



# Higher-order pathwise theory of fluctuations in stochastic homogenization

Mitia Duerinckx<sup>1,2</sup> · Felix Otto<sup>3</sup>

Received: 7 March 2019 / Revised: 24 October 2019 / Published online: 28 November 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

We consider linear elliptic equations in divergence form with stationary random coefficients of integrable correlations. We characterize the fluctuations of a macroscopic observable of a solution to relative order  $\frac{d}{2}$ , where  $d$  is the spatial dimension; the fluctuations turn out to be Gaussian. As for previous work on the leading order, this higher-order characterization relies on a pathwise proximity of the macroscopic fluctuations of a general solution to those of the (higher-order) correctors, via a (higher-order) two-scale expansion injected into the “homogenization commutator”, thus confirming the scope of this notion. This higher-order generalization sheds a clearer light on the algebraic structure of the higher-order versions of correctors, flux correctors, two-scale expansions, and homogenization commutators. It reveals that in the same way as this algebra provides a higher-order theory for microscopic spatial oscillations, it also provides a higher-order theory for macroscopic random fluctuations, although both phenomena are not directly related. We focus on the model framework of an underlying Gaussian ensemble, which allows for an efficient use of (second-order) Malliavin calculus for stochastic estimates. On the technical side, we introduce annealed Calderón–Zygmund estimates for the elliptic operator with random coefficients, which conveniently upgrade the known quenched large-scale estimates.

**Keywords** Stochastic homogenization · Higher-order homogenization · Linear elliptic equation · Fluctuations · Homogenization commutator

---

✉ Mitia Duerinckx  
mduerinc@ulb.ac.be

Felix Otto  
otto@mis.mpg.de

<sup>1</sup> Laboratoire de Mathématique d’Orsay, UMR 8628, Université Paris-Sud, Orsay, France

<sup>2</sup> Département de Mathématique, Université Libre de Bruxelles, Brussels, Belgium

<sup>3</sup> Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

## Contents

1	Introduction	626
2	Higher-order theory of oscillations	633
3	Higher-order homogenization commutators	642
4	Reminder on Malliavin calculus	649
5	Representation formulas	653
6	Annealed Calderón–Zygmund theory	661
7	Proof of the higher-order pathwise result	673
8	Proof of the convergence of the covariance structure	678
9	Proof of the normal approximation result	680
	Appendix A. More Malliavin calculus	688
	References	690

## 1 Introduction

### 1.1 General overview

Let  $\mathbf{a}$  be a random coefficient field on  $\mathbb{R}^d$  that is stationary and ergodic and satisfies the boundedness and ellipticity properties

$$|\mathbf{a}(x)\xi| \leq |\xi|, \quad \xi \cdot \mathbf{a}(x)\xi \geq \lambda|\xi|^2, \quad \text{for all } x, \xi \in \mathbb{R}^d, \quad (1.1)$$

for some  $\lambda > 0$ , and denote by  $(\Omega, \mathcal{F}, \mathbb{P})$  the underlying probability space. Given 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 Lax–Milgram solutions to the rescaled problems

$$-\nabla \cdot \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla u_\varepsilon = \nabla \cdot f \quad \text{in } \mathbb{R}^d, \quad (1.2)$$

where the rescaled coefficients  $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)$  vary on the “microscopic” scale  $\varepsilon$ , and we study “macroscopic” observables of the form  $\int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon$  with  $g \in C_c^\infty(\mathbb{R}^d)^d$  deterministic. Qualitative homogenization theory [34,45] states that almost surely  $\int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon \rightarrow \int_{\mathbb{R}^d} g \cdot \nabla \bar{u}$  as  $\varepsilon \downarrow 0$ , where  $\nabla \bar{u}$  solves the (deterministic) homogenized equation

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = \nabla \cdot f \quad \text{in } \mathbb{R}^d,$$

and the homogenized coefficients  $\bar{\mathbf{a}} \in \mathbb{R}^{d \times d}$  are given by  $\bar{\mathbf{a}} e_i = \mathbb{E}[\mathbf{a}(\nabla \varphi_i + e_i)]$  for  $1 \leq i \leq d$ , in terms of the corrector  $\varphi_i$ , that is, the unique (up to a random additive constant) almost sure solution of the corrector equation,

$$-\nabla \cdot \mathbf{a}(\nabla \varphi_i + e_i) = 0 \quad \text{in } \mathbb{R}^d, \quad (1.3)$$

in the class of functions the gradient of which is stationary, centered (that is,  $\mathbb{E}[\nabla \varphi_i] = 0$ ), and has finite second moment. In other words, the field-flux constitutive relation  $\nabla w_\varepsilon \mapsto \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla w_\varepsilon$  is replaced on the macroscopic scale by the effective relation  $\nabla \bar{w} \mapsto \bar{\mathbf{a}} \nabla \bar{w}$ .

A well-travelled question in homogenization is to further characterize the oscillations of the solution  $u_\varepsilon$ . As expected from formal asymptotics, the so-called two-scale

expansion<sup>1</sup>  $(1 + \varepsilon\varphi_i(\frac{\cdot}{\varepsilon})\nabla_i)\bar{u}$  precisely captures the oscillations of  $u_\varepsilon$  to order  $O(\varepsilon)$  (in dimension  $d > 2$ ), in the sense of

$$\|\nabla(u_\varepsilon - (1 + \varepsilon\varphi_i(\frac{\cdot}{\varepsilon})\nabla_i)\bar{u})\|_{L^2(\mathbb{R}^d \times \Omega)} = O(\varepsilon).$$

This expansion has a natural geometric interpretation: by definition (1.3), the corrector  $\varphi_i$  defines the correction of Euclidean coordinates  $x \mapsto x_i$  into  $\mathbf{a}$ -harmonic coordinates  $x \mapsto x_i + \varphi_i(x)$ ,

$$-\nabla \cdot \mathbf{a}\nabla(x_i + \varphi_i) = 0,$$

so that the expansion  $(1 + \varepsilon\varphi_i(\frac{\cdot}{\varepsilon})\nabla_i)\bar{u}$  amounts to locally correcting the homogenized solution  $\bar{u}$  in terms of oscillating  $\mathbf{a}$ -harmonic coordinates. Such expansions are naturally pursued to higher order: while the (centered) first-order corrector  $\varphi = \varphi^1$  is characterized by the property that  $(1 + \varphi_i\nabla_i)\bar{\ell}$  is  $\mathbf{a}$ -harmonic for all affine functions  $\bar{\ell}$ , the (centered) second-order corrector  $\varphi^2$  is characterized by the property that for all quadratic polynomials  $\bar{q}$  the expansion  $(1 + \varphi_i\nabla_i + \varphi_{ij}^2\nabla_{ij}^2)\bar{q}$  fully captures the oscillations imposed by the heterogeneous elliptic operator  $-\nabla \cdot \mathbf{a}\nabla$  in the sense of  $-\nabla \cdot \mathbf{a}\nabla(1 + \varphi_i\nabla_i + \varphi_{ij}^2\nabla_{ij}^2)\bar{q}$  being deterministic (see Sect. 2 for a precise statement). The second-order two-scale expansion  $(1 + \varepsilon\varphi_i(\frac{\cdot}{\varepsilon})\nabla_i + \varepsilon^2\varphi_{ij}^2(\frac{\cdot}{\varepsilon})\nabla_{ij}^2)\bar{u}_\varepsilon^2$  then captures the oscillations of  $u_\varepsilon$  to order  $O(\varepsilon^2)$  (in dimension  $d > 4$ ), where  $\bar{u}_\varepsilon^2$  is a proxy for a solution to the generically ill-posed higher-order homogenized equation

$$-\nabla \cdot (\bar{\mathbf{a}} + \varepsilon\bar{\mathbf{a}}_i^2\nabla_i)\nabla\bar{U}_\varepsilon^2 = \nabla \cdot f \quad \text{in } \mathbb{R}^d,$$

in terms of the second-order homogenized coefficients  $\bar{\mathbf{a}}^2e_j := \mathbb{E}[\mathbf{a}(\nabla\varphi_{ij}^2 + \varphi_ie_j)]$ . In other words, the effective field-flux constitutive relation  $\nabla\bar{w} \mapsto \bar{\mathbf{a}}\nabla\bar{w}$  is refined to its second-order version  $\nabla\bar{w} \mapsto (\bar{\mathbf{a}} + \varepsilon\bar{\mathbf{a}}_i^2\nabla_i)\nabla\bar{w}$ . The proxy  $\bar{u}_\varepsilon^2$  for a second-order homogenized solution can e.g. be chosen in form of  $\bar{u}_\varepsilon^2 := \bar{u} + \varepsilon\bar{u}^2$  with  $\bar{u}^2$  solving

$$-\nabla \cdot (\bar{\mathbf{a}}\nabla\bar{u}^2 + \bar{\mathbf{a}}_i^2\nabla_i\nabla\bar{u}) = 0 \quad \text{in } \mathbb{R}^d.$$

Such a description of oscillations of  $u_\varepsilon$  via higher-order two-scale expansions is classical in the periodic setting [11]. It also holds in the random setting for large enough dimension and for decaying enough correlations [9,24] (cf. [3,20,22] for the first-order level); under appropriate mixing conditions, higher-order expansions allow to capture oscillations up to the critical order  $O(\varepsilon^{d/2})$  (with a logarithmic correction in even dimensions). The limitation is due to the fact that, in the random setting, not all higher-order correctors are well-behaved:  $\varphi^n$  can be chosen stationary with finite moments only for  $n < \frac{d}{2}$ . Precise statements and short proofs are included in Sect. 2 below in the Gaussian setting.

While periodic homogenization boils down to the description of oscillations of  $u_\varepsilon$ , stochastic homogenization in addition means studying the random fluctuations of the macroscopic observables  $\int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon$ . It was recently shown [25] that the centered

<sup>1</sup> We systematically use Einstein’s summation rule on repeated indices.

rescaled observables  $\varepsilon^{-d/2} \int_{\mathbb{R}^d} g \cdot (\nabla u_\varepsilon - \mathbb{E}[\nabla u_\varepsilon])$  converge in law to a Gaussian. Note that this does not enter the above theory of oscillations since accuracy there is precisely limited to order  $O(\varepsilon^{d/2})$ . We may naturally look for a finer description of this convergence by means of a two-scale expansion. As observed in [25], however, the limiting variance of  $\varepsilon^{-d/2} \int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon$  generically differs from that of  $\varepsilon^{-d/2} \int_{\mathbb{R}^d} g \cdot \nabla(1 + \varepsilon\varphi_i(\frac{\cdot}{\varepsilon})\nabla_i)\bar{u}$ , which shows that two-scale expansion cannot be applied naïvely when it comes to fluctuations. In [19], we unravelled the full mechanism behind this observation by means of the so-called homogenization commutator, which led to a new pathwise theory of fluctuations in stochastic homogenization (in agreement with the pathwise heuristics in [25]). More precisely, fluctuations of  $\nabla u_\varepsilon$  were shown to coincide with fluctuations of a suitable deterministic (Helmholtz) projection of the corresponding homogenization commutator, which behaves like a Gaussian white noise on large scales and for which the two-scale expansion is accurate at leading order. In the present article we explain how this result naturally extends to higher order in parallel to the known theory of oscillations: we obtain a full characterization of fluctuations of  $\varepsilon^{-d/2} \int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon$  to order  $O(\varepsilon^{d/2})$  (with again a logarithmic correction in even dimensions). The main emphasis is on unravelling the suitable higher-order commutator structure.

The homogenization commutator can be viewed as a commutator between large-scale averaging and the field-flux relation; on first-order level it takes the form

$$\Xi_\varepsilon^1[\nabla w_\varepsilon] := (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}})\nabla w_\varepsilon.$$

This expression first appeared in [5] (see also [7]) and is particularly natural since  $H$ -convergence for (1.2) (cf. [39]) is precisely equivalent to weak convergence of  $\Xi_\varepsilon[\nabla u_\varepsilon]$  to 0 (cf. [19]), thus summarizing the whole qualitative homogenization theory. This convergence property is the mathematical formulation of the Hill–Mandel relation in mechanics [27, 28] and can further be made quantitative [19]. For a higher-order theory, we need a higher-order extension of the commutator. In view of the second-order effective field-flux relation  $\nabla \bar{w} \mapsto (\bar{\mathbf{a}} + \varepsilon \bar{\mathbf{a}}_i^2 \nabla_i)\nabla \bar{w}$ , the second-order commutator is naturally defined as

$$\Xi_\varepsilon^2[\nabla w_\varepsilon] := (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}} - \varepsilon \bar{\mathbf{a}}_i^2 \nabla_i)\nabla w_\varepsilon,$$

and a higher-order Hill–Mandel relation can indeed be formulated in these terms (cf. Lemma 3.3 below). Note that  $w_\varepsilon \mapsto \Xi_\varepsilon^1[\nabla w_\varepsilon]$  and  $\Xi_\varepsilon^2[\nabla w_\varepsilon]$  are viewed as first- and second-order differential operators, respectively. Next, we define suitable two-scale expansions of these homogenization commutators. For the first order, we inject the first-order two-scale expansion of  $w_\varepsilon$  into  $\Xi_\varepsilon^1[\nabla \cdot]$  and truncate the obtained differential operator at first order, thus defining the following first-order standard commutator,

$$\Xi_\varepsilon^{\circ,1}[\nabla \bar{w}] := \nabla_i \bar{w} (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}})(\nabla \varphi_i(\frac{\cdot}{\varepsilon}) + e_i).$$

Similarly, injecting the second-order two-scale expansion of  $w_\varepsilon$  into  $\Xi_\varepsilon^2[\nabla \cdot]$  and truncating the obtained differential operator at second order, we define the following second-order standard commutator,

$$\begin{aligned} \Xi_\varepsilon^{\circ,2}[\nabla\bar{w}] &:= \nabla_i\bar{w} \left( \mathbf{a}(\tfrac{\cdot}{\varepsilon}) - \bar{\mathbf{a}} - \varepsilon\bar{\mathbf{a}}_j^2\nabla_j \right) (\nabla\varphi_i(\tfrac{\cdot}{\varepsilon}) + e_i) \\ &+ \varepsilon\nabla_{ij}^2\bar{w} \left( \left( \mathbf{a}(\tfrac{\cdot}{\varepsilon}) - \bar{\mathbf{a}} - \varepsilon\bar{\mathbf{a}}_l^2\nabla_l \right) (\nabla\varphi_{ij}^2(\tfrac{\cdot}{\varepsilon}) + \varphi_i(\tfrac{\cdot}{\varepsilon})e_j) - \bar{\mathbf{a}}_j^2(\nabla\varphi_i(\tfrac{\cdot}{\varepsilon}) + e_i) \right). \end{aligned}$$

The standard commutators  $\bar{w} \mapsto \Xi_\varepsilon^{\circ,1}[\nabla\bar{w}]$  and  $\Xi_\varepsilon^{\circ,2}[\nabla\bar{w}]$  are first- and second-order differential operators, respectively, with  $\varepsilon$ -rescaled (distributional) stationary random coefficients. These stationary differential operators are viewed as intrinsic quantities, with coefficients given by suitable combinations of correctors and of their gradients. The above definitions are efficiently summarized in the form

$$\begin{aligned} \Xi_\varepsilon^{\circ,1}[\nabla\bar{w}](x) &= \Xi_\varepsilon^1[\nabla(1 + \varepsilon\varphi_i(\tfrac{\cdot}{\varepsilon})\nabla_i)T_x^1\bar{w}](x), \\ \Xi_\varepsilon^{\circ,2}[\nabla\bar{w}](x) &= \Xi_\varepsilon^2[\nabla(1 + \varepsilon\varphi_i(\tfrac{\cdot}{\varepsilon})\nabla_i + \varepsilon^2\varphi_{ij}^2(\tfrac{\cdot}{\varepsilon})\nabla_{ij}^2)T_x^2\bar{w}](x), \end{aligned}$$

where  $T_x^1\bar{w}$  and  $T_x^2\bar{w}$  denote the first- and second-order Taylor polynomials of  $\bar{w}$  at the basepoint  $x$ , respectively. Note that these quantities are centered:  $\mathbb{E}[\Xi_\varepsilon^{\circ,n}[\nabla\bar{w}]] = 0$  for  $n = 1, 2$  and any smooth deterministic field  $\bar{w}$ .

Pursuing these constructions to higher order, our main result is threefold and is summarized as follows, under appropriate mixing conditions:

- (i) Up to order  $O(\varepsilon^{d/2})$  in the fluctuation scaling, the fluctuations of  $\nabla u_\varepsilon$  are determined by those of the corresponding higher-order commutators; cf. Theorem 1(i) below.
- (ii) Up to order  $O(\varepsilon^{d/2})$  in the fluctuation scaling (with a logarithmic correction in even dimensions), the two-scale expansions of the commutators are accurate: the fluctuations of the commutators are equivalent to those of the corresponding standard commutators; cf. Theorem 1(ii). This fully extends the pathwise theory of [19] to higher order.
- (iii) The standard commutators are approximately local functions of the coefficients  $\mathbf{a}$ . This allows to infer a quantitative characterization of their limiting covariance structure, as well as a quantitative central limit theorem (CLT); cf. Theorem 1(iii)–(iv). Similarly as for the coefficient field  $\mathbf{a}$  itself, while at first order the scaling limit is given by a Gaussian white noise, corrections converge to derivatives of white noise.

### 1.2 Main results

For simplicity, we focus on the model setting of a Gaussian coefficient field, in which case a powerful Malliavin calculus is available on the probability space, substantially simplifying the analysis. More precisely, we set

$$\mathbf{a}(x) := a_0(G(x)), \tag{1.4}$$

where  $G$  is some  $\mathbb{R}^\kappa$ -valued centered stationary Gaussian random field on  $\mathbb{R}^d$  constructed on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and where  $a_0 \in C_b^2(\mathbb{R}^\kappa)^{d \times d}$  is such that the ellipticity and boundedness assumptions (1.1) are satisfied. We use scalar notation

$a'_0, a''_0$  for derivatives of  $a_0$ . In addition, we assume that the Gaussian field  $G$  has integrable correlations: more precisely, it has the form  $G = c_0 * \xi$ , where  $\xi$  denotes  $\kappa$ -dimensional white noise on  $\mathbb{R}^d$  and where the kernel  $c_0: \mathbb{R}^d \rightarrow \mathbb{R}^{\kappa \times \kappa}$  is chosen even (that is,  $c_0(x) = c_0(-x)$  for all  $x$ ) and satisfies the integrability condition

$$\int_{\mathbb{R}^d} \left( \sup_{B(x)} |c_0| \right) dx \leq 1, \tag{1.5}$$

where  $B(x)$  is the unit ball centered at  $x$ . The covariance function of  $G$  is then given by  $c := c_0 * c_0$  and satisfies  $\int_{\mathbb{R}^d} (\sup_{B(x)} |c|) dx \leq 1$ . The case of non-integrable correlations could be treated similarly, but it would lead to different scalings, cf. [16]. Our main result is as follows.

**Theorem 1** *Consider the above Gaussian setting with integrable correlations (1.5) and set  $\ell := \lceil \frac{d}{2} \rceil$ , that is, the smallest integer  $\geq \frac{d}{2}$ . For all  $1 \leq n \leq \ell$ , let  $\bar{u}_\varepsilon^n$  and  $\bar{v}_\varepsilon^n$  be the  $n$ th-order homogenized solutions of the primal and dual equations, respectively, as in Definition 2.7 below, and let the  $n$ th-order homogenization commutator  $\Xi_\varepsilon^n$  and its standard version  $\Xi_\varepsilon^{\circ,n}$  be as in Definitions 3.1 and 3.4. Then the following hold for all  $0 < \varepsilon \leq \frac{1}{2}$ .*

- (i) Reduction to commutators: For all  $f, g \in C_c^\infty(\mathbb{R}^d)^d$ ,  $1 \leq n \leq \ell$ , and  $p < \infty$ ,

$$N_p \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^n \cdot \Xi_\varepsilon^n [\nabla u_\varepsilon] \right) \lesssim_{p,f,g} \varepsilon^n,$$

where  $N_p(X) := \mathbb{E} [|X - \mathbb{E}[X]|^p]^{\frac{1}{p}}$ .

- (ii) Two-scale expansion of commutators: For all  $f, g \in C_c^\infty(\mathbb{R}^d)^d$ ,  $1 \leq n \leq \ell$ , and  $p < \infty$ ,

$$N_p \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot (\Xi_\varepsilon^n [\nabla u_\varepsilon] - \Xi_\varepsilon^{\circ,n} [\nabla \bar{u}_\varepsilon^n]) \right) \lesssim_{p,f,g} \varepsilon^n \mu_{d,n}(\frac{1}{\varepsilon}),$$

where

$$\varepsilon^n \mu_{d,n}(\frac{1}{\varepsilon}) := \begin{cases} \varepsilon^n & : n < \ell, \\ \varepsilon^{\frac{d}{2}} |\log \varepsilon|^{\frac{1}{2}} & : n = \ell \text{ and } d \text{ even,} \\ \varepsilon^{\frac{d}{2}} & : n = \ell \text{ and } d \text{ odd.} \end{cases} \tag{1.6}$$

- (iii) Local covariance structure: Further assume that  $\int_{\mathbb{R}^d} |y|^\ell |c(y)| dy \leq 1$ . There exists an  $(m + 4)$ th-order tensor  $\mathcal{Q}_{l'k'lk}^m$  for all  $l', k', l, k \geq 0$  with  $l' + k' + l + k = m \leq \ell - 1$ , such that for all  $g, \nabla \bar{w} \in C_c^\infty(\mathbb{R}^d)^d$  and  $1 \leq n \leq \ell$ , setting

$$\mathcal{Q}_\varepsilon^n [g, \nabla \bar{w}] := \sum_{m=0}^{n-1} \varepsilon^m \sum_{\substack{l', k', l, k \geq 0 \\ l' + k' + l + k = m}} \mathcal{Q}_{l'k'lk}^m \odot \int_{\mathbb{R}^d} \nabla^{l'} g \otimes \nabla^{k'} \nabla \bar{w} \otimes \nabla^l g \otimes \nabla^k \nabla \bar{w},$$

where  $\odot$  denotes the complete contraction of tensors of the same order, there holds

$$\left| \text{Var} \left[ \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \Xi_\varepsilon^{\circ, n} [\nabla \bar{w}] \right] - \mathcal{Q}_\varepsilon^n [g, \nabla \bar{w}] \right| \lesssim_{g, \bar{w}} \varepsilon^n \mu_{d, n} \left( \frac{1}{\varepsilon} \right).$$

(iv) Normal approximation: For all  $g, \nabla \bar{w} \in C_c^\infty(\mathbb{R}^d)^d$  and  $1 \leq n \leq \ell$ ,

$$d_{\mathcal{N}} \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \Xi_\varepsilon^{\circ, n} [\nabla \bar{w}] \right) \lesssim_{g, \bar{w}} \frac{\varepsilon^{\frac{d}{2}} |\log \varepsilon| (|\log \varepsilon| \log^2 |\log \varepsilon|)^{\frac{1}{2d} n(n-1)}}{\text{Var} \left[ \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \Xi_\varepsilon^{\circ, n} [\nabla \bar{w}] \right]},$$

where we have set

$$d_{\mathcal{N}}(X) := W_2 \left( \frac{X - \mathbb{E}[X]}{\text{Var}[X]^{\frac{1}{2}}}; \mathcal{N} \right) + d_{\text{TV}} \left( \frac{X - \mathbb{E}[X]}{\text{Var}[X]^{\frac{1}{2}}}; \mathcal{N} \right),$$

where  $W_2(\cdot; \mathcal{N})$  and  $d_{\text{TV}}(\cdot; \mathcal{N})$  denote the 2-Wasserstein and the total variation distances to a standard Gaussian law, respectively.

In the estimates above, the notation  $\lesssim_\gamma$  (with parameters  $\gamma$ ) stands for  $\leq$  up to a multiplicative constant  $C_\gamma \geq 1$  that only depends on  $d, \lambda, \|a_0\|_{W^{2, \infty}}$ , and on the parameters  $\gamma$ , through an upper bound on suitable (weighted) Sobolev norms of  $\gamma$  if  $\gamma$  is a function.  $\diamond$

Choosing  $n = \ell$ , we are thus led to an intrinsic description of the fluctuations of macroscopic observables  $\int_{\mathbb{R}^d} g \cdot (\nabla u_\varepsilon - \mathbb{E}[\nabla u_\varepsilon])$  with accuracy  $O(\varepsilon^d)$  (with a logarithmic correction in even dimensions), that is, up to the square of the CLT order. In particular, non-Gaussian corrections are only expected beyond that order.

**Corollary 1** *With the assumptions and notation of Theorem 1, for all  $f, g \in C_c^\infty(\mathbb{R}^d)^d$ , in terms of the intrinsic quantity*

$$X_\varepsilon^\circ(g, f) := \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^\ell \cdot (\Xi_\varepsilon^{\circ, \ell} [\nabla \bar{u}_\varepsilon^\ell] - \mathbb{E}[\Xi_\varepsilon^{\circ, \ell} [\nabla \bar{u}_\varepsilon^\ell]]),$$

there holds for all  $p < \infty$ ,

$$\mathbb{E} \left[ \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot (\nabla u_\varepsilon - \mathbb{E}[\nabla u_\varepsilon]) - X_\varepsilon^\circ(g, f) \right)^p \right]^{\frac{1}{p}} \lesssim_{p, f, g} \varepsilon^{\frac{d}{2}} \times \begin{cases} |\log \varepsilon|^{\frac{1}{2}} & : d \text{ even,} \\ 1 & : d \text{ odd,} \end{cases}$$

and, under the non-degeneracy condition  $\mathcal{Q}_{0000} \odot \int_{\mathbb{R}^d} (\nabla \bar{v} \otimes \nabla \bar{u})^{\otimes 2} \neq 0$ ,

$$d \left( X_\varepsilon^\circ(g, f); \mathcal{Q}_\varepsilon^\ell [\nabla \bar{v}_\varepsilon^\ell, \nabla \bar{u}_\varepsilon^\ell]^{\frac{1}{2}} \mathcal{N} \right) \lesssim_{f, g} \varepsilon^{\frac{d}{2}} |\log \varepsilon|^{1+\rho} (\log |\log \varepsilon|)^{2\rho}, \tag{1.7}$$

where  $d(\cdot; \cdot)$  stands for the sum of the 2-Wasserstein and the total variation distances, where  $\mathcal{N}$  denotes a standard Gaussian random variable, and where

$$\rho := \frac{1}{8d} \times \begin{cases} d(d-2) & : d \text{ even,} \\ (d+1)(d-1) & : d \text{ odd.} \end{cases} \quad \diamond$$

Another question that one may formulate concerns the description of fluctuations of higher-order correctors—although this is somewhat artificial since only combinations of correctors in form of two-scale expansions are intrinsically relevant. As briefly described in Remark 3.6, the proof of the above main result can be adapted in a straightforward way to also describe the fluctuations of  $\varepsilon^{n-\frac{d}{2}}\varphi^n(\frac{\cdot}{\varepsilon})$  and  $\varepsilon^{n-\frac{d}{2}}\nabla\varphi^{n+1}(\frac{\cdot}{\varepsilon})$  with accuracy  $O(\varepsilon^{\frac{d}{2}})$  for all  $n \leq \ell - 1$  (with a logarithmic correction in even dimensions).

## Outline of the article

In Sect. 2, we start by recalling the theory of oscillations in stochastic homogenization in terms of higher-order correctors and two-scale expansions. Section 3 is devoted to the definition and elementary properties of higher-order homogenization commutators, including the (easy) proof of item (i) of Theorem 1 as well as the statement of quantitative higher-order Hill–Mandel relations. Next, we turn to the rest of the proof of Theorem 1, which is based on the following three main ingredients.

- *Malliavin calculus*: In the present Gaussian setting, given a random variable  $X$ , a Poincaré inequality holds in the form  $\text{Var}[X] \leq C \mathbb{E}[\|DX\|_{\mathfrak{H}}^2]$  (cf. [15,29]), where the Malliavin derivative  $DX$  encodes the infinitesimal variation of  $X$  with respect to changes in the underlying Gaussian field  $G$  and where  $\|\cdot\|_{\mathfrak{H}}$  is the Hilbert norm associated with the covariance structure of  $G$ . Similarly, the approximate normality of  $X$  can be estimated by the size of  $D^2X$  with help of a so-called second-order Poincaré inequality [14,43]. With these functional analytic tools at hand, we are thus reduced to estimating infinitesimal variations of the quantities of interest with respect to  $G$ , thus somehow linearizing the dependence on  $G$ , which is then particularly amenable to PDE methods. On a more technical level, rather than using ad hoc derivatives with respect to the coefficient field as in previous works in stochastic homogenization, we use here the full power of the Gaussian Malliavin calculus. Relevant definitions and results are recalled in Sect. 4.
- *Representation formulas*: In view of applying the above tools from Malliavin calculus, we consider the infinitesimal variations of the quantities of interest, for which we establish suitable representation formulas. This requires a good understanding of the algebra behind higher-order correctors and commutators, and is performed in Sect. 5.
- *PDE ingredients*: It remains to estimate the infinitesimal variations of interest from their representation formulas. Based on the large-scale Lipschitz regularity theory for  $\alpha$ -harmonic functions as developed in [5–7,21], we introduce in Sect. 6 a new annealed Calderón–Zygmund theory for linear elliptic equations in divergence form with random coefficients. This constitutes an upgrade of the quenched large-scale Calderón–Zygmund theory of [3,4,21] (see also [31] for a direct approach);



it happens to be particularly well suited for our purposes in this article and is of independent interest.

Finally, with these three ingredients at hand, we turn to the proof of items (ii), (iii), and (iv) of Theorem 1 in Sects. 7–9, respectively.

**Notation**

- We denote by  $C \geq 1$  any constant that only depends on  $d, \lambda$ , and  $\|a_0\|_{W^{2,\infty}}$ . We use the notation  $\lesssim$  (resp.  $\gtrsim$ ) for  $\leq C \times$  (resp.  $\geq \frac{1}{C} \times$ ) up to such a multiplicative constant  $C$ . We write  $\simeq$  when both  $\lesssim$  and  $\gtrsim$  hold. We add subscripts to  $C, \lesssim, \gtrsim, \simeq$  in order to indicate dependence on other parameters. If the subscript is a function, then it is understood as dependence on an upper bound on suitable (weighted) Sobolev norms of the function.
- The ball centered at  $x$  of radius  $r$  in  $\mathbb{R}^d$  is denoted by  $B_r(x)$ , and we simply write  $B(x)$  if  $r = 1$ .
- For a function  $f$  and  $1 \leq p < \infty$ , we write  $[f]_p(x) := (f_{B(x)} |f|^p)^{1/p}$  for the local moving  $L^p$  average, and similarly  $[f]_\infty(x) := \sup_{B(x)} |f|$ .
- For a  $k$ th-order tensor  $T$  we define for any fixed indices  $i_1, \dots, i_k$ ,

$$\text{Sym}_{i_1 \dots i_k} T_{i_1 \dots i_k} := \frac{1}{k!} \sum_{\sigma \in \mathcal{S}_k} T_{i_{\sigma(1)} \dots i_{\sigma(k)}}$$

with  $\mathcal{S}_k$  denoting the set of permutations of the set  $\{1, \dots, k\}$ . Although this notation may seem redundant at this point, it will prove very useful in the sequel.

- For  $r \in \mathbb{R}$  we denote by  $\lceil r \rceil$  the smallest integer  $\geq r$  and by  $\lfloor r \rfloor$  the largest integer  $\leq r$ . For  $r, s \in \mathbb{R}$  we write  $r \vee s := \max\{r, s\}$  and  $r \wedge s := \min\{r, s\}$ . We set  $\langle x \rangle := (1 + |x|^2)^{1/2}$  for  $x \in \mathbb{R}^d$ , and  $\langle \nabla \rangle$  denotes the corresponding pseudo-differential operator.

