ROBUSTNESS OF THE PATHWISE STRUCTURE OF FLUCTUATIONS IN STOCHASTIC HOMOGENIZATION

MITIA DUERINCKX, ANTOINE GLORIA, AND FELIX OTTO

Abstract. We consider a linear elliptic system in divergence form with random coefficients and study the random fluctuations of large-scale averages of the field and the flux of the solution operator. In the context of the random conductance model, we developed in a previous work a theory of fluctuations based on the notion of homogenization commutator: we proved that the two-scale expansion of this special quantity is accurate at leading order in the fluctuation scaling when averaged on large scales (as opposed to the two-scale expansion of the solution operator taken separately) and that the large-scale fluctuations of the field and the flux of the solution operator can be recovered from those of the commutator. This implies that the large-scale fluctuations of the commutator of the corrector drive all other large-scale fluctuations to leading order, which we refer to as the *pathwise structure* of fluctuations in stochastic homogenization. In the present contribution we extend this result in two directions: we treat *continuum* elliptic (possibly non-symmetric) systems and allow for strongly *correlated* coefficient fields. Our main result shows in this general setting that the two-scale expansion of the homogenization commutator is still accurate to leading order when averaged on large scales, which illustrates the robustness of the pathwise structure of fluctuations.

Contents

1. Introduction

This article constitutes the second part of a series that develops a theory of fluctuations in stochastic homogenization of linear elliptic (non-necessarily symmetric) systems in divergence form. In the first part [8], we provided a complete picture of our theory (with optimal error estimates and convergence rates) in the simplified framework of the random conductance model. We proved three main results: the pathwise structure of fluctuations, their asymptotic normality, and the identification of the limiting covariance structure. In the present contribution, we focus on the fundamental pathwise aspect of the theory, that is, the accuracy of the two-scale expansion for large-scale fluctuations of the so-called homogenization commutator, and we extend its validity to continuum (non-symmetric) systems with strongly correlated coefficient fields. More precisely, we cover the general setting of coefficient fields satisfying multiscale functional inequalities as introduced in [6, 7], and therefore treat all the models considered in the reference textbook [19] on heterogenous materials. We take this as a sign of the robustness of the pathwise structure. Questions regarding the scaling limit of the standard homogenization commutator require more detailed probabilistic assumptions and are addressed in the forthcoming contribution [5] in the case of correlated Gaussian fields (see below for an informal discussion of these results). In [9], we further explain how this whole theory of fluctuations naturally extends to higher orders. We refer to the introduction of the companion article [8] for a general discussion of the literature on fluctuations in stochastic homogenization (a short discussion of the key pathwise structure is given at the end of this introduction).

Let $\boldsymbol{a}$ be a stationary and ergodic random coefficient field on $\mathbb{R}^d$ that is bounded in the sense of

$$|\boldsymbol{a}(x)\xi| \le |\xi| \qquad \text{for all } \xi \in \mathbb{R}^d \text{ and } x \in \mathbb{R}^d, \tag{1.1}$$

and satisfies the ellipticity property

$$\int_{\mathbb{R}^d} \nabla u \cdot \boldsymbol{a} \nabla u \ge \lambda \int_{\mathbb{R}^d} |\nabla u|^2 \qquad \text{for all } u \in C_c^\infty(\mathbb{R}^d), \tag{1.2}$$

for some $\lambda > 0$; this notion of functional coercivity is weaker than pointwise ellipticity for systems. Throughout the article we use scalar notation, but no iota in the proofs would change for systems under assumptions (1.1) and (1.2). For all $\varepsilon > 0$, we set $\boldsymbol{a}_\varepsilon := \boldsymbol{a}(\frac{\cdot}{\varepsilon})$, and for a deterministic vector field $f \in C_c^\infty(\mathbb{R}^d)^d$ we consider the random family $(\nabla u_\varepsilon)_{\varepsilon>0}$ of unique gradient solutions in $\mathrm{L}^2(\mathbb{R}^d)^d$ of the rescaled problems

$$-\nabla \cdot \boldsymbol{a}_\varepsilon \nabla u_\varepsilon \,=\, \nabla \cdot f. \tag{1.3}$$

(The choice of considering an equation on the whole space rather than on a bounded set allows us to focus on fluctuations in the bulk, and avoid effects of boundary layers. The choice of taking a right-hand side in divergence form allows to treat all dimensions at once.) It is known since the pioneering work of Papanicolaou and Varadhan [17] and Kozlov [15] that, almost surely, ∇u_ε converges weakly in $\mathrm{L}^2(\mathbb{R}^d)^d$ as $\varepsilon \downarrow 0$ to the unique gradient solution $\nabla \bar u \in \mathrm{L}^2(\mathbb{R}^d)^d$ of

$$-\nabla \cdot \bar{\boldsymbol{a}} \nabla \bar u \,=\, \nabla \cdot f, \tag{1.4}$$

where $\bar{\boldsymbol{a}}$ is a deterministic and constant matrix that only depends on the law of $\boldsymbol{a}$. More precisely, for any direction $e \in \mathbb{R}^d$, the projection $\bar{\boldsymbol{a}}e$ is the expectation of the flux of the corrector in the direction e,

$$\bar{\boldsymbol{a}} e \,=\, \mathbb{E}\left[\boldsymbol{a}(\nabla \phi_e + e)\right], \tag{1.5}$$

where the corrector ϕ_e is the unique (up to a random additive constant) almost-sure solution of the corrector equation in $\mathbb{R}^d$,

$$-\nabla \cdot \boldsymbol{a}(\nabla \phi_e + e) \,=\, 0,$$

in the class of functions the gradient of which is stationary, has finite second moment, and has zero expectation. We denote by $\phi = (\phi_i)_{i=1}^d$ the vector field the entries of which are the correctors ϕ_i in the canonical directions e_i of $\mathbb{R}^d$.

In [8], we developed a complete theory of fluctuations in stochastic homogenization for the random conductance model (see also [14] for heuristic arguments). The key in our theory is to focus on the so-called *homogenization commutator* of the solution,

$$\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon, \tag{1.6}$$

and to study its relation to the *standard homogenization commutator* $\Xi := (\Xi_i)_{i=1}^d$, where the solution u_ε is replaced by $\boldsymbol{a}$-harmonic coordinates $x \mapsto x_i + \phi_i(x)$,

$$\Xi_i := \boldsymbol{a}(\nabla \phi_i + e_i) - \bar{\boldsymbol{a}}(\nabla \phi_i + e_i), \qquad \Xi_{ij} := (\Xi_i)_j. \tag{1.7}$$

In the framework of [8], we showed the following three crucial properties (which we reformulate here in the non-symmetric continuum setting):

(I) First and most importantly, the two-scale expansion of the homogenization commutator of the solution

$$\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon - \mathbb{E}\left[\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon\right] \approx \Xi_i(\tfrac{\cdot}{\varepsilon}) \nabla_i \bar{u} \tag{1.8}$$

is accurate in the fluctuation scaling in the sense that for all $g \in C_c^\infty(\mathbb{R}^d)^d$ and $q < \infty$,

$$\mathbb{E}\left[\Big| \int_{\mathbb{R}^d} g \cdot \big(\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon - \mathbb{E}\left[\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon\right]\big) - \int_{\mathbb{R}^d} g \cdot \Xi_i(\tfrac{\cdot}{\varepsilon}) \nabla_i \bar{u} \Big|^q\right]^{\frac{1}{q}} \\ \lesssim_{f,g,q} \varepsilon\, \mathbb{E}\left[\Big| \int_{\mathbb{R}^d} g \cdot \Xi_i(\tfrac{\cdot}{\varepsilon}) \nabla_i \bar{u} \Big|^q\right]^{\frac{1}{q}}, \tag{1.9}$$

up to a $|\log \varepsilon|$ factor in the critical dimension $d = 2$. This property is highly nontrivial and is due to the special form of the commutator (1.6).

(II) Second, both the fluctuations of the field ∇u_ε and of the flux $\boldsymbol{a}_\varepsilon \nabla u_\varepsilon$ can be recovered through *deterministic* projections of the fluctuations of the homogenization commutator (1.6), which shows that no information is lost by passing to the commutator. More precisely, the following elementary identities are easily checked,

$$\begin{aligned} \int_{\mathbb{R}^d} g \cdot (\nabla u_\varepsilon - \nabla \bar{u}) &= -\int_{\mathbb{R}^d} (\bar{\mathcal{P}}_H^* g) \cdot (\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon), \\ \int_{\mathbb{R}^d} g \cdot (\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla \bar{u}) &= \int_{\mathbb{R}^d} (\bar{\mathcal{P}}_L^* g) \cdot (\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon), \end{aligned} \tag{1.10}$$

in terms of the Helmholtz and Leray projections in $\mathrm{L}^2(\mathbb{R}^d)^d$,

$$\begin{aligned} \bar{\mathcal{P}}_H &:= \nabla(\nabla \cdot \bar{\boldsymbol{a}} \nabla)^{-1} \nabla \cdot, \qquad & \bar{\mathcal{P}}_L &:= \mathrm{Id} - \bar{\mathcal{P}}_H \bar{\boldsymbol{a}}, \\ \bar{\mathcal{P}}_H^* &:= \nabla(\nabla \cdot \bar{\boldsymbol{a}}^* \nabla)^{-1} \nabla \cdot, \qquad & \bar{\mathcal{P}}_L^* &:= \mathrm{Id} - \bar{\mathcal{P}}_H \bar{\boldsymbol{a}}^*, \end{aligned} \tag{1.11}$$

where $\bar{\boldsymbol{a}}^*$ denotes the transpose of $\bar{\boldsymbol{a}}$. Similarly, the fluctuations of the field $\nabla \phi$ and of the flux $\boldsymbol{a} \nabla \phi$ of the corrector are also determined by those of the standard commutator Ξ itself: indeed, the definition of Ξ yields $-\nabla \cdot \bar{\boldsymbol{a}} \nabla \phi_i = \nabla \cdot \Xi_i$ and

$\boldsymbol{a}(\nabla\phi_i+e_i)-\bar{\boldsymbol{a}}e_i = \Xi_i+\bar{\boldsymbol{a}}\nabla\phi_i$, to the effect of $\nabla\phi_i = -\bar{\mathcal{P}}_H\Xi_i$ and $\boldsymbol{a}(\nabla\phi_i+e_i)-\bar{\boldsymbol{a}}e_i = (\mathrm{Id}-\bar{\boldsymbol{a}}\bar{\mathcal{P}}_H)\Xi_i$ in the stationary sense, hence formally,

$$\begin{aligned} \int_{\mathbb{R}^d} F : \nabla\phi(\tfrac{\cdot}{\varepsilon}) &= -\int_{\mathbb{R}^d} \bar{\mathcal{P}}_H^* F : \Xi(\tfrac{\cdot}{\varepsilon}), \\ \int_{\mathbb{R}^d} F : \big(\boldsymbol{a}_\varepsilon(\nabla\phi(\tfrac{\cdot}{\varepsilon})+\mathrm{Id})-\bar{\boldsymbol{a}}\big) &= \int_{\mathbb{R}^d} \mathcal{P}_L^* F : \Xi(\tfrac{\cdot}{\varepsilon}), \end{aligned} \tag{1.12}$$

where $\bar{\mathcal{P}}_H^*$ and $\bar{\mathcal{P}}_L^*$ act on the second index of the tensor field F, and where we use the notation $M : M' = \sum_{1\le i,j\le d} M_{ij}M'_{ij}$ for the double inner product of matrices M, M'. A suitable sense to these formal identities is given in Corollary 1.

(III) Third, the standard homogenization commutator Ξ is an approximately local function of the coefficients $\boldsymbol{a}$, which allows to infer the large-scale behavior of Ξ from the large-scale behavior of $\boldsymbol{a}$ itself. This locality property is best seen when formally computing the so-called "vertical" derivatives of Ξ with respect to $\boldsymbol{a}$: Letting ϕ^* denote the corrector associated with the pointwise transpose field $\boldsymbol{a}^*$, and letting σ^* denote the corresponding flux corrector (cf. (2.10)), we obtain (cf. [8, equation (1.10)] and (2.15) below)

$$\begin{aligned} \frac{\partial}{\partial\boldsymbol{a}(x)}\Xi_{ij} = (\nabla\phi_j^*+e_j)\cdot\frac{\partial\boldsymbol{a}}{\partial\boldsymbol{a}(x)}(\nabla\phi_i+e_i) \\ -\nabla\cdot\bigg(\phi_j^*\frac{\partial\boldsymbol{a}}{\partial\boldsymbol{a}(x)}(\nabla\phi_i+e_i)\bigg)-\nabla\cdot\bigg((\phi_j^*\boldsymbol{a}+\sigma_j^*)\frac{\partial\nabla\phi_i}{\partial\boldsymbol{a}(x)}\bigg). \end{aligned}$$

In view of $\frac{\partial\boldsymbol{a}}{\partial\boldsymbol{a}(x)} = \delta(\cdot-x)$, the first right-hand side term reveals an exactly local dependence upon $\boldsymbol{a}$. The second term is exactly local as well, but since it is written in divergence form its contribution is negligible when integrating on large scales. The only non-local effect comes from the last term due to $\frac{\partial\nabla\phi}{\partial\boldsymbol{a}}$, which is given by the mixed derivative of the Green's function for $-\nabla\cdot\boldsymbol{a}\nabla$ and thus is expected to have only borderline integrable decay. However, it also appears inside a divergence, hence it is negligible when integrated on large scales.

Let us comment on the structure of fluctuations revealed in (I)–(II). Together with the two-scale expansion (1.9) of commutators, identities (1.10) and (1.12) imply that the fluctuations of ∇u_ε, $\boldsymbol{a}_\varepsilon\nabla u_\varepsilon$, $\nabla\phi(\tfrac{\cdot}{\varepsilon})$, and $\boldsymbol{a}_\varepsilon\nabla\phi(\tfrac{\cdot}{\varepsilon})$ are determined to leading order by those of $\Xi(\tfrac{\cdot}{\varepsilon})$, with error estimated in a strong norm in probability. We chose to refer to this key property as the "pathwise" structure of fluctuations in analogy with the language of SPDEs in order to emphasize that this result does not only compare probability laws of different objects (possibly constructed on different probability spaces), but compares these objects for the same realizations of the randomness (for the same "paths"), here in form of an error estimate at the level of stretched exponential moments. As emphasized in [8], besides its theoretical importance, this *pathwise structure* is bound to affect multi-scale computing and uncertainty quantification in an essential way. This result is indeed of the complexity-reducing type of the central results in homogenization, as it provides a description of fluctuations of a general solution by means of an off-line procedure using the standard commutator Ξ in form of a two-scale expansion. Next, in case of a weakly correlated coefficient field $\boldsymbol{a}$, we expect from property (III) that $\Xi(\tfrac{\cdot}{\varepsilon})$ displays the CLT scaling and that $\varepsilon^{-d/2}\Xi(\tfrac{\cdot}{\varepsilon})$ converges to a white noise; the pathwise structure (I)–(II) then allows to recover the known scaling limit results for the different quantities of interest in stochastic homogenization, as indeed shown in [8] for the random conductance model.

In the present contribution, we focus on the pathwise structure (I)–(II), in particular on the error estimate for the two-scale expansion (1.8) of the homogenization commutator, and we mainly consider the class of Gaussian coefficient fields with a covariance function that decays at infinity at some fixed (yet arbitrary) algebraic rate $(1+|x|)^{-\beta}$ parametrized by $\beta > 0$. We show that properties (I)–(II) still hold for this whole Gaussian class, which illustrates the surprising robustness of the pathwise structure with respect to the large-scale behavior of the homogenization commutator. Indeed, in dimension $d = 1$ (in which case the quantities under investigation are simpler and explicit[1]), two typical behaviors have been identified in terms of the scaling limit of the standard homogenization commutator Ξ, depending on the parameter β (cf. [4]),

- For $\beta > d = 1$: The standard commutator Ξ displays the CLT scaling and its rescaling $\varepsilon^{-\frac{1}{2}}\Xi(\frac{\cdot}{\varepsilon})$ converges in law to a non-degenerate white noise (Gaussian fluctuations, local limiting covariance structure), but the convergence rate is arbitrarily slow as β gets closer to $d = 1$.
- For $0 < \beta < d = 1$: The suitable rescaling $\varepsilon^{-\frac{\beta}{2}}\Xi(\frac{\cdot}{\varepsilon})$ converges along a subsequence to a fractional Gaussian field (Gaussian fluctuations, nonlocal limiting covariance structure, potentially no uniqueness of the limit). Note that a different, non-Gaussian behavior may also occur in degenerate cases (cf. [13, 16] and second item in Remark 2.1).

