

## Optimal artificial boundary conditions based on second-order correctors for three dimensional random elliptic media

Jianfeng Lu, Felix Otto & Lihan Wang

**To cite this article:** Jianfeng Lu, Felix Otto & Lihan Wang (2024) Optimal artificial boundary conditions based on second-order correctors for three dimensional random elliptic media, Communications in Partial Differential Equations, 49:7-8, 609-670, DOI: [10.1080/03605302.2024.2374568](https://doi.org/10.1080/03605302.2024.2374568)

**To link to this article:** <https://doi.org/10.1080/03605302.2024.2374568>



Published online: 17 Jul 2024.



Submit your article to this journal [↗](#)



Article views: 79



View related articles [↗](#)



View Crossmark data [↗](#)



# Optimal artificial boundary conditions based on second-order correctors for three dimensional random elliptic media

Jianfeng Lu, Felix Otto, and Lihan Wang

Mathematics Department, Duke University, Durham, NC, USA; Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany; Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

## ABSTRACT

We are interested in numerical algorithms for computing the electrical field generated by a charge distribution localized on scale  $\ell$  in an infinite heterogeneous medium, in a situation where the medium is only known in a box of diameter  $L \gg \ell$  around the support of the charge. We propose a boundary condition that with overwhelming probability is (near) optimal with respect to scaling in terms of  $\ell$  and  $L$ , in the setting where the medium is a sample from a stationary ensemble with a finite range of dependence (set to be unity and with the assumption that  $\ell \gg 1$ ). The boundary condition is motivated by quantitative stochastic homogenization that allows for a multipole expansion. This work extends, the algorithm in which is optimal in two dimension, and thus we need to take quadrupoles, next to dipoles, into account. This in turn relies on stochastic estimates of second-order, next to first-order, correctors. These estimates are provided for finite range ensembles under consideration, based on an extension of the semi-group approach of Gloria and Otto.

## ARTICLE HISTORY

Received 4 June 2022  
Accepted 26 June 2024

## KEYWORDS

Stochastic homogenization;  
artificial boundary condition;  
correctors; multipoles;  
random media

## 1. Introduction and main results

Consider a conducting medium as described by a symmetric  $\lambda$ -uniform coefficient field  $a = a(x)$  in  $d$ -dimensional space, that is, for any  $x, \xi \in \mathbb{R}^d$

$$\lambda |\xi|^2 \leq \xi \cdot a(x) \xi \leq |\xi|^2. \quad (1)$$

Consider a localized charge distribution that is overall neutral, as described by a compactly supported dipole density  $g$ . Let us give a sense to its characteristic scale  $\ell$  by assuming that it is of the form

$$g(x) = \hat{g}\left(\frac{x}{\ell}\right) \quad (2)$$

for some sufficiently smooth  $\hat{g}$  supported in the unit ball. We are interested in the field  $\nabla u$  the charge generates, which is the decaying solution of the elliptic divergence-form equation

$$\nabla \cdot (a \nabla u + g) = 0 \quad \text{in } \mathbb{R}^d. \quad (3)$$

In this paper, we address the following question: Suppose we only know the medium  $a$  in some box  $Q_{2L} := (-2L, 2L)^d$ , to what precision may we infer the value of  $\nabla u$ ? What is a practical algorithm to retrieve it? Heuristically, one expects  $\nabla u(x)$  to decay as a dipole, i.e., like  $(\frac{\ell}{|x|})^d$ , thus we expect that changing the coefficient field  $a$  outside the box  $Q_{2L}$  will affect  $\nabla u$  to order  $(\frac{\ell}{L})^d$ , and imposing homogeneous Dirichlet conditions on  $\partial Q_L$  would do no worse – and this would be the end of the story and the paper.

In this paper, however, we consider a more specific situation, namely when  $a$  is sampled from a stationary ensemble  $\langle \cdot \rangle$ , which puts us into the context of stochastic homogenization. More precisely, we shall assume that  $\langle \cdot \rangle$  is of finite range, which we set to be unity without loss of generality. This means that for two sets  $D$  and  $D' \subset \mathbb{R}^d$  with distance larger than 1, the restrictions  $a|_D$  and  $a|_{D'}$  are independent. What information may we retrieve in this case? Consider again changing the coefficient field outside  $Q_{2L}$  (now with  $L \gg 1$ ) but keeping the statistical ensemble, heuristically, we expect that due to stochastic cancellations, the impact on  $\nabla u$  reduces to  $\frac{1}{\sqrt{L^d}}(\frac{\ell}{L})^d$ . Indeed, the additional attenuation factor  $\frac{1}{\sqrt{L^d}}$  comes from the Central Limit Theorem (CLT) scaling involving the square root of the relevant volume, non-dimensionalized by the correlation length. In fact, this precision on inferring the value of  $\nabla u$  cannot be improved, as the following lower bound on the variance of  $\nabla u$  from previous work [2], conditioned on the restriction  $a|_{Q_{2L}}$ , shows, which we expect to hold for generic ensembles.