**2 Higher-order theory of oscillations**

In this section, as a prelude to the study of higher-order fluctuations, we describe the classical higher-order theory of oscillations in stochastic homogenization in terms of two-scale expansions, including some new results and shorter arguments. We start by recalling the definition of higher-order correctors, homogenized coefficients, fluxes, and flux correctors as motivated by formal two-scale expansions (e.g. [11,30]). A simple iterative way to view this definition is as follows: for  $n \geq 1$ , assuming that the previous correctors  $\varphi^1, \dots, \varphi^{n-1}$  are well-defined as centered stationary objects, the next corrector gradient  $\nabla \varphi^n$  is the unique centered stationary object that is characterized by the property that for all  $n$ th-order polynomials  $\bar{q}$  the corrected polynomial  $F^n[\bar{q}] := \bar{q} + \sum_{k=1}^n \varphi_{i_1 \dots i_k}^k \nabla_{i_1 \dots i_k}^k \bar{q}$  captures oscillations of the heterogeneous operator in the sense that  $\nabla \cdot \mathbf{a} \nabla F^n[\bar{q}]$  is deterministic. In that case, there must hold [cf. (2.9)]

$$\nabla \cdot \mathbf{a} \nabla F^n[\bar{q}] = \nabla \cdot \left( \sum_{k=1}^{n-1} \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \right) \nabla \bar{q},$$

which in addition characterizes the higher-order homogenized coefficients  $\bar{a}^2, \dots, \bar{a}^{n-1}$  next to the first-order one  $\bar{a}^1 := \bar{a}$ . In other words, the higher-order corrector expansion  $F^n[\cdot]$  precisely intertwines the heterogeneous elliptic operator with its higher-order homogenized version. Note that the above only characterizes the symmetric part  $\text{Sym}_{i_1 \dots i_n} \varphi_{i_1 \dots i_n}$  of correctors, which is indeed the only relevant quantity in view of two-scale expansions. It is however convenient to define correctors with a suitable non-symmetric part in such a way that the corresponding fluxes can be written in terms of skew-symmetric flux correctors, which is a well-known crucial tool in view of error estimates (e.g. [30, p. 27]). As stated below, in contrast with the periodic case, only  $\ell := \lceil \frac{d}{2} \rceil$  correctors can be defined, with the  $\ell - 1$  first of them being stationary; this number would be further reduced in the case of coefficient fields with stronger correlations (e.g. [9,20]).

**Definition 2.1** (*Higher-order correctors*) Given  $\ell := \lceil \frac{d}{2} \rceil$ , we inductively define the correctors  $(\varphi^n)_{0 \leq n \leq \ell}$ , the homogenized coefficients  $(\bar{a}^n)_{1 \leq n \leq \ell}$ , the fluxes  $(q^n)_{1 \leq n \leq \ell}$ , and the flux correctors  $(\sigma^n)_{0 \leq n \leq \ell}$  as follows:

- $\varphi^0 := 1$  and for all  $1 \leq n \leq \ell$  we define  $\varphi^n := (\varphi^n_{i_1 \dots i_n})_{1 \leq i_1, \dots, i_n \leq d}$  with  $\varphi^n_{i_1 \dots i_n}$  a scalar field that satisfies

$$\begin{aligned}
 -\nabla \cdot \mathbf{a} \nabla \varphi^n_{i_1 \dots i_n} &= \nabla \cdot ((\mathbf{a} \varphi^n_{i_1 \dots i_{n-1}} - \sigma^{n-1}_{i_1 \dots i_{n-1}}) e_{i_n}), \\
 \nabla \varphi^n_{i_1 \dots i_n} &\text{ stationary, } \mathbb{E} [|\nabla \varphi^n_{i_1 \dots i_n}|^2] < \infty, \quad \mathbb{E} [\nabla \varphi^n_{i_1 \dots i_n}] = 0.
 \end{aligned}$$

For  $n < \ell$  we can choose  $\varphi^n$  itself stationary with  $\mathbb{E} [\varphi^n] = 0$ , while for  $n = \ell$  we choose the anchoring  $\varphi^\ell(0) = 0$ .

- For all  $1 \leq n \leq \ell$  we define  $\bar{a}^n := (\bar{a}^n_{i_1 \dots i_{n-1}})_{1 \leq i_1, \dots, i_{n-1} \leq d}$  with  $\bar{a}^n_{i_1 \dots i_{n-1}}$  the matrix given by

$$\bar{a}^n_{i_1 \dots i_{n-1}} e_{i_n} := \mathbb{E} [\mathbf{a} (\nabla \varphi^n_{i_1 \dots i_n} + \varphi^n_{i_1 \dots i_{n-1}} e_{i_n})]. \tag{2.1}$$

- For all  $1 \leq n \leq \ell$  we define  $q^n := (q^n_{i_1 \dots i_n})_{1 \leq i_1, \dots, i_n \leq d}$  with  $q^n_{i_1 \dots i_n}$  the vector field given by

$$q^n_{i_1 \dots i_n} := \mathbf{a} \nabla \varphi^n_{i_1 \dots i_n} + (\mathbf{a} \varphi^n_{i_1 \dots i_{n-1}} - \sigma^{n-1}_{i_1 \dots i_{n-1}}) e_{i_n} - \bar{a}^n_{i_1 \dots i_{n-1}} e_{i_n},$$

where the definition of  $\bar{a}^n$  and the normalization of  $\sigma^{n-1}$  below ensure  $\mathbb{E} [q^n] = 0$ .

- $\sigma^0 := 0$  and for all  $1 \leq n \leq \ell$  we define  $\sigma^n := (\sigma^n_{i_1 \dots i_n})_{1 \leq i_1, \dots, i_n \leq d}$  with  $\sigma^n_{i_1 \dots i_n}$  a skew-symmetric matrix field that satisfies

$$\begin{aligned}
 -\Delta \sigma^n_{i_1 \dots i_n} &= \nabla \times q^n_{i_1 \dots i_n}, \quad \nabla \cdot \sigma^n_{i_1 \dots i_n} = q^n_{i_1 \dots i_n}, \\
 \nabla \sigma^n_{i_1 \dots i_n} &\text{ stationary, } \mathbb{E} [|\nabla \sigma^n_{i_1 \dots i_n}|^2] < \infty, \quad \mathbb{E} [\nabla \sigma^n_{i_1 \dots i_n}] = 0,
 \end{aligned}$$

with the notation  $(\nabla \times X)_{ij} := \nabla_i X_j - \nabla_j X_i$  for a vector field  $X$  and with the notation  $(\nabla \cdot Y)_i := \nabla_j Y_{ij}$  for a matrix field  $Y$ . For  $n < \ell$  we can choose  $\sigma^n$  itself stationary with  $\mathbb{E} [\sigma^n] = 0$ , while for  $n = \ell$  we choose the anchoring  $\sigma^\ell(0) = 0$ .

◇

We first state that the above definition of higher-order correctors indeed makes sense. This easily follows from [9, proof of Proposition 9 and Lemma 12] together with Lemma 7.1 below in the considered Gaussian setting (see also [3,20,21]), and the proof is omitted. Note that the optimal stochastic integrability (that is, the optimal constants  $c_n$ 's below) is not required in our analysis, hence is not addressed here.

**Proposition 2.2** (Higher-order correctors [3,9,20,21]) *All the quantities in Definition 2.1 exist and are uniquely defined. In addition, there exists a sequence  $(c_n)_{0 \leq n \leq \ell}$  of positive numbers with  $c_0 = \frac{1}{2}$  such that for all  $0 \leq n \leq \ell$  and  $1 \leq p < \infty$ ,*

$$|\bar{a}^n| \lesssim 1, \quad \mathbb{E} \left[ \left[ \nabla \varphi^n \right]_2^p \right]^{\frac{1}{p}} \lesssim p^{c_{n-1}}, \quad \mathbb{E} \left[ \left[ \varphi^n \right]_2^p(x) \right]^{\frac{1}{p}} + \mathbb{E} \left[ \left[ \sigma^n \right]_2^p(x) \right]^{\frac{1}{p}} \lesssim p^{c_n} \mu_{d,n}(x),$$

where in line with (1.6) we have set

$$\mu_{d,n}(x) := \begin{cases} 1 & : n < \ell, \\ \log^{\frac{1}{2}}(2 + |x|) & : n = \ell \text{ and } d \text{ even}, \\ 1 + |x|^{\frac{1}{2}} & : n = \ell \text{ and } d \text{ odd}. \end{cases} \quad \diamond$$

For later reference we also state the following result on the fluctuation scaling of correctors, showing that higher-order correctors have stronger correlations hence worse fluctuation scalings, as first emphasized in [24].

**Lemma 2.3** (Fluctuation scaling of higher-order correctors) *For all  $g \in C_c^\infty(\mathbb{R}^d)$ ,  $0 \leq n \leq \ell - 1$ , and  $1 \leq p < \infty$ ,*

$$\mathbb{E} \left[ \left| \int_{\mathbb{R}^d} g(x) (\nabla \varphi^{n+1}, \varphi^n, \sigma^n) \left( \frac{x}{\varepsilon} \right) dx \right|^p \right]^{\frac{1}{p}} \lesssim_p \varepsilon^{\frac{d}{2}-n} \|g\|_{L^{\frac{2d}{d+2n}}(\mathbb{R}^d)}. \quad \diamond$$

If  $\mathbf{a}$  is replaced by the pointwise transpose field  $\mathbf{a}^*$ , we write  $(\varphi^{*,n})_{0 \leq n \leq \ell}$ ,  $(\bar{\mathbf{a}}^{*,n})_{1 \leq n \leq \ell}$ ,  $(q^{*,n})_{1 \leq n \leq \ell}$ , and  $(\sigma^{*,n})_{0 \leq n \leq \ell}$  for the corresponding objects. While it is well-known that  $\bar{\mathbf{a}}^{*,1} = (\bar{\mathbf{a}}^1)^*$ , the following lemma extends this relation to higher order.

**Lemma 2.4** (Symmetries of homogenized coefficients) *For all  $1 \leq n \leq \ell$ ,*<sup>2</sup>

$$\text{Sym}_{i_1 \dots i_n} (e_j \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n}) = (-1)^{n+1} \text{Sym}_{i_1 \dots i_n} (e_{i_n} \cdot \bar{\mathbf{a}}_{j i_1 \dots i_{n-2}}^{*,n} e_{i_{n-1}}). \quad (2.2)$$

In particular, if  $\mathbf{a}$  is symmetric and if  $n$  is even, there holds

$$\text{Sym}_{i_1 \dots i_{n+1}} (e_{i_{n+1}} \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n}) = 0. \quad \diamond$$

**Remark 2.5** The proof of Lemma 2.4 proceeds by letting the successive inverse operators  $(\nabla \cdot \mathbf{a} \nabla)^{-1} \nabla \cdot$  defining  $(\nabla \varphi^n, \varphi^{n-1})$  “migrate” to the test function  $\mathbf{a}$  in the

<sup>2</sup> For  $n = 1$  the notation  $\bar{\mathbf{a}}_{j i_1 \dots i_{n-2}}^{*,n} e_{i_{n-1}}$  stands for  $\bar{\mathbf{a}}^{*,1} e_j$ , and we use a similar unifying notation throughout in the sequel.

definition (2.1) of  $\bar{a}^n$ . Using the algebra of correctors, a full migration precisely leads to the dual expression  $\bar{a}^{*,n}$  and proves the relation (2.2). Rather stopping this migration argument at an intermediate step, we obtain for all  $1 \leq m < n \leq \ell$ ,

$$\text{Sym}_{i_1 \dots i_n} (e_j \cdot \bar{a}^n_{i_1 \dots i_{n-1}} e_{i_n}) = (-1)^{m+1} \text{Sym}_{i_1 \dots i_n} \mathbb{E} \left[ \nabla \varphi_{j i_n \dots i_{n-m+1}}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-m}}^{n-m} - \varphi_{j i_n \dots i_{n-m+2}}^{*,m} e_{i_{n-m+1}} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-m-1}}^{n-m-1} e_{i_{n-m}} \right].$$

Note that this turns into (2.2) for  $m = n - 1$ . We deduce for  $n = 2m + 1$  odd,

$$\text{Sym}_{i_1 \dots i_n} (e_j \cdot \bar{a}^n_{i_1 \dots i_{n-1}} e_{i_n}) = (-1)^{m+1} \text{Sym}_{i_1 \dots i_n} \mathbb{E} \left[ \nabla \varphi_{j i_n \dots i_{m+2}}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{m+1}}^{m+1} - \varphi_{j i_n \dots i_{m+3}}^{*,m} e_{i_{m+2}} \cdot \mathbf{a} \varphi_{i_1 \dots i_m}^m e_{i_{m+1}} \right],$$

and for  $n = 2m$  even, further using the algebra of correctors to symmetrize the expression,

$$\begin{aligned} & \text{Sym}_{i_1 \dots i_n} (e_j \cdot \bar{a}^n_{i_1 \dots i_{n-1}} e_{i_n}) \\ &= (-1)^{m+1} \text{Sym}_{i_1 \dots i_n} \mathbb{E} \left[ (\nabla \varphi_{j i_n \dots i_{m+2}}^{*,m} + \varphi_{j i_n \dots i_{m+3}}^{*,m-1} e_{i_{m+2}}) \cdot \mathbf{a} \varphi_{i_1 \dots i_m}^m e_{i_{m+1}} - \varphi_{j i_n \dots i_{m+2}}^{*,m} e_{i_{m+1}} \cdot \mathbf{a} (\nabla \varphi_{i_1 \dots i_m}^m + \varphi_{i_1 \dots i_{m-1}}^{m-1} e_{i_m}) \right]. \end{aligned}$$

This shows that  $\bar{a}^n$  can be defined in terms of correctors  $(\nabla \varphi^{s+1}, \varphi^s)$  with  $s \leq \lfloor \frac{n}{2} \rfloor$  only. This simple observation was independently made in [1, Theorem 3.5] (see also [47, Lemma 5.2.6]) and has the following striking consequence: Choosing a periodicization in law  $\mathbf{a}_L$  of the coefficient field  $\mathbf{a}$  and recalling that the periodic Poincaré inequality ensures that all the correctors  $\varphi_L^n$  are defined as stationary objects so that all the corresponding homogenized coefficients  $\bar{a}_L^n$  are well-defined, the above computation entails that  $\text{Sym}_{i_1 \dots i_n} (\bar{a}_{L:i_1 \dots i_{n-1}}^n e_{i_n})$  has a limit as  $L \uparrow \infty$  for all  $n \leq d$  if  $d$  is odd and for all  $n < d$  if  $d$  is even. This partially solves a weak version of the Bourgain–Spencer conjecture [12,33] as stated in [17, Section 4.2]. Refinements in this direction are postponed to a future work.  $\diamond$

We turn to the accuracy of two-scale expansions in the homogenization regime. Given  $f \in L^2(\mathbb{R}^d)^d$ , we consider the unique Lax–Milgram solution  $\nabla u_\varepsilon$  of the rescaled elliptic PDE (1.2),

$$-\nabla \cdot \mathbf{a}(\frac{\cdot}{\varepsilon}) \nabla u_\varepsilon = \nabla \cdot f \quad \text{in } \mathbb{R}^d.$$

Standard two-scale expansion techniques [11] formally suggest

$$u_\varepsilon = \sum_{k=0}^n \varepsilon^k \varphi_{i_1 \dots i_k}^k(\frac{\cdot}{\varepsilon}) \nabla_{i_1 \dots i_k}^k \bar{U}_\varepsilon^n + O(\varepsilon^{n+1}),$$

where  $\bar{U}_\varepsilon^n$  satisfies the  $n$ th-order homogenized equation

$$-\nabla \cdot \left( \sum_{k=1}^n \varepsilon^{k-1} \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \right) \nabla \bar{U}_\varepsilon^n = \nabla \cdot f \quad \text{in } \mathbb{R}^d. \tag{2.3}$$

The effective field-flux constitutive relation  $\nabla \bar{w} \mapsto \bar{\mathbf{a}} \nabla \bar{w}$  is thus refined into the higher-order relation

$$\nabla \bar{w} \mapsto \sum_{k=1}^n \varepsilon^{k-1} \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{w}, \tag{2.4}$$

which includes dispersive corrections. However, Eq. (2.3) is ill-posed in general for  $n$  even and a suitable proxy needs to be devised (cf. also [32]). We start by introducing a notation for two-scale expansions.

**Definition 2.6** (*Higher-order two-scale expansions*) For  $0 \leq n \leq \ell$ , given a deterministic smooth function  $\bar{w}$ , its  $n$ th-order two-scale expansion  $F_\varepsilon^n[\bar{w}]$  is defined as

$$F_\varepsilon^n[\bar{w}] := \sum_{k=0}^n \varepsilon^k \varphi_{i_1 \dots i_k}^k \left(\frac{\cdot}{\varepsilon}\right) \nabla_{i_1 \dots i_k}^k \bar{w}.$$

At the level of gradients, we similarly define for  $0 \leq n \leq \ell$ ,

$$E_\varepsilon^n[\nabla \bar{w}] := \sum_{k=0}^{n-1} \varepsilon^k \left( \nabla \varphi_{i_1 \dots i_{k+1}}^{k+1} + \varphi_{i_1 \dots i_k}^k e_{i_{k+1}} \right) \left(\frac{\cdot}{\varepsilon}\right) \nabla_{i_1 \dots i_{k+1}}^{k+1} \bar{w}. \quad \diamond$$

In particular, note that the quantities  $\nabla F_\varepsilon^n[\bar{w}]$  and  $E_\varepsilon^n[\nabla \bar{w}]$  are related as follows: for all  $0 \leq n \leq \ell$ ,

$$\nabla F_\varepsilon^n[\bar{w}] = E_\varepsilon^n[\nabla \bar{w}] + \varepsilon^n \varphi_{i_1 \dots i_n}^n \nabla_{i_1 \dots i_n}^n \bar{w}, \tag{2.5}$$

where all coefficients of  $E_\varepsilon^n[\nabla \bar{w}]$  are stationary. We turn to the definition of a suitable proxy for the ill-posed higher-order homogenized equation (2.3). For later reference we simultaneously consider the dual equations.

**Definition 2.7** (*Higher-order homogenized equations*) Given  $f, g \in C_c^\infty(\mathbb{R}^d)^d$ , let  $\nabla u_\varepsilon$  and  $\nabla v_\varepsilon$  be the unique Lax–Milgram solutions of the rescaled elliptic PDEs

$$-\nabla \cdot \mathbf{a} \left(\frac{\cdot}{\varepsilon}\right) \nabla u_\varepsilon = \nabla \cdot f, \quad -\nabla \cdot \mathbf{a}^* \left(\frac{\cdot}{\varepsilon}\right) \nabla v_\varepsilon = \nabla \cdot g \quad \text{in } \mathbb{R}^d. \tag{2.6}$$

For  $1 \leq n \leq \ell$ , we define the corresponding  $n$ th-order homogenized solutions

$$\bar{u}_\varepsilon^n := \sum_{k=1}^n \varepsilon^{k-1} \bar{u}^k, \quad \bar{v}_\varepsilon^n := \sum_{k=1}^n \varepsilon^{k-1} \bar{v}^k,$$

where  $\nabla \tilde{u}^1 = \nabla \bar{u}$  and  $\nabla \tilde{v}^1 = \nabla \bar{v}$  are the unique Lax–Milgram solutions in  $\mathbb{R}^d$  of the first-order homogenized equations,

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = \nabla \cdot f, \quad -\nabla \cdot \bar{\mathbf{a}}^* \nabla \bar{v} = \nabla \cdot g,$$

and where for  $2 \leq n \leq \ell$  we inductively define the  $n$ th-order corrections  $\nabla \tilde{u}^n$  and  $\nabla \tilde{v}^n$  as the unique Lax–Milgram solutions in  $\mathbb{R}^d$  of

$$\begin{aligned} -\nabla \cdot \bar{\mathbf{a}} \nabla \tilde{u}^n &= \nabla \cdot \sum_{k=2}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \tilde{u}^{n+1-k}, \\ -\nabla \cdot \bar{\mathbf{a}}^* \nabla \tilde{v}^n &= \nabla \cdot \sum_{k=2}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \tilde{v}^{n+1-k}. \end{aligned} \quad \diamond$$

With the above notation, the main result from the higher-order corrector theory in stochastic homogenization takes the following guise. For  $q = 2$ , it is a straightforward consequence of the definition of higher-order correctors and of the corrector estimates of Proposition 2.2. It extends [3,9,20,22] to higher order and completes [24, Theorem 1.6] up to the optimal order with finer norms. For  $n = \ell$ , the obtained maximal accuracy is  $O(\varepsilon^{d/2})$  (with a logarithmic correction in even dimensions), indicating that the study of fluctuations goes beyond the theory of oscillations.

**Proposition 2.8** (Accuracy of higher-order two-scale expansions) *For all  $f \in C_c^\infty(\mathbb{R}^d)^d$ ,  $1 \leq n \leq \ell$ , and  $1 < q \leq p < \infty$ ,*

$$\mathbb{E} \left[ \left\| \left[ \nabla(u_\varepsilon - F_\varepsilon^n[\tilde{u}_\varepsilon^n]) \right]_2 \right\|_{L^q(\mathbb{R}^d)}^p \right]^{\frac{1}{p}} \lesssim_{q,p} \varepsilon^n \mu_{d,n}(\tfrac{1}{\varepsilon}) \left\| \mu_{d,n}[\langle \nabla \rangle^{2n-1} f] \right\|_{L^q(\mathbb{R}^d)}. \tag{2.7}$$

$\diamond$

### 2.1 Proof of Lemma 2.3

The result is easily obtained based on Malliavin calculus and on the annealed Calderón–Zygmund theory of Sect. 6. More precisely, for  $0 \leq n \leq \ell - 1$ , we start from Proposition 4.1(iii) together with (4.2) in the form

$$\begin{aligned} &\mathbb{E} \left[ \left| \int_{\mathbb{R}^d} g(\nabla \varphi^{n+1}, \varphi^n, \sigma^n) \right|^p \right]^{\frac{1}{p}} \\ &\lesssim_p \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} g(\nabla D\varphi^{n+1}, D\varphi^n, D\sigma^n) \right]_1^2 \right)^{\frac{p}{2}} \right]^{\frac{1}{p}}, \end{aligned}$$

and the conclusion follows after  $\varepsilon$ -rescaling from the estimates of Lemma 7.1 on the Malliavin derivatives of higher-order correctors.  $\square$

### 2.2 Proof of Lemma 2.4

As explained in Remark 2.5, the argument consists in letting the successive inverse operators  $(-\nabla \cdot \mathbf{a} \nabla)^{-1} \nabla \cdot$  defining the correctors  $(\nabla \varphi^n, \varphi^{n-1})$  migrate to the test function  $\mathbf{a}$  in the definition of the homogenized coefficient

$$\bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n} = \mathbb{E} \left[ \mathbf{a} (\nabla \varphi_{i_1 \dots i_n}^n + \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n}) \right].$$

Taking into account the particular algebraic structure of the correctors, a full migration precisely leads to the dual expression  $\bar{\mathbf{a}}^{*,n}$  up to additional terms involving higher-order flux correctors, which disappear when taking the symmetric part. We split the proof into two steps.

*Step 1.* Migration process: proof that for all  $1 \leq m < n \leq \ell$ ,

$$\begin{aligned} & \mathbb{E} \left[ \nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_n}^n - \varphi_{j_1 \dots j_{m-1}}^{*,m-1} e_{j_m} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} \right] \\ &= -\mathbb{E} \left[ \nabla \varphi_{j_1 \dots j_m i_n}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} - \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-2}}^{n-2} e_{i_{n-1}} \right] \\ & \quad - \mathbb{E} \left[ e_{i_n} \cdot \sigma_{j_1 \dots j_{m-1}}^{*,m-1} e_{j_m} \varphi_{i_1 \dots i_{n-1}}^{n-1} + \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \sigma_{i_1 \dots i_{n-2}}^{n-2} e_{i_{n-1}} \right]. \end{aligned} \tag{2.8}$$

We abundantly use the properties of higher-order correctors (cf. Definition 2.1). The equation for  $\varphi^n$ , in conjunction with the stationarity of all involved objects, and the skew-symmetry of  $\sigma_{i_1 \dots i_{n-1}}^{n-1}$  yield

$$\begin{aligned} \mathbb{E} \left[ \nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_n}^n \right] &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot (\mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} - \sigma_{i_1 \dots i_{n-1}}^{n-1}) e_{i_n} \right] \\ &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} + \varphi_{j_1 \dots j_m}^{*,m} (\nabla \cdot \sigma_{i_1 \dots i_{n-1}}^{n-1}) \cdot e_{i_n} \right], \end{aligned}$$

hence, using the equation  $\nabla \cdot \sigma^{n-1} = q^{n-1}$  and the choice  $\mathbb{E} [\varphi^{*,m}] = 0$ ,

$$\begin{aligned} \mathbb{E} \left[ \nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_n}^n \right] &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} + \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} \right. \\ & \quad \left. + \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-2}}^{n-2} e_{i_{n-1}} - \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \sigma_{i_1 \dots i_{n-2}}^{n-2} e_{i_{n-1}} \right]. \end{aligned}$$

We focus on rewriting the second right-hand side term. Using the equation for  $\varphi^{*,m+1}$  and the skew-symmetry of  $\sigma_{j_1 \dots j_m}^{*,m}$ , we find

$$\begin{aligned} & \mathbb{E} \left[ \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} \right] \\ &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m i_n}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} - e_{i_n} \cdot \sigma_{j_1 \dots j_m}^{*,m} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} \right] \\ &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m i_n}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} + (\nabla \cdot \sigma_{j_1 \dots j_m}^{*,m}) \cdot \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} \right], \end{aligned}$$

hence, using the equation for  $\nabla \cdot \sigma^{*,m}$  and the choice  $\mathbb{E}[\varphi^{n-1}] = 0$ ,

$$\begin{aligned} & \mathbb{E} \left[ \varphi_{j_1 \dots j_m}^{*,m} e_{i_n} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} \right] \\ &= \mathbb{E} \left[ -\nabla \varphi_{j_1 \dots j_m i_n}^{*,m+1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_{n-1}}^{n-1} + \nabla \varphi_{j_1 \dots j_m}^{*,m} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} \right. \\ & \quad \left. + \varphi_{j_1 \dots j_{m-1}}^{*,m-1} e_{j_m} \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} - \sigma_{j_1 \dots j_{m-1}}^{*,m-1} e_{j_m} \cdot \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} \right], \end{aligned}$$

and the claim (2.8) follows.

*Step 2. Conclusion.*

For  $1 \leq n \leq \ell$ , the definition of  $\bar{\mathbf{a}}^n$  and the equation for  $\varphi^{*,1}$  yield

$$e_j \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n} = -\mathbb{E} \left[ \nabla \varphi_j^{*,1} \cdot \mathbf{a} \nabla \varphi_{i_1 \dots i_n}^n - e_j \cdot \mathbf{a} \varphi_{i_1 \dots i_{n-1}}^{n-1} e_{i_n} \right].$$

Iterating identity (2.8) then leads to

$$\begin{aligned} e_j \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n} &= (-1)^n \mathbb{E} \left[ \nabla \varphi_{j i_n \dots i_2}^{*,n} \cdot \mathbf{a} \nabla \varphi_{i_1}^1 - \varphi_{j i_n \dots i_3}^{*,n-1} e_{i_2} \cdot \mathbf{a} e_{i_1} \right] \\ &+ \sum_{l=0}^{n-2} (-1)^l \mathbb{E} \left[ e_{i_{n-l}} \cdot \sigma_{j i_n \dots i_{n-l+2}}^{*,l} e_{i_{n-l+1}} \varphi_{i_1 \dots i_{n-l-1}}^{n-l-1} + \varphi_{j i_n \dots i_{n-l+1}}^{*,l+1} e_{i_{n-l}} \cdot \sigma_{i_1 \dots i_{n-l-2}}^{n-l-2} e_{i_{n-l-1}} \right]. \end{aligned}$$

By skew-symmetry of  $\sigma_{j i_n \dots i_{n-l+2}}^{*,l}$  and  $\sigma_{i_1 \dots i_{n-l-2}}^{n-l-2}$  for all  $l$ , we deduce

$$\text{Sym}_{i_1 \dots i_n} e_j \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n} = (-1)^n \text{Sym}_{i_1 \dots i_n} \mathbb{E} \left[ \nabla \varphi_{j i_n \dots i_2}^{*,n} \cdot \mathbf{a} \nabla \varphi_{i_1}^1 - \varphi_{j i_n \dots i_3}^{*,n-1} e_{i_2} \cdot \mathbf{a} e_{i_1} \right],$$

hence, using the equation for  $\varphi^1$  and the definition of  $\bar{\mathbf{a}}^{*,n}$ ,

$$\begin{aligned} \text{Sym}_{i_1 \dots i_n} e_j \cdot \bar{\mathbf{a}}_{i_1 \dots i_{n-1}}^n e_{i_n} &= (-1)^{n+1} \text{Sym}_{i_1 \dots i_n} \mathbb{E} \left[ e_{i_1} \cdot \mathbf{a}^* \left( \nabla \varphi_{j i_n \dots i_2}^{*,n} + \varphi_{j i_n \dots i_3}^{*,n-1} e_{i_2} \right) \right] \\ &= (-1)^{n+1} \text{Sym}_{i_1 \dots i_n} e_{i_1} \cdot \bar{\mathbf{a}}_{j i_n \dots i_3}^{*,n} e_{i_2}, \end{aligned}$$

and the conclusion follows. □

### 2.3 Proof of Proposition 2.8

We focus on the case  $\varepsilon = 1$  and drop it from all subscripts in the notation, while the final result is obtained after  $\varepsilon$ -rescaling. We split the proof into two steps.

*Step 1.* Equation for  $F_n[\bar{w}]$ : proof that for all  $\bar{w} \in C_c^\infty(\mathbb{R}^d)$  and  $n \geq 0$ ,

$$\begin{aligned} \nabla \cdot \mathbf{a} \nabla F^n[\bar{w}] &= \nabla \cdot \left( \sum_{k=1}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{w} \right) \\ &+ \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right). \end{aligned} \tag{2.9}$$



We argue by induction. The claim is obvious for  $n = 0$ . Now, if it holds for some  $n \geq 0$ , we deduce

$$\begin{aligned} \nabla \cdot \mathbf{a} \nabla F^{n+1}[\bar{w}] &= \nabla \cdot \left( \sum_{k=1}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{w} \right) \\ &+ \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right) + \nabla \cdot \mathbf{a} \nabla \left( \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right). \end{aligned} \tag{2.10}$$

The definition of  $\sigma_{i_1 \dots i_{n+1}}^{n+1}$  (cf. Definition 2.1) yields

$$\begin{aligned} \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right) &= \nabla \cdot \left( (\nabla \cdot \sigma_{i_1 \dots i_{n+1}}^{n+1}) \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right) \\ &- \nabla \cdot \left( \mathbf{a} \nabla \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right) + \nabla \cdot \left( \bar{\mathbf{a}}_{i_1 \dots i_n}^{n+1} \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right). \end{aligned}$$

Hence, using the skew-symmetry of  $\sigma_{i_1 \dots i_{n+1}}^{n+1}$  and decomposing

$$\nabla \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} = \nabla \left( \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right) - \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w},$$

we obtain

$$\begin{aligned} \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right) &= \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_{n+1}}^{n+1} - \sigma_{i_1 \dots i_{n+1}}^{n+1}) \nabla \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right) \\ &- \nabla \cdot \mathbf{a} \nabla \left( \varphi_{i_1 \dots i_{n+1}}^{n+1} \nabla_{i_1 \dots i_{n+1}}^{n+1} \bar{w} \right) + \nabla \cdot \left( \bar{\mathbf{a}}_{i_1 \dots i_n}^{n+1} \nabla \nabla_{i_1 \dots i_n}^n \bar{w} \right). \end{aligned}$$

Injecting this into (2.10) leads to the claim (2.9) at level  $n + 1$ .