In particular, the pathwise result is shown to hold in both examples whereas the rescaled standard commutator does not necessarily converge to white noise or may even not converge at all. The identification of the scaling limit of the standard commutator is thus a separate question and is addressed in [5] in all dimensions for the whole range of values of $\beta > 0$, combining Malliavin calculus with techniques developed in [10]. More precisely, this work extends [4] to dimensions $d > 1$ in the following sense,

- For $\beta > d$: The rescaled commutator $\varepsilon^{-\frac{d}{2}}\Xi(\frac{\cdot}{\varepsilon})$ converges in law to a generically non-degenerate white noise.
- For $\beta < d$: The rescaled commutator $\varepsilon^{-\frac{\beta}{2}}\Xi(\frac{\cdot}{\varepsilon})$ converges along a subsequence to a generically non-degenerate fractional Gaussian field. Different limits can indeed be reached in general, unless the covariance function has a self-similar profile at infinity.

These results illustrate the fact that the standard commutator Ξ is an approximately local function of the random coefficient field $\boldsymbol{a}$ (cf. (III) above), which essentially allows to relate the scaling limit of the commutator with the scaling limit of the coefficient field itself (as in dimension $d = 1$). Interestingly, this also shows that the pathwise structure of fluctuations can in general not be reduced to a quantitative joint convergence in law since there might not even be any convergence in law to talk about in the first place.

Although we focus here for shortness on the model case of Gaussian coefficient fields, the arguments that we provide in this contribution are robust enough to cover the general setting of mutiscale functional inequalities introduced and studied in [6, 7], and therefore to treat all the models of random coefficient fields considered in the reference textbook [19]

[1] In dimension $d = 1$, the homogenization commutator indeed simply takes the form $\Xi(x) = \bar{\boldsymbol{a}}(1 - \frac{\bar{\boldsymbol{a}}}{\boldsymbol{a}(x)})$, which is exactly local with respect to $\boldsymbol{a}$.

on heterogenous materials (see indeed third item of Remark 2.1). This makes the results of this contribution not only of theoretical but also of practical interest.

Let us conclude this introduction with a short discussion of the recent literature concerning (I)–(III); we refer to [8, Section 1.4] for more detail. The pathwise structure (I)–(II) of fluctuations, which we extend here to the continuum setting with long-range correlations, was first formulated and proved by us in [8] for the random conductance model. A related form of (I)–(II) was conjectured in [14] within the variational and renormalization framework of [3, 1, 2], but it has not been made rigorous yet (nor does it appear in the textbook [2]). A variational quantity related to the standard commutator can be first traced back to [3], whereas its canonical form (1.7) used here was independently introduced in [1, 2] and [8] (there motivated by the seminal works of Murat and Tartar). The locality property of the standard commutator Ξ and its convergence to white noise were first established in [8] for the random conductance model, and in [1, 12] for the continuum setting with a finite range of dependence assumption, while the case of long-range correlations is first considered in our companion article [5].

2. Main results and structure of the proof

2.1. Notation and statement of the main results. For some $k \geq 1$ let a be an $\mathbb{R}^k$-valued Gaussian random field, constructed on a probability space $(\Omega, \mathbb{P})$ (with expectation $\mathbb{E}$), which is stationary and centered, and thus characterized by its covariance function

$$c(x) := \mathbb{E}\left[a(x) \otimes a(0)\right], \qquad c : \mathbb{R}^d \to \mathbb{R}^{k \times k}.$$

We assume that the covariance function decays algebraically at infinity in the sense that there exist $\beta, C_0 > 0$ such that for all $x \in \mathbb{R}^d$,

$$\frac{1}{C_0}(1+|x|)^{-\beta} \;\leq\; |c(x)| \;\leq\; C_0(1+|x|)^{-\beta}. \tag{2.1}$$

Given a map $h \in C^1_b(\mathbb{R}^k)^{d\times d}$, we define $\boldsymbol{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ by $\boldsymbol{a}(x) = h(a(x))$, and assume that it satisfies the conditions (1.1) and (1.2) almost surely. We then (abusively) call the coefficient field $\boldsymbol{a}$ *Gaussian with parameter* $\beta > 0$. If $\boldsymbol{a}$ is Gaussian with parameter β, then $\boldsymbol{a}$ is ergodic, hence we have existence and uniqueness of correctors ϕ and of the homogenized coefficients $\bar{\boldsymbol{a}}$ (cf. Lemma 2.2 below). From a technical point of view, we shall rely on (and frequently refer to) results and methods developed in [6, 10, 11].

Throughout the article, we use the notation $\lesssim_{(\dots)}$ (resp. $\gtrsim_{(\dots)}$) for $\leq C\times$ (resp. $\geq C\times$), where the multiplicative constant C depends on $d, \lambda, \beta, \|\nabla h\|_{\mathrm{L}^\infty}$, on the constant C_0 in (2.1), and on the additional parameters "$(\dots)$" if any. We write $\simeq_{(\dots)}$ when both $\lesssim_{(\dots)}$ and $\gtrsim_{(\dots)}$ hold. In an assumption, we use the notation $\ll_{(\dots)}$ for $\leq \frac{1}{C}\times$ for some (large enough) constant $C \simeq_{(\dots)} 1$.

We now define a string of random functionals that encode the fluctuations of the different objects of interest. The notation I is reserved to functionals involving the solution operator, and the notation J to functionals involving correctors; the subscript $_0$ is reserved to commutators, the subscript $_1$ to fields, and the subscript $_2$ to fluxes. We consider the fluctuations of the commutator $\boldsymbol{a}_\varepsilon \nabla u_\varepsilon - \bar{\boldsymbol{a}} \nabla u_\varepsilon$, of the field ∇u_ε, and of the flux $\boldsymbol{a}_\varepsilon \nabla u_\varepsilon$ of the solution to (1.3), as encoded by the (centered) random bilinear functionals $I_0^\varepsilon : (f, g) \mapsto I_0^\varepsilon(f, g)$,

$I_1^\varepsilon:(f,g)\mapsto I_1^\varepsilon(f,g)$, and $I_2^\varepsilon:(f,g)\mapsto I_2^\varepsilon(f,g)$ defined for all $f,g\in C_c^\infty(\mathbb{R}^d)^d$ by

$$\begin{aligned} I_0^\varepsilon(f,g) &:= \int_{\mathbb{R}^d} g\cdot(\boldsymbol{a}_\varepsilon\nabla u_\varepsilon-\bar{\boldsymbol{a}}\nabla u_\varepsilon-\mathbb{E}\left[\boldsymbol{a}_\varepsilon\nabla u_\varepsilon-\bar{\boldsymbol{a}}\nabla u_\varepsilon\right]),\\ I_1^\varepsilon(f,g) &:= \int_{\mathbb{R}^d} g\cdot\nabla(u_\varepsilon-\mathbb{E}\left[u_\varepsilon\right]),\\ I_2^\varepsilon(f,g) &:= \int_{\mathbb{R}^d} g\cdot\big(\boldsymbol{a}_\varepsilon\nabla u_\varepsilon-\mathbb{E}\left[\boldsymbol{a}_\varepsilon\nabla u_\varepsilon\right]\big).\end{aligned}$$

Likewise, we consider the fluctuations of the standard commutator $\Xi=(\boldsymbol{a}-\bar{\boldsymbol{a}})(\nabla\phi+\mathrm{Id})$, of the corrector field $\nabla\phi$, and of the corrector flux $\boldsymbol{a}(\nabla\phi+\mathrm{Id})$ as encoded by the (centered) random linear functionals $J_0^\varepsilon:F\mapsto J_0^\varepsilon(F)$, $J_1^\varepsilon:F\mapsto J_1^\varepsilon(F)$, and $J_2^\varepsilon:F\mapsto J_2^\varepsilon(F)$ defined for all $F\in C_c^\infty(\mathbb{R}^d)^{d\times d}$ by

$$\begin{aligned} J_0^\varepsilon(F) &:= \int_{\mathbb{R}^d} F(x):\Xi(\tfrac{x}{\varepsilon})\,dx,\\ J_1^\varepsilon(F) &:= \int_{\mathbb{R}^d} F(x):\nabla\phi(\tfrac{x}{\varepsilon})\,dx,\\ J_2^\varepsilon(F) &:= \int_{\mathbb{R}^d} F(x):\big(\boldsymbol{a}_\varepsilon(x)(\nabla\phi(\tfrac{x}{\varepsilon})+\mathrm{Id})-\bar{\boldsymbol{a}}\big)\,dx.\end{aligned}$$

We first prove the following boundedness result for J_0^ε, establishing the suitable β-dependent scaling for the fluctuations of the homogenization commutator (see also [11, Theorem 1]). More precisely, in the spirit of (III), this shows that large-scale averages of the standard commutator have the same scaling as large-scale averages of the coefficient field $\boldsymbol{a}$ itself (cf. [6, Proposition 1.5]); in the case of integrable correlations, this coincides with the CLT scaling $\varepsilon^{d/2}$.

Proposition 1 (Fluctuation scaling)**.** *Let $d\ge 1$, assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta>0$, define $\pi_*:\mathbb{R}^+\to\mathbb{R}^+$ by*

$$\pi_*(t) := \begin{cases} (1+t)^\beta & : \ \beta<d,\\ (1+t)^d\,\frac{1}{\log(2+t)} & : \ \beta=d,\\ (1+t)^d & : \ \beta>d,\end{cases} \tag{2.2}$$

and define the rescaled functional

$$\widehat{J}_0^\varepsilon := \pi_*(\tfrac{1}{\varepsilon})^{\frac12}J_0^\varepsilon.$$

For all $0<\varepsilon\le 1$, $F\in C_c^\infty(\mathbb{R}^d)^{d\times d}$, $0<p-1\ll 1$, and $\alpha>\frac{d-(\beta\wedge d)}{2}+d\frac{p-1}{2p}$, we have

$$|\widehat{J}_0^\varepsilon(F)| \le \mathcal{C}_{\alpha,p}^{\varepsilon,F}\left(\|w_1^\alpha F\|_{\mathrm{L}^{2p}}+\mathbb{1}_{\beta\le d}\|[F]_2\|_{\mathrm{L}^p}\right), \tag{2.3}$$

where $w_1(z):=1+|z|$, $[F]_2(x):=(\fint_{B(x)}|F|^2)^{\frac12}$, and where $\mathcal{C}_{\alpha,p}^{\varepsilon,F}$ is a random variable with stretched exponential moments: there exists $\gamma_1\simeq 1$ such that

$$\sup_{0<\varepsilon<1}\mathbb{E}\left[\exp\Big(\frac{1}{C_{\alpha,p}}(\mathcal{C}_{\alpha,p}^{\varepsilon,F})^{\gamma_1}\Big)\right] \le 2$$

for some (deterministic) constant $C_{\alpha,p}\simeq_{\alpha,p}1$. ◊

Our next main result establishes the accuracy of the two-scale expansion (1.8) for large-scale averages of the homogenization commutator in the suitable fluctuation scaling. This error is encoded by the following (centered) random bilinear functional,

$$E^\varepsilon(f,g) := \int_{\mathbb{R}^d} g\cdot\big(\boldsymbol{a}_\varepsilon\nabla u_\varepsilon-\bar{\boldsymbol{a}}\nabla u_\varepsilon-\mathbb{E}\left[\boldsymbol{a}_\varepsilon\nabla u_\varepsilon-\bar{\boldsymbol{a}}\nabla u_\varepsilon\right]\big)-\int_{\mathbb{R}^d} g\cdot\Xi(\tfrac{\cdot}{\varepsilon})\nabla\bar u$$
$$= I_0^\varepsilon(f,g)-J_0^\varepsilon(\nabla\bar u\otimes g) \;=\; I_0^\varepsilon(f,g)-J_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes g). \quad (2.4)$$

More precisely, we show that the typical scaling of this error $I_0^\varepsilon(f,g)-J_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes g)$ is an order $\varepsilon\mu_*(\frac1\varepsilon)$ (cf. (2.5)) smaller than the typical scaling of large-scale averages of the commutator $J_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes g)$ itself as given in (2.3). In view of the generic non-degeneracy result in [5], this can be rewritten as a relative error estimate in form of (1.9). This property summarizes the pathwise structure of fluctuations and is the key part of our theory.

Theorem 1 (Pathwise structure of fluctuations)**.** *Let $d\ge1$, assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta>0$, let π_* be defined in (2.2), and define $\mu_*:\mathbb{R}_+\to\mathbb{R}_+$ by*

$$\mu_*(r) := \begin{cases} 1 & : \ \beta>2,\ d>2,\\ \log^{\frac12}(2+r) & : \ \beta>2,\ d=2,\\ \sqrt{1+r} & : \ \beta>1,\ d=1,\\ \log^{\frac12}(2+r) & : \ \beta=2,\ d>2,\\ \log(2+r) & : \ \beta=2,\ d=2,\\ \sqrt{1+r}\,\log^{\frac12}(2+r) & : \ \beta=1,\ d=1,\\ (1+r)^{1-\frac\beta2} & : \ \beta<2,\ d\ge2,\ \text{or}\ \beta<1,\ d=1. \end{cases} \quad (2.5)$$

Set $\mu_*(z):=\mu_*(|z|)$, recall the notation $w_1(z):=1+|z|$, and consider the rescaled error functional $\widehat{E}^\varepsilon:=\pi_*(\frac1\varepsilon)^{\frac12}E^\varepsilon$. For all $0<\varepsilon\le1$, $f,g\in C_c^\infty(\mathbb{R}^d)^d$, $0<p-1\ll1$, and $\alpha>\frac{d-(\beta\wedge d)}{2}+d\frac{p-1}{4p}$, we have

$$|\widehat{E}^\varepsilon(f,g)| \le \varepsilon\mu_*(\tfrac1\varepsilon)\,\mathcal{C}^{\varepsilon,f,g}_{\alpha,p}\Big(\|\mu_*\nabla f\|_{\mathrm{L}^4}\|w_1^\alpha g\|_{\mathrm{L}^{4p}}+\|\mu_*\nabla g\|_{\mathrm{L}^4}\|w_1^\alpha f\|_{\mathrm{L}^{4p}}$$
$$+\mathbb{1}_{\beta\le d}\big(\|\mu_*\nabla f\|_{\mathrm{L}^2}\|g\|_{\mathrm{L}^2\cap\mathrm{L}^{2p}}+\|\mu_*\nabla g\|_{\mathrm{L}^2}\|f\|_{\mathrm{L}^2\cap\mathrm{L}^{2p}}\big)\Big), \quad (2.6)$$

where $\mathcal{C}^{\varepsilon,f,g}_{\alpha,p}$ is a random variable with stretched exponential moments: there exists $\gamma_2\simeq1$ such that

$$\sup_{0<\varepsilon<1}\mathbb{E}\left[\exp\Big(\frac{1}{C_{\alpha,p}}(\mathcal{C}^{\varepsilon,f,g}_{\alpha,p})^{\gamma_2}\Big)\right] \le 2$$

for some (deterministic) constant $C_{\alpha,p}\simeq_{\alpha,p}1$. ◊

Remark 2.1.

- The exponents γ_1 and γ_2 in the above results can be made explicit; we do not pursue this direction since the values obtained in the proofs are not expected to be optimal.
- The ε-scaling in the above results is believed to be optimal. The rescaling in the definition of $\widehat{J}_0^\varepsilon$ and $\widehat{E}^\varepsilon$ is natural since it precisely coincides with the scaling of large-scale averages of the coefficient field $\boldsymbol{a}$ itself. For some non-generic examples, the bound (2.3) may however overestimate the variance. In dimension $d=1$, one may indeed construct explicit Gaussian coefficient fields $\boldsymbol{a}$ such that fluctuations of the

homogenization commutator J_0^ε are of smaller order than what (2.3) predicts [18, 13, 16], in which case the suitable rescaling of J_0^ε has a non-Gaussian limit. In such situations, the pathwise property (2.6) (or its higher-order pathwise version as in [9]) might still provide relevant information. General necessary and sufficient conditions for the sharpness of (2.3) are provided in [5].

- The proofs of the above results are robust enough to cover the general setting of multiscale functional inequalities introduced in [6, 7]. In the case of functional inequalities with oscillation, we may indeed use Cauchy-Schwarz' inequality and an energy estimate to replace the perturbed functions $\tilde\phi$ and $\nabla\tilde u$ appearing in the representation formula (3.3) below by their unperturbed versions ϕ and ∇u. This allows to conclude whenever the weight has a superalgebraic decay (see indeed [11, proof of Theorem 4]). If one is only interested in Gaussian coefficient fields, one may replace the use of functional inequalities by a direct use of the Brascamp-Lieb inequality in terms of Malliavin calculus, which allows to shorten some of the proofs (and improve the norms of the test functions F, f, g), cf. [9]. $\Diamond$

In view of the identities (1.10) and (1.12), the above pathwise result implies that the large-scale fluctuations of I_0^ε, I_1^ε, I_2^ε, J_1^ε, and J_2^ε are driven by the fluctuations of J_0^ε in a pathwise sense (see [8, Corollary 2.4] for details).

