r^2}\fint_{B_r} |\overline{u}_r(x) - v_r(x)|^2 \, dx \le \frac{C}{(C')^2} M_r^2 r^{-\alpha/2} \le CM^2 r^{-\alpha/2},$$

provided that (3.12) holds. Up to a redefinition of $\alpha > 0$, this is (3.14), with α depending additionally on (γ, s).

Step 5. We observe that, for every $r \in \left[\mathcal{X}(1+M)^{\frac{2d}{\theta}}, \frac{1}{8}R\right]$ such that (3.12) holds,

$$\frac{1}{r}\left(\fint_{B_r} |u(x) - \overline{u}_r(x)|^2 \, dx\right)^{\frac{1}{2}} \le C_1\left[r^{-\alpha} M + R^{\frac{d}{p}} r^{1-\frac{d}{p}}\left(\fint_{B_R} |f(x)|^p \, dx\right)^{\frac{1}{p}}\right], \tag{3.15}$$

where we write $C_1 = C$ for future reference. Indeed, this follows immediately from (3.10) and (3.14).

Step 6. Let $C'' := 8^{1+d/2} + C_2(1 + 2C_1)$, where $C_2 = C(\sigma, 1 - d/p, \alpha)$ is the constant given by Proposition 3.3 (and σ was defined in Step 1, above). We show that, for every $r \in \left[\mathcal{X}(1+M)^{\frac{2d}{\theta}}, \frac{1}{2}R\right]$,

$$\frac{1}{r}\inf_{a\in\mathbb{R}^d}\left(\fint_{B_r} |u(x) - a|^2 \, dx\right)^{\frac{1}{2}} \le C''M. \tag{3.16}$$

We argue by induction. The estimate is obvious for $r \in [R/8, R/2]$. Suppose that (3.16) holds for every radius $r \in [4r_0, R/2]$, with r_0 such that $r_0 \ge \mathcal{X}(1 + M)^{\frac{2d}{\theta}}$. Then (3.12) holds for every $r \in [r_0, R/8]$, provided we choose $C' \ge C''/c$. In view of (3.15), an application of Proposition 3.3 yields that, for every $r \in [r_0, R/2]$,

$$\frac{1}{r}\inf_{a\in\mathbb{R}^d}\left(\fint_{B_r} |u(x) - a|^2 \, dx\right)^{\frac{1}{2}} \le C_2(M + 2C_1M) \le C''M.$$

This completes the proof of (3.16).

Step 7. We conclude. By the Caccioppoli inequality (Proposition B.2), (3.16) and the Hölder inequality, we have for every $r \in \left[\mathcal{X}(1+M)^{\frac{2d}{\theta}}, \frac{1}{2}R\right]$,

$$\begin{aligned}\fint_{B_r} |\nabla u(x)|^2 \, dx &\le C\left(K_0^2 + \frac{1}{r^2}\inf_{a\in\mathbb{R}} \fint_{B_{2r}} |u(x)-a|^2 \, dx + r^2 \fint_{B_{2r}} |f(x)|^2 \, dx\right) \\ &\le CM^2 + CR^{\frac{2d}{p}} r^{2-\frac{2d}{p}} \left(\fint_{B_R} |f(x)|^p \, dx\right)^{\frac{2}{p}} \\ &\le CM^2.\end{aligned}$$

This completes the proof of the theorem. □

3.3. **The Campanato $C^{1,\alpha}$–type iteration.** In this subsection we give the proofs of Proposition 3.3 and Lemma 3.4.

Proof of Proposition 3.3. Step 1. We first argue that we may assume (3.7) to hold for every $r \in [r_0, R/2]$. Indeed, assume that the proposition is proved with this stronger assumption. Then we can use the result with R replaced by $R/4$ to get (3.8) for $r \in [r_0, R/8]$, up to a redefinition of C. That (3.8) holds for $r \in [R/8, R]$ is obvious (up to a redefinition of C). We may therefore assume that (3.7) holds for every $r \in [r_0, R/2]$.

Throughout, C and c denote positive constants depending only on (α, σ, γ) which may vary in each occurrence. Define $s_0 := R$ and, for each $j \in \mathbb{N}$, set $s_j := \sigma^{j-1}R/4$. Pick $m \in \mathbb{N}$ such that $s_{m+1} < r_0/2 \le s_m$. Denote, for each $j \in \{0, \dots, m\}$,

$$F_j := \frac{1}{s_j}\inf_{l\in\mathcal{L}} \left(\fint_{B_{s_j}} |u(x)-l(x)|^2 \, dx\right)^{\frac{1}{2}},$$

$$H_j := \frac{1}{s_j}\inf_{a\in\mathbb{R}} \left(\fint_{B_{s_j}} |u(x)-a|^2 \, dx\right)^{\frac{1}{2}}.$$

Step 2. We show that for every $j \in \{1, \dots, m\}$,

$$F_{j+1} \le \frac{1}{2}F_j + Cs_j^{-\alpha}(K + H_{j-1}) + CLs_j^{\gamma}. \tag{3.17}$$

By (3.7) (and $\sigma \le 1/2$) that there exists $v \in \mathcal{A}(s_j, \sigma)$ such that

$$\frac{1}{s_j}\left(\fint_{B_{s_j}} |u(x)-v(x)|^2 \, dx\right)^{\frac{1}{2}} \le s_j^{-\alpha}\left(K + H_{j-1}\right) + Ls_j^{\gamma}. \tag{3.18}$$

By the triangle inequality,

$$F_{j+1} \le \frac{1}{s_{j+1}}\left(\fint_{B_{s_{j+1}}} |u(x)-v(x)|^2 \, dx\right)^{\frac{1}{2}} + \frac{1}{s_{j+1}}\inf_{l\in\mathcal{L}}\left(\fint_{B_{s_{j+1}}} |v(x)-l(x)|^2 \, dx\right)^{\frac{1}{2}}.$$

The first term is bounded by $\sigma^{d/2-1} \le C$ times the left side of (3.18), while since $v \in \mathcal{A}(s_j, \sigma)$, the second term is bounded by

$$\frac{1}{2}\frac{1}{s_j}\inf_{l\in\mathcal{L}}\left(\fint_{B_{s_j}} |v(x)-l(x)|^2 \, dx\right)^{\frac{1}{2}} \le \frac{1}{2}F_j + \frac{1}{2}\frac{1}{s_j}\left(\fint_{B_{s_j}} |u(x)-v(x)|^2 \, dx\right)^{\frac{1}{2}},$$

and we use (3.18) again to estimate the right-most term. This yields (3.17).

Step 3. We show that, for every $j \in \{0, \dots, m-1\}$,

$$F_{j+1} \le \frac{1}{2} F_j + C s_j^{-\alpha} \left(K + H_0 + \sum_{i=0}^{j} F_i \right) + C L s_j^{\gamma}. \tag{3.19}$$

Select $p_j \in \mathbb{R}^d$ such that

$$F_j = \frac{1}{s_j} \inf_{a \in \mathbb{R}^d} \left(\fint_{B_{s_j}} |u(x) - p_j \cdot x - a|^2 \, dx \right)^{\frac{1}{2}}.$$

The triangle inequality yields, for every $j \in \{0, \dots, m\}$,

$$F_j \le H_j \le 2|p_j| + F_j. \tag{3.20}$$

Moreover, for any $a, b \in \mathbb{R}$,

$$\begin{aligned} |p_j| &\le \frac{C}{s_j} \left(\fint_{B_{s_j}} |p_j \cdot x|^2 \, dx \right)^{\frac{1}{2}} \le \frac{C}{s_j} \left(\fint_{B_{s_j}} |p_j \cdot x + a - b|^2 \, dx \right)^{\frac{1}{2}} \\ &\le \frac{C}{s_j} \left(\fint_{B_{s_j}} |u(x) - p_j \cdot x - a|^2 \, dx \right)^{\frac{1}{2}} + \frac{C}{s_j} \left(\fint_{B_{s_j}} |u(x) - b|^2 \, dx \right)^{\frac{1}{2}}, \end{aligned}$$

so that

$$|p_j| \le C(F_j + H_j) \le C H_j. \tag{3.21}$$

Similarly, for every $j \in \{0, \dots, m-1\}$,

$$|p_{j+1} - p_j| \le \frac{C}{s_j} \left(\fint_{B_{s_j}} |(p_{j+1} - p_j) \cdot x|^2 \, dx \right)^{\frac{1}{2}} \le C(F_{j+1} + F_j). \tag{3.22}$$

By iterating (3.22) and using (3.21) with $j = 0$, we get

$$|p_j| \le H_0 + C \sum_{i=0}^{j} F_i. \tag{3.23}$$

By (3.17), (3.20) and (3.23), we obtain (3.19).

Step 4. We show that, for every $j \in \{0, \dots, m\}$,

$$F_j \le 2^{-j} F_0 + C s_j^{-\alpha} (K + H_0) + C L \left(s_j^{\gamma} + R^{\gamma} s_j^{-\alpha} \right). \tag{3.24}$$

We argue by strong induction. Fix $A, B \ge 1$ to be selected below and suppose that $k \in \{0, \dots, m-1\}$ is such that, for every $j \in \{0, \dots, k\}$,

$$F_j \le 2^{-j} F_0 + A s_j^{-\alpha} (K + H_0) + L \left(A s_j^{\gamma} + B R^{\gamma} s_j^{-\alpha} \right). \tag{3.25}$$

Using (3.19) and this assumption (and then $F_0 \le H_0$), we find that

$$\begin{aligned} F_{k+1} &\le \frac{1}{2}\left(2^{-k}F_0 + As_k^{-\alpha}(K+H_0) + L\left(As_k^{\gamma} + BR^{\gamma}s_k^{-\alpha}\right)\right) + CLs_k^{\gamma} \\ &\quad + Cs_k^{-\alpha}\left(K + H_0 + \sum_{j=0}^{k}\left(2^{-j}F_0 + As_j^{-\alpha}(K+H_0) + L\left(As_j^{\gamma} + BR^{\gamma}s_j^{-\alpha}\right)\right)\right) \\ &\le 2^{-(k+1)}F_0 + s_{k+1}^{-\alpha}(K+H_0)\left(\frac{1}{2}A + CAs_k^{-\alpha} + C\right) \\ &\quad + Ls_{k+1}^{\gamma}\left(\frac{1}{2}\sigma^{-\gamma}A + C\right) + LR^{\gamma}s_{k+1}^{-\alpha}\left(\frac{1}{2}B + CA + CBs_k^{-\alpha}\right). \end{aligned}$$

Now suppose in addition that $k \le n$ with n such that $Cs_n^{-\alpha} \le \frac{1}{4}$. Then using this and $\sigma^{\gamma} \ge \frac{2}{3}$, we may select A large enough that

$$\frac{1}{2}A + CAs_k^{-\alpha} + C \le \frac{3}{4}A + C \le A \quad \text{and} \quad \frac{1}{2}\sigma^{-\gamma}A + C \le A$$

and then select B large enough that

$$\frac{1}{2}B + CA + CBs_k^{-\alpha} \le B.$$

We obtain

$$F_{k+1} \le 2^{-(k+1)}F_0 + As_{k+1}^{-\alpha}(K+H_0) + ALs_{k+1}^{\gamma} + BLR^{\gamma}s_{k+1}^{-\alpha}.$$

By induction, we obtain (3.25) for every $j \le n$. For every $j \in \{n+1,\dots,m\}$, we have $1 \le s_j/s_m \le C$ and hence $F_j \le CF_n$. We thus obtain (3.25) for every $j \in \{0,\dots,m\}$. This completes the proof of (3.24).

Step 5. We conclude the argument. To obtain (3.8), we need to estimate H_j. According to (3.20), (3.23) and (3.24), we have

$$\begin{aligned} H_j \le F_j + 2|p_j| &\le 2H_0 + C\sum_{i=0}^{j} F_i \\ &\le 2H_0 + C\sum_{i=0}^{j}\left(2^{-i}F_0 + s_i^{-\alpha}(K+H_0) + L\left(s_i^{\gamma} + R^{\gamma}s_i^{-\alpha}\right)\right) \\ &\le 2H_0 + CF_0 + Cs_j^{-\alpha}(K+H_0) + CLR^{\gamma}\left(1 + s_j^{-\alpha}\right) \\ &\le C\left(H_0 + s_j^{-\alpha}K + LR^{\gamma}\right). \end{aligned}$$

This completes the proof of the proposition. □

Proof of Lemma 3.4. Denote $l_0(x) := u(0) + x \cdot \nabla u(0)$. Then the hypothesis of the lemma yields, for every $r \in (0, R/2]$ and $s \in (0, r/2]$,

$$\begin{aligned} \left(\fint_{B_s} |u(x) - l_0(x)|^2 \, dx\right)^{\frac{1}{2}} &\le \operatorname*{osc}_{x \in B_s} |u(x) - l_0(x)| \\ &\le s^{1+\alpha}\left[\nabla u\right]_{C^{0,\alpha}(B_s)} \\ &\le Ks^{1+\alpha}r^{-1-\beta}\inf_{l \in \mathcal{L}}\left(\fint_{B_r} |u(x) - l(x)|^2 \, dx\right)^{\frac{1}{2}}. \end{aligned}$$

Taking $s = \sigma r$, this gives

$$\frac{1}{s}\left(\fint_{B_s} |u(x) - l_0(x)|^2 \, dx\right)^{\frac{1}{2}} \le K\sigma^\alpha \left(\frac{1}{r}\inf_{l\in\mathcal{L}}\left(\fint_{B_{2r}} |u(x) - l(x)|^2 \, dx\right)^{\frac{1}{2}}\right).$$

Taking $\sigma \in (0, \frac{1}{2}]$ small enough that $K\sigma^\alpha \le \frac{1}{2}$, we obtain the lemma. □

4. The subadditive quantities

We introduce the subadditive quantity μ_0 and superadditive quantity μ which play the central role in the proof of Theorem 3.1. We review their basic properties and give some notation needed in the rest of the paper.

Throughout this section, we let $C \ge 1$ and $c \in (0, 1]$ denote positive constants which may vary in each occurrence and depend only on (d, Λ).

4.1. **Definition of the subadditive quantities.** We denote the space of L^2 solenoidal vector fields on a bounded Lipschitz domain $U \subseteq \mathbb{R}^d$ by

$$L^2_{\mathrm{sol}}(U) := \left\{\mathbf{f} \in L^2(U;\mathbb{R}^d) \,:\, \int_U \mathbf{f}(x)\cdot\nabla\phi(x)\,dx = 0 \quad \text{for every} \quad \phi \in H^1_0(U)\right\}.$$

The subspace of $L^2_{\mathrm{sol}}(U)$ consisting of solenoidal vector fields with zero normal component on ∂U is

$$L^2_{\mathrm{sol},0}(U) := \left\{\mathbf{f} \in L^2_{\mathrm{sol}}(U) \,:\, \int_U \mathbf{f}(x)\cdot\nabla\phi(x)\,dx = 0 \quad \text{for every} \quad \phi \in H^1(U)\right\}.$$

Note that $L^2_{\mathrm{sol},0}(U)$ is the L^2 closure of the vector fields in $L^2_{\mathrm{sol}}(U)$ with compact support in U.

For every $p, q, p^*, q^* \in \mathbb{R}^d$ and bounded Lipschitz domain $U \subseteq \mathbb{R}^d$, we define

$$\mu(U, q^*, p^*) := \inf_{u\in H^1(U),\, \mathbf{g}\in L^2_{\mathrm{sol}}(U)} \fint_U \left(F(\nabla u(x), \mathbf{g}(x), x) - q^*\cdot\nabla u(x) - p^*\cdot\mathbf{g}(x)\right) dx$$

and

$$\mu_0(U, p, q) := \inf_{v\in H^1_0(U),\, \mathbf{h}\in L^2_{\mathrm{sol},0}(U)} \fint_U F(p + \nabla v(x), q + \mathbf{h}(x), x)\,dx.$$

Since the admissible set for μ_0 has more constraints, we see that, for $p, q, p^*, q^* \in \mathbb{R}^d$,

$$\mu(U, q^*, p^*) \le p\cdot q^* + p^*\cdot q + \mu_0(U, p, q). \tag{4.1}$$

Up to normalization, $\mu(\cdot, q^*, p^*)$ is *superadditive* and $\mu_0(\cdot, p, q)$ is *subadditive*. Precisely, if $U, U_1, \ldots, U_k$ are bounded domains satisfying

$$U_1, \ldots, U_k \quad \text{are pairwise disjoint and} \quad |U \setminus (U_1 \cup \cdots \cup U_k)| = 0, \tag{4.2}$$

then

$$\mu(U, q^*, p^*) \ge \sum_{i=1}^k \frac{|U_i|}{|U|}\mu(U_i, q^*, p^*) \quad \text{and} \quad \mu_0(U, p, q) \le \sum_{i=1}^k \frac{|U_i|}{|U|}\mu_0(U_i, p, q). \tag{4.3}$$

This is due to the fact that an approximate minimizer for $\mu(U, q^*, p^*)$ gives, by restriction, a candidate for the minimizer of $\mu(U_i, q^*, p^*)$ for each subdomain U_i and, conversely, a minimizing candidate for $\mu_0(U, p, q)$ may be constructed by

assembling each of the minimizers of $\mu(U_i,p,q)$ as these agree on the boundaries of the subdomains (up to additive constants).

We next observe that μ and μ_0 are uniformly bounded on the support of $\mathbb{P}$: for every bounded open $U\subseteq\mathbb{R}^d$ and $p^*,q^*\in\mathbb{R}^d$, we have

$$-C\left(K_0^2+|p^*|^2+|q^*|^2\right)\le \mu(U,q^*,p^*)\le CK_0^2, \quad \mathbb{P}\text{–a.s.} \tag{4.4}$$

To see this, we first use (P1) and the proof of Lemma 2.14 to obtain that the minimum of $F(\cdot,\cdot,x)$ occurs at a point $(p_0(x),q_0(x))$ satisfying

$$|p_0(x)|^2+|q_0(x)|^2\le CK_0^2.$$

Moreover, by (1.11) and (P1),

$$-K_0^2\le F(p_0(x),q_0(x),x)\le 0,$$

so by (1.9) and (1.10),

$$\frac{1}{2\Lambda}\left(|p-p_0(x)|^2+|q-q_0(x)|^2\right)-K_0^2\le F(p,q,x)\le \frac{\Lambda}{2}\left(|p-p_0(x)|^2+|q-q_0(x)|^2\right).$$

Hence by Young's inequality, we have $\mathbb{P}$–a.s.,

$$\frac{1}{\Lambda}\left(|p|^2+|q|^2\right)-CK_0^2\le F(p,q,x)\le \Lambda\left(|p|^2+|q|^2\right)+CK_0^2. \tag{4.5}$$

To get the right inequality of (4.4), we test with the zero function in the definition of μ and use the previous inequality. To get the left inequality of (4.4), we use the previous inequality and Young's inequality to get, $\mathbb{P}$-a.s.,

$$F(p,q,x)-p\cdot q^*-q\cdot p^*\ge -C\left(K_0^2+|p^*|^2+|q^*|^2\right).$$

In particular, from (4.4) we have the bound

$$|\mu(U,q^*,p^*)|\le C\left(K_0^2+|p^*|^2+|q^*|^2\right) \quad \mathbb{P}\text{–a.s.} \tag{4.6}$$

Similarly, $\mathbb{P}$-a.s., we have the following uniform estimates on μ_0, for every bounded open $U\subseteq\mathbb{R}^d$ and $p,q\in\mathbb{R}^d$:

$$\frac{1}{\Lambda}\left(|p|^2+|q|^2\right)-CK_0^2\le \mu_0(U,p,q)\le \Lambda\left(|p|^2+|q|^2\right)+CK_0^2. \tag{4.7}$$

The right inequality was obtained by testing with the zero function and using (4.5). The left inequality follows from (4.5) and Jensen's inequality. Finally, note that

$$\mu_0(U,p,q)\ge p\cdot q. \tag{4.8}$$

Indeed, for each bounded and Lipschitz $U\subseteq\mathbb{R}^d$,

$$\begin{aligned}
\mu_0(U,p,q) &= \inf_{v\in H_0^1(U),\,\mathbf{h}\in L^2_{\mathrm{sol},0}(U)} \fint_U F(p+\nabla v(x),q+\mathbf{h}(x),x)\,dx\\
&\ge \inf_{v\in H_0^1(U),\,\mathbf{h}\in L^2_{\mathrm{sol},0}(U)} \fint_U (p+\nabla v(x))\cdot(q+\mathbf{h}(x))\,dx\\
&= p\cdot q.
\end{aligned}$$

4.2. **Notation for cubes.** Throughout the rest of the paper, we work with a triadic cube decomposition of $\mathbb{R}^d$. For each $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$, we set

$$\square_n(x) := 3^n \lfloor 3^{-n} x \rfloor + \square_n,$$

where we recall that $\square_n$ is the cube $(-3^n/2, 3^n/2)^d$. For each $n \in \mathbb{N}$, the family $\{\square_n(x) : x \in \mathbb{R}^d\} = \{z + \square_n : z \in 3^n\mathbb{Z}^d\}$ forms a disjoint partition of $\mathbb{R}^d$ (up to a Lebesgue null set). Notice in particular that the cube $\square_n(x)$ is not centered at x (unless $x \in 3^n\mathbb{Z}^d$). Rather, it is the unique cube which contains x with side length 3^n and centered at an element of $3^n\mathbb{Z}^d$.

In order to work with the uniform mixing condition (P3), it is sometimes convenient to delete a thin mesoscopic boundary strip from the cubes $\square_n(x)$ so that the cubes are separated from each other. We denote these *trimmed cubes* by

$$\boxdot_n := \left(-\frac{1}{2}\left(3^n - 3^{n/(1+\beta)}\right), \frac{1}{2}\left(3^n - 3^{n/(1+\beta)}\right)\right)^d \subseteq \mathbb{R}^d,$$

where we recall that $\beta > 0$ is the exponent appearing in condition (P3). For every $x \in \mathbb{R}^d$, we let $\boxdot_n(x) = 3^n \lfloor 3^{-n} x \rfloor + \boxdot_n$. It is clear that, for every $x, y \in \mathbb{R}^d$,

$$\boxdot_n(x) \neq \boxdot_n(y) \quad \Longrightarrow \quad \operatorname{dist}\left(\boxdot_n(x), \boxdot_n(y)\right) \geq 3^{n/(1+\beta)}. \tag{4.9}$$

The proportion of volume we have sculpted from $\square_n$ to create $\boxdot_n$ is relatively small:

$$\frac{\left|\square_n \setminus \boxdot_n\right|}{\left|\square_n\right|} \leq C 3^{-n\beta/(1+\beta)}. \tag{4.10}$$

4.3. **Definition of $\overline{\mu}$ and $\overline{\mu}_0$.** Since each triadic cube $\square_{n+1}$ is the disjoint union of 3^d subcubes of the form $z + \square_n$, with $z \in \mathbb{Z}^d$, we obtain from the super- and subadditivity properties and stationarity that $\mathbb{E}\left[\mu(\square_n, q^*, p^*)\right]$ is nondecreasing in n and $\mathbb{E}\left[\mu_0(\square_n, p, q)\right]$ is nonincreasing in n. Therefore,

$$\lim_{n\to\infty} \mathbb{E}\left[\mu(\square_n, q^*, p^*)\right] = \overline{\mu}(q^*, p^*) := \sup_{n\in\mathbb{N}} \mathbb{E}\left[\mu(\square_n, q^*, p^*)\right] \tag{4.11}$$

and

$$\lim_{n\to\infty} \mathbb{E}\left[\mu_0(\square_n, p, q)\right] = \overline{\mu}_0(p, q) := \inf_{n\in\mathbb{N}} \mathbb{E}\left[\mu_0(\square_n, p, q)\right]. \tag{4.12}$$

Note that, by (4.1), we have

$$\overline{\mu}(q^*, p^*) \leq \overline{\mu}_0(p, q) + p \cdot q^* + p^* \cdot q. \tag{4.13}$$

Next we observe that μ and μ_0 are monotonic with respect to the trimmed cubes, up to a small error:

$$\mu(\boxdot_{n+m}, q^*, p^*) \geq 3^{-dm} \sum_{\boxdot_n(x) \subseteq \boxdot_{n+m}} \mu(\boxdot_n(x), q^*, p^*) - C 3^{-n\beta/(1+\beta)} \left(K_0^2 + |p^*|^2 + |q^*|^2\right) \tag{4.14}$$

and

$$\mu_0(\boxdot_{n+m}, p, q) \leq 3^{-dm} \sum_{\boxdot_n(x) \subseteq \boxdot_{n+m}} \mu_0(\boxdot_n(x), p, q) + C 3^{-n\beta/(1+\beta)} \left(K_0^2 + |p|^2 + |q|^2\right). \tag{4.15}$$

To obtain these, we make the obvious partition of $\square_{n+m}$ into the union of the trimmed cubes $\square_n(x) \subseteq \square_{n+m}$ and the extra trimmed region and apply (4.6), (4.7) and (4.10). In particular, by taking expectations and using stationarity, we get

$$\mathbb{E}\left[\mu(\square_{n+1}, q^*, p^*)\right] \geq \mathbb{E}\left[\mu(\square_n(x), q^*, p^*)\right] - C3^{-n\beta/(1+\beta)}\left(K_0^2 + |p^*|^2 + |q^*|^2\right) \tag{4.16}$$

and

$$\mathbb{E}\left[\mu_0(\square_{n+1}, p, q)\right] \leq \mathbb{E}\left[\mu_0(\square_n(x), p, q)\right] + C3^{-n\beta/(1+\beta)}\left(K_0^2 + |p|^2 + |q|^2\right). \tag{4.17}$$

4.4. **The minimizing pairs for μ and μ_0.** Throughout the rest of this paper, we denote the minimizing pair of $\mu(U, q^*, p^*)$ by $(u, \mathbf{g}) = (u(\cdot, U, q^*, p^*), \mathbf{g}(\cdot, U, q^*, p^*))$. To be precise, $(u, \mathbf{g})$ are uniquely defined for each $p^*, q^* \in \mathbb{R}^d$ and bounded open subject $U \subseteq \mathbb{R}^d$ by the conditions that $(u, \mathbf{g}) \in H^1(U) \times L^2_{\mathrm{sol}}(U)$

$$\begin{cases} \mu(U, q^*, p^*) = \fint_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - q^* \cdot \nabla u(x) - p^* \cdot \mathbf{g}(x)\right) dx, \\ \fint_V u(x)\, dx = 0 \quad \text{for every connected component } V \text{ of } U. \end{cases}$$

The existence of the minimizing pair for μ is immediately obtained from the direct method since the uniform convexity of F guarantees that the functional is weakly lowersemicontinuous. Uniqueness is also a consequence of uniform convexity.

It is immediate from (4.5) and (4.6) that

$$\fint_U \left(|\nabla u(x, U, q^*, p^*)|^2 + |\mathbf{g}(x, U, q^*, p^*)|^2\right) dx \leq C\left(K_0^2 + |p^*|^2 + |q^*|^2\right) \quad \mathbb{P}\text{–a.s.} \tag{4.18}$$

We denote the minimizing pair of $\mu_0(U, p, q)$ by $(v, \mathbf{h})$, and it is uniquely defined by $(v, \mathbf{h}) \in H^1_0(U) \times L^2_{\mathrm{sol},0}(U)$ and

$$\mu_0(U, p, q) = \fint_U F\left(p + \nabla v(x), q + \mathbf{h}(x), x\right) dx. \tag{4.19}$$

We sometimes write $v = v(\cdot, U, p, q)$ and $\mathbf{h} = \mathbf{h}(\cdot, U, p, q))$ to display the dependence on (U, p, q). The existence and uniqueness of the minimizing pair for μ_0 is a straightforward application of the direct method, just as for μ. The analogous bound to (4.18) is

$$\fint_U \left(|\nabla v(x, U, p, q)|^2 + |\mathbf{h}(x, U, p, q)|^2\right) dx \leq C\left(K_0^2 + |p|^2 + |q|^2\right) \quad \mathbb{P}\text{–a.s.}, \tag{4.20}$$

which follows immediately from (4.5) and (4.7).

Remark 4.1. The quantities μ and μ_0 are inherently variational and may not have a simple description in terms of the PDE. We remark that neither u nor v is necessarily a solution of the equation (1.1), nor is $\mathbf{g}$ or $\mathbf{h}$ necessarily the flux of u or v, respectively. For $(v, \mathbf{h})$, this is because the minimization problem is too restrictive, requiring $\mathbf{h} \in L^2_{\mathrm{sol},0}(U)$ rather than $\mathbf{h} \in L^2_{\mathrm{sol}}(U)$. For $(u, \mathbf{g})$, it is

because there is no boundary condition, and in particular the minimizer may have boundary conditions other than an affine function with slope p^*.

4.5. **Two convexity lemmas.** Uniform convexity enters into our arguments exclusively through the following two lemmas. We recall that Ω was defined in (1.12).