*Step 2. Conclusion.*

Let  $n \geq 1$  be fixed. Applying (2.9) with  $\bar{w} := \bar{u}^n$  and subtracting the Eq. (2.6) for  $u$ , we obtain

$$\begin{aligned} & - \nabla \cdot \mathbf{a} \nabla (u - F^n[\bar{u}^n]) \\ &= \nabla \cdot \left( f + \sum_{k=1}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{u}^n \right) + \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{u}^n \right). \end{aligned} \tag{2.11}$$

We now examine the equation satisfied by  $\bar{u}^n = \sum_{k=1}^n \tilde{u}^k$ : summing the defining equations for  $(\tilde{u}^k)_{1 \leq k \leq n}$  (cf. Definition 2.7), we find

$$- \nabla \cdot \bar{\mathbf{a}}^1 \nabla \bar{u}^n = \nabla \cdot f + \nabla \cdot \left( \sum_{k=2}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \sum_{l=k}^n \tilde{u}^{l+1-k} \right),$$

or equivalently,

$$- \nabla \cdot \left( \sum_{k=1}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \right) \nabla \bar{u}^n = \nabla \cdot f - \nabla \cdot \left( \sum_{k=2}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \sum_{l=n+2-k}^n \tilde{u}^l \right).$$

(This is to be compared with the formal ill-posed Eq. (2.3).) Inserting this into (2.11) yields

$$\begin{aligned}
 -\nabla \cdot \mathbf{a} \nabla (u - F^n[\bar{u}^n]) &= \nabla \cdot \left( \sum_{k=2}^n \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \sum_{l=n+2-k}^n \tilde{u}^l \right) \\
 &\quad + \nabla \cdot \left( (\mathbf{a} \varphi_{i_1 \dots i_n}^n - \sigma_{i_1 \dots i_n}^n) \nabla \nabla_{i_1 \dots i_n}^n \bar{u}^n \right). \tag{2.12}
 \end{aligned}$$

An energy estimate and the corrector estimates of Proposition 2.2 then imply for all  $p < \infty$ ,

$$\begin{aligned}
 \mathbb{E} \left[ \left\| \nabla (u - F^n[\bar{u}^n]) \right\|_{L^2(\mathbb{R}^d)}^p \right]^{\frac{1}{p}} &\lesssim_p \|\mu_{d,n} \nabla^{n+1} \bar{u}^n\|_{L^2(\mathbb{R}^d)} \\
 &\quad + \sum_{k=2}^n \sum_{l=n+2-k}^n \|\nabla^k \tilde{u}^l\|_{L^2(\mathbb{R}^d)} \\
 &\lesssim_p \sum_{k=n}^{2n-1} \|\mu_{d,n} \nabla^k f\|_{L^2(\mathbb{R}^d)},
 \end{aligned}$$

where the last estimate follows from the weighted Calderón–Zygmund theory for the constant-coefficient equations defining  $(\tilde{u}^k)_{1 \leq k \leq n}$ . The conclusion (2.7) in  $L^2(\mathbb{R}^d)$  follows after  $\varepsilon$ -rescaling. (Note that for the sake of shortness in the statement the norms of  $\nabla^k f$  for  $n \leq k \leq 2n - 1$  are replaced by the norm of  $\langle \nabla \rangle^{2n-1} f$ , which is no longer scale invariant.) The corresponding result in  $L^q(\mathbb{R}^d)$  for general  $1 < q < \infty$  requires to replace the energy estimate for (2.12) by some  $L^q$  theory, as provided e.g. by Theorem 6.1 below; details are omitted.  $\square$

### 3 Higher-order homogenization commutators

As discovered in [5,19], the homogenization commutator  $\Xi_\varepsilon^1[\nabla u_\varepsilon] := (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}}^1) \nabla u_\varepsilon$ , which takes the form of a commutator between large-scale averaging and the field-flux constitutive relation, plays a key role in the study of fluctuations in stochastic homogenization. Indeed, while the two-scale expansion of the solution  $\nabla u_\varepsilon$  is not accurate in the fluctuation scaling [25], the expansion of its commutator  $\Xi_\varepsilon^1[\nabla u_\varepsilon]$  is accurate [19]. The leading-order fluctuations are then governed by the standard homogenization commutator, that is, the commutator of the corrector field,  $\Xi_\varepsilon^{\circ,1}[\nabla \bar{u}^1] := (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}}^1) (\nabla \varphi_i^1(\frac{\cdot}{\varepsilon}) + e_i) \nabla_i \bar{u}^1$ . For higher-order fluctuations, a suitable higher-order correction of the homogenization commutator  $\Xi_\varepsilon^1[\nabla u_\varepsilon]$  is defined as follows.

**Definition 3.1** (*Higher-order homogenization commutators*) For  $0 \leq n \leq \ell$ , given a random function  $w$ , its  $n$ th-order homogenization commutator  $\Xi_\varepsilon^n[\nabla w]$  is the (distributional) random vector field

$$\Xi_\varepsilon^n[\nabla w] := \left( \mathbf{a}(\frac{\cdot}{\varepsilon}) - \sum_{k=1}^n \varepsilon^{k-1} \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \right) \nabla w. \quad \diamond$$

This definition is natural in view of the higher-order effective field-flux constitutive relation (2.4). Note that the symmetries of the homogenized coefficients (cf. Lemma 2.4) and an integration by parts lead to the following duality relation for all smooth and compactly supported random functions  $w, w'$ ,

$$\int_{\mathbb{R}^d} \nabla w' \cdot \Xi_\varepsilon^n[\nabla w] = \int_{\mathbb{R}^d} \Xi_\varepsilon^{*,n}[\nabla w'] \cdot \nabla w. \tag{3.1}$$

Similarly as for the first-order commutator [19, (1.8)], we state that the higher-order fluctuations of the field  $\nabla u_\varepsilon$  are determined by those of the higher-order commutators. While at first order we have the exact relation  $\int_{\mathbb{R}^d} g \cdot (\nabla u_\varepsilon - \nabla \bar{u}) = \int_{\mathbb{R}^d} \nabla \bar{v} \cdot \Xi_\varepsilon^1[\nabla u_\varepsilon]$ , this does no longer hold at higher orders due to the choice of a proxy for solutions of the generically ill-posed higher-order homogenized equations (2.3) (cf. (3.4) below). A corresponding formula can be deduced for the flux  $\mathbf{a}(\frac{\cdot}{\varepsilon})\nabla u_\varepsilon$ . This key principle corresponds to item (i) in Theorem 1.

**Proposition 3.2** (Reduction to commutators) *For all  $1 \leq n \leq \ell$  and  $p < \infty$ ,*

$$\begin{aligned} N_p \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \nabla u_\varepsilon - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^n \cdot \Xi_\varepsilon^n[\nabla u_\varepsilon] \right) \\ \lesssim_p \varepsilon^n \|\langle \nabla \rangle^{2(n-1)} g\|_{L^4(\mathbb{R}^d)} \|f\|_{L^4(\mathbb{R}^d)}. \end{aligned}$$

where we recall the notation  $N_p(X) = \mathbb{E} [|X - \mathbb{E}[X]|^p]^{\frac{1}{p}}$ . ◇

Next, we wish to emphasize that the above definition of the higher-order commutator  $\Xi_\varepsilon^n$  leads to a higher-order Hill–Mandel relation (e.g. [51] in the mechanics literature). First recall that the first-order commutator  $\Xi_\varepsilon^1[\nabla u_\varepsilon] := (\mathbf{a} - \bar{\mathbf{a}})(\frac{\cdot}{\varepsilon})\nabla u_\varepsilon$  was a natural quantity to consider in [5, 19] since the very definition of  $H$ -convergence [39] is equivalent to the weak convergence of  $\Xi_\varepsilon^1[\nabla u_\varepsilon]$  to 0, which is the mathematical formulation of the classical Hill–Mandel relation [27, 28]. Interestingly, this can be made quantitative: a direct computation for the solution of (2.6) yields

$$e_j \cdot \Xi_\varepsilon^1[\nabla u_\varepsilon] = -\varepsilon \nabla_k \left( (\mathbf{a}^* \varphi_j^{*,1} - \sigma_j^{*,1})(\frac{\cdot}{\varepsilon}) e_k \cdot \nabla u_\varepsilon + \varphi_j^{*,1}(\frac{\cdot}{\varepsilon}) f_k \right) + \nabla \varphi_j^{*,1}(\frac{\cdot}{\varepsilon}) \cdot f,$$

where the right-hand side is checked to be of order  $O(\varepsilon)$  in a weak norm if  $d > 2$ . This property extends to higher order as follows.

**Lemma 3.3** (Higher-order Hill–Mandel relation) *For all  $f \in C_c^\infty(\mathbb{R}^d)^d$  and  $1 \leq n \leq \ell$ ,*

$$\begin{aligned} e_j \cdot \Xi_\varepsilon^n[\nabla u_\varepsilon] &= (-\varepsilon)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n})(\frac{\cdot}{\varepsilon}) e_{i_n} \cdot \nabla u_\varepsilon + \varphi_{j i_1 \dots i_{n-1}}^{*,n}(\frac{\cdot}{\varepsilon}) f_{i_n} \right) \\ &+ \sum_{k=0}^{n-1} (-\varepsilon)^k \nabla_{i_1 \dots i_k}^k \left( (\nabla \varphi_{j i_1 \dots i_k}^{*,k+1} + \mathbb{1}_{k>0} \varphi_{j i_1 \dots i_{k-1}}^{*,k} e_{i_k})(\frac{\cdot}{\varepsilon}) \cdot f \right), \end{aligned} \tag{3.2}$$

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

$$N_p \left( \int_{\mathbb{R}^d} g \cdot \Xi_\varepsilon^n[\nabla u_\varepsilon] \right) \lesssim_p \varepsilon^n \mu_{d,n}(\tfrac{1}{\varepsilon}) \left( \|\mu_{d,n} \nabla^n g\|_{L^2(\mathbb{R}^d)} + \|g\|_{H^{\frac{d}{2}} \cap L^\infty(\mathbb{R}^d)} \right) \| [f]_{2d} \|_{L^2(\mathbb{R}^d)}.$$

◇

Finally, the relevant higher-order two-scale expansion of the higher-order commutator  $\Xi_\varepsilon^n[\nabla u_\varepsilon]$  takes the form  $\Xi_\varepsilon^{\circ,n}[\nabla \bar{u}_\varepsilon^n]$ , where the so-called standard commutator  $\Xi_\varepsilon^{\circ,n}[\nabla \cdot]$  is the  $n$ th-order differential operator with  $\varepsilon$ -rescaled (distributional) stationary random coefficients that is obtained by inserting the  $n$ th-order two-scale expansion  $F_\varepsilon^n[\cdot]$  into the commutator  $\Xi_\varepsilon^n[\nabla \cdot]$  and by truncating the obtained differential operator at order  $n$ . In other words, the standard commutator  $\Xi_\varepsilon^{\circ,n}[\nabla \cdot]$  is the  $n$ th-order linear differential operator characterized by  $\Xi_\varepsilon^{\circ,n}[\nabla \bar{q}] = \Xi_\varepsilon^n[\nabla F_\varepsilon^n[\bar{q}]]$  for all  $n$ th-order polynomials  $\bar{q}$ . This is efficiently expressed as follows.

**Definition 3.4** (Standard higher-order homogenization commutators) For  $0 \leq n \leq \ell$ , given a smooth deterministic function  $\bar{w}$ , we consider its  $n$ th-order Taylor polynomial with basepoint  $x$ ,

$$T_x^n \bar{w}(y) := \bar{w}(x) + \sum_{1 \leq |\alpha| \leq n} \frac{(y-x)^\alpha}{\alpha!} \nabla^\alpha \bar{w}(x),$$

where we use standard multi-index notation, and the  $n$ th-order standard homogenization commutator  $\Xi_\varepsilon^{\circ,n}[\nabla \bar{w}]$  is then defined as the (distributional) random vector field

$$\Xi_\varepsilon^{\circ,n}[\nabla \bar{w}](x) := \Xi_\varepsilon^n[\nabla F_\varepsilon^n[T_x^n \bar{w}]](x) \stackrel{(2.5)}{=} \Xi_\varepsilon^n[E_\varepsilon^n[\nabla T_x^n \bar{w}]](x).$$

(Henceforth by a slight abuse of notation we similarly define  $\Xi^n[H]$  as in Definition 3.1 even when  $H$  is not a gradient field.) ◇

This definition ensures  $\mathbb{E}[\Xi_\varepsilon^{\circ,n}[\nabla \bar{w}]] = 0$  for all  $0 \leq n \leq \ell$  and  $\bar{w} \in C_c^\infty(\mathbb{R}^d)$ . Note that  $\Xi_\varepsilon^{\circ,n}[\nabla \bar{w}]$  in general differs from  $\Xi_\varepsilon^n[E_\varepsilon^n[\nabla \bar{w}]]$  whenever  $n > 1$ , even though  $E_\varepsilon^n[\nabla T_x^n \bar{w}](x) = E_\varepsilon^n[\nabla \bar{w}](x)$ . An explicit formula is as follows.

**Lemma 3.5** (Explicit formula for  $\Xi_\varepsilon^{\circ,n}$ ) For all  $\bar{w} \in C_c^\infty(\mathbb{R}^d)$  and  $1 \leq n \leq \ell$ ,

$$\begin{aligned} \Xi_\varepsilon^{\circ,n}[\nabla \bar{w}] &= \Xi_\varepsilon^n[E_\varepsilon^n[\nabla \bar{w}]] \\ &+ \sum_{k=1}^{n-1} \varepsilon^k (\text{Sym}_{i_1 \dots i_k} \bar{a}_{i_1 \dots i_k}^{k+1}) \sum_{s=0}^{n-1} \varepsilon^s \sum_{l=0}^{k+s-n} \binom{k}{l} \nabla_{i_1 \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w} \\ &\times \nabla_{i_1 \dots i_l}^l \left( (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \binom{\cdot}{\varepsilon} \right). \end{aligned}$$

◇

**Remark 3.6** While Proposition 3.2 reduces higher-order fluctuations of the field  $\nabla u_\varepsilon$  to corresponding fluctuations of its higher-order homogenization commutator, we briefly argue that fluctuations of higher-order correctors can similarly be reduced to

fluctuations of higher-order standard commutators. A simple adaptation of the proof of Proposition 3.2 yields for  $0 \leq n \leq \ell - 1$ ,

$$N_p \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot E_\varepsilon^n [\nabla \bar{w}] - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^n \cdot \Xi_\varepsilon^n [E_\varepsilon^n [\nabla \bar{w}]] \right) \lesssim_{p,f,g} \varepsilon^n.$$

Inserting the formula of Lemma 3.5 and the definition of  $E_\varepsilon^n$  (cf. Definition 2.6), the expression on the left-hand side can, after straightforward computations, be rewritten as follows,

$$\begin{aligned} & \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot E_\varepsilon^n [\nabla \bar{w}] - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^n \cdot \Xi_\varepsilon^n [E_\varepsilon^n [\nabla \bar{w}]] \\ &= \sum_{s=0}^{n-1} \varepsilon^{s-\frac{d}{2}} \int_{\mathbb{R}^d} T_{\varepsilon; j_1 \dots j_{s+1}}^{n,s} (f, g) \cdot (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) (\dot{\varepsilon}) \\ & \quad - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v}_\varepsilon^n \cdot \Xi_\varepsilon^{\circ,n} [\nabla \bar{w}], \end{aligned} \tag{3.3}$$

in terms of

$$\begin{aligned} T_{\varepsilon; j_1 \dots j_{s+1}}^{n,s} (f, g) &:= g \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w} + \sum_{k=1}^{n-1} \varepsilon^k (-1)^{k+s-n} \sum_{r=0}^{k+s-n} \binom{k}{r} \binom{k-1-r}{k+s-n-r} \\ & \quad \times (\text{Sym}_{i_1 \dots i_k} \bar{a}_{i_1 \dots i_k}^{k+1})^* (\nabla \nabla_{i_1 \dots i_r}^r \bar{v}_\varepsilon^n) (\nabla_{i_{r+1} \dots i_k}^{k-r} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}). \end{aligned}$$

If we wish to describe only first-order fluctuations, we may replace  $T_\varepsilon^{n,s}(f, g)$  by  $g \nabla^{s+1} \bar{w}$  and  $\nabla \bar{v}_\varepsilon^n$  by  $\nabla \bar{v}$ , and use Lemma 2.3 to estimate the remainder, to the effect of

$$N_p \left( \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot E_\varepsilon^n [\nabla \bar{w}] - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v} \cdot \Xi_\varepsilon^{\circ,n} [\nabla \bar{w}] \right) \lesssim_{p,g,\bar{w}} \varepsilon,$$

hence, recalling Definition 2.6 and taking differences, for all  $0 \leq n \leq \ell - 1$ ,

$$\begin{aligned} & N_p \left( \varepsilon^{n-1-\frac{d}{2}} \int_{\mathbb{R}^d} g \nabla_{j_1 \dots j_n}^n \bar{w} \cdot (\nabla \varphi_{j_1 \dots j_n}^n + \varphi_{j_1 \dots j_{n-1}}^{n-1} e_{j_n}) (\dot{\varepsilon}) \right. \\ & \quad \left. - \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} \nabla \bar{v} \cdot (\Xi_\varepsilon^{\circ,n} [\nabla \bar{w}] - \Xi_\varepsilon^{\circ,n-1} [\nabla \bar{w}]) \right) \lesssim_{p,g,\bar{w}} \varepsilon. \end{aligned}$$

This characterizes the first-order fluctuations of  $\text{Sym}_{j_1 \dots j_n} (\nabla \varphi_{j_1 \dots j_n}^n + \varphi_{j_1 \dots j_{n-1}}^{n-1} e_{j_n})$  in terms of (differences of) standard commutators. Note that fluctuations of the latter are easily jointly characterized by repeating the analysis of Sects. 8 and 9, and are seen to converge to suitable linear combinations of derivatives of Gaussian white noise. From here we may inductively infer first-order fluctuations of  $\text{Sym}_{j_1 \dots j_n} \varphi_{j_1 \dots j_n}^n$  for all  $n \leq \ell - 1$ , while corresponding higher-order fluctuations are similarly extracted from (3.3).  $\diamond$

### 3.1 Proof of Proposition 3.2

We focus on the case  $\varepsilon = 1$  and drop it from all subscripts in the notation, while the final result is obtained after  $\varepsilon$ -rescaling. We split the proof into three steps.

Step 1. Proof of

$$\int_{\mathbb{R}^d} g \cdot \nabla u - \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot f = \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot \Xi^n[\nabla u] + \int_{\mathbb{R}^d} \left( \sum_{k=2}^n \sum_{j=n+2-k}^n \bar{a}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \tilde{v}^j \right) \cdot \nabla u. \tag{3.4}$$

Testing equation  $-\nabla \cdot \bar{a}^{*,1} \nabla \bar{v}^1 = \nabla \cdot g$  with  $u$  and testing equation  $-\nabla \cdot a \nabla u = \nabla \cdot f$  with  $\bar{v}^n$ , we find

$$\int_{\mathbb{R}^d} g \cdot \nabla u - \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot f = \int_{\mathbb{R}^d} (a^* \nabla \bar{v}^n - \bar{a}^{*,1} \nabla \bar{v}^1) \cdot \nabla u.$$

Noting that the definition of  $\Xi^n[\nabla u]$  (cf. Definition 3.1) together with (3.1) yields after integration by parts

$$\int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot \Xi^n[\nabla u] = \int_{\mathbb{R}^d} \left( a^* \nabla \bar{v}^n - \sum_{k=1}^n \bar{a}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{v}^n \right) \cdot \nabla u,$$

we deduce

$$\int_{\mathbb{R}^d} g \cdot \nabla u - \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot f = \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot \Xi^n[\nabla u] + \int_{\mathbb{R}^d} \left( \sum_{k=1}^n \bar{a}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \bar{v}^n - \bar{a}^{*,1} \nabla \bar{v}^1 \right) \cdot \nabla u.$$

Decomposing  $\bar{v}^n := \sum_{j=1}^n \tilde{v}^j$  and using the defining equations for  $(\tilde{v}^j)_{2 \leq j \leq n}$  (cf. Definition 2.7) in the resumed form

$$\nabla \cdot \sum_{k=1}^n \sum_{j=1}^{n+1-k} \bar{a}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \tilde{v}^j = \nabla \cdot \bar{a}^{*,1} \nabla \bar{v}^1,$$

the claim follows.

Step 2. CLT scaling: for all  $g \in C_c^\infty(\mathbb{R}^d)^d$ ,

$$\mathbb{E} \left[ \left( \int_{\mathbb{R}^d} g \cdot (\nabla u - \mathbb{E}[\nabla u]) \right)^p \right]^{\frac{1}{p}} \lesssim_p \|g\|_2 \|f\|_2 \leq \|g\|_{L^4(\mathbb{R}^d)} \|f\|_{L^4(\mathbb{R}^d)}.$$

This result is classical [18,20,37]. A short proof based on Malliavin calculus and on the annealed Calderón–Zygmund theory of Sect. 6 is easily obtained as in (7.9); details are omitted.

Step 3. Conclusion.

Since  $\int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot f$  is deterministic, the identity of Step 1 yields

$$\begin{aligned} & \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} g \cdot (\nabla u - \mathbb{E}[\nabla u]) - \int_{\mathbb{R}^d} \nabla \bar{v}^n \cdot (\mathfrak{E}^n[\nabla u] - \mathbb{E}[\mathfrak{E}^n[\nabla u]]) \right)^p \right]^{\frac{1}{p}} \\ & \lesssim \sum_{k=2}^n \sum_{j=n+2-k}^n \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \bar{a}_{i_1 \dots i_{k-1}}^{*,k} \nabla \nabla_{i_1 \dots i_{k-1}}^{k-1} \tilde{v}^j \cdot (\nabla u - \mathbb{E}[\nabla u]) \right)^p \right]^{\frac{1}{p}}. \end{aligned}$$

Applying the CLT scaling result of Step 2 and the Calderón–Zygmund theory for the constant-coefficient equations defining  $(\tilde{v}^k)_{1 \leq k \leq n}$ , this expression is estimated by

$$\lesssim_p \sum_{k=n}^{2(n-1)} \|\nabla^k g\|_{L^4(\mathbb{R}^d)} \|f\|_{L^4(\mathbb{R}^d)},$$

and the conclusion follows after  $\varepsilon$ -rescaling. □

### 3.2 Proof of Lemma 3.3

We focus on the case  $\varepsilon = 1$  and drop it from all subscripts in the notation, while the final result is obtained after  $\varepsilon$ -rescaling. We split the proof into two steps.

Step 1. Proof that for all  $1 \leq n \leq \ell$ ,

$$\begin{aligned} e_j \cdot \mathfrak{E}^n[\nabla u] &= (-1)^n \nabla_{i_1 \dots i_n}^n \left( (a^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla u + \varphi_{j i_1 \dots i_{n-1}}^{*,n} f_{i_n} \right) \\ &+ \sum_{k=0}^{n-1} (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\nabla \varphi_{j i_1 \dots i_k}^{*,k+1} + \mathbb{1}_{k>0} \varphi_{j i_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot f \right), \end{aligned} \tag{3.5}$$

which coincides with (3.2) after  $\varepsilon$ -rescaling.

First, by Lemma 2.4, the  $n$ th-order commutator can alternatively be written as

$$e_j \cdot \mathfrak{E}^n[\nabla u] = a^* e_j \cdot \nabla u - \sum_{k=1}^n (-1)^{k-1} \nabla_{i_1 \dots i_{k-1}}^{k-1} (\bar{a}_{j i_1 \dots i_{k-2}}^{*,k} e_{i_{k-1}} \cdot \nabla u). \tag{3.6}$$

It remains to show that for all  $0 \leq n \leq \ell$ ,

$$\begin{aligned} a^* e_j \cdot \nabla u &= (-1)^n \nabla_{i_1 \dots i_n}^n \left( (a^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla u \right) \\ &+ \sum_{k=1}^n (-1)^{k-1} \nabla_{i_1 \dots i_{k-1}}^{k-1} (\varphi_{j i_1 \dots i_{k-1}}^{*,k} \nabla \cdot a \nabla u) \\ &+ \sum_{k=1}^n (-1)^{k-1} \nabla_{i_1 \dots i_{k-1}}^{k-1} (\bar{a}_{j i_1 \dots i_{k-2}}^{*,k} e_{i_{k-1}} \cdot \nabla u), \end{aligned} \tag{3.7}$$

while the result (3.5) indeed follows after injecting the equation  $-\nabla \cdot \mathbf{a} \nabla u = \nabla \cdot f$  and writing  $\varphi_{j_{i_1} \dots i_{k-1}}^{*,k} \nabla \cdot f = \nabla \cdot (\varphi_{j_{i_1} \dots i_{k-1}}^{*,k} f) - \nabla \varphi_{j_{i_1} \dots i_{k-1}}^{*,k} \cdot f$ . We prove (3.7) by induction. It is obvious for  $n = 0$  (recall that for  $n = 0$  the notation  $\mathbf{a}^* \varphi_{j_{i_1} \dots i_{n-1}}^{*,n} e_{i_n}$  stands for  $\mathbf{a}^* e_j$ ). Assume that it holds for some  $n \geq 0$ . In order to deduce it at level  $n + 1$ , it suffices to prove that

$$\begin{aligned} & \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_{i_1} \dots i_{n-1}}^{*,n} - \sigma_{j_{i_1} \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla u \right) \\ &= -\nabla_{i_1 \dots i_n}^{n+1} \left( (\mathbf{a}^* \varphi_{j_{i_1} \dots i_n}^{*,n+1} - \sigma_{j_{i_1} \dots i_n}^{*,n+1}) e_l \cdot \nabla u \right) \\ & \quad + \nabla_{i_1 \dots i_n}^n (\varphi_{j_{i_1} \dots i_n}^{*,n+1} \nabla \cdot \mathbf{a} \nabla u) + \nabla_{i_1 \dots i_n}^n (\bar{\mathbf{a}}_{j_{i_1} \dots i_{n-1}}^{*,n+1} e_{i_n} \cdot \nabla u). \end{aligned} \tag{3.8}$$

The equation for  $\nabla \cdot \sigma^{*,n+1}$  (cf. Definition 2.1) yields

$$\begin{aligned} & \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_{i_1} \dots i_{n-1}}^{*,n} - \sigma_{j_{i_1} \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla w \right) = \nabla_{i_1 \dots i_n}^n \left( (\nabla \cdot \sigma_{j_{i_1} \dots i_n}^{*,n+1}) \cdot \nabla w \right) \\ & \quad - \nabla_{i_1 \dots i_n}^n (\nabla \varphi_{j_{i_1} \dots i_n}^{*,n+1} \cdot \mathbf{a} \nabla w) + \nabla_{i_1 \dots i_n}^n (\bar{\mathbf{a}}_{j_{i_1} \dots i_{n-1}}^{*,n+1} e_{i_n} \cdot \nabla w), \end{aligned}$$

and the claim (3.8) follows from the skew-symmetry of  $\sigma_{j_{i_1} \dots i_n}^{*,n+1}$ .

*Step 2. Conclusion.*

For  $g \in C_c^\infty(\mathbb{R}^d)^d$ , combining (3.5) with the energy estimate  $\|\nabla u\|_{L^2(\mathbb{R}^d)} \lesssim \|f\|_{L^2(\mathbb{R}^d)}$  and with the corrector estimates of Proposition 2.2, we find for all  $p < \infty$ ,

$$\begin{aligned} & \mathbb{E} \left[ \left| \int_{\mathbb{R}^d} g \cdot \Xi^n[\nabla u] \right|^p \right]^{\frac{1}{p}} \lesssim_p \|\mu_{d,n} \nabla^n g\|_{L^2(\mathbb{R}^d)} \|f\|_{L^2(\mathbb{R}^d)} \\ & \quad + \sum_{k=0}^{n-1} \mathbb{E} \left[ \left| \int_{\mathbb{R}^d} (\nabla \varphi_{j_{i_1} \dots i_k}^{*,k+1} + \mathbf{1}_{k>0} \varphi_{j_{i_1} \dots i_{k-1}}^{*,k} e_{i_k}) \cdot f \nabla_{i_1 \dots i_k}^k g_j \right|^p \right]^{\frac{1}{p}}. \end{aligned}$$

In order to optimally estimate the last contribution, we must exploit stochastic cancellations in form of Lemma 2.3, which yields after  $\varepsilon$ -rescaling, for all  $1 \leq n \leq \ell$ ,

$$\begin{aligned} & \mathbb{E} \left[ \left| \int_{\mathbb{R}^d} g \cdot \Xi_\varepsilon^n[\nabla u_\varepsilon] \right|^p \right]^{\frac{1}{p}} \lesssim_p \varepsilon^n \mu_{d,n} \left(\frac{1}{\varepsilon}\right) \\ & \quad \times \left( \|\mu_{d,n} \nabla^n g\|_{L^2(\mathbb{R}^d)} \|f\|_{L^2(\mathbb{R}^d)} + \sum_{k=0}^{n-1} \|[f \nabla^k g]_2\|_{L^{\frac{2d}{d+2k}}(\mathbb{R}^d)} \right). \end{aligned}$$

Since  $\ell - 1 \leq \frac{1}{2}(d - 1)$ , using Hölder’s and Sobolev’s inequalities in the form

$$\begin{aligned} \max_{0 \leq k \leq \ell-1} \|[f \nabla^k g]_2\|_{L^{\frac{2d}{d+2k}}(\mathbb{R}^d)} & \leq \max_{0 \leq k \leq \ell-1} \|[f]_{\frac{2d}{d-2k}}\|_{L^2(\mathbb{R}^d)} \|\nabla^k g\|_{L^{\frac{d}{k}}(\mathbb{R}^d)} \\ & \lesssim \|[f]_2\|_{L^2(\mathbb{R}^d)} \|g\|_{H^{\frac{d}{2}} \cap L^\infty(\mathbb{R}^d)}, \end{aligned}$$

the conclusion follows. □



### 3.3 Proof of Lemma 3.5

First note that, given a linear differential operator  $O$  of order  $m$ , say  $O[\bar{w}] = \sum_{k=0}^m P_{j_1 \dots j_k}^k(x) \nabla_{j_1 \dots j_k}^k \bar{w}$ , the characterizing property of the Taylor polynomial yields

$$O[T_x^n \bar{w}](x) = \sum_{k=0}^{m \wedge n} P_{j_1 \dots j_k}^k(x) \nabla_{j_1 \dots j_k}^k \bar{w}(x), \tag{3.9}$$

which amounts to truncating the differential operator  $O$  at order  $n$ . Now observe that  $\Xi_\varepsilon^n[E_\varepsilon^n[\nabla \cdot]]$  is a linear differential operator of order  $2n - 1$  with (distributional) stationary random coefficients. Composing the explicit definitions of  $E_\varepsilon^n$  and  $\Xi_\varepsilon^n$  (cf. Definitions 2.6 and 3.1), we obtain

$$\begin{aligned} \Xi_\varepsilon^n[E_\varepsilon^n[\nabla \bar{w}]] &= (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}}^1) E_\varepsilon^n[\nabla \bar{w}] \\ &- \sum_{k=2}^n \varepsilon^{k-1} \bar{\mathbf{a}}_{i_1 \dots i_{k-1}}^k \nabla_{i_1 \dots i_{k-1}}^{k-1} \sum_{s=0}^{n-1} \varepsilon^s (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) (\frac{\cdot}{\varepsilon}) \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}. \end{aligned}$$

Hence, relabelling  $k - 1$  as  $k$  in the second right-hand side contribution, using the symmetry in  $i_1, \dots, i_k$ , and expanding the multiple derivative  $\nabla_{i_1 \dots i_k}^k$  by using the general Leibniz rule,

$$\begin{aligned} \Xi_\varepsilon^n[E_\varepsilon^n[\nabla \bar{w}]] &= (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}}^1) E_\varepsilon^n[\nabla \bar{w}] \\ &- \sum_{k=1}^{n-1} \varepsilon^k (\text{Sym}_{i_1 \dots i_k} \bar{\mathbf{a}}_{i_1 \dots i_k}^{k+1}) \sum_{s=0}^n \varepsilon^s \sum_{l=0}^k \binom{k}{l} \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w} \\ &\times \nabla_{i_1 \dots i_l}^l \left( (\mathbb{1}_{s \neq n} \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) (\frac{\cdot}{\varepsilon}) \right). \end{aligned} \tag{3.10}$$

Now applying (3.9) to the definition of  $\Xi_\varepsilon^{o,n}[\nabla \bar{w}]$  (cf. Definition 3.4) together with the above formula for  $\Xi_\varepsilon^n[E_\varepsilon^n[\nabla \bar{w}]]$ , we are led to

$$\begin{aligned} \Xi_\varepsilon^{o,n}[\nabla \bar{w}] &= (\mathbf{a}(\frac{\cdot}{\varepsilon}) - \bar{\mathbf{a}}^1) E_\varepsilon^n[\nabla \bar{w}] \\ &- \sum_{k=1}^{n-1} \varepsilon^k (\text{Sym}_{i_1 \dots i_k} \bar{\mathbf{a}}_{i_1 \dots i_k}^{k+1}) \sum_{s=0}^n \varepsilon^s \sum_{l=k+s+1-n}^k \binom{k}{l} \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w} \\ &\times \nabla_{i_1 \dots i_l}^l \left( (\mathbb{1}_{s \neq n} \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) (\frac{\cdot}{\varepsilon}) \right). \end{aligned}$$

Comparing this with (3.10) yields the conclusion. □

### 4 Reminder on Malliavin calculus

In this section, we recall some basic definitions of the Malliavin calculus with respect to the Gaussian field  $G$  (e.g. [36,42,44] for details). For simplicity, we focus on the explicit

description of the Malliavin derivative on a simple dense subspace of random variables (cf.  $\mathcal{R}$  hereafter), where it is seen to coincide with the formal  $L^2$ -gradient with respect to  $G$ . Next, we give a precise statement of the different tools from Malliavin calculus that will be used in the sequel. We emphasize that the reader should not be intimidated by the somehow abstract appearance of the theory: for our application to stochastic homogenization all (averaged) quantities of interest happen to be Malliavin smooth, and their Malliavin derivatives are systematically computed by duality arguments so that in the end everything is expressed in terms of the (explicit) Malliavin derivative of the coefficient field  $\mathbf{a}$  itself.