Corollary 1 ([8])**.** *Let $d\ge 2$, assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta>0$, let π_* and μ_* be defined by (2.2) and (2.5), let $\bar{\mathcal{P}}_H$, $\bar{\mathcal{P}}_H^*$, and $\bar{\mathcal{P}}_L^*$ be as in (1.11), and recall the rescaled functionals*

$$\widehat{I}_i^\varepsilon := \pi_*(\tfrac1\varepsilon)^{\frac12} I_i^\varepsilon, \qquad \widehat{J}_i^\varepsilon := \pi_*(\tfrac1\varepsilon)^{\frac12} J_i^\varepsilon, \qquad i=0,1,2.$$

For all $\varepsilon>0$ and $f,g\in C_c^\infty(\mathbb{R}^d)^d$, we have for all $0<p-1\ll 1$ and $\alpha>\frac{d-(\beta\wedge d)}{2}+d\frac{p-1}{4p}$,

$$\begin{aligned}|\widehat{I}_0^\varepsilon(f,g)-\widehat{J}_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes g)|+|\widehat{I}_1^\varepsilon(f,g)-\widehat{J}_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes \bar{\mathcal{P}}_H^* g)|+|\widehat{I}_2^\varepsilon(f,g)+\widehat{J}_0^\varepsilon(\bar{\mathcal{P}}_H f\otimes \bar{\mathcal{P}}_L^* g)|\\ \le \varepsilon\mu_*(\tfrac1\varepsilon)\,\mathcal{C}_{\alpha,p}^{\varepsilon,f,g}\Big(\|\mu_*\nabla f\|_{\mathrm{L}^4}\|w_1^\alpha g\|_{\mathrm{L}^{4p}}+\|\mu_*\nabla g\|_{\mathrm{L}^4}\|w_1^\alpha f\|_{\mathrm{L}^{4p}}\\ +\mathbb{1}_{\beta\le d}\big(\|\mu_*\nabla f\|_{\mathrm{L}^2}\|g\|_{\mathrm{L}^2\cap\mathrm{L}^{2p}}+\|\mu_*\nabla g\|_{\mathrm{L}^2}\|f\|_{\mathrm{L}^2\cap\mathrm{L}^{2p}}\big)\Big),\end{aligned}$$

where $\mathcal{C}_{\alpha,p}^{\varepsilon,f,g}$ is a random variable with stretched exponential moments independent of ε as in the statement of Theorem 1. In addition, for all $\varepsilon>0$ and $F\in C_c^\infty(\mathbb{R}^d)^{d\times d}$, we have almost surely

$$J_1^\varepsilon(F)=-J_0^\varepsilon(\bar{\mathcal{P}}_H^*F), \qquad J_2^\varepsilon(F)=J_0^\varepsilon(\bar{\mathcal{P}}_L^*F),$$

where in particular we may give an almost sure meaning to $J_0^\varepsilon(\bar{\mathcal{P}}_H^*F)$ and $J_0^\varepsilon(\bar{\mathcal{P}}_L^*F)$ for all $F\in C_c^\infty(\mathbb{R}^d)^{d\times d}$, even when $\bar{\mathcal{P}}_H^*F$ and $\bar{\mathcal{P}}_L^*F$ do not have integrable decay. $\Diamond$

2.2. **Structure of the proof.** We describe the string of arguments that leads to Proposition 1 and Theorem 1. Next to the corrector ϕ, we first need to recall the notion of flux corrector σ, which was recently introduced in the stochastic setting in [10, Lemma 1] and allows to put the equation for the two-scale homogenization error in divergence form (cf. (3.19) below). The extended corrector (ϕ,σ) is only defined up to an additive constant, and we henceforth choose the anchoring $\fint_B(\phi,\sigma)=0$ on the centered unit ball B.

Lemma 2.2 ([10])**.** *Let the coefficient field $\boldsymbol{a}$ be stationary and ergodic. Then there exist two random tensor fields $(\phi_i)_{1\le i\le d}$ and $(\sigma_{ijk})_{1\le i,j,k\le d}$ with the following properties: The* *gradient fields* $\nabla\phi_i$ *and* $\nabla\sigma_{ijk}$ *are stationary*[2] *and have finite second moments and vanishing expectations:*

$$\mathbb{E}\left[|\nabla\phi_i|^2\right] \le \frac{1}{\lambda^2}, \qquad \sum_{j,k=1}^d \mathbb{E}\left[|\nabla\sigma_{ijk}|^2\right] \le 4d\Big(\frac{1}{\lambda^2}+1\Big), \qquad \mathbb{E}\left[\nabla\phi_i\right] = \mathbb{E}\left[\nabla\sigma_{ijk}\right] = 0. \tag{2.7}$$

Moreover, for all i*, the field* $\sigma_i := (\sigma_{ijk})_{1\le j,k\le d}$ *is skew-symmetric, that is,*

$$\sigma_{ijk} = -\sigma_{ikj}. \tag{2.8}$$

Finally, the following equations are satisfied a.s. in the distributional sense on $\mathbb{R}^d$*,*

$$\begin{aligned} -\nabla\cdot\boldsymbol{a}(\nabla\phi_i+e_i) &= 0, && (2.9)\\ \nabla\cdot\sigma_i &= q_i - \mathbb{E}\left[q_i\right], && (2.10)\\ -\triangle\sigma_{ijk} &= \partial_j q_{ik} - \partial_k q_{ij}, \end{aligned}$$

where $q_i = (q_{ij})_{1\le j\le d}$ *is given by* $q_i := \boldsymbol{a}(\nabla\phi_i+e_i)$*, and where the (distributional) divergence of a tensor field is defined as* $(\nabla\cdot\sigma_i)_j := \sum_{k=1}^d \nabla_k\sigma_{ijk}$*.* ◊

The proofs of Proposition 1 and Theorem 1 are based on the combination of three main ingredients:

- a sensitivity calculus combined with functional inequalities for Gaussian ensembles [6, 7];
- the bounds on correctors proved in [11];
- a duality argument combined with the large-scale (weighted) Calderón-Zygmund estimates of [10].

In the case when the coefficients satisfy a finite range of dependence assumption rather than a functional inequality, we do not have a convenient sensitivity calculus at our disposal, and this first ingredient can be replaced by a semi-group approach that provides a convenient disintegration of scales; this is postponed to a forthcoming work.

The sensitivity calculus measures the influence of changes of the coefficient field $\boldsymbol{a}$ on random variables $X = X(\boldsymbol{a})$ via the functional (or Malliavin-type) derivative $\partial^{\mathrm{fct}}X(x) = \frac{\partial X}{\partial \boldsymbol{a}}(\boldsymbol{a},x)$, that is, the $\mathrm{L}^2(\mathbb{R}^d)^{d\times d}$-gradient of X with respect to $\boldsymbol{a}$. We recall that this functional derivative is characterized as follows, for any compactly supported perturbation $\boldsymbol{b}\in\mathrm{L}^\infty(\mathbb{R}^d)^{d\times d}$,

$$\int_{\mathbb{R}^d} \partial^{\mathrm{fct}}X(\boldsymbol{a},x) : \boldsymbol{b}(x)\,dx \;:=\; \lim_{t\downarrow 0}\frac{1}{t}\big(X(\boldsymbol{a}+t\boldsymbol{b}) - X(\boldsymbol{a})\big). \tag{2.11}$$

This quantity measures the sensitivity of the random variable $X = X(\boldsymbol{a})$ with respect to changes in the coefficient field. This sensitivity calculus is a building block to control the variance and the entropy of X via functional inequalities in the probability space [6]. A crucial role is played by the parameter $\beta > 0$ that characterizes the decay of the covariance function of $\boldsymbol{a}$, and we define as follows a weighted norm $|\!|\!|\cdot|\!|\!|_\beta^2$ on random fields G, depending on $\beta > 0$,

$$|\!|\!|G|\!|\!|_\beta^2 \;:=\; \int_1^\infty \|G\|_\ell^2\,\ell^{-\beta-1}\,d\ell, \tag{2.12}$$

[2]That is, $\nabla\phi_i(\boldsymbol{a};\cdot+z) = \nabla\phi_i(\boldsymbol{a}(\cdot+z);\cdot)$ and $\nabla\sigma_{ijk}(\boldsymbol{a};\cdot+z) = \nabla\sigma_{ijk}(\boldsymbol{a}(\cdot+z);\cdot)$ a.e. in $\mathbb{R}^d$, for all shift vectors $z\in\mathbb{R}^d$.

where for all $\ell \ge 1$

$$\|G\|_\ell^2 \;:=\; \ell^{-d}\int_{\mathbb{R}^d}\Big(\int_{B_\ell(z)}|G|\Big)^2 dz. \tag{2.13}$$

As shown in [6, Proposition 2.4], in the integrable case $\beta > d$, we can drop the integral over ℓ, in which case

$$\|\!|G|\!\|_\beta^2 \;\simeq\; \|\!|G|\!\|^2 \;:= \|G\|_1^2. \tag{2.14}$$

In these terms, we may formulate the following multiscale logarithmic Sobolev inequality for the Gaussian coefficient field $\boldsymbol{a}$. In view of (2.14), for $\beta > d$, this reduces to the standard logarithmic Sobolev inequality (LSI). The proof is based on a corresponding Brascamp-Lieb inequality (cf. [7, Theorem 3.1]).

Lemma 2.3 ([7]). *Assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta > 0$. Then for all random variables $X = X(\boldsymbol{a})$,*

$$\operatorname{Ent}\big[X^2\big] \;:=\; \mathbb{E}\left[X^2\log X^2\right] - \mathbb{E}\left[X^2\right]\mathbb{E}\left[\log X^2\right] \;\lesssim\; \mathbb{E}\Big[\|\!|\partial^{\mathrm{fct}}X|\!\|_\beta^2\Big]. \qquad \Diamond$$

Our general strategy for the proof of Proposition 1 and Theorem 1 consists in estimating the weighted norm (2.12) of the functional derivatives of $J_0^\varepsilon(F)$ and of $E^\varepsilon(f,g)$. The following lemma provides a useful representation formula for these functional derivatives, in particular relying on the specific structure of homogenization commutators. This is a continuum version of [8, Lemma 3.2]. By scaling, it is enough to consider $\varepsilon = 1$, and we write for simplicity $J_0 := J_0^1$ and $E := E^1$.

Lemma 2.4 (Representation formulas). *Let the coefficient field $\boldsymbol{a}$ be Gaussian with parameter $\beta > 0$. For all $f \in C_c^\infty(\mathbb{R}^d)^d$, let $\nabla u := \nabla u_1$ denote the solution of (1.3) (with $\varepsilon = 1$), let $\nabla\bar u$ denote the solution of (1.4), and define the two-scale expansion error $w_f := u - (1+\phi_i\nabla_i)\bar u$. Then, for all $F \in C_c^\infty(\mathbb{R}^d)^{d\times d}$,*

$$\partial^{\mathrm{fct}}J_0(F) = (F_{ij}\,e_j + \nabla S_i)\otimes(\nabla\phi_i + e_i), \tag{2.15}$$

and for all $g \in C_c^\infty(\mathbb{R}^d)^d$,

$$\begin{aligned}\partial^{\mathrm{fct}}E(f,g) = g_j\,(\nabla\phi_j^* + e_j)\otimes(\nabla w_f + \phi_i\nabla\nabla_i\bar u) + (\phi_j^*\,\nabla g_j + \nabla r)\otimes\nabla\bar u \\ - \big(\phi_j^*\,\nabla(g_j\nabla_i\bar u) + \nabla R_i\big)\otimes(\nabla\phi_i + e_i),\end{aligned} \tag{2.16}$$

where the auxiliary fields $\nabla S = (\nabla S_i)_{i=1}^d$, ∇r, and $\nabla R = (\nabla R_i)_{i=1}^d$ are the gradient solutions in $\mathrm{L}^2(\mathbb{R}^d)^d$ of

$$-\nabla\cdot\boldsymbol{a}^*\nabla S_i \;=\; \nabla\cdot\big(F_{ij}(\boldsymbol{a}^* - \bar{\boldsymbol{a}}^*)e_j\big), \tag{2.17}$$

$$-\nabla\cdot\boldsymbol{a}^*\nabla r \;=\; \nabla\cdot\big((\phi_j^*\boldsymbol{a}^* - \sigma_j^*)\nabla g_j\big), \tag{2.18}$$

$$-\nabla\cdot\boldsymbol{a}^*\nabla R_i \;=\; \nabla\cdot\big((\phi_j^*\boldsymbol{a}^* - \sigma_j^*)\nabla(g_j\nabla_i\bar u)\big), \tag{2.19}$$

and $\boldsymbol{a}^*$ denotes the pointwise transpose coefficient field of $\boldsymbol{a}$, and (ϕ^*, σ^*) denotes the corresponding extended corrector (recall that $\overline{\boldsymbol{a}^*} = \bar{\boldsymbol{a}}^*$). $\Diamond$

Before we turn to the (technical) estimates of $\|\!|\partial^{\mathrm{fct}}J_0(F)|\!\|_\beta$ and $\|\!|\partial^{\mathrm{fct}}E(f,g)|\!\|_\beta$, let us give an informal discussion of the scalings of the terms appearing in (2.15) and (2.16). To keep this discussion short, assume that $\nabla\phi, \nabla\sigma$ are bounded (which only holds after taking stochastic moments), that $|\phi(x)| + |\sigma(x)| \lesssim \mu_*(|x|)$ (which again only holds after taking stochastic moments), and that the Helmholtz projections associated with $-\nabla\cdot\boldsymbol{a}^*\nabla$ (and used to define S_i, r, and R_i via (2.17)–(2.19)) enjoy perfectly local bounds in the sense that

$$-\nabla\cdot\tilde{\boldsymbol{a}}\nabla z=\nabla\cdot Z \qquad\Longrightarrow\qquad |\nabla z(x)|\lesssim|Z(x)|\quad\text{for all }x\in\mathbb{R}^d,$$

with $\tilde{\boldsymbol{a}}=\boldsymbol{a}$ or $\bar{\boldsymbol{a}}$ (which even in the homogeneous case $\tilde{\boldsymbol{a}}=\bar{\boldsymbol{a}}$ would only hold after taking suitable Lebesgue norms in view of the Calderón-Zygmund theory). For $J_0(F)$, equation (2.17) would then yield the pointwise bound $|\partial^{\text{fct}}J_0(F)|\lesssim|F|$, hence

$$\|\!|\partial^{\text{fct}}J_0(F)\|\!|_\beta^2 \;\lesssim\; \int_1^\infty \ell^{-d}\int_{\mathbb{R}^d}\Big(\int_{B_\ell(z)}|F|\Big)^2dz\;\ell^{-\beta-1}\,d\ell.$$

To estimate the right-hand side, assume that F is compactly supported in B_R for some $R>0$, so that

$$\ell^{-d}\int_{\mathbb{R}^d}\Big(\int_{B_\ell(z)}|F|\Big)^2dz\lesssim(\ell^d\mathbb{1}_{\ell\le R}+R^d\mathbb{1}_{\ell>R})\Big(\int_{\mathbb{R}^d}|F|^2\Big),$$

which after integration yields, in view of (2.2),

$$\|\!|\partial^{\text{fct}}J_0(F)\|\!|_\beta^2\lesssim\Big(\int_1^R\ell^{d-\beta-1}d\ell+R^d\int_R^\infty\ell^{-\beta-1}d\ell\Big)\Big(\int_{\mathbb{R}^d}|F|^2\Big)\lesssim R^d\pi_*^{-1}(R)\int_{\mathbb{R}^d}|F|^2.$$

Replacing F by $\varepsilon^dF(\frac{\cdot}{\varepsilon})$, hence replacing R by $\frac1\varepsilon R$, we conclude by LSI,

$$\mathrm{Var}\Big[\pi_*(\tfrac1\varepsilon)^{\frac12}J_0^\varepsilon(F)\Big]\;\lesssim_R\;\|F\|^2_{\mathrm{L}^2(\mathbb{R}^d)},$$

as claimed in Proposition 1 (with a slightly stronger norm of the test function F).