Lemma 4.2. *For every $F \in \Omega$, bounded domain $U \subseteq \mathbb{R}^d$, $w_1, w_2 \in H^1(U)$ and $\mathbf{f}_1, \mathbf{f}_2 \in L^2_{\mathrm{sol}}(U)$,*

$$\begin{aligned}&\fint_U \left(|\nabla w_1(x) - \nabla w_2(x)|^2 + |\mathbf{f}_1(x) - \mathbf{f}_2(x)|^2 \right) dx \\ &\leq 4\Lambda \left(\fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx + \fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx - 2\mu(U,0,0) \right).\end{aligned}$$

Proof. Denote $w := \frac{1}{2}w_1 + \frac{1}{2}w_2$ and $\mathbf{f} := \frac{1}{2}\mathbf{f}_1 + \frac{1}{2}\mathbf{f}_2$ and compute, using (1.9):

$$\begin{aligned}\mu(U,0,0) &\leq \fint_U F(\nabla w(x), \mathbf{f}(x), x)\, dx \\ &\leq \frac{1}{2}\fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx + \frac{1}{2}\fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx \\ &\quad - \frac{1}{8\Lambda}\fint_U \left(|\nabla w_1(x) - \nabla w_2(x)|^2 + |\mathbf{f}_1(x) - \mathbf{f}_2(x)|^2 \right) dx.\end{aligned}$$

A rearrangement of this inequality yields the lemma. □

The next lemma is a converse of the previous one and follows from (1.10).

Lemma 4.3. *For every $F \in \Omega$, bounded domain $U \subseteq \mathbb{R}^d$, $w_1, w_2 \in H^1(U)$ and $\mathbf{f}_1, \mathbf{f}_2 \in L^2_{\mathrm{sol}}(U)$,*

$$\begin{aligned}\fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx &\leq 2\fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx - \mu(U,0,0) \\ &\quad + \frac{\Lambda}{4}\fint_U \left(|\nabla w_1(x) - \nabla w_2(x)|^2 + |\mathbf{f}_1(x) - \mathbf{f}_2(x)|^2 \right) dx.\end{aligned}$$

Proof. Set $w := 2w_2 - w_1$ and $\mathbf{f} := 2\mathbf{f}_2 - \mathbf{f}_1$, so that $w_2 = \frac{1}{2}w + \frac{1}{2}w_1$ and $\mathbf{f}_2 = \frac{1}{2}\mathbf{f} + \frac{1}{2}\mathbf{f}_1$. Using (1.10), we observe that

$$\begin{aligned}&\fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx \\ &\quad \geq \frac{1}{2}\fint_U F(\nabla w(x), \mathbf{f}(x), x)\, dx + \frac{1}{2}\fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx \\ &\qquad - \frac{\Lambda}{8}\fint_U \left(|\nabla w(x) - \nabla w_1(x)|^2 + |\mathbf{f}(x) - \mathbf{f}_1(x)|^2 \right) dx.\end{aligned}$$

Using

$$\mu(U,0,0) \leq \fint_U F(\nabla w(x), \mathbf{f}(x), x)\, dx$$

and rearranging, we get the lemma. □

Remark 4.4. We also obtain versions of the previous two lemmas for functions with affine boundary data. Under the additional assumption that $w_1 + w_2 \in$

$H^1_0(U)$ and $\mathbf{f}_1(x) + \mathbf{f}_2(x) \in L^2_{\mathrm{sol},0}(U)$, the conclusion of Lemma 4.2 may be improved to

$$\begin{aligned}(4.21)\quad \fint_U \left(|\nabla w_1(x) - \nabla w_2(x)|^2 + |\mathbf{f}_1(x) - \mathbf{f}_2(x)|^2\right) dx \\ \le 4\Lambda\bigg(\fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx \\ + \fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx - 2\mu_0(U,0,0)\bigg).\end{aligned}$$

Likewise, under the additional assumptions that $2w_1 - w_2 \in H^1_0(U)$ and $2\mathbf{f}_1 - \mathbf{f}_2 \in L^2_{\mathrm{sol},0}(U)$, the conclusion of Lemma 4.3 may be strengthened to

$$\begin{aligned}(4.22)\quad \fint_U F(\nabla w_1(x), \mathbf{f}_1(x), x)\, dx \le 2 \fint_U F(\nabla w_2(x), \mathbf{f}_2(x), x)\, dx - \mu_0(U,0,0) \\ + \frac{\Lambda}{4} \fint_U \left(|\nabla w_1(x) - \nabla w_2(x)|^2 + |\mathbf{f}_1(x) - \mathbf{f}_2(x)|^2\right) dx.\end{aligned}$$

Remark 4.5. Appropriate versions of Lemmas 4.2 and 4.3 as well as (4.21) and (4.22) for arbitrary $(q^*, p^*), (p, q) \in \mathbb{R}^d \times \mathbb{R}^d$ rather than $(0,0)$, are immediate consequences of the former by applying the results to the operators

$$(p,q) \mapsto F(\hat{p} + p, \hat{q} + q, x) \quad \text{and} \quad (p,q) \mapsto F(p,q,x) - q^* \cdot p - p^* \cdot q.$$

Indeed, while these operators do not in general belong to Ω because they do not satisfy (1.11), they do satisfy the other conditions (1.8), (1.9) and (1.10), and meanwhile (1.11) was not used in the proofs of Lemmas 4.2 and 4.3. We make use of this fact without further comment.

4.6. **Some basic estimates on the triadic cubes.** We next present some inequalities which allow us to compare μ and μ_0 in the trimmed versus the untrimmed cubes. We have, for every $p^*, q^* \in \mathbb{R}^d$ and $n \in \mathbb{N}$,

$$(4.23)\qquad \mu(\boxdot_n, q^*, p^*) \le \mu(\square_n, q^*, p^*) + C3^{-n\beta/(1+\beta)} \left(K_0^2 + |p^*|^2 + |q^*|^2\right).$$

Indeed, it follows from superadditivity that

$$\mu(\square_n, q^*, p^*) \ge \frac{|\boxdot_n|}{|\square_n|} \mu(\boxdot_n, q^*, p^*) + \frac{|\square_n \setminus \boxdot_n|}{|\square_n|} \mu(\square_n \setminus \boxdot_n, q^*, p^*),$$

so that

$$\mu(\square_n, q^*, p^*) - \mu(\boxdot_n, q^*, p^*) \ge \frac{|\square_n \setminus \boxdot_n|}{|\square_n|} \left(\mu(\boxdot_n, q^*, p^*) + \mu(\square_n \setminus \boxdot_n, q^*, p^*)\right),$$

and (4.23) follows using (4.6) and (4.10). Similarly, for μ_0 we have

$$(4.24)\qquad \mu_0(\boxdot_n, p, q) \ge \mu_0(\square_n, p, q) - C3^{-n\beta/(1+\beta)} \left(K_0^2 + |p|^2 + |q|^2\right).$$

We do not have an almost sure inequality which bounds $\mu(\square_n, q^*, p^*)$ by $\mu(\boxdot_n, q^*, p^*)$, but using stationarity, we can show such an inequality in expectation:

$$\begin{aligned}(4.25)\quad \mathbb{E}\left[\mu(\square_n, q^*, p^*)\right] \le \mathbb{E}\left[\mu(\boxdot_{n+1}, q^*, p^*)\right] \\ + C\left(\mathbb{E}\left[\mu(\boxdot_{n+1}, q^*, p^*)\right] - \mathbb{E}\left[\mu(\boxdot_n, q^*, p^*)\right]\right) + C3^{-\frac{n\beta}{1+\beta}} \left(K_0^2 + |p^*|^2 + |q^*|^2\right).\end{aligned}$$

To see this, notice that $\boxdot_{n+1}$ can be partitioned into 3^d+1 disjoint connected sets, which consist of the untrimmed cube $\square_n$, the 3^d-1 trimmed cubes of the form $\boxdot_n(x) \subseteq \boxdot_{n+1}$ which do not intersect $\square_n$ and a trimmed region of measure at most $C3^{-n\beta/(1+\beta)}|\boxdot_{n+1}|$. Using the superadditivity of μ with respect to this partition, taking expectations, using stationary and (4.6), we arrive at

$$\begin{aligned}(4.26)\quad 3^d\mathbb{E}\left[\mu(\boxdot_{n+1},q^*,p^*)\right] \geq \mathbb{E}\left[\mu(\square_n,q^*,p^*)\right] + (3^d-1)\mathbb{E}\left[\mu(\boxdot_n,q^*,p^*)\right] \\ - C3^{-\frac{n\beta}{1+\beta}}\left(K_0^2+|p^*|^2+|q^*|^2\right),\end{aligned}$$

from which (4.25) follows. Using (4.23), we may slightly improve (4.25) to

$$\begin{aligned}(4.27)\quad &\mathbb{E}\left[\mu(\square_n,q^*,p^*)\right] \leq \mathbb{E}\left[\mu(\boxdot_n,q^*,p^*)\right] \\ &+ C\left(\mathbb{E}\left[\mu(\boxdot_{n+1},q^*,p^*)\right] - \mathbb{E}\left[\mu(\square_n,q^*,p^*)\right]\right) + C3^{-\frac{n\beta}{1+\beta}}\left(K_0^2+|p^*|^2+|q^*|^2\right).\end{aligned}$$

This also permits us to compare the minimizers of μ in the trimmed and untrimmed cubes. Writing $U := \boxdot_n \cup (\square_n \setminus \boxdot_n)$ and applying Lemma 4.2, we have

$$\begin{aligned}&\fint_U |(\nabla u,\mathbf{g})(x,U,p^*,q^*) - (\nabla u,\mathbf{g})(x,\square_n,p^*,q^*)|^2\,dx \\ &\quad\leq C\left(\mu(\square_n,q^*,p^*) - \mu(U,q^*,p^*)\right) \\ &\quad\leq C\left(\mu(\square_n,q^*,p^*) - \frac{|\boxdot_n|}{|\square_n|}\mu(\boxdot_n,q^*,p^*) - \frac{|\square_n\setminus\boxdot_n|}{|\square_n|}\mu(\square_n\setminus\boxdot_n,q^*,p^*)\right) \\ &\quad\leq C\left(\mu(\square_n,q^*,p^*) - \mu(\boxdot_n,q^*,p^*)\right) + C\left(K_0^2+|p^*|^2+|q^*|^2\right)3^{-\frac{n\beta}{1+\beta}}.\end{aligned}$$

Taking expectations and using (4.27) and the fact that

$$(\nabla u,\mathbf{g})(x,\boxdot_n,p^*,q^*) = (\nabla u,\mathbf{g})(x,U,p^*,q^*)|_{\boxdot_n},$$

we obtain that

$$\begin{aligned}(4.28)\quad &\mathbb{E}\left[\fint_{\boxdot_n} |(\nabla u,\mathbf{g})(x,\boxdot_n,p^*,q^*) - (\nabla u,\mathbf{g})(x,\square_n,p^*,q^*)|^2\,dx\right] \\ &\leq C\left(\mathbb{E}\left[\mu(\boxdot_{n+1},q^*,p^*)\right] - \mathbb{E}\left[\mu(\square_n,q^*,p^*)\right]\right) + C3^{-\frac{n\beta}{1+\beta}}\left(K_0^2+|p^*|^2+|q^*|^2\right).\end{aligned}$$

4.7. **Some further properties of μ and μ_0.** For every $U \subseteq \mathbb{R}^d$, we have that

$$(4.29)\qquad (p,q) \mapsto \mu_0(U,p,q) - \frac{1}{2\Lambda}\left(|p|^2+|q|^2\right) \quad \text{is convex}$$

and

$$(4.30)\qquad (p,q) \mapsto \mu_0(U,p,q) - \frac{\Lambda}{2}\left(|p|^2+|q|^2\right) \quad \text{is concave.}$$

The first claim follows from (4.21), Remark 4.5 and Jensen's inequality. The second claim follows from (4.22) and Remark 4.5. Combined with (4.7), these imply that $\mu_0(U,\cdot,\cdot)$ satisfies the continuity estimate

$$\begin{aligned}(4.31)\quad &|\mu_0(U,p_1,q_1) - \mu_0(U,p_2,q_2)| \\ &\qquad\leq C\left(K_0+|p_1|+|p_2|+|q_1|+|q_2|\right)\left(|p_1-p_2|+|q_1-q_2|\right).\end{aligned}$$

The analogous inequality also holds for μ:

$$|\mu(U,q_1^*,p_1^*)-\mu(U,q_2^*,p_2^*)| \leq C\left(K_0+|p_1^*|+|p_2^*|+|q_1^*|+|q_2^*|\right)\left(|p_1^*-p_2^*|+|q_1^*-q_2^*|\right). \tag{4.32}$$

To see this, we use $(u,\mathbf{g})(\cdot,U,q_1^*,p_1^*)$ as a candidate for attaining the infimum in the definition of $\mu(U,q_2^*,p_2^*)$ and use (4.18).

4.8. **The homogenized coefficients $\overline{F}$ and functional $\overline{\mathcal{J}}$.** We define the homogenized variational coefficients $\overline{F}$ by

$$\overline{F}(p,q) := \overline{\mu}_0(p,q).$$

It is then immediate from (4.7), (4.12), (4.29) and (4.30) that $\overline{F}$ grows quadratically and is uniformly convex and $C^{1,1}$ in both variables, that is,

$$\frac{1}{\Lambda}\left(|p|^2+|q|^2\right)-CK_0^2 \leq \overline{F}(p,q) \leq \Lambda\left(|p|^2+|q|^2\right)+CK_0^2, \tag{4.33}$$

$$(p,q)\mapsto \overline{F}(p,q)-\frac{1}{2\Lambda}\left(|p|^2+|q|^2\right) \quad \text{is convex} \tag{4.34}$$

and

$$(p,q)\mapsto \overline{F}(p,q)-\frac{\Lambda}{2}\left(|p|^2+|q|^2\right) \quad \text{is concave.} \tag{4.35}$$

Likewise, (4.31) implies that

$$\left|\overline{F}(p_1,q_1)-\overline{F}(p_2,q_2)\right| \leq C\left(K_0+|p_1|+|p_2|+|q_1|+|q_2|\right)\left(|p_1-p_2|+|q_1-q_2|\right) \tag{4.36}$$

and hence

$$\left|\nabla\overline{F}(p,q)\right| \leq C(K_0+|p|+|q|). \tag{4.37}$$

From (4.34) and (4.35) we also deduce that

$$\left|\nabla\overline{F}(p_1,q_1)-\nabla\overline{F}(p_2,q_2)\right| \leq C\left(|p_1-p_2|+|q_1-q_2|\right). \tag{4.38}$$

By (4.8), for every $p,q\in\mathbb{R}^d$,

$$\overline{F}(p,q)\geq p\cdot q. \tag{4.39}$$

In order to see that $\overline{F}$ variationally represents a uniformly monotone and Lipschitz vector field (which we would then denote by $\overline{\mathbf{a}}$), it suffices by Lemma 2.13 to check that, for each $p\in\mathbb{R}^d$,

$$\inf_{q\in\mathbb{R}^d}\left(\overline{F}(p,q)-p\cdot q\right)=0. \tag{4.40}$$

This is a consequence of (4.39) and the duality between $\overline{\mu}$ and $\overline{\mu}_0$, for which we require the results in Section 6. We thus postpone the demonstration of (4.40) and the construction of the homogenized coefficients $\overline{\mathbf{a}}$ to Section 6.5 (see Proposition 6.6).

We define, in each bounded Lipschitz domain $U \subseteq \mathbb{R}^d$, the functional

$$\overline{\mathcal{J}}[u,u^*] := \inf\left\{ \int_U \left(\overline{F}(\nabla u(x),\mathbf{g}(x)) - \nabla u(x)\cdot \mathbf{g}(x)\right) dx \right. \\ \left. : \mathbf{g}\in L^2(U;\mathbb{R}^d),\ -\nabla\cdot \mathbf{g} = u^* \right\}. \tag{4.41}$$

Modulo the proof of (4.40), we have shown that the theory described in Section 2.2 applies to $\overline{\mathcal{J}}$.

5. Structure of the proof of Theorem 3.1

In this section, we reduce the proof of Theorem 3.1 to two ingredients. The first is the core issue, namely the convergence of the subadditive quantities μ and μ_0 defined in the previous section. The second is a general, deterministic fact that follows from an oscillating test function argument and essentially allows to recover Theorem 3.1 from the convergence of the subadditive quantities.

5.1. Two ingredients in the proof of Theorem 3.1.

We first present the results concerning the convergence of the quantities μ and μ_0, which are proved in the next section. As we will see, the quantities $\overline{\mu}$ and $\overline{F}$ are dual convex functions and thus each pair (p,q) is dual to $\nabla\overline{F}(p,q)$. For every $(p,q)\in\mathbb{R}^d\times\mathbb{R}^d$ and bounded domain $U\subseteq\mathbb{R}^d$, we denote

$$\mathcal{E}(U,p,q) := \left|\mu_0(U,p,q) - \overline{F}(p,q)\right| + \left|\mu(U,\nabla\overline{F}(p,q)) - \nabla\overline{F}(p,q)\cdot(p,q) + \overline{F}(p,q)\right|.$$

The goal is to show that, for some exponent $\alpha>0$,

$$\mathbb{P}\left[\mathcal{E}(\square_n,p,q)\gtrsim 3^{-n\alpha}\right]\ll 1.$$

We are interested in bounding the probability on the left side as strongly as possible – as opposed to finding the optimal exponent α, which is much less important for our purposes. In addition, we need an estimate which also possesses some uniformity in (p,q) and allows for translations of the cubes.

The following theorem possesses each of these desired properties. It can be compared to [3, Theorem 3.1]. Its proof is the focus of Section 6.

Theorem 5.1 (Convergence of the subadditive quantities)**.** *Fix $M,R\ge 1$, $\theta\in(0,\beta)$ and $\tau\ge 1$. There exist $s_0(d,\Lambda,\beta,\theta,\tau)\ge 1$ and $C(d,\Lambda,\beta,C_3,\theta,\tau)\ge 1$ such that, for every $n\in\mathbb{N}$, $s\in[s_0,\infty)$ and $t\in[1,\infty)$,*

$$\mathbb{P}\left[\sup_{p,q\in B_{M3^{n\tau/s}}}\ \sup_{y\in B_R}\frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2+|p|^2+|q|^2)3^{-n\theta/s}}\ge t\right]\le CR^dM^{2d}t^{-s}. \tag{5.1}$$

Moreover, if (P4) holds, then for some $\alpha(d,\Lambda,\beta)>0$ and with C depending additionally on (γ,C_4), we have the following stronger estimate: for every

exponent $s \in (0, d\gamma/(d+\gamma))$, $n \in \mathbb{N}$ *and* $t \geq 1$,

$$\mathbb{P}\left[\sup_{p,q \in B_{M3^n}} \sup_{y \in B_{R3^n}} \frac{\mathcal{E}(y + \square_n, p, q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\alpha(d\gamma/(d+\gamma)-s)}} \geq Ct \right] \leq CR^d M^{2d} \exp\left(-3^{ns} t\right). \tag{5.2}$$

The second ingredient in the proof of Theorem 3.1 is the following proposition (which can be compared to [3, Proposition 4.1]) the proof of which is the focus of Section 7. It is a purely deterministic statement which asserts that estimates on $\mathcal{E}$ on mesoscopic cubes (of exactly the sort appearing in Theorem 5.1) imply quenched estimates for the error in homogenization for general macroscopic Dirichlet problems. Thus it essentially reduces Theorem 3.1 to Theorem 5.1.

Proposition 5.2 (Deterministic bounds for Dirichlet problems)**.** *For every bounded Lipschitz domain* $U_0 \subseteq \mathbb{R}^d$ *and* $\delta > 0$*, there exist* $C(d, \Lambda, U_0, \beta, C_3, \delta)$ *and exponents* $\alpha(d, \Lambda, \beta, \delta) > 0$ *and* $\rho(d, \Lambda, \delta) > 1$ *such that the following holds. For every* $F \in \Omega$*,* $m, n \in \mathbb{N}$*,* $U := 3^{n+m} U_0$ *and* $f \in W^{1,2+\delta}(U)$*, if* $u, \overline{u} \in f + H_0^1(U)$ *are such that*

$$\mathcal{J}[u, 0] = \overline{\mathcal{J}}[\overline{u}, 0] = 0, \tag{5.3}$$

then for every integer l *satisfying* $n \leq l \leq m + n$*,*

$$3^{-2(n+m)} \fint_U |u(x) - \overline{u}(x)|^2 \, dx \leq CM^2 \left(\mathcal{E}'_{n,m,M} + 3^{-2m} + 3^{-\alpha(n+m-l)} \right), \tag{5.4}$$

where we denote

$$M := K_0 + \left(\fint_U |\nabla f(x)|^{2+\delta} \, dx \right)^{\frac{1}{2+\delta}} \tag{5.5}$$

and

$$\mathcal{E}'_{n,m,M} = \mathcal{E}'_{n,m,M}(U_0) := \left(\fint_U \left(\sup_{p,q \in B_{CM3^{dm/2}}} \frac{\mathcal{E}(x + \square_n, p, q)}{K_0^2 + |p|^2 + |q|^2} \right)^{\rho} dx \right)^{\frac{1}{\rho}}. \tag{5.6}$$

5.2. **The proof of Theorem 3.1.** We next show that Theorem 3.1 is a consequence of the previous two statements.

Proof of Theorem 3.1. Step 1. We verify the weaker estimate, i.e. (3.1) and (3.4). We allow $C(d, \Lambda, \beta, C_3, \delta, U_0, \theta) \geq 1$ to vary in each occurrence.

Fix $R \geq 1$. Also set $\theta' := \theta + \frac{1}{2}(\beta - \theta)$ and fix $\tau \geq 1$ and $s \geq s_0$ to be selected below. We may suppose without loss of generality that $R = 3^k$ for some $k \in \mathbb{N}$. Write $k = n + m$ with $n, m \in \mathbb{N}$ and m chosen as large as possible subject to the constraint

$$\frac{md}{2} \leq \frac{n\tau}{s},$$

and pick $l := n + \lfloor m/2 \rfloor$. Observe that, with these choices of n, m and l, the last factor in the second term on the right side of (5.4) is estimated by

$$3^{-2m} + 3^{-\alpha(n+m-l)} \leq CR^{-\theta'/s},$$

provided that τ and s_0 are chosen sufficiently large. Indeed, since

$$\alpha(n+m-l) \sim \frac{\alpha\tau}{2sd}n, \qquad \frac{\theta'}{s}(n+m) \sim \left(1+\frac{\tau}{sd}\right)\frac{\theta'}{s}n \qquad (n\to\infty),$$

it suffices to check that

$$\frac{\alpha\tau}{2d} > \left(1+\frac{\tau}{sd}\right)\theta',$$

and this is satisfied provided that $\tau > 2d\theta'/\alpha$ and $s_0 \geq C$ is sufficiently large.

Therefore, Proposition 5.2 gives

$$R^{-2}\fint_U |u(x)-\overline{u}(x)|^2\,dx \leq CM^2\mathcal{E}'_{n,m,M} + CM^2R^{-\theta'/s},$$

where $\mathcal{E}'_{n,m,M}$ is defined in (5.6). By Jensen's inequality and stationarity, we find that, for every $s \geq \rho$, where $\rho(d,\Lambda,\delta) > 1$ is the exponent in the statement of Proposition 5.2, we have

$$\begin{aligned}
\mathbb{E}\left[\left(\mathcal{E}'_{n,m,M_0}\right)^s\right] &= \mathbb{E}\left[\left(\fint_U \left(\sup_{p,q\in B_{CM_0 3^{md/2}}} \frac{\mathcal{E}(x+\square_n,p,q)}{K_0^2+|p|^2+|q|^2}\right)^\rho dx\right)^{\frac{s}{\rho}}\right] \\
&\leq \mathbb{E}\left[\fint_U \left(\sup_{p,q\in B_{CM_0 3^{md/2}}} \frac{\mathcal{E}(x+\square_n,p,q)}{K_0^2+|p|^2+|q|^2}\right)^s dx\right] \\
&\leq \mathbb{E}\left[\left(\sup_{x\in B_{\sqrt{d}}}\sup_{p,q\in B_{CM_0 3^{md/2}}} \frac{\mathcal{E}(x+\square_n,p,q)}{K_0^2+|p|^2+|q|^2}\right)^s\right].
\end{aligned}$$

We next use the fact that, for a nonnegative random variable X,

$$\mathbb{E}[X^s] = \int_0^\infty t^{s-1}\mathbb{P}[X>t]\,dt,$$

and apply Theorem 5.1 (replacing s by $s+1$ and θ by θ' there, and increasing s_0 as necessary) to obtain, for every $s \geq s_0(d,\Lambda,\beta,\theta,\delta)$,

$$\mathbb{E}\left[\left(\sup_{x\in B_{\sqrt{d}}}\sup_{p,q\in B_{CM_0 3^{md/2}}} \frac{\mathcal{E}(\square_n,p,q)}{K_0^2+|p|^2+|q|^2}\right)^s\right] \leq CM_0^{2d}3^{-n\theta'}.$$

Setting

$$\mathcal{X}_{s,n,M_0} := C\left(1+3^{n\theta'/s}\mathcal{E}'_{n,m,M_0}\right),$$

we therefore obtain, under the assumption that $M \leq M_0$, the estimate

$$R^{-2}\fint_U |u(x)-\overline{u}(x)|^2\,dx \leq \mathcal{X}_{s,n,M_0}M^2 3^{-n\theta'/s},$$

with $\mathcal{X}_{s,n,M_0}$ satisfying $\mathbb{E}[\mathcal{X}^s_{s,n,M_0}] \leq CM_0^{2d}$. We now define

$$\mathcal{X}_{s,M_0} := \sup_{j\in\mathbb{N}} 3^{-j(\beta-\theta)/4s}\mathcal{X}_{s,j,M_0}.$$

Clearly, $\mathbb{E}[\mathcal{X}^s_{s,M_0}] \leq CM_0^{2d}$. Note also that the definition of $\mathcal{X}_{s,M_0}$ does not depend on R, and

$$R^{-2}\fint_U |u(x)-\overline{u}(x)|^2\,dx \leq \mathcal{X}_{s,M_0}M^2 3^{-n\theta''/s}, \tag{5.7}$$

where $\theta'' := \theta + \frac{1}{4}(\beta - \theta)$. If s_0 is large enough (depending on the appropriate quantities), then by our choice of n and m, we have $3^{-n\theta''} \le R^{-\theta}$. To remove the dependence of $\mathcal{X}_{s,M_0}$ on M_0, we define

$$\mathcal{X}_s := \sup_{k\in\mathbb{N}} 2^{-2dk/s}\mathcal{X}_{s+1,2^k}.$$

It is clear that (3.1) holds and, translating (5.7) in terms of $\mathcal{X}_s$ yields (3.4). The proof is now complete.

Step 2. Under assumption (P4), we verify the stronger estimates (3.5), (3.6). Here we fix an exponent $s \in (0, d\gamma/(d+\gamma))$ and allow C and c to depend additionally on (γ, C_4, s). We pick s_1 and s_2 such that $s < s_1 < s_2 < d\gamma/(d+\gamma)$ with the gaps between these bounded by c. As above we may suppose that $R = 3^k$ for some $k \in \mathbb{N}$. We write $k = n + m$ where $n, m \in \mathbb{N}$ are chosen so that m is as large as possible such that $n \ge m$, and

$$(m+n)s_1 < ns_2 \quad \text{and} \quad md\gamma/(d+\gamma) \le n\alpha\left(d\gamma/(d+\gamma) - s_2\right). \tag{5.8}$$

Note that, in addition to R, the integers m and n depend only on s. Choosing $l = n + \lfloor m/2 \rfloor$ in (5.4) yields

$$\begin{aligned} R^{-2} \fint_U |u(x) - \overline{u}(x)|^2 \, dx &\le CM^2\mathcal{E}'_{n,m,M} + CM^2R^{-\alpha/s} \\ &\le CM^2\left(1 + \mathcal{Y}_{s,M_0}R^{-s}\right)R^{-\alpha/s}, \end{aligned} \tag{5.9}$$

provided that $M \le M_0$, where we have defined

$$\mathcal{Y}_{s,M_0} := \sup_{n\in\mathbb{N}}\left(R^s\left(3^{md\gamma/(d+\gamma)}\mathcal{E}'_{n,m,M_0} - C\right)_+\right).$$

(Here and in what follows we think of R and m as functions of n.) We next study the integrability of $\mathcal{Y}_{s,M_0}$. According to (5.2), for every $t \ge 1$,

$$\mathbb{P}\left[\mathcal{E}'_{n,m,M_0} \ge C3^{-(n+m)\alpha(d\gamma/(d+\gamma)-s_2)}t\right] \le CM_0^{2d}\exp\left(-3^{(n+m)s_2}t\right).$$

Using (5.8) and rearranging this leads to the bound

$$\mathbb{P}\left[R^s\left(3^{md\gamma/(d+\gamma)}\mathcal{E}'_{n,m,M_0} - C\right)_+ \ge Ct\right] \le CM_0^{2d}\exp\left(-c3^{-(s_1-s)}t\right).$$

Taking a union bound and summing this over $n \in \mathbb{N}$ yields

$$\mathbb{P}\left[\mathcal{Y}_{s,M_0} \ge Ct\right] \le CM_0^{2d}\exp\left(-ct\right).$$

Defining $\mathcal{Y}_s := c\sup_k k\mathcal{Y}_{s,2^k}$ and integrating the previous inequality, we obtain (3.5). The inequality (3.6) is obtained by expressing (5.9) in terms of $\mathcal{Y}_s$. This completes the proof of the theorem. □

6. Convergence of the subadditive quantities

In this section, we prove Theorem 5.1. We focus the majority of our effort to obtain the following slightly weaker statement. In the final subsection, we derive Theorem 5.1 from it.