The Gaussian random field  $G$  can be seen as a random Schwartz distribution, that is, as a random element in  $\mathcal{S}'(\mathbb{R}^d)^\kappa$ . Indeed, for all  $\zeta_1, \zeta_2 \in C_c^\infty(\mathbb{R}^d)^\kappa$ , we define  $G(\zeta_1), G(\zeta_2)$  (or formally  $\int_{\mathbb{R}^d} G \zeta_1, \int_{\mathbb{R}^d} G \zeta_2$ ) as centered Gaussian random variables with covariance

$$\text{Cov} [G(\zeta_1); G(\zeta_2)] := \iint_{\mathbb{R}^d \times \mathbb{R}^d} \zeta_1(x) \cdot c(x - y) \zeta_2(y) \, dx dy.$$

We define  $\mathfrak{H}$  as the closure of  $C_c^\infty(\mathbb{R}^d)^\kappa$  for the (semi)norm

$$\|\zeta_1\|_{\mathfrak{H}}^2 := \langle \zeta_1, \zeta_1 \rangle_{\mathfrak{H}}, \quad \langle \zeta_1, \zeta_2 \rangle_{\mathfrak{H}} := \iint_{\mathbb{R}^d \times \mathbb{R}^d} \zeta_1(x) \cdot c(x - y) \zeta_2(y) \, dx dy. \tag{4.1}$$

The space  $\mathfrak{H}$  (up to taking the quotient with respect to the kernel of  $\|\cdot\|_{\mathfrak{H}}$ ) is a separable Hilbert space. Under the integrability condition (1.5), it obviously contains  $L^2(\mathbb{R}^d)^\kappa$  and there holds

$$\|\zeta\|_{\mathfrak{H}}^2 \lesssim \int_{\mathbb{R}^d} |\zeta|^2. \tag{4.2}$$

In view of the isometry relation  $\text{Cov} [G(\zeta_1); G(\zeta_2)] = \langle \zeta_1, \zeta_2 \rangle_{\mathfrak{H}}$ , the random field  $G$  is said to be an isonormal Gaussian process over  $\mathfrak{H}$ .

Without loss of generality, we work under the minimality assumption  $\mathcal{F} = \sigma(G)$ , which implies that the linear subspace

$$\mathcal{R} := \left\{ g(G(\zeta_1), \dots, G(\zeta_n)) : n \in \mathbb{N}, g \in C_c^\infty(\mathbb{R}^n), \zeta_1, \dots, \zeta_n \in C_c^\infty(\mathbb{R}^d)^\kappa \right\} \subset L^2(\Omega)$$

is dense in  $L^2(\Omega)$ . This allows to define operators and prove properties on the simpler subspace  $\mathcal{R}$  before extending them to  $L^2(\Omega)$  by density. For  $r \geq 1$  we similarly define

$$\mathcal{R}(\mathfrak{H}^{\otimes r}) := \left\{ \sum_{i=1}^n \psi_i X_i : n \in \mathbb{N}, X_1, \dots, X_n \in \mathcal{R}, \psi_1, \dots, \psi_n \in \mathfrak{H}^{\otimes r} \right\} \subset L^2(\Omega; \mathfrak{H}^{\otimes r}),$$

which is dense in  $L^2(\Omega; \mathfrak{H}^{\otimes r})$ . For a random variable  $X \in \mathcal{R}$ , say  $X = g(G(\zeta_1), \dots, G(\zeta_n))$ , we define its Malliavin derivative  $DX \in L^2(\Omega; \mathfrak{H})$  as

$$DX = \sum_{i=1}^n \zeta_i \partial_i g(G(\zeta_1), \dots, G(\zeta_n)), \tag{4.3}$$

which coincides with the formal  $L^2$  gradient of  $X$  with respect to  $G$ . For an element  $X \in \mathcal{R}(\mathfrak{H}^{\otimes r})$  with  $r \geq 1$ , say  $X = \sum_{i=1}^n \psi_i X_i$ , the Malliavin derivative  $DX \in L^2(\Omega; \mathfrak{H}^{\otimes(r+1)})$  is defined as  $DX = \sum_{i=1}^n \psi_i \otimes DX_i$ . For  $j \geq 1$ , we iteratively define the  $j$ th-order Malliavin derivative  $D^j: \mathcal{R}(\mathfrak{H}^{\otimes r}) \rightarrow L^2(\Omega; \mathfrak{H}^{\otimes(r+j)})$  for all  $r \geq 0$ . For all  $r, m \geq 0$ , we then set

$$\begin{aligned} \|X\|_{\mathbb{D}^{m,2}(\mathfrak{H}^{\otimes r})}^2 &:= \langle X, X \rangle_{\mathbb{D}^{m,2}(\mathfrak{H}^{\otimes r})}, \\ \langle X, Y \rangle_{\mathbb{D}^{m,2}(\mathfrak{H}^{\otimes r})} &:= \mathbb{E}[\langle X, Y \rangle_{\mathfrak{H}^{\otimes r}}] + \sum_{j=1}^m \mathbb{E}[\langle D^j X, D^j Y \rangle_{\mathfrak{H}^{\otimes(r+j)}}], \end{aligned}$$

we define the Malliavin–Sobolev space  $\mathbb{D}^{m,2}(\mathfrak{H}^{\otimes r})$  as the closure of  $\mathcal{R}(\mathfrak{H}^{\otimes r})$  for this norm, and we extend the Malliavin derivatives  $D^j$  by density to these spaces.

Next, we define a divergence operator  $D^*$  as the adjoint of the Malliavin derivative  $D$  and we start with its domain: for  $r \geq 1$ , we set

$$\begin{aligned} \text{dom}(D^*|_{L^2(\Omega; \mathfrak{H}^{\otimes r})}) &:= \left\{ X \in L^2(\Omega; \mathfrak{H}^{\otimes r}) : \exists C < \infty \text{ such that} \right. \\ &\quad \left. \mathbb{E}[\langle DY, X \rangle_{\mathfrak{H}^{\otimes r}}]^2 \leq C \mathbb{E}[\|Y\|_{\mathfrak{H}^{\otimes(r-1)}}^2] \text{ for all } Y \in \mathbb{D}^{1,2}(\Omega; \mathfrak{H}^{\otimes(r-1)}) \right\}, \end{aligned}$$

and for all  $X \in \text{dom}(D^*|_{L^2(\Omega; \mathfrak{H}^{\otimes r})})$  we define the divergence  $D^*X \in L^2(\Omega; \mathfrak{H}^{\otimes(r-1)})$  by duality via the relation

$$\mathbb{E}[\langle Y, D^*X \rangle_{\mathfrak{H}^{\otimes(r-1)}}] = \mathbb{E}[\langle DY, X \rangle_{\mathfrak{H}^{\otimes r}}] \quad \text{for all } Y \in \mathbb{D}^{1,2}(\mathfrak{H}^{\otimes(r-1)}). \tag{4.4}$$

One can show that  $\text{dom}(D^*|_{L^2(\Omega; \mathfrak{H}^{\otimes r})}) \supset \mathbb{D}^{1,2}(\mathfrak{H}^{\otimes r})$  (e.g. [44, Proposition 1.3.1]). Combining the Malliavin derivative and the associated divergence, we construct the so-called Ornstein–Uhlenbeck operator (or infinite-dimensional Laplacian)

$$\mathcal{L} := D^*D,$$

which is an essentially self-adjoint positive operator. The explicit actions of  $D^*$  and  $\mathcal{L}$  on  $\mathcal{R}$  are easily computed (e.g. [42, p. 34]): for  $X \in \mathcal{R}(\mathfrak{H}^{\otimes r})$ , say  $X = \sum_{i=1}^n \psi_i X_i$ , we have  $\mathcal{L}X = \sum_{i=1}^n \psi_i \mathcal{L}X_i$ , while for  $X_i = g(G(\zeta_1), \dots, G(\zeta_n))$ ,

$$\mathcal{L}X_i = \sum_{j=1}^n G(\zeta_j) \partial_j g(G(\zeta_1), \dots, G(\zeta_n)) - \sum_{j,k=1}^n \langle \zeta_j, \zeta_k \rangle_{\mathfrak{H}} \partial_{jk}^2 g(G(\zeta_1), \dots, G(\zeta_n)).$$

In particular, a direct computation (e.g. [42, p. 35]) leads to the crucial commutator relation

$$D\mathcal{L} = (1 + \mathcal{L})D. \tag{4.5}$$

In addition, one checks that  $\mathcal{L}$  is shift-invariant. While the positivity of  $\mathcal{L}$  ensures that the inverse operator  $(1 + \mathcal{L})^{-1}$  is contractive on  $L^2(\Omega)$ , it is easily checked to be also contractive on  $L^p(\Omega)$  for all  $p > 1$  (cf. e.g. [38, Proposition 3.2]). Although we focus here on the explicit description of operators on their core  $\mathcal{R}$ , the action of  $D, D^*, \mathcal{L}$  on general  $L^2$  random variables are naturally described in terms of their Wiener chaos expansion (e.g. [42,44]).

Based on the above definitions, we state the following proposition collecting various useful results for the fine analysis of functionals of the Gaussian field  $G$ . Item (i) is classical [15,29,40], as well as item (iii) (e.g. [8] and references therein). Item (ii) is well-known in the discrete Gaussian setting [26] (cf. also [38,50]). Item (iv) in total variation distance (and similarly in 1-Wasserstein distance) is a consequence of Stein’s method: it was first obtained in the discrete setting by Chatterjee [14], while the present Malliavin analogue is due to [41,43] (of which we give a slightly different formulation here). The corresponding result in 2-Wasserstein distance is different and is due to [35]. Note the particular role of the so-called Stein kernel  $\langle DX, (1 + \mathcal{L})^{-1}DX \rangle_{\mathfrak{H}}$  in items (ii) and (iv). A short proof is included in “Appendix A” for the reader’s convenience.

**Proposition 4.1** [14,26,29,35,41,43]

(i) First-order Poincaré inequality For all  $X \in L^2(\Omega)$ ,

$$\text{Var}[X] \leq \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^2 \right].$$

(ii) Helffer–Sjöstrand identity For all  $X, Y \in \mathbb{D}^{1,2}(\Omega)$ ,

$$\text{Cov}[X; Y] = \mathbb{E} \left[ \langle DX, (1 + \mathcal{L})^{-1}DY \rangle_{\mathfrak{H}} \right]. \tag{4.6}$$

(iii) Logarithmic Sobolev inequality For all  $X \in L^2(\Omega)$ ,

$$\text{Ent}[X^2] := \mathbb{E} \left[ X^2 \log \frac{X^2}{\mathbb{E}[X^2]} \right] \leq 2 \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^2 \right],$$

hence in particular, for all  $p \geq 1$ ,

$$\mathbb{E} \left[ |X - \mathbb{E}[X]|^{2p} \right] \leq (2p + 1)^p \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^{2p} \right].$$

(iv) Second-order Poincaré inequality For all  $X \in L^2(\Omega)$  with  $\mathbb{E}[X] = 0$  and  $\text{Var}[X] = 1$ ,

$$\begin{aligned} W_2(X; \mathcal{N}) \vee d_{\text{TV}}(X; \mathcal{N}) &\leq 2 \text{Var} \left[ \langle DX, (1 + \mathcal{L})^{-1}DX \rangle_{\mathfrak{H}} \right]^{\frac{1}{2}} \\ &\leq 3 \mathbb{E} \left[ \|D^2X\|_{\text{op}}^4 \right]^{\frac{1}{4}} \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^4 \right]^{\frac{1}{4}}, \end{aligned}$$

where  $W_2(\cdot; \mathcal{N})$  and  $d_{TV}(\cdot; \mathcal{N})$  denote the 2-Wasserstein and the total variation distance to a standard Gaussian law, respectively, and where the operator norm of  $D^2X$  is defined by

$$\|D^2X\|_{\text{op}} := \sup_{\substack{\zeta, \zeta' \in \mathfrak{H} \\ \|\zeta\|_{\mathfrak{H}} = \|\zeta'\|_{\mathfrak{H}} = 1}} \langle D^2X, \zeta \otimes \zeta' \rangle_{\mathfrak{H} \otimes \mathfrak{H}}. \tag{4.7}$$

◇

**Remark 4.2** Malliavin-type calculus has also been successfully developed in some non-Gaussian settings, to which most of our conclusions could thus be extended. For instance, in the case of an i.i.d. discrete random field  $G := (G_x)_{x \in \mathbb{Z}^d}$ , a corresponding calculus is based on the so-called Glauber derivative  $D_x X = \mathbb{E}[X | G_x] - X$  (e.g. [19]). Another important example is that of a Poisson point process  $\xi$  on  $\mathbb{R}^d$ , in which case the Malliavin derivative is chosen as the difference operator  $D_x X(\xi) = X(\xi \cup \{x\}) - X(\xi \setminus \{x\})$  (e.g. [46]).

◇

### 5 Representation formulas

As recalled in Proposition 4.1 above, tools from Malliavin calculus allow to linearize the dependence on the underlying Gaussian field  $G$  and reduce various subtle questions about quantities of interest to the estimation of their Malliavin derivative, that is, of their local dependence with respect to  $G$ . In this spirit, the key ingredient for items (ii) and (iii) of Theorem 1 consists in suitable representation formulas for the Malliavin derivative of higher-order commutators.

- *Locality of higher-order commutators*

The key property of homogenization commutators for fluctuations is their improved locality with respect to the coefficient field. While this was first discovered at first order in [19] (see also [5,23,25]), formula (5.1) below indicates that the higher-order commutator  $\Xi_\varepsilon^n[\nabla u_\varepsilon]$  behaves like a local quantity at the level of its Malliavin derivative up to a higher-order error  $O(\varepsilon^n)$ . More precisely, the formula takes the form

$$e_j \cdot D \Xi_\varepsilon^n[\nabla u_\varepsilon] = \sum_{k=0}^{n-1} (-1)^k \varepsilon^k \nabla_{i_1 \dots i_k}^k \left( (\nabla \varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1} i_k}^{*,k}) \left( \frac{\cdot}{\varepsilon} \right) \cdot D \mathbf{a} \left( \frac{\cdot}{\varepsilon} \right) \nabla u_\varepsilon \right) + O(\varepsilon^n),$$

where the main right-hand side terms are indeed local in view of  $D_z \mathbf{a} = a'_0(G(z)) \delta(\cdot - z)$ , and where the error can be checked to be of order  $O(\varepsilon^n)$  in  $L^2(\Omega; \mathfrak{H})$  when integrated with a smooth test function. While the full formula (5.1) below is given for simplicity for  $\varepsilon = 1$ , the  $\varepsilon$ -scaling is transparent by counting the number of derivatives that fall on test functions. The proof consists in exploiting the algebraic structure of higher-order correctors and commutators in order to make this clean local structure appear.

Since the standard commutator  $\Xi_\varepsilon^{\circ,n}[\nabla\bar{w}]$  is obtained by applying  $\Xi_\varepsilon^n[\nabla\cdot]$  to the two-scale expansion  $F_\varepsilon^n[\nabla\bar{w}]$  whenever  $\bar{w}$  is an  $n$ th-order polynomial, we may deduce a corresponding formula (5.2) for its Malliavin derivative,

$$e_j \cdot D \Xi_\varepsilon^{\circ,n}[\nabla\bar{w}] = \sum_{k=0}^{n-1} (-1)^k \varepsilon^k \nabla_{i_1 \dots i_k}^k \left( (\nabla\varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \binom{\cdot}{\varepsilon} \cdot D\mathbf{a} \binom{\cdot}{\varepsilon} E_\varepsilon^n[\nabla\bar{w}] \right) + O(\varepsilon^n).$$

In view of Proposition 4.1(ii), this improved locality property provides a simple access to the higher-order characterization of the asymptotic covariance structure of large-scale averages of commutators, cf. Sect. 8.

• *Accuracy of the two-scale expansion of higher-order commutators:*

Comparing the formulas for Malliavin derivatives of the commutator  $\Xi_\varepsilon^n[\nabla u_\varepsilon]$  and of its standard version  $\Xi_\varepsilon^{\circ,n}[\nabla\bar{w}]$ , and choosing  $\bar{w} = \bar{u}_\varepsilon^n$  the homogenized solution, we find for the difference,

$$e_j \cdot D(\Xi_\varepsilon^n[\nabla u_\varepsilon] - \Xi_\varepsilon^{\circ,n}[\nabla\bar{u}_\varepsilon^n]) = O(\varepsilon^n) + \sum_{k=0}^{n-1} (-1)^k \varepsilon^k \nabla_{i_1 \dots i_k}^k \left( (\nabla\varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \binom{\cdot}{\varepsilon} \cdot D\mathbf{a} \binom{\cdot}{\varepsilon} (\nabla u_\varepsilon - E_\varepsilon^n[\nabla\bar{u}_\varepsilon^n]) \right),$$

where we exactly recognize the higher-order two-scale expansion error  $\nabla u_\varepsilon - E_\varepsilon^n[\nabla\bar{u}_\varepsilon^n]$ , which is known to be of small order  $O(\varepsilon^n)$  in view of Proposition 2.8. This constitutes the core of the proof of the higher-order pathwise result, cf. Sect. 7.

We turn to the precise statement of the representation formulas. These are understood in a weak sense: when integrated in space with a smooth test function, the quantities in consideration are indeed Malliavin smooth and the formulas make sense after integrations by parts (see also Lemmas 7.1 and 9.1).

**Proposition 5.1** (Formulas for first Malliavin derivatives) *Let  $\varepsilon = 1$  and drop it from all subscripts in the notation. For all  $1 \leq n \leq \ell$  there holds for any random function  $w$  with  $\nabla \cdot \mathbf{a} \nabla w$  deterministic,*

$$e_j \cdot D \Xi^n[\nabla w] = \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla\varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \cdot D\mathbf{a} \nabla w \right) + (-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla D w \right), \tag{5.1}$$

and also, for any smooth deterministic function  $\bar{w}$ ,

$$e_j \cdot D \Xi^{\circ,n}[\nabla\bar{w}] = \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla\varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \cdot D\mathbf{a} E^n[\nabla\bar{w}] \right) - \sum_{k=0}^n (-1)^k \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \nabla_{i_1 \dots i_l}^l$$

$$\begin{aligned} & \times \left( \text{Sym}_{i_1 \dots i_k} \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k} \right) \cdot \mathbf{Da} \left( \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}} \right) \right) \\ & + (-1)^n \sum_{s=0}^{n-1} \sum_{l=s+1}^n \binom{n}{l} \left( \nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w} \right) \nabla_{i_1 \dots i_l}^l \\ & \times \left( \text{Sym}_{i_1 \dots i_n} \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D \left( \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}} \right) \right). \end{aligned} \tag{5.2}$$

In particular,

$$\begin{aligned} & e_j \cdot D \left( \Xi^n[\nabla u] - \Xi^{\circ,n}[\nabla \bar{u}^n] \right) \\ & = \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k} \right) \cdot \mathbf{Da} \left( \nabla u - E^n[\nabla \bar{u}^n] \right) \right) \\ & + \sum_{k=0}^n (-1)^k \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \binom{k}{l} \left( \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{u}^n \right) \nabla_{i_1 \dots i_l}^l \\ & \times \left( \text{Sym}_{i_1 \dots i_k} \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k} \right) \cdot \mathbf{Da} \left( \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}} \right) \right) \\ & + (-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot D \left( \nabla u - E^n[\nabla \bar{u}^n] \right) \right) \\ & + (-1)^n \sum_{s=0}^{n-1} \sum_{l=0}^s \binom{n}{l} \left( \nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{u}^n \right) \nabla_{i_1 \dots i_l}^l \\ & \times \left( \text{Sym}_{i_1 \dots i_n} \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D \left( \nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}} \right) \right). \end{aligned} \tag{5.3}$$

◇

In view of the approximate normality result of Theorem 1(iv), the second-order Poincaré inequality of Proposition 4.1(iv) requires to further estimate the second Malliavin derivative of higher-order standard commutators, and we establish the following formula. Although some of the terms are proportional to  $D^2 \mathbf{a}$ , hence display an exact locality, some others unavoidably take the form  $\mathbf{Da} D \nabla \varphi^1$ , hence are only approximately local. This is due to the nonlinear dependence of the solution operator on the coefficient field, and is the reason why some logarithms are lost in the convergence rate to normality in Theorem 1(iv), cf. Sect. 9.

**Proposition 5.2** (Formulas for second Malliavin derivatives) *Let  $\varepsilon = 1$  and drop it from all subscripts in the notation. For all  $1 \leq n \leq \ell$ , there holds for any smooth deterministic function  $\bar{w}$ ,*

$$\begin{aligned} & e_j \cdot D^2 \Xi^{\circ,n}[\nabla \bar{w}] \\ & = \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k} \right) \cdot D^2 \mathbf{a} E^n[\nabla \bar{w}] \right) \end{aligned}$$

$$\begin{aligned}
 & - \sum_{k=0}^n (-1)^k \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 & \times \left( \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D^2 \mathbf{a} (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right) \\
 & + 2 \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D \mathbf{a} D E^n [\nabla \bar{w}] \right) \\
 & - 2 \sum_{k=0}^n (-1)^k \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 & \times \left( \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D \mathbf{a} D (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right) \\
 & + (-1)^n \sum_{s=0}^{n-1} \sum_{l=s+1}^n \binom{n}{l} (\nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 & \times \left( \text{Sym}_{i_1 \dots i_n} ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n}) \cdot D^2 (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right).
 \end{aligned}$$

**5.1 Proof of Proposition 5.1**

We split the proof into three steps.

*Step 1.* Proof that for any random function  $w$  with  $\nabla \cdot \mathbf{a} \nabla w$  deterministic there holds for all  $0 \leq n \leq \ell$ ,

$$\begin{aligned}
 e_j \cdot D \Xi^n [\nabla w] & = \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D \mathbf{a} \nabla w \right) \\
 & + (-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla D w \right). \tag{5.4}
 \end{aligned}$$

We argue by induction. The result is obvious for  $n = 0$  (recall  $\Xi^0 [\nabla w] = \mathbf{a} \nabla w$ , cf. Definition 3.1). Assume that it holds for some  $n \geq 0$ . In order to deduce it at level  $n + 1$ , we appeal to the alternative definition (3.6) of the commutator, so that it suffices to prove

$$\begin{aligned}
 & \nabla_{i_1 \dots i_n}^n ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla D w) \\
 & = -\nabla_{i_1 \dots i_n}^{n+1} ((\mathbf{a}^* \varphi_{j_1 \dots i_n}^{*,n+1} - \sigma_{j_1 \dots i_n}^{*,n+1}) e_l \cdot \nabla D w) - \nabla_{i_1 \dots i_n}^{n+1} (\varphi_{j_1 \dots i_n}^{*,n+1} e_l \cdot D \mathbf{a} \nabla w) \\
 & + \nabla_{i_1 \dots i_n}^n (\nabla \varphi_{j_1 \dots i_n}^{*,n+1} \cdot D \mathbf{a} \nabla w) + \nabla_{i_1 \dots i_n}^n (\bar{\mathbf{a}}_{j_1 \dots i_{n-1}}^{*,n+1} e_{i_n} \cdot \nabla D w). \tag{5.5}
 \end{aligned}$$

The definition of  $\sigma^{*,n+1}$  (cf. Definition 2.1) yields

$$\begin{aligned}
 & \nabla_{i_1 \dots i_n}^n ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot \nabla D w) = \nabla_{i_1 \dots i_n}^n ((\nabla \cdot \sigma_{j_1 \dots i_n}^{*,n+1}) \cdot \nabla D w) \\
 & - \nabla_{i_1 \dots i_n}^n (\nabla \varphi_{j_1 \dots i_n}^{*,n+1} \cdot \mathbf{a} \nabla D w) + \nabla_{i_1 \dots i_n}^n (\bar{\mathbf{a}}_{j_1 \dots i_{n-1}}^{*,n+1} e_{i_n} \cdot \nabla D w).
 \end{aligned}$$



Then using the skew-symmetry of  $\sigma_{j_1 \dots i_n}^{*,n+1}$  and the equation  $D(\nabla \cdot \mathbf{a} \nabla w) = 0$  in the form

$$\begin{aligned} \nabla \varphi_{j_1 \dots i_n}^{*,n+1} \cdot \mathbf{a} \nabla D w &= \nabla \cdot (\varphi_{j_1 \dots i_n}^{*,n+1} \mathbf{a} \nabla D w) + \varphi_{j_1 \dots i_n}^{*,n+1} \nabla \cdot D \mathbf{a} \nabla w \\ &= \nabla \cdot (\varphi_{j_1 \dots i_n}^{*,n+1} \mathbf{a} \nabla D w) \\ &\quad + \nabla \cdot (\varphi_{j_1 \dots i_n}^{*,n+1} D \mathbf{a} \nabla w) - \nabla \varphi_{j_1 \dots i_n}^{*,n+1} \cdot D \mathbf{a} \nabla w, \end{aligned}$$

the claim (5.5) follows.

*Step 2.* Proof that for any random function  $w$  with  $\nabla \cdot \mathbf{a} \nabla w$  deterministic and for any smooth deterministic function  $\bar{w}$  there holds for all  $0 \leq n, m \leq \ell$ ,

$$\begin{aligned} e_j \cdot D \Xi^n [\nabla w - E^m [\nabla \bar{w}]] &= \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \cdot D \mathbf{a} (\nabla w - E^m [\nabla \bar{w}]) \right) \\ &\quad + \sum_{k=0}^n (-1)^k \sum_{s=1}^m \sum_{l=0}^{k+s-m-1} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_l}^l \\ &\quad \times \left( \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1} i_k}^{*,k} e_{i_k}) \cdot D \mathbf{a} (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1} j_s}^{s-1} e_{j_s}) \right) \\ &\quad + (-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1} i_n}^{*,n} - \sigma_{j_1 \dots i_{n-1} i_n}^{*,n}) e_{i_n} \cdot D (\nabla w - E^m [\nabla \bar{w}]) \right) \\ &\quad + (-1)^n \sum_{s=1}^m \sum_{l=0}^{n+s-m-1} \binom{n}{l} (\nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_l}^l \\ &\quad \times \left( \text{Sym}_{i_1 \dots i_n} ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1} i_n}^{*,n} - \sigma_{j_1 \dots i_{n-1} i_n}^{*,n}) e_{i_n}) \cdot D (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1} j_s}^{s-1} e_{j_s}) \right) \\ &\quad + \sum_{k=1}^{n-1} (-1)^k \sum_{s=1}^m \sum_{l=0}^{k+s-m-1} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_l}^l \\ &\quad \times \left( \text{Sym}_{i_1 \dots i_k} (\bar{\mathbf{a}}_{j_1 \dots i_{k-1} i_k}^{*,k+1} e_{i_k}) \cdot D (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1} j_s}^{s-1} e_{j_s}) \right), \end{aligned} \tag{5.6}$$

where we recall that by a slight abuse of notation we similarly define  $\Xi^n[H]$  as in Definition 3.1 even if  $H$  is not a gradient field.