We now turn to the two-scale expansion error, for which (2.18) and (2.19) would yield the following pointwise bound, under the simplifying assumptions,

$$|\partial^{\text{fct}}E(f,g)|\;\lesssim\;|g|(|\nabla w_f|+\mu_*|\nabla^2\bar u|)+\mu_*|\nabla g||\nabla\bar u|.$$

Let us further reformulate the right-hand side. On the one hand, since $-\nabla\cdot\bar{\boldsymbol{a}}\nabla\bar u=\nabla\cdot f$, the simplifying assumptions yield the pointwise bounds $|\nabla\bar u|\lesssim|f|$ and $|\nabla^2\bar u|\lesssim|\nabla f|$. On the other hand, the function w_f satisfies the equation $-\nabla\cdot\boldsymbol{a}\nabla w_f=\nabla\cdot((\boldsymbol{a}\phi-\sigma)\nabla^2\bar u)$, (cf. [11, Remark 3]) so that our simplifying assumptions yield this time $|\nabla w_f|\lesssim\mu_*|\nabla^2\bar u|\lesssim\mu_*|\nabla f|$. This leads to the bound

$$|\partial^{\text{fct}}E(f,g)|\;\lesssim\;\mu_*(|g||\nabla f|+|\nabla g||f|).$$

As above, after rescaling, and using $\mu_*(|\frac{x}{\varepsilon}|)\lesssim\mu_*(|x|)\mu_*(\frac1\varepsilon)$, we conclude by LSI,

$$\mathrm{Var}\Big[\pi_*(\tfrac1\varepsilon)^{\frac12}E^\varepsilon(f,g)\Big]\;\lesssim\;\varepsilon^2\mu_*^2(\tfrac1\varepsilon)\big(\|\mu_*\nabla g\|^2_{\mathrm{L}^4(\mathbb{R}^d)}\|f\|^2_{\mathrm{L}^4(\mathbb{R}^d)}+\|\mu_*\nabla f\|^2_{\mathrm{L}^4(\mathbb{R}^d)}\|g\|^2_{\mathrm{L}^4(\mathbb{R}^d)}\big),$$

as claimed in Theorem 1 (with slightly stronger norms of the test functions). Note that the additional factor ε^2 comes from the gradients ∇f and ∇g in the bound of the functional derivative (which indeed both yield an ε factor by rescaling).

The rigorous proof of Proposition 1 and Theorem 1 amounts to taking care of the fact that the above simplifying assumptions only hold in weaker forms. More precisely:

- Bounds on the correctors only hold for stretched exponential moments and not pointwise (cf. Lemma 2.8 below), so that the bounds on $|\partial^{\text{fct}}J_0(F)$ and $|\partial^{\text{fct}}E(f,g)|$ will not hold pointwise but for stretched exponential moments.

- More importantly, the Helmholtz projection never enjoys pointwise bounds, which must be weakened in two ways. First, for the homogeneous operator $-\nabla\cdot\bar{\boldsymbol{a}}\nabla$, we must resort to the boundedness of the Helmholtz projection in L^p spaces for $1<p<\infty$ (Calderón-Zygmund estimates). Second, for the heterogeneous operator $-\nabla\cdot\boldsymbol{a}\nabla$, regularity theory can only hold on large scales [2, 10], so that Calderón-Zygmund estimates must be locally averaged at some random scale r_* (cf. Lemma 2.7); we will have to get rid of this random local average at some point using Hölder's inequality and a small weight (see e.g. the second right-hand side factor in (2.20)). Finally, when the corrector does not have uniformly bounded moments (that is, when it grows at infinity), we further need to resort to weighted Calderón-Zygmund estimates (cf. Lemma 2.7(c)); see e.g. the weight μ_* in the third right-hand side factor in (2.22).

Estimates on $\|\!|\partial^{\mathrm{fct}}J_0(F)|\!\|_\beta$ and $\|\!|\partial^{\mathrm{fct}}E(f,g)|\!\|_\beta$ are obtained in the following two technical propositions. As above, we prove the estimate for $\varepsilon=1$ and then argue by scaling. Since we need some flexibility in the weights, some estimates involve a parameter $R\geq1$. This parameter is arbitrary and should be thought of as being $R=\frac{1}{\varepsilon}$ for the proof of the main results (similarly as in the above informal discussion). Henceforth we write $\int$ instead of $\int_{\mathbb{R}^d}$ for simplicity.

Proposition 2.5 (Main estimates I)**.** *Let the coefficient field $\boldsymbol{a}$ be Gaussian with parameter $\beta>0$. Let π_* and μ_* be defined by* (2.2) *and* (2.5)*, respectively, and let the random field r_* be the minimal radius of Lemma 2.7 below. For $F\in C^\infty_c(\mathbb{R}^d)$ we denote by $[F]_2(x):=(\fint_{B(x)}|F|^2)^{\frac12}$ the moving local quadratic average, and for $R\geq1$ we set $w_R(x):=\frac{|x|}{R}+1$. Then the following hold:*

(i) If $\beta>d$, we have for all $R\geq1$, $0<\alpha-d\ll1$, and $0<p-1\ll_\alpha1$,

$$\|\!|\partial^{\mathrm{fct}}J_0(F)|\!\|^2\lesssim_{\alpha,p}r_*(0)^{\alpha\frac{p-1}{p}}\Big(\int r_*^{d\frac{p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{p}}\Big(\int w_R^{\alpha(p-1)}|F|^{2p}\Big)^{\frac1p}. \tag{2.20}$$

(ii) If $\beta\leq d$, we have for all $R\geq1$, $0<\gamma<\beta$, $0<\alpha-d\ll1$, and $0<p-1\ll_{\gamma,\alpha}1$,

$$\begin{aligned}\|\!|\partial^{\mathrm{fct}}J_0(F)|\!\|_\beta^2\lesssim_{\alpha,p}{}&R^d\pi_*(R)^{-1}[\mathrm{RHS}(2.20)]+R^{2d-\beta-\frac{2d}{p}}r_*(0)^{2d\frac{p-1}{p}}\Big(\int r_*^{d(1-\frac{p}{2})}[F]_2^p\Big)^{\frac2p}\\&+R^{d-\beta}r_*(0)^{d-\gamma+\alpha\frac{p-1}{p}}\Big(\int r_*^{\gamma\frac{p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{p}}\Big(\int w_R^{(d-\gamma)p+\alpha(p-1)}|F|^{2p}\Big)^{\frac1p},\end{aligned} \tag{2.21}$$

where we use the short-hand notation [RHS(2.20)] *for the right-hand side of* (2.20)*.* ◊

Proposition 2.6 (Main estimates II)**.** *Let the coefficient field $\boldsymbol{a}$ be Gaussian with parameter $\beta>0$. Let π_* and μ_* be defined by* (2.2) *and* (2.5)*, respectively, and let the random fields r_* and $\mathcal{C}$ be defined in Lemmas 2.7 and 2.8 below. For $F\in C^\infty_c(\mathbb{R}^d)^d$ we denote by $[F]_\infty(x):=\sup_{B(x)}|F|$ the moving local supremum, and we recall the notation $w_R(x)=\frac{|x|}{R}+1$. Then the following hold:*

(i) If $\beta>d$, we have for all $R\geq1$, $0<\alpha-d\ll1$, and $0<p-1\ll_\alpha1$,

$$\|\!|\partial^{\mathrm{fct}}E(f,g)|\!\|^2\lesssim_{\alpha,p}r_*(0)^{\alpha\frac{p-1}{2p}}\Big(\int r_*^{2d\frac{2p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{2p}}$$

$$\times\bigg(\Big(\Big(\int \mathcal{C}^4\mu_*^4[\nabla^2\bar u]_\infty^4\Big)^{\frac12}\Big(\int w_R^{\alpha(p-1)}[g]_\infty^{4p}\Big)^{\frac1{2p}} + \Big(\int \mathcal{C}^4\mu_*^4[\nabla g]_\infty^4\Big)^{\frac12}\Big(\int w_R^{\alpha(p-1)}(|f|+[\nabla\bar u]_\infty)^{4p}\Big)^{\frac1{2p}}\bigg). \tag{2.22}$$

(ii) If $\beta\le d$, we have for all $R\ge1$, $0\le\gamma<\beta$, $0<\alpha-d\ll1$, and $0<p-1\ll_\alpha 1$,

$$\begin{aligned}\|\partial^{\mathrm{fct}}E(f,g)\|_\beta^2 \lesssim_{\alpha,p}\ & R^d\pi_*(R)^{-1}[\mathrm{RHS}(2.22)]\\ &+R^{-\beta}\Big(\int \mathcal{C}^2\mu_*^2[\nabla g]_\infty^2\Big)\bigg(\Big(\int |f|^2\Big)+(Rr_*(0))^{d\frac{p-1}{p}}\Big(\int[\nabla\bar u]_\infty^{2p}\Big)^{\frac1p}\bigg)\\ &+R^{-\beta}\Big(\int \mathcal{C}^2\mu_*^2[\nabla^2\bar u]_\infty^2\Big)\bigg(\Big(\int r_*^d[g]_\infty^2\Big)+(Rr_*(0))^{d\frac{p-1}{p}}\Big(\int[g]_\infty^{2p}\Big)^{\frac1p}\bigg)\\ &+R^{d-\beta}r_*(0)^{d-\gamma+\alpha\frac{p-1}{2p}}\Big(\int r_*^{\gamma\frac{2p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{2p}}\\ &\quad\times\bigg(\Big(\Big(\int \mathcal{C}^4\mu_*^4[\nabla g]_\infty^4\Big)^{\frac12}\Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)}(|f|+[\nabla\bar u]_\infty)^{4p}\Big)^{\frac1{2p}}\\ &\qquad+\Big(\int \mathcal{C}^4\mu_*^4[\nabla^2\bar u]_\infty^4\Big)^{\frac12}\Big(\int r_*^{2pd}w_R^{2p(d-\gamma)+\alpha(p-1)}[g]_\infty^{4p}\Big)^{\frac1{2p}}\bigg),\end{aligned} \tag{2.23}$$

where we use the short-hand notation [RHS(2.22)] *for the right-hand side of* (2.22). ♢

The proofs of Propositions 2.5 and 2.6 rely on two further ingredients: large-scale weighted Calderón-Zygmund estimates and moment bounds on the extended corrector (ϕ,σ) (which are at the origin of the scaling μ_* in the estimates). We start by recalling the former, which follows from [10, Theorem 1, Corollary 4, and Corollary 5] (see also [2, Section 7]).

Lemma 2.7 ([10])**.** *Assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta>0$, and let π_* be as in* (2.2)*. There exists a stationary, $\frac18$-Lipschitz continuous random field $r_*\ge1$ (the so-called* minimal radius*), satisfying for some (deterministic) constant $C\simeq1$,*

$$\mathbb{E}\Big[\exp\Big(\frac1C\pi_*(r_*)\Big)\Big]\le 2, \tag{2.24}$$

such that the following properties hold a.s.,

(a) Mean-value property:
For any $\boldsymbol{a}$-harmonic function u in B_R (that is, $-\nabla\cdot\boldsymbol{a}\nabla u=0$ in B_R), we have for all radii $r_*(0)\le r\le R$,

$$\fint_{B_r}|\nabla u|^2\lesssim\fint_{B_R}|\nabla u|^2. \tag{2.25}$$

Applied to the extended corrector of Lemma 2.2, this yields for all $\ell\ge1$ and $x\in\mathbb{R}^d$,

$$\int_{B_\ell(x)}|\nabla(\phi,\sigma)|^2\lesssim\big(\ell+r_*(x)\big)^d. \tag{2.26}$$

(b) Large-scale Calderón-Zygmund estimates:
Set $B_*(x):=B_{r_*(x)}(x)$, and more generally $B_{\ell*}(x):=B_{\ell+r_*(x)}(x)$. For all $1<p<\infty$,

for all (sufficiently fast) decaying scalar fields u and vector fields g related in $\mathbb{R}^d$ by

$$-\nabla\cdot a\nabla u=\nabla\cdot g,$$

we have

$$\int\Big(\fint_{B_*(x)}|\nabla u|^2\Big)^{\frac{p}{2}}dx\lesssim_p\int\Big(\fint_{B_*(x)}|g|^2\Big)^{\frac{p}{2}}dx. \tag{2.27}$$

(c) Large-scale weighted Calderón-Zygmund estimates:
For all $2\le p<\infty$, $0\le\gamma<d(p-1)$, and for all non-decreasing radial weights $w\ge 1$ satisfying

$$w(r)\le w(r')\le\Big(\frac{r'}{r}\Big)^{\gamma}w(r)\qquad\text{for all } 0\le r\le r',$$

we have for all u and g as in (b) above,

$$\int\Big(\fint_{B_*(x)}|\nabla u|^2\Big)^{\frac{p}{2}}w_*(x)\,dx\lesssim_{p,\gamma}\int\Big(\fint_{B_*(x)}|g|^2\Big)^{\frac{p}{2}}w_*(x)\,dx, \tag{2.28}$$

where $w_*(x):=w(|x|+r_*(0))$. ◊

Whereas the minimal radius r_* quantifies the sublinearity of the extended corrector at infinity [10], the precise growth of the latter is estimated as follows (cf. [11, Theorem 2]).

Lemma 2.8 ([10, 11])**.** *Assume that the coefficient field $\boldsymbol{a}$ is Gaussian with parameter $\beta>0$, let μ_* be as in (2.5), and let r_* be as in Lemma 2.7. Then the extended corrector (ϕ,σ) defined in Lemma 2.2 satisfies for all $x\in\mathbb{R}^d$,*

$$\Big(\fint_{B(x)}|(\phi,\sigma)|^2\Big)^{\frac12}\le\mathcal{C}(x)\mu_*(x), \tag{2.29}$$

where $\mathcal{C}\ge 1$ is a 1-Lipschitz continuous random field with stretched exponential moments: there exist $\gamma\simeq_\beta 1$ and $C_\gamma\simeq_\gamma 1$ such that

$$\mathbb{E}\Big[\exp\Big(\frac{1}{C_\gamma}\mathcal{C}^\gamma\Big)\Big]\le 2. \tag{2.30}$$

◊

In order to reformulate integrals in a form well-suited to apply (weighted) large-scale Calderón-Zygmund estimates, we display below an auxiliary lemma that takes advantage of the Lipschitz continuity of r_*.

Lemma 2.9. *Let $\|\cdot\|_\ell$ be defined in (2.13) and let r_* be as in Lemma 2.7. For all U,V and $\ell\ge 1$, we have*

$$\|UV\|_\ell^2\lesssim\int\Big(\int_{B_{2\ell*}(x)}|U|^2\Big)\Big(\fint_{B_*(x)}|V|^2\Big)dx, \tag{2.31}$$

and the refined estimate

$$\|UV\|_\ell^2\lesssim\int_{|x|\ge\ell}\Big(\int_{B_{2\ell*}(x)}|U|^2\Big)\Big(\fint_{B_*(x)}|V|^2\Big)dx+\Big(\int_{B_{7\ell*}(0)}\Big(\fint_{\bar B_*(x)}|U|^2\Big)^{\frac12}\Big(\fint_{\bar B_*(x)}|V|^2\Big)^{\frac12}dx\Big)^2, \tag{2.32}$$

where we recall $B_*(x)=B_{r_*(x)}(x)$, $B_{2\ell*}(x)=B_{2\ell+r_*(x)}(x)$, and $B_{7\ell*}(0)=B_{7\ell+r_*(0)}(0)$, and where we have set $\bar B_*(x):=B_{5r_*(x)}(x)$. ◊

3. Proof of the representation formulas and of the main estimates

3.1. Proof of Lemma 2.4: Representation formulas. We first introduce some notation. Let $\boldsymbol{a}$ and $\tilde{\boldsymbol{a}}$ be two (admissible) coefficient fields, and set $\delta\boldsymbol{a} := \tilde{\boldsymbol{a}} - \boldsymbol{a}$. For all random variables (or fields) $F = F(\boldsymbol{a})$, we set $\tilde{F} := F(\tilde{\boldsymbol{a}})$ and $\delta F := \tilde{F} - F$. We then denote by (ϕ, σ), (ϕ^*, σ^*), $(\tilde{\phi}, \tilde{\sigma})$, and $(\tilde{\phi}^*, \tilde{\sigma}^*)$ the extended correctors associated with $\boldsymbol{a}$, $\boldsymbol{a}^*$, $\tilde{\boldsymbol{a}}$, and $\tilde{\boldsymbol{a}}^*$, respectively.

Step 1. Proof of identity (2.15).
The definition (1.7) of Ξ_{ij} yields

$$\delta J_0(F) = \int F_{ij} \delta \Xi_{ij} = \int F_{ij}\, e_j \cdot (\boldsymbol{a} - \bar{\boldsymbol{a}}) \nabla \delta\phi_i + \int F_{ij}\, e_j \cdot \delta\boldsymbol{a} (\nabla \tilde{\phi}_i + e_i).$$

Using the definition (2.17) of the auxiliary field S as well as the corrector equation (2.9) for ϕ_i and $\tilde{\phi}_i$ in the form

$$\nabla \cdot \boldsymbol{a} \nabla \delta \phi_i = -\nabla \cdot \delta\boldsymbol{a} (\nabla \tilde{\phi}_i + e_i), \tag{3.1}$$

the first right-hand side term above can be rewritten as

$$\int F_{ij}\, e_j \cdot (\boldsymbol{a} - \bar{\boldsymbol{a}}) \nabla \delta \phi_i = -\int \nabla S_i \cdot \boldsymbol{a} \nabla \delta\phi_i = \int \nabla S_i \cdot \delta\boldsymbol{a} (\nabla \tilde{\phi}_i + e_i),$$

and the conclusion (2.15) follows from the definition (2.11) of the functional derivative.