Proposition 6.1. *Fix $(q^*, p^*) \in \mathbb{R}^d \times \mathbb{R}^d$. There exist a unique pair*

$$(\overline{P}(q^*, p^*), \overline{Q}(q^*, p^*)) \in \mathbb{R}^d \times \mathbb{R}^d$$

such that

$$\overline{\mu}(q^*, p^*) + p^* \cdot \overline{Q} + q^* \cdot \overline{P} = \overline{\mu}_0(\overline{P}, \overline{Q}) \tag{6.1}$$

and, for every $\theta \in (0, \beta)$, an exponent $s_0(d, \Lambda, \beta, \theta) > 0$ and $C(d, \Lambda, \beta, C_3, \theta) > 0$ such that, for all $s \in [s_0, \infty)$, $R \geq 1$, $n \in \mathbb{N}$ and $t \geq 1$,

$$\mathbb{P}\left[\sup_{y \in B_R} \mathcal{E}_*(y + \square_n, q^*, p^*) \geq C(K_0^2 + |p^*|^2 + |q^*|^2) 3^{-n\theta/s} t\right] \leq C R^d t^{-s}, \tag{6.2}$$

where we denote

$$\mathcal{E}_*(U, q^*, p^*) := |\mu(U, q^*, p^*) - \overline{\mu}(q^*, p^*)| + \left|\mu_0(U, \overline{P}, \overline{Q}) - \overline{\mu}_0(\overline{P}, \overline{Q})\right|. \tag{6.3}$$

Moreover, if we assume (P4), then, for every exponent $s \in (0, d\gamma/(d+\gamma))$ and with C depending additionally on (s, γ, C_4), we have the following stronger estimate: for all $R \geq 1$, $n \in \mathbb{N}$ and $t \geq 1$,

$$\mathbb{P}\left[\sup_{y \in B_R} \mathcal{E}_*(y + \square_n, q^*, p^*) \geq C(K_0^2 + |p^*|^2 + |q^*|^2) 3^{-n \min\{\alpha, d\gamma/(d+\gamma) - s\}} t\right] \leq C R^d \exp\left(-3^{ns} t\right). \tag{6.4}$$

6.1. **Reduction to the case $p^* = q^* = 0$.** It suffices to prove Proposition 6.1 for $p^* = q^* = 0$. Indeed, given $p^*, q^* \in \mathbb{R}^d$, let

$$F_{p^*,q^*}(p, q, x) := F(p + p^*, q + q^*, x) - p \cdot q^* - q \cdot p^* - p^* \cdot q^*.$$

Note that F_{p^*,q^*} belongs to Ω and satisfies the same assumptions as F, except that the constant K_0 in (P1) must be replaced by $K_0 + |p^*| + |q^*|$. Applying Proposition 6.1 to F_{p^*,q^*} (to be more precise, to the pushforward of $\mathbb{P}$ under the map $F \mapsto F_{p^*,q^*}$) at $(0, 0)$ then yields the general result.

6.2. **Flatness of minimizers.** The purpose of this subsection is to prove Lemma 6.3 (stated below), which is the key step in the proof of Proposition 6.1. The main idea is to show that if the difference in the expectation of μ between two successive triadic scales is small, then the minimizers of μ must be very flat. This invites a comparison to corresponding minimizers of μ_0, allowing us to show that the expectation of μ is close to its limit.

In order to lighten notation, we simply write $\mu(U)$ for $\mu(U, 0, 0)$, $\overline{\mu}$ for $\overline{\mu}(0, 0)$, and $(u, \mathbf{g})(\cdot, U)$ for $(u, \mathbf{g})(\cdot, U, 0, 0)$. We denote the spatial averages of the minimizing pair $(u, \mathbf{g})(\cdot, U)$ of $\mu(U)$ by

$$P(U) := \fint_U \nabla u(x, U)\, dx \quad \text{and} \quad Q(U) := \fint_U \mathbf{g}(x, U)\, dx,$$

and we use the shorthand notation

$$\overline{P}_n := \mathbb{E}\left[P(\square_n)\right] \quad \text{and} \quad \overline{Q}_n := \mathbb{E}\left[Q(\square_n)\right].$$

Throughout this section, we denote the difference between the expected values of μ at the two successive triadic scales $n+1$ and n (with the smaller cube trimmed) by

$$\tau_n := \mathbb{E}\left[\mu(\square_{n+1})\right] - \mathbb{E}\left[\mu(\boxdot_n)\right]. \tag{6.5}$$

and we denote errors which will accumulate due to mixing and trimming at scale n by

$$\kappa_n := K_0^2 3^{-n\beta/8(1+\beta)} + K_0^2 3^{-n/3}.$$

The following adaptation of [3, Lemma 3.2] is the first step in the argument for Proposition 6.1. The lemma states that the variances of $P(\boxdot_n)$ and $Q(\boxdot_n)$ are controlled by a multiple of $\tau_n + \kappa_n$. The proof uses the weak mixing condition (P3) for the first time in the paper. In preparation for its application, we rephrase (P3) in terms of the following covariance estimate (this is proved in Appendix A, see (A.2)): for every $U, V \subseteq \mathbb{R}^d$, $\mathcal{F}_U$–measurable random variable X and $\mathcal{F}_V$–measurable random variable Y, we have

$$\operatorname{cov}\left[X, Y\right] \le 4C_3 \|X\|_\infty \|Y\|_\infty (1 + \operatorname{dist}(U,V))^{-\beta}, \tag{6.6}$$

where $\|X\|_\infty$ denotes the $\mathbb{P}$–essential supremum of $|X|$:

$$\|X\|_\infty := \inf\left\{\lambda > 0 : \mathbb{P}\left[|X| > \lambda\right] = 0\right\}.$$

Lemma 6.2. *There exist $C(d,\Lambda,\beta) \ge 1$ and $C'(d,\Lambda,\beta,C_3) \ge 1$ such that, for every $n \in \mathbb{N}$,*

$$\mathbb{E}\left[\left|P(\boxdot_n) - \overline{P}_n\right|^2\right] + \mathbb{E}\left[\left|Q(\boxdot_n) - \overline{Q}_n\right|^2\right] \le C\tau_n + C'\kappa_n.$$

Proof. In this argument, C denotes a constant depending only on (d,Λ,β) and C' denotes a constant depending only on (d,Λ,β,C_3); these may vary in each occurrence. Fix $n \in \mathbb{N}$, a unit direction $e \in \partial B_1$, a smooth vector field $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^d$ and a smooth function $\varphi : \mathbb{R}^d \to \mathbb{R}$ satisfying the following:

$$\begin{cases} \mathbf{f} \text{ and } \varphi \text{ have compact support in } \boxdot_{n+1}, \\ \mathbf{f} = \nabla\varphi = e \quad \text{in } \square_n, \\ \operatorname{div}\mathbf{f} = 0, \\ |\mathbf{f}| + |\nabla\varphi| \le C \quad \text{in } \mathbb{R}^d. \end{cases}$$

Since $\mathbf{f}$ and $\mathbf{g}(\cdot,\square_{n+1})$ are divergence-free, we have

$$\fint_{\square_{n+1}} \mathbf{f}(x)\cdot\nabla u(x,\square_{n+1})\,dx = \fint_{\square_{n+1}} \nabla\varphi(x)\cdot\mathbf{g}(x,\square_{n+1})\,dx = 0. \tag{6.7}$$

Step 1. We localize $\nabla u(\cdot,\square_{n+1})$ and $\mathbf{g}(\cdot,\square_{n+1})$ to the trimmed subcubes. The precise claim is that

$$\mathbb{E}\left[\fint_{U\cup V} \left(|\nabla u(x,\square_{n+1}) - \nabla u(x,U\cup V)|^2 + |\mathbf{g}(x,\square_{n+1}) - \mathbf{g}(x,U\cup V)|^2\right)dx\right] \le C\tau_n + C\kappa_n, \tag{6.8}$$

where

$$U := \bigcup_{z\in\{-3^n,0,3^n\}^d} (z + \boxdot_n) \qquad \text{and} \qquad V := \square_{n+1} \setminus \overline{U},$$

and $\overline{U}$ denotes the closure of U. There are 3^d+1 connected components of $U\cup V$, which are V and the cubes of the form $\boxdot_n(x)\subseteq\square_{n+1}$. Thus for every $x\in\square_{n+1}$, we have

$$\begin{cases}(u,\mathbf{g})(\cdot,U\cup V)|_{\boxdot_n(x)}=(u,\mathbf{g})(\cdot,U)|_{\boxdot_n(x)}=(u,\mathbf{g})(\cdot,\boxdot_n(x)),\\ (u,\mathbf{g})(\cdot,U\cup V)|_V=(u,\mathbf{g})(\cdot,V).\end{cases}\tag{6.9}$$

In particular,

$$\mu(U)=3^{-d}\sum_{\boxdot_n(x)\subseteq\square_{n+1}}\mu(\boxdot_n(x))$$

and

$$\mu(U\cup V)=\frac{|V|}{|U\cup V|}\mu(V)+\frac{|U|}{|U\cup V|}\mu(U).\tag{6.10}$$

By stationarity,

$$\mathbb{E}\left[\mu(U)\right]=\mathbb{E}\left[\mu(\boxdot_n)\right].$$

Notice that, by (4.10), the volume of V is small in proportion to that of $\square_{n+1}$:

$$|V|\le C3^{-n\beta/(1+\beta)}\left|\square_{n+1}\right|.\tag{6.11}$$

It follows from (6.10) and (4.6) that

$$\mu(U\cup V)=\mu(U)+\frac{|V|}{|U\cup V|}\left(\mu(V)-\mu(U)\right)\ge\mu(U)-CK_0^2 3^{-n\beta/(1+\beta)}.$$

By Lemma 4.2 and the previous inequality,

$$\begin{aligned}\mathbb{E}\left[\fint_{U\cup V}\left(|\nabla u(x,\square_{n+1})-\nabla u(x,U\cup V)|^2+|\mathbf{g}(x,\square_{n+1})-\mathbf{g}(x,U\cup V)|^2\right)dx\right]&\\ \le C\left(\mathbb{E}\left[\mu(\square_{n+1})\right]-\mathbb{E}\left[\mu(U\cup V)\right]\right)&\\ \le C\left(\mathbb{E}\left[\mu(\square_{n+1})\right]-\mathbb{E}\left[\mu(U)\right]+CK_0^2 3^{-n\beta/(1+\beta)}\right)&\\ =C(\tau_n+C\kappa_n),&\end{aligned}$$

and this is (6.8).

Step 2. Using (6.7), (6.8) and (6.9), we compute

$$\begin{aligned}\operatorname{var}\left[\int_V\mathbf{f}(x)\cdot\nabla u(x,V)\,dx+\sum_{\boxdot_n(y)\subseteq\square_{n+1}}\int_{\boxdot_n(y)}\mathbf{f}(x)\cdot\nabla u(x,\boxdot_n(y))\,dx\right]&\\ =\operatorname{var}\left[\int_{U\cup V}\mathbf{f}(x)\cdot\nabla u(x,U\cup V)\,dx\right]&\\ \le 2\operatorname{var}\left[\int_{\square_{n+1}}\mathbf{f}(x)\cdot\nabla u(x,\square_{n+1})\,dx\right]+C(\tau_n+C\kappa_n)\left|\square_{n+1}\right|^2&\\ =C(\tau_n+C\kappa_n)\left|\square_{n+1}\right|^2.&\end{aligned}$$

By (4.18) and (6.11), we have

$$\int_V|\nabla u(x,V)|^2\,dx\le C|V|K_0^2\le CK_0^2 3^{-n\beta/(1+\beta)}\left|\square_{n+1}\right|=C\kappa_n\left|\square_{n+1}\right|.$$

The previous two inequalities yield

$$\operatorname{var}\left[\sum_{\square_n(y)\subseteq\square_{n+1}} \fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\right] \le C(\tau_n + C\kappa_n).$$

Now we expand the variance using the identity

$$\begin{aligned}\operatorname{var}&\left[\sum_{\square_n(y)\subseteq\square_{n+1}} \fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\right]\\ &= \sum_{\square_n(y),\square_n(z)\subseteq\square_{n+1}} \operatorname{cov}\left[\fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\,,\right.\\ &\qquad\qquad \left.\fint_{\square_n(z)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(z))\,dx\right].\end{aligned}$$

Recall that by (4.18) and the definition of $\mathbf{f}$, the random variables appearing in the covariances above are $\mathbb{P}$-a.s. bounded by CK_0. In view of (4.9), we can apply the mixing condition (P3) in the form of (6.6) to obtain, for every $\square_n(y)\neq\square_n(z)$,

$$\begin{aligned}\left|\operatorname{cov}\left[\fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\ ,\ \fint_{\square_n(z)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(z))\,dx\right]\right|&\\ \le C'K_0^2 3^{-n\beta/(1+\beta)} = C'K_0^2\kappa_n.&\end{aligned}$$

To sum up, we have proved the estimate

$$\sum_{\square_n(y)\subseteq\square_{n+1}} \operatorname{var}\left[\fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\right] \le C(\tau_n + C'\kappa_n).$$

The previous inequality and the fact that $\mathbf{f}\equiv e$ in $\square_n$ imply

$$\begin{aligned}\operatorname{var}\left[e\cdot P(\square_n)\right] &= \operatorname{var}\left[\fint_{\square_n} \mathbf{f}(x)\cdot\nabla u(x,\square_n)\,dx\right]\\ &\le \sum_{\square_n(y)\subseteq\square_{n+1}} \operatorname{var}\left[\fint_{\square_n(y)} \mathbf{f}(x)\cdot\nabla u(x,\square_n(y))\,dx\right] \le C(\tau_n + C'\kappa_n).\end{aligned}$$

An almost identical argument, starting from the second equality of (6.7) rather than the first and replacing each occurrence of $\mathbf{f}(x)\cdot\nabla u$ by $\nabla\varphi(x)\cdot\mathbf{g}$, gives the estimate

$$\operatorname{var}\left[e\cdot Q(\square_n)\right] \le C(\tau_n + C'\kappa_n).$$

Summing the previous two inequalities over $e\in\{e_1,\ldots,e_d\}$ yields the lemma. □

Note that (4.28) permits us to obtain the conclusion of Lemma 6.2 for the untrimmed cubes: there exists $C(d,\Lambda,\beta)\ge 1$ and $C'(d,\Lambda,\beta,C_3)\ge 1$ such that, for every $n\in\mathbb{N}$,

$$\mathbb{E}\left[\left|P(\square_n)-\overline{P}_n\right|^2\right] + \mathbb{E}\left[\left|Q(\square_n)-\overline{Q}_n\right|^2\right] \le C\tau_n + C'\kappa_n. \tag{6.12}$$

The next lemma is the key step in the proof of Proposition 6.1. It allows us to estimate the expected difference between μ and its limit $\overline{\mu}$ by the expected difference between μ at two successive scales.

Lemma 6.3. *There exist $C(d,\Lambda,\beta) \geq 1$ and $C'(d,\Lambda,\beta,C_3) \geq 1$ such that, for every $n \in \mathbb{N}$,*

$$\mathbb{E}\left[\mu_0(\boxdot_{2n}, \overline{P}_n, \overline{Q}_n)\right] \leq \mathbb{E}\left[\mu(\square_n)\right] + C(\tau_n + C'\kappa_n). \tag{6.13}$$

Proof. The convention for the constants C and C' is the same here as in the previous lemma.

Fix $n \in \mathbb{N}$. In order to estimate the quantity $\mu_0(\boxdot_{2n}, \overline{P}_n, \overline{Q}_n)$ from above, we construct a candidate for minimizing the energy which satisfies the appropriate affine boundary conditions. Precisely, it suffices to exhibit $(v, \mathbf{h}) \in H^1_0(\boxdot_{2n}) \times L^2_{\mathrm{sol},0}(\boxdot_{2n})$ satisfying

$$\mathbb{E}\left[\fint_{\boxdot_{2n}} F\left(\overline{P}_n + \nabla v(x), \overline{Q}_n + \mathbf{h}(x), x\right) dx \right] \leq \mathbb{E}\left[\mu(\square_n)\right] + C(\tau_n + C'\kappa_n). \tag{6.14}$$

Step 1. The construction of the candidate $(v, \mathbf{h}) \in H^1_0(\boxdot_{2n}) \times L^2_{\mathrm{sol},0}(\boxdot_{2n})$, which we build by patching the minimizers for μ on the overlapping family of cubes $\{z + \square_{n+1} : z \in 3^n\mathbb{Z}^d\}$. We first build a partition of unity subordinate to these cubes by setting

$$\psi(y) := \int_{\square_n} \psi_0(y - x)\, dx,$$

where $\psi_0 \in C^\infty_c(\mathbb{R}^d)$ is a smooth, even function satisfying

$$0 \leq \psi_0 \leq C3^{-dn}, \quad \int_{\mathbb{R}^d} \psi_0(x)\, dx = 1, \quad |\nabla \psi_0| \leq C3^{-(d+1)n}, \quad \operatorname{supp} \psi_0 \subseteq \square_n.$$

We note that ψ is smooth, even and supported in $\square_{n+1}$, satisfies $0 \leq \psi \leq 1$, and the translates of it form a partition of unity: for each $x \in \mathbb{R}^d$,

$$\sum_{z \in 3^n\mathbb{Z}^d} \psi(x - z) = 1. \tag{6.15}$$

Moreover, we have

$$\sup_{x \in \mathbb{R}^d} |\nabla \psi(x)| \leq C3^{-n}. \tag{6.16}$$

We also introduce two smooth cutoff functions $\xi, \zeta \in C^\infty_c(\square_{2n})$ satisfying

$$\begin{aligned} &0 \leq \xi \leq 1, \quad \xi \equiv 1 \text{ on } \{x \in \square_{2n} : \operatorname{dist}(x, \partial\square_{2n}) > C3^{2n/(1+\delta)}\}, \\ &\qquad \xi \equiv 0 \text{ on } \left\{x \in z + \square_n : z \in 3^n\mathbb{Z}^d,\ z + \square_{n+1} \not\Subset \boxdot_{2n}\right\}, \quad |\nabla \xi| \leq C3^{-2n/(1+\delta)} \end{aligned} \tag{6.17}$$

where $\delta \in (0, \beta]$ will be selected below in Step 7, and

$$\begin{aligned} &0 \leq \zeta \leq 1, \quad \zeta \equiv 1 \text{ on } \left\{x \in z + \square_n : z \in 3^n\mathbb{Z}^d,\ z + \square_{n+3} \subseteq \square_{2n}\right\}, \\ &\qquad \zeta \equiv 0 \text{ on } \left\{x \in z + \square_{n+1} : z \in 3^n\mathbb{Z}^d,\ z + \square_{n+1} \not\subseteq \square_{2n}\right\}, \quad |\nabla \zeta| \leq C3^{-n}. \end{aligned} \tag{6.18}$$

To construct v, we first define a vector field $\mathbf{f} \in L^2(\mathbb{R}^d; \mathbb{R}^d)$ by

$$\mathbf{f}(x) := \zeta(x) \sum_{z \in 3^n\mathbb{Z}^d} \psi(x - z)\left(\nabla u(x, z + \square_{n+1}) - \overline{P}_n\right). \tag{6.19}$$

Since $\mathbf{f}$ is not necessarily the gradient of an H^1 function, due to the errors made by introducing the partition of unity and the cutoff function, we need to take its Helmholtz-Hodge projection. We may write

$$\mathbf{f} = \overline{\mathbf{f}} + \nabla w - \nabla \cdot \mathbf{S} \quad \text{in } \square_{2n}, \tag{6.20}$$

where

$$\overline{\mathbf{f}} := \fint_{\square_{2n}} \mathbf{f}(x)\,dx,$$

$w \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ is defined as the unique solution of

$$\begin{cases} -\Delta w = -\nabla\cdot\mathbf{f} & \text{in } \square_{2n}, \\ \fint_{\square_{2n}} w(x)\,dx = 0, \\ w \text{ is } \square_{2n}\text{–periodic}, \end{cases}$$

and $\mathbf{S}$ is valued in the skew-symmetric matrices and has entries $S_{ij} \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ uniquely determined (up to an additive constant) by

$$\begin{cases} -\Delta S_{ij} = \partial_j f_i - \partial_i f_j & \text{in } \square_{2n}, \\ S_{ij} \text{ is } \square_{2n}\text{–periodic}. \end{cases}$$

Here f_i denotes the ith entry of $\mathbf{f}$ and $\nabla\cdot\mathbf{S}$ is the vector field with entries $\sum_{j=1}^d \partial_j S_{ij}$. Indeed, one may check via a straightforward computation that each component of the vector field $\mathbf{f} - \nabla w + \nabla\cdot\mathbf{S}$ is harmonic and therefore constant by periodicity. This constant must be $\overline{\mathbf{f}}$ since ∇w and $\nabla\cdot\mathbf{S}$ have zero mean in $\square_{2n}$. This confirms (6.20).

We finally define $v \in H^1_0(\boxdot_{2n})$ by setting

$$v(x) := \xi(x)w(x), \quad x \in \square_{2n}.$$

Note that the cutoff function ξ is supported in $\boxdot_{2n}$ and thus we indeed have $v \in H^1_0(\boxdot_{2n})$. Below we will argue that ∇v is expected to be close to $\mathbf{f}$ in $L^2(\boxdot_{2n})$ due to the fact that, as we will show, w, $\overline{\mathbf{f}}$ and $\nabla\cdot\mathbf{S}$ each have a small expected L^2 norm.

We proceed similarly to construct $\mathbf{h} \in L^2_{\mathrm{sol},0}(\boxdot_{2n})$. Define $\mathbf{k} \in L^2(\mathbb{R}^d;\mathbb{R}^d)$ by

$$\mathbf{k}(x) := \zeta(x) \sum_{z \in 3^n\mathbb{Z}^d} \psi(x-z)\left(\mathbf{g}(x, z+\square_{n+1}) - \overline{Q}_n\right), \tag{6.21}$$

Since $\mathbf{k}$ is not necessarily solenoidal, we remove its divergence part via the Helmholtz-Hodge projection. As above, we may write

$$\mathbf{k} = \overline{\mathbf{k}} + \nabla h - \nabla\cdot\mathbf{T}$$

where

$$\overline{\mathbf{k}} := \fint_{\square_{2n}} \mathbf{k}(x)\,dx,$$

$h \in H^1_{\mathrm{per}}(\square_{2n})$ is the unique solution of

$$\begin{cases} -\Delta h = -\nabla\cdot\mathbf{k} & \text{in } \square_{2n}, \\ \fint_{\square_{2n}} h(x)\,dx = 0, \\ h \text{ is } \square_{2n}\text{–periodic}, \end{cases}$$

and $\mathbf{T}$ is a skew-symmetric matrix-valued field with entries in $H^1_{\mathrm{per}}(\square_{2n})$. We finally define $\mathbf{h} \in L^2_{\mathrm{sol},0}(\boxdot_{2n})$ by setting

$$\mathbf{h}(x) := \xi(x)\left(\nabla\cdot\mathbf{T}\right)(x) - \nabla\widetilde{h},$$

where $\widetilde{h} \in H^1(\boxdot_{2n})$ is defined (also uniquely up to a constant) as the solution of

$$\begin{cases} -\Delta \widetilde{h} = -\nabla \xi(x) \cdot (\nabla \cdot \mathbf{T}) & \text{in } \boxdot_{2n}, \\ \partial_\nu \widetilde{h} = 0 & \text{on } \partial \boxdot_{2n}. \end{cases}$$

It is clear that $\mathbf{h} \in L^2_{\mathrm{sol},0}(\boxdot_{2n})$. Below we will argue that $|\mathbf{k} - \mathbf{h}|$ has small expected $L^2(\boxdot_{2n})$ norm.

This completes the construction of $(v, \mathbf{h}) \in H^1_0(\boxdot_{2n}) \times L^2_{\mathrm{sol},0}(\boxdot_{2n})$. The rest of the argument is focused on the proof of (6.14).

Step 2. We show that, for every $z \in 3^n\mathbb{Z}^d \cap \square_{2n}$,

$$\mathbb{E}\bigg[\fint_{z+\square_n} \Big(\Big| \mathbf{f}(x) - \zeta(x) \big(\nabla u(x, z+\square_{n+1}) - \overline{P}_n \big) \Big|^2 + \Big| \mathbf{k}(x) - \zeta(x) \big(\mathbf{g}(x, z+\square_{n+1}) - \overline{Q}_n \big) \Big|^2 \Big) \, dx \bigg] \le C\tau_n. \tag{6.22}$$

By Lemma 4.2 we have, for every $z \in 3^n\mathbb{Z}^d$,

$$\begin{aligned} &\sum_{y \in \{-3^n, 0, 3^n\}^d} \fint_{z+y+\square_n} \left| (\nabla u, \mathbf{g})(x, z+\square_{n+1}) - (\nabla u, \mathbf{g})(x, z+y+\square_n) \right|^2 dx \\ &\quad \le C \sum_{y \in \{-3^n, 0, 3^n\}^d} \left(\fint_{z+y+\square_n} F\left((\nabla u, \mathbf{g})(x, z+\square_{n+1}), x \right) dx - \mu(y+z+\square_n) \right) \\ &\quad = C \left(\mu(z+\square_{n+1}) - \sum_{y \in \{-3^n, 0, 3^n\}^d} \mu(z+y+\square_n) \right). \end{aligned}$$

Taking expectations and using the triangle inequality, we find that, for every $z \in 3^n\mathbb{Z}^d$ and $y \in \{-3^n, 0, 3^n\}^d$,

$$\mathbb{E}\bigg[\fint_{z+\square_n} \left| (\nabla u, \mathbf{g})(x, z+y+\square_{n+1}) - (\nabla u, \mathbf{g})(x, z+\square_{n+1}) \right|^2 dx \bigg] \le C\tau_n. \tag{6.23}$$

The bound on the first term on the left of (6.22) is obtained from the previous inequality and the following identity, which holds for every $z \in 3^n\mathbb{Z}^d$ and $x \in z + \square_n$ by the definition of $\mathbf{f}$:

$$\begin{aligned} &\mathbf{f}(x) - \zeta(x) \big(\nabla u(x, z+\square_{n+1}) - \overline{P}_n \big) \\ &\qquad = \zeta(x) \sum_{y \in \{-3^n, 0, 3^n\}^d} \psi(x-y) \left(\nabla u(x, z+y+\square_{n+1}) - \nabla u(x, z+\square_{n+1}) \right). \end{aligned}$$

The bound on the second term on the left of (6.22) follows from (6.23) and a similar identity.

Step 3. We claim that

$$\mathbb{E}\left[\left| \overline{\mathbf{f}} \right|^2 \right] + \mathbb{E}\left[\left| \overline{\mathbf{k}} \right|^2 \right] \le C\tau_n + C'\kappa_n. \tag{6.24}$$

In view of (6.18), it is convenient to denote

$$\mathcal{Z}_n := \left\{ z \in 3^n\mathbb{Z}^d \,:\, z + \square_{n+3} \subseteq \square_{2n} \right\}. \tag{6.25}$$

Observe that $(3^n\mathbb{Z}^d \cap \square_{2n}) \setminus \mathcal{Z}_n$ has $C3^{n(d-1)}$ elements. By Lemma 6.2 (or more precisely (6.12)), (6.22) and (4.18), we have

$$\begin{aligned}
\mathbb{E}\left[\left|\overline{\mathbf{f}}\right|^2\right] &\le \mathbb{E}\left[3^{-nd} \sum_{z\in 3^n\mathbb{Z}^d\cap\square_{2n}} \left| \fint_{z+\square_n} \mathbf{f}(x)\,dx \right|^2\right] \\
&\le 2\mathbb{E}\left[3^{-nd} \sum_{z\in 3^n\mathbb{Z}^d\cap\square_{2n}} \left| \fint_{z+\square_n} \zeta(x)\left(\nabla u(x, z+\square_{n+1})\,dx - \overline{P}_n\right)\right|^2\right] + C\tau_n \\
&\le 2\mathbb{E}\left[3^{-nd} \sum_{z\in \mathcal{Z}_n} \left| \fint_{z+\square_n} \nabla u(x, z+\square_{n+1})\,dx - \overline{P}_n\right|^2\right] + CK_0^2 3^{-n} + C\tau_n \\
&\le C\tau_n + C'\kappa_n.
\end{aligned}$$

An analogous calculation gives the estimate for $\mathbb{E}\left[\left|\overline{\mathbf{k}}\right|^2\right]$.