Let  $n \geq 0$  be fixed. We argue by induction on  $m$ . For  $m = 0$ , the result (5.6) coincides with (5.4) (recall  $E^0[\nabla \bar{w}] = 0$ , cf. Definition 2.6). Now assume that (5.6) holds for some  $m \geq 0$  and let us deduce it at the level  $m + 1$ . Given a polynomial  $\bar{q}$  of order  $m + 1$ , noting that  $\nabla^{m+2} \bar{q}$  vanishes, identity (2.9) ensures that  $\nabla \cdot \mathbf{a} \nabla F^{m+1}[\bar{q}]$  is deterministic. We then apply (5.6) with  $w = F^{m+1}[\bar{q}]$  and  $\bar{w} = \bar{q}$ . Abundantly using that  $\nabla^{m+1} \bar{q}$  is constant and noting that identity (2.5) together with the definition of  $E^m$  yields

$$\begin{aligned} \nabla F^{m+1}[\bar{q}] - E^m[\nabla \bar{q}] &= E^{m+1}[\nabla \bar{q}] - E^m[\nabla \bar{q}] \\ &= (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{q}, \end{aligned}$$

this leads to

$$\begin{aligned}
& e_j \cdot D \Xi^n [\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}] \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{q} \\
&= \sum_{k=0}^n (-1)^k (\nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{q}) \nabla_{i_1 \dots i_k}^k \\
&\times \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot Da (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right) \\
&+ \sum_{k=0}^n (-1)^k \sum_{s=1}^m \binom{k}{k+s-m-1} (\nabla_{i_{k+s-m} \dots i_k}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{q}) \nabla_{i_1 \dots i_{k+s-m-1}}^{k+s-m-1} \\
&\times \left( \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot Da (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right) \\
&+ (-1)^n (\nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{q}) \nabla_{i_1 \dots i_n}^n \\
&\times \left( (a^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot D (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right) \\
&+ (-1)^n \sum_{s=1}^m \binom{n}{n+s-m-1} (\nabla_{i_{n+s-m} \dots i_n}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{q}) \nabla_{i_1 \dots i_{n+s-m-1}}^{n+s-m-1} \\
&\times \left( \text{Sym}_{i_1 \dots i_n} (a^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot D (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right) \\
&+ \sum_{k=1}^{n-1} (-1)^k \sum_{s=1}^m \binom{k}{k+s-m-1} (\nabla_{i_{k+s-m} \dots i_k}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{q}) \nabla_{i_1 \dots i_{k+s-m-1}}^{k+s-m-1} \\
&\times \left( \text{Sym}_{i_1 \dots i_k} (\bar{a}_{j_1 \dots i_{k-1}}^{*,k+1} e_{i_k}) \cdot D (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right).
\end{aligned}$$

Note that both sides of this identity depend linearly on the (constant and deterministic) symmetric tensor  $\nabla^{m+1} \bar{q}$ . Hence, we may replace it by the (deterministic) symmetric tensor  $\nabla^{m+1} \bar{w}$ , to the effect of

$$\begin{aligned}
& e_j \cdot D \Xi^n [\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}] \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w} \\
&= \sum_{k=0}^n (-1)^k (\nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w}) \nabla_{i_1 \dots i_k}^k \\
&\times \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot Da (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right) \\
&+ \sum_{k=0}^n (-1)^k \sum_{s=1}^m \binom{k}{k+s-m-1} (\nabla_{i_{k+s-m} \dots i_k}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_{k+s-m-1}}^{k+s-m-1} \\
&\times \left( \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot Da (\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right) \\
&+ (-1)^n (\nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w}) \nabla_{i_1 \dots i_n}^n \\
&\times \left( (a^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \cdot D (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right)
\end{aligned}$$

$$\begin{aligned}
 &+ (-1)^n \sum_{s=1}^m \binom{n}{n+s-m-1} (\nabla_{i_{n+s-m} \dots i_n}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_{n+s-m-1}}^{n+s-m-1} \\
 &\times \left( \text{Sym}_{i_1 \dots i_n} \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D(\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right) \\
 &+ \sum_{k=1}^{n-1} (-1)^k \sum_{s=1}^m \binom{k}{k+s-m-1} (\nabla_{i_{k+s-m} \dots i_k}^{m+1-s} \nabla_{j_1 \dots j_s}^s \bar{w}) \nabla_{i_1 \dots i_{k+s-m-1}}^{k+s-m-1} \\
 &\times \left( \text{Sym}_{i_1 \dots i_k} \left( \bar{\mathbf{a}}_{j_1 \dots i_{k-1}}^{*,k+1} e_{i_k} \right) \cdot D(\nabla \varphi_{j_1 \dots j_s}^s + \varphi_{j_1 \dots j_{s-1}}^{s-1} e_{j_s}) \right). \tag{5.7}
 \end{aligned}$$

By definition of  $E^m$  and  $\Xi^n$  (cf. Definitions 2.6 and 3.1) together with Lemma 2.4, using the general Leibniz rule, we find

$$\begin{aligned}
 &e_j \cdot D \Xi^n [\nabla w - E^{m+1}[\nabla \bar{w}]] \\
 &= e_j \cdot D \Xi^n [\nabla w - E^m[\nabla \bar{w}]] - e_j \cdot D \Xi^n \left[ (\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w} \right] \\
 &= e_j \cdot D \Xi^n [\nabla w - E^m[\nabla \bar{w}]] - e_j \cdot D \Xi^n \left[ \nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}} \right] \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w} \\
 &+ \sum_{k=1}^{n-1} (-1)^k \sum_{l=0}^{k-1} \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 &\times \left( \text{Sym}_{i_1 \dots i_k} \left( \bar{\mathbf{a}}_{j_1 \dots i_{k-1}}^{*,k+1} e_{i_k} \right) \cdot D(\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right).
 \end{aligned}$$

Injecting identities (5.6) and (5.7) into this equality, using again the definition of  $E^m$  (cf. Definition 2.6) and the general Leibniz rule in the forms

$$\begin{aligned}
 &\sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D\mathbf{a}(\nabla w - E^m[\nabla \bar{w}]) \right) \\
 &= \sum_{k=0}^n (-1)^k \nabla_{i_1 \dots i_k}^k \left( (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot D\mathbf{a}(\nabla w - E^{m+1}[\nabla \bar{w}]) \right) \\
 &+ \sum_{k=0}^n (-1)^k \sum_{l=0}^k \binom{k}{l} (\nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 &\times \left( \text{Sym}_{i_1 \dots i_k} \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k} \right) \cdot D\mathbf{a}(\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right),
 \end{aligned}$$

and

$$\begin{aligned}
 &(-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D(\nabla w - E^m[\nabla \bar{w}]) \\
 &= (-1)^n \nabla_{i_1 \dots i_n}^n \left( (\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D(\nabla w - E^{m+1}[\nabla \bar{w}])
 \end{aligned}$$

$$\begin{aligned}
 &+ (-1)^n \sum_{l=0}^n \binom{n}{l} (\nabla_{i_1 \dots i_k}^{n-l} \nabla_{j_1 \dots j_{m+1}}^{m+1} \bar{w}) \nabla_{i_1 \dots i_l}^l \\
 &\times \left( \text{Sym}_{i_1 \dots i_n} \left( (\mathbf{a}^* \varphi_{j_1 \dots j_{n-1}}^{*,n} - \sigma_{j_1 \dots j_{n-1}}^{*,n}) e_{i_n} \right) \cdot D(\nabla \varphi_{j_1 \dots j_{m+1}}^{m+1} + \varphi_{j_1 \dots j_m}^m e_{j_{m+1}}) \right),
 \end{aligned} \tag{5.8}$$

and recombining the terms, the result (5.6) follows at level  $m + 1$ . For instance, the increment of the first triple sum over  $(k, s, l)$  splits into three contributions,

$$\begin{aligned}
 &\sum_{k=0}^n \sum_{s=1}^{m+1} \sum_{l=0}^{k+s-(m+1)-1} - \sum_{k=0}^n \sum_{s=1}^m \sum_{l=0}^{k+s-m-1} \\
 &= \sum_{k=0}^n \sum_{l=0}^k \mathbb{1}_{s=m+1} - \sum_{k=0}^n \mathbb{1}_{s=m+1} - \sum_{k=0}^n \sum_{s=1}^m \mathbb{1}_{l=k+s-m-1}.
 \end{aligned}$$

Step 3. Conclusion.

Applying (5.4) yields (5.1). We now turn to (5.3) and first note that by definition of  $\Xi^{\circ,n}$  (cf. Definition 3.4) and by (2.5) we have

$$(\Xi^n [E^n [\nabla \bar{u}^n]] - \Xi^{\circ,n} [\nabla \bar{u}^n])(x) = \Xi^n [E^n [\nabla (\bar{u}^n - T_x^n \bar{u}^n)]](x),$$

so that by definition of  $E^n$  (cf. Definition 2.6) and by (3.6) we obtain

$$\begin{aligned}
 &e_j \cdot (\Xi^n [E^n [\nabla \bar{u}^n]] - \Xi^{\circ,n} [\nabla \bar{u}^n]) \\
 &= - \sum_{k=1}^{n-1} (-1)^k \sum_{s=0}^{n-1} \sum_{l=0}^{k+s-n} \binom{k}{l} (\nabla_{i_1 \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{u}^n) \nabla_{i_1 \dots i_l}^l \\
 &\quad \times \left( \text{Sym}_{i_1 \dots i_k} (\bar{\mathbf{a}}_{j_1 \dots j_{k-1}}^{*,k+1} e_{i_k}) \cdot (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right).
 \end{aligned}$$

We now apply  $D$  to this identity, add (5.6) with  $m = n$ , with  $(w, \bar{w})$  replaced by  $(u, \bar{u}^n)$ , and with  $s$  replaced by  $s + 1$ , and so obtain (5.3). Finally, (5.2) follows by subtracting (5.3) (with  $(u, \bar{u}^n) = (w, \bar{w})$ ) from (5.1) and inserting the formula for  $DE^n [\nabla \bar{u}^n]$  already used in (5.8) for  $m = n$ . □

### 5.2 Proof of Proposition 5.2

The conclusion follows by repeating the proof of Proposition 5.1, taking advantage of the same algebraic identities, but now further keeping track of the mixed term in  $D^2(\mathbf{a} \nabla w) = \mathbf{a} \nabla D^2 w + 2D\mathbf{a} \nabla D w + D^2 \mathbf{a} \nabla w$ . □

### 6 Annealed Calderón–Zygmund theory

In this section, we establish the following new annealed Calderón–Zygmund estimate for linear elliptic equations in divergence form with random coefficients. This constitutes a useful upgrade of the quenched large-scale Calderón–Zygmund estimates of [3,4,21]. The merit of (6.2)–(6.4) below is that the stochastic  $L^q$  norm is *inside* the spatial  $L^p$  norm; while there is a tiny loss in stochastic integrability, the spatial integrability is as in the constant-coefficient case.

**Theorem 6.1** *Consider the Helmholtz projection  $T := \nabla(\nabla \cdot \mathbf{a}\nabla)^{-1}\nabla \cdot \mathbf{a}$ . There exists a  $\frac{1}{8}$ -Lipschitz stationary field  $r_* \geq 1$  on  $\mathbb{R}^d$  such that for all  $h \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$  and  $1 < q \leq p < \infty$  there holds*

$$\left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |Th|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}} \lesssim_{p,q} \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}}, \tag{6.1}$$

where  $B_*(x) := B_{r_*(x)}(x)$ . In addition, in the Gaussian setting with integrable correlations (1.5), the stationary random field  $r_*$  satisfies  $\mathbb{E} \left[ \exp\left(\frac{1}{C}r_*^d\right) \right] \leq 2$  for some  $C \simeq 1$ . In particular, for all  $1 < p, q < \infty$  and  $0 < \delta \leq \frac{1}{2}$ ,

$$\| [Th]_2 \|_{L^p(\mathbb{R}^d; L^q(\Omega))} \lesssim_{p,q} \delta^{-\left(\frac{1}{p \wedge q \wedge 2} - \frac{1}{p \vee q \vee 2}\right)} |\log \delta|^{2\left|\frac{1}{q} - \frac{1}{p}\right|} \| [h]_2 \|_{L^p(\mathbb{R}^d; L^{q+\delta}(\Omega))}. \tag{6.2}$$

◇

We shall also make use of the following consequence of this annealed estimate for the corresponding Riesz potentials. As opposed to the above, this can alternatively be deduced from the annealed Green’s function estimates in [37].

**Corollary 6.2** *Consider the Riesz potentials  $U_{e,\circ} := \nabla \Delta^{-1}e$  and  $U_e := \nabla(\nabla \cdot \mathbf{a}\nabla)^{-1}e \cdot \mathbf{a}$  with  $|e| = 1$ . For all  $h \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$  and  $1 < p, q < \infty$  with  $\frac{dp}{d+p} > 1$ ,*

$$\| [U_{e,\circ}h]_2 \|_{L^p(\mathbb{R}^d; L^q(\Omega))} \lesssim_p \| [h]_2 \|_{L^{\frac{dp}{d+p}}(\mathbb{R}^d; L^q(\Omega))}, \tag{6.3}$$

and for all  $0 < \delta \leq \frac{1}{2}$ ,

$$\| [U_e h]_2 \|_{L^p(\mathbb{R}^d; L^q(\Omega))} \lesssim_{p,q} \delta^{-\left(\frac{1}{p \wedge q \wedge 2} - \frac{1}{p \vee q \vee 2}\right)} |\log \delta|^{2\left|\frac{1}{q} - \frac{1}{p}\right|} \| [h]_2 \|_{L^{\frac{dp}{d+p}}(\mathbb{R}^d; L^{q+\delta}(\Omega))} \tag{6.4}$$

◇

### 6.1 Proof of Theorem 6.1

Our argument is based on a multiple use of the following version of the Calderón–Zygmund lemma due to Shen [48, Theorem 2.1] (cf. also [49, Theorem 2.4]), based on ideas by Caffarelli and Peral [13]. For a ball  $B$ , we set  $B = B(x_B, r_B)$  and abusively write  $\alpha B := B(x_B, \alpha r_B)$  for  $\alpha > 0$ .

**Lemma 6.3** [13,48] *Let  $1 \leq p_0 < p_1 \leq \infty$ ,  $C_0 > 0$ , and  $f, g \in L^{p_0} \cap L^{p_1}(\mathbb{R}^d)$ . Assume that for all balls  $B \subset \mathbb{R}^d$  there exist measurable functions  $f_{B,0}$  and  $f_{B,1}$  such that there hold  $f = f_{B,0} + f_{B,1}$  on  $B$  and*

$$\begin{aligned} \left( \int_B |f_{B,0}|^{p_0} \right)^{\frac{1}{p_0}} &\leq C_0 \left( \int_{C_0 B} |g|^{p_0} \right)^{\frac{1}{p_0}}, \\ \left( \int_{\frac{1}{C_0} B} |f_{B,1}|^{p_1} \right)^{\frac{1}{p_1}} &\leq C_0 \left( \int_B |f_{B,1}|^{p_0} \right)^{\frac{1}{p_0}}. \end{aligned}$$

Then, for all  $p_0 < q < p_1$ ,

$$\left( \int_{\mathbb{R}^d} |f|^q \right)^{\frac{1}{q}} \lesssim_{C_0, p_0, q, p_1} \left( \int_{\mathbb{R}^d} |g|^q \right)^{\frac{1}{q}}. \quad \diamond$$

Combining this lemma with the quenched large-scale Lipschitz regularity theory due to [6,7,21] leads to the following quenched large-scale Calderón–Zygmund estimate. This slightly shortens the proof of [4] and [21, Corollary 4] (cf. also [3, Section 7]).

**Proposition 6.4** (Quenched large-scale Calderón–Zygmund estimate) *There exists a stationary random field  $r_*$  as in the statement of Theorem 6.1 such that for all  $h \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$  and  $1 < p < \infty$ ,*

$$\left( \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |Th|^2 \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \lesssim_p \left( \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |h|^2 \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}},$$

where we recall the notation  $B_*(x) := B_{r_*(x)}(x)$ . \(\diamond\)

Before proceeding to the proof, we state the following useful properties of averages on the family of balls  $\{B_*(x)\}_{x \in \mathbb{R}^d}$ . The point is that these balls have a radius that varies with their center  $x$ .

**Lemma 6.5** *For any measurable function  $f$  on  $\mathbb{R}^d$ , the following estimates hold.*

(i) *For all balls  $B \subset \mathbb{R}^d$  with  $r_B \geq \frac{1}{4}r_*(x_B)$ , we have*

$$\int_B |f| \lesssim \int_{2B} \left( \int_{B_*(x)} |f| \right) dx, \tag{6.5}$$

hence, in particular,

$$\int_{\mathbb{R}^d} |f| \simeq \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |f| \right) dx. \tag{6.6}$$

(ii) For all balls  $B \subset \mathbb{R}^d$  and  $x_0 \in B$  with  $r_B \leq \frac{1}{4}r_*(x_0)$ , we have

$$\left( \int_{B_*(x_0)} |f| \right)^{\frac{1}{2}} \lesssim \int_{5B} \left( \int_{B_*(x)} |f| \right)^{\frac{1}{2}} dx, \tag{6.7}$$

hence, in particular,

$$\int_{B_*(x_0)} |f| \lesssim \int_{5B} \left( \int_{B_*(x)} |f| \right) dx. \tag{6.8}$$

◇

We may now proceed to the proof of Proposition 6.4, which amounts to combining the quenched large-scale Lipschitz regularity theory of [6,7,21] together with Lemma 6.3.

**Proof of Proposition 6.4** As in [21, Step 1 of the proof of Corollary 4], the random field  $r_*$  is chosen as the largest  $\frac{1}{8}$ -Lipschitz lower bound on the so-called minimal radius defined in [21] (what is denoted here by  $r_*$  thus corresponds to the notation  $\underline{r}_*$  in [21]). We split the proof into two steps.

*Step 1.* For all balls  $B \subset \mathbb{R}^d$ , decomposing  $Th = \nabla w_{B,0} + \nabla w_{B,1}$  with

$$-\nabla \cdot \mathbf{a} \nabla w_{B,0} = \nabla \cdot (\mathbf{a}h \mathbb{1}_B), \quad -\nabla \cdot \mathbf{a} \nabla w_{B,1} = \nabla \cdot (\mathbf{a}h \mathbb{1}_{\mathbb{R}^d \setminus B}), \tag{6.9}$$

we prove for all  $2 \leq p \leq \infty$ ,

$$\left( \int_B \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right) dx \right)^{\frac{1}{2}} \lesssim \left( \int_{2B} \left( \int_{B_*(x)} |h|^2 \right) dx \right)^{\frac{1}{2}}, \tag{6.10}$$

$$\left( \int_{\frac{1}{2}B} \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \lesssim \left( \int_B \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right) dx \right)^{\frac{1}{2}}, \tag{6.11}$$

We start with the proof of (6.10), and we first consider the case  $r_B \geq \frac{1}{4}r_*(x_B)$ . An energy estimate on the equation for  $w_{B,0}$  yields

$$\int_{\mathbb{R}^d} |\nabla w_{B,0}|^2 \lesssim \int_B |h|^2, \tag{6.12}$$

hence, combined with (6.5) and (6.6),

$$\begin{aligned} \int_B \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right) dx &\leq |B|^{-1} \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right) dx \\ &\lesssim \int_B |h|^2 \lesssim \int_{2B} \left( \int_{B_*(x)} |h|^2 \right) dx, \end{aligned}$$

which proves (6.10). Next, we consider the case  $r_B \leq \frac{1}{4}r_*(x_B)$ . For  $x_0 \in B$ , using the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  in form of  $|B_*(x_0)| \simeq |B_*(x_B)|$ , using the energy estimate (6.12), and applying (6.8) (with  $x_0$  and  $B$  replaced by  $x_B$  and  $\frac{1}{10}B$ ), we find

$$\int_{B_*(x_0)} |\nabla w_{B,0}|^2 \lesssim |B_*(x_B)|^{-1} \int_B |h|^2 \lesssim \int_{B_*(x_B)} |h|^2 \lesssim \int_{\frac{1}{2}B} \left( \int_{B_*(x)} |h|^2 \right) dx. \tag{6.13}$$

(The choice of  $\frac{1}{2}B$  in the right-hand side does not matter here and is only made for later reference.) Integrating this estimate over  $x_0 \in B$ , the conclusion (6.10) follows. We turn to the proof of (6.11), which is a consequence of quenched large-scale Lipschitz regularity theory, noting that  $w_{B,1}$  is indeed  $\alpha$ -harmonic in  $B$ . It is obviously sufficient to prove for all  $x_0 \in \frac{1}{12}B$ ,

$$\int_{B_*(x_0)} |\nabla w_{B,1}|^2 \lesssim \int_B \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right) dx. \tag{6.14}$$

If  $r_*(x_0) \geq \frac{1}{3}r_B$ , this follows from (6.8) (with  $B$  replaced by  $\frac{1}{12}B$ ). Let now  $x_0 \in \frac{1}{12}B$  be fixed with  $r_*(x_0) \leq \frac{1}{3}r_B$ . Since the ball  $B(x_0, r_*(x_0) + \frac{1}{12}r_B)$  is contained in  $\frac{1}{2}B$ , where  $w_{B,1}$  is  $\alpha$ -harmonic, the quenched large-scale Lipschitz regularity theory (e.g. [21, Theorem 1]) implies

$$\int_{B_*(x_0)} |\nabla w_{B,1}|^2 \lesssim \int_{B(x_0, r_*(x_0) + \frac{1}{12}r_B)} |\nabla w_{B,1}|^2 \lesssim \int_{\frac{1}{2}B} |\nabla w_{B,1}|^2.$$

Since for  $r_*(x_0) \leq \frac{1}{3}r_B$  the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  yields  $r_*(x_B) \leq r_*(x_0) + \frac{1}{96}r_B \leq 2r_B$ , the conclusion (6.14) follows from (6.5) (with  $B$  replaced by  $\frac{1}{2}B$ ).

*Step 2. Conclusion.*

Combining (6.10) and (6.11), and applying Lemma 6.3 with

$$\begin{aligned} p_0 = 2 \leq p_1 \leq \infty, \quad f(x) &= \left( \int_{B_*(x)} |Th|^2 \right)^{\frac{1}{2}}, \quad g(x) = \left( \int_{B_*(x)} |h|^2 \right)^{\frac{1}{2}}, \\ f_{B,0}(x) &= \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{1}{2}}, \quad f_{B,1}(x) = \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}}, \end{aligned}$$



we conclude for all  $2 \leq p < \infty$ ,

$$\left( \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |Th|^2 \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}} \lesssim_p \left( \int_{\mathbb{R}^d} \left( \int_{B_*(x)} |h|^2 \right)^{\frac{p}{2}} dx \right)^{\frac{1}{p}}.$$

A standard duality argument allows to deduce the corresponding result for  $1 < p \leq 2$  (e.g. [21, Step 7 of the proof of Corollary 4]).  $\square$

We need two additional ingredients for the proof of Theorem 6.1. The first is a deterministic regularity result for  $\alpha$ -harmonic functions. It follows e.g. from [10, proof of Lemma 4], but a short proof is included below for the reader’s convenience.

**Lemma 6.6** *If  $w$  is  $\alpha$ -harmonic on a ball  $B$ , there holds*

$$\left( \int_{\frac{1}{2}B} |\nabla w|^2 \right)^{\frac{1}{2}} \lesssim \int_B |\nabla w|. \quad \diamond$$

The second ingredient is needed to pass from statements on averages at the scale  $r_*$  to corresponding statements on unit scale. This can be done at no loss in the spatial, but a small controlled loss in the stochastic integrability.

**Lemma 6.7** *Let the stationary random field  $r_*$  satisfy  $\mathbb{E} \left[ \exp\left(\frac{1}{C} r_*^d\right) \right] \leq 2$  for some  $C \simeq 1$ . For all  $f \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))$  and  $1 \leq q \leq p < \infty$ , there hold*

(i) *for all  $r > q$ ,*

$$\begin{aligned} & \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}} \\ & \lesssim_{p,q} \left( \frac{1}{q} - \frac{1}{r} \right)^{-\left(\frac{1}{q}-\frac{1}{2}\right)_+} L \left( \frac{1}{q} - \frac{1}{r} \right)^{\frac{1}{q}-\frac{1}{p}} \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^{\frac{p}{r}} \right] \right)^{\frac{1}{p}}; \end{aligned}$$

(ii) *for all  $r < q$ ,*

$$\begin{aligned} & \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}} \\ & \gtrsim_{p,q} \left( \frac{1}{r} - \frac{1}{q} \right)^{\left(\frac{1}{2}-\frac{1}{p}\right)_+} L \left( \frac{1}{r} - \frac{1}{q} \right)^{-\left(\frac{1}{q}-\frac{1}{p}\right)} \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^{\frac{p}{r}} \right] \right)^{\frac{1}{p}}; \end{aligned}$$

where we have set  $L(t) := \log(2 + \frac{1}{t})$ .  $\diamond$

With these ingredients at hand, we turn the proof of Theorem 6.1. The argument is based on a second application of Lemma 6.3 starting with the quenched large-scale Calderón–Zygmund estimate of Proposition 6.4.

**Proof of Theorem 6.1** Let  $r_*$  be chosen as in the proof of Proposition 6.4 above. We split the proof into two steps.

*Step 1.* For all balls  $B \subset \mathbb{R}^d$ , decomposing  $Th = \nabla w_{B,0} + \nabla w_{B,1}$  as in (6.9), we prove for all  $1 < q < \infty$  and  $1 \leq p_1 \leq \infty$ ,

$$\left( \int_B \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{q}{2}} dx \right]^{\frac{1}{q}} \lesssim_q \left( \int_{6B} \mathbb{E} \left[ \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} dx \right]^{\frac{1}{q}}, \tag{6.15}$$

$$\left( \int_{\frac{1}{24}B} \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{q}{2}} dx \right]^{\frac{1}{p_1}} \lesssim_q \left( \int_B \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{q}{2}} dx \right]^{\frac{1}{q}}. \tag{6.16}$$

We start with the proof of (6.15). It suffices to show

$$\int_B \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{q}{2}} dx \lesssim_q \int_{6B} \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} dx, \tag{6.17}$$

since the conclusion (6.15) then follows by taking the expectation. First consider the case  $r_B \leq \frac{1}{4}r_*(x_B)$ . Integrating (6.13) over  $x_0 \in B$  yields

$$\int_B \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{q}{2}} dx \lesssim_q \left( \int_{\frac{1}{2}B} \left( \int_{B_*(x)} |h|^2 \right) dx \right)^{\frac{q}{2}}. \tag{6.18}$$

The claim (6.17) follows from this estimate together with Jensen’s inequality if  $q \geq 2$ , but a different argument is needed if  $q < 2$ . For  $x_0 \in \frac{1}{2}B$ , since the condition  $r_B \leq \frac{1}{4}r_*(x_B)$  and the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  imply  $r_B \leq \frac{1}{4}r_*(x_0) + \frac{1}{64}r_B$ , hence  $r_B \leq \frac{1}{2}r_*(x_0)$ , it follows from (6.7) (with  $B$  replaced by  $\frac{1}{2}B$ ) that

$$\left( \int_{B_*(x_0)} |h|^2 \right)^{\frac{1}{2}} \lesssim \int_{\frac{5}{2}B} \left( \int_{B_*(x)} |h|^2 \right)^{\frac{1}{2}} dx.$$

Combined with (6.18), this leads to (6.17) by Jensen’s inequality. It remains to consider the case  $r_B \geq \frac{1}{4}r_*(x_B)$ . Applying Proposition 6.4 (with  $q$  playing the role of  $p$ ) to the equation for  $w_{B,0}$  yields

$$\int_{\mathbb{R}^d} \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{q}{2}} dx \lesssim_q \int_{\mathbb{R}^d} \left( \int_{B_*(x)} \mathbb{1}_B |h|^2 \right)^{\frac{q}{2}} dx. \tag{6.19}$$

For  $x \in B$ , the balls  $B_*(x)$  and  $B$  are disjoint whenever  $|x - x_B| \geq r_*(x) + r_B$ . Since the condition  $r_B \geq \frac{1}{4}r_*(x_B)$  and the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  imply

$$r_*(x) + r_B \leq r_*(x_B) + r_B + \frac{1}{8}|x - x_B| \leq 5r_B + \frac{1}{8}|x - x_B|,$$

we deduce that  $B_*(x)$  and  $B$  are disjoint whenever  $|x - x_B| \geq 5r_B + \frac{1}{8}|x - x_B|$ , that is, whenever  $|x - x_B| \geq 6r_B$ . This implies

$$\int_{\mathbb{R}^d} \left( \int_{B_*(x)} \mathbb{1}_B |h|^2 \right)^{\frac{q}{2}} dx \leq \int_{6B} \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} dx,$$

so that the claim (6.17) follows from (6.19).

We turn to the proof of (6.16). For  $2 \leq q \leq \infty$ , taking the  $\frac{q}{2}$ th power of (6.14), using Jensen’s inequality, and taking the expectation, we easily obtain (6.16). In order to conclude for all values  $1 \leq q \leq \infty$ , it suffices to establish the following improved version of (6.14): for all  $x_0 \in \frac{1}{24}B$ ,

$$\left( \int_{B_*(x_0)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}} \lesssim \int_B \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}} dx. \tag{6.20}$$

If  $r_*(x_0) \geq \frac{1}{6}r_B$ , this already follows from (6.7) (with  $B$  replaced by  $\frac{1}{24}B$ ). Let now  $x_0 \in \frac{1}{24}B$  be fixed with  $r_*(x_0) \leq \frac{1}{6}r_B$ . Since  $w_{B,1}$  is  $\alpha$ -harmonic on  $B$  and since the ball  $B(x_0, r_*(x_0) + \frac{1}{24}r_B)$  is contained in  $\frac{1}{4}B$ , the quenched large-scale Lipschitz regularity theory (e.g. [21, Theorem 1]) implies

$$\left( \int_{B_*(x_0)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{B(x_0, r_*(x_0) + \frac{1}{24}r_B)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{\frac{1}{4}B} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}},$$

and Lemma 6.6 (with  $B$  replaced by  $\frac{1}{2}B$ ) then leads to

$$\left( \int_{B_*(x_0)} |\nabla w_{B,1}|^2 \right)^{\frac{1}{2}} \lesssim \int_{\frac{1}{2}B} |\nabla w_{B,1}|.$$

Since for  $r_*(x_0) \leq \frac{1}{6}r_B$  the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  yields  $r_*(x_B) \leq r_*(x_0) + \frac{1}{192}r_B \leq 2r_B$ , the conclusion (6.20) follows from (6.5) (with  $B$  replaced by  $\frac{1}{2}B$ ) and from Jensen’s inequality.

**Step 2. Conclusion.**

Combining (6.15) and (6.16), and applying Lemma 6.3 with

$$1 < p_0 = q \leq p_1 < \infty, \quad f(x) = \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w|^2 \right)^{\frac{q}{2}} \right]^{\frac{1}{q}}, \quad g(x) = \mathbb{E} \left[ \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} \right]^{\frac{1}{q}},$$

$$f_{B,0}(x) = \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w_{B,0}|^2 \right)^{\frac{q}{2}} \right]^{\frac{1}{q}}, \quad f_{B,1}(x) = \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w_{B,1}|^2 \right)^{\frac{q}{2}} \right]^{\frac{1}{q}},$$

we deduce for all  $1 < q \leq p < \infty$ ,

$$\left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |\nabla w|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}} \lesssim_{p,q} \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |h|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \right)^{\frac{1}{p}},$$

which is the main stated result (6.1). It remains to post-process this result by rewriting both sides of the inequality: combining (6.1) with the different bounds stated in Lemma 6.7, and noting that for  $q \leq p$  there holds  $(\frac{1}{q} - \frac{1}{2})_+ + (\frac{1}{2} - \frac{1}{p})_+ = \frac{1}{q \wedge 2} - \frac{1}{p \vee 2}$ , the conclusion (6.2) follows for  $q \leq p$ , while the corresponding result for  $p \leq q$  is deduced by duality.  $\square$

**6.2 Proof of Corollary 6.2**

For  $d > 1$ , appealing to the integral representation

$$U_{e, \circ} g(x) = C \int_{\mathbb{R}^d} \frac{y}{|y|^d} e \cdot g(x - y) dy,$$

the triangle inequality yields

$$\| [U_{e, \circ} g]_2 \|_{L^p(\mathbb{R}^d; L^q(\Omega))} \lesssim \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \frac{\mathbb{E}[|g|_2^q(x - y)]^{\frac{1}{q}}}{|y|^{d-1}} dy \right)^p dx \right)^{\frac{1}{p}},$$

and the result (6.3) follows from the Hardy–Littlewood–Sobolev inequality. Next, writing  $U_e = T a^{-1} U_{e, \circ} a$ , the estimate (6.4) follows from the combination of (6.2) and (6.3).  $\square$

**6.3 Proof of Lemma 6.5**

We start with the proof of (i), and thus let  $r_B \geq \frac{1}{4} r_*(x_B)$ . For  $|x - z| \leq \frac{1}{2} r_*(z)$ , the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  implies  $|x - z| \leq \frac{1}{2} r_*(x) + \frac{1}{16} |x - z|$ , hence  $|x - z| \leq r_*(x)$ , and similarly  $|B_*(x)| \simeq |B_*(z)|$ . These observations lead to the lower bound

$$\begin{aligned} \int_{2B} \left( \int_{B_*(x)} |f| \right) dx &= \int_{\mathbb{R}^d} |f(z)| \left( \int_{2B} \frac{\mathbb{1}_{|x-z| \leq r_*(x)}}{|B_*(x)|} dx \right) dz \\ &\geq \int_{\mathbb{R}^d} |f(z)| \left( \int_{2B} \frac{\mathbb{1}_{|x-z| \leq \frac{1}{2} r_*(z)}}{|B_*(x)|} dx \right) dz \gtrsim \int_{\mathbb{R}^d} |f(z)| \frac{|2B \cap \frac{1}{2} B_*(z)|}{|B_*(z)|} dz. \end{aligned} \tag{6.21}$$

Since  $r_*(x_B) \leq 4r_B$ , we note that for all  $z \in B$  the ball  $\frac{1}{5} B_*(z)$  is included in

$$B(x_B, r_B + \frac{1}{5} r_*(z)) \subset B(x_B, r_B + \frac{1}{5} r_*(x_B) + \frac{1}{40} r_B) \subset 2B,$$

and (6.5) follows. The upper bound in (6.6) is a consequence (6.5) with  $r_B \uparrow \infty$ , so that it remains to establish the lower bound. For that purpose, we write again

$$\int_{\mathbb{R}^d} \left( \int_{B_*(x)} |f| \right) dx = \int_{\mathbb{R}^d} |f(z)| \left( \int_{\mathbb{R}^d} \frac{\mathbb{1}_{|x-z| \leq r_*(x)}}{|B_*(x)|} dx \right) dz.$$

Since for  $|x - z| \leq r_*(x)$  the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  implies  $r_*(x) \simeq r_*(z)$ , the last integral in bracket is  $\lesssim 1$ , and the desired lower bound follows.