**Theorem 1.1.** [2, Theorem 2] *There exists a stationary, unit-range ensemble  $\langle \cdot \rangle$  supported on  $a$ 's satisfying (1) with the following property: Consider the solution  $u$  of (3), where  $g$  is of the form (2) for some  $\ell$  and  $\hat{g}$ , then there exists a radius  $R$  such that for any  $\omega = \frac{1}{R^d} \hat{\omega}(\frac{x}{R})$  for some sufficiently smooth  $\hat{\omega}$  supported in<sup>1</sup>  $B_1$  with  $\int \hat{\omega} = 1$ ,*

$$\langle |\int \omega \nabla u - \langle \int \omega \nabla u | a|_{Q_{2L}} \rangle|^2 \rangle^{\frac{1}{2}} \geq \frac{1}{C} (\frac{\ell}{L})^d (\frac{1}{L})^{\frac{d}{2}} \quad \text{provided} \quad \frac{L}{C} \geq \ell \geq C.$$

Here the radius  $R$  and the constant  $C$  depend only on the ensemble,  $\hat{g}$ , and  $\hat{\omega}$ .

In [2], a practical algorithm, that saturates this scaling for  $d = 2$ , was proposed and analyzed. The error of such an algorithm is  $O((\frac{\ell}{L})^d (\frac{1}{L})^{1-})$  in any dimension. In this paper, we tackle the more physically relevant case of  $d = 3$ , which requires a substantial modification of the algorithm and its analysis.

We will propose a deterministic algorithm, [Algorithm 1](#), that involves the realization  $a$  only in terms of its restriction  $a|_{Q_{2L}}$ . The algorithm saturates the lower bound of [Theorem 1.1](#) in terms of scaling. More specifically, by solving a couple of auxiliary boundary value problems with homogeneous Dirichlet boundary conditions on  $Q_{2L}$ ,  $Q_{\frac{7}{4}L}$ , and  $Q_{\frac{3}{2}L}$ , this algorithm constructs Dirichlet boundary data  $u_L$  on  $\partial Q_L$ , which in turn defines  $u^{(L)}$ , the output of the algorithm, by solving

$$-\nabla \cdot a \nabla u^{(L)} = \nabla \cdot g \text{ in } Q_L, \quad u^{(L)} = u_L \text{ on } \partial Q_L. \quad (4)$$

For  $L, \ell, R \gg 1$  and with overwhelming probability, this algorithm saturates the lower bound of [Theorem 1.1](#), with the little caveat that the CLT exponent,  $\frac{3}{2}$  for  $d = 3$ , has to be replaced by  $\beta < \frac{3}{2}$ . Part of the probabilistic nature of the statement is contained in the random radius

---

<sup>1</sup>Throughout this work we use  $B_R$  to denote the ball  $\{|x| \leq R\}$ .

$r_{**}$ , which can be interpreted as the scale from which onwards stochastic homogenization is effective.

**Theorem 1.2.** *Let  $d = 3$  and  $\langle \cdot \rangle$  be a stationary, unit-range ensemble supported on  $a$ 's satisfying (1). Let  $g$  be of the form (2) for some  $\ell$  and  $\hat{g}$ , let  $u$  denote the solution of (3), and  $u^{(L)}$  be the output of Algorithm 1 for some  $L \geq \ell$ . Then for any  $\beta < \frac{3}{2}$ , there exists a random radius  $r_{**}$  such that conditioning on  $\ell \geq r_{**}$ , with probability at least  $1 - \exp(-L^{1/C})$ , we have for any  $R \in [r_{**}, L]$ ,<sup>2</sup>*

$$\left( \int_{B_R} |\nabla(u^{(L)} - u)|^2 \right)^{\frac{1}{2}} \leq C \left( \frac{\ell}{L} \right)^d \left( \frac{r_{**}}{L} \right)^\beta \quad \text{provided} \quad \frac{L}{C} \geq \ell \geq C.$$

Moreover, the radius  $r_{**}$  satisfies

$$\langle \exp(r_{**}^\beta) \rangle \leq C. \quad (5)$$

Here  $C$  denotes a constant that depends only on  $\lambda$ ,  $\hat{g}$ ,  $\beta$  and  $\hat{\omega}$  appearing in Theorem 1.1.