Step 4. We show that

$$\mathbb{E}\left[3^{-4n} \fint_{\square_{2n}} |w(x)|^2\,dx\right] \le C'K_0^2 3^{-n\beta/(1+\beta)}. \tag{6.26}$$

Let $\phi \in H^2_{\mathrm{loc}}(\mathbb{R}^d)$ denote the unique solution of

$$\begin{cases} -\Delta\phi = w \quad \text{in } \mathbb{R}^d, \\ \fint_{\square_{2n}} \phi(x)\,dx = 0, \\ \phi \quad \text{is } \square_{2n}\text{–periodic.} \end{cases}$$

By integration by parts, we have the identities

$$\begin{aligned}
\int_{\square_{2n}} |\nabla\nabla\phi(x)|^2\,dx = \int_{\square_{2n}} |w(x)|^2\,dx &= \int_{\square_{2n}} \nabla\phi(x)\cdot\mathbf{f}(x)\,dx \\
&= \int_{\square_{2n}} \nabla\phi(x)\cdot\left(\mathbf{f}(x) - \mathbb{E}\left[\fint_{\square_n} \mathbf{f}(x)\,dx\right]\right)dx.
\end{aligned} \tag{6.27}$$

We need a second mesoscale, given by an integer $k \in (n, 2n)$ to be selected below. In what follows, we denote $(\nabla\phi)_z := \fint_{z+\square_k} \nabla\phi(x)\,dx$ and $\sum_z = \sum_{z\in 3^k\mathbb{Z}^d\cap\square_{2n}}$ as well as $\widetilde{\mathbf{f}} := \mathbf{f} - \mathbb{E}\left[\fint_{\square_n} \mathbf{f}(x)\,dx\right]$. Now we estimate, by the Poincaré inequality:

$$\begin{aligned}
\int_{\square_{2n}} |w(x)|^2\,dx &= \int_{\square_{2n}} \nabla\phi(x)\cdot\widetilde{\mathbf{f}}(x)\,dx \\
&= \sum_z \left(\int_{z+\square_k} (\nabla\phi(x) - (\nabla\phi)_z)\cdot\widetilde{\mathbf{f}}(x)\,dx + (\nabla\phi)_z \cdot \int_{z+\square_k} \widetilde{\mathbf{f}}(x)\,dx\right) \\
&\le C\sum_z 3^k \left(\int_{z+\square_k} |\nabla\nabla\phi(x)|^2\,dx\right)^{\frac12} \left(\int_{z+\square_k} |\mathbf{f}(x)|^2\,dx\right)^{\frac12} \\
&\qquad + C\sum_z (\nabla\phi)_z \cdot \int_{z+\square_k} \widetilde{\mathbf{f}}(x)\,dx.
\end{aligned}$$

To estimate the first sum on the right side of the previous inequality, we use the discrete Hölder inequality, (6.27) and (4.18), and then Young's inequality:

$$\begin{aligned}
&\sum_z 3^k \left(\int_{z+\square_n} |\nabla\nabla\phi(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{z+\square_n} \left|\widetilde{\mathbf{f}}(x)\right|^2 dx \right)^{\frac{1}{2}} \\
&\quad \le 3^k \left(\sum_z \int_{z+\square_n} |\nabla\nabla\phi(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\sum_z \int_{z+\square_n} \left|\widetilde{\mathbf{f}}(x)\right|^2 dx \right)^{\frac{1}{2}} \\
&\quad = 3^k \left(\int_{\square_{2n}} |\nabla\nabla\phi(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{\square_{2n}} \left|\widetilde{\mathbf{f}}(x)\right|^2 dx \right)^{\frac{1}{2}} \\
&\quad \le 3^k \left(\int_{\square_{2n}} |w(x)|^2 \, dx \right)^{\frac{1}{2}} \left(CK_0^2 \, |\square_{2n}| \right)^{\frac{1}{2}} \\
&\quad \le \frac{1}{4} \int_{\square_{2n}} |w(x)|^2 \, dx + CK_0^2 \, |\square_{2n}| \, 3^{2k}.
\end{aligned}$$

We next estimate the expectation of the second sum. Using Hölder's inequality in two different forms, we get

$$\begin{aligned}
&\mathbb{E}\left[\sum_z (\nabla\phi)_z \cdot \int_{z+\square_k} \widetilde{\mathbf{f}}(x) \, dx \right] \\
&\quad = \sum_z \mathbb{E}\left[(\nabla\phi)_z \cdot \int_{z+\square_k} \widetilde{\mathbf{f}}(x) \, dx \right] \\
&\quad \le \sum_z \mathbb{E}\left[|(\nabla\phi)_z|^2 \right]^{\frac{1}{2}} \mathbb{E}\left[\left(\int_{z+\square_k} \widetilde{\mathbf{f}}(x) \, dx \right)^2 \right]^{\frac{1}{2}} \\
&\quad \le \left(\sum_z \mathbb{E}\left[|(\nabla\phi)_z|^2 \right] \right)^{\frac{1}{2}} \left(\sum_z \mathbb{E}\left[\left(\int_{z+\square_k} \widetilde{\mathbf{f}}(x) \, dx \right)^2 \right] \right)^{\frac{1}{2}}.
\end{aligned} \tag{6.28}$$

For the first factor on the right side of the previous inequality, we have, by the Poincaré inequality and (6.27),

$$\begin{aligned}
\sum_z \mathbb{E}\left[|(\nabla\phi)_z|^2 \right] = \mathbb{E}\left[\sum_z |(\nabla\phi)_z|^2 \right] &\le C3^{-kd} \, \mathbb{E}\left[\int_{\square_{2n}} |\nabla\phi(x)|^2 \, dx \right] \\
&\le C3^{-kd+4n} \, \mathbb{E}\left[\int_{\square_{2n}} |\nabla\nabla\phi(x)|^2 \, dx \right] \\
&= C3^{-kd+4n} \, \mathbb{E}\left[\int_{\square_{2n}} |w(x)|^2 \, dx \right].
\end{aligned} \tag{6.29}$$

In preparation to estimate the second factor, we use the mixing condition in the form of (6.6) to get

$$\begin{aligned}
\mathbb{E}\left[\left(\int_{\square_k} \widetilde{\mathbf{f}}(x) \, dx \right)^2 \right] &= \mathbb{E}\left[\sum_{y,y' \in 3^n\mathbb{Z}^d \cap \square_k} \int_{y+\square_n} \widetilde{\mathbf{f}}(x) \, dx \int_{y'+\square_n} \widetilde{\mathbf{f}}(x) \, dx \right] \\
&\le C' \left(CK_0 \, |\square_n| \right)^2 \sum_{y,y' \in 3^n\mathbb{Z}^d \cap \square_k} (1 + |y - y'|)^{-\beta} \\
&= C' K_0^2 3^{2dk - \beta(k-n)}
\end{aligned}$$

By stationarity, the same estimate holds with $z+\square_k$ in place of $\square_k$ provided that the cube $z+\square_k$ does not touch $\partial\square_{2n}$. For the cubes which do touch the boundary of the macroscopic cube (and thus intersect the support of ζ), we use the following cruder, deterministic bound given by (4.18):

$$\left(\int_{\square_k} \widetilde{\mathbf{f}}(x)\,dx\right)^2 \le C3^{dk}\int_{\square_k}\left|\widetilde{\mathbf{f}}(x)\right|^2\,dx \le CK_0^2 3^{2dk}$$

Combining these, using that there are at most $C3^{(2n-k)(d-1)}$ cubes of the form $z+\square_k$ which touch the boundary of $\square_n$, we get

$$\sum_z \mathbb{E}\left[\left(\int_{z+\square_k}\widetilde{\mathbf{f}}(x)\,dx\right)^2\right] \le C'K_0^2 3^{2dn+dk-\beta(k-n)} + CK_0^2 3^{2dn+dk-(2n-k)}. \tag{6.30}$$

We may now estimate the right side of (6.28) using applying (6.29), (6.30) and Young's inequality. The result is:

$$\begin{aligned}\mathbb{E}\left[\sum_z (\nabla\phi)_z\cdot\int_{z+\square_k}\widetilde{\mathbf{f}}(x)\,dx\right] \\ \le \frac14\mathbb{E}\left[\int_{\square_{2n}}|w(x)|^2\,dx\right] + K_0^2\left|\square_{2n}\right|3^{4n}\left(C'3^{-\beta(k-n)}+C3^{-(2n-k)}\right)\end{aligned} \tag{6.31}$$

Combining the above inequalities now yields

$$\mathbb{E}\left[\int_{\square_{2n}}|w(x)|^2\,dx\right] \le K_0^2 3^{4n}\left|\square_{2n}\right|\left(C'3^{-\beta(k-n)}+C3^{-(2n-k)}\right).$$

Taking finally k to be the nearest integer to $(2n+n\beta)/(1+\beta)$, we obtain (6.26).

Step 5. We estimate the expected contribution of $|\nabla h|^2$ by a computation which bears a resemblance to the one in the previous step. The claim is that

$$\mathbb{E}\left[\fint_{\square_{2n}}|\nabla h(x)|^2\,dx\right] \le C(\tau_n+\kappa_n). \tag{6.32}$$

Here we use the abbreviations $\mathbf{g}_z := \mathbf{g}(\cdot, z+\square_{n+1})$, $\psi_z := \psi(\cdot - z)$, $\sum_z := \sum_{z\in 3^n\mathbb{Z}^d\cap\square_{2n}}$ and $\fint_z := \fint_{z+\square_{n+1}}$. Observe that, in the sense of distributions, we have

$$\nabla\cdot\mathbf{k} = \sum_z \zeta\nabla\psi_z\cdot\left(\mathbf{g}_z-\overline{Q}_n\right) + \sum_z \psi_z\nabla\zeta\cdot\left(\mathbf{g}_z-\overline{Q}_n\right) \quad \text{in } \square_{2n}. \tag{6.33}$$

In particular, the right-hand side belongs to $L^2(\square_{2n})$ and thus $h\in H^2_{\mathrm{per}}(\square_{2n})$. Using the identity

$$\sum_z \zeta(x)\nabla\psi_z(x) = 0, \quad x\in\mathbb{R}^d, \tag{6.34}$$

which follows from (6.15) and (6.18), we may re-express the previous identity, for $x\in\square_{2n}$, as

$$\begin{aligned}(\nabla\cdot\mathbf{k})(x) = \sum_z \zeta(x)\nabla\psi_z(x)\cdot\left(\mathbf{g}_z(x)-\overline{Q}_n-\mathbf{k}(x)\right) \\ + \sum_z \psi_z(x)\nabla\zeta(x)\cdot\left(\mathbf{g}_z(x)-\overline{Q}_n\right).\end{aligned} \tag{6.35}$$

For each $i \in \{1, \dots, d\}$, let $\phi_i \in H^3_{\mathrm{loc}}(\mathbb{R}^d)$ denote the unique solution of

$$\begin{cases} -\Delta \phi_i = \partial_i h & \text{in } \square_{2n}, \\ \fint_{\square_{2n}} \phi_i(x)\,dx = 0, \\ \phi_i \text{ is } \square_{2n}\text{–periodic.} \end{cases}$$

Then integrating by parts, we find

$$\int_{\square_{2n}} |\nabla \nabla \phi_i(x)|^2\,dx = \int_{\square_{2n}} |\partial_i h(x)|^2\,dx = \int_{\square_{2n}} \partial_i \phi_i(x)\,(\nabla \cdot \mathbf{k})\,(x)\,dx. \tag{6.36}$$

We continue the computation by substituting (6.35) and using the notation $(\partial_i \phi_i)_z := \fint_{z+\square_{n+1}} \partial_i \phi_i(x)\,dx$:

$$\begin{aligned} &\int_{\square_{2n}} \partial_i \phi_i(x)\,(\nabla \cdot \mathbf{k})\,(x)\,dx \\ &\quad = \sum_z \int_z \left(\partial_i \phi_i(x) - (\partial_i \phi_i)_z\right) \zeta(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right) dx \\ &\qquad + \sum_z (\partial_i \phi_i)_z \int_z \zeta(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right) dx \\ &\qquad + \sum_z \int_z \partial_i \phi_i(x) \psi_z(x) \nabla \zeta(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n\right) dx. \end{aligned}$$

We put the second sum on the right side of the previous expression into a more convenient form using (6.21), (6.15), integration by parts and (6.34) twice:

$$\begin{aligned} &\sum_z (\partial_i \phi_i)_z \int_z \zeta(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right) dx \\ &\quad = \sum_{y,z} (\partial_i \phi_i)_z \int_z \zeta(x) \psi_y(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n - \zeta(x)\left(\mathbf{g}_y(x) - \overline{Q}_n\right)\right) dx \\ &\quad = \sum_{y,z} \left((\partial_i \phi_i)_z - (\partial_i \phi_i)_y\right) \int_z -(\zeta(x))^2 \psi_y(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_y(x) - \overline{Q}_n\right) dx \\ &\qquad + \sum_z (\partial_i \phi_i)_z \int_z -\psi_z(x) \nabla \zeta(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n\right) dx. \end{aligned}$$

Combining the previous two identities yields

$$\begin{aligned} &\int_{\square_{2n}} \partial_i \phi_i(x)\,(\nabla \cdot \mathbf{k})\,(x)\,dx \\ &\quad = \sum_z \int_z \left(\partial_i \phi_i(x) - (\partial_i \phi_i)_z\right) \zeta(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right) dx \\ &\qquad + \sum_{y,z} \left((\partial_i \phi_i)_y - (\partial_i \phi_i)_z\right) \int_z (\zeta(x))^2 \psi_y(x) \nabla \psi_z(x) \cdot \left(\mathbf{g}_y(x) - \overline{Q}_n\right) dx \\ &\qquad + \sum_z \int_z \left(\partial_i \phi_i(x) - (\partial_i \phi_i)_z\right) \psi_z(x) \nabla \zeta(x) \cdot \left(\mathbf{g}_z(x) - \overline{Q}_n\right) dx. \end{aligned} \tag{6.37}$$

We estimate the three sums on the right side of (6.37) in a similar fashion, each in turn. For the first sum, we use (6.16), the Hölder, discrete Hölder and

Poincaré inequalities, (6.36) and Young's inequality to deduce that

$$\begin{aligned}
&\sum_z \int_z (\partial_i\phi_i(x) - (\partial_i\phi_i)_z)\,\zeta(x)\left(\nabla\psi_z(x)\cdot\left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right)\right)dx \\
&\qquad \le C\left(\sum_z \int_z |\nabla\nabla\phi_i(x)|^2\,dx\right)^{\frac12}\left(\sum_z \int_z \left|\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right|^2 dx\right)^{\frac12} \\
&\qquad \le \frac14 \int_{\square_{2n}} |\partial_i h(x)|^2\,dx + C\sum_z \int_z \left|\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right|^2 dx.
\end{aligned}$$

Taking expectations and using (6.22), we get

$$\begin{aligned}
(6.38)\quad &\mathbb{E}\left[\sum_z \int_z (\partial_i\phi_i(x) - (\partial_i\phi_i)_z)\left(\nabla\psi_z(x)\cdot\left(\mathbf{g}_z(x) - \overline{Q}_n - \mathbf{k}(x)\right)\right)dx\right] \\
&\qquad\qquad \le \frac14 \mathbb{E}\left[\int_{2n} |\partial_i h(x)|^2\,dx\right] + C\,|\square_{2n}|\,\tau_n.
\end{aligned}$$

For the second sum, we notice that each entry vanishes unless $y \in z + \square_{n+2}$, there are at most C such entries y in the sum for any given entry z, and for such y and z, the Poincaré inequality gives

$$\left|(\partial_i\phi_i)_y - (\partial_i\phi_i)_z\right|^2 \le C3^{2n}|\,\square_n\,|^{-1}\int_{z+\square_{n+3}} |\nabla\nabla\phi_i(x)|^2\,dx.$$

Using this, (6.16), (6.36) and (4.18), the Hölder and Young inequalities and the fact that $3^n\mathbb{Z}^d \cap \square_{2n}$ has $C3^{nd}$ elements, we get

$$\begin{aligned}
&\sum_{y,z}\left((\partial_i\phi_i)_y - (\partial_i\phi_i)_z\right)\int_z (\zeta(x))^2\psi_y(x)\nabla\psi_z(x)\cdot\left(\mathbf{g}_y(x) - \overline{Q}_n\right)dx \\
&\qquad \le C\left(3^{2n}|\,\square_n\,|^{-1}\sum_z \int_{z+\square_{n+3}} |\nabla\nabla\phi_i(x)|^2\,dx\right)^{\frac12}\left(\sum_z 3^{-2n}CK_0^2|\,\square_n\,|\right)^{\frac12} \\
&\qquad \le \frac14\int_{\square_{2n}} |\partial_i h(x)|^2\,dx + CK_0^2 3^{nd}.
\end{aligned}$$

For the third sum on the right side of (6.37), we proceed in almost the same way as for the first two, except that rather than use (6.16) we use the estimate for $\nabla\zeta$ in (6.18) and the fact that $\nabla\zeta$ vanishes except if $z \notin \mathcal{Z}_n$ (recall that $\mathcal{Z}_n$ is defined in (6.25)) and there are at most $C3^{n(d-1)}$ such elements in the sum. We obtain:

$$\begin{aligned}
&\sum_z \int_z (\partial_i\phi_i(x) - (\partial_i\phi_i)_z)\,\psi_z(x)\nabla\zeta(x)\cdot\left(\mathbf{g}_z(x) - \overline{Q}_n\right)dx \\
&\qquad \le C\left(3^{2n}\sum_z \int_z |\nabla\nabla\phi_i(x)|^2\,dx\right)^{\frac12}\left(\sum_{z\notin\mathcal{Z}_n} 3^{-2n}CK_0^2|\,\square_n\,|\right)^{\frac12} \\
&\qquad \le \frac14\int_{\square_{2n}} |\partial_i h(x)|^2\,dx + CK_0^2 3^{n(2d-1)}.
\end{aligned}$$

Combining the previous two inequalities with (6.36), (6.37) and (6.38) yields

$$\mathbb{E}\left[\int_{\square_{2n}} |\partial_i h(x)|^2\,dx\right] \le CK_0^2 3^{n(2d-1)} + C\,|\square_{2n}|\,\tau_n.$$

Dividing by $|\square_{2n}|$ gives (6.32).

Step 6. We estimate the expected size of $|\nabla \cdot \mathbf{S}|^2$, using a computation which is closely analogous to the one in Step 5. Indeed, the main difference from Step 5 is essentially notational. The estimate we will show is

$$\mathbb{E}\left[\fint_{\square_{2n}} |\nabla \cdot \mathbf{S}(x)|^2 \, dx \right] \leq C(\tau_n + \kappa_n). \tag{6.39}$$

As in the previous step, we use the abbreviations $u_z := u(\cdot, z + \square_{n+1})$, $\psi_z := \psi(\cdot - z)$, $\sum_z := \sum_{z \in 3^n \mathbb{Z}^d \cap \square_{2n}}$ and $\int_z := \int_{z+\square_{n+1}}$ Observe that, in the sense of distributions, for every $i, j \in \{1, \ldots, d\}$, we have

$$\begin{aligned} \partial_j f_i - \partial_i f_j = \sum_z \zeta \left(\partial_j \psi_z (\partial_i u_z - \overline{P}_{n,i}) - \partial_i \psi_z (\partial_j u_z - \overline{P}_{n,j}) \right) \\ + \sum_z \psi_z \left(\partial_j \zeta (\partial_i u_z - \overline{P}_{n,i})) - \partial_i \zeta (\partial_j u_z - \overline{P}_{n,j}) \right) \quad \text{in } \square_{2n} . \end{aligned}$$

The right side belongs to $L^2(\square_{2n})$, thus $\partial_j f_i - \partial_i f_j \in L^2(\square_{2n})$ and $S_{ij} \in H^2_{\mathrm{loc}}(\mathbb{R}^d)$. Using (6.34), we may also express the previous identity slightly differently as

$$\begin{aligned} (\partial_j f_i - \partial_i f_j)(x) = \sum_z \zeta(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \mathbf{f}(x) \right]_{ij} \\ + \sum_z \psi_z \left[\nabla \zeta, \nabla u_z - \overline{P}_n \right]_{ij} \quad \text{in } \square_{2n}, \end{aligned} \tag{6.40}$$

where we henceforth use the notation

$$[\mathbf{v}, \mathbf{w}]_{ij} := v_j w_i - v_i w_j$$

for indices $i, j \in \{1, \ldots, d\}$ and vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$ with entries (v_i) and (w_i), respectively. Next we define, for each $i \in \{1, \ldots, d\}$,

$$\sigma_i := -(\nabla \cdot \mathbf{S})_i = -\sum_{j=1}^d \partial_j S_{ij}.$$

It is evident that $\sigma_i \in H^1_{\mathrm{per}}(\square_{2n})$ and σ_i is a solution of the equation

$$-\Delta \sigma_i = -\sum_{j=1}^d \partial_j \left(\partial_j f_i - \partial_i f_j \right) \quad \text{in } \square_{2n} .$$

Since σ_i has zero mean in $\square_{2n}$, there exists $\rho_i \in H^3_{\mathrm{loc}}(\mathbb{R}^d)$, which is unique up to an additive constant, satisfying

$$\begin{cases} -\Delta \rho_i = \sigma_i & \text{in } \mathbb{R}^d, \\ \rho_i \text{ is } \square_{2n}\text{–periodic.} \end{cases}$$

We have the identities

$$\int_{\square_{2n}} |\nabla \nabla \rho_i(x)|^2 \, dx = \int_{\square_{2n}} |\sigma_i(x)|^2 \, dx = \int_{\square_{2n}} \nabla \rho_i(x) \cdot \nabla \sigma_i(x) \, dx. \tag{6.41}$$

Integrating by parts and using the equation for σ_i, we obtain

$$\int_{\square_{2n}} |\sigma_i(x)|^2 \, dx = \int_{\square_{2n}} \nabla \rho_i(x) \cdot \nabla \sigma_i(x) \, dx = \sum_{j=1}^d \int_{\square_{2n}} \partial_j \rho_i(x) \left(\partial_j f_i - \partial_i f_j \right) dx.$$

To further shorten the notation, in each of the following expressions we keep the sum over j implicit (note that i is not summed over) and set $(\partial_j \rho_i)_z :=$

$\fint_{z+\square_{n+1}} \partial_j \rho_i(x)\,dx$. Continuing then the computation by substituting (6.40), we obtain

$$\begin{aligned}
&\int_{\square_{2n}} \partial_j \rho_i(x) \left(\partial_j f_i - \partial_i f_j\right) dx \\
&\quad = \sum_z \int_z \left(\partial_j \rho_i(x) - \left(\partial_j \rho_i\right)_z\right) \zeta(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \mathbf{f}(x)\right]_{ij} dx \\
&\qquad + \sum_z \left(\partial_j \rho_i\right)_z \int_z \zeta(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \mathbf{f}(x)\right]_{ij} dx \\
&\qquad + \sum_z \int_z \partial_j \rho_i(x) \psi_z(x) \left[\nabla \zeta(x), \nabla u_z(x) - \overline{P}_n\right]_{ij} dx.
\end{aligned}$$

We put the second sum on the right side into a more convenient form via (6.19), (6.15), integration by parts and (6.34):

$$\begin{aligned}
&\sum_z \left(\partial_j \rho_i\right)_z \int_z \zeta(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \mathbf{f}(x)\right]_{ij} dx \\
&\quad = \sum_{y,z} \left(\partial_j \rho_i\right)_z \int_z \zeta(x) \psi_y(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \zeta(x)(\nabla u_y(x) - \overline{P}_n)\right]_{ij} dx \\
&\quad = \sum_{y,z} \left(\left(\partial_j \rho_i\right)_z - \left(\partial_j \rho_i\right)_y\right) \int_z -(\zeta(x))^2 \psi_y(x) \left[\nabla \psi_z(x), \nabla u_y(x) - \overline{P}_n\right]_{ij} dx \\
&\qquad + \sum_z \left(\partial_j \rho_i\right)_z \int_z -\psi_z(x) \left[\nabla \zeta(x), \nabla u_z(x) - \overline{P}_n\right]_{ij} dx.
\end{aligned}$$

Combining this with the previous identity, we get

$$\begin{aligned}
&\int_{\square_{2n}} |\sigma_i(x)|^2 \, dx \\
&\quad = \sum_z \int_z \left(\partial_j \rho_i(x) - \left(\partial_j \rho_i\right)_z\right) \zeta(x) \left[\nabla \psi_z(x), \nabla u_z(x) - \overline{P}_n - \mathbf{f}(x)\right]_{ij} dx \\
&\qquad + \sum_{y,z} \left(\left(\partial_j \rho_i\right)_y - \left(\partial_j \rho_i\right)_z\right) \int_z (\zeta(x))^2 \psi_y(x) \left[\nabla \psi_z(x), \nabla u_y(x) - \overline{P}_n\right]_{ij} dx \\
&\qquad + \sum_z \int_z \left(\partial_j \rho_i(x) - \left(\partial_j \rho_i\right)_z\right) \psi_z(x) \left[\nabla \zeta(x), \nabla u_z(x) - \overline{P}_n\right]_{ij} dx.
\end{aligned}$$

We may now compare this identity with (6.37) and observe that the three sums on the right side are similar to those on the right side of (6.37). In fact, following the arguments in Step 5 (with obvious substitutions, changing for instance ϕ_i to ρ_i, $\mathbf{g}_z - \overline{Q}_n$ to $\nabla u_z - \overline{P}_n$ and $\mathbf{k}$ to $\mathbf{f}$), we may bound these three sums in the same way. This completes the argument for (6.39).

Step 7. We show that the effect of the cutoff ξ in the definitions of v and $\mathbf{h}$ is expected to be small: precisely,

$$\begin{aligned}
\mathbb{E}\left[\fint_{\square_{2n}} |\nabla v(x) - \nabla w(x)|^2 \, dx\right] + \mathbb{E}\left[\fint_{\square_{2n}} |\mathbf{h}(x) - \nabla \cdot \mathbf{T}(x)|^2 \, dx\right] \\
\leq C\tau_n + C'\kappa_n.
\end{aligned} \tag{6.42}$$

We use the identity

$$\nabla v(x) - \nabla w(x) = w(x) \nabla \xi(x) + (\xi(x) - 1) \nabla w(x)$$

and (6.17) to obtain

$$\fint_{\square_{2n}} |\nabla v(x) - \nabla w(x)|^2 \, dx \leq C3^{-4n/(1+\delta)} \fint_{\square_{2n}} |w(x)|^2 \, dx + C \fint_{\square_{2n}} |\xi(x) - 1|^2 |\nabla w(x)|^2 \, dx. \tag{6.43}$$

The expectation of the first integral on the right side is controlled by (6.26):

$$\mathbb{E}\left[3^{-4n/(1+\delta)} \fint_{\square_{2n}} |w(x)|^2 \, dx \right] \leq CK_0^2 3^{-n\beta/(1+\beta)+4n\delta/(1+\delta)} \leq C\kappa_n,$$

provided we select

$$\delta := \frac{1}{16}\beta.$$

For the expectation of the second integral on the right side of (6.43), we recall from (6.17) that $\xi \equiv 1$ except in

$$D := \left\{ x \in \square_{2n} \,:\, \operatorname{dist}(x, \partial\square_{2n}) > C3^{2n/(1+\delta)} \right\}.$$

Therefore, using that D intersects at most $C3^{n(d-2\delta/(1+\delta))}$ subcubes of the form $z + \square_{n+1}$, with $z \in 3^n\mathbb{Z}^d$, and applying (4.18), (6.24) and (6.39), we obtain

$$\begin{aligned} \mathbb{E}\left[\fint_{\square_{2n}} |\xi(x) - 1|^2 |\nabla w(x)|^2 \, dx \right] &\leq \frac{1}{|\square_{2n}|} \mathbb{E}\left[\int_D |\nabla w(x)|^2 \, dx \right] \\ &\leq \frac{C}{|\square_{2n}|} \mathbb{E}\left[\int_D |\mathbf{f}(x)|^2 \, dx + \int_{\square_{2n}} |\mathbf{f}(x) - \nabla w(x)|^2 \, dx \right] \\ &\leq \frac{C}{|\square_{2n}|} \left(CK_0^2 3^{n(d-2\delta/(1+\delta))} |\square_n| + C|\square_{2n}| \left(\tau_n + C'\kappa_n \right) \right) \\ &\leq C\tau_n + C'\kappa_n. \end{aligned} \tag{6.44}$$

Combining the previous inequality with (6.26) and (6.43), we obtain the desired estimate for the first term on the left of (6.42).