Step 2. Proof of identity (2.16).
We start by giving a suitable representation formula for the functional derivative of the homogenization commutator of the solution $(\boldsymbol{a} - \bar{\boldsymbol{a}})\nabla u$. By the property (2.10) of the flux corrector σ_j^* in the form $(\boldsymbol{a}^* - \bar{\boldsymbol{a}}^*) e_j = -\boldsymbol{a}^* \nabla \phi_j^* + \nabla \cdot \sigma_j^*$, and by its skew-symmetry (2.8) in the form $(\nabla \cdot \sigma_j^*) \cdot \nabla \delta u = -\nabla \cdot (\sigma_j^* \nabla \delta u)$, we find

$$\begin{aligned} \delta\big(e_j \cdot (\boldsymbol{a} - \bar{\boldsymbol{a}}) \nabla u\big) &= e_j \cdot \delta \boldsymbol{a} \nabla \tilde{u} + e_j \cdot (\boldsymbol{a} - \bar{\boldsymbol{a}}) \nabla \delta u \\ &= e_j \cdot \delta\boldsymbol{a} \nabla \tilde{u} - \nabla \cdot (\sigma_j^* \nabla \delta u) - \nabla \phi_j^* \cdot \boldsymbol{a} \nabla \delta u. \end{aligned}$$

Equation (1.3) for u and $\tilde{u}$ in the form

$$-\nabla \cdot \boldsymbol{a} \nabla \delta u = \nabla \cdot \delta\boldsymbol{a} \nabla \tilde{u} \tag{3.2}$$

allows us to rewrite the last right-hand side term as

$$\begin{aligned} -\nabla \phi_j^* \cdot \boldsymbol{a} \nabla \delta u &= -\nabla \cdot (\phi_j^* \boldsymbol{a} \nabla \delta u) + \phi_j^* \nabla \cdot \boldsymbol{a} \nabla \delta u \\ &= -\nabla \cdot (\phi_j^* \boldsymbol{a} \nabla \delta u) - \phi_j^* \nabla \cdot \delta \boldsymbol{a} \nabla \tilde{u} \\ &= -\nabla \cdot (\phi_j^* \boldsymbol{a} \nabla \delta u) - \nabla \cdot (\phi_j^* \delta\boldsymbol{a} \nabla \tilde{u}) + \nabla \phi_j^* \cdot \delta \boldsymbol{a} \nabla \tilde{u}. \end{aligned}$$

Hence, we conclude

$$\delta\big(e_j \cdot (\boldsymbol{a} - \bar{\boldsymbol{a}}) \nabla u\big) = (\nabla \phi_j^* + e_j) \cdot \delta\boldsymbol{a} \nabla \tilde{u} - \nabla \cdot \big((\phi_j^* \boldsymbol{a} + \sigma_j^*) \nabla \delta u\big) - \nabla \cdot (\phi_j^* \delta \boldsymbol{a} \nabla \tilde{u}),$$

and similarly, replacing $x \mapsto u(x)$ by $x \mapsto \phi_i(x) + x_i$,

$$\delta \Xi_{ij} = (\nabla \phi_j^* + e_j) \cdot \delta\boldsymbol{a} (\nabla \tilde{\phi}_i + e_i) - \nabla \cdot \big((\phi_j^* \boldsymbol{a} + \sigma_j^*) \nabla \delta \phi_i\big) - \nabla \cdot \big(\phi_j^* \delta \boldsymbol{a} (\nabla \tilde{\phi}_i + e_i)\big).$$

Considering $\delta\big(e_j\cdot(\boldsymbol{a}-\bar{\boldsymbol{a}})\nabla u\big)-\nabla_i\bar u\,\delta\Xi_{ij}$ and multiplying by g_j, we are led to

$$\begin{aligned}\delta E(f,g)=\int g_j(\nabla\phi_j^*+e_j)\cdot\delta\boldsymbol{a}\big(\nabla\tilde u-(\nabla\tilde\phi_i+e_i)\nabla_i\bar u\big)\\+\int\phi_j^*\nabla g_j\cdot\delta\boldsymbol{a}\nabla\tilde u-\int\phi_j^*\nabla(g_j\nabla_i\bar u)\cdot\delta\boldsymbol{a}(\nabla\tilde\phi_i+e_i)\\+\int\nabla g_j\cdot(\phi_j^*\boldsymbol{a}+\sigma_j^*)\nabla\delta u-\int\nabla(g_j\nabla_i\bar u)\cdot(\phi_j^*\boldsymbol{a}+\sigma_j^*)\nabla\delta\phi_i.\end{aligned}$$

For the first right-hand side term we use the definition of w_f in form of $\nabla u-(\nabla\phi_i+e_i)\nabla_i\bar u=\nabla w_f+\phi_i\nabla\nabla_i\bar u$, whereas for the last two RHS terms we use the definitions (2.18) and (2.19) of the auxiliary fields r and R, combined with equations (3.1) and (3.2), so that

$$\begin{aligned}\delta E(f,g)=\int g_j(\nabla\phi_j^*+e_j)\cdot\delta\boldsymbol{a}(\nabla\tilde w_f+\tilde\phi_i\nabla\nabla_i\bar u)\\+\int\phi_j^*\nabla g_j\cdot\delta\boldsymbol{a}\nabla\tilde u-\int\phi_j^*\nabla(g_j\nabla_i\bar u)\cdot\delta\boldsymbol{a}(\nabla\tilde\phi_i+e_i)\\+\int\nabla r\cdot\delta\boldsymbol{a}\nabla\tilde u-\int\nabla R_i\cdot\delta\boldsymbol{a}(\nabla\tilde\phi_i+e_i),\end{aligned}\qquad(3.3)$$

and the conclusion (2.16) follows from the definition (2.11) of the functional derivative.

3.2. **Proof of Lemma 2.9.** We first recall the following equivalence for all non-negative functions h,

$$\int h\simeq\int\Big(\fint_{B_*(x)}h\Big),\qquad(3.4)$$

cf. [10, Proof of Corollary 4, Step 5]. Estimate (2.31) is a consequence of (3.4) in form of

$$\begin{aligned}\|UV\|_\ell^2=\ell^{-d}\int\Big(\int_{B_\ell(x)}|UV|\Big)^2dx\ &\le\ \int\Big(\int_{B_\ell(x)}|U|^2\Big)\Big(\fint_{B_\ell(x)}|V|^2\Big)\,dx\\&\le\ \int\bigg(\fint_{B_\ell(x)}|V(y)|^2\Big(\int_{B_{2\ell}(y)}|U|^2\Big)dy\bigg)\,dx\\&=\ \int|V(x)|^2\Big(\int_{B_{2\ell}(x)}|U|^2\Big)\,dx\\&\lesssim\ \int\Big(\int_{B_{2\ell*}(x)}|U|^2\Big)\Big(\fint_{B_*(x)}|V|^2\Big)\,dx.\end{aligned}$$

We now turn to the proof of (2.32). We distinguish the generic case $r_*(0)\le\ell$ from the non-generic case $r_*(0)>\ell$, and we start with the latter. By the $\frac18$-Lipschitz continuity of r_* and the assumption $r_*(0)>\ell\ge1$, we have

$$|x|\le\ell\quad\Longrightarrow\quad\frac78r_*(0)\le r_*(0)-\frac18\ell\le r_*(x)\le r_*(0)+\frac18\ell\le\frac98r_*(0)$$

and

$$|x|\le\frac{\ell}{2}\quad\Longrightarrow\quad\frac{15}{16}r_*(0)\le r_*(x)\le\frac{17}{16}r_*(0)\quad\Longrightarrow\quad B_{4r_*(0)}(0)\subset\bar B_*(x)$$

(where $\bar{B}_*(x) = B_{5r_*(x)}(x)$, since $\frac{15}{16}\times 5 - \frac{1}{2} = \frac{67}{16} > 4$), so that

$$\begin{aligned}
&\int_{|x|\le\ell}\Big(\int_{B_{2\ell*}(x)}|U|^2\Big)\Big(\fint_{B_*(x)}|V|^2\Big)\,dx\\
&\lesssim\ r_*(0)^{-d}\,\ell^d\Big(\int_{B_{(3+\frac{9}{8})r_*(0)}(0)}|U|^2\Big)\Big(\int_{B_{(1+\frac{9}{8})r_*(0)}(0)}|V|^2\Big)\\
&\le\ r_*(0)^{-d}\bigg(\int_{B_{\frac{\ell}{2}}(0)}\Big(\int_{\bar{B}_*(x)}|U|^2\Big)^{\frac12}\Big(\int_{\bar{B}_*(x)}|V|^2\Big)^{\frac12}\bigg)^2\\
&\lesssim\ \bigg(\int_{B_{\ell*}(0)}\Big(\fint_{\bar{B}_*(x)}|U|^2\Big)^{\frac12}\Big(\fint_{\bar{B}_*(x)}|V|^2\Big)^{\frac12}\bigg)^2.
\end{aligned}$$

Combined with (2.31), this yields the conclusion (2.32) for $r_*(0) > \ell$. We turn to the generic case $r_*(0) \le \ell$. We split the integral over $\mathbb{R}^d$ into the far-field contribution $|x| \ge 4\ell$ and the near-field contribution $|x| < 4\ell$. For the former, we proceed as above,

$$\begin{aligned}
\ell^{-d}\int_{|x|\ge 4\ell}\Big(\int_{B_\ell(x)}|UV|\Big)^2dx\ &\le\ \int_{|x|\ge 4\ell}\bigg(\fint_{B_\ell(x)}|V(y)|^2\Big(\int_{B_{2\ell}(y)}|U|^2\Big)\,dy\bigg)\,dx\\
&\le\ \int_{|x|\ge 3\ell}|V(x)|^2\Big(\int_{B_{2\ell}(x)}|U|^2\Big)\,dx\\
&\lesssim\ \int\bigg(\fint_{B_*(x)}|V(y)|^2\Big(\int_{B_{2\ell}(y)}|U|^2\Big)\mathbb{1}_{|y|\ge 3\ell}\,dy\bigg)dx,
\end{aligned}$$

where the last bound follows from (3.4). By the $\frac{1}{8}$-Lipschitz continuity of r_* and the assumption $r_*(0) \le \ell$, we infer that the condition $|x| < \ell$ implies $r_*(x) \le r_*(0) + \frac{1}{8}\ell < 2\ell$, hence $B_*(x) \subset B_{3\ell}(0)$. The above inequality then reduces to

$$\begin{aligned}
\ell^{-d}\int_{|x|\ge 4\ell}\Big(\int_{B_\ell(x)}|UV|\Big)^2dx\ &\lesssim\ \int_{|x|\ge\ell}\bigg(\fint_{B_*(x)}|V(y)|^2\Big(\int_{B_{2\ell}(y)}|U|^2\Big)\,dy\bigg)\,dx\\
&\le\ \int_{|x|\ge\ell}\Big(\int_{B_{2\ell*}(x)}|U|^2\Big)\Big(\fint_{B_*(x)}|V|^2\Big)\,dx. \qquad (3.5)
\end{aligned}$$

We turn to the near-field contribution $|x| < 4\ell$. We start with the trivial estimate

$$\ell^{-d}\int_{|x|<4\ell}\Big(\int_{B_\ell(x)}|UV|\Big)^2dx \lesssim \Big(\int_{B_{5\ell}(0)}|UV|\Big)^2,$$

and we use (3.4) in form of

$$\int_{B_{5\ell}(0)}|UV| \lesssim \int\Big(\fint_{B_*(x)}|UV|\mathbb{1}_{|y|<5\ell}\,dy\Big)\,dx.$$

By the $\frac{1}{8}$-Lipschitz continuity of r_* and the assumption $r_*(0) \le \ell$, we infer that the condition $|x| > 7\ell$ implies $B_*(x) \cap B_{5\ell}(0) = \varnothing$, hence

$$\int_{B_{5\ell}(0)}|UV| \lesssim \int_{|x|<7\ell}\Big(\fint_{B_*(x)}|UV|\Big)\,dx.$$

The Cauchy-Schwarz' inequality then leads to

$$\ell^{-d}\int_{|x|<4\ell}\Big(\int_{B_\ell(x)}|UV|\Big)^2dx \lesssim \bigg(\int_{|x|<7\ell}\Big(\fint_{B_*(x)}|U|^2\Big)^{\frac12}\Big(\fint_{B_*(x)}|V|^2\Big)^{\frac12}dx\bigg)^2,$$

and (2.32) follows in combination with (3.5) in the generic case $r_*(0) \le \ell$.

3.3. **Proof of Proposition 2.5: Main estimates I.** We split the proof into two main steps, first addressing the case of the standard LSI ($\beta > d$), and then turning to the general multiscale case ($\beta \le d$). Let $R \ge 1$ be arbitrary.

Step 1. Proof of (2.20) for standard LSI ($\beta > d$).
Since for standard LSI ($\beta > d$) we have $|||\partial^{\rm fct} J_0(F)|||_\beta \lesssim |||\partial^{\rm fct} J_0(F)|||_1$ (cf. (2.14)), it suffices to prove the following estimate: for all $\ell \ge 1$, $\alpha > d$, and $p > 1$ with $\alpha(p-1) < d(2p-1)$,

$$|||\partial^{\rm fct} J_0(F)|||_\ell^2 \lesssim_{\alpha,p} \ell^d r_*(0)^{\alpha\frac{p-1}{p}} \Big(\int r_*^{d\frac{p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \Big(\int w_R^{\alpha(p-1)} |F|^{2p}\Big)^{\frac1p}. \tag{3.6}$$

Starting point is formula (2.15), which, by (2.31) in Lemma 2.9 for $U = \nabla\phi + \mathrm{Id}$ and $V = |F| + |\nabla S|$, and by the mean-value property (2.26), implies

$$\begin{aligned} |||\partial^{\rm fct} J_0(F)|||_\ell^2 &\lesssim \int \big(\ell + r_*(x)\big)^d \Big(\fint_{B_*(x)} |F|^2 + |\nabla S|^2\Big)\,dx \\ &\lesssim \ell^d \int r_*(x)^d \Big(\fint_{B_*(x)} |F|^2 + |\nabla S|^2\Big)\,dx. \end{aligned} \tag{3.7}$$

This yields the conclusion (3.6) in combination with the following estimate applied for $s = q = 1$ and $\nabla v = \nabla S$ (cf. (2.17)): If $\nabla v \in \mathrm{L}^2(\mathbb{R}^d)^d$ is the gradient solution of $-\nabla\cdot \boldsymbol{a}\nabla v = \nabla\cdot h$ with $h \in C_c^\infty(\mathbb{R}^d)^d$, then for all $s \ge 0$, $q \ge 1$, $\alpha > d$, and $p > 1$ with $\alpha(p-1) < d(2pq-1)$,

$$\begin{aligned} &\int r_*(x)^{ds} \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2\Big)^q dx \\ &\qquad \lesssim_{\alpha,p,q,s} r_*(0)^{\alpha\frac{p-1}{p}} \Big(\int r_*^{ds\frac{p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \Big(\int w_R^{\alpha(p-1)} [h]_2^{2pq}\Big)^{\frac1p}, \end{aligned} \tag{3.8}$$

where for $s = 0$ we may even choose $p = 1$, in which case (3.8) is replaced by

$$\int \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2\Big)^q dx \lesssim_q \int [h]_2^{2q}, \tag{3.9}$$

which we state and prove here for future reference only.

Here comes the argument for (3.8). For all $\alpha > d$ and $p > 1$, we smuggle in the weight $w_{R*}(x) := w_R(|x| + r_*(0))$ to the power $\alpha\frac{p-1}{p}$, and use Hölder's inequality with exponents $(\frac{p}{p-1}, p)$, so that

$$\begin{aligned} \int r_*(x)^{ds} \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2\Big)^q dx &\lesssim_s \Big(\int r_*^{ds\frac{p}{p-1}} w_{R*}^{-\alpha}\Big)^{\frac{p-1}{p}} \\ &\quad \times \Big(\int w_{R*}(x)^{\alpha(p-1)} \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2\Big)^{pq} dx\Big)^{\frac1p}, \end{aligned}$$

where the first right-hand side sum is bounded by $\int r_*^{ds\frac{p}{p-1}} w_R^{-\alpha}$ since $w_R \le w_{R*}$. Provided that $\alpha(p-1) < d(2pq-1)$, we may apply the large-scale weighted Calderón-Zygmund estimate of Lemma 2.7(c) to the equation for ∇v, to the effect of

$$\int r_*(x)^{ds}\Big(\fint_{B_*(x)}|h|^2+|\nabla v|^2\Big)^q dx \lesssim_{\alpha,p,q,s} \Big(\int r_*^{ds\frac{p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \times\Big(\int w_{R*}(x)^{\alpha(p-1)}\Big(\fint_{B_*(x)}|h|^2\Big)^{pq}dx\Big)^{\frac1p}.$$

The claim (3.8) then follows from the bound $w_{R*}(x)\le r_*(0)\inf_{B_*(x)}w_R$, Jensen's inequality, and (3.4). For $s=0$, we appeal to the large-scale (not weighted) Calderón-Zygmund estimate of Lemma 2.7(b), which amounts to choosing $p=1$ in the above.