We believe that with some additional work, it is possible to derive an “a posteriori” style result similar to [2], that is, we could define some computable  $r_{**}^{(L)}$  that plays the same role as  $r_{**}$  in Theorem 1.2. We also comment here that Theorem 1.2 holds with any  $\beta < 2$  for  $d \geq 4$ , and the algorithm is thus also near-optimal when  $d = 4$ . Obtaining the optimal algorithm for  $d \geq 5$  requires computing correctors of order three or higher and we do not discuss it here.

Before we further discuss ideas of Algorithm 1 and the proof of the theorem, let us first compare our result with the previous work [2]. A main difference lies in the introduction of the functions  $\psi_T^{(L)}$  and the corresponding coefficients  $c_T^{(L)}$ , which are approximations of second-order correctors and quadrupoles that will be introduced below. These are available for  $d > 2$  and indeed necessary for the algorithm to (almost) reach the CLT-scaling  $\beta < \frac{3}{2}$ . This aspect of difference will be discussed in more details in Section 2.1. A more technical difference is that Algorithm 1 uses a different approximation  $\phi_T^{(L)}$  of the first-order corrector  $\phi$ , compare (6) to [2, (18)]. The difference lies in the massive term  $\frac{1}{T}\phi_T^{(L)}$ . The reason of such change is discussed in Section 2.2.

---

**Algorithm 1** Optimal algorithm for the approximate solution  $u^{(L)}$  in  $Q_L$

---

- 1: For  $\beta \in (1, \frac{3}{2})$  set  $\varepsilon = \frac{3}{2} - \beta$  and  $T = L^{2(1-\varepsilon)}$ . For  $i = 1, \dots, d$ , solve for the approximate first-order corrector  $\phi_{i,T}^{(L)}$ :

$$\frac{1}{T}\phi_{i,T}^{(L)} - \nabla \cdot a \nabla \phi_{i,T}^{(L)} = \nabla \cdot a e_i \text{ in } Q_{2L}, \quad \phi_{i,T}^{(L)} = 0 \text{ on } \partial Q_{2L}. \quad (6)$$

- 2: Calculate the approximate homogenized coefficients via

$$a_h^{(L)} e_i = \int \omega q_{i,T}^{(L)}, \quad (7)$$

where

$$q_{i,T}^{(L)} := a(e_i + \nabla \phi_{i,T}^{(L)}) \quad (8)$$

and  $\omega(x) = \frac{1}{L^d} \hat{\omega}(\frac{x}{L})$  with  $\hat{\omega}$  as in Theorem 1.1.

---

<sup>2</sup>Here  $f$  denotes spatial average.

- 3: Find  $\tilde{u}_h^{(L)}$  on  $\partial Q_L$ :

$$\tilde{u}_h^{(L)} = G_h^{(L)} * (\nabla \cdot g) = \int dy G_h^{(L)}(x - y) \nabla \cdot g(y), \quad (9)$$

where  $G_h^{(L)}(x) := \frac{1}{4\pi |(a_h^{(L)})^{-1/2}x|}$  is the Green function for the constant-coefficient operator  $-\nabla \cdot a_h^{(L)} \nabla$ .

- 4: Solve for approximate first-order flux correctors  $\sigma_{i,T}^{(L)} = \{\sigma_{ijk,T}^{(L)}\}_{j,k}$ :

$$\frac{1}{T} \sigma_{ijk,T}^{(L)} - \Delta \sigma_{ijk,T}^{(L)} = \partial_j q_{ik,T}^{(L)} - \partial_k q_{ij,T}^{(L)} \text{ in } Q_{\frac{7}{4}L}, \quad \sigma_{ijk,T}^{(L)} = 0 \text{ on } \partial Q_{\frac{7}{4}L}. \quad (10)$$