The argument for estimating the second term on the left side of (6.42) is similar. We start from the identity

$$\mathbf{h}(x) - \nabla \cdot \mathbf{T}(x) = (\xi(x) - 1) \nabla \cdot \mathbf{T}(x) - \nabla \widetilde{h}(x)$$

and observe that the equation for $\widetilde{h}$ can be written as

$$-\Delta \widetilde{h} = -\nabla \cdot ((\xi - 1)\nabla \cdot \mathbf{T}) \quad \text{in } \boxdot_{2n},$$

and therefore we have

$$\int_{\boxdot_{2n}} \left| \nabla \widetilde{h}(x) \right|^2 dx \leq C \int_{\boxdot_{2n}} |\xi(x) - 1|^2 |\nabla \cdot \mathbf{T}(x)|^2 \, dx.$$

It thus suffices to show that

$$\mathbb{E}\left[\int_{\square_{2n}} |\xi(x) - 1|^2 |\nabla \cdot \mathbf{T}(x)|^2 \, dx \right] \leq C\tau_n + C'\kappa_n.$$

The proof of this estimate is very similar to (6.44), and so we omit the details.

Step 8. We estimate the expected difference in $L^2(z+\square_n)$ between $(\nabla v, \mathbf{h})$ and $(\nabla u, \mathbf{g})(\cdot, z+\square_{n+1})$ for each $z \in 3^n\mathbb{Z}^d \cap \boxdot_{2n}$. The claim is that

$$\mathbb{E}\left[3^{-dn} \sum_{z\in 3^n\mathbb{Z}^d\cap\boxdot_{2n}} \fint_{z+\square_n} \left|\nabla v(x) - \nabla u(x, z+\square_{n+1}) + \overline{P}_n\right|^2 dx\right] + \mathbb{E}\left[3^{-dn} \sum_{z\in 3^n\mathbb{Z}^d\cap\boxdot_{2n}} \fint_{z+\square_n} \left|\mathbf{h}(x) - \mathbf{g}(x, z+\square_{n+1}) + \overline{Q}_n\right|^2 dx\right] \le C(\tau_n + C'\kappa_n). \tag{6.45}$$

Indeed, for each $z \in 3^n\mathbb{Z}^d \cap \boxdot_{2n}$ and $x \in z+\square_n$, we have

$$\begin{aligned}\nabla v(x) - \nabla u(x, z+\square_{n+1}) + \overline{P}_n &\\ = (\nabla v(x) - \nabla w(x)) + \left(\mathbf{f}(x) - \nabla u(x, z+\square_{n+1}) + \overline{P}_n\right) - \left(\overline{\mathbf{f}} - \nabla\cdot\mathbf{S}\right).&\end{aligned}$$

Thus the estimate for the first term on the left of (6.45) is a consequence of (6.22) (note that $\zeta \equiv 1$ on $z+\square_{n+1}$ for every $z \in 3^n\mathbb{Z}^d \cap \boxdot_{2n}$), (6.24), (6.39) and (6.42). The estimate for the other term follows immediately from (6.22), (6.24), (6.32) and (6.42).

Step 9. We complete the argument by deriving (6.14). By Lemma 4.3, we have, for each $z \in 3^n\mathbb{Z}^d \cap \boxdot_{2n}$,

$$\begin{aligned}\fint_{z+\square_n} & F(\overline{P}_n + \nabla v(x), \overline{Q}_n + \mathbf{h}(x), x)\,dx \\ &\le 2\fint_{z+\square_n} F(\nabla u(x, z+\square_{n+1}), \mathbf{g}(x, z+\square_{n+1}), x)\,dx - \mu(z+\square_n) \\ &\quad + C\fint_{z+\square_n} \Big(\left|\nabla v(x) - \nabla u(x, z+\square_{n+1}) + \overline{P}_n\right|^2 \\ &\qquad + \left|\mathbf{h}(x) - \mathbf{g}(x, z+\square_{n+1}) + \overline{Q}_n\right|^2\Big)\,dx.\end{aligned}$$

In view of (6.17), it is convenient to denote $\mathcal{Z}_n' := \{z \in 3^n\mathbb{Z}^d : z+\square_{n+1} \notin \boxdot_{2n}\}$ and $U := \cup_{z\in\mathcal{Z}_n'}(z+\square_n)$. Note that ξ vanishes on $\boxdot_{2n}\setminus U$ and thus v and $\mathbf{h}$ do as well. Observe also that

$$|\boxdot_{2n}\setminus U| \le C3^{-n}|\boxdot_{2n}| \quad \text{and} \quad \left||\mathcal{Z}_n'|\,|\square_n| - |\boxdot_{2n}|\right| \le C3^{-n}|\boxdot_{2n}|.$$

Now we take the expectation of the previous inequality, sum over $z \in \mathcal{Z}_n'$, using (4.5), (4.6), (4.18), (6.23), Lemma 4.3, (6.45), stationarity and the above observations to obtain

$$\begin{aligned}\mathbb{E}\left[\int_{\boxdot_{2n}} F(\overline{P}_n + \nabla v(x), \overline{Q}_n + \mathbf{h}(x), x)\,dx\right] &\\ \le \sum_{z\in\mathcal{Z}_n'} \mathbb{E}\left[\int_{z+\square_n} F(\overline{P}_n + \nabla v(x), \overline{Q}_n + \mathbf{h}(x), x)\,dx\right] &\\ + \mathbb{E}\left[\int_{\boxdot_{2n}\setminus U} F(\overline{P}_n, \overline{Q}_n, x)\,dx\right] &\\ \le |\boxdot_{2n}|\left(\mathbb{E}\left[\mu(\square_n)\right] + C(\tau_n + C'\kappa_n)\right).&\end{aligned}$$

Dividing by $|\boxdot_{2n}|$ yields (6.14) and completes the proof of the lemma. □

6.3. **Proof of Proposition 6.1.** We use the flatness theory developed in the previous subsection to prove Proposition 6.1. We begin by iterating Lemma 6.3 to obtain a rate of convergence for $\mathbb{E}\left[\mu(\boxdot_n)\right]$ and $\mathbb{E}\left[\mu_0(\boxdot_n, \overline{P}, \overline{Q})\right]$ as $n \to \infty$ to their common limit $\overline{\mu}$.

Lemma 6.4. *There exist $\alpha(d, \Lambda, \beta) > 0$, $C(d, \Lambda, \beta, C_3) \geq 1$ and $\overline{P}, \overline{Q} \in \mathbb{R}^d$ such that*

$$\overline{\mu} = \overline{\mu}_0(\overline{P}, \overline{Q}) \tag{6.46}$$

and, for every $n \in \mathbb{N}$,

$$\left|\mathbb{E}\left[\mu(\boxdot_n)\right] - \overline{\mu}\right| + \left|\mathbb{E}\left[\mu_0(\boxdot_n, \overline{P}, \overline{Q})\right] - \overline{\mu}\right| \leq CK_0^2 3^{-n\alpha}. \tag{6.47}$$

Proof. We split the proof into three steps. The convention for constants is different from the previous subsection. Here $C \geq 1$ denotes a constant depending only on (d, Λ, β, C_3), while $\widetilde{C} \geq 1$ and $0 < \widetilde{c} < 1$, which are used only in Step 1, are allowed to depend on only (d, Λ, β). These may vary in each occurrence.

Step 1. We show that there exists $\alpha(d, \Lambda, \beta) > 0$ such that, for every $n \in \mathbb{N}$,

$$\overline{\mu} - \mathbb{E}\left[\mu(\boxdot_n)\right] \leq CK_0^2 3^{-n\alpha}. \tag{6.48}$$

By (4.12), (4.13) and Lemma 6.3, we have

$$\overline{\mu} \leq \mathbb{E}\left[\mu_0(\boxdot_{2n}, \overline{P}_n, \overline{Q}_n)\right] \leq \mathbb{E}\left[\mu(\square_n)\right] + \left(\widetilde{C}\tau_n + C\kappa_n\right), \tag{6.49}$$

which by (4.27) can be upgraded to

$$\overline{\mu} \leq \mathbb{E}\left[\mu(\boxdot_n)\right] + \left(\widetilde{C}\tau_n + C\kappa_n\right).$$

Let $\mu_n := \mathbb{E}\left[\mu(\boxdot_n)\right]$. Recalling the definition of τ_n in (6.5) and using (4.23) yields

$$\overline{\mu} \leq \mathbb{E}\left[\mu(\boxdot_n)\right] + \widetilde{C}\left(\mathbb{E}\left[\mu(\boxdot_{n+1})\right] - \mathbb{E}\left[\mu(\boxdot_n)\right]\right) + C\kappa_n.$$

Denoting $\widetilde{\mu}_n := \overline{\mu} - \mathbb{E}\left[\mu(\boxdot_n)\right]$, we can rewrite this as

$$\widetilde{\mu}_{n+1} - \widetilde{\mu}_n \leq -\widetilde{C}^{-1}\widetilde{\mu}_n + CK_0^2 3^{-n\beta/(1+\beta)}.$$

Letting $\widetilde{c} := 1 - \widetilde{C}^{-1} < 1$, we arrive at

$$\frac{\widetilde{\mu}_{n+1}}{\widetilde{c}^{n+1}} - \frac{\widetilde{\mu}_n}{\widetilde{c}^n} \leq \frac{C}{\widetilde{c}^{n+1}} K_0^2 3^{-n\beta/(1+\beta)}.$$

Summing the telescopic series, we get the discrete Duhamel formula

$$\widetilde{\mu}_n \leq \widetilde{c}^n \widetilde{\mu}_0 + CK_0^2 \sum_{k=0}^{n-1} \widetilde{c}^{n-k-1} 3^{-k\beta/(1+\beta)},$$

and (6.48) follows since $\widetilde{\mu}_0 \leq CK_0^2$.

Notice that by (4.16) and (4.23), (6.48) implies that for every $n \in \mathbb{N}$,

$$\left|\mathbb{E}\left[\mu(\boxdot_n)\right] - \overline{\mu}\right| + \left|\mathbb{E}\left[\mu(\square_n)\right] - \overline{\mu}\right| \leq CK_0^2 3^{-n\alpha}. \tag{6.50}$$

In particular, $\tau_n \leq CK_0^2 3^{-n\alpha}$ by the triangle inequality.

Step 2. We argue that there exist $\overline{P}, \overline{Q} \in \mathbb{R}^d$ such that

$$\left|\overline{P} - \overline{P}_n\right| + \left|\overline{Q} - \overline{Q}_n\right| \leq CK_0^2 3^{-n\alpha}. \tag{6.51}$$

By Lemma 4.2,

$$\int_{\square_n} |\nabla u(x,\square_{n+1}) - \nabla u(x,\square_n)|^2\,dx \\ \leq C\left(\int_{\square_n} F(\nabla u(x,\square_{n+1}),\mathbf{g}(x,\square_{n+1}),x)\,dx - \mu(\square_n)\right).$$

By (6.23), the latter is bounded by a multiple of τ_n. By (6.23) and (6.50), we thus have

$$|\mathbb{E}[P(\square_{n+1})] - \mathbb{E}[P(\square_n)]| + |\mathbb{E}[Q(\square_{n+1})] - \mathbb{E}[Q(\square_n)]| \leq CK_0^2 3^{-n\alpha}.$$

This implies that $\{\mathbb{E}[P(\square_n)]\}_{n\in\mathbb{N}}$ and $\{\mathbb{E}[Q(\square_n)]\}_{n\in\mathbb{N}}$ are Cauchy sequences in $\mathbb{R}^d$, and thus there exist $\overline{P},\overline{Q}\in\mathbb{R}^d$ such that

$$\mathbb{E}[P(\square_n)] \to \overline{P} \quad \text{and} \quad \mathbb{E}[Q(\square_n)] \to \overline{Q} \quad \text{as } n\to\infty. \tag{6.52}$$

Moreover, we have

$$\left|\overline{P} - \mathbb{E}[P(\square_n)]\right| \leq \sum_{k=n}^{+\infty} |\mathbb{E}[P(\square_{k+1})] - \mathbb{E}[P(\square_k)]| \leq CK_0^2 3^{-n\alpha},$$

and similarly

$$\left|\overline{Q} - \mathbb{E}[Q(\square_n)]\right| \leq CK_0^2 3^{-n\alpha}.$$

To complete the proof of (6.51), it remains to show that

$$\left|\overline{P}_n - \mathbb{E}[P(\square_n)]\right| + \left|\overline{Q}_n - \mathbb{E}[Q(\square_n)]\right| \leq CK_0^2 3^{-n\alpha}.$$

This follows from (6.12) and (6.50).

For future reference, we notice that (4.18) also implies that

$$\left|\overline{P}\right| + \left|\overline{Q}\right| \leq CK_0^2. \tag{6.53}$$

Step 3. By redefining $\alpha(d,\Lambda,\beta) > 0$, if necessary, we argue that, for every $n\in\mathbb{N}$,

$$\left|\mathbb{E}\left[\mu_0(\boxdot_n,\overline{P},\overline{Q})\right] - \mathbb{E}[\mu(\boxdot_n)]\right| \leq CK_0^2 3^{-n\alpha}. \tag{6.54}$$

We have

$$\mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P},\overline{Q})\right] - \mathbb{E}[\mu(\boxdot_{2n})] \\ \leq \left|\mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P},\overline{Q})\right] - \mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P}_n,\overline{Q}_n)\right]\right| \\ + \left(\mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P}_n,\overline{Q}_n)\right] - \mathbb{E}[\mu(\boxdot_n)]\right) + \left(\mathbb{E}[\mu(\boxdot_n)] - \mathbb{E}[\mu(\boxdot_{2n})]\right).$$

The first term is estimated by (4.31), (6.51) and (6.53):

$$\left|\mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P},\overline{Q})\right] - \mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P}_n,\overline{Q}_n)\right]\right| \leq CK_0^2 3^{-n\alpha}.$$

The second term is estimated by (6.49) and (6.50) and the third term by (6.50). Combining these, we deduce

$$\left|\mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P},\overline{Q})\right] - \mathbb{E}[\mu(\boxdot_{2n})]\right| = \mathbb{E}\left[\mu_0(\boxdot_{2n},\overline{P},\overline{Q})\right] - \mathbb{E}[\mu(\boxdot_{2n})] \leq CK_0^2 3^{-n\alpha}.$$

Replacing α with $\alpha/2$ yields (6.54).

Notice that (6.54) implies $\overline{\mu} = \overline{\mu}_0(\overline{P}, \overline{Q})$, that is, we have proved (6.46). It also implies, by (6.54) and the triangle inequality, that

$$\left|\mathbb{E}\left[\mu_0(\square_n, \overline{P}, \overline{Q})\right] - \overline{\mu}\right| \leq CK_0^2 3^{-n\alpha}. \tag{6.55}$$

This completes the proof of the lemma. □

We next upgrade the convergence from the previous lemma, using the mixing conditions, to obtain estimates on stochastic moments of $|\mu - \overline{\mu}|$ and $|\mu_0 - \overline{\mu}|$ in the triadic cubes. Recall that $\mathcal{E}_*$ is defined in (6.3).

Lemma 6.5. *For every $\theta \in (0, \beta)$, there exist an exponent $s_0(d, \Lambda, \beta, \alpha) \geq 1$ and a constant $C(d, \Lambda, \beta, C_3, \alpha) \geq 1$ such that, for every $n \in \mathbb{N}$ and $s \geq s_0$,*

$$\mathbb{E}\left[|\mathcal{E}_*(\square_n)|^s\right] \leq \left(CK_0^2\right)^s 3^{-\theta n}. \tag{6.56}$$

Under the additional assumption that (P4) holds, there exist $\alpha(d, \Lambda, \beta, \gamma) > 0$ and $C(d, \Lambda, \beta, C_3, \gamma, C_4) \geq 1$ such that, for every $n \in \mathbb{N}$,

$$\mathbb{E}\left[\exp\left(3^{nd\gamma/(d+\gamma)} \frac{\mathcal{E}_*(\square_n)}{CK_0^2}\right)\right] \leq \exp\left(C3^{n(d\gamma/(d+\gamma)-\alpha)}\right). \tag{6.57}$$

Proof. In this argument, we drop the dependence of μ_0 on $(\overline{P}, \overline{Q})$ for convenience and denote

$$\sigma_n := K_0^2 3^{-n\alpha}, \tag{6.58}$$

where the exponent $\alpha > 0$ implicit in σ_n depends only on (d, Λ, β) and may change in each occurrence. Here C depends only on (d, Λ, β, C_3) and may vary in each occurrence.

The first four steps are concerned with the proof of the first statement. Note that (6.56) is much easier to prove under slightly stronger (although still relatively weak) mixing conditions than (P3), such as for example "ψ-mixing." We have to work a bit because "α–mixing" is a very weak condition.

Step 1. We claim that, for every $n \in \mathbb{N}$, at least one of the following two estimates must hold:

$$\mathbb{E}\left[(\overline{\mu} - \mu(\square_n))_+\right]^2 \leq \frac{5}{8}\mathbb{E}\left[(\overline{\mu} - \mu(\square_n))_+^2\right] + C\sigma_n^2, \tag{6.59}$$

or

$$\mathbb{E}\left[(\mu_0(\square_n) - \overline{\mu})_+\right]^2 \leq \frac{5}{8}\mathbb{E}\left[(\mu_0(\square_n) - \overline{\mu})_+^2\right] + C\sigma_n^2. \tag{6.60}$$

(In fact, one can replace 5/8 by any number larger than 1/2.) We proceed with the proof of this alternative by noting that, since $\mu_0(\square_n) - \mu(\square_n) \geq 0$, Chebyshev's inequality yields, for every $t > 0$,

$$\mathbb{P}\left[\mu_0(\square_n) - \mu(\square_n) \geq t\sigma_n\right] \leq (t\sigma_n)^{-1}\mathbb{E}\left[\mu_0(\square_n) - \mu(\square_n)\right].$$

By Lemma 6.4, taking C sufficiently large, we obtain that

$$\mathbb{P}\left[\mu_0(\square_n) - \mu(\square_n) \geq C\sigma_n\right] \leq \frac{1}{8}.$$

It follows that

$$\mathbb{P}\left[\mu(\square_n) \leq \overline{\mu} - C\sigma_n \quad \text{and} \quad \mu_0(\square_n) \geq \overline{\mu} + C\sigma_n\right] \leq \frac{1}{8}.$$

This implies

$$\min\left\{\mathbb{P}\left[\mu(\boxdot_n) \le \overline{\mu} - C\sigma_n\right], \mathbb{P}\left[\mu_0(\boxdot_n) \ge \overline{\mu} + C\sigma_n\right]\right\} \le \frac{9}{16}. \tag{6.61}$$

From (6.61), we obtain the desired claim observing that for every nonnegative random variable X and $s > 0$,

$$\mathbb{E}\left[X\right] \le \mathbb{E}\left[X\mathbf{1}_{X>s}\right] + s \le \left(\mathbb{P}\left[X > s\right]\mathbb{E}\left[X^2\right]\right)^{1/2} + s$$

and applying Young's inequality.

Step 2. We show using the mixing condition (P3) that

$$\begin{aligned}\mathbb{E}\left[(\mu - \mu(\boxdot_{n+1}))_+^2\right] \\ \le \frac{6}{5}\left(3^{-d}\operatorname{var}\left[(\overline{\mu} - \mu(\boxdot_n))_+\right] + \mathbb{E}\left[(\mu - \mu(\boxdot_n))_+\right]^2\right) + C\sigma_n^2,\end{aligned} \tag{6.62}$$

and

$$\begin{aligned}\mathbb{E}\left[(\mu_0(\boxdot_{n+1}) - \overline{\mu})_+^2\right] \\ \le \frac{6}{5}\left(3^{-d}\operatorname{var}\left[(\mu_0(\boxdot_n) - \overline{\mu})_+\right] + \mathbb{E}\left[(\mu_0(\boxdot_n) - \overline{\mu})_+\right]^2\right) + C\sigma_n^2.\end{aligned} \tag{6.63}$$

(In fact, one can replace 6/5 by any number larger than 1.) The arguments for (6.62) and (6.63) are almost identical, so we only prove (6.62). Recall from (4.14) that

$$\mu(\boxdot_{n+1}) \ge 3^{-d}\sum_{\boxdot_n(x)\subseteq\boxdot_n}\mu(\boxdot_n(x)) - C\sigma_n,$$

hence

$$(\mu - \mu(\boxdot_{n+1}))_+ \le \underbrace{3^{-d}\sum_{\boxdot_n(x)\subseteq\boxdot_{n+1}}(\mu - \mu(\boxdot_n(x)))_+}_{=:\mathcal{S}_n} + C\sigma_n.$$

We have

$$\mathbb{E}\left[\mathcal{S}_n^2\right] = \operatorname{var}\left[\mathcal{S}_n\right] + \mathbb{E}\left[\mathcal{S}_n\right]^2,$$

and by stationarity, $\mathbb{E}[\mathcal{S}_n] = \mathbb{E}\left[(\mu - \mu(\boxdot_n))_+\right]$. In order to estimate the variance, we use (P3) in the form given by (6.6), (4.6) and (4.9) to find that, for every $\boxdot_n(x) \neq \boxdot_n(y)$,

$$\left|\operatorname{cov}\left[(\overline{\mu} - \mu(\boxdot_n(x)))_+, (\overline{\mu} - \mu(\boxdot_n(y)))_+\right]\right| \le CK_0^4 3^{-n\beta/(1+\beta)} \le C\sigma_n^2,$$

and therefore

$$\begin{aligned}\operatorname{var}[\mathcal{S}_n] = 3^{-2d}\sum_{\boxdot_n(x),\boxdot_n(y)\subseteq\boxdot_{n+1}}\operatorname{cov}\left[(\overline{\mu} - \mu(\boxdot_n(x)))_+, (\overline{\mu} - \mu(\boxdot_n(y)))_+\right] \\ \le 3^{-d}\operatorname{var}\left[(\overline{\mu} - \mu(\boxdot_n))_+\right] + C\sigma_n^2.\end{aligned}$$

Summarizing and using Young's inequality, we obtain (6.62).

Step 3. We show that for every $n \in \mathbb{N}$, at least one of the following two inequalities holds:

$$\mathbb{E}\left[(\overline{\mu} - \mu(\boxdot_{n+1}))_+^2\right] \le \frac{9}{10}\mathbb{E}\left[(\overline{\mu} - \mu(\boxdot_n))_+^2\right] + C\sigma_n^2 \tag{6.64}$$

or

$$\mathbb{E}\left[(\mu_0(\boxdot_{n+1})-\overline{\mu})_+^2\right] \le \frac{9}{10}\,\mathbb{E}\left[(\mu_0(\boxdot_{n+1})-\overline{\mu})_+^2\right] + C\sigma_n^2. \tag{6.65}$$

In fact, we claim that (6.59) implies (6.64) and (6.60) implies (6.65), and thus the alternative follows from the one in Step 1. Indeed, by (6.62), we have

$$\begin{aligned}
&\mathbb{E}\left[(\mu-\mu(\boxdot_{n+1}))_+^2\right]\\
&\quad\le \frac{6}{5}\left(3^{-d}\operatorname{var}\left[(\overline{\mu}-\mu(\boxdot_n))_+\right]+\mathbb{E}\left[(\mu-\mu(\boxdot_n))_+\right]^2\right)+C\,\sigma_n^2\\
&\quad= \frac{6}{5}\left(3^{-d}\mathbb{E}\left[(\overline{\mu}-\mu(\boxdot_n))_+^2\right]+(1-3^{-d})\mathbb{E}\left[(\mu-\mu(\boxdot_n))_+\right]^2\right)+C\,\sigma_n^2\\
&\quad\le \frac{6}{5}\left(3^{-d}+\frac{5}{8}(1-3^{-d})\right)\mathbb{E}\left[(\overline{\mu}-\mu(\boxdot_n))_+^2\right]+C\,\sigma_n^2,
\end{aligned}$$

where we assumed (6.59) to hold in the last step. This gives (6.64). We omit the proof of the second implication, which is very similar.

Step 4. We show that, for each fixed $n\in\mathbb{N}$,

$$\min\left\{\mathbb{E}\left[(\overline{\mu}-\mu(\boxdot_n))_+^2\right],\mathbb{E}\left[(\mu_0(\boxdot_n)-\overline{\mu})_+^2\right]\right\}\le C\sigma_n^2. \tag{6.66}$$

Let A_1 and A_2, respectively, be the subset of $\mathbb{N}$ consisting of n for which (6.64) and (6.65) hold. We have $\mathbb{N}=A_1\cup A_2$ and thus, for each $n\in\mathbb{N}$, at least one of the set $A_1\cap\{1,\dots,n\}$ and $A_2\cap\{1,\dots,n\}$ has at least $n/2$ elements. If the former, then by (4.16), (6.64), (4.6) and a simple computation, we have

$$\begin{aligned}
\mathbb{E}\left[(\overline{\mu}-\mu(\boxdot_n))_+^2\right] &\le \left(\frac{9}{10}\right)^{n/2}\mathbb{E}\left[(\overline{\mu}-\mu(\boxdot_0))_+^2\right]+C\sigma_n^2\\
&\le CK_0^4\left(\frac{9}{10}\right)^{n/2}+C\sigma_n^2.
\end{aligned}$$

If it is rather $A_2\cap\{1,\dots,n\}$, then we have a similar bound for $\mathbb{E}\left[(\mu_0(\boxdot_n)-\overline{\mu})_+^2\right]$. Since the first term on the right side is bounded by $C\sigma_n^2$ after a redefinition of α, we obtain (6.66).

Step 5. We show that

$$\mathbb{E}\left[\mathcal{E}_*(\boxdot_n)\right]\le CK_0^2 3^{-n\alpha}. \tag{6.67}$$

Observe, using $\mu_0(U)\le\mu(U)$ again, that we have, for any bounded domain $U\subseteq\mathbb{R}^d$,

$$|\mu_0(U)-\overline{\mu}|\le\mu_0(U)-\mu(U)+|\overline{\mu}-\mu(U)|$$

and

$$|\overline{\mu}-\mu(U)|\le\mu_0(U)-\mu(U)+|\mu_0(U)-\overline{\mu}|.$$

Combining the previous two inequalities for $U=\boxdot_n$, taking their expectation and applying Lemma 6.4 (in particular (6.54)), we get

$$\begin{aligned}
&\max\left\{\mathbb{E}\left[|\mu_0(\boxdot_n)-\overline{\mu}|\right],\mathbb{E}\left[|\overline{\mu}-\mu(\boxdot_n)|\right]\right\}\\
&\qquad\le\min\left\{\mathbb{E}\left[|\mu(\boxdot_n)-\overline{\mu}|\right],\mathbb{E}\left[|\mu_0(\boxdot_n)-\overline{\mu}|\right]\right\}+\mathbb{E}\left[\mu_0(\boxdot_n)-\mu(\boxdot_n)\right].
\end{aligned}$$

Note that a centered random variable X satisfies $\mathbb{E}[X_+] = \mathbb{E}[X_-]$, so that $\mathbb{E}[|X|] = 2\mathbb{E}[X_+]$. We use this observation and apply Lemma 6.4 twice to obtain

$$\begin{aligned}\mathbb{E}\left[\left|\overline{\mu} - \mu(\boxdot_n)\right|\right] &\le \mathbb{E}\left[\left|\mathbb{E}\left[\mu(\boxdot_n)\right] - \mu(\boxdot_n)\right|\right] + C\sigma_n \\ &= 2\mathbb{E}\left[\left(\mathbb{E}\left[\mu(\boxdot_n)\right] - \mu(\boxdot_n)\right)_+\right] + C\sigma_n \\ &\le 2\mathbb{E}\left[\left(\overline{\mu} - \mu(\boxdot_n)\right)_+\right] + C\sigma_n.\end{aligned}$$

Similarly, we also have

$$\mathbb{E}\left[\left|\mu_0(\boxdot_n) - \overline{\mu}\right|\right] \le 2\mathbb{E}\left[\left(\mu_0(\boxdot_n) - \overline{\mu}\right)_+\right] + C\sigma_n.$$

The previous three inequalities and (6.66) yield

$$\begin{aligned}\mathbb{E}\left[\mathcal{E}_*(\boxdot_n)\right] &\le 2\max\left\{\mathbb{E}\left[\left|\mu_0(\boxdot_n) - \overline{\mu}\right|\right], \mathbb{E}\left[\left|\overline{\mu} - \mu(\boxdot_n)\right|\right]\right\} \\ &\le 4\min\left\{\mathbb{E}\left[\left(\overline{\mu} - \mu(\boxdot_n)\right)_+\right], \mathbb{E}\left[\left(\mu_0(\boxdot_n) - \overline{\mu}\right)_+\right]\right\} + C\sigma_n \\ &\le C\sigma_n.\end{aligned}$$

This completes the proof of (6.67).

Step 6. We upgrade the convergence rate in (6.67) to the optimal one for large moments, using Lemma A.2, thereby completing the proof of the first statement of the lemma. The claim is that, for every $\theta \in (0,\beta)$, there exist an exponent $s_0(d,\Lambda,\beta,\theta) \ge 1$ and $C(d,\Lambda,\beta,\theta,C_3) \ge 1$ such that, for every $s \ge s_0$,

$$\mathbb{E}\left[\left|\mathcal{E}_*(\boxdot_n)\right|^s\right] \le \left(CK_0^2\right)^s 3^{-\theta n}. \tag{6.68}$$

In this step, we allow the constant C to depend additionally on θ. We also allow $\alpha(d,\Lambda,\beta) > 0$ to vary in each occurrence, as usual.