Step 2. Proof of (2.21) in the general multiscale case ($\beta\le d$).
The combination of (2.12) with (3.6) is not enough to prove (2.21), and we have to refine (3.6) in the regime $\ell\ge R$. By (2.32) in Lemma 2.9 and the mean-value property (2.26),

$$\|\partial^{\mathrm{fct}}J_0(F)\|_\ell^2 \lesssim \int_{|x|\ge\ell}(\ell+r_*(x))^d\Big(\fint_{B_*(x)}|F|^2+|\nabla S|^2\Big)dx + \bigg(\int_{B_{7\ell*}(0)}\Big(\fint_{\bar B_*(x)}|F|^2+|\nabla S|^2\Big)^{\frac12}\bigg)^2. \quad (3.10)$$

Let now $\nabla v\in \mathrm{L}^2(\mathbb{R}^d)^d$ be the gradient solution of $-\nabla\cdot\boldsymbol{a}\nabla v=\nabla\cdot h$ with $h\in C_c^\infty(\mathbb{R}^d)^d$. In the following two substeps we estimate the far-field and near-field contributions separately.

Substep 2.1. Far-field estimate: For all $s\ge0$, $q\ge1$, $0\le\gamma\le ds$, $\alpha>d$, and $p>1$ with $(ds-\gamma)p+\alpha(p-1)<d(2pq-1)$,

$$\int_{|x|\ge\ell}(\ell+r_*(x))^{ds}\Big(\fint_{B_*(x)}|h|^2+|\nabla v|^2\Big)^q dx \lesssim_{\gamma,\alpha,p,q,s} \ell^\gamma R^{ds-\gamma}r_*(0)^{ds-\gamma+\alpha\frac{p-1}{p}} \times\Big(\int r_*^{\gamma\frac{p}{p-1}}w_R^{-\alpha}\Big)^{\frac{p-1}{p}}\Big(\int w_R^{(ds-\gamma)p+\alpha(p-1)}[h]_2^{2pq}\Big)^{\frac1p}. \quad (3.11)$$

We smuggle in the weight w_{R*} to the power $\alpha\frac{p-1}{p}$ and the weight w_{1*} to the power $ds-\gamma$, and use Hölder's inequality with exponents $(\frac{p}{p-1},p)$, to the effect of

$$\int_{|x|\ge\ell}(\ell+r_*(x))^{ds}\Big(\fint_{B_*(x)}|h|^2+|\nabla v|^2\Big)^q dx \lesssim_s \Big(\int_{|x|\ge\ell}w_{R*}^{-\alpha}w_{1*}^{-(ds-\gamma)\frac{p}{p-1}}(\ell+r_*)^{ds\frac{p}{p-1}}\Big)^{\frac{p-1}{p}} \times\Big(\int w_{R*}^{\alpha(p-1)}w_{1*}^{(ds-\gamma)p}\Big(\fint_{B_*(x)}|h|^2+|\nabla v|^2\Big)^{pq}dx\Big)^{\frac1p}.$$

In the first right-hand side factor, we use the bound $w_{1*}(x) \gtrsim \ell + r_*(x)$ for $|x| \ge \ell$, while in the second right-hand side factor we use $w_{1*} \le R w_{R*}$. The above then leads to

$$\int_{|x|\ge\ell} \big(\ell + r_*(x)\big)^{ds} \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2 \Big)^q dx \lesssim_{\gamma,s} R^{ds-\gamma} \Big(\int_{|x|\ge\ell} w_{R*}^{-\alpha} (\ell + r_*)^{\gamma\frac{p}{p-1}} \Big)^{\frac{p-1}{p}} \times \Big(\int w_{R*}^{(ds-\gamma)p+\alpha(p-1)} \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2 \Big)^{pq} dx \Big)^{\frac{1}{p}}.$$

Provided $(ds-\gamma)p+\alpha(p-1) < d(2pq-1)$, we may apply the large-scale weighted Calderón-Zygmund estimate of Lemma 2.7(c) to the equation for ∇v. Using the bound $w_{R*}(x) \le r_*(0) \inf_{B_*(x)} w_R$, Jensen's inequality, and (3.4), the conclusion (3.11) follows.

Substep 2.2. Near-field estimate: For all $\ell \ge 1$ and $p > 1$,

$$\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} |h|^2 + |\nabla v|^2 \Big)^{\frac12} \lesssim_p \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int r_*^{d(1-\frac{p}{2})_+} [h]_2^p \Big)^{\frac1p}. \tag{3.12}$$

Indeed, by Hölder's inequality with exponents $(\frac{p}{p-1}, p)$,

$$\begin{aligned} &\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} |h|^2 + |\nabla v|^2 \Big)^{\frac12} \\ &\lesssim \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} |h|^2 + |\nabla v|^2 \Big)^{\frac{p}{2}} \Big)^{\frac1p} \\ &\lesssim \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int \Big(\fint_{B_*(x)} |h|^2 + |\nabla v|^2 \Big)^{\frac{p}{2}} \Big)^{\frac1p}, \end{aligned}$$

where we pass from $\bar B_*(x) = B_{6r_*(x)}(x)$ to $B_*(x) = B_{r_*(x)}(x)$. The claim (3.12) follows from the large-scale Calderón-Zygmund estimate of Lemma 2.7(b) applied to v.

Substep 2.3. Conclusion.
For all $\ell \ge 1$, $0 \le \gamma \le d$, $\alpha > d$, and $p > 1$ with $(d-\gamma)p + \alpha(p-1) < d(2p-1)$ and $p \le 2$, the combination of Substeps 2.1–2.2 with (3.10) yields the following improvement of (3.6),

$$\begin{aligned} \|\partial^{\rm fct} J_0(F)\|_\ell^2 \lesssim_{\gamma,\alpha,p} &\ \ell^{2d\frac{p-1}{p}} \Big(\int r_*^{d(1-\frac{p}{2})} [F]_2^p \Big)^{\frac2p} \\ &+ \ell^\gamma R^{d-\gamma} r_*(0)^{d-\gamma+\alpha\frac{p-1}{p}} \Big(\int r_*^{\gamma\frac{p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{p}} \Big(\int w_R^{(d-\gamma)p+\alpha(p-1)} |F|^{2p} \Big)^{\frac1p}. \end{aligned} \tag{3.13}$$

We appeal to (2.12), which we combine wtih (3.6) for $\ell \le R$ and with (3.13) for $\ell > R$. Provided that $0 \le \gamma < \beta$, we compute

$$\begin{aligned} \int_1^R \ell^{d-1-\beta} d\ell &\lesssim \left\{ \begin{array}{lll} R^{d-\beta} & : & \beta < d \\ \log R & : & \beta = d \end{array} \right\} \simeq R^d \pi_*(R)^{-1}, \\ \int_R^\infty \ell^{-1-\beta} d\ell &\lesssim R^{-\beta}, \\ R^{d-\gamma} \int_R^\infty \ell^{-1-\beta+\gamma} d\ell &\lesssim R^{d-\gamma-(\beta-\gamma)} = R^{d-\beta}, \end{aligned}$$

and the conclusion (2.21) follows.

3.4. **Proof of Proposition 2.6: Main estimates II.** By (2.16) in Lemma 2.4, we have $\partial^{\mathrm{fct}} E(f,g) = G_1 + G_2 + G_3$ with

$$G_1 := g_j\,(\nabla\phi_j^* + e_j) \otimes (\nabla w_f + \phi_i \nabla\nabla_i \bar u), \qquad G_2 := (\phi_j^*\,\nabla g_j + \nabla r) \otimes \nabla u,$$
$$G_3 := -\big(\phi_j^*\,\nabla(g_j \nabla_i \bar u) + \nabla R_i\big) \otimes (\nabla\phi_i + e_i),$$

so that it suffices to estimate the norms of each of the G_i's separately. We split the proof into two main steps: we first address the case of the standard LSI ($\beta > d$), and then turn to the general multiscale case ($\beta \le d$). Let $R \ge 1$ be arbitrary.

Step 1. Proof of (2.22) for standard LSI ($\beta > d$).
Since for standard LSI ($\beta > d$) we have $|||\partial^{\mathrm{fct}} J_0(F)|||_\beta \lesssim \|\partial^{\mathrm{fct}} J_0(F)\|_1$, it suffices to establish the following estimates: for all $\ell \ge 1$, $0 < \alpha - d \ll 1$, and $0 < p - 1 \ll_\alpha 1$,

$$\|G_1\|_\ell^2 \;\lesssim_{\alpha,p}\; \ell^d \Big(\int r_*^{4d} [g]_\infty^4\Big)^{\frac12} \Big(\int \mathcal{C}^4 \mu_*^4 [\nabla^2 \bar u]_\infty^4\Big)^{\frac12}, \tag{3.14}$$

$$\|G_2\|_\ell^2 \;\lesssim_{\alpha,p}\; \ell^d r_*(0)^{\alpha\frac{p-1}{2p}} \Big(\int r_*^{d\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{2p}} \times \Big(\int \mathcal{C}^4 \mu_*^4 [\nabla g]_\infty^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} |f|^{4p}\Big)^{\frac{1}{2p}}, \tag{3.15}$$

$$\|G_3\|_\ell^2 \;\lesssim_{\alpha,p}\; \ell^d r_*(0)^{\alpha\frac{p-1}{p}} \Big(\int r_*^{d\frac{p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \times \Big(\int w_R^{\alpha(p-1)} \mathcal{C}^{2p} \mu_*^{2p} [\nabla(g\nabla\bar u)]_\infty^{2p}\Big)^{\frac1p}. \tag{3.16}$$

Indeed, replacing p by $\frac{2p}{p+1}$, estimating $[\nabla(g\nabla\bar u)]_\infty \lesssim [\nabla g]_\infty [\nabla \bar u]_\infty + [g]_\infty [\nabla^2 \bar u]_\infty$, and using Hölder's inequality, the estimate (3.16) easily leads to

$$\|G_3\|_\ell^2 \lesssim_{\alpha,p} \ell^d r_*(0)^{\alpha\frac{p-1}{2p}} \Big(\int r_*^{d\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{2p}} \times \bigg(\Big(\int \mathcal{C}^4 \mu_*^4 [\nabla^2 \bar u]_\infty^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} [g]_\infty^{4p}\Big)^{\frac{1}{2p}} + \Big(\int \mathcal{C}^4 \mu_*^4 [\nabla g]_\infty^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} [\nabla \bar u]_\infty^{4p}\Big)^{\frac{1}{2p}}\bigg),$$

so that (2.22) follows in combination with (3.14) and (3.15). We address the estimates (3.14)–(3.16) separately, and split the proof into three substeps.

Substep 1.1. Proof of (3.14).
By (2.31) in Lemma 2.9,

$$\|G_1\|_\ell^2 \lesssim \int \Big(\int_{B_{2\ell*}(x)} |g|^2 |\nabla\phi + \mathrm{Id}\,|^2\Big) \Big(\fint_{B_*(x)} (|\nabla w_f| + |\phi||\nabla^2 \bar u|)^2\Big) dx$$

which by Cauchy-Schwarz' inequality turns into

$$\|G_1\|_\ell^2 \lesssim \Big(\int \Big(\int_{B_{2\ell*}(x)} |g|^2|\nabla\phi+\mathrm{Id}\,|^2\Big)^2 dx\Big)^{\frac12} \times \Big(\int \Big(\fint_{B_*(x)} (|\nabla w_f|+|\phi||\nabla^2\bar u|)^2\Big)^2 dx\Big)^{\frac12}. \tag{3.17}$$

We start by treating the first right-hand side factor. Taking the local supremum of g, using the mean-value property (2.26), the Lipschitz continuity of r_*, Jensen's inequality, and (3.4), we obtain

$$\begin{aligned}\int \Big(\int_{B_{2\ell*}(x)} |g|^2|\nabla\phi+\mathrm{Id}\,|^2\Big)^2 dx &\lesssim \int \Big(\int_{B_{2\ell*}(x)} r_*^d[g]_\infty^2\Big)^2 dx\\ &\lesssim \int \Big(\fint_{B_{2\ell*}(x)} r_*^d(\ell+r_*)^d[g]_\infty^2\Big)^2 dx\\ &\le \int r_*^{2d}(\ell+r_*)^{2d}[g]_\infty^4\\ &\lesssim \ell^{2d}\int r_*^{4d}[g]_\infty^4.\end{aligned} \tag{3.18}$$

We turn to the second right-hand side factor in (3.17). Note that the two-scale expansion error w_f satisfies the following equation (cf. [11, proof of Theorem 3]),

$$-\nabla\cdot\boldsymbol{a}\nabla w_f = \nabla\cdot\big((\boldsymbol{a}\phi_j+\sigma_j)\nabla\nabla_j\bar u\big). \tag{3.19}$$

By (3.9) with $q=2$ applied to w_f, we obtain after taking local suprema of $\nabla^2\bar u$, and controlling correctors by Lemma 2.8,

$$\int \Big(\fint_{B_*(x)} (|\nabla w_f|+|\phi||\nabla^2\bar u|)^2\Big)^2 \lesssim \int [(|\phi|+|\sigma|)\nabla^2\bar u]_2^4 \lesssim \int \mathcal{C}^4\mu_*^4[\nabla^2\bar u]_\infty^4. \tag{3.20}$$

Combined with (3.17) and (3.18), this yields the conclusion (3.14).

Substep 1.2. Proof of (3.15).
By Lemma 2.9 in form of (2.31) and Cauchy-Schwarz' inequality,

$$\|G_2\|_\ell^2 \lesssim \Big(\int \Big(\fint_{B_*(x)} (|\phi||\nabla g|+|\nabla r|)^2\Big)^2\Big)^{\frac12}\Big(\int \Big(\int_{B_{2\ell*}(x)} |\nabla u|^2\Big)^2\Big)^{\frac12}. \tag{3.21}$$

We start with the first right-hand side factor. By (3.9) with $q=2$ applied to the solution r of (2.18), we obtain after taking local suprema of ∇g and controlling correctors by Lemma 2.8,

$$\int \Big(\fint_{B_*(x)} (|\phi||\nabla g|+|\nabla r|)^2\Big)^2 \lesssim \int [(|\phi|+|\sigma|)\nabla g]_2^4 \lesssim \int \mathcal{C}^4\mu_*^4[\nabla g]_\infty^4. \tag{3.22}$$

We turn to the second right-hand side factor in (3.21). By the Lipschitz continuity of r_*, Jensen's inequality, and (3.4),

$$\int \Big(\int_{B_{2\ell*}(x)} |\nabla u|^2\Big)^2 \lesssim \int \Big(\fint_{B_{2\ell*}(x)} (\ell+r_*)^d|\nabla u|^2\Big)^2 \lesssim \int (\ell+r_*)^{2d}\Big(\fint_{B_*(x)} |\nabla u|^2\Big)^2.$$

By (3.8) with $s = q = 2$ applied to the solution u of (1.3), we deduce for all $\alpha > d$ and $p > 1$ with $\alpha(p-1) < d(4p-1)$,

$$\int (\ell + r_*)^{2d} \Big(\fint_{B_*(x)} |\nabla u|^2 \Big)^2 \lesssim_{\alpha,p} \ell^{2d} r_*(0)^{\alpha \frac{p-1}{p}} \Big(\int r_*^{d \frac{2p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{p}} \Big(\int w_R^{\alpha(p-1)} |f|^{4p} \Big)^{\frac{1}{p}}. \quad (3.23)$$

Combined with (3.21) and (3.22), this yields (3.15).