We turn to the proof of (ii). By Jensen’s inequality, it suffices to establish (6.7). Let  $x_0 \in B$  with  $r_B \leq \frac{1}{4}r_*(x_0)$ . Noting that for all  $x \in B$  the ball  $B_*(x_0)$  is contained in  $B(x, r_*(x_0) + 2r_B)$ , we find

$$\left( \int_{B_*(x_0)} |f| \right)^{\frac{1}{2}} \lesssim \int_B \left( \int_{B(x, r_*(x_0) + 2r_B)} |f| \right)^{\frac{1}{2}} dx.$$

The ball  $B_{4r_B}$  at the origin can be covered by  $N \simeq 1$  balls of radius  $r_B$ . Denote by  $(z_i)_{i=1}^N \subset B_{4r_B}$  their centers. The collection  $(B(z_i, r_*(x_0) - r_B))_{i=1}^N$  then covers  $B_{r_*(x_0) + 2r_B}$ . By subadditivity of the square root, we may decompose

$$\begin{aligned} \left( \int_{B_*(x_0)} |f| \right)^{\frac{1}{2}} &\lesssim \int_B \left( \sum_{i=1}^N \int_{B(x+z_i, r_*(x_0) - r_B)} |f| \right)^{\frac{1}{2}} dx \\ &\lesssim \sum_{i=1}^N \int_B \left( \int_{B(x+z_i, r_*(x_0) - r_B)} |f| \right)^{\frac{1}{2}} dx, \end{aligned}$$

hence, noting that  $z_i + B \subset 5B$  for all  $i$ ,

$$\left( \int_{B_*(x_0)} |f| \right)^{\frac{1}{2}} \lesssim \int_{5B} \left( \int_{B(x, r_*(x_0) - r_B)} |f| \right)^{\frac{1}{2}} dx.$$

For  $x \in 5B$ , the  $\frac{1}{8}$ -Lipschitz property of  $r_*$  yields

$$r_*(x_0) - r_B \leq r_*(x) + \frac{1}{8}|x - x_0| - r_B \leq r_*(x) + \frac{3}{4}r_B - r_B \leq r_*(x),$$

and (iii) follows. □

**6.4 Proof of Lemma 6.6**

By scaling and translation invariance, we may assume that  $B$  is the unit ball at the origin and that  $\int_B w = 0$ . Choose a smooth cut-off function  $\chi$  with  $\chi = 1$  in  $\frac{1}{2}B$  and  $\chi = 0$  outside  $B$ . Caccioppoli’s inequality yields

$$\begin{aligned} \left( \int_{\mathbb{R}^d} |\nabla(\chi w)|^2 \right)^{\frac{1}{2}} &\leq \left( \int_{\mathbb{R}^d} \chi^2 |\nabla w|^2 \right)^{\frac{1}{2}} \\ &+ \left( \int_{\mathbb{R}^d} |\nabla \chi|^2 |w|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{\mathbb{R}^d} |\nabla \chi|^2 |w|^2 \right)^{\frac{1}{2}}. \end{aligned} \tag{6.22}$$

Sobolev’s inequality combined with Poincaré’s inequality (with vanishing boundary condition) then implies for some suitable  $p = p(d) > 2$ ,

$$\begin{aligned} \left(\int_{\mathbb{R}^d} |\chi w|^p\right)^{\frac{1}{p}} &\lesssim \left(\int_{\mathbb{R}^d} |\nabla(\chi w)|^2 + |\chi w|^2\right)^{\frac{1}{2}} \lesssim \left(\int_{\mathbb{R}^d} |\nabla(\chi w)|^2\right)^{\frac{1}{2}} \\ &\stackrel{(6.22)}{\lesssim} \left(\int_{\mathbb{R}^d} |\nabla\chi|^2 |w|^2\right)^{\frac{1}{2}}. \end{aligned} \tag{6.23}$$

Choosing  $q := 2\frac{p-1}{p-2}$  and writing  $\chi = \xi^q$ , we deduce by Hölder’s inequality,

$$\begin{aligned} \left(\int_{\mathbb{R}^d} |\nabla\chi|^2 |w|^2\right)^{\frac{1}{2}} &\lesssim \left(\int_{\mathbb{R}^d} |\xi^{q-1} w|^2\right)^{\frac{1}{2}} = \left(\int_{\mathbb{R}^d} |\xi^q w|^{\frac{p}{p-1}} |w|^{\frac{p-2}{p-1}}\right)^{\frac{1}{2}} \\ &\leq \left(\int_{\mathbb{R}^d} |\chi w|^p\right)^{\frac{1}{2(p-1)}} \left(\int_B |w|\right)^{\frac{1}{q}} \stackrel{(6.23)}{\lesssim} \left(\int_{\mathbb{R}^d} |\nabla\chi|^2 |w|^2\right)^{\frac{q-1}{2q}} \left(\int_B |w|\right)^{\frac{1}{q}}, \end{aligned}$$

hence, by Young’s inequality,

$$\left(\int_{\mathbb{R}^d} |\nabla\chi|^2 |w|^2\right)^{\frac{1}{2}} \lesssim \int_B |w|.$$

Injecting this into (6.22) and using Poincaré’s inequality (with vanishing mean-value), we conclude

$$\left(\int_{\frac{1}{2}B} |\nabla w|^2\right)^{\frac{1}{2}} \leq \left(\int_{\mathbb{R}^d} |\nabla(\chi w)|^2\right)^{\frac{1}{2}} \lesssim \int_B |w| \lesssim \int_B |\nabla w|,$$

as desired. □

### 6.5 Proof of Lemma 6.7

*Step 1.* Proof of the upper bound (i).

Let  $1 \leq q \leq p < \infty$ . We start with the following deterministic version of the claimed upper bound: for all  $R \geq 1$  we have

$$\left(\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_{B_R(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx\right)^{\frac{1}{p}} \lesssim R^{d(\frac{1}{q}-\frac{1}{2})+} \left(\int_{\mathbb{R}^d} \mathbb{E} [ [f]_2^q ]^{\frac{p}{q}}\right)^{\frac{1}{p}}. \tag{6.24}$$

If  $q \geq 2$ , Jensen’s inequality on  $t \mapsto t^{\frac{q}{2}}$  yields

$$\left(\int_{B_R(x)} |f|^2\right)^{\frac{q}{2}} \lesssim \left(\int_{B_R(x)} [f]_2^2\right)^{\frac{q}{2}} \lesssim \int_{B_R(x)} [f]_2^q. \tag{6.25}$$

If  $q \leq 2$ , covering the ball  $B_R$  by  $N \simeq R^d$  balls of radius  $\frac{1}{2}$  with centers  $(z_i)_{i=1}^N \subset B_{R-\frac{1}{4}}$  such that  $\min_{i \neq j} |z_i - z_j| \geq \frac{1}{2}$ , it follows from the discrete  $\ell^2$ - $\ell^q$  inequality that

$$\begin{aligned} \left(\int_{B_R(x)} |f|^2\right)^{\frac{q}{2}} &\leq \left(\sum_{i=1}^N \int_{B_{\frac{1}{2}}(x+z_i)} |f|^2\right)^{\frac{q}{2}} \leq \sum_{i=1}^N \left(\int_{B_{\frac{1}{2}}(x+z_i)} |f|^2\right)^{\frac{q}{2}} \\ &\leq \sum_{i=1}^N \int_{B_{\frac{1}{4}}(x+z_i)} \left(\int_{B(y)} |f|^2\right)^{\frac{q}{2}} dy \lesssim \sum_{i=1}^N \int_{B_{\frac{1}{4}}(x+z_i)} [f]_2^q, \end{aligned}$$

and hence, since the balls  $(B_{\frac{1}{4}}(x + z_i))_{i=1}^N$  are disjoint and included in  $B_R(x)$ ,

$$\left(\int_{B_R(x)} |f|^2\right)^{\frac{q}{2}} \lesssim \int_{B_R(x)} [f]_2^q. \tag{6.26}$$

These two estimates (6.25) and (6.26) lead to

$$\left(\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_{B_R(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx\right)^{\frac{1}{p}} \lesssim R^{d(\frac{1}{q}-\frac{1}{2})+} \left(\int_{\mathbb{R}^d} \mathbb{E}\left[\int_{B_R(x)} [f]_2^q dx\right)^{\frac{p}{q}}\right)^{\frac{1}{p}},$$

and the claim (6.24) follows from Jensen’s inequality on  $t \mapsto t^{\frac{p}{q}}$ .

We turn to the proof of (i). Conditioning with respect to the value of  $r_*(x)$  on dyadic scale and using Jensen’s inequality on  $t \mapsto t^{\frac{p}{q}}$  with  $\sum_{n=0}^{\infty} 2^{-nd\delta} \simeq \delta^{-1}$ , we find for  $0 < \delta \leq 1$ ,

$$\begin{aligned} &\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_{B_*(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx \\ &\lesssim \int_{\mathbb{R}^d} \mathbb{E}\left[\sum_{n=0}^{\infty} \mathbb{1}_{2^{n-1} < r_*(x) \leq 2^{n+1}} \left(\int_{B_{2^{n+1}}(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx \\ &\lesssim \int_{\mathbb{R}^d} \mathbb{E}\left[\sum_{n=0}^{\infty} 2^{-nd\delta} r_*(x)^{d\delta} \mathbb{1}_{2^{n-1} < r_*(x) \leq 2^{n+1}} \left(\int_{B_{2^{n+1}}(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx \\ &\lesssim \delta^{1-\frac{p}{q}} \sum_{n=0}^{\infty} 2^{nd\delta(\frac{p}{q}-1)} \int_{\mathbb{R}^d} \mathbb{E}\left[\mathbb{1}_{2^{n-1} < r_*(x) \leq 2^{n+1}} \left(\int_{B_{2^{n+1}}(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx. \end{aligned}$$

Taking advantage of the  $\frac{1}{8}$ -Lipschitz property of  $r_*$ , we can rewrite

$$\begin{aligned} &\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_{B_*(x)} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx \\ &\lesssim \delta^{1-\frac{p}{q}} \sum_{n=0}^{\infty} 2^{nd\delta(\frac{p}{q}-1)} \int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_{B_{2^{n+1}}(x)} \mathbb{1}_{2^{n-1} < r_* \leq 2^{n+2}} |f|^2\right)^{\frac{q}{2}}\right]^{\frac{p}{q}} dx, \end{aligned}$$

and applying (6.24) with  $R = 2^{n+1}$  then leads to

$$\int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \lesssim \delta^{1-\frac{p}{q}} \sum_{n=0}^{\infty} (2^{ndp})^{\left(\frac{1}{q}-\frac{1}{2}\right)_+ + \delta\left(\frac{1}{q}-\frac{1}{p}\right)} \int_{\mathbb{R}^d} \mathbb{E} \left[ \mathbb{1}_{2^{n-1}-2 \leq r_* \leq 2^{n+2}} [f]_2^q \right]^{\frac{p}{q}}.$$

Hence, for  $r > q$ , using Hölder’s inequality and the moment bounds on the stationary  $r_*$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \\ & \lesssim \delta^{1-\frac{p}{q}} \sum_{n=0}^{\infty} (2^{ndp})^{\left(\frac{1}{q}-\frac{1}{2}\right)_+ + \delta\left(\frac{1}{q}-\frac{1}{p}\right)} \mathbb{P} \left[ r_* \geq 2^{n-1} - 2 \right]^{p\frac{r-q}{rq}} \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^r \right]^{\frac{p}{r}} \\ & \lesssim_{q,p} \delta^{1-\frac{p}{q}} \left( \frac{rq}{r-q} \right)^{p\left(\left(\frac{1}{q}-\frac{1}{2}\right)_+ + \delta\left(\frac{1}{q}-\frac{1}{p}\right)\right)} \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^r \right]^{\frac{p}{r}}, \end{aligned}$$

where in the last estimate we used  $\sum_n (2^{nd})^\alpha e^{-2^{nd}\theta} \lesssim_{\alpha,d} \theta^{-\alpha}$ . Optimizing  $\frac{1}{\delta}\theta^{-\delta}$  in  $0 < \delta \leq 1$  through  $\delta = (\log \frac{1}{\theta})^{-1}$  yielding  $\sim \log \frac{1}{\theta}$ , the desired upper bound follows.

*Step 2. Proof of the lower bound (ii).*

Let  $1 \leq r \leq p \leq \infty$ . We start with the following deterministic version of the claimed lower bound: for all  $R \geq 1$  we have

$$\left( \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_R(x)} |f|^2 \right)^{\frac{r}{2}} \right]^{\frac{p}{r}} dx \right)^{\frac{1}{p}} \gtrsim R^{-d\left(\frac{1}{2}-\frac{1}{p}\right)_+} \left( \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^r \right]^{\frac{p}{r}} \right)^{\frac{1}{p}}. \tag{6.27}$$

If  $2 \leq r \leq p$ , using the discrete  $\ell^r-\ell^2$  inequality and the discrete  $\ell^p-\ell^r$  inequality, we indeed find

$$\begin{aligned} & \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_R(x)} |f|^2 \right)^{\frac{r}{2}} \right]^{\frac{p}{r}} dx \gtrsim R^{-dp\left(\frac{1}{2}-\frac{1}{r}\right)} \int_{\mathbb{R}^d} \mathbb{E} \left[ \int_{B_R(x)} [f]_2^r \right]^{\frac{p}{r}} dx \\ & \gtrsim R^{-dp\left(\frac{1}{2}-\frac{1}{p}\right)} \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^r \right]^{\frac{p}{r}}. \end{aligned}$$

If  $r \leq 2 \leq p$ , the triangle inequality in  $L^{\frac{2}{r}}(B_R(x))$  and the discrete  $\ell^p-\ell^2$  inequality lead to

$$\begin{aligned} & \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_R(x)} |f|^2 \right)^{\frac{r}{2}} \right]^{\frac{p}{r}} dx \gtrsim \int_{\mathbb{R}^d} \left( \int_{B_R(x)} \mathbb{E} \left[ [f]_2^r \right]^{\frac{2}{r}} \right)^{\frac{p}{2}} dx \\ & \gtrsim R^{-dp\left(\frac{1}{2}-\frac{1}{p}\right)} \int_{\mathbb{R}^d} \mathbb{E} \left[ [f]_2^r \right]^{\frac{p}{r}}. \end{aligned}$$



Finally, if  $r \leq p \leq 2$ , using again the triangle inequality in  $L^{\frac{2}{r}}(B_R(x))$  and Jensen’s inequality on  $t \mapsto t^{\frac{2}{r}}$ , we obtain

$$\int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_R(x)} |f|^2 \right)^{\frac{r}{2}} \right]^{\frac{p}{r}} dx \gtrsim \int_{\mathbb{R}^d} \left( \int_{B_R(x)} \mathbb{E} [ |f|_2^r ]^{\frac{2}{r}} \right)^{\frac{p}{2}} dx \gtrsim \int_{\mathbb{R}^d} \mathbb{E} [ |f|_2^r ]^{\frac{p}{r}}.$$

This proves the claim (6.27) in all cases.

We turn to the proof of (ii). Conditioning with respect to the value of  $r_*(x)$  on dyadic scale and using Jensen’s inequality on  $t \mapsto t^{\frac{p}{r}}$  with  $\sum_{n=0}^{\infty} 2^{-nd\delta} \simeq \delta^{-1}$ , we find for  $0 < \delta \leq 1$ ,

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [ |f|_2^r ]^{\frac{p}{r}} &\lesssim \int_{\mathbb{R}^d} \mathbb{E} \left[ \sum_{n=0}^{\infty} 2^{-nd\delta} r_*(x)^{d\delta} \mathbb{1}_{2^{n-1} < r_*(x) \leq 2^{n+1-1}} [f]_2^r \right]^{\frac{p}{r}} dx \\ &\lesssim \delta^{1-\frac{p}{r}} \sum_{n=0}^{\infty} 2^{nd\delta(\frac{p}{r}-1)} \int_{\mathbb{R}^d} \mathbb{E} [ \mathbb{1}_{2^{n-1} < r_*(x) \leq 2^{n+1-1}} [f]_2^r(x) ]^{\frac{p}{r}} dx. \end{aligned}$$

Applying (6.27) with  $R = (2^n - 2) \vee 1$  and taking advantage of the  $\frac{1}{8}$ -Lipschitz property of  $r_*$ , we deduce

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [ |f|_2^r ]^{\frac{p}{r}} &\lesssim \delta^{1-\frac{p}{r}} \sum_{n=0}^{\infty} (2^{ndp})^{(\frac{1}{2}-\frac{1}{p})_+ + \delta(\frac{1}{r}-\frac{1}{p})} \int_{\mathbb{R}^d} \mathbb{E} [ \mathbb{1}_{2^{n-1}-2 \leq r_*(x) \leq 2^{n+2}} \left( \int_{B_*(x)} |f|^2 \right)^{\frac{r}{2}} ]^{\frac{p}{r}} dx. \end{aligned}$$

Hence, for  $q > r$ , using Hölder’s inequality and the moment bounds on  $r_*$ , and proceeding as for (i),

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [ |f|_2^r ]^{\frac{p}{r}} &\lesssim \delta^{1-\frac{p}{r}} \sum_{n=0}^{\infty} (2^{ndp})^{(\frac{1}{2}-\frac{1}{p})_+ + \delta(\frac{1}{r}-\frac{1}{p})} \mathbb{P} [ r_* \geq 2^{n-1} - 2 ]^{p\frac{q-r}{qr}} \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx \\ &\lesssim_{q,p} \delta^{1-\frac{p}{r}} \left( \frac{qr}{q-r} \right)^p \left( \frac{1}{2} - \frac{1}{p} \right)_+ + \delta \left( \frac{1}{r} - \frac{1}{p} \right) \int_{\mathbb{R}^d} \mathbb{E} \left[ \left( \int_{B_*(x)} |f|^2 \right)^{\frac{q}{2}} \right]^{\frac{p}{q}} dx. \end{aligned}$$

Optimizing in  $0 < \delta \leq 1$  yields the desired lower bound. □

### 7 Proof of the higher-order pathwise result

In order to establish the accuracy of the two-scale expansion of higher-order commutators (hence, the higher-order pathwise structure), we appeal to the representation formula (5.3) established in Proposition 5.1 and the conclusion easily follows from the following estimates on the Malliavin derivative of higher-order correctors.

**Lemma 7.1** *Let  $T$ ,  $U_e$ , and  $U_{e,\circ}$  denote the operators defined in Theorem 6.1 and Corollary 6.2 with  $\mathbf{a}$  replaced by its adjoint  $\mathbf{a}^*$ . For all  $g \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$  and  $0 \leq s \leq \ell - 1$ ,*

$$\begin{aligned}
 J_s(g)^2 &:= \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} g \cdot \mathbf{a} \nabla D \varphi^{s+1} \right]_1^2 + \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} g \cdot \mathbf{a} D \varphi^s \right]_1^2 + \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} g \cdot D \sigma^s \right]_1^2 \\
 &\lesssim \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \| [U^k T^\beta g]_2 ([\nabla \varphi^{s+1-k}]_2 + [\varphi^{s-k}]_2) \|_{L^2(\mathbb{R}^d)}^2,
 \end{aligned}
 \tag{7.1}$$

where  $U^k$  denotes any product of the form  $U_1 \dots U_k$  with  $U_j \in \{U_{e_i}, U_{e_i,\circ}\}_{1 \leq i \leq d}$  for all  $j$ , and where we implicitly sum over all possible such products. Furthermore, for all  $\delta > 0$  and  $2 \leq p < \infty$ ,

$$\mathbb{E} [J_s(g)^p]^{1/p} \lesssim_{p,\delta} \| [g]_2 \|_{L^{\frac{2d}{d+2s}}(\mathbb{R}^d; L^{p+\delta}(\Omega))}. \tag{7.2} \diamond$$

**Proof** Denote by  $J_{s,1}(g)^2$ ,  $J_{s,2}(g)^2$ , and  $J_{s,3}(g)^2$  the three left-hand side terms in (7.1), hence  $J_s(g)^2 = J_{s,1}(g)^2 + J_{s,2}(g)^2 + J_{s,3}(g)^2$ . We split the proof into two steps.

*Step 1.* Proof that for all  $g \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$  and  $0 \leq s \leq \ell - 1$ ,

$$J_{s,1}(g) \lesssim J_{s,2}(Tg) + J_{s,3}(Tg) + \| [Tg]_2 ([\nabla \varphi^{s+1}]_2 + [\varphi^s]_2) \|_{L^2(\mathbb{R}^d)}, \tag{7.3}$$

$$J_{s,2}(g) = J_{s-1,1}(Ug), \tag{7.4}$$

$$J_{s,3}(g) \lesssim J_{s-1}(U_\circ g) + \| [U_\circ g]_2 ([\nabla \varphi^s]_2 + [\varphi^{s-1}]_2) \|_{L^2(\mathbb{R}^d)}. \tag{7.5}$$

We start with (7.3). Taking the Malliavin derivative of the equation for  $\varphi^{s+1}$  (cf. Definition 2.1) and testing it with  $(\nabla \cdot \mathbf{a}^* \nabla)^{-1} \nabla \cdot (\mathbf{a}^* g)$ , we obtain by definition of  $Tg = \nabla(\nabla \cdot \mathbf{a}^* \nabla)^{-1} \nabla \cdot (\mathbf{a}^* g)$ ,

$$\begin{aligned}
 \left| \int_{\mathbb{R}^d} g \cdot \mathbf{a} \nabla D_z \varphi_{i_1 \dots i_{s+1}}^{s+1} \right| &= \left| \int_{\mathbb{R}^d} Tg \cdot \mathbf{a} \nabla D_z \varphi_{i_1 \dots i_{s+1}}^{s+1} \right| \\
 &\lesssim |Tg(z)| (|\nabla \varphi_{i_1 \dots i_{s+1}}^{s+1}(z)| + |\varphi_{i_1 \dots i_s}^s(z)|) \\
 &\quad + \left| \int_{\mathbb{R}^d} Tg \cdot \mathbf{a} D_z \varphi_{i_1 \dots i_s}^s e_{i_{s+1}} \right| + \left| \int_{\mathbb{R}^d} Tg \cdot D_z \sigma_{i_1 \dots i_s}^s e_{i_{s+1}} \right|,
 \end{aligned}$$

where for the first right-hand side term we used that by definition (1.4) the Malliavin derivative of  $\mathbf{a}$  takes the form

$$D_z \mathbf{a} = a'_0(G(z)) \delta(\cdot - z). \tag{7.6}$$

The claim (7.3) follows by applying  $(\int_{\mathbb{R}^d} (\int_{B(x)} \cdot dz)^2 dx)^{1/2}$  to both sides of the estimate. Next, by definition of  $U_{e_{i_{s+1}}} g = \nabla(\nabla \cdot \mathbf{a}^* \nabla)^{-1} (e_{i_{s+1}} \cdot \mathbf{a}^* g)$ , we find

$$\left| \int_{\mathbb{R}^d} g \cdot \mathbf{a} D_z \varphi_{i_1 \dots i_s}^s e_{i_{s+1}} \right| = \left| \int_{\mathbb{R}^d} U_{e_{i_{s+1}}} g \cdot \mathbf{a} \nabla D_z \varphi_{i_1 \dots i_s}^s \right|,$$

so that (7.4) follows directly. We turn to (7.5). Writing  $g = \Delta \Delta^{-1}g$ , integrating by parts, and using the Malliavin derivative of the equation for  $\Delta \sigma^s$  (cf. Definition 2.1), we find

$$\begin{aligned} \left| \int_{\mathbb{R}^d} g \cdot D_z \sigma_{i_1 \dots i_s}^s e_{i_{s+1}} \right| &\leq \left| \int_{\mathbb{R}^d} (\nabla_i \Delta^{-1} g_i) e_{i_{s+1}} \cdot D_z q_{i_1 \dots i_s}^s \right| \\ &\quad + \left| \int_{\mathbb{R}^d} (\nabla_{i_{s+1}} \Delta^{-1} g_i) e_i \cdot D_z q_{i_1 \dots i_s}^s \right|, \end{aligned}$$

and the claim (7.5) follows from taking the Malliavin derivative of the definition of  $q^s$  and writing  $\nabla \Delta^{-1} g_i = U_{e_i, \circ} g$ .

**Step 2. Conclusion.**

Using the notation introduced in the statement, Step 1 yields for all  $g \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$ ,

$$\begin{aligned} J_s(g) &\lesssim_p J_{s-1}(U \cdot g) + J_{s-1}(U \cdot Tg) + \left\| [Tg]_2 ([\nabla \varphi^{s+1}]_2 + [\varphi^s]_2) \right\|_{L^2(\mathbb{R}^d)} \\ &\quad + \left\| ([U \circ g]_2 + [U \circ Tg]_2) ([\nabla \varphi^s]_2 + [\varphi^{s-1}]_2) \right\|_{L^2(\mathbb{R}^d)}. \end{aligned}$$

Since by definition  $TU = U$  and  $TU_\circ = U_\circ$ , the decomposition (7.1) follows by induction. Next, combining (7.1) with the corrector estimates of Proposition 2.2, we find for all  $g \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))^d$ ,  $0 \leq s \leq \ell - 1$ ,  $2 \leq p < \infty$ , and  $\delta > 0$ ,

$$\mathbb{E} \left[ J_s(g)^p \right]^{\frac{1}{p}} \lesssim_{p,\delta} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \| [U^k T^\beta g]_2 \|_{L^2(\mathbb{R}^d; L^{p+\delta}(\Omega))},$$

where the anchoring in  $s = 0$  follows from (7.3) because of  $J_{0,2} = J_{0,3} = 0$  in view of  $\varphi^0 = 1$  and  $\sigma^0 = 0$ . Here we have used that for a stationary  $\psi$  we have by the triangle inequality in  $L^{\frac{p}{2}}(\Omega)$  and Hölder’s inequality,

$$\mathbb{E} \left[ \|\psi g\|_{L^2(\mathbb{R}^d)}^p \right]^{\frac{1}{p}} \leq \|\psi g\|_{L^2(\mathbb{R}^d; L^p(\Omega))} \leq \mathbb{E} \left[ |\psi|^{\frac{p(p+\delta)}{\delta}} \right]^{\frac{\delta}{p(p+\delta)}} \|g\|_{L^2(\mathbb{R}^d; L^{p+\delta}(\Omega))}.$$

Now, for  $0 \leq s \leq \ell - 1$ ,  $2 \leq p < \infty$ ,  $0 \leq k \leq s$ ,  $\beta \in \{0, 1\}$ , and  $\delta > 0$ , we apply the bound of Corollary 6.2  $k$  times and the one of Theorem 6.1 once, increasing at every application the stochastic integrability by  $\frac{\delta}{k+1}$ , to the effect of

$$\| [U^k T^\beta g]_2 \|_{L^2(\mathbb{R}^d; L^{p+\delta}(\Omega))} \lesssim_{p,\delta} \| [g]_2 \|_{L^{\frac{2d}{d+2k}}(\mathbb{R}^d; L^{p+2\delta}(\Omega))},$$

given that the restriction  $k \leq s \leq \ell - 1$  indeed ensures  $\frac{2d}{d+2k} > 1$ , and the conclusion follows. □

With these estimates at hand, we may now turn to the proof of Theorem 1(ii), that is, of the higher-order pathwise result.