Fix $n, m \in \mathbb{N}$. Also fix $k \in \mathbb{N}$, which will be selected below and only depends on (d,Λ,β,θ). In particular, $k \le C$. Using the subadditivity of $(\overline{\mu} - \mu(\cdot))_+$ in the form of (4.14) and then applying Lemma A.2 and (P3), we obtain

$$\begin{aligned}&\mathbb{E}\left[\left(\overline{\mu} - \mu(\boxdot_{n+m})\right)_+^{2k}\right] \\ &\quad\le \mathbb{E}\left[\left(3^{-dm}\sum_{\boxdot_n(x)\subseteq\boxdot_{n+m}}\left(\overline{\mu} - \mu(\boxdot_n(x))\right)_+ + CK_0^2 3^{-n\beta/(1+\beta)}\right)^{2k}\right] \\ &\quad\le 2\mathbb{E}\left[\left(3^{-dm}\sum_{\boxdot_n(x)\subseteq\boxdot_{n+m}}\left(\overline{\mu} - \mu(\boxdot_n(x))\right)_+\right)^{2k}\right] + \left(CK_0^2\right)^{2k} 3^{-2kn\beta/(1+\beta)} \\ &\quad\le CK_0^{4k}\left(\max\left\{\mathbb{E}\left[\frac{\left(\overline{\mu} - \mu(\boxdot_n(x))\right)_+}{CK_0^2}\right], 3^{-dm}\right\}^{2k} + 3^{-\beta n} + 3^{-kn\alpha}\right).\end{aligned}$$

Inserting (6.67), we obtain

$$\mathbb{E}\left[\left(\overline{\mu} - \mu(\boxdot_{n+m})\right)_+^{2k}\right] \le CK_0^{4k}\left(\left(3^{-n\alpha} + 3^{-dm}\right)^{2k} + 3^{-\beta n}\right).$$

A slight reformulation of the previous inequality (replace $n+m$ by n and n by $n-m$) yields, for every $n, m \in \mathbb{N}$ with $m \le n$,

$$\mathbb{E}\left[\left(\overline{\mu} - \mu(\boxdot_n)\right)_+^{2k}\right] \le CK_0^{4k}\left(\left(3^{-(n-m)\alpha} + 3^{-dm}\right)^{2k} + 3^{-\beta(n-m)}\right).$$

We may extend this inequality to $m \in \mathbb{R}_+$ with $m \le n$ by adjusting the constant C. We now select m as a function of n so that $\beta(n-m) = \theta n$, that is, $m(n) := (\beta-\theta)n/\beta$ and we thereby obtain, for every $n \in \mathbb{N}$,

$$\mathbb{E}\left[\left(\overline{\mu} - \mu(\square_n)\right)_+^{2k}\right] \le CK_0^{4k}\left(\left(3^{-n\theta\alpha/\beta} + 3^{-nd(\beta-\theta)/\beta}\right)^{2k} + 3^{-\theta n}\right).$$

We now select $k \in \mathbb{N}$ to be the smallest positive integer satisfying

$$2k \max\left\{\frac{\theta\alpha}{\beta}, \frac{d(\beta-\theta)}{\beta}\right\} \ge \theta.$$

Thus, as we had promised, $k \le C$. We deduce that, for every $n \in \mathbb{N}$,

$$\mathbb{E}\left[\left(\overline{\mu} - \mu(\square_n)\right)_+^{2k}\right] \le CK_0^{4k} Ck 3^{-\theta n} \le CK_0^{4k} 3^{-\theta n}.$$

Fix $s_0 := 2k$ and observe that we may use the previous inequality and (4.6) to obtain, for every $s \ge s_0$,

$$\mathbb{E}\left[\left(\overline{\mu} - \mu(\square_n)\right)_+^{s}\right] \le (CK_0^2)^{s-2k}\,\mathbb{E}\left[\left(\overline{\mu} - \mu(\square_n)\right)_+^{2k}\right] \le (CK_0^2)^s 3^{-\theta n}.$$

By a very similar argument, we also obtain the bound

$$\mathbb{E}\left[\left(\mu_0(\square_{n+m}, \overline{P}, \overline{Q}) - \overline{\mu}\right)_+^{s}\right] \le (CK_0^2)^s 3^{-\theta n}.$$

We next use the fact that, due to $\mu(U) \le \mu_0(U, \overline{P}, \overline{Q})$, we have

(6.69)
$$\mathcal{E}_*(U) = \left|\overline{\mu} - \mu(U)\right| + \left|\mu_0(U, \overline{P}, \overline{Q}) - \overline{\mu}\right| \le 2\left(\overline{\mu} - \mu(U)\right)_+ + 2\left(\mu_0(U, \overline{P}, \overline{Q}) - \overline{\mu}\right)_+$$

and combining this with the previous inequalities to obtain (6.68).

Step 7. In this last step, we upgrade the convergence to exponential moments under the additional assumption that (P4) holds. This is the first and only place in the paper that we use (P4), and is analogous to the previous step except that we use the much stronger Lemma A.3 in place of Lemma A.2.

The convention for the constants in this step is different from the rest of the proof: here we allow C to depend on the parameters (γ, C_4) in (P4) in addition to (d, Λ, β, C_3), and α may depend also on γ in addition to (d, Λ, β).

Fix $n, m \in \mathbb{N}$ and $0 < t \le (C_1K_0^2)^{-1}3^{dm}$, where the previous $C_1 := C \ge 1$ is fixed large enough that, for all bounded Lipschitz domains $U \subseteq \mathbb{R}^d$,

$$\mathbb{P}\left[\left(\overline{\mu} - \mu(U)\right)_+ \le C_1K_0^2\right] = 1. \tag{6.70}$$

Now we compute, using (4.14) and Lemma A.3:

$$\begin{aligned}
&\log \mathbb{E}\left[\exp\left(t\left(\overline{\mu} - \mu(\square_{n+m})\right)_+\right)\right] \\
&\quad\le \log \mathbb{E}\left[\exp\left(t3^{-dm} \sum_{\square_n(x)\subseteq\square_{n+m}} \left(\overline{\mu} - \mu(\square_n(x))\right)_+\right)\right] + CtK_0^2 3^{-n\beta/(1+\beta)} \\
&\quad= \log \mathbb{E}\left[\prod_{\square_n(x)\subseteq\square_{n+m}} \exp\left(t3^{-dm}\left(\overline{\mu} - \mu(\square_n(x))\right)_+\right)\right] + CK_0^2 t 3^{-n\alpha} \\
&\quad\le C\left(t\mathbb{E}\left[\left(\overline{\mu} - \mu(\square_n)\right)_+\right] + K_0^2 3^{dm}\exp\left(-3^{n\gamma}\right)\exp\left(CK_0^2 t\right)\right) + CK_0^2 t 3^{-n\alpha}.
\end{aligned}$$

We now apply (6.67) to obtain, for every $n, m \in \mathbb{N}$ and $0 < t \le (CK_0^2)^{-1}3^{dm}$,

$$\frac{1}{t}\log \mathbb{E}\left[\exp\left(t\left(\overline{\mu} - \mu(\boxdot_{n+m})\right)_+\right)\right] \le CK_0^2 3^{-n\alpha} + CK_0^2 t^{-1}3^{dm}\exp\left(CK_0^2 t - 3^{n\gamma}\right).$$

A very similar computation yields

$$\frac{1}{t}\log \mathbb{E}\left[\exp\left(t\left(\mu_0(\boxdot_{n+m}, \overline{P}, \overline{Q}) - \overline{\mu}\right)_+\right)\right] \le CK_0^2 3^{-n\alpha} + CK_0^2 t^{-1}3^{dm}\exp\left(CK_0^2 t - 3^{n\gamma}\right).$$

Combining these inequalities and using (6.69), we obtain

$$\frac{1}{t}\log \mathbb{E}\left[\exp\left(t\mathcal{E}_*(\boxdot_{n+m})\right)\right] \le CK_0^2 3^{-n\alpha} + CK_0^2 t^{-1}3^{dm}\exp\left(CK_0^2 t - 3^{n\gamma}\right).$$

A reformulation of this inequality (as in Step 6, we replace $n+m$ with n and n with $n-m$) yields, for every $n, m \in \mathbb{N}$ with $m \le n$ and $t > 0$ as above,

$$\frac{1}{t}\log \mathbb{E}\left[\exp\left(t\mathcal{E}_*(\boxdot_n)\right)\right] \le CK_0^2 3^{-(n-m)\alpha} + CK_0^2 t^{-1}3^{dm}\exp\left(CK_0^2 t - 3^{(n-m)\gamma}\right).$$

We may extend the previous inequality to $m \in \mathbb{R}_+$ with $m \le n$ by adjusting the constants C. We now select $m := n\gamma/(d+\gamma)$ and $t := (CK_0^2)^{-1}3^{dm}$, with $C \ge C_1$ chosen sufficiently large that the expression inside the exponential on the right side of the previous inequality above can be estimated by

$$CK_0^2 t - 3^{(n-m)\gamma} \le \frac{1}{2}3^{dm} - 3^{(n-m)\gamma} = -\frac{1}{2}3^{(n-m)\gamma}.$$

Substituting, we get

$$\log \mathbb{E}\left[\exp\left(3^{dm}\frac{\mathcal{E}_*(\boxdot_n)}{CK_0^2}\right)\right] \le C3^{dm-(n-m)\alpha} + C3^{dm}\exp\left(-\frac{1}{2}\cdot 3^{nd\gamma/(d+\gamma)}\right) \le C3^{dm-(n-m)\alpha}.$$

This is (6.57). □

We next complete the proof of Proposition 6.1. Most of the heavy lifting has already been accomplished, and the statement of Lemma 6.5 is already quite close to that of Proposition 6.1. What is left is to allow for arbitrary translations of the cubes.

Proof of Proposition 6.1. Here C is a positive constant depending only on (d, Λ, β, C_3) while α depends only on (d, Λ, β), and these may vary in each occurrence. As above, we drop the dependence of μ_0 on $(\overline{P}, \overline{Q})$ and let σ_n be denoted by (6.58), where the exponent α implicit in σ_n may change in each occurrence.

For $y \in \mathbb{R}^d$, let $[y]$ be the nearest element of $\mathbb{Z}^d$ to y (where we resort to the lexicographical ordering by the indices in case of nonuniqueness). Observe that

$$[y] + \boxdot_n \subseteq y + \square_n \quad \text{and} \quad \left|(y+\square_n) \setminus ([y]+\boxdot_n)\right| \le C3^{-n\beta/(1+\beta)}|\,\square_n\,|.$$

Similar to the proof of (4.23), it follows from (4.3) and (4.6) that, for each $y \in \mathbb{R}^d$ and $n \in \mathbb{N}$,

$$\mu(y+\square_n) \ge \mu([y]+\boxdot_n) - CK_0^2 3^{-n\beta/(1+\beta)} \ge \mu([y]+\boxdot_n) - C\sigma_n \quad \mathbb{P}\text{–a.s.}$$

In particular, for every $R \geq 1$,

$$\sup_{y \in B_R} \left(\overline{\mu} - \mu(y + \square_n)\right)_+ \leq \sup_{z \in B_{CR}} \left(\overline{\mu} - \mu(z + \boxdot_n)\right)_+ + C\sigma_n \quad \mathbb{P}\text{–a.s.}$$

Similarly,

$$\sup_{y \in B_R} \left(\mu_0(y + \square_n) - \overline{\mu}\right)_+ \leq \sup_{z \in B_{CR}} \left(\mu_0(z + \boxdot_n) - \overline{\mu}\right)_+ + C\sigma_n \quad \mathbb{P}\text{–a.s.}$$

The previous two inequalities and (6.69) imply

$$\sup_{y \in B_R} \mathcal{E}_*(y + \square_n) \leq C \sup_{z \in B_{CR}} \mathcal{E}_*(z + \boxdot_n) + C\sigma_n \quad \mathbb{P}\text{–a.s.}$$

Hence by a union bound and stationarity, we get, for any $\zeta \geq C\sigma_n$,

$$\mathbb{P}\left[\sup_{y \in B_R} \mathcal{E}_*(y + \square_n) \geq \zeta\right] \leq |B_{CR}|\, \mathbb{P}\left[\mathcal{E}_*(\boxdot_n) \geq C\zeta\right] \leq CR^d\, \mathbb{P}\left[\mathcal{E}_*(\boxdot_n) \geq C\zeta\right].$$

To prove the first statement of the proposition, it remains to show that, for any $\theta \in (0, \beta)$, there exists $s_0(d, \Lambda, \beta, \theta) \geq 1$ such that, for every $s \in [s_0, \infty)$ and $t \geq 1$,

$$\mathbb{P}\left[\mathcal{E}_*(\boxdot_n) \geq CK_0^2 3^{-n\theta/s} t\right] \leq Ct^{-s}. \tag{6.71}$$

This follows at once from Chebyshev's inequality and (6.56).

To prove the second assertion of the proposition, we have left to show that, under assumption (P4), that for any $t \geq 1$ and $s \in (0, d\gamma/(d+\gamma))$, we have

$$\mathbb{P}\left[\mathcal{E}_*(\boxdot_n) \geq CK_0^2 3^{-n \min\{\alpha, d\gamma/(d+\gamma) - s\}} t\right] \leq C \exp\left(-\frac{1}{2} 3^{ns} t\right), \tag{6.72}$$

where the constants α and C have appropriate dependence as stated in the proposition. Fix $s \in (0, d\gamma/(d+\gamma))$, $n \in \mathbb{N}$ and $t \geq 1$ and compute, using Chebyshev's inequality and (6.57):

$$\begin{aligned}\mathbb{P}\left[3^{nd\gamma/(d+\gamma)} \frac{\mathcal{E}_*(\boxdot_n)}{CK_0^2} \geq 3^{ns} t\right] &\leq \exp\left(-3^{ns} t\right) \mathbb{E}\left[\exp\left(3^{nd\gamma/(d+\gamma)} \frac{\mathcal{E}_*(\boxdot_n)}{CK_0^2}\right)\right] \\ &\leq \exp\left(-3^{ns} t + C 3^{n(d\gamma/(d+\gamma) - \alpha)}\right).\end{aligned}$$

Suppose that $s \geq (d\gamma/(d+\gamma) - \alpha) + c$. Then the expression in the exponential in the last expression is at most $-\frac{1}{2} 3^{ns} t$ for $n \geq C$, and thus, for such s, we obtain, for every $n \in \mathbb{N}$ and $t \geq 1$,

$$\mathbb{P}\left[\frac{\mathcal{E}_*(\boxdot_n)}{CK_0^2} \geq 3^{-n(d\gamma/(d+\gamma) - s)} t\right] \leq C \exp\left(-\frac{1}{2} 3^{ns} t\right).$$

By applying this inequality for $s' := (d\gamma/(d+\gamma) - \alpha) + c$ and adjusting the constant C, we obtain, for every $0 < s \leq s'$,

$$\mathbb{P}\left[\frac{\mathcal{E}_*(\boxdot_n)}{CK_0^2} \geq 3^{-nc} t\right] \leq C \exp\left(-\frac{1}{2} 3^{ns'} t\right) \leq C \exp\left(-\frac{1}{2} 3^{ns} t\right).$$

After combining the previous two inequalities and possibly redefining α, we obtain (6.72). □

6.4. **The proof of Theorem 5.1.** We next give the proof of Theorem 5.1, which follows from Proposition 6.1 and some union bounds.

Proof of Theorem 5.1. According to (4.1) and (6.1), we have

$$\overline{F}(p,q) = \overline{\mu}_0(p,q) = \inf_{p^*,q^* \in \mathbb{R}^d} \left(\overline{\mu}(q^*,p^*) - p \cdot q^* - p^* \cdot q \right).$$

That is, $\overline{F}$ is the Legendre-Fenchel transform of $-\overline{\mu}$. It follows that, if $(q^*,p^*) = \nabla \overline{F}(p,q)$, then

$$\mathcal{E}_*(U,q^*,p^*) = \mathcal{E}(U,p,q) = \left| \mu(U,q^*,p^*) - \overline{\mu}(q^*,p^*) \right| + \left| \mu_0(U,p,q) - \overline{\mu}_0(U,p,q) \right|.$$

Thus we see that the error estimate (6.2) is already close to the desired conclusion (5.1), we just need to obtain some uniformity in (p,q) and in large scales.

Step 1. We prove (5.1). Fix $\theta \in (0,\beta)$ and $\tau \geq 1$. According to (4.31), (4.32), (4.36) and (4.38), the quantity $\mathcal{E}$ is continuous in (p,q): that is, for every $p_1,p_2,q_1,q_2 \in \mathbb{R}^d$ and bounded domain $U \subseteq \mathbb{R}^d$,

$$\begin{aligned} (6.73) \quad & \left| \mathcal{E}(U,p_1,q_1) - \mathcal{E}(U,p_2,q_2) \right| \\ & \qquad \leq C \left(K_0 + |p_1| + |q_1| + |p_2| + |q_2| \right) \left(|p_1 - p_2| + |q_1 - q_2| \right) \quad \mathbb{P}\text{–a.s.} \end{aligned}$$

This yields that

$$\begin{aligned} (6.74) \quad & \sup_{p,q \in B_{M3^{n\tau/s}}} \frac{\mathcal{E}(\square_n,p,q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\theta/s}} \\ & \qquad \leq C + \sup_{p,q \in B_{M3^{n\tau/s}} \cap G_{n,s}} \frac{\mathcal{E}(\square_n,p,q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\theta/s}} \quad \mathbb{P}\text{–a.s.}, \end{aligned}$$

where we have set

$$G_{n,s} := 3^{-n\theta/s} \mathbb{Z}^d \subseteq \mathbb{R}^d.$$

Observe that the number of elements in the set $B_{M3^{n\tau/s}} \cap G_{n,s}$ is

$$\left| B_{M3^{n\theta/s}} \cap G_{n,s} \right| = C M^d 3^{nd(\theta+\tau)/s}.$$

By a union bound and Proposition 6.1 applied with exponent $\theta + \frac{1}{2}(\beta - \theta)$, we have, for every $n \in \mathbb{N}$, $t \geq 1$ and $s \geq s_0(d,\Lambda,\beta,\theta) \geq 1$,

$$\begin{aligned} & \mathbb{P}\left[\sup_{p,q \in B_{M3^{n\tau/s}}} \sup_{y \in B_R} \frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\theta/s}} \geq Ct \right] \\ & \quad \leq \left| B_{M3^{n\theta/s}} \cap G_{n,s} \right|^2 \cdot \sup_{p,q \in \mathbb{R}^d} \mathbb{P}\left[\sup_{y \in B_R} \frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\theta/s}} \geq t \right] \\ & \quad \leq C R^d M^{2d} 3^{2nd(\theta+\tau)/s} \left(3^{n(\beta-\theta)/2s} t \right)^{-s} \\ & \quad = C R^d M^{2d} 3^{2nd(\theta+\tau)/s - n(\beta-\theta)/2} t^{-s}. \end{aligned}$$

By making $s_0(d,\Lambda,\beta,\theta,\tau)$ larger, we may assume that $2d(\theta+\tau)/s \leq \frac{1}{2}(\beta-\theta)$ and then we get, for every $n \in \mathbb{N}$, $t \geq 1$ and $s \geq s_0(d,\Lambda,\beta,\theta,\tau)$,

$$\mathbb{P}\left[\sup_{p,q \in B_{M3^{n\tau/s}}} \sup_{y \in B_R} \frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2 + |p|^2 + |q|^2) 3^{-n\theta/s}} \geq Ct \right] \leq C R^d M^{2d} t^{-s}.$$

This is (5.1).

Step 2. We prove (5.2) under assumption (P4). The argument is almost the same as in Step 1, only easier. Following the argument there, the result of the first union bound can be improved by using (6.4) rather than (6.2): we get, for every $s \in (0, d\gamma/(d+\gamma))$, $R \geq 1$, $n \in \mathbb{N}$ and $t \geq 1$,

$$\begin{aligned}
&\mathbb{P}\left[\sup_{p,q\in B_{M3^n}} \sup_{y\in B_{R3^n}} \frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2+|p|^2+|q|^2)3^{-n\alpha}} \geq Ct\right] \\
&\quad \leq |B_{M3^n}\cap G_n|^2 \cdot \sup_{p,q\in\mathbb{R}^d} \mathbb{P}\left[\sup_{y\in B_{R3^n}} \frac{\mathcal{E}(y+\square_n,p,q)}{(K_0^2+|p|^2+|q|^2)3^{-n\alpha_s}} \geq t\right] \\
&\quad \leq CR^d M^{2d} 3^{3nd} \exp\left(-3^{ns}t\right),
\end{aligned}$$

where we have set $\alpha_s := \min\{\alpha, d\gamma/(d+\gamma) - s\}$ and $G_n := 3^{-n\alpha_s/2}(\mathbb{Z}^d \times \mathbb{Z}^d)$. For convenience we may assume that $\alpha_s \leq 1$ so that $|B_{M3^n} \cap G_n| \leq 3^{3nd}$. Finally, we may absorb the factor 3^{3nd} into the constant C at the cost of slightly shrinking the exponent s. Since $\alpha_s \geq \alpha(d\gamma/(d+\gamma) - s)$ for some $\alpha(d,\Lambda,\beta) > 0$, we thereby obtain the result. □

6.5. **Construction of the homogenized coefficients $\overline{\mathbf{a}}$.** We conclude this section by verifying that $\overline{F}$ is the representative of a Lipschitz, uniformly monotone vector field $\overline{\mathbf{a}}$.

Proposition 6.6. *There exists a Lipschitz, uniformly monotone vector field $\overline{\mathbf{a}}$ which is variationally represented by $\overline{F}$ and satisfies, for some $C(\Lambda) \geq 1$,*

$$\begin{cases}
|\overline{\mathbf{a}}(0)| \leq CK_0, \\
|\overline{\mathbf{a}}(p_1) - \overline{\mathbf{a}}(p_2)| \leq 4\Lambda\,|p_1 - p_2|, \\
(\overline{\mathbf{a}}(p_1) - \overline{\mathbf{a}}(p_2)) \cdot (p_1 - p_2) \geq \dfrac{1}{4\Lambda}\,|p_1 - p_2|^2.
\end{cases} \tag{6.75}$$

Proof. According to Lemma 2.13, to prove the existence of a vector field $\overline{\mathbf{a}}$ represented by $\overline{F}$, it suffices to show that, for each $p \in \mathbb{R}^d$,

$$\inf_{q\in\mathbb{R}^d} (\overline{F}(p,q) - p\cdot q) = 0. \tag{6.76}$$

Fix $p_0 \in \mathbb{R}^d$. We first claim that there exists unique $q, q^* \in \mathbb{R}^d$ such that

$$\nabla\overline{F}(p_0,q) = (q^*, p_0). \tag{6.77}$$

To see this, observe that $q \mapsto D_q F(p_0,q)$ is bijective on $\mathbb{R}^d$ by (4.33) and the uniform convexity of $\overline{F}$, where $D_q\overline{F}$ denotes the gradient of $\overline{F}$ in the variable q. Thus there exists a unique $q \in \mathbb{R}^d$ for which $D_q F(p_0,q) = p_0$. We may then identify $q^* = D_p F(p_0,q)$.

According to Proposition 2.15, we have

$$\begin{aligned}
0 = \inf\Bigg\{ \fint_{\square_n} &(F(p_0 + \nabla w(x), \mathbf{f}(x), x) - (p_0 + \nabla w(x))\cdot \mathbf{f}(x))\, dx \\
&: w \in H^1_0(\square_n),\ \mathbf{f} \in L^2_{\mathrm{sol}}(\square_n)\Bigg\}.
\end{aligned}$$

It is immediate from this expression that

$$\mu(\square_n, q^*, p_0) + q^* \cdot p_0 \le 0 \le \mu_0(\square_n, p_0, q) - p_0 \cdot q.$$

Passing to the limit $n \to \infty$ yields

$$\overline{\mu}(q^*, p_0) + q^* \cdot p_0 \le 0 \le \overline{\mu}_0(p_0, q) - p_0 \cdot q. \tag{6.78}$$

In the notation of Proposition 6.1, the identify (6.77) asserts that

$$(p_0, q) = \left(\overline{P}(p_0, q^*), \overline{Q}(p_0, q^*)\right).$$

Thus according to (6.1),

$$\overline{F}(p_0, q) = \overline{\mu}_0(p_0, q) = \overline{\mu}(q^*, p_0) + (q + q^*) \cdot p_0. \tag{6.79}$$

Combining (6.78) and (6.79) yields

$$\overline{F}(p_0, q) = \overline{\mu}_0(p_0, q) = p_0 \cdot q.$$

This gives (6.76), as the other inequality was proved already in (4.39).

To conclude, we notice that the second and third lines of (6.75) are immediate consequences of the second assertion of Lemma 2.13 and (4.34). The first line of (6.75) follows from (4.33) and Lemma 2.14. □

7. The proof of Proposition 5.2

In this section, we show that the L^2 difference between the solutions of the Dirichlet problem (on a given macroscopic domain) for the heterogeneous and homogeneous equations is controlled by the convergence rate of μ to its limit $\overline{\mu}$ on mesoscopic subdomains. The argument is a generalization of the one for [3, Proposition 4.1], although here the proof is much simpler due to the fact that the variational formulation we use is very convenient to work with – even compared to the more classical variational formulation (for equations which admit one). In particular, since the variational problem for $\overline{\mathcal{J}}$ is always null, we have only to prove the analogue of the easier of the two energy estimates from [3, Appendix A].

For the rest of the section, we fix a bounded Lipschitz domain $U_0 \subseteq \mathbb{R}^d$, $m, n \in \mathbb{N}$ and set $U := 3^{n+m} U_0$. We also fix $F \in \Omega$, $\delta > 0$, $f \in W^{1,2+\delta}(U)$ and let M be defined by (5.5). We denote by $(u, \mathbf{g})$ and $(\overline{u}, \overline{\mathbf{g}})$ the unique elements of $(f + H_0^1(U)) \times L^2_{\mathrm{sol}}(U)$ such that

$$0 = \int_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - \nabla u(x) \cdot \mathbf{g}(x)\right) dx \tag{7.1}$$

and

$$0 = \int_U \left(\overline{F}\left(\nabla \overline{u}(x), \overline{\mathbf{g}}(x), x\right) - \nabla \overline{u}(x) \cdot \overline{\mathbf{g}}(x)\right) dx. \tag{7.2}$$

In particular, u and $\overline{u}$ satisfy (5.3). Note that for $u, \overline{u} \in f + H_0^1(U)$, we can use the substitutions

$$\fint_U \nabla u(x) \cdot \mathbf{g}(x)\, dx = \fint_U \nabla f(x) \cdot \mathbf{g}(x)\, dx,$$
$$\fint_U \nabla \overline{u}(x) \cdot \overline{\mathbf{g}}(x)\, dx = \fint_U \nabla f(x) \cdot \overline{\mathbf{g}}(x)\, dx,$$

which reveal the uniform convexity of the minimization problems. As in the statement of Proposition 5.2, we consider $n+m$ to be the macroscopic (triadic) scale and n to denote a mesoscopic scale.

The convention for constants in this section is as follows: $\alpha > 0$ is an exponent depending on $(d, \Lambda, \beta, \delta)$, while the constants $C \geq 1$ and $0 < c \leq 1$ depend on $(d, \Lambda, U_0, \beta, C_3, \delta)$. We allow each of these parameters to vary in each occurrence.

We prove Proposition 5.2 by an energy comparison argument. The idea is to remove the microscopic fluctuations from $(u, \mathbf{g})$ by a mesoscopic spatial average, thereby obtaining $(\widetilde{u}, \widetilde{\mathbf{g}})$. The main step is to show that the $\overline{F}$–energy of $(\widetilde{u}, \widetilde{\mathbf{g}})$ is not much larger than the F–energy of $(u, \mathbf{g})$. This is stated precisely below in (7.15). Since by construction, we will also have that $\mathbf{g} - \widetilde{\mathbf{g}} \in L^2_{\mathrm{sol},0}(U)$, we will be able to deduce that

$$
\begin{aligned}
0 &= \fint_U \left(F(\nabla u(x), \mathbf{g}(x), x) - \nabla f(x) \cdot \mathbf{g}(x)\right) dx \\
&= \fint_U \left(\overline{F}(\nabla \widetilde{u}(x), \widetilde{\mathbf{g}}(x)) - \nabla f(x) \cdot \widetilde{\mathbf{g}}(x)\right) dx \ - \ \text{a small error.}
\end{aligned}
$$

This allows us to conclude that $\nabla \widetilde{u}$ and $\nabla \overline{u}$ are close in $L^2(U)$, as are $\widetilde{\mathbf{g}}$ and $\overline{\mathbf{g}}$. By construction, u and $\widetilde{u}$ are also close in $L^2(U)$ since the latter is the spatial average of the former, which belongs to $W^{1,2+\delta}$ for some $\delta > 0$ by the Meyers estimate. Hence, by the triangle inequality, we deduce that u and $\overline{u}$ are close in $L^2(U)$.