Substep 1.3. Proof of (3.16).
By (2.31) in Lemma 2.9 and the mean-value property (2.26), we find

$$\|G_3\|_\ell^2 \lesssim \int \big(\ell + r_*(x)\big)^d \Big(\fint_{B_*(x)} \big(|\phi^*| |\nabla(g\nabla \bar u)| + |\nabla R| \big)^2 \Big) dx.$$

By (3.8) with $s = q = 1$ applied to the solution R of (2.19), we deduce for all $\alpha > d$ and $p > 1$ with $\alpha(p-1) < d(2p-1)$,

$$\|G_3\|_\ell^2 \lesssim \ell^d r_*(0)^{\alpha \frac{p-1}{p}} \Big(\int r_*^{d \frac{p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{p}} \Big(\int w_R^{\alpha(p-1)} [\phi^* \nabla(g\nabla \bar u)]_2^{2p} \Big)^{\frac{1}{p}}.$$

Taking local suprema of $\nabla(g\nabla \bar u)$ and using Lemma 2.8 to control correctors, (3.16) follows.

Step 2. Proof of (2.23) in the general multiscale case ($\beta \le d$).
As in Step 2 of the proof of Proposition 2.5, we need to refine (3.14)–(3.16) in the range $\ell > R$. More precisely, we shall establish that for all $\ell \ge 1$, $0 \le \gamma \le d$, $0 < \alpha - d \ll 1$, and $0 < p - 1 \ll_\alpha 1$,

$$\begin{aligned} \|G_1\|_\ell^2 \lesssim_{\gamma,\alpha,p} & \Big(\int \mathcal{C}^2 \mu_*^2 [\nabla^2 \bar u]_\infty^2 \Big) \Big(\int r_*^d [g]_\infty^2 \Big) \\ & + \ell^\gamma R^{d-\gamma} r_*(0)^{d-\gamma+\alpha\frac{p-1}{2p}} \Big(\int r_*^{\gamma \frac{2p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{2p}} \\ & \times \Big(\int r_*^{2pd} w_R^{2p(d-\gamma)+\alpha(p-1)} [g]_\infty^{4p} \Big)^{\frac{1}{2p}} \Big(\int \mathcal{C}^4 \mu_*^4 [\nabla^2 \bar u]_\infty^4 \Big)^{\frac12}, \quad (3.24) \\ \|G_2\|_\ell^2 \lesssim_{\gamma,\alpha,p} & \Big(\int \mathcal{C}^2 \mu_*^2 [\nabla g]_\infty^2 \Big) \Big(\int |f|^2 \Big) \\ & + \ell^\gamma R^{d-\gamma} r_*(0)^{d-\gamma+\alpha\frac{p-1}{2p}} \Big(\int r_*^{\gamma \frac{2p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{2p}} \\ & \times \Big(\int \mathcal{C}^4 \mu_*^4 [\nabla g]_\infty^4 \Big)^{\frac12} \Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)} |f|^{4p} \Big)^{\frac{1}{2p}}, \quad (3.25) \\ \|G_3\|_\ell^2 \lesssim_{\gamma,\alpha,p} & \ell^{2d\frac{p-1}{p}} r_*(0)^{2d\frac{p-1}{p}} \Big(\int \mathcal{C}^p \mu_*^p [\nabla(g\nabla\bar u)]_\infty^p \Big)^{\frac{2}{p}} \\ & + \ell^\gamma R^{d-\gamma} r_*(0)^{d-\gamma+\alpha\frac{p-1}{p}} \Big(\int r_*^{\gamma \frac{p}{p-1}} w_R^{-\alpha} \Big)^{\frac{p-1}{p}} \\ & \times \Big(\int w_R^{p(d-\gamma)+\alpha(p-1)} \mathcal{C}^{2p} \mu_*^{2p} [\nabla(g\nabla\bar u)]_\infty^{2p} \Big)^{\frac{1}{p}}. \quad (3.26) \end{aligned}$$

Replacing p by $\frac{2p}{p+1}$, estimating $[\nabla(g\nabla\bar{u})]_\infty \lesssim [\nabla g]_\infty[\nabla\bar{u}]_\infty + [g]_\infty[\nabla^2\bar{u}]_\infty$, and using Hölder's inequality, the estimate (3.26) easily leads to

$$\begin{aligned}\|G_3\|_\ell^2 \lesssim_{\gamma,\alpha,p} \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\Big(\int \mathcal{C}^2\mu_*^2[\nabla g]_\infty^2\Big)\Big(\int [\nabla\bar{u}]_\infty^{2p}\Big)^{\frac{1}{p}} \\ + \Big(\int \mathcal{C}^2\mu_*^2[\nabla^2\bar{u}]_\infty^2\Big)\Big(\int [g]_\infty^{2p}\Big)^{\frac{1}{p}}\Big) \\ + \ell^\gamma R^{d-\gamma} r_*(0)^{d-\gamma+\alpha\frac{p-1}{2p}} \Big(\int r_*^{\gamma\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{2p}} \\ \times \Big(\Big(\int \mathcal{C}^4\mu_*^4[\nabla g]_\infty^4\Big)^{\frac{1}{2}} \Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)}[\nabla\bar{u}]_\infty^{4p}\Big)^{\frac{1}{2p}} \\ + \Big(\int \mathcal{C}^4\mu_*^4[\nabla^2\bar{u}]_\infty^4\Big)^{\frac{1}{2}} \Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)}[g]_\infty^{4p}\Big)^{\frac{1}{2p}}\Big).\end{aligned}$$

Starting from (2.12) and appealing to (3.14)–(3.16) for $\ell \le R$ and to (3.24)–(3.26) for $\ell > R$, we obtain the desired estimate (2.23) after arguing as in Substep 2.3 of the proof of Proposition 2.5. The rest of this step is split into three parts and is dedicated to the proof of (3.24)–(3.26).

Substep 2.1. Proof of (3.24).
By (2.32) in Lemma 2.9 and Cauchy-Schwarz' inequality,

$$\begin{aligned}\|G_1\|_\ell^2 \lesssim \Big(\int_{|x|\ge\ell} (\ell + r_*(x))^{2d}\Big(\fint_{B_*(x)} |g|^2|\nabla\phi + \mathrm{Id}\,|^2\Big)^2 dx\Big)^{\frac{1}{2}} \\ \times \Big(\int \Big(\fint_{B_{2\ell*}(x)} (|\nabla w_f| + |\phi||\nabla^2\bar{u}|)^2\Big)^2 dx\Big)^{\frac{1}{2}} \\ + \Big(\int_{B_{7\ell*}(0)} \Big(\fint_{\bar{B}_*(x)} |g|^2|\nabla\phi + \mathrm{Id}\,|^2\Big) dx\Big)\Big(\int_{B_{7\ell*}(0)} \Big(\fint_{\bar{B}_*(x)} (|\nabla w_f| + |\phi||\nabla^2\bar{u}|)^2\Big) dx\Big).\end{aligned}$$

First, we take local suprema of g, apply Lemma 2.8 to control correctors, and use (3.4), to the effect of

$$\int_{B_{7\ell*}(0)} \Big(\fint_{\bar{B}_*(x)} |g|^2|\nabla\phi + \mathrm{Id}\,|^2\Big) dx \lesssim \int r_*^d[g]_\infty^2.$$

Second, using (3.4) and the energy estimate for (3.19), taking local suprema of $\nabla^2\bar{u}$, and using Lemma 2.8 to control correctors, we find

$$\begin{aligned}\int_{B_{7\ell*}(0)} \Big(\fint_{\bar{B}_*(x)} (|\nabla w_f| + |\phi||\nabla^2\bar{u}|)^2\Big) dx \lesssim \int (|\nabla w_f| + |\phi||\nabla^2\bar{u}|)^2 \\ \lesssim \int (|\phi| + |\sigma|)^2|\nabla^2\bar{u}|^2 \lesssim \int \mathcal{C}^2\mu_*^2[\nabla^2\bar{u}]_\infty^2.\end{aligned}$$

Third, appealing to (3.11) with $s = q = 2$ and $|h| = |g||\nabla\phi + \mathrm{Id}\,|$, we obtain for all $0 \le \gamma \le d$, $\alpha > d$, and $p > 1$ with $2p(d-\gamma) + \alpha(p-1) < d(4p-1)$,

$$\int_{|x|\ge\ell} \big(\ell + r_*(x)\big)^{2d}\Big(\int_{B_*(x)} |g|^2|\nabla\phi + \mathrm{Id}\,|^2\Big)^2 dx \lesssim_{\gamma,\alpha,p} \ell^{2\gamma} R^{2(d-\gamma)} r_*(0)^{2(d-\gamma)+\alpha\frac{p-1}{p}} \times \Big(\int r_*^{\gamma\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)} [g(\nabla\phi + \mathrm{Id})]_2^{4p}\Big)^{\frac1p},$$

while the mean-value property (2.26) yields $[g(\nabla\phi + \mathrm{Id})]_2^{4p} \lesssim r_*^{2pd}[g]_\infty^{4p}$. The conclusion (3.24) then follows from the combination of the above estimates with (3.20) (with r_* replaced by $2\ell + r_*$).

Substep 2.2. Proof of (3.25).
By (2.32) in Lemma 2.9 and Cauchy-Schwarz' inequality,

$$\|G_2\|_\ell^2 \lesssim \Big(\int_{|x|\ge\ell} \big(\ell + r_*(x)\big)^{2d}\Big(\fint_{B_*(x)} |\nabla u|^2\Big)^2 dx\Big)^{\frac12} \times \Big(\int \Big(\fint_{B_{2\ell*}(x)} (|\phi||\nabla g| + |\nabla r|)^2\Big)^2 dx\Big)^{\frac12} + \Big(\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} |\nabla u|^2\Big) dx\Big)\Big(\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} (|\phi||\nabla g| + |\nabla r|)^2\Big) dx\Big).$$

First, using (3.4), the energy estimate for (2.18), taking local suprema of g, and applying Lemma 2.8 to control correctors,

$$\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} (|\phi||\nabla g| + |\nabla r|)^2\Big) dx \lesssim \int (|\phi||\nabla g| + |\nabla r|)^2 \lesssim \int (|\phi| + |\sigma|)^2|\nabla g|^2 \lesssim \int \mathcal{C}^2\mu_*^2[\nabla g]_\infty^2.$$

Second, the energy estimate for (1.3) yields

$$\int_{B_{7\ell*}(0)} \Big(\fint_{\bar B_*(x)} |\nabla u|^2\Big) dx \lesssim \int |\nabla u|^2 \lesssim \int |f|^2.$$

Third, appealing to (3.11) with $s = q = 2$ and $v = u$, we obtain for all $0 \le \gamma \le d$, $\alpha > d$, and $p > 1$ with $2p(d-\gamma) + \alpha(p-1) < d(4p-1)$,

$$\int_{|x|\ge\ell} \big(\ell + r_*(x)\big)^{2d}\Big(\fint_{B_*(x)} |\nabla u|^2\Big)^2 dx \lesssim_{\gamma,\alpha,p} \ell^{2\gamma} R^{2(d-\gamma)} r_*(0)^{2(d-\gamma)+\alpha\frac{p-1}{p}} \times \Big(\int r_*^{\gamma\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{p}} \Big(\int w_R^{2p(d-\gamma)+\alpha(p-1)} |f|^{4p}\Big)^{\frac1p}.$$

Finally, applying (3.9) with $q = 2$ (with r_* replaced by $2\ell + r_*$) to the solution r of (2.18), taking local suprema of g, and applying Lemma 2.8 to control correctors,

$$\int \Big(\fint_{B_{2\ell*}(x)} (|\phi||\nabla g| + |\nabla r|)^2\Big)^2 dx \lesssim \int [(|\phi| + |\sigma|)\nabla g]_2^4 \lesssim \int \mathcal{C}^4\mu_*^4[\nabla g]_\infty^4.$$

The combination of these four estimates yields the conclusion (3.25).

Substep 2.3. Proof of (3.26).
By (2.32) in Lemma 2.9, Cauchy-Schwarz' inequality, and the mean-value property (2.26),

$$\|G_3\|_\ell^2 \lesssim \int_{|x|\geq \ell} (\ell + r_*(x))^d \Big(\fint_{B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)\, dx \\ + \Big(\int_{B_{7\ell *}(0)} \Big(\fint_{\bar B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)^{\frac12} dx \Big)^2, \quad (3.27)$$

We start with the second right-hand side term. By Hölder's inequality with exponents $(\frac{p}{p-1}, p)$,

$$\int_{B_{7\ell *}(0)} \Big(\fint_{\bar B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)^{\frac12} dx \\ \lesssim \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int_{B_{7\ell *}(0)} \Big(\fint_{\bar B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)^{\frac p2} dx \Big)^{\frac1p} \\ \lesssim \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int \Big(\fint_{B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)^{\frac p2} dx \Big)^{\frac1p},$$

where we passed from $\bar B_*(x) = B_{5r_*(x)}(x)$ to $B_*(x) = B_{r_*(x)}(x)$. Appealing to the large-scale Caderón-Zygmund estimate of Lemma 2.7(b) with exponent $1 < p \leq 2$ applied to the solution R of (2.19), taking local suprema of $\nabla(g\nabla\bar u)$, and using Lemma 2.8 to control correctors, we deduce

$$\int_{B_{7\ell *}(0)} \Big(\fint_{\bar B_*(x)} \big(|\phi^*||\nabla(g\nabla\bar u)| + |\nabla R|\big)^2 \Big)^{\frac12} dx \lesssim \ell^{d\frac{p-1}{p}} r_*(0)^{d\frac{p-1}{p}} \Big(\int \mathcal{C}^p \mu_*^p [\nabla(g\nabla\bar u)]_\infty^p \Big)^{\frac1p}.$$

We turn to the first right-hand side term in (3.27), apply (3.11) with $s = q = 1$ to (2.19), take local suprema of $\nabla(g\nabla\bar u)$, and use Lemma 2.8 to control correctors. The conclusion (3.26) follows.

4. Proof of the main results

We mainly focus on the proof of the statements for the standard LSI ($\beta > d$), and quickly argue how to adapt the argument to general multiscale LSI ($\beta \leq d$) in the last step.

Step 1. Proof of Proposition 1 for standard LSI ($\beta > d$).
Let $F \in C_c^\infty(\mathbb{R}^d)^{d\times d}$. Starting point is (2.20) in Proposition 2.5. By Hölder's inequality, the triangle inequality in probability, and the stationarity of r_*, we obtain for all $R \geq 1$, $0 < \alpha - d \ll 1$, $0 < p - 1 \ll_\alpha 1$, and $q \gg \frac{1}{p-1}$,

$$\mathbb{E}\Big[\|\!|\partial^{\mathrm{fct}} J_0(F)|\!\|^{2q} \Big]^{\frac1q} \lesssim_{\alpha,p} \mathbb{E}\Big[\Big(r_*^{d+\alpha\frac{p-1}{p}} \Big)^q \Big]^{\frac1q} R^{d\frac{p-1}{p}} \Big(\int w_R^{\alpha(p-1)} |F|^{2p} \Big)^{\frac1p}.$$

Replacing F by $\varepsilon^{\frac d2} F(\varepsilon\cdot)$ and choosing $R = \frac1\varepsilon$, this yields

$$\mathbb{E}\Big[\|\!|\partial^{\mathrm{fct}} \widehat J_0^\varepsilon(F)|\!\|^{2q} \Big]^{\frac1q} \lesssim_{\alpha,p} \mathbb{E}\Big[\Big(r_*^{d+\alpha\frac{p-1}{p}} \Big)^q \Big]^{\frac1q} \Big(\int w_1^{\alpha(p-1)} |F|^{2p} \Big)^{\frac1p}. \quad (4.1)$$

We now recall the following implication (which follows from multiscale LSI in form of the moment bounds in [6, Proposition 3.1(i)]; see also [11, Step 1 of the proof of Theorem 1]): for all random variables Y_1, Y_2, given $q_0 \geq 1$ and $\kappa > 0$,

$$\mathbb{E}\left[\|\partial^{\mathrm{fct}} Y_1\|_\beta^{2q}\right]^{\frac{1}{q}} \leq \mathbb{E}\left[Y_2^q\right]^{\frac{1}{q}} \quad \text{for all } q \geq q_0, \quad \text{and} \quad \mathbb{E}\left[\exp(Y_2^\kappa)\right] \leq 2$$
$$\Longrightarrow \quad \exists\, C \simeq_{q_0,\kappa} 1:\ \mathbb{E}\left[\exp\left(\frac{1}{C} Y_1^{2\frac{\kappa}{1+\kappa}}\right)\right] \leq 2. \quad (4.2)$$

Using this property and the moment bound of Lemma 2.7 for r_*, the estimate (4.1) leads to the conclusion (2.3).

Step 2. Proof of Theorem 1 for standard LSI ($\beta > d$).
Let $f, g \in C_c^\infty(\mathbb{R}^d)$. We split the proof into two substeps: we first improve (2.22) to avoid local suprema in the estimate, and then turn to (2.6) itself.