**Proof of Theorem 1(ii)** Let  $1 \leq p < \infty$ . The starting point is the  $L^p$  inequality of Proposition 4.1(iii) together with (4.2), in the form

$$Q := \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} g \cdot (\Xi^n[\nabla u] - \Xi^{\circ,n}[\nabla \bar{u}^n]) - \int_{\mathbb{R}^d} g \cdot \mathbb{E} [\Xi^n[\nabla u] - \Xi^{\circ,n}[\nabla \bar{u}^n]] \right)^{2p} \right]^{\frac{1}{2p}}$$

$$\lesssim_p \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} g \cdot D(\Xi^n[\nabla u] - \Xi^{\circ,n}[\nabla \bar{u}^n]) \right]_1^2 \right)^p \right]^{\frac{1}{2p}}.$$

Inserting identity (5.3) of Proposition 5.1, integrating by parts in the  $z$ -integral to put derivatives on  $g$  and  $\bar{u}^n$ , using (7.6), and using the corrector estimates of Proposition 2.2, we deduce for all  $1 \leq n \leq \ell$ ,

$$Q \lesssim_p \sum_{k=0}^n \|\mu_{d,k}[\nabla^k g]_\infty\|_{L^4(\mathbb{R}^d)} \|\nabla u - E^n[\nabla \bar{u}^n]\|_{L^4(\mathbb{R}^d; L^{4p}(\Omega))}$$

$$+ \sum_{k=0}^n \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k+s+1-l} \bar{u}^n)]_\infty\|_{L^2(\mathbb{R}^d)}$$

$$+ \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S^1 \cdot \nabla D u \right]_1^2 \right)^p \right]^{\frac{1}{2p}} + \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S^1 \cdot D E^n[\nabla \bar{u}^n] \right]_1^2 \right)^p \right]^{\frac{1}{2p}}$$

$$+ \sum_{s=0}^{n-1} \sum_{l=0}^s \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S_{j_1 \dots j_{s+1}}^{2;l,s} \cdot D(\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right]_1^2 \right)^p \right]^{\frac{1}{2p}}, \tag{7.7}$$

where we have set

$$S^1 := \nabla_{i_1 \dots i_n}^n g_j (\mathbf{a}^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n}) e_{i_n},$$

$$S_{j_1 \dots j_{s+1}}^{2;l,s} := \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{u}^n) \text{Sym}_{i_1 \dots i_n} ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n}).$$

We separately treat the different right-hand side terms in (7.7). First, from equation (2.12), formula (2.5), the annealed Calderón–Zygmund estimate of Theorem 6.1, and the corrector estimates in Proposition 2.2, we obtain

$$\|\nabla u - E^n[\nabla \bar{u}^n]\|_{L^4(\mathbb{R}^d; L^{4p}(\Omega))}$$

$$\lesssim_p \|\mu_{d,n}[\nabla^{n+1} \bar{u}^n]_\infty\|_{L^4(\mathbb{R}^d)} + \sum_{k=2}^n \sum_{l=n+2-k}^n \|\nabla^k \bar{u}^l\|_{L^4(\mathbb{R}^d)}. \tag{7.8}$$

Second, recalling the notation  $T = \nabla(\nabla \cdot \mathbf{a}^* \nabla)^{-1} \nabla \cdot \mathbf{a}^*$  from Theorem 6.1 and using the equation for  $u$  in the form  $-\nabla \cdot \mathbf{a} \nabla D_z u = \nabla \cdot D_z \mathbf{a} \nabla u$ , we find

$$\begin{aligned} \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S^1 \cdot \nabla Du \right]_1^2 \right)^p \right]^{\frac{1}{2p}} &= \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} T((\mathbf{a}^*)^{-1} S^1) \cdot \mathbf{a} \nabla Du \right]_1^2 \right)^p \right]^{\frac{1}{2p}} \\ &\lesssim \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} [T((\mathbf{a}^*)^{-1} S^1)]_2^2 [\nabla u]_2^2 \right)^p \right]^{\frac{1}{2p}} \\ &\leq \|T((\mathbf{a}^*)^{-1} S^1)\|_2 \|L^4(\mathbb{R}^d; L^{4p}(\Omega))\| \|\nabla u\|_2 \|L^4(\mathbb{R}^d; L^{4p}(\Omega))\|, \end{aligned}$$

hence, using Theorem 6.1 to bound the operator  $T$  and to estimate  $\nabla u$ , and applying the corrector estimates of Proposition 2.2, we deduce

$$\begin{aligned} \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S^1 \cdot \nabla Du \right]_1^2 \right)^p \right]^{\frac{1}{2p}} &\lesssim_{p,\delta} \|S^1\|_2 \|L^4(\mathbb{R}^d; L^{4p+\delta}(\Omega))\| \|f\|_2 \|L^4(\mathbb{R}^d) \\ &\lesssim \|\mu_{d,n}[\nabla^n g]\|_\infty \|L^4(\mathbb{R}^d)\| \|f\|_2 \|L^4(\mathbb{R}^d)\|. \end{aligned} \tag{7.9}$$

Third, decomposing

$$D_z E^n [\nabla \bar{u}^n] = \sum_{k=0}^{n-1} (D_z \nabla \varphi_{i_1 \dots i_{k+1}}^{k+1} + D_z \varphi_{i_1 \dots i_k}^k e_{i_{k+1}}) \nabla_{i_1 \dots i_{k+1}}^{k+1} \bar{u}^n,$$

applying (7.2), and using the corrector estimates of Proposition 2.2,

$$\begin{aligned} \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S^1 \cdot D E^n [\nabla \bar{u}^n] \right]_1^2 \right)^p \right]^{\frac{1}{2p}} &\lesssim_{p,\delta} \sum_{k=0}^{n-1} \| [S^1 \nabla^{k+1} \bar{u}^n]_2 \|_{L^{\frac{2d}{d+2k}}(\mathbb{R}^d; L^{2p+\delta}(\Omega))} \\ &\lesssim_{p,\delta} \sum_{k=0}^{n-1} \| \mu_{d,n}[\nabla^n g \nabla^{k+1} \bar{u}^n]_\infty \|_{L^{\frac{2d}{d+2k}}(\mathbb{R}^d)}. \end{aligned}$$

Fourth, applying (7.2) and using the corrector estimates of Proposition 2.2,

$$\begin{aligned} &\sum_{s=0}^{n-1} \sum_{l=0}^s \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S_{j_1 \dots j_{s+1}}^{2;l,s} \cdot D(\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right]_1^2 \right)^p \right]^{\frac{1}{2p}} \\ &\lesssim_{p,\delta} \sum_{s=0}^{n-1} \sum_{l=0}^s \| [S^{2;l,s}]_2 \|_{L^{\frac{2d}{d+2s}}(\mathbb{R}^d; L^{2p+\delta}(\Omega))} \\ &\lesssim_{p,\delta} \sum_{s=0}^{n-1} \sum_{l=0}^s \| \mu_{d,n}[\nabla^l (g \nabla^{n+s+1-l} \bar{u}^n)]_\infty \|_{L^{\frac{2d}{d+2s}}(\mathbb{R}^d)}. \end{aligned}$$

Collecting the above estimates on the right-hand side terms of (7.7), we see that all of them are bilinear in  $g$  and  $f$ . Crucial for us is their scaling in the underlying lengthscale, which has three origins: the (total) number of derivatives (to be counted

negatively), the spatial integrability exponent  $p$  (to be counted in form of  $\frac{d}{p}$ ), and the weights  $\mu_{d,n}$  (cf. Proposition 2.2). For instance, in view of (7.8), the first right-hand side term in (7.7) has scaling  $\leq L^{-n+\frac{d}{4}+\frac{d}{4}}\mu_{d,n}(L) = L^{-n+\frac{d}{2}}\mu_{d,n}(L)$ , where  $L$  stands for the lengthscale, because  $\nabla\tilde{u}^l$  has the same scaling as  $\nabla^{l-1}f$  (cf. Definition 2.7). Let us take the last term as a second example: the scaling of its  $(s, l)$ -contribution is  $L^{-(n+s)+\frac{d+2s}{2}}\mu_{d,n}(L) = L^{-n+\frac{d}{2}}\mu_{d,n}(L)$ . In fact, all terms have a scaling at most  $L^{-n+\frac{d}{2}}\mu_{d,n}(L)$ , which after  $\varepsilon$ -rescaling ( $L = \varepsilon^{-1}$ ) turns into the desired estimate.  $\square$

### 8 Proof of the convergence of the covariance structure

Combining the representation formula (5.2) of Proposition 5.1 with the Helffer–Sjöstrand identity of Proposition 4.1(ii), we establish Theorem 1(iii), that is, the convergence of the covariance structure of the higher-order standard homogenization commutator.

**Proof of Theorem 1(iii)** Combining the Helffer–Sjöstrand identity of Proposition 4.1(ii) with the representation formula (5.2) tested with  $g \in C_c^\infty(\mathbb{R}^d)^d$ , we obtain, using (7.6) and recalling that  $(1 + \mathcal{L})^{-1}$  has operator norm  $\leq 1$ ,

$$\begin{aligned} & \left| \text{Var} \left[ \int_{\mathbb{R}^d} g \cdot \Xi^{\circ,n}[\nabla\bar{w}] \right] - \mathbb{E} \left[ \langle S, (1 + \mathcal{L})^{-1} S \rangle_{\mathfrak{H}} \right] \right| \\ & \leq 2 \mathbb{E} \left[ \|S\|_{\mathfrak{H}}^2 \right]^{\frac{1}{2}} \mathbb{E} \left[ \|R\|_{\mathfrak{H}}^2 \right]^{\frac{1}{2}} + \mathbb{E} \left[ \|R\|_{\mathfrak{H}}^2 \right], \end{aligned} \tag{8.1}$$

where we have set

$$S(z) := \sum_{k=0}^n \nabla_{i_1 \dots i_k}^k g_j(z) \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1}}^{*,k} e_{i_k} \right)(z) \cdot a'_0(G(z)) E^n[\nabla\bar{w}](z),$$

and

$$\begin{aligned} R(z) := & - \sum_{k=0}^n (-1)^k \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} (-1)^l \binom{k}{l} \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w})(z) \\ & \times \text{Sym}_{i_1 \dots i_k} \left( \mathbb{1}_{k \neq n} \nabla \varphi_{j i_1 \dots i_k}^{*,k+1} + \varphi_{j i_1 \dots i_{k-1}}^{*,k} e_{i_k} \right)(z) \cdot a'_0(G(z)) (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}})(z) \\ & + (-1)^n \sum_{s=0}^{n-1} \sum_{l=s+1}^n (-1)^l \binom{n}{l} \int_{\mathbb{R}^d} \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \\ & \times \text{Sym}_{i_1 \dots i_n} \left( (a^* \varphi_{j i_1 \dots i_{n-1}}^{*,n} - \sigma_{j i_1 \dots i_{n-1}}^{*,n}) e_{i_n} \right) \cdot D_z (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}). \end{aligned}$$

Appealing to the definition of  $E^n[\nabla\bar{w}]$  (cf. Definition 2.6), to (4.2), and to the corrector estimates of Proposition 2.2, we find

$$\mathbb{E} \left[ \|S\|_{\mathfrak{H}}^2 \right]^{\frac{1}{2}} \lesssim \sum_{k=0}^n \sum_{l=0}^{n-1} \|\mu_{d,n}[\nabla^k g]_{\infty}[\nabla^{l+1} \bar{w}]_{\infty}\|_{L^2(\mathbb{R}^d)},$$

and similarly,

$$\begin{aligned} \mathbb{E} \left[ \|R\|_{\mathfrak{H}}^2 \right]^{\frac{1}{2}} &\lesssim \sum_{k=0}^n \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k+s+1-l} \bar{w})]_{\infty}\|_{L^2(\mathbb{R}^d)} \\ &+ \sum_{s=0}^{n-1} \sum_{l=s+1}^n \mathbb{E} \left[ \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} S_{j_1 \dots j_{s+1}}^{2;l,s} \cdot D_z(\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}) \right]_1^2 \right]^{\frac{1}{2}}, \end{aligned}$$

in terms of

$$S_{j_1 \dots j_{s+1}}^{2;l,s} := \nabla_{i_1 \dots i_l}^l (g_{j_{l+1} \dots j_n} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \text{Sym}_{i_1 \dots i_n} ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n}).$$

Then applying (7.2) and again using the corrector estimates of Proposition 2.2, we deduce for all  $\delta > 0$ ,

$$\begin{aligned} \mathbb{E} \left[ \|R\|_{\mathfrak{H}}^2 \right]^{\frac{1}{2}} &\lesssim_{\delta} \sum_{k=0}^n \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k+s+1-l} \bar{w})]_{\infty}\|_{L^2(\mathbb{R}^d)} \\ &+ \sum_{s=0}^{n-1} \sum_{l=s+1}^n \| [S^{2;l,s}]_2 \|_{L^{\frac{2d}{d+2s}}(\mathbb{R}^d; L^{2+\delta}(\Omega))} \\ &\lesssim_{\delta} \sum_{k=0}^n \sum_{s=n-k}^{n-1} \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k+s+1-l} \bar{w})]_{\infty}\|_{L^2(\mathbb{R}^d)} \\ &+ \sum_{s=0}^{n-1} \sum_{l=s+1}^n \|\mu_{d,n}[\nabla^l (g \nabla^{n+s+1-l} \bar{w})]_{\infty}\|_{L^{\frac{2d}{d+2s}}(\mathbb{R}^d)}. \end{aligned}$$

As in the proof of Theorem 1(ii) (cf. Sect. 7), it is easily checked that these expressions in  $(g, \nabla \bar{w})$  have the desired length scaling, hence lead to the claimed error estimate after  $\varepsilon$ -rescaling.

Now that the right-hand side of (8.1) is properly estimated, we turn to a suitable decomposition of  $\mathbb{E} [\langle S, (1 + \mathcal{L})^{-1} S \rangle_{\mathfrak{H}}]$ . Inserting the definition of  $E^n[\nabla \bar{w}]$  (cf. Definition 2.6) into  $S$ , using the definition (4.1) of the scalar product in  $\mathfrak{H}$ , expanding the product, and focusing on terms of order  $\varepsilon^k$  with  $k \leq n - 1$ , we write

$$\begin{aligned} &\left| \mathbb{E} [\langle S, (1 + \mathcal{L})^{-1} S \rangle_{\mathfrak{H}}] - \sum_{l,l',k,k'=0}^{n-1} \mathbb{1}_{l+l'+k+k' \leq n-1} \right. \\ &\times \left. \iint_{\mathbb{R}^d \times \mathbb{R}^d} c(z' - z) \nabla_{i_1 \dots i_{l'}}^{l'} g_{j'}(z') \nabla_{j_1 \dots j_{k'+1}}^{k'+1} \bar{w}(z') \nabla_{i_1 \dots i_l}^l g_j(z) \nabla_{j_1 \dots j_{k+1}}^{k+1} \bar{w}(z) \right. \end{aligned}$$

$$\begin{aligned} & \times \mathbb{E} \left[ T_{j'_1 \dots i'_l j_1 \dots j_{k+1}}^{l',k'}(z') (1 + \mathcal{L})^{-1} T_{j_1 \dots i_l j_1 \dots j_{k+1}}^{l,k}(z) \right] dz dz' \Big| \\ & \lesssim \sum_{l,l'=0}^n \sum_{k,k'=0}^{n-1} \mathbb{1}_{l+l'+k+k' \geq n} \|\mu_{d,l'}[\nabla^{l'} g]_{\infty} [\nabla^{k'+1} \bar{w}]_{\infty}\|_{L^2(\mathbb{R}^d)} \|\mu_{d,l}[\nabla^l g]_{\infty} [\nabla^{k+1} \bar{w}]_{\infty}\|_{L^2(\mathbb{R}^d)}, \end{aligned}$$

where we have set

$$T_{j_1 \dots i_l j_1 \dots j_{k+1}}^{l,k} := (\mathbb{1}_{l \neq n} \nabla \varphi_{j_1 \dots i_l}^{*,l+1} + \varphi_{j_1 \dots i_{l-1}}^{*,l} e_{i_l}) \cdot a'_0(G) (\nabla \varphi_{j_1 \dots j_{k+1}}^{k+1} + \varphi_{j_1 \dots j_k}^k e_{j_{k+1}}).$$

Combining this with (8.1) and with the above estimates on  $\mathbb{E} [\|S\|_{\mathcal{S}}^2]^{\frac{1}{2}}$  and  $\mathbb{E} [\|R\|_{\mathcal{S}}^2]^{\frac{1}{2}}$  and appealing to the stationarity of  $T_{j_1 \dots i_l j_1 \dots j_{k+1}}^{l,k}$  (recalling that  $\mathcal{L}$  commutes with shifts), we deduce after  $\varepsilon$ -rescaling,

$$\begin{aligned} & \left| \text{Var} \left[ \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \Xi_{\varepsilon}^{\circ,n} [\nabla \bar{w}] \right] - \sum_{m=0}^{n-1} \varepsilon^m \sum_{\substack{l,l',k,k' \geq 0 \\ l+l'+k+k'=m}} \right. \\ & \times \iint_{\mathbb{R}^d \times \mathbb{R}^d} c(y) \nabla_{i'_1 \dots i'_l}^{l'} g_{j'}(z + \varepsilon y) \nabla_{j'_1 \dots j'_{k'+1}}^{k'+1} \bar{w}(z + \varepsilon y) \nabla_{i_1 \dots i_l}^l g_j(z) \nabla_{j_1 \dots j_{k+1}}^{k+1} \bar{w}(z) \\ & \left. \times \mathbb{E} \left[ T_{j'_1 \dots i'_l j'_1 \dots j'_{k'+1}}^{l',k'}(y) (1 + \mathcal{L})^{-1} T_{j_1 \dots i_l j_1 \dots j_{k+1}}^{l,k}(0) \right] dy dz \right| \lesssim_{n,f,g} \varepsilon^n \mu_{d,n} \left(\frac{1}{\varepsilon}\right). \end{aligned}$$

Appealing to the boundedness of  $(1 + \mathcal{L})^{-1}$  and to the corrector estimates of Proposition 2.2 in the form

$$\left| \mathbb{E} \left[ T_{j'_1 \dots i'_l j'_1 \dots j'_{k'+1}}^{l',k'}(y) (1 + \mathcal{L})^{-1} T_{j_1 \dots i_l j_1 \dots j_{k+1}}^{l,k}(0) \right] \right| \lesssim 1,$$

the conclusion follows after Taylor-expanding  $\nabla^{l'} g_{j'}(z + \varepsilon y) \nabla^{k'+1} \bar{w}(z + \varepsilon y)$  to order  $\varepsilon^{n-m}$  with the integrability assumption  $\int_{\mathbb{R}^d} |y|^n |c(y)| dy \leq 1$ .  $\square$

### 9 Proof of the normal approximation result

In order to establish the asymptotic normality of the higher-order standard homogenization commutators, we combine the representation formula of Proposition 5.2 with the second-order Poincaré inequality of Proposition 4.1(iv). We then need the following version of Lemma 7.1 for the estimation of second Malliavin derivatives of higher-order correctors.

**Lemma 9.1** *For all  $\zeta \in C_c^\infty(\mathbb{R}^d)$ ,  $g \in C_c^\infty(\mathbb{R}^d)^d$ ,  $0 \leq s \leq \ell - 1$ , and  $p \geq 1$ ,*

$$\begin{aligned} K_s(g, \zeta) & := \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x) \zeta(y) \left( \int_{\mathbb{R}^d} g \cdot \mathbf{a} \nabla D_{xy}^2 \varphi^{s+1} \right) dx dy \right| \\ & + \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x) \zeta(y) \left( \int_{\mathbb{R}^d} g \cdot \mathbf{a} D_{xy}^2 \varphi^s \right) dx dy \right| \end{aligned}$$



$$\begin{aligned}
 & + \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} g \cdot D_{xy}^2 \sigma^s \right) dx dy \right| \\
 \lesssim & \|[\zeta]_\infty\|_{L^{\frac{2p}{p-1}}(\mathbb{R}^d)}^2 \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \| [U^k T^\beta g]_2 ([\nabla \varphi^{s+1-k}]_2 + [\varphi^{s-k}]_2) \|_{L^p(\mathbb{R}^d)} \\
 & + \|[\zeta]_\infty\|_{L^2(\mathbb{R}^d)} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} J_{s-k}(\zeta(\mathbf{a}^*)^{-1} a'_0(G)^* U^k T^\beta g),
 \end{aligned}$$

where the notation is taken from Lemma 7.1. ◇

**Proof** Denote by  $K_{s,1}(g, \zeta)$ ,  $K_{s,2}(g, \zeta)$ , and  $K_{s,3}(g, \zeta)$  the three left-hand side terms in the claimed estimate, hence  $K_s(g, \zeta) = K_{s,1}(g, \zeta) + K_{s,2}(g, \zeta) + K_{s,3}(g, \zeta)$ . We prove that for all  $g \in C_c^\infty(\mathbb{R}^d; L^\infty(\mathbb{R}^d))^d$ ,  $\zeta \in C_c^\infty(\mathbb{R}^d)$ ,  $0 \leq s \leq \ell - 1$ , and  $p \geq 1$ ,

$$\begin{aligned}
 K_{s,1}(g, \zeta) \lesssim & K_{s,2}(Tg, \zeta) + K_{s,3}(Tg, \zeta) + \|[\zeta]_\infty\|_{L^2(\mathbb{R}^d)} J_s(\zeta(\mathbf{a}^*)^{-1} a'_0(G)^* Tg) \\
 & + \|[\zeta]_\infty\|_{L^{\frac{2p}{p-1}}(\mathbb{R}^d)}^2 \|([\nabla \varphi^{s+1}]_2 + [\varphi^s]_2)[Tg]_2\|_{L^p(\mathbb{R}^d)}, \tag{9.1}
 \end{aligned}$$

$$K_{s,2}(g, \zeta) = K_{s-1,1}(Ug, \zeta), \tag{9.2}$$

$$\begin{aligned}
 K_{s,3}(g, \zeta) \lesssim & K_{s-1}(U_0g, \zeta) + \|[\zeta]_\infty\|_{L^2(\mathbb{R}^d)} J_{s-1}(\zeta(\mathbf{a}^*)^{-1} a'_0(G)^* U_0g) \\
 & + \|[\zeta]_\infty\|_{L^{\frac{2p}{p-1}}(\mathbb{R}^d)}^2 \|([\nabla \varphi^s]_2 + [\varphi^{s-1}]_2)[U_0g]_2\|_{L^p(\mathbb{R}^d)}. \tag{9.3}
 \end{aligned}$$

These estimates combine to

$$\begin{aligned}
 K_s(g, \zeta) \lesssim & \sum_{\beta \in \{0,1\}} \left( K_{s-1}(U.T^\beta g, \zeta) \right. \\
 & + \|[\zeta]_\infty\|_{L^{\frac{2p}{p-1}}(\mathbb{R}^d)} \|([\nabla \varphi^{s+1}]_2 + [\varphi^s]_2)[Tg]_2 + ([\nabla \varphi^s]_2 + [\varphi^{s-1}]_2)[U_0.T^\beta g]_2\|_{L^p(\mathbb{R}^d)} \\
 & \left. + \|[\zeta]_\infty\|_{L^2(\mathbb{R}^d)} \left( J_s(\zeta(\mathbf{a}^*)^{-1} a'_0(G)^* Tg) + J_{s-1}(\zeta(\mathbf{a}^*)^{-1} a'_0(G)^* U_0.T^\beta g) \right) \right).
 \end{aligned}$$

Noting as in the proof of Lemma 7.1 that  $TU = U$  and  $TU_0 = U_0$ , the conclusion follows by induction.

Let  $0 \leq s \leq \ell - 1$ . We start with the proof of (9.1). Taking the second Malliavin derivative of the equation for  $\varphi^{s+1}$  (cf. Definition 2.1), we find

$$\begin{aligned}
 -\nabla \cdot \mathbf{a} \nabla D_{xy}^2 \varphi_{i_1 \dots i_{s+1}}^{s+1} & = \nabla \cdot ((\mathbf{a} D_{xy}^2 \varphi_{i_1 \dots i_s}^s - D_{xy}^2 \sigma_{i_1 \dots i_s}^s) e_{i_{s+1}}) \\
 & + \nabla \cdot (D_x \mathbf{a} (\nabla D_y \varphi_{i_1 \dots i_{s+1}}^{s+1} + D_y \varphi_{i_1 \dots i_s}^s e_{i_{s+1}})) + \nabla \cdot (D_y \mathbf{a} (\nabla D_x \varphi_{i_1 \dots i_{s+1}}^{s+1} + D_x \varphi_{i_1 \dots i_s}^s e_{i_{s+1}})) \\
 & + \nabla \cdot (D_{xy}^2 \mathbf{a} (\nabla \varphi_{i_1 \dots i_{s+1}}^{s+1} + \varphi_{i_1 \dots i_s}^s e_{i_{s+1}})).
 \end{aligned}$$

Testing with  $(\nabla \cdot \mathbf{a}^* \nabla)^{-1} \nabla \cdot (\mathbf{a}^* g)$ , recalling the definition of  $T$ , and using (7.6) and

$$D_{xy}^2 \mathbf{a} = a''_0(G(x)) \delta(\cdot - x) \delta(x - y), \tag{9.4}$$

we find after integration against  $\zeta \otimes \zeta$ ,

$$\begin{aligned} & \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} g \cdot \mathbf{a} \nabla D_{xy}^2 \varphi_{i_1 \dots i_{s+1}}^{s+1} \right) dx dy \right| \\ & \lesssim \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} Tg \cdot \left( \mathbf{a} D_{xy}^2 \varphi_{i_1 \dots i_s}^s - D_{xy}^2 \sigma_{i_1 \dots i_s}^s \right) e_{i_{s+1}} \right) dx dy \right| \\ & \quad + \int_{\mathbb{R}^d} |\zeta(y)| \left| \int_{\mathbb{R}^d} \zeta Tg \cdot a'_0(G) (\nabla D_y \varphi_{i_1 \dots i_{s+1}}^{s+1} + D_y \varphi_{i_1 \dots i_s}^s e_{i_{s+1}}) dy \right| \\ & \quad + \int_{\mathbb{R}^d} [\zeta]_{\infty}^2 [Tg]_2 ([\nabla \varphi^{s+1}]_2 + [\varphi^s]_2), \end{aligned} \tag{9.5}$$

and the claim (9.1) follows from Hölder’s inequality. Next, (9.2) follows from the definition of  $U$ . We turn to (9.3). Writing  $g = \Delta \Delta^{-1} g$ , integrating by parts, and using the equation for  $\Delta \sigma^s$  (cf. Definition 2.1), we find

$$\begin{aligned} & \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} g \cdot D_{xy}^2 \sigma_{i_1 \dots i_s}^s e_{i_{s+1}} \right) dx dy \right| \\ & \leq \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} (\nabla_i \Delta^{-1} g_i) e_{i_{s+1}} \cdot D_{xy}^2 q_{i_1 \dots i_s}^s \right) dx dy \right| \\ & \quad + \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x)\zeta(y) \left( \int_{\mathbb{R}^d} (\nabla_{i_{s+1}} \Delta^{-1} g_i) e_i \cdot D_{xy}^2 q_{i_1 \dots i_s}^s \right) dx dy \right|. \end{aligned}$$

Hence, by the definition of  $q^s$  to which we apply  $D_{xy}^2$ , we obtain an analogous structure as in (9.5), and the claim (9.3) follows. □

With these estimates at hand, we now turn to the proof of Theorem 1(iv), that is, of the normal approximation result for the higher-order standard homogenization commutators.

**Proof of Theorem 1(iv)** We split the proof into two steps.

*Step 1.* Proof that for all  $8 \leq p, q < \infty$  and  $1 \leq n \leq \ell$ ,

$$\begin{aligned} & \mathbb{E} \left[ \left\| D^2 \int_{\mathbb{R}^d} g \cdot \mathbb{E}^{\circ, n} [\nabla \bar{w}] \right\|_{\text{op}}^4 \right]^{\frac{1}{4}} \\ & \lesssim \sum_{k=0}^n \sum_{s=0}^{n-1} p^{c_k+c_s} \left( \|\mu_{d,k}[\nabla^k g \nabla^{s+1} \bar{w}]\|_{L^p(\mathbb{R}^d)} \right. \\ & \quad \left. + \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k-l+s+1} \bar{w})]\|_{L^p(\mathbb{R}^d)} \right) \\ & \quad + \sum_{k=0}^n \sum_{s=0}^{n-1} \sum_{m=0}^s p^{c_k+c_s-m} (p \log^2 p)^{\frac{m(m+1)}{2d}} \left( \|\mu_{d,k}[\nabla^k g \nabla^{s+1} \bar{w}]\|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d)} \right) \end{aligned}$$

$$\begin{aligned}
 & + \sum_{l=0}^{k+s-n} \left\| \mu_{d,k} [\nabla^l (g \nabla^{k-l+s+1} \bar{w})]_\infty \right\|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d)} \\
 & + C_q \sum_{s=0}^{n-1} \sum_{l=s+1}^n \left\| \mu_{d,n} [\nabla^l (g \nabla^{n-l+s+1} \bar{w})]_\infty \right\|_{L^{\frac{dq}{2d+sq}}(\mathbb{R}^d)}.
 \end{aligned}$$

In order to not overestimate the logarithmic correction in the final result, we have to be specific here on the  $p$ -dependence; however, this is only necessary in the leading-order terms (that is, in the first two terms). Let  $1 \leq n \leq \ell$ . Applying the representation formula of Proposition 5.2, using (7.6) and (9.4), inserting the definition of  $E_\varepsilon^n[\nabla \bar{w}]$  (cf. Definition 2.6), and ordering terms by the number of Malliavin derivatives on  $\mathbf{a}$ , we find

$$D^2 \int_{\mathbb{R}^d} g \cdot \Xi^{\circ,n}[\nabla \bar{w}] = U_1 + U_2 + U_3, \tag{9.6}$$

in terms of

$$\begin{aligned}
 U_1(x, y) & := \delta(x - y) S^1(x), \\
 U_2(x, y) & := \tilde{U}_2(x, y) + \tilde{U}_2(y, x), \\
 \tilde{U}_2(x, y) & := \sum_{s=0}^{n-1} S_{j_1 \dots j_{s+1}}^{2;s}(x) \cdot D_y(\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}})(x), \\
 U_3(x, y) & := \sum_{s=0}^{n-1} \int_{\mathbb{R}^d} S_{j_1 \dots j_{s+1}}^{3;s} \cdot D_{xy}^2(\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}),
 \end{aligned}$$

where we identify the operators  $U_i$  with their kernels and we have set

$$\begin{aligned}
 S^1 & := \sum_{k=0}^n \sum_{s=0}^{n-1} \left( (\nabla_{i_1 \dots i_k}^k g_j \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \right. \\
 & \quad - \sum_{l=0}^{k+s-n} \binom{k}{l} (-1)^{k-l} \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \\
 & \quad \times \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) \cdot a_0''(G) (\nabla \varphi_{j_1 \dots j_{s+1}}^{s+1} + \varphi_{j_1 \dots j_s}^s e_{j_{s+1}}), \\
 S_{j_1 \dots j_{s+1}}^{2;s} & := \sum_{k=0}^n \left( (\nabla_{i_1 \dots i_k}^k g_j \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \right. \\
 & \quad - \sum_{l=0}^{k+s-n} \binom{k}{l} (-1)^{k+l} \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_k}^{k-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \\
 & \quad \times \text{Sym}_{i_1 \dots i_k} (\mathbb{1}_{k \neq n} \nabla \varphi_{j_1 \dots i_k}^{*,k+1} + \varphi_{j_1 \dots i_{k-1}}^{*,k} e_{i_k}) a_0'(G), \\
 S_{j_1 \dots j_{s+1}}^{3;s} & := \sum_{l=s+1}^n \binom{n}{l} (-1)^{n+l} \nabla_{i_1 \dots i_l}^l (g_j \nabla_{i_{l+1} \dots i_n}^{n-l} \nabla_{j_1 \dots j_{s+1}}^{s+1} \bar{w}) \\
 & \quad \times \text{Sym}_{i_1 \dots i_n} ((\mathbf{a}^* \varphi_{j_1 \dots i_{n-1}}^{*,n} - \sigma_{j_1 \dots i_{n-1}}^{*,n}) e_{i_n}),
 \end{aligned}$$

with the implicit understanding that the contribution of  $\sum_{l=0}^{k+s-n}$  vanishes if  $k < n - s$ . We analyze the three right-hand side terms in (9.6) separately and we start with  $U_1$ . Decomposing the covariance function as  $c = c_0 * c_0$  so that  $\|\zeta\|_{\mathfrak{S}} = \|c_0 * \zeta\|_{L^2(\mathbb{R}^d)}$ , the definition (4.7) of the norm  $\|\cdot\|_{\text{op}}$  is equivalent to

$$\|U_1\|_{\text{op}} = \sup_{\|\zeta\|_{L^2(\mathbb{R}^d)} = \|\zeta'\|_{L^2(\mathbb{R}^d)} = 1} \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} (c_0 * \zeta)(x) U_1(x, y) (c_0 * \zeta')(y) dx dy \right|. \tag{9.7}$$

For  $p \geq 4$ , inserting the special form of  $U_1(x, y) = \delta(x - y) S^1(x)$  and noting that the discrete  $\ell^{\frac{2p}{p-2}} - \ell^2$  inequality and the assumption (1.5) imply

$$\|[c_0 * \zeta]_{\infty}\|_{L^{\frac{2p}{p-2}}(\mathbb{R}^d)} \lesssim \|[c_0 * \zeta]_{\infty}\|_{L^2(\mathbb{R}^d)} \leq \|\zeta\|_{L^2(\mathbb{R}^d)}, \tag{9.8}$$

we get

$$\|U_1\|_{\text{op}} \lesssim \|[S^1]_1\|_{L^p(\mathbb{R}^d)},$$

hence,

$$\mathbb{E} \left[ \|U_1\|_{\text{op}}^4 \right]^{\frac{1}{4}} \leq \mathbb{E} \left[ \|U_1\|_{\text{op}}^p \right]^{\frac{1}{p}} \lesssim \|[S^1]_1\|_{L^p(\mathbb{R}^d; L^p(\Omega))}.$$

By definition of  $S^1$ , appealing to the corrector estimates of Proposition 2.2, we deduce

$$\begin{aligned} \mathbb{E} \left[ \|U_1\|_{\text{op}}^4 \right]^{\frac{1}{4}} &\lesssim \sum_{k=0}^n \sum_{s=0}^{n-1} p^{c_k+c_s} \left( \|\mu_{d,k}[\nabla^k g \nabla^{s+1} \bar{w}]\|_{L^p(\mathbb{R}^d)} \right. \\ &\quad \left. + \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k-l+s+1} \bar{w})]\|_{L^p(\mathbb{R}^d)} \right). \end{aligned}$$