7.1. **Construction of the approximating pair.** To prove Proposition 5.2, it suffices to consider the case in which $m \geq 5$ and $n+3 \leq l < m+n$, since we can then recover the general case by adjusting the constant C in (5.4). The integer l is used to measure an additional mesoscopic scale that gives the thickness of a boundary strip to be removed from U in the construction. Note that

$$
c3^{d(n+m)} \leq |U| \leq C3^{d(n+m)}. \tag{7.3}
$$

We define Lipschitz subdomains $V^\circ \subseteq V \subseteq W \subseteq U$ by

$$
\begin{aligned}
V^\circ &:= \bigcup \left\{ z + \square_n \,:\, z \in 3^n \mathbb{Z}^d,\ z + \square_{l+4} \subseteq U \right\}, \\
V &:= \bigcup \left\{ z + \square_n \,:\, z \in 3^n \mathbb{Z}^d,\ z + \square_{l+2} \subseteq U \right\}, \\
W &:= \bigcup \left\{ z + \square_n \,:\, z \in 3^n \mathbb{Z}^d,\ z + \square_{n+3} \subseteq U \right\}.
\end{aligned}
$$

Since U_0 is Lipschitz, we have

$$
|U \setminus V^\circ| \leq C3^{l-(n+m)}|U|. \tag{7.4}
$$

Since $n + 3 \leq l$, we have

$$
V \subseteq \bigcup_{z \in 3^n \mathbb{Z}^d \cap W} (z + \square_n). \tag{7.5}
$$

Observe that $\operatorname{dist}(V^\circ, \partial V) \geq 3^l$ and select a cutoff function $\eta \in C_c^\infty(U)$ satisfying

$$
0 \leq \eta \leq 1, \quad \eta \equiv 1 \text{ on } V^\circ + \square_n, \quad \eta \equiv 0 \text{ on } U \setminus V, \quad \sup_{x \in U} |\nabla \eta(x)| \leq C3^{-l}. \tag{7.6}
$$

We next construct the approximating pair $(\widetilde{u},\widetilde{\mathbf{g}}) \in (f+H^1_0(U)) \times L^2_{\mathrm{sol}}(U)$ by modifying $(u,\mathbf{g})$. Define the mesoscopic spatial averages of $(u,\mathbf{g})$ by

$$u_{\mathrm{sa}}(y) := \fint_{y+\square_n} u(x)\,dx \quad \text{and} \quad \mathbf{g}_{\mathrm{sa}}(y) := \fint_{y+\square_n} \mathbf{g}(x)\,dx, \quad y \in V. \tag{7.7}$$

Observe that $(u_{\mathrm{sa}},\mathbf{g}_{\mathrm{sa}}) \in H^1(V) \times L^2_{\mathrm{sol}}(V)$. We get our approximate pair by smoothly interpolating between $(u_{\mathrm{sa}},\mathbf{g}_{\mathrm{sa}})$ and $(u,\mathbf{g})$ near the boundary of V:

$$\begin{cases} \widetilde{u}(x) := \eta(x) u_{\mathrm{sa}}(x) + (1-\eta(x))u(x), \\ \widetilde{\mathbf{g}}(x) := \eta(x)\mathbf{g}_{\mathrm{sa}}(x) + (1-\eta(x))\mathbf{g}(x) - \nabla h(x), \end{cases}$$

where $h \in H^1(U)$ is the (unique, up to a constant) solution of the Neumann problem

$$\begin{cases} \Delta h = \nabla \cdot (\eta \mathbf{g}_{\mathrm{sa}} + (1-\eta)\mathbf{g}) & \text{in } U, \\ \partial_\nu h = 0 & \text{on } \partial U. \end{cases} \tag{7.8}$$

Observe that we indeed have $(\widetilde{u},\widetilde{\mathbf{g}}) \in (f+H^1_0(U)) \times L^2_{\mathrm{sol}}(U)$.

We conclude this subsection by giving some L^p estimates needed below. First, according to the Meyers estimate (Proposition B.6), there exists an exponent $\delta_0(d,\Lambda,\delta) \in (0,\delta]$ such that

$$\left(\fint_U |\nabla u(x)|^{2+\delta_0}\,dx \right)^{1/(2+\delta_0)} + \left(\fint_U |\nabla \overline{u}(x)|^{2+\delta_0}\,dx \right)^{1/(2+\delta_0)} \le CM. \tag{7.9}$$

Note that $\mathbf{g}(x) = \mathbf{a}(\nabla u(x),x)$ and $\overline{\mathbf{g}} = \overline{\mathbf{a}}(\nabla \overline{u})$ by Lemma 2.13 and Proposition 2.15. By (1.9), (P1), and (6.75), we have that

$$|\mathbf{a}(p,x)| \le C\left(K_0 + |p|\right) \quad \text{and} \quad |\overline{\mathbf{a}}(p)| \le C\left(K_0+|p|\right).$$

We deduce that, a.e. in U,

$$|\mathbf{g}| \le C\left(K_0 + |\nabla u|\right) \quad \text{and} \quad |\overline{\mathbf{g}}| \le C\left(K_0 + |\nabla \overline{u}|\right) \tag{7.10}$$

and thus

$$\left(\fint_U |\mathbf{g}(x)|^{2+\delta_0}\,dx \right)^{1/(2+\delta_0)} + \left(\fint_U |\overline{\mathbf{g}}(x)|^{2+\delta_0}\,dx \right)^{1/(2+\delta_0)} \le CM. \tag{7.11}$$

The inequalities above, together with the Hölder inequality, give a bound on the energy density of $(u,\mathbf{g})$ and $(\overline{u},\overline{\mathbf{g}})$ in the boundary strip $U \setminus V^\circ$. Indeed, we have

$$\begin{aligned} \frac{1}{|U|} \int_{U\setminus V^\circ} \left(|\nabla u(x)|^2 + |\nabla \overline{u}(x)|^2 + |\mathbf{g}(x)|^2 + |\overline{\mathbf{g}}(x)|^2 \right) dx \\ \le CM^2 \left(\frac{|U\setminus V^\circ|}{|U|} \right)^{\delta_0/(2+\delta_0)} \le CM^2 3^{-\alpha(n+m-l)}. \end{aligned} \tag{7.12}$$

They also imply, by the Cauchy-Schwarz inequality, that for every $y \in V$,

$$|\nabla u_{\mathrm{sa}}(y)|^2 \le \left(\fint_{y+\square_n} |\nabla u(x)|\,dx \right)^2 \le \frac{|U|}{|\square_n|} \fint_U |\nabla u(x)|^2\,dx \le C3^{dm} M^2, \tag{7.13}$$

and similarly, using (7.11),

$$|\mathbf{g}_{\mathrm{sa}}(y)| \le C3^{dm/2} M. \tag{7.14}$$

7.2. **The energy estimate.** As we will see, Proposition 5.2 is essentially a consequence of the uniform convexity of $\overline{F}$ and the following estimate, which we prove in this subsection:

$$\begin{aligned}(7.15)\quad &\fint_U \overline{F}\left(\nabla \widetilde{u}(x), \widetilde{\mathbf{g}}(x)\right) dx - \fint_U F\left(\nabla u(x), \mathbf{g}(x), x\right) dx \\ &\leq CM^2 \left(\fint_U \left(\sup_{p,q \in B_{CM3^{dm/2}}} \frac{\mathcal{E}(x+\square_n, p, q)}{K_0^2 + |p|^2 + |q|^2} \right)^{\frac{2+\delta_0}{\delta_0}} dx \right)^{\frac{\delta_0}{2+\delta_0}} + CM^2 3^{-\alpha(n+m-l)}.\end{aligned}$$

Here $\delta_0(d, \Lambda, \delta) > 0$ is defined in the previous subsection and given by the Meyers estimate.

We now present the derivation of the claimed energy estimate.

Proof of (7.15). We break the argument into several steps.

Step 1. We begin with the main part of the argument for (7.15). For $x \in V$, let $(p^*(x), q^*(x))$ denote the dual pair (via $\overline{F}$) to $(\nabla u_{\mathrm{sa}}(x), \mathbf{g}_{\mathrm{sa}}(x) - \nabla h(x))$, that is:

$$(p^*(x), q^*(x)) := \nabla \overline{F}\left(\nabla u_{\mathrm{sa}}(x), \mathbf{g}_{\mathrm{sa}}(x) - \nabla h(x)\right).$$

Using $(u, \mathbf{g})$ as a minimizer candidate for $\mu(y + \square_n, p^*(y), q^*(y))$, we have

$$\begin{aligned}&\mu(y+\square_n, p^*(y), q^*(y)) \\ &\quad \leq \fint_{y+\square_n} \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - q^*(y) \cdot \nabla u(x) - p^*(y) \cdot \mathbf{g}(x)\right) dx \\ &\quad = \fint_{y+\square_n} F\left(\nabla u(x), \mathbf{g}(x), x\right) dx - \left(q^*(y), p^*(y)\right) \cdot \left(\nabla u_{\mathrm{sa}}(y), \mathbf{g}_{\mathrm{sa}}(y)\right).\end{aligned}$$

By (6.1),

$$\begin{aligned}\left(q^*(y), p^*(y)\right) \cdot \left(\nabla u_{\mathrm{sa}}(y), \mathbf{g}_{\mathrm{sa}}(y)\right) &= \overline{F}\left(\nabla u_{\mathrm{sa}}(y), \mathbf{g}_{\mathrm{sa}}(y) - \nabla h(y)\right) \\ &\quad + p^*(y) \cdot \nabla h(y) - \overline{\mu}(p^*(y), q^*(y))\end{aligned}$$

and therefore we deduce that, for every $y \in V$,

$$\begin{aligned}(7.16)\quad &\overline{F}\left(\nabla u_{\mathrm{sa}}(y), \mathbf{g}_{\mathrm{sa}}(y) - \nabla h(y)\right) - \fint_{y+\square_n} F\left(\nabla u(x), \mathbf{g}(x), x\right) dx \\ &\qquad \leq -p^*(y) \cdot \nabla h(y) + \overline{\mu}(p^*(y), q^*(y)) - \mu(y+\square_n, p^*(y), q^*(y)) \\ &\qquad \leq -p^*(y) \cdot \nabla h(y) + \mathcal{E}(y+\square_n, p^*(y), q^*(y)).\end{aligned}$$

Integrating this over V°, we obtain

$$\begin{aligned}&\int_{V^\circ} \overline{F}\left(\nabla u_{\mathrm{sa}}(x), \mathbf{g}_{\mathrm{sa}}(x) - \nabla h(x)\right) dx - \int_{V^\circ} \fint_{x+\square_n} F\left(\nabla u(y), \mathbf{g}(y), y\right) dy\, dx \\ &\qquad \leq \int_{V^\circ} |p^*(x)|\, |\nabla h(x)|\, dx + \int_{V^\circ} \mathcal{E}(x+\square_n, p^*(x), q^*(x))\, dx.\end{aligned}$$

This is already quite close to (7.15). The rest of the argument is mainly concerned with using (7.9) to show that the contributions to the energy in the mesoscopic boundary strip $U \setminus V^\circ$ are negligible, to bound $|\nabla h|$ and to put the last term involving $\mathcal{E}(x+\square_n, p^*(x), q^*(x))$ into a form resembling the right side of (7.15).

In fact, to complete the proof of (7.15), it suffices to verify the following three inequalities:

$$\left| \fint_U F\left(\nabla u(x), \mathbf{g}(x), x\right) \, dx - \frac{1}{|U|} \int_{V^\circ} \fint_{x+\square_n} F\left(\nabla u(y), \mathbf{g}(y), y\right) \, dy \, dx \right| + \left| \fint_U \overline{F}\left(\nabla \widetilde{u}(x), \widetilde{\mathbf{g}}(x)\right) \, dx - \frac{1}{|U|} \int_{V^\circ} \overline{F}\left(\nabla u_{\mathrm{sa}}(x), \mathbf{g}_{\mathrm{sa}}(x) - \nabla h(x)\right) \, dx \right| \leq CM^2 3^{-\alpha(n+m-l)}, \tag{7.17}$$

$$\fint_U |p^*(x)| \, |\nabla h(x)| \, dx \leq CM^2 3^{-\alpha(n+m-l)}, \tag{7.18}$$

and

$$\frac{1}{|U|} \int_{V^\circ} \mathcal{E}(x+\square_n, p^*(x), q^*(x)) \, dx \leq CM^2 \left(\fint_U \left(\sup_{p,q \in B_{CM3^{dm/2}}} \frac{\mathcal{E}(x+\square_n, p, q)}{K_0^2 + |p|^2 + |q|^2} \right)^{\frac{2+\delta}{\delta}} dx \right)^{\frac{\delta}{2+\delta}}. \tag{7.19}$$

Step 2. We next show that

$$\fint_V \left(3^{-n} |u_{\mathrm{sa}}(x) - u(x)| + |\nabla u_{\mathrm{sa}}(x)| + |\mathbf{g}_{\mathrm{sa}}(x)|\right)^{2+\delta_0} dx \leq CM^{2+\delta_0}. \tag{7.20}$$

We first estimate the second term in the integrand using (7.9):

$$\fint_V |\nabla u_{\mathrm{sa}}(x)|^{2+\delta_0} \, dx = \fint_V \left| \fint_{x+\square_n} \nabla u(y) \, dy \right|^{2+\delta_0} dx \leq \fint_V \fint_{x+\square_n} |\nabla u(y)|^{2+\delta_0} \, dy \, dx \leq \frac{|U|}{|V|} \fint_U |\nabla u(x)|^{2+\delta_0} \, dx \leq CM^{2+\delta_0}.$$

The third term in the integrand is handled similarly.

To estimate the first term in the integrand, we use the expression

$$\begin{aligned} &\fint_{y+\square_n} |u_{\mathrm{sa}}(x) - u(x)|^{2+\delta_0} \, dx \\ &\quad = \fint_{y+\square_n} \left| u(x) - \fint_{x+\square_n} u(z) \, dz \right|^{2+\delta_0} dx \\ &\quad \leq C \fint_{y+\square_n} \left| u(x) - \fint_{y+\square_n} u(z) \, dz \right|^{2+\delta_0} dx \\ &\qquad + C \fint_{y+\square_n} \left| \fint_{x+\square_n} u(z) \, dz - \fint_{y+\square_n} u(z) \, dz \right|^{2+\delta_0} dx. \end{aligned}$$

We use the Poincaré inequality to bound the first term on the right side:

$$\fint_{y+\square_n} \left| u(x) - \fint_{y+\square_n} u(z) \, dz \right|^{2+\delta_0} dx \leq C 3^{n(2+\delta_0)} \fint_{y+\square_n} |\nabla u(x)|^{2+\delta_0} \, dx,$$

and to bound the second term on the right side, we compute, for all $y \in V$ and $x \in y + \square_n$,

$$\left| \fint_{x+\square_n} u(z)\,dz - \fint_{y+\square_n} u(z)\,dz \right| = \left| \fint_{\square_n} \int_0^1 (x-y)\cdot\nabla u(tx+(1-t)y+z)\,dt\,dz \right| \leq C3^n \fint_{y+\square_{n+1}} |\nabla u(z)|\,dz.$$

Assembling these yields

$$\fint_{y+\square_n} |u_{\mathrm{sa}}(x) - u(x)|^{2+\delta_0}\,dx \leq C3^{n(2+\delta_0)} \fint_{y+\square_{n+2}} |\nabla u(x)|^{2+\delta_0}\,dx$$

and then integrating over $y \in V$ yields the desired estimate of the first term in the integrand in (7.20). This completes the proof of (7.20).

Step 3. We use (7.9) and (7.20) to obtain

$$\fint_U \left(|\nabla \widetilde{u}(x)|^{2+\delta_0} + |\widetilde{\mathbf{g}}(x)|^{2+\delta_0} \right) dx \leq CM^{2+\delta_0}. \tag{7.21}$$

Differentiating the expression for $\widetilde{u}$, we get

$$\nabla\widetilde{u}(x) = \nabla\eta(x)\,(u_{\mathrm{sa}}(x) - u(x)) + \eta(x)\,(\nabla u_{\mathrm{sa}}(x) - \nabla u(x)) + \nabla u(x).$$

Thus the desired estimate for $\nabla\widetilde{u}$ follows from (7.9), (7.6) (recall that $n \leq l$) and (7.20). The estimate for $\widetilde{\mathbf{g}}$ is similar, but we have the extra term ∇h. To estimate this, we note that

$$\nabla\cdot(\eta\mathbf{g}_{\mathrm{sa}} + (1-\eta)\mathbf{g}) = \nabla\cdot((1-\eta)\,(\mathbf{g}_{\mathrm{sa}} - \mathbf{g}))$$

and apply for instance [15, Theorem 1.2], in view of (7.20), to get

$$\fint_U |\nabla h(x)|^{2+\delta_0}\,dx \leq C \fint_U |\mathbf{g}_{\mathrm{sa}}(x) - \mathbf{g}(x)|^{2+\delta_0}\,dx \leq CM^{2+\delta_0}. \tag{7.22}$$

This completes the proof of (7.21). For future reference we observe also that, by (4.37), (7.20) and (7.22), we have

$$\fint_V \left(|p^*(x)| + |q^*(x)|\right)^{2+\delta_0} dx \leq CM^{2+\delta_0}. \tag{7.23}$$

Step 4. We prove (7.18). First we notice that we can bound ∇h more brutally (compared to (7.22)) in L^2 by exploiting that $(1-\eta)$ vanishes in V and using the Hölder inequality. Using [15, Theorem 1.2] with exponent 2 rather than $2+\delta_0$ and (7.4), we get

$$\begin{aligned} \fint_U |\nabla h(x)|^2\,dx &\leq C\frac{1}{|U|}\int_{U\setminus V} |\mathbf{g}_{\mathrm{sa}}(x) - \mathbf{g}(x)|^2\,dx \\ &\leq C\left(\frac{|U\setminus V|}{|U|}\right)^{\delta_0/(2+\delta_0)} \left(\fint_U |\mathbf{g}_{\mathrm{sa}}(x) - \mathbf{g}(x)|^{2+\delta}\,dx\right)^{2/(2+\delta_0)} \\ &\leq C3^{-\alpha(n+m-l)}M^2. \end{aligned} \tag{7.24}$$

The previous inequality, (7.23) and the Hölder inequality yield (7.18).

Step 5. We give the proof of (7.19), which follows from an application of the Hölder inequality and (7.23):

$$\begin{aligned}
&\frac{1}{|U|}\int_{V^\circ} \mathcal{E}(x+\square_n, p^*(x), q^*(x))\,dx \\
&\quad\le \frac{1}{|U|}\int_{V^\circ} \frac{\mathcal{E}(x+\square_n, p^*(x), q^*(x))}{K_0^2+|p^*(x)|^2+|q^*(x)|^2}\left(K_0^2+|p^*(x)|^2+|q^*(x)|^2\right)\,dx \\
&\quad\le C\frac{1}{|U|}\left(\int_{V^\circ}\left(K_0^2+|p^*(x)|^2+|q^*(x)|^2\right)^{\frac{2+\delta}{2}}\,dx\right)^{\frac{2}{2+\delta}} \\
&\qquad\qquad \times\left(\int_{V^\circ}\left(\frac{\mathcal{E}(x+\square_n, p^*(x), q^*(x))}{K_0^2+|p^*(x)|^2+|q^*(x)|^2}\right)^{\frac{2+\delta}{\delta}}\,dx\right)^{\frac{\delta}{2+\delta}} \\
&\quad\le CM^2\left(\fint_U\left(\sup_{p,q\in B_{CM3^{dm/2}}}\frac{\mathcal{E}(x+\square_n,p,q)}{K_0^2+|p|^2+|q|^2}\right)^{\frac{2+\delta}{\delta}}\,dx\right)^{\frac{\delta}{2+\delta}}.
\end{aligned}$$

In the last line, we introduced the supremum with the help of (7.13) and (7.14), but we also of course need a similar estimate for $|\nabla h|$. To get the latter, we use the fact that h is harmonic in $V^\circ+\square_n$ and then apply Cauchy-Schwarz with the help of (7.22), similar to the derivation of (7.13) and (7.11) above, to obtain, for every $y\in V^\circ$,

$$|\nabla h(y)|^2 \le C\left(\fint_{y+\square_n}|\nabla h(x)|\,dx\right)^2 \le CM^2 3^{md}.$$

This completes the proof of (7.19).

Step 6. We give the proof of (7.17). To estimate the second term on the left side of (7.17), we use the identity

$$\begin{aligned}
&\fint_U \overline{F}\left(\nabla\widetilde{u}(x),\widetilde{\mathbf{g}}(x)\right)\,dx - \frac{1}{|U|}\int_{V^\circ}\overline{F}\left(\nabla u_{\mathrm{sa}}(x),\mathbf{g}_{\mathrm{sa}}(x)-\nabla h(x)\right)\,dx \\
&\qquad\qquad = \frac{1}{|U|}\int_{U\setminus V^\circ}\overline{F}\left(\nabla\widetilde{u}(x),\widetilde{\mathbf{g}}(x)\right)\,dx
\end{aligned}$$

and by (P1), (7.21) and the Hölder inequality,

$$\frac{1}{|U|}\left|\int_{U\setminus V^\circ}\overline{F}\left(\nabla\widetilde{u}(x),\widetilde{\mathbf{g}}(x)\right)\,dx\right| \le CM^2\left(\frac{|U\setminus V^\circ|}{|U|}\right)^{\delta_0/(2+\delta_0)} = CM^2 3^{-\alpha(n+m-l)}.$$

The first term on the left side of (7.17) is handled similarly, using (7.9) in place of (7.20) and (7.21). Set

$$V^{\circ\circ} := \left\{x\in V^\circ \,:\, x+\square_n\subseteq V^\circ\right\}, \tag{7.25}$$

and observe that $|U \setminus V^{\circ\circ}| \le C3^{-(n+m-l)}|U|$. We obtain

$$\begin{aligned}\left| \fint_U F\left(\nabla u(x), \mathbf{g}(x), x\right) dx - \frac{1}{|U|} \int_{V^\circ} \fint_{x+\square_n} F\left(\nabla u(y), \mathbf{g}(y), y\right) dy\, dx \right| \\ \le \frac{1}{|V^\circ|} \int_{U \setminus V^{\circ\circ}} \left|F\left(\nabla u(x), \mathbf{g}(x), x\right)\right| dx \\ \le CM^2 3^{-\alpha(n+m-l)},\end{aligned}$$

which completes the proof of (7.17) and hence of (7.15). □

7.3. **The proof of Proposition 5.2.** We now derive Proposition 5.2 as a consequence of (7.15).

Proof of Proposition 5.2. By construction, $\mathbf{g}-\widetilde{\mathbf{g}} \in L^2_{\mathrm{sol},0}(U)$ and $u, \widetilde{u} \in f + H^1_0(U)$. This implies that

$$\int_U \nabla u(x) \cdot \mathbf{g}(x)\, dx = \int_U \nabla \widetilde{u}(x) \cdot \widetilde{\mathbf{g}}(x)\, dx = \int_U \nabla f(x) \cdot \mathbf{g}(x) = \int_U \nabla f(x) \cdot \widetilde{\mathbf{g}}(x).$$

According to the previous line, (7.1) and (7.15),

$$\begin{aligned}(7.26)\qquad &\fint_U \left(\overline{F}\left(\nabla \widetilde{u}(x), \widetilde{\mathbf{g}}(x)\right) - \nabla f(x) \cdot \widetilde{\mathbf{g}}(x)\right) dx \\ &\qquad \le \fint_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - \nabla f(x) \cdot \mathbf{g}(x)\right) dx + C\mathcal{E}'' = C\mathcal{E}'',\end{aligned}$$

where $\mathcal{E}'_{n,m,M}$ is as in (5.6) with $\rho := (2+\delta_0)/\delta_0$ and we denote

$$\mathcal{E}'' := M^2\left(\mathcal{E}'_{n,m,M} + 3^{-\alpha(m+n-l)}\right).$$

By (7.2), Proposition 2.15 and the fact that $\overline{u} \in f + H^1_0(U)$, we have

$$\begin{aligned}0 &= \fint_U \left(\overline{F}\left(\nabla \overline{u}(x), \overline{\mathbf{g}}(x)\right) - \nabla \overline{u}(x) \cdot \overline{\mathbf{g}}(x)\right) dx \\ &= \fint_U \left(\overline{F}\left(\nabla \overline{u}(x), \overline{\mathbf{g}}(x)\right) - \nabla f(x) \cdot \overline{\mathbf{g}}(x)\right) dx \\ &= \inf_{(u', \mathbf{g}') \in (f + H^1_0(U)) \times L^2_{\mathrm{sol},0}(U)} \fint_U \left(\overline{F}\left(\nabla u'(x), \mathbf{g}'(x)\right) - \nabla f(x) \cdot \mathbf{g}'(x)\right) dx.\end{aligned}$$

The functional inside the infimum on the last line is uniformly convex in the variable $(u', \mathbf{g}') \in (f + H^1_0(U)) \times L^2_{\mathrm{sol},0}(U)$, and therefore the previous display and (7.26) imply that

$$\fint_U \left(|\nabla \overline{u}(x) - \nabla \widetilde{u}(x)|^2 + |\overline{\mathbf{g}}(x) - \widetilde{\mathbf{g}}(x)|^2\right) dx \le C\mathcal{E}''$$

(c.f. the proof of Lemma 4.2). The Poincaré inequality then yields

$$3^{-2(n+m)} \fint_U |\overline{u}(x) - \widetilde{u}(x)|^2\, dx \le C\mathcal{E}''.$$

Using the identity $u - \widetilde{u} = \eta\,(u_{\mathrm{sa}} - u)$ and applying (7.20), we also have

$$3^{-2n} \fint_U |u(x) - \widetilde{u}(x)|^2\, dx \le CM^2.$$

The previous two lines and the triangle inequality imply

$$3^{-2(n+m)} \fint_U |u(x) - \overline{u}(x)|^2\, dx \le C\mathcal{E}'' + CM^2 3^{-2m},$$

which is (5.4). This completes the proof of the proposition. □

Appendix A. Consequences of the mixing conditions

In this section we review some basic consequences of the mixing conditions which are used in several key arguments in the paper. So that we may handle both mixing conditions at once, we assume that $\mathbb{P}$ is α-mixing with a rate given by a decreasing function $\phi : [0,\infty) \to [0,1]$: for all Borel subsets $U, V \subseteq \mathbb{R}^d$ and events $A \in \mathcal{F}_U$ and $B \in \mathcal{F}_V$, we have

$$|\mathbb{P}[A \cap B] - \mathbb{P}[A]\mathbb{P}[B]| \le \phi(\operatorname{dist}(U,V)). \tag{A.1}$$

We next reformulate this condition equivalently as a statement about the difference between the expectation of a product and the product of expectations. Throughout, we denote the $\mathbb{P}$-essential supremum of a random variable $|X|$ by

$$\|X\|_\infty := \inf\{\lambda > 0 : \mathbb{P}[|X| > \lambda] = 0\}.$$

Lemma A.1. *Assume that $\mathbb{P}$ satisfies (A.1). Fix $M, D > 0$. Consider Borel subsets $U_1, \dots, U_k \subseteq \mathbb{R}^d$ such that $\operatorname{dist}(U_i, U_j) \ge D$ for every $i \ne j$. Let $X_1, \dots, X_k$ be random variables on Ω such that, for each $i \in \{1, \dots, k\}$, X_i is $\mathcal{F}_{U_i}$–measurable. Then*

$$\left| \mathbb{E}\left[\prod_{i=1}^k X_i \right] - \prod_{i=1}^k \mathbb{E}[X_i] \right| \le 4(k-1) \left(\prod_{i=1}^k \|X_i\|_\infty \right) \phi(D).$$

Proof. It suffices by induction to prove the result for $k = 2$. We need to show that

$$\operatorname{cov}[X,Y] \le 4\|X\|_\infty \|Y\|_\infty \phi(\operatorname{dist}(U,V)), \tag{A.2}$$

provided that X is $\mathcal{F}_U$–measureable and Y is $\mathcal{F}_V$–measureable. To get this, we compute

$$\begin{aligned} \operatorname{cov}[X,Y] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \int_{-\|X\|_\infty}^{\|X\|_\infty} \int_{-\|Y\|_\infty}^{\|Y\|_\infty} \left(\mathbb{P}[X > s \text{ and } Y > t] - \mathbb{P}[X > s]\mathbb{P}[Y > t] \right) dt\, ds \end{aligned}$$

and observe that the integrand is bounded by $\phi(\operatorname{dist}(U,V))$. □

In the next two lemmas, we put Lemma A.1 in a more convenient form for its application in the proof of Lemma 6.5. The notation here for the cubes is the same as in Section 4.2.