Substep 2.1. Improvement of (2.22): for all $R \geq 1$, $0 < \alpha - d \ll 1$, and $0 < p - 1 \ll_\alpha 1$,

$$\begin{aligned}\|\partial^{\mathrm{fct}} E(f,g)\|^2 \lesssim_{\alpha,p}\ & r_*(0)^{\alpha\frac{p-1}{2p}} \left(\int r_*^{2d\frac{2p}{p-1}} w_R^{-\alpha}\right)^{\frac{p-1}{2p}} \\ & \times \left(\left(\int \mathcal{C}^4 \mu_*^4 (|\nabla f| + |\nabla^2 \bar u|)^4\right)^{\frac12} \left(\int w_R^{\alpha(p-1)} |g|^{4p}\right)^{\frac{1}{2p}}\right. \\ & \left. + \left(\int \mathcal{C}^4 \mu_*^4 |\nabla g|^4\right)^{\frac12} \left(\int w_R^{\alpha(p-1)} (|f| + |\nabla \bar u|)^{4p}\right)^{\frac{1}{2p}}\right). \quad (4.3)\end{aligned}$$

We first apply (2.22) to the averaged functions f_1 and g_1 defined by $f_1(x) := ⨍_{B(x)} f$ and $g_1(x) := ⨍_{B(x)} g$. Noting that $[f_1]_\infty \lesssim ⨍_{B_2(x)} |f|$ and that the solution $\bar u_1$ of the homogenized equation (1.4) with averaged right-hand side f_1 is given by $\bar u_1(x) = ⨍_{B(x)} \bar u$, and using the Lipschitz continuity of $\mathcal{C}$, we obtain for all $0 < \alpha - d \ll 1$ and $0 < p - 1 \ll_\alpha 1$,

$$\begin{aligned}\|\partial^{\mathrm{fct}} E(f,g)\|^2 \lesssim_{\alpha,p}\ & \|\partial^{\mathrm{fct}} E(f_1 - f, g_1)\|^2 + \|\partial^{\mathrm{fct}} E(f, g_1 - g)\|^2 \\ & + r_*(0)^{\alpha\frac{p-1}{2p}} \left(\int r_*^{2d\frac{2p}{p-1}} w_R^{-\alpha}\right)^{\frac{p-1}{2p}} \left(\left(\int \mathcal{C}^4 \mu_*^4 |\nabla^2 \bar u|^4\right)^{\frac12} \left(\int w_R^{\alpha(p-1)} |g|^{4p}\right)^{\frac{1}{2p}}\right. \\ & \left. + \left(\int \mathcal{C}^4 \mu_*^4 |\nabla g|^4\right)^{\frac12} \left(\int w_R^{\alpha(p-1)} (|f| + |\nabla \bar u|)^{4p}\right)^{\frac{1}{2p}}\right). \quad (4.4)\end{aligned}$$

It remains to estimate the first two right-hand side terms of (4.4), which we will prove to be small not because the two-scale expansion is accurate, but because $f_1 - f$ and $g_1 - g$ are small themselves after rescaling. Arguing as in the proof of (2.16), we have the alternative formula

$$\partial^{\mathrm{fct}} E(f,g) = g \otimes \nabla u - g \otimes \nabla \bar u - g \nabla_i \bar u \otimes (\nabla \phi_i + e_i) + \nabla t \otimes \nabla u - \nabla T_i \otimes (\nabla \phi_i + e_i),$$

where the auxiliary fields ∇t and ∇T_i are the gradient solutions in $\mathrm{L}^2(\mathbb{R}^d)^d$ of

$$\begin{aligned} -\nabla \cdot \boldsymbol{a}^* \nabla t &= \nabla \cdot ((\boldsymbol{a}^* - \bar{\boldsymbol{a}}^*) g), \\ -\nabla \cdot \boldsymbol{a}^* \nabla T_i &= \nabla \cdot ((\boldsymbol{a}^* - \bar{\boldsymbol{a}}^*) g \nabla_i \bar u). \end{aligned}$$

Using this decomposition and arguing as in the proof of Proposition 2.5, we obtain for all $R \geq 1$, $0 < \alpha - d \ll 1$, and $0 < p - 1 \ll_\alpha 1$,

$$\|\partial^{\mathrm{fct}} E(f,g)\|^2 \lesssim r_*(0)^{\alpha\frac{p-1}{2p}} \Big(\int r_*^{d\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{2p}} \\ \times \min\bigg\{\Big(\int (|f| + |\nabla \bar{u}|)^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} |g|^{4p}\Big)^{\frac{1}{2p}} ; \\ \Big(\int |g|^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} (|f| + |\nabla \bar{u}|)^{4p}\Big)^{\frac{1}{2p}}\bigg\}.$$

Using in addition that $|f - f_1| \leq \int_0^1 \fint_{tB} |\nabla f(\cdot + y)| dy dt$, this turns into

$$\|\partial^{\mathrm{fct}} E(f_1 - f, g_1)\| + \|\partial^{\mathrm{fct}} E(f, g_1 - g)\| \\ \lesssim r_*(0)^{\alpha\frac{p-1}{2p}} \Big(\int r_*^{d\frac{2p}{p-1}} w_R^{-\alpha}\Big)^{\frac{p-1}{2p}} \bigg(\Big(\int (|\nabla f| + |\nabla^2 \bar{u}|)^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} |g|^{4p}\Big)^{\frac{1}{2p}} \\ + \Big(\int |\nabla g|^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} (|f| + |\nabla \bar{u}|)^{4p}\Big)^{\frac{1}{2p}}\bigg).$$

Combining this with (4.4) leads to the conclusion (4.3).

Substep 2.2. Conclusion.
By Hölder's inequality, the triangle inequality in probability, and the stationarity of r_*, the estimate (4.3) leads to the following: for all $R \geq 1$, $0 < \alpha - d \ll 1$, $0 < p - 1 \ll_\alpha 1$, and $q \gg \frac{1}{p-1}$,

$$\mathbb{E}\Big[\|\partial^{\mathrm{fct}} E(f,g)\|^{2q}\Big]^{\frac1q} \lesssim_{\alpha,p} \mathbb{E}\bigg[\Big(r_*^{2d+\alpha\frac{p-1}{2p}} \mathcal{C}^2\Big)^q\bigg]^{\frac1q} \\ \times R^{\frac{d}{2}(1-\frac1p)} \bigg(\Big(\int \mu_*^4 (|\nabla f| + |\nabla^2 \bar{u}|)^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} |g|^{4p}\Big)^{\frac{1}{2p}} \\ + \Big(\int \mu_*^4 |\nabla g|^4\Big)^{\frac12} \Big(\int w_R^{\alpha(p-1)} (|f| + |\nabla \bar{u}|)^{4p}\Big)^{\frac{1}{2p}}\bigg).$$

We then apply the standard weighted Calderón-Zygmund theory to the constant-coefficient equation (1.4) for $\bar{u}$, and replace f and g by $\varepsilon^{\frac{d}{4}} f(\varepsilon\cdot)$ and $\varepsilon^{\frac{d}{4}} g(\varepsilon\cdot)$. For the choice $R = \frac1\varepsilon$, and by the bound $\mu_*(\frac{\cdot}{\varepsilon}) \lesssim \mu_*(\frac1\varepsilon)\mu_*(\cdot)$, this implies

$$\mathbb{E}\Big[\|\partial^{\mathrm{fct}} \widehat{E}^\varepsilon(f,g)\|^{2q}\Big]^{\frac1q} \lesssim_{\alpha,p} \mathbb{E}\bigg[\Big(r_*^{2d+\alpha\frac{p-1}{2p}} \mathcal{C}^2\Big)^q\bigg]^{\frac1q} \\ \times \varepsilon^2 \mu_*(\tfrac1\varepsilon)^2 \bigg(\Big(\int \mu_*^4 |\nabla f|^4\Big)^{\frac12} \Big(\int w_1^{\alpha(p-1)} |g|^{4p}\Big)^{\frac{1}{2p}} \\ + \Big(\int \mu_*^4 |\nabla g|^4\Big)^{\frac12} \Big(\int w_1^{\alpha(p-1)} |f|^{4p}\Big)^{\frac{1}{2p}}\bigg). \quad (4.5)$$

We now recall the following version of Hölder's inequality: for all random variables Y_1, Y_2, given $\kappa_1, \kappa_2 > 0$,

$$\mathbb{E}\left[\exp\left(Y_1^{\kappa_1}\right)\right] \le 2 \quad \text{and} \quad \mathbb{E}\left[\exp\left(Y_2^{\kappa_2}\right)\right] \le 2 \\ \Longrightarrow \quad \exists C \simeq_{\kappa_1,\kappa_2} 1: \ \mathbb{E}\left[\exp\left(\frac{1}{C}(Y_1Y_2)^{\frac{\kappa_1\kappa_2}{\kappa_1+\kappa_2}}\right)\right] < \infty. \quad (4.6)$$

Using this property, the moment bounds of Lemmas 2.7 and 2.8 for r_* and $\mathcal{C}$ yield, for all $\eta > 0$, $\mathbb{E}\big[\exp(\frac{1}{C_\eta}(r_*^{2d}\mathcal{C}^2)^{\frac14-\eta})\big] \le 2$ for some $C_\eta \simeq_\eta 1$. Combining this with (4.5), property (4.2) yields the conclusion (2.6).

Step 3. General multiscale LSI ($\beta \le d$).
We start with Proposition 1. By Hölder's inequality, the triangle inequality in probability, and the stationarity of r_*, the estimate (2.21) in Proposition 2.5 leads to the following: for all $R \ge 1$, $0 < \gamma < \beta$, $0 < \alpha - d \ll 1$, $0 < p-1 \ll_{\gamma,\alpha} 1$, and $q \gg \frac{1}{p-1}$,

$$\mathbb{E}\left[\|\!|\partial^{\mathrm{fct}} J_0(F)|\!\|_\beta^{2q}\right]^{\frac1q} \lesssim_{\gamma,\alpha,p} \mathbb{E}\left[\left(r_*^{d+\alpha\frac{p-1}{p}}\right)^q\right]^{\frac1q} R^{2d}\pi_*(R)^{-1} \\ \times\left(R^{-\frac dp}\left(\int w_R^{(d-\gamma)p+\alpha(p-1)}|F|^{2p}\right)^{\frac1p} + R^{-\frac{2d}{p}}\left(\int [F]_2^p\right)^{\frac2p}\right).$$

Replacing F by $\varepsilon^d\pi_*(\frac1\varepsilon)^{\frac12}F(\varepsilon\cdot)$ and choosing $R = \frac1\varepsilon$, this yields

$$\mathbb{E}\left[\|\!|\partial^{\mathrm{fct}} \widehat{J}_0^\varepsilon(F)|\!\|_\beta^{2q}\right]^{\frac1q} \lesssim_{\gamma,\alpha,p} \mathbb{E}\left[\left(r_*^{d+\alpha\frac{p-1}{p}}\right)^q\right]^{\frac1q} \\ \times\left(\left(\int w_1^{(d-\gamma)p+\alpha(p-1)}|F|^{2p}\right)^{\frac1p} + \left(\int [F]_2^p\right)^{\frac2p}\right).$$

The combination of this estimate with property (4.2) and with the moment bound of Lemma 2.7 for r_* implies the desired estimate (2.3).

We finally turn to Theorem 1. Arguing as in Substep 2.1 above, we may get rid of local suprema in the estimate (2.23) in Proposition 2.6. Using then Hölder's inequality, the triangle inequality in probability, and the stationarity of r_*, we obtain the following: for all $R \ge 1$, $0 \le \gamma < \beta$, $0 < \alpha - d \ll 1$, $0 < p-1 \ll_\alpha 1$, and $q \gg \frac{1}{p-1}$,

$$\mathbb{E}\left[\|\!|\partial^{\mathrm{fct}} E(f,g)|\!\|^{2q}\right]^{\frac1q} \lesssim_{\gamma,\alpha,p} \mathbb{E}\left[\left(r_*^{2d+\alpha\frac{p-1}{2p}}\mathcal{C}^2\right)^q\right]^{\frac1q} \\ \times\Bigg(R^{d+\frac d2(1-\frac1p)}\pi_*(R)^{-1}\bigg(\left(\int \mu_*^4(|\nabla f| + |\nabla^2\bar u|)^4\right)^{\frac12}\left(\int w_R^{2p(d-\gamma)+\alpha(p-1)}|g|^{4p}\right)^{\frac{1}{2p}} \\ + \left(\int \mu_*^4|\nabla g|^4\right)^{\frac12}\left(\int w_R^{2p(d-\gamma)+\alpha(p-1)}(|f| + |\nabla\bar u|)^{4p}\right)^{\frac{1}{2p}}\bigg) \\ + R^{-\beta}\left(\int \mu_*^2(|\nabla f| + |\nabla^2\bar u|)^2\right)\left(\left(\int |g|^2\right) + R^{d-\frac dp}\left(\int |g|^{2p}\right)^{\frac1p}\right) \\ + R^{-\beta}\left(\int \mu_*^2|\nabla g|^2\right)\left(\left(\int (|f| + |\nabla\bar u|)^2\right) + R^{d-\frac dp}\left(\int (|f| + |\nabla\bar u|)^{2p}\right)^{\frac1p}\right)\Bigg).$$

Since in dimension $d \geq 2$ the weights μ_*^2 and μ_*^4 always belong to the Muckenhoupt classes A_2 and A_4, respectively, we may apply the standard weighted Calderón-Zygmund theory to the constant-coefficient equation (1.4) for $\bar{u}$ in order to simplify the above right-hand side. Replacing then f and g by $\pi_*(\frac{1}{\varepsilon})^{\frac{1}{4}}\varepsilon^{\frac{d}{2}}f(\varepsilon\cdot)$ and $\pi_*(\frac{1}{\varepsilon})^{\frac{1}{4}}\varepsilon^{\frac{d}{2}}g(\varepsilon\cdot)$, choosing $R=\frac{1}{\varepsilon}$, and using the bound $\mu_*(\frac{\cdot}{\varepsilon}) \lesssim \mu_*(\frac{1}{\varepsilon})\mu_*(\cdot)$, the conclusion (2.6) follows as in Substep 2.2.

Acknowledgements

The work of MD is supported by F.R.S.-FNRS through a Research Fellowship and by the CNRS-Momentum program. AG acknowledges financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2014-2019 Grant Agreement QUANTHOM 335410).

References

[1] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, 208:999–1154, 2017.

[2] S. Armstrong, T. Kuusi, and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2019.

[3] S. Armstrong, C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(2):423–481, 2016.

[4] G. Bal, J. Garnier, S. Motsch, and V. Perrier. Random integrals and correctors in homogenization. *Asymptot. Anal.*, 59(1-2):1–26, 2008.

[5] M. Duerinckx, J. Fischer, and A. Gloria. Scaling limit of the homogenization commutator for correlated Gaussian coefficient fields. Preprint, arXiv:1910.04088.

[6] M. Duerinckx and A. Gloria. Multiscale functional inequalities in probability: Concentration properties. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 17:133–157, 2020.

[7] M. Duerinckx and A. Gloria. Multiscale functional inequalities in probability: Constructive approach. *Ann. Henri Lebesgue*, 2020. In press.

[8] M. Duerinckx, A. Gloria, and F. Otto. The structure of fluctuations in stochastic homogenization. *Commun. Math. Phys.*, 2020. In press.

[9] M. Duerinckx and F. Otto. Higher-order pathwise theory of fluctuations in stochastic homogenization. *Stoch. PDE: Anal. Comp.*, 2019.

[10] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. *Milan J. Math.*, 88(1):99–170, 2020.

[11] A. Gloria, S. Neukamm, and F. Otto. Quantitative stochastic homogenization for correlated fields. Preprint, arXiv:1910.05530.

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

[13] Y. Gu and G. Bal. Random homogenization and convergence to integrals with respect to the Rosenblatt process. *J. Differential Equations*, 253(4):1069–1087, 2012.

[14] Y. Gu and J.-C. Mourrat. Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.*, 14(1):452–481, 2016.

[15] S. M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.

[16] A. Lechiheb, I. Nourdin, G. Zheng, and E. Haouala. Convergence of random oscillatory integrals in the presence of long-range dependence and application to homogenization. *Probab. Math. Statist.*, 38(2):271–286, 2018.

[17] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[18] M. S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete*, 50(1):53–83, 1979.

[19] S. Torquato. *Random heterogeneous materials*, volume 16 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002.

(Mitia Duerinckx) Université Paris-Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, 91405 Orsay, France & Université Libre de Bruxelles, Département de Mathématique, 1050 Brussels, Belgium

Email address: mduerinc@ulb.ac.be

(Antoine Gloria) Sorbonne Université, CNRS, Université de Paris, Laboratoire Jacques-Louis Lions, 75005 Paris, France & Université Libre de Bruxelles, Département de Mathématique, 1050 Brussels, Belgium

Email address: antoine.gloria@upmc.fr

(Felix Otto) Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Email address: otto@mis.mpg.de