We turn to  $U_2$ . By symmetry, it suffices to estimate the norm of  $\tilde{U}^2$ . From (9.7) and  $\|[c_0 * \zeta]_{\infty}\|_{L^2(\mathbb{R}^d)} \leq \|\zeta\|_{L^2(\mathbb{R}^d)}$ , we see

$$\|\tilde{U}_2\|_{\text{op}}^2 \lesssim \sup_{\|\zeta\|_{\infty} = 1} \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} \tilde{U}_2(x, \cdot) \zeta(x) dx \right]_1^2.$$

Taking the expectation and arguing by duality, we find

$$\mathbb{E} \left[ \|\tilde{U}_2\|_{\text{op}}^4 \right]^{\frac{1}{4}} \lesssim \sup_{\|\zeta\|_{\infty} = 1} M(\zeta), \tag{9.9}$$

where

$$M(\zeta) := \mathbb{E} \left[ \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} \tilde{U}_2(x, \cdot) \zeta(x) dx \right]_1^2 \right]^{\frac{1}{2}}.$$

Let  $\zeta \in C_c^\infty(\mathbb{R}^d; L^\infty(\Omega))$  be fixed with  $\|[\zeta]_\infty\|_{L^4(\Omega; L^2(\mathbb{R}^d))} = 1$ . Noting that Jensen’s inequality yields

$$\|[\zeta]_\infty\|_{L^2(\mathbb{R}^d; L^2(\Omega))} = \|[\zeta]_\infty\|_{L^2(\Omega; L^2(\mathbb{R}^d))} \leq \|[\zeta]_\infty\|_{L^4(\Omega; L^2(\mathbb{R}^d))},$$

and that the discrete  $\ell^4$ – $\ell^2$  inequality gives

$$\|[\zeta]_\infty\|_{L^4(\mathbb{R}^d; L^4(\Omega))} \lesssim \|[\zeta]_\infty\|_{L^4(\Omega; L^2(\mathbb{R}^d))},$$

we deduce by interpolation, for all  $2 \leq r \leq 4$ ,

$$\|[\zeta]_\infty\|_{L^r(\mathbb{R}^d; L^r(\Omega))} \lesssim \|[\zeta]_\infty\|_{L^4(\Omega; L^2(\mathbb{R}^d))} = 1. \tag{9.10}$$

By definition of  $\tilde{U}_2$  and by Lemma 7.1 in the form (7.1), we get

$$M(\zeta) \lesssim \sum_{s=0}^{n-1} \sum_{m=0}^s \sum_{\beta \in \{0,1\}} \| [U^m T^\beta ((a^*)^{-1} \zeta S^{2;s}) ]_2 ([\nabla \varphi^{s+1-m}]_2 + [\varphi^{s-m}]_2) \|_{L^2(\mathbb{R}^d; L^2(\Omega))},$$

and the corrector estimates of Proposition 2.2 yield for all  $p \geq 1$ ,

$$M(\zeta) \lesssim \sum_{s=0}^{n-1} \sum_{m=0}^s p^{cs-m} \sum_{\beta \in \{0,1\}} \| [U^m T^\beta ((a^*)^{-1} \zeta S^{2;s}) ]_2 \|_{L^2(\mathbb{R}^d; L^{\frac{2p}{p-1}}(\Omega))}. \tag{9.11}$$

For  $p \geq 2$ , applying  $m$  times the bound of Corollary 6.2 and once the one of Theorem 6.1 with  $\delta = \frac{1}{m+1} \frac{2p}{(p-1)(p-2)} \sim \frac{1}{p}$  in order to pass from the stochastic integrability  $\frac{2p}{p-1}$  to  $\frac{2p}{p-2}$ , we obtain the following,

$$M(\zeta) \lesssim \sum_{s=0}^{n-1} \sum_{m=0}^s p^{cs-m} (p \log^2 p)^{\frac{m(m+1)}{2d}} \|[\zeta S^{2;s}]_2\|_{L^{\frac{2d}{d+2m}}(\mathbb{R}^d; L^{\frac{2p}{p-2}}(\Omega))},$$

where the exponent  $\frac{m(m+1)}{d}$  of  $p \log^2 p$  comes from the large prefactors in Corollary 6.2 and Theorem 6.1 and is due to the successive deviations of spatial integrability from 2. Hölder’s inequality together with (9.10) yields for  $p \geq 8$ ,

$$\begin{aligned} \|[\zeta S^{2;s}]_2\|_{L^{\frac{2d}{d+2m}}(\mathbb{R}^d; L^{\frac{2p}{p-2}}(\Omega))} &\leq \|[\zeta]_\infty\|_{L^{\frac{2p}{p-4}}(\mathbb{R}^d; L^{\frac{2p}{p-4}}(\Omega))} \| [S^{2;s}]_2 \|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d; L^p(\Omega))} \\ &\lesssim \| [S^{2;s}]_2 \|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d; L^p(\Omega))}. \end{aligned}$$

By definition of  $S^{2;s}$ , using the corrector estimates of Proposition 2.2, we conclude

$$\begin{aligned} \mathbb{E} \left[ \|\tilde{U}_2\|_{\text{op}}^4 \right]^{\frac{1}{4}} &\lesssim \sum_{s=0}^{n-1} \sum_{m=0}^s p^{c_s-m} (p \log^2 p)^{\frac{m(m+1)}{2d}} \sum_{k=0}^n p^{c_k} \\ &\left( \|\mu_{d,k}[\nabla^k g \nabla^{s+1} \bar{w}]\|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d)} \right. \\ &\left. + \sum_{l=0}^{k+s-n} \|\mu_{d,k}[\nabla^l (g \nabla^{k-l+s+1} \bar{w})]\|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d)} \right). \end{aligned} \tag{9.12}$$

It remains to analyze  $U_3$ . For  $p \geq 2$ , we first appeal to the definition of  $U_3$  and to Lemma 9.1 (with  $p$  replaced by  $\frac{p}{2}$ ) in the form

$$\begin{aligned} &\left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \zeta(x) U_3(x, y) \zeta(y) dx dy \right| \\ &\lesssim \|\zeta\|_{\infty}^2 \left\| \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \left\| [U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})]_2 (\nabla \varphi^{s+1-k} \mathbf{1}_2 + [\varphi^{s-k} \mathbf{1}_2]) \right\|_{L^{\frac{p}{2}}(\mathbb{R}^d)} \right\|_{L^{\frac{2p}{p-2}}(\mathbb{R}^d)} \\ &+ \|\zeta\|_{\infty} \left\| \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} J_{s-k} (\zeta (\mathbf{a}^*)^{-1} a'_0(G)^* U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})) \right\|_{L^2(\mathbb{R}^d)}. \end{aligned}$$

From (9.7) and (9.8), we deduce

$$\begin{aligned} \|U_3\|_{\text{op}} &\lesssim \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \left\| [U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})]_2 (\nabla \varphi^{s+1-k} \mathbf{1}_2 + [\varphi^{s-k} \mathbf{1}_2]) \right\|_{L^{\frac{p}{2}}(\mathbb{R}^d)} \\ &+ \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \sup_{\|\zeta\|_{\infty} \| \zeta \|_{L^2(\mathbb{R}^d)} = 1} J_{s-k} (\zeta (\mathbf{a}^*)^{-1} a'_0(G)^* U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})). \end{aligned}$$

Taking the  $L^4(\Omega)$  norm and using the corrector estimates of Proposition 2.2 in the first right-hand side term, and arguing by duality as in (9.9) and repeating the proof of (9.12) for the second right-hand side term, we obtain for all  $8 \leq p < \infty$ ,

$$\begin{aligned} \mathbb{E} \left[ \|U_3\|_{\text{op}}^4 \right]^{\frac{1}{4}} &\lesssim p \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \left\| [U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})]_2 \right\|_{L^{\frac{p}{2}}(\mathbb{R}^d; L^p(\Omega))} \\ &+ \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{\beta \in \{0,1\}} \sum_{m=0}^{s-k} \left\| [U^k T^\beta ((\mathbf{a}^*)^{-1} S^{3;s})]_2 \right\|_{L^{\frac{dp}{2d+mp}}(\mathbb{R}^d; L^p(\Omega))}. \end{aligned}$$

A multiple use of the annealed bounds of Theorem 6.1 and Corollary 6.2 then leads to

$$\begin{aligned} \mathbb{E} \left[ \|U_3\|_{\text{op}}^4 \right]^{\frac{1}{4}} &\lesssim_p \sum_{s=0}^{n-1} \sum_{k=0}^s \sum_{m=0}^{s-k} \| [S^{3;s}]_2 \|_{L^{\frac{dp}{2d+(k+m)p}}(\mathbb{R}^d; L^{2p}(\Omega))} \\ &\lesssim_p \sum_{s=0}^{n-1} \| [S^{3;s}]_2 \|_{L^{\frac{dp}{2d+sp}}(\mathbb{R}^d; L^{2p}(\Omega))}. \end{aligned}$$

By definition of  $S^{3;s}$ , using the corrector estimates of Proposition 2.2, we conclude

$$\mathbb{E} \left[ \|U_3\|_{\text{op}}^4 \right]^{\frac{1}{4}} \lesssim_p \sum_{s=0}^{n-1} \sum_{l=s+1}^n \| \mu_{d,n} [\nabla^l (g \nabla^{n-l+s+1} \bar{w})]_{\infty} \|_{L^{\frac{dp}{2d+sp}}(\mathbb{R}^d)}.$$

Gathering the above estimates and replacing  $p$  by  $q$  in the estimate for  $U^3$ , the claim follows.

*Step 2. Conclusion.*

By scaling, the conclusion of Step 1 yields for all  $8 \leq p, q < \infty$  and  $1 \leq n \leq \ell$ ,

$$\begin{aligned} \varepsilon^{-\frac{d}{2}} \mathbb{E} \left[ \left\| D^2 \int_{\mathbb{R}^d} g \cdot \Xi_{\varepsilon}^{\circ,n} [\nabla \bar{w}] \right\|_{\text{op}}^4 \right]^{\frac{1}{4}} &\lesssim_{g,\bar{w}} \varepsilon^{\frac{d}{2} - \frac{d}{p}} \sum_{k=0}^n \sum_{s=0}^{n-1} \mu_{d,k} \left(\frac{1}{\varepsilon}\right) \varepsilon^{k+s} p^{c_k+c_s} \\ &+ \varepsilon^{\frac{d}{2} - \frac{2d}{p}} \sum_{k=0}^n \sum_{s=0}^{n-1} \sum_{m=0}^s \mu_{d,k} \left(\frac{1}{\varepsilon}\right) \varepsilon^{k+s-m} p^{c_k+c_s-m} (p \log^2 p)^{\frac{m(m+1)}{2d}} + C_q \varepsilon^{\frac{d}{2} - \frac{2d}{q}} \mu_{d,n} \left(\frac{1}{\varepsilon}\right) \varepsilon^n, \end{aligned}$$

hence, choosing  $p = |\log \varepsilon|$  and  $q = 8 \vee 4d$ , and recalling that  $c_0 = \frac{1}{2}$ ,

$$\varepsilon^{-\frac{d}{2}} \mathbb{E} \left[ \left\| D^2 \int_{\mathbb{R}^d} g \cdot \Xi_{\varepsilon}^{\circ,n} [\nabla \bar{w}] \right\|_{\text{op}}^4 \right]^{\frac{1}{4}} \lesssim_{f,g} \varepsilon^{\frac{d}{2}} |\log \varepsilon| (|\log \varepsilon| \log^2 |\log \varepsilon|)^{\frac{1}{2d}n(n-1)}.$$

In terms of  $X_{\varepsilon}^n := \varepsilon^{-\frac{d}{2}} \int_{\mathbb{R}^d} g \cdot \Xi_{\varepsilon}^{\circ,n} [\nabla \bar{w}]$ , we now apply Proposition 4.1(iv) in the form

$$d_{\mathcal{N}}(X_{\varepsilon}^n) \lesssim \frac{\mathbb{E} \left[ \|D^2 X_{\varepsilon}^n\|_{\text{op}}^4 \right]^{\frac{1}{4}} \mathbb{E} \left[ \|DX_{\varepsilon}^n\|_{\mathfrak{H}}^4 \right]^{\frac{1}{4}}}{\text{Var} [X_{\varepsilon}^n]}$$

Inserting the above estimate on  $\|D^2 X_{\varepsilon}^n\|_{\text{op}}$  and noting that the proof of Theorem 1(iii) yields  $\mathbb{E} \left[ \|DX_{\varepsilon}^n\|_{\mathfrak{H}}^4 \right] \lesssim_{f,g} 1$ , the conclusion follows. □

**Acknowledgements** The work of MD was supported by F.R.S.-FNRS and by the CNRS-Momentum program.

## Appendix A. More Malliavin calculus

In this appendix, we include for completeness a short, essentially self-contained proof of Proposition 4.1.

**Proof of Proposition 4.1** We split the proof into four steps.

*Step 1.* Proof of (i).

In terms of the Ornstein–Uhlenbeck semigroup  $e^{-t\mathcal{L}}$  (e.g. [44, Section 1.4]), we may write

$$\text{Var}[X] = - \int_0^\infty \partial_t \mathbb{E} \left[ (e^{-t\mathcal{L}} X)^2 \right] dt$$

where Mehler’s formula (e.g. [44, (1.67)]) indeed ensures that  $\mathbb{E} \left[ (e^{-t\mathcal{L}} X)^2 \right] \rightarrow \mathbb{E}[X]^2$  as  $t \uparrow \infty$ . Computing the derivative in  $t$  yields

$$\text{Var}[X] = 2 \int_0^\infty \mathbb{E} \left[ (e^{-t\mathcal{L}} X) \mathcal{L}(e^{-t\mathcal{L}} X) \right] dt = 2 \int_0^\infty \mathbb{E} \left[ \|De^{-t\mathcal{L}} X\|_{\mathfrak{H}}^2 \right] dt,$$

hence, appealing to the commutation relation (4.5) in the form  $De^{-t\mathcal{L}} = e^{-t} e^{-t\mathcal{L}} D$ ,

$$\text{Var}[X] = 2 \int_0^\infty e^{-2t} \mathbb{E} \left[ \|e^{-t\mathcal{L}} DX\|_{\mathfrak{H}}^2 \right] dt,$$

and the positivity of  $\mathcal{L}$  leads to the conclusion,

$$\text{Var}[X] \leq 2 \int_0^\infty e^{-2t} \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^2 \right] dt = \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^2 \right].$$

*Step 2.* Proof of (ii).

Let  $X, Y \in \mathbb{D}^{1,2}$  with  $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ . First note that the Poincaré inequality (i) implies that the restriction of  $\mathcal{L}$  to  $L^2(\Omega)/\mathbb{C} := \{U \in L^2(\Omega) : \mathbb{E}[U] = 0\}$  is invertible. In particular, there exists  $Z \in L^2(\Omega)$  such that  $Y = \mathcal{L}Z$ . By definition of the adjoint  $D^*$  (cf. (4.4)), we then find

$$\mathbb{E}[XY] = \mathbb{E}[X\mathcal{L}Z] = \mathbb{E}[XD^*DZ] = \mathbb{E}[\langle DX, DZ \rangle_{\mathfrak{H}}]. \quad (\text{A.1})$$

Appealing to the commutator relation (4.5) in the form  $DY = D\mathcal{L}Z = (1 + \mathcal{L})DZ$ , we conclude

$$\mathbb{E}[XY] = \mathbb{E} \left[ \langle DX, (1 + \mathcal{L})^{-1}DY \rangle_{\mathfrak{H}} \right].$$

*Step 3.* Proof of (iii).

Similarly as in Step 1, in terms of the Ornstein–Uhlenbeck semigroup  $e^{-t\mathcal{L}}$ , we may write



$$\begin{aligned} \text{Ent}[X^2] &= - \int_0^\infty \partial_t \mathbb{E} \left[ (e^{-t\mathcal{L}} X^2) \log(e^{-t\mathcal{L}} X^2) \right] dt \\ &= \int_0^\infty \mathbb{E} \left[ (\mathcal{L} e^{-t\mathcal{L}} X^2) \log(e^{-t\mathcal{L}} X^2) \right] dt = \int_0^\infty \mathbb{E} \left[ \frac{\|D e^{-t\mathcal{L}} X^2\|_{\mathfrak{H}}^2}{e^{-t\mathcal{L}} X^2} \right] dt, \end{aligned}$$

hence, by the commutation relation (4.5) in the form  $D e^{-t\mathcal{L}} = e^{-t} e^{-t\mathcal{L}} D$ ,

$$\text{Ent}[X^2] = \int_0^\infty e^{-2t} \mathbb{E} \left[ \frac{\|e^{-t\mathcal{L}} D X^2\|_{\mathfrak{H}}^2}{e^{-t\mathcal{L}} X^2} \right] dt.$$

Appealing to Mehler’s formula for the Ornstein–Uhlenbeck semigroup  $e^{-t\mathcal{L}}$  (e.g. [44, (1.67)]), the Cauchy–Schwarz inequality leads to

$$\|e^{-t\mathcal{L}} D X^2\|_{\mathfrak{H}}^2 = 4 \|e^{-t\mathcal{L}}(XDX)\|_{\mathfrak{H}}^2 \leq 4 (e^{-t\mathcal{L}} X^2)(e^{-t\mathcal{L}} \|DX\|_{\mathfrak{H}}^2),$$

so that the above becomes

$$\text{Ent}[X^2] \leq 4 \int_0^\infty e^{-2t} \mathbb{E} \left[ e^{-t\mathcal{L}} \|DX\|_{\mathfrak{H}}^2 \right] dt \leq 2 \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^2 \right],$$

which proves the logarithmic Sobolev inequality. Next, higher integrability is deduced by integration as e.g. in [2, Theorem 3.4]: we write

$$\mathbb{E} \left[ |X|^{2p} \right]^{\frac{1}{p}} - \mathbb{E} \left[ X^2 \right] = \int_1^p \frac{1}{q^2} \mathbb{E} \left[ |X|^{2q} \right]^{\frac{1}{q}-1} \text{Ent} \left[ |X|^{2q} \right] dq,$$

so that the logarithmic Sobolev inequality implies

$$\begin{aligned} \mathbb{E} \left[ |X|^{2p} \right]^{\frac{1}{p}} - \mathbb{E} \left[ X^2 \right] &\leq 2 \int_1^p \frac{1}{q^2} \mathbb{E} \left[ |X|^{2q} \right]^{\frac{1}{q}-1} \mathbb{E} \left[ \|D|X|^q\|_{\mathfrak{H}}^2 \right] dq \\ &= 2 \int_1^p \mathbb{E} \left[ |X|^{2q} \right]^{\frac{1}{q}-1} \mathbb{E} \left[ |X|^{2(q-1)} \|DX\|_{\mathfrak{H}}^2 \right] dq \\ &\leq 2 \int_1^p \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^{2q} \right]^{\frac{1}{q}} dq \leq 2p \mathbb{E} \left[ \|DX\|_{\mathfrak{H}}^{2p} \right]^{\frac{1}{p}}, \end{aligned}$$

and the conclusion follows from the Poincaré inequality (i).

*Step 4. Proof of (iv).*

Let  $X \in L^2(\Omega)$  with  $\mathbb{E}[X] = 0$  and  $\text{Var}[X] = 1$ . For  $h \in L^\infty(\mathbb{R})$ , we define its Stein transform  $S_h$  as the solution of Stein’s equation

$$S'_h(x) - x S_h(x) = h(x) - \mathbb{E}[h(\mathcal{N})].$$

As in Step 2, there exists  $Z \in L^2(\Omega)$  such that  $Y = \mathcal{L}Z$ . We then compute

$$\mathbb{E}[h(X)] - \mathbb{E}[h(\mathcal{N})] = \mathbb{E}[S'_h(X) - XS_h(X)] = \mathbb{E}[S'_h(X) - (D^*DZ)S_h(X)],$$

and hence, integrating by parts and using (A.1) in the form  $\mathbb{E}[\langle DZ, DX \rangle_{\mathfrak{H}}] = \text{Var}[X] = 1$ ,

$$\begin{aligned} |\mathbb{E}[h(X)] - \mathbb{E}[h(\mathcal{N})]| &= |\mathbb{E}[S'_h(X)(1 - \langle DZ, DX \rangle_{\mathfrak{H}})]| \\ &\leq \|S'_h\|_{L^\infty} \text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}]^{\frac{1}{2}}. \end{aligned} \tag{A.2}$$

Noting that  $\|S'_h\|_{L^\infty} \leq 2\|h\|_{L^\infty}$  (e.g. [42, Theorem 3.3.1]) and taking the supremum over all  $h \in L^\infty(\mathbb{R})$ , we deduce

$$d_{\text{TV}}(X; \mathcal{N}) \leq 2 \text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}]^{\frac{1}{2}}.$$

Noting that for  $h$  Lipschitz-continuous there holds  $\|S'_h\|_{L^\infty} \leq \sqrt{\frac{2}{\pi}}\|h'\|_{L^\infty}$  (e.g. [42, Proposition 3.5.1]), we can deduce a similar bound on the 1-Wasserstein distance. The corresponding bound on the 2-Wasserstein distance takes the form

$$W_2(X; \mathcal{N}) \leq \mathbb{E} \left[ |\langle DZ, DX \rangle_{\mathfrak{H}} - 1|^2 \right]^{\frac{1}{2}} = \text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}]^{\frac{1}{2}};$$

its proof is of a different nature and is based on an optimal transport argument in density space (cf. [35, Proposition 3.1]).

It remains to estimate the variance  $\text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}]$ . For that purpose, we apply the first-order Poincaré inequality (i) in the form

$$\text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}] \leq \mathbb{E} \left[ \|\langle D^2Z, DX \rangle_{\mathfrak{H}} + \langle DZ, D^2X \rangle_{\mathfrak{H}}\|_{\mathfrak{H}}^2 \right].$$

Noting that the commutation relation (4.5) leads to  $DX = (1 + \mathcal{L})DZ$  and  $D^2X = (2 + \mathcal{L})D^2Z$ , we deduce

$$\begin{aligned} \text{Var}[\langle DZ, DX \rangle_{\mathfrak{H}}] &\leq \mathbb{E} \left[ \|\langle D^2X, ((1 + \mathcal{L})^{-1} + (2 + \mathcal{L})^{-1})DX \rangle_{\mathfrak{H}}\|_{\mathfrak{H}}^2 \right] \\ &\leq \mathbb{E} \left[ \|D^2X\|_{\text{op}}^4 \right]^{\frac{1}{4}} \mathbb{E} \left[ \|\langle ((1 + \mathcal{L})^{-1} + (2 + \mathcal{L})^{-1})DX \rangle_{\mathfrak{H}}\|_{\mathfrak{H}}^4 \right]^{\frac{1}{4}}, \end{aligned}$$

Noting as in [38, Proposition 3.2] that  $(1 + \mathcal{L})^{-1}$  and  $(2 + \mathcal{L})^{-1}$  have operator norms on  $L^4(\Omega)$  bounded by 1 and  $\frac{1}{2}$ , respectively, the conclusion follows.  $\square$

### References

1. Abdulle, A., Pouchon, T.: Effective models and numerical homogenization for wave propagation in heterogeneous media on arbitrary timescales. Preprint. [arXiv:1905.09062](https://arxiv.org/abs/1905.09062)

2. Aida, S., Stroock, D.: Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1**(1), 75–86 (1994)
3. Armstrong, S., Kuusi, T., Mourrat, J.-C.: *Quantitative Stochastic Homogenization and Large-Scale Regularity*, Volume 352 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham (2019)
4. Armstrong, S.N., Daniel, J.-P.: Calderón–Zygmund estimates for stochastic homogenization. *J. Funct. Anal.* **270**(1), 312–329 (2016)
5. Armstrong, S.N., Kuusi, T., Mourrat, J.-C.: The additive structure of elliptic homogenization. *Invent. Math.* **208**, 999–1154 (2017)
6. Armstrong, S.N., Mourrat, J.-C.: Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.* **219**(1), 255–348 (2016)
7. Armstrong, S.N., Smart, C.K.: Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér.* (4) **49**(2), 423–481 (2016)
8. Bakry, D.: L'hypercontractivité et son utilisation en théorie des semigroupes. In: *Lectures on Probability Theory (Saint-Flour., 1992)*, Volume 1581 of *Lecture Notes in Mathematics*, pp. 1–114. Springer, Berlin (1994)
9. Bella, P., Fehrman, B., Fischer, J., Otto, F.: Stochastic homogenization of linear elliptic equations: higher-order error estimates in weak norms via second-order correctors. *SIAM J. Math. Anal.* **49**(6), 4658–4703 (2017)
10. Bella, P., Giunti, A., Otto, F.: Quantitative stochastic homogenization: local control of homogenization error through corrector. In: *Mathematics and Materials*, Volume 23 of *IAS/Park City Mathematics Series*, pp. 301–327. American Mathematical Society, Providence (2017)
11. Bensoussan, A., Lions, J.-L., Papanicolaou, G.: *Asymptotic Analysis for Periodic Structures*, Volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam (1978)
12. Bourgain, J.: On a homogenization problem. *J. Stat. Phys.* **172**(2), 314–320 (2018)
13. Caffarelli, L.A., Peral, I.: On  $W^{1,p}$  estimates for elliptic equations in divergence form. *Commun. Pure Appl. Math.* **51**(1), 1–21 (1998)
14. Chatterjee, S.: Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Relat. Fields* **143**(1–2), 1–40 (2009)
15. Chernoff, H.: A note on an inequality involving the normal distribution. *Ann. Probab.* **9**(3), 533–535 (1981)
16. Duerinckx, M., Fischer, J., Gloria, A.: Scaling limit of the homogenization commutator for Gaussian coefficient fields. Preprint. [arXiv:1910.04088](https://arxiv.org/abs/1910.04088)
17. Duerinckx, M., Gloria, A., Lemm, M.: A remark on a surprising result by Bourgain in homogenization. *Commun. Partial Differ. Equ.* **44**(12), 1345–1357 (2019)
18. Duerinckx, M., Gloria, A., Otto, F.: Robustness of the pathwise structure of fluctuations in stochastic homogenization. Preprint. [arXiv:1807.11781](https://arxiv.org/abs/1807.11781)
19. Duerinckx, M., Gloria, A., Otto, F.: The structure of fluctuations in stochastic homogenization. Preprint. [arXiv:1602.01717](https://arxiv.org/abs/1602.01717)
20. Gloria, A., Neukamm, S., Otto, F.: Quantitative estimates in stochastic homogenization for correlated coefficient fields. Preprint. [arXiv:1910.05530](https://arxiv.org/abs/1910.05530)
21. Gloria, A., Neukamm, S., Otto, F.: A regularity theory for random elliptic operators. Preprint. [arXiv:1409.2678](https://arxiv.org/abs/1409.2678)
22. Gloria, A., Neukamm, S., Otto, F.: An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *M2AN Math. Model. Numer. Anal.* **48**(2), 325–346 (2014)
23. Gloria, A., Otto, F.: The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. Preprint. [arXiv:1510.08290](https://arxiv.org/abs/1510.08290)
24. Gu, Y.: High order correctors and two-scale expansions in stochastic homogenization. *Probab. Theory Relat. Fields* **169**(3–4), 1221–1259 (2017)
25. Gu, Y., Mourrat, J.-C.: Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.* **14**(1), 452–481 (2016)
26. Helffer, B., Sjöstrand, J.: On the correlation for Kac-like models in the convex case. *J. Stat. Phys.* **74**(1–2), 349–409 (1994)
27. Hill, R.: Elastic properties of reinforced solids: some theoretical principles. *J. Mech. Phys. Solids* **11**, 357–372 (1963)
28. Hill, R.: On constitutive macro-variables for heterogeneous solids at finite strain. *Proc. R. Soc. Lond. A* **326**, 131–146 (1972)

29. Houdré, C., Pérez-Abreu, V.: Covariance identities and inequalities for functionals on Wiener and Poisson spaces. *Ann. Probab.* **23**(1), 400–419 (1995)
30. Jikov, V.V., Kozlov, S.M., Oleĭnik, O.A.: *Homogenization of Differential Operators and Integral Functionals*. Springer, Berlin (1994). Translated from Russian by G. A. Iosif'yan
31. Josien, M., Otto, F.: Introduction to stochastic homogenization: oscillations and fluctuations. In preparation (2019)
32. Kamotski, V., Matthies, K., Smyshlyaev, V.P.: Exponential homogenization of linear second order elliptic PDEs with periodic coefficients. *SIAM J. Math. Anal.* **38**(5), 1565–1587 (2006/07)
33. Kim, J., Lemm, M.: On the averaged Green's function of an elliptic equation with random coefficients. Preprint. [arXiv:1804.10260](https://arxiv.org/abs/1804.10260)
34. Kozlov, S.M.: The averaging of random operators. *Mat. Sb. (N.S.)* **109**(151)(2):188–202, 327 (1979)
35. Ledoux, M., Nourdin, I., Peccati, G.: Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.* **25**(1), 256–306 (2015)
36. Malliavin, P.: *Stochastic Analysis*, Volume 313 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (1997)
37. Marahrens, D., Otto, F.: Annealed estimates on the Green's function. *Probab. Theory Relat. Fields* **163**(3–4), 527–573 (2015)
38. Mourrat, J.-C., Otto, F.: Correlation structure of the corrector in stochastic homogenization. *Ann. Probab.* **44**(5), 3207–3233 (2016)
39. Murat, F., Tartar, L.: *H*-convergence. In: *Topics in the Mathematical Modelling of Composite Materials*, Volume 31 of Progress Nonlinear Differential Equations Applications, pp. 21–43. Birkhäuser, Boston (1997)
40. Nash, J.: Continuity of solutions of parabolic and elliptic equations. *Am. J. Math.* **80**, 931–954 (1958)
41. Nourdin, I., Peccati, G.: Stein's method on Wiener chaos. *Probab. Theory Relat. Fields* **145**(1–2), 75–118 (2009)
42. Nourdin, I., Peccati, G.: *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality*, Volume 192 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge (2012)
43. Nourdin, I., Peccati, G., Reinert, G.: Second order Poincaré inequalities and CLTs on Wiener space. *J. Funct. Anal.* **257**(2), 593–609 (2009)
44. Nualart, D.: *The Malliavin Calculus and Related Topics*, 2nd edn. Springer, Berlin (2006)
45. Papanicolaou, G.C., Varadhan, S.R.S.: Boundary value problems with rapidly oscillating random coefficients. In: *Random Fields*, Vol. I, II (Esztergom, 1979), Volume 27 of Colloq. Math. Soc. János Bolyai, pp. 835–873. North-Holland, Amsterdam (1981)
46. Peccati, G., Reitzner, M. (eds.): *Stochastic Analysis for Poisson Point Processes*, Volume 7 of Bocconi & Springer Series. Bocconi University Press, Springer, Berlin (2016)
47. Pouchon, T.: *Effective models and numerical homogenization methods for long time wave propagation in heterogeneous media*. Ph.D. thesis, École Polytechnique Fédérale de Lausanne (2017)
48. Shen, Z.: The  $L^p$  boundary value problems on Lipschitz domains. *Adv. Math.* **216**, 212–254 (2007)
49. Shen, Z.: The Calderón-Zygmund lemma revisited. In: *Lectures on the Analysis of Nonlinear Partial Differential Equations. Part 2*, Volume 2 of Morningside Lecture Notes in Mathematics, pp. 203–224. International Press, Somerville (2012)
50. Sjöstrand, J.: Correlation asymptotics and Witten Laplacians. *Algebra Anal.* **8**(1), 160–191 (1996)
51. Tran, T.-H., Monchiet, V., Bonnet, G.: A micromechanics-based approach for the derivation of constitutive elastic coefficients of strain-gradient media. *Int. J. Solids Struct.* **49**(5), 783–792 (2012)