Lemma A.2. *Assume that $\mathbb{P}$ satisfies (P2) and (A.1). Fix $n, k \in \mathbb{N}$ and let X be an $\mathcal{F}_{\square_n}$–measurable random variable satisfying $0 \le X \le 1$. Let X_z denote the translation of X by $z \in \mathbb{Z}^d$. Then there exists $C(d,k) \ge 1$ such that, for every $m \in \mathbb{N}$,*

$$\mathbb{E}\left[\left(3^{-dm} \sum_{z \in 3^n \mathbb{Z}^d \cap \square_{n+m}} X_z \right)^{2k} \right] \le C\left(\max\left\{ \mathbb{E}[X], 3^{-dm} \right\}^{2k} + \phi(3^n) \right).$$

Proof. We first separate the subcubes into 3^d distinct groups. We have

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{z\in 3^n\mathbb{Z}^d\cap \square_{n+m}} X_z\right)^{2k}\right]^{\frac{1}{2k}} &= \mathbb{E}\left[\left(\sum_{y\in\{-3^n,0,3^n\}^d}\ \sum_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} X_{y+z}\right)^{2k}\right]^{\frac{1}{2k}}\\ &\le \sum_{y\in\{-3^n,0,3^n\}^d} \mathbb{E}\left[\left(\sum_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} X_{y+z}\right)^{2k}\right]^{\frac{1}{2k}}\\ &= 3^d\,\mathbb{E}\left[\left(\sum_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} X_z\right)^{2k}\right]^{\frac{1}{2k}}.\end{aligned}$$

Therefore it suffices to analyze the sum on the right side. The benefit of the previous computation is that each of the distinct cubes in the collection $\{z+\square_n : z\in 3^{n+1}\mathbb{Z}^d\}$ are separated by a distance of more than 3^n.

We follow the classic method of moments: we expand the sum by writing

$$\mathbb{E}\left[\left(\sum_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} X_z\right)^{2k}\right] = \sum_{z_1,\dots,z_{2k}\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \mathbb{E}\left[X_{z_1}\cdots X_{z_{2k}}\right]$$

and proceed by analyzing each term in the sum according to how many distinct entries it has. We write $(z_1,\dots,z_{2k})\in H_j$ if it contains exactly j distinct entries. For convenience, set $N := |3^{n+1}\mathbb{Z}^d\cap \square_{n+m}|$ and note that $c3^{dm}\le N\le C3^{dm}$.

We next apply Lemma A.1 to each element $(z_1,\dots,z_{2k})\in H_j$ to separate its j distinct entries and use the crude bound $\mathbb{E}[X^s]\le \mathbb{E}[X]$, for $s\ge 1$, which follows from the $\mathbb{P}$–a.s. bound $|X|\le 1$. We get:

$$\mathbb{E}\left[X_{z_1}\cdots X_{z_{2k}}\right]\le \mathbb{E}[X]^j + 4j\phi(3^n) \quad \text{for every } (z_1,\dots,z_{2k})\in H_j.$$

We next estimate the number of elements in H_j, which is a simple combinatorics exercise. A crude upper bound is

$$|H_j|\le \binom{N}{j} j^{2k} \le \frac{N^j j^{2k}}{j!},$$

which we see from the fact that there are $\binom{N}{j}$ different ways to select j distinct elements from a finite set of size N, and then at most j^{2k} ways to make a $2k$-tuple from them. Using Stirling's inequality

$$j!\ge j^j\exp(-j),$$

we obtain

$$|H_j|\le \exp(j)N^j(2k)^{2k-j}.$$

We deduce that

$$\begin{aligned}\sum_{z_1,\dots,z_{2k}\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \mathbb{E}\left[X_{z_1}\cdots X_{z_{2k}}\right] &= \sum_{j=1}^{2k} \sum_{(z_1,\dots,z_{2k})\in H_j} \mathbb{E}\left[X_{z_1}\cdots X_{z_{2k}}\right] \\ &\le \sum_{j=1}^{2k} |H_j| \left(\mathbb{E}[X]^j + 4j\phi(3^n)\right) \\ &\le \sum_{j=1}^{2k} \exp(j) N^j (2k)^{2k-j} \left(\mathbb{E}[X]^j + 4j\phi(3^n)\right).\end{aligned}$$

Set $\theta := \max\left\{4kN^{-1}\exp(-1), \mathbb{E}\left[X\right]\right\}$ and observe that

$$\begin{aligned}&\sum_{j=1}^{2k} \exp(j) N^j (2k)^{2k-j} \left(\mathbb{E}[X]^j + 4j\phi(3^n)\right) \\ &\qquad \le \sum_{j=1}^{2k} \exp(j) N^j (2k)^{2k-j} \left(\theta^j + 4(2k)\phi(3^n)\right) \\ &\qquad \le 2N^{2k}\exp(2k)\left(\theta^{2k} + 4k\phi(3^n)\right).\end{aligned}$$

This completes the argument. □

We next give the analogous statement to Lemma A.2 for exponential moments.

Lemma A.3. *Assume that $\mathbb{P}$ satisfies (P2) and (A.1). Fix $n \in \mathbb{N}$ and let X be an $\mathcal{F}_{\square_n}$–measurable random variable satisfying $0 \le X \le 1$. Let X_z denote the translation of X by $z \in \mathbb{Z}^d$. Then there exists $C(d) \ge 1$ such that, for every $m \in \mathbb{N}$ and $t \in [0,1]$,*

$$\log \mathbb{E}\left[\prod_{z\in 3^n\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_z\right)\right] \le C3^{dm}\left(t\mathbb{E}\left[X\right] + \phi(3^n)\exp\left(Ct3^{dm}\right)\right).$$

Proof. As in the previous lemma, we use that the family $\{z+\square_n : z \in 3^{n+1}\mathbb{Z}^d\}$ consists of disjoint cubes separated by a distance greater than 3^n. Fix $m \in \mathbb{N}$ and $t \ge 0$ and compute:

$$\begin{aligned}&\log \mathbb{E}\left[\prod_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_z\right)\right] \\ &\qquad \le \log\left(\prod_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \mathbb{E}\left[\exp(tX_z)\right] + C3^{dm}\exp\left(Ct3^{dm}\right)\phi(3^n)\right) \\ &\qquad \le \sum_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \log \mathbb{E}\left[\exp(tX_z)\right] + C3^{dm}\exp\left(Ct3^{dm}\right)\phi(3^n) \\ &\qquad = C3^{dm}\log \mathbb{E}\left[\exp(tX)\right] + C3^{dm}\exp\left(Ct3^{dm}\right)\phi(3^n).\end{aligned}$$

In the above string of inequalities, we used Lemma A.1 in the first line, the elementary inequality $\log(s+t) \le t + \log s$ (which is valid for every $s \ge 1$ and $t \ge 0$) in the second line and stationarity in the third line. Next we use Hölder's

inequality and stationarity once more to obtain, for every $t \geq 0$,

$$\begin{aligned}\log \mathbb{E}\left[\prod_{z\in 3^n\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_z\right)\right] &= \log \mathbb{E}\left[\prod_{y\in\{-3^n,0,3^n\}^d}\ \prod_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_{y+z}\right)\right] \\ &\leq \log \prod_{y\in\{-3^n,0,3^n\}^d} \mathbb{E}\left[\prod_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \exp\left(3^d tX_{y+z}\right)\right]^{3^{-d}} \\ &= \log \mathbb{E}\left[\prod_{z\in 3^{n+1}\mathbb{Z}^d\cap \square_{n+m}} \exp\left(3^d tX_z\right)\right].\end{aligned}$$

Combining the above inequalities yields, for every $m \in \mathbb{N}$ and $t \geq 0$,

$$\log \mathbb{E}\left[\prod_{z\in 3^n\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_z\right)\right] \leq C3^{dm}\left(\log \mathbb{E}\left[\exp\left(t3^d X\right)\right] + \phi(3^n)\exp\left(Ct3^{dm}\right)\right).$$

We now use the elementary inequalities

$$\begin{cases} \exp(s) \leq 1 + Cs & \text{for every } 0 \leq s \leq 3^d, \\ \log(1+s) \leq s & \text{for every } s \geq 0, \end{cases}$$

and the fact that $0 \leq X \leq 1$ to obtain, for every $m \in \mathbb{N}$ and $0 \leq t \leq 1$,

$$\log \mathbb{E}\left[\prod_{z\in 3^n\mathbb{Z}^d\cap \square_{n+m}} \exp\left(tX_z\right)\right] \leq C3^{dm}\left(t\mathbb{E}\left[X\right] + \phi(3^n)\exp\left(Ct3^{dm}\right)\right).$$

This completes the proof. □

Appendix B. Basic energy estimates

In this appendix we record some regularity estimates needed in the paper, in particular the Caccioppoli inequality and both local and global versions of the Meyers estimate.

The statements here are entirely deterministic, so throughout we fix $F \in \Omega$ such that F satisfies the inequality in (P1), that is,

$$\begin{aligned}\text{(B.1)}\qquad \frac{1}{2\Lambda}\left(|p|^2 + |q|^2\right) - K_0(1 + |p| + |q|) &\leq F(p,q,x) \\ &\leq \frac{\Lambda}{2}\left(|p|^2 + |q|^2\right) + K_0(1 + |p| + |q|).\end{aligned}$$

We also fix a bounded Lipschitz domain U and let $\mathcal{J} : H^1(U) \times H^{-1}(U) \to \mathbb{R}$ be the functional defined in (1.13).

We begin with a simple L^2 energy estimate.

Proposition B.1. *Suppose $u, v \in H^1(U)$ and $u^*, v^* \in H^{-1}(U)$ are such that $u - v \in H_0^1(U)$ and*

$$\text{(B.2)}\qquad \mathcal{J}\left[u, u^*\right] = \mathcal{J}\left[v, v^*\right] = 0.$$

Then there exists a constant $C(d, \Lambda, U) \geq 1$ such that

$$\left\|u - v\right\|_{H^1(U)} \leq C\left\|u^* - v^*\right\|_{H^{-1}(U)}.$$

Proof. Select $\mathbf{h} \in L^2(U;\mathbb{R}^d)$ to satisfy

$$-\nabla\cdot\mathbf{h} = v^* \quad \text{and} \quad 0 = \fint_U \left(F\left(\nabla v(x),\mathbf{h}(x),x\right) - \nabla v(x)\cdot\mathbf{h}(x)\right)dx \tag{B.3}$$

and recall that, by Proposition 2.15,

$$F\left(\nabla v(x),\mathbf{h}(x),x\right) - \nabla v(x)\cdot\mathbf{h}(x) \quad \text{a.e. in } U. \tag{B.4}$$

Let $h \in H^1(U)$ denote the (unique up to a constant) solution of

$$\begin{cases} -\Delta h = u^* - v^* & \text{in } U, \\ \partial_\nu h = |\partial U|^{-1}\langle 1, u^* - v^*\rangle & \text{on } \partial U. \end{cases}$$

According to [15, Theorem 1.2], we have

$$\|\nabla h\|_{L^2(U)} \le C\,\|u^* - v^*\|_{H^{-1}(U)}. \tag{B.5}$$

Using (B.1), (1.11) and (B.4), we have that

$$\begin{aligned} \mathcal{J}\left[v,u^*\right] &\le \fint_U \left(F\left(\nabla v(x),\mathbf{h}(x)-\nabla h(x),x\right) - \nabla v(x)\cdot\left(\mathbf{h}(x)-\nabla h(x)\right)\right)dx \\ &\le C\fint_U |\nabla h(x)|^2\,dx \le C\,\|u^*-v^*\|_{H^{-1}(U)}^2. \end{aligned}$$

Now we use (B.1), the uniform convexity of $w \mapsto \mathcal{J}\left[w,u^*\right]$ on $u + H_0^1(U)$, the assumption $\mathcal{J}\left[u,u^*\right] = 0$ and the Poincaré inequality to obtain the lemma. □

We next verify the interior Caccioppoli inequality. The variational formulation yields a particularly simple proof.

Proposition B.2. *Fix a domain V such that $\overline{V} \subseteq U$. Suppose that $u \in H^1(U)$ and $u^* \in H^{-1}(U)$ satisfy*

$$\mathcal{J}\left[u,u^*\right] = 0.$$

Then there exists a constant $C(d,\Lambda,U,V) \ge 1$ such that

$$\|\nabla u\|_{L^2(V)} \le C\left(K_0 + \|u\|_{L^2(U)} + \|u^*\|_{H^{-1}(U)}\right). \tag{B.6}$$

Proof. By Proposition B.1, it suffices to consider the case $u^* = 0$. Select a solenoidal vector field $\mathbf{g} \in L^2_{\mathrm{sol}}(U)$ to satisfy

$$0 = \fint_U \left(F\left(\nabla u(x),\mathbf{g}(x),x\right) - \nabla u(x)\cdot\mathbf{g}(x)\right)dx.$$

By Proposition 2.15,

$$F\left(\nabla u(x),\mathbf{g}(x),x\right) = \nabla u(x)\cdot\mathbf{g}(x) \quad \text{a.e. in } U. \tag{B.7}$$

Set $v := (1+\eta^2)u$, where $\eta \in C_c^\infty(\mathbb{R}^d)$ is a smooth test function satisfying

$$0 \le \eta \le 1, \quad \eta \equiv 1 \text{ on } \overline{V}, \quad \operatorname{supp}\eta \subseteq U, \quad \sup_{x\in U}|\nabla\eta(x)| \le C. \tag{B.8}$$

Using that $\mathbf{g}$ is solenoidal and $u - v \in H_0^1(U)$ as well as (B.7), we obtain

$$\begin{aligned} 0 &= \int_U \mathbf{g}(x)\cdot\left(\nabla v(x) - \nabla u(x)\right)dx \\ &= \int_U \mathbf{g}(x)\cdot\left(\eta^2(x)\nabla u(x) + 2u(x)\eta(x)\nabla\eta(x)\right)dx \\ &\ge \int_U \left(F(\nabla u(x),\mathbf{g}(x),x)\eta^2(x) - C\eta(x)\,|\mathbf{g}(x)|\,|u(x)|\right)dx. \end{aligned}$$

Using that F satisfies (B.1) and applying Young's inequality, we get

$$\begin{aligned}\int_U \left(|\nabla u(x)|^2 + |\mathbf{g}(x)|^2\right)\eta^2(x)\,dx &\le CK_0^2 + C\int_U F(\nabla u(x), \mathbf{g}(x), x)\eta^2(x)\,dx \\ &\le CK_0^2 + C\int_U \eta(x)\,|\mathbf{g}(x)|\,|u(x)|\,dx \\ &\le CK_0^2 + \frac{1}{2}\int_U |\mathbf{g}(x)|^2\,\eta^2(x) + C\int_U |u(x)|^2\,dx.\end{aligned}$$

A rearrangement yields the proposition. □

The proof of the Meyers estimate follows the argument of Giaquinta and Giusti [18] which requires a version of the Gehring lemma in order to obtain some improvement of integrability. We use the one given in Giaquinta and Modica [19, Proposition 5.1] which can also be found in [20, Theorem 6.6].

Lemma B.3. *Fix a domain V such that $\overline{V} \subseteq U$. Let $K > 0$, $\delta > 0$, $0 < m < 1$, $f \in L^1(U)$ and $g \in L^{1+\delta}(U)$. Suppose that, for every $r > 0$ and $x \in \mathbb{R}^d$ such that $B_{2r}(x) \subseteq U$, we have*

$$\fint_{B_r(x)} |f(x)|\,dx \le K\left(\left(\fint_{B_{2r}(x)} |f(x)|^m\,dx\right)^{\frac{1}{m}} + \fint_{B_{2r}(x)} |g(x)|\,dx\right).$$

Then there exist $\delta_0(d, m, K, \delta) \in (0, \delta]$ and a constant $C(d, m, K, U, V) \ge 1$ such that $f \in L^{1+\delta_0}(V)$ and

$$\int_V |f(x)|^{1+\delta_0}\,dx \le C\left(\left(\int_U |f(x)|\,dx\right)^{1+\delta_0} + \int_U |g(x)|^{1+\delta_0}\,dx\right).$$

We next present a version of the interior Meyers estimate.

Proposition B.4. *Suppose that $u \in H^1(U)$ satisfies*

$$\mathcal{J}[u, 0] = 0.$$

Fix a domain $V \subseteq \mathbb{R}^d$ such that $\overline{V} \subseteq U$. Then there exist $\delta_0(d, \Lambda) > 0$ and $C(d, \Lambda, U, V) \ge 1$ such that

$$\|\nabla u\|_{L^{2+\delta_0}(V)} \le C\left(K_0 + \|u\|_{L^2(U)}\right).$$

Proof. Fix $x \in U$ and $r > 0$ such that $x + 2r\square_0 \subseteq U$. By (a properly rescaled version of) Proposition B.2 and the Poincaré-Sobolev inequality (cf. [20, Theorem 3.15]),

$$\begin{aligned}\fint_{x+r\square_0} |\nabla u(x)|^2\,dx &\le C\left(K_0^2 + r^{-2}\fint_{x+2r\square_0} \left|u(x) - \fint_{x+2r\square_0} u(y)\,dy\right|^2\,dx\right) \\ &\le C\left(K_0^2 + \left(\fint_{x+2r\square_0} |\nabla u(x)|^{\frac{2d}{d+2}}\,dx\right)^{\frac{d+2}{d}}\right).\end{aligned}$$

Note that the constant C here depends only on (d, Λ). Thus the hypotheses of Lemma B.3 hold for the function $(K_0^2 + |\nabla u|^2)$ with $m = d/(d+2)$ and $K = C(d, \Lambda)$, and so an application of the lemma yields the result. □

The rest of this appendix is concerned with obtaining a version of Meyers' estimate which is valid up to the boundary of a Lipschitz domain. For this we adapt the classical argument which can be found for example in Giusti [20]. We begin with a variant of the Caccioppoli inequality which holds for balls centered in U but which may intersect ∂U.

Proposition B.5. *Suppose that $f \in H^1(U)$ and that $u \in f + H_0^1(U)$ satisfies*

$$\mathcal{J}[u,0] = 0.$$

Then there exists a constant $C(d,\Lambda,U) \geq 1$ such that, for every $z \in U$ and $r > 0$

$$\|\nabla u\|_{L^2(B_r(z)\cap U)} \leq C\left(K_0 + r^{-1}\|u-f\|_{L^2(B_{2r}(z)\cap U)} + \|\nabla f\|_{L^2(B_{2r}(z)\cap U)}\right). \tag{B.9}$$

Proof. Select a solenoidal vector field $\mathbf{g} \in L^2_{\mathrm{sol}}(U)$ to satisfy

$$0 = \fint_U \left(F\left(\nabla u(x), \mathbf{g}(x), x\right) - \nabla u(x)\cdot \mathbf{g}(x)\right) dx$$

and recall that, by Proposition 2.15,

$$F\left(\nabla u(x), \mathbf{g}(x), x\right) = \nabla u(x)\cdot \mathbf{g}(x) \quad \text{a.e. in } U. \tag{B.10}$$

We consider the test function $v := u + \eta^2(u-f)$, where $\eta \in C_c^\infty(B_{2r}(z))$ is a smooth cutoff function satisfying

$$0 \leq \eta \leq 1, \quad \eta \equiv 1 \text{ on } \overline{B}_r(z), \quad \operatorname{supp}\eta \subseteq B_{2r}(z), \quad \sup_{x\in B_{2r}(z)} |\nabla \eta(x)| \leq Cr^{-1}.$$

Using that $\mathbf{g}$ is solenoidal, $u - v \in H_0^1(B_{2r}(z)\cap U)$ and (B.10), we get

$$\begin{aligned}
0 &= \int_{B_{2r}(z)\cap U} \mathbf{g}(x)\cdot\left(\nabla v(x) - \nabla u(x)\right) dx \\
&= \int_{B_{2r}(z)\cap U} \mathbf{g}(x)\cdot\left(\eta^2(x)(\nabla u(x) - \nabla f(x)) + 2(u(x)-f(x))\eta(x)\nabla\eta(x)\right) dx \\
&\geq \int_{B_{2r}(z)\cap U} F(\nabla u(x), \mathbf{g}(x), x)\eta^2(x)\,dx - \int_{B_{2r}(z)\cap U} \eta^2(x)\,|\mathbf{g}(x)|\,|\nabla f(x)|\,dx \\
&\qquad - Cr^{-1}\int_{B_{2r}(z)\cap U} \eta(x)\,|\mathbf{g}(x)|\,|u(x)-f(x)|\,dx.
\end{aligned}$$

Using (B.1) and Young's inequality, we get

$$\begin{aligned}
&\int_{B_{2r}(z)\cap U} \left(|\nabla u(x)|^2 + |\mathbf{g}(x)|^2\right)\eta^2(x)\,dx \\
&\qquad \leq CK_0^2 + C\int_{B_{2r}(z)\cap U} F(\nabla u(x), \mathbf{g}(x), x)\eta^2(x)\,dx \\
&\qquad \leq CK_0^2 + C\int_{B_{2r}(z)\cap U} \eta^2(x)\,|\mathbf{g}(x)|\,|\nabla f(x)|\,dx \\
&\qquad\qquad + Cr^{-1}\int_{B_{2r}(z)\cap U} \eta(x)\,|\mathbf{g}(x)|\,|u(x)-f(x)|\,dx \\
&\qquad \leq CK_0^2 + C\int_{B_{2r}(z)\cap U} |\mathbf{g}(x)|^2\,\eta^2(x)\,dx + C\int_{B_{2r}(z)\cap U} \eta^2(x)|\nabla f(x)|^2\,dx \\
&\qquad\qquad + Cr^{-2}\int_{B_{2r}(z)\cap U} |u(x)-f(x)|^2\,dx.
\end{aligned}$$

This completes the proof of the proposition. □

Proposition B.6. *Fix $\delta > 0$, $f \in W^{1,2+\delta}(U)$ and $u \in f + H_0^1(U)$ such that*

$$\mathcal{J}[u, 0] = 0.$$

Then there exist $\delta_0(d, \Lambda, \delta) \in (0, \delta]$ and $C(d, \Lambda, \delta, U) \geq 1$ such that

$$\|\nabla u\|_{L^{2+\delta_0}(U)} \leq C\left(\|u\|_{L^2(U)} + \|f\|_{W^{1,2+\delta}(U)}\right).$$

Proof. According to the Sobolev extension theorem for bounded Lipschitz domains (cf. [1, Theorem 4.32]), we may assume that $f \in W^{1,2+\delta}(\mathbb{R}^d)$ and

$$\|f\|_{W^{1,2+\delta}(\mathbb{R}^d)} \leq C\|f\|_{W^{1,2+\delta}(U)}. \tag{B.11}$$

We may also extend u to $\mathbb{R}^d$ by setting $u := f$ in $\mathbb{R}^d \setminus U$.

We first claim that, for every $z \in \mathbb{R}^d$ and $r > 0$, we have

$$\begin{aligned} &\fint_{B_r(z)} |\nabla u(x) - \nabla f(x)|^2 \, dx \\ &\leq C\left(K_0^2 + \left(\fint_{B_{2r}(z)} |\nabla u(x) - \nabla f(x)|^{\frac{2d}{d+2}} \, dx\right)^{\frac{d+2}{d}} + \fint_{B_{2r}(z)} |\nabla f(x)|^2 \, dx\right), \end{aligned} \tag{B.12}$$

We split the argument into two cases: $B_{2r}(z) \setminus U = \varnothing$ or $B_{2r}(z) \setminus U \neq \varnothing$. In the former case, we use the interior Caccioppoli estimate to obtain, as in the proof of Proposition B.4, that

$$\fint_{B_r(z)} |\nabla u(x)|^2 \, dx \leq C\left(K_0^2 + \left(\fint_{B_{2r}(z)} |\nabla u(x)|^{\frac{2d}{d+2}} \, dx\right)^{\frac{d+2}{d}}\right),$$

which of course implies (B.12). If $B_{2r}(z) \setminus U \neq \varnothing$, then, since U is Lipschitz, we have

$$|B_{4r}(z) \setminus U| \geq c|B_{4r}(z)|.$$

Since $u - f \equiv 0$ on $\mathbb{R}^d \setminus U$, the Sobolev-Poincaré inequality (the version we need can be found in [20, Theorem 3.16]) gives

$$r^{-2} \fint_{B_{4r}(z)} |u(x) - f(x)|^2 \, dx \leq C\left(\fint_{B_{4r}(z)} |\nabla u(x) - \nabla f(x)|^{\frac{2d}{d+2}}\right)^{\frac{d+2}{2d}}.$$

Then according to (B.9), for every $z \in \mathbb{R}^d$ and $r > 0$,

$$\|\nabla u - \nabla f\|_{L^2(B_r(z))} \leq C\left(K_0 + r^{-1}\|u - f\|_{L^2(B_{2r}(z))} + \|\nabla f\|_{L^2(B_{2r}(z))}\right).$$

The combination of the previous two inequalities gives (B.12), except that the balls on the right side have radius $4r$ rather than $2r$. This can of course be removed by a simple covering argument.

Applying Lemma B.3 (with $V = U$ and $U = U + B_1$), in view of (B.11) and (B.12), now gives the result. □

References

[1] R. A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] S. N. Armstrong and Z. Shen. Lipschitz estimates in almost-periodic homogenization. Preprint, arXiv:1409.2094.

[3] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér.* to appear, arXiv:1406.0996.

[4] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.

[5] M. Avellaneda and F.-H. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8-9):897–910, 1991.

[6] H. H. Bauschke and X. Wang. The kernel average for two convex functions and its application to the extension and representation of monotone operators. *Trans. Amer. Math. Soc.*, 361(11):5947–5965, 2009.

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

[8] J. Ben-Artzi, D. Marahrens, and S. Neukamm. Moment bounds on the corrector of stochastic homogenization of non-symmetric elliptic finite difference equations. Preprint, arXiv:1407.6984.

[9] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aii, A971–A974, 1976.

[10] J. G. Conlon and A. Naddaf. On homogenization of elliptic equations with random coefficients. *Electron. J. Probab.*, 5:no. 9, 58 pp. (electronic), 2000.

[11] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.

[12] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.

[13] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[14] S. Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.

[15] J. Geng and Z. Shen. The Neumann problem and Helmholtz decomposition in convex domains. *J. Funct. Anal.*, 259(8):2147–2164, 2010.

[16] N. Ghoussoub. *Self-dual partial differential systems and their variational principles*. Springer Monographs in Mathematics. Springer, New York, 2009.

[17] N. Ghoussoub, A. Moameni, and R. Z. Sáiz. Homogenization of maximal monotone vector fields via selfdual variational calculus. *Adv. Nonlinear Stud.*, 11(2):323–360, 2011.

[18] M. Giaquinta and E. Giusti. On the regularity of the minima of variational integrals. *Acta Math.*, 148:31–46, 1982.

[19] M. Giaquinta and G. Modica. Regularity results for some classes of higher order nonlinear elliptic systems. *J. Reine Angew. Math.*, 311/312:145–169, 1979.

[20] E. Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[21] A. Gloria and D. Marahrens. Annealed estimates on the Green functions and uncertainty quantification, Preprint.

[22] A. Gloria, S. Neukamm, and F. Otto. An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *ESAIM Math. Model. Numer. Anal.*, 48(2):325–346, 2014.

[23] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.
[24] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators, Preprint, arXiv:1409.2678.
[25] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.
[26] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.
[27] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization, Preprint.
[28] C. Kenig, F. Lin, and Z. Shen. Convergence rates in L^2 for elliptic homogenization problems. *Arch. Rational Mech. Anal.*, 203(3):1009–1036, 2012.
[29] C. Kenig, F. Lin, and Z. Shen. Homogenization of elliptic systems with Neumann boundary conditions. *J. Amer. Math. Soc.*, 26:901–937, 2013.
[30] A. Lamacz, S. Neukamm, and F. Otto. Moment bounds for the corrector in stochastic homogenization of a percolation model. Preprint, arXiv:1406.5723.
[31] D. Marahrens and F. Otto. Annealed estimates on the Green's function. Preprint, arXiv:1304.4408.
[32] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):294–327, 2011.
[33] A Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, Unpublished preprint.
[34] B. Nayroles. Deux théorèmes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aiv, A1035–A1038, 1976.
[35] J. Nolen. Normal approximation for a random elliptic equation. *Probab. Theory Related Fields*, 2014 (to appear).
[36] J.-P. Penot. The relevance of convex analysis for the study of monotonicity. *Nonlinear Anal.*, 58(7-8):855–871, 2004.
[37] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
[38] A. Visintin. Scale-transformations and homogenization of maximal monotone relations with applications. *Asymptot. Anal.*, 82(3-4):233–270, 2013.
[39] A. Visintin. Variational formulation and structural stability of monotone equations. *Calc. Var. Partial Differential Equations*, 47(1-2):273–317, 2013.

(S.N. Armstrong) Ceremade (UMR CNRS 7534), Université Paris-Dauphine, Paris, France
E-mail address: armstrong@ceremade.dauphine.fr

(J.-C. Mourrat) Ecole normale supérieure de Lyon, CNRS, Lyon, France
E-mail address: jean-christophe.mourrat@ens-lyon.fr