- 5: Solve for approximate second-order correctors  $\psi_{ij,T}^{(L)}$ :

$$\frac{1}{T} \psi_{ij,T}^{(L)} - \nabla \cdot a \nabla \psi_{ij,T}^{(L)} = \nabla \cdot (\phi_{i,T}^{(L)} a - \sigma_{i,T}^{(L)}) e_j \text{ in } Q_{\frac{3}{2}L}, \quad \psi_{ij,T}^{(L)} = 0 \text{ on } \partial Q_{\frac{3}{2}L}. \quad (11)$$

- 6: For the indices

$$(i, j) \in \mathcal{J} = \{(1, 2), (1, 3), (2, 3), (2, 2), (3, 3)\}, \quad (12)$$

calculate

$$c_{ij,T}^{(L)} = - \int g \cdot \nabla \left( \sum_{k=1}^3 \phi_{k,T}^{(L)} \partial_k v_{h,ij}^{(L)} + (2 - \delta_{ij})(\psi_{ij,T}^{(L)} - \frac{a_{hij}^{(L)}}{a_{h11}^{(L)}} \psi_{11,T}^{(L)}) \right), \quad (13)$$

where  $v_{h,ij}^{(L)}$  denote the  $a_h^{(L)}$ -harmonic polynomials

$$v_{h,ij}^{(L)} = (1 - \frac{1}{2} \delta_{ij})(x_i x_j - \frac{a_{hij}^{(L)}}{a_{h11}^{(L)}} x_1^2). \quad (14)$$

- 7: Obtain  $u_h^{(L)}$  as

$$u_h^{(L)} = \tilde{u}_h^{(L)} + \sum_{i=1}^3 \left( \int g \cdot \nabla \phi_{i,T}^{(L)} \right) \partial_i G_h^{(L)} + \sum_{(i,j) \in \mathcal{J}} c_{ij,T}^{(L)} \partial_{ij} G_h^{(L)}. \quad (15)$$

- 8: Solve for  $u^{(L)}$  (here and for the rest of the paper we adopt Einstein's summation convention for repeated indices):

$$-\nabla \cdot a \nabla u^{(L)} = \nabla \cdot g \text{ in } Q_L, \quad u^{(L)} = (1 + \phi_{i,T}^{(L)} \partial_i + \psi_{ij,T}^{(L)} \partial_{ij}) u_h^{(L)} \text{ on } \partial Q_L. \quad (16)$$

## Related works

Quantitative stochastic homogenization, which dates back to Yurinskii [4], has been intensively studied during the past decade. Naddaf and Spencer [5, 6] introduced the notion of spectral-gap inequality and captured the CLT-type optimal scaling in stochastic homogenization under the condition of small ellipticity contrast, which is removed in [7–9] for discrete coefficients. The spectral-gap inequality is then refined to logarithmic Sobolev inequality in [10] and further generalized to ensembles with potentially thick correlation tails in [11, 12]. Another approach by Armstrong and Smart [13], Armstrong, Kuusi and Mourrat [14, 15]

uses a variational approach and obtain similar results. See [16] for a nice monograph. The Gaussianity of the energy of the solution was first identified by Nolen [17], and has been extended to homogenized coefficients by Gloria and Nolen [18], both for the representative volume element method. The covariance structure of correctors have been studied in [19], based on the annealed Green's function estimates in [10]. This is extended to study the fluctuation of homogenization error in [20]. Scaling limit of the correctors have been studied in [21]. Second-order correctors have been studied in [22, 23]. Correctors of higher order have been constructed in [24] up to order of  $\lceil \frac{d}{2} \rceil$  with suboptimal error estimate in two-scale expansion, which is improved to be optimal in [25]. Higher-order correctors up to order  $d$  were recently constructed in [26] using a distributional formulation. In a perturbative regime, it is possible to define homogenized coefficients up to order  $2d$  [27] by using a breakthrough result of Bourgain [28], in its refined form established in [29].

Developing efficient numerical algorithms has been a major motivation behind the study of quantitative stochastic homogenization, see e.g., the review [30]. Let us just mention a few directions here: Quantitative error estimates for computation of effective coefficients in stochastic homogenization have been studied in [31–34], where strategies using different boundary conditions or massive terms have been studied. Representative volume method, a popular approach used by engineers, are systematically analyzed in [35, 36]. Iterative multigrid methods have been studied in [37–39]. There have also been abundant research in numerical homogenization where the coefficients do not necessarily arise from random setting. An approach using an embedded corrector problem for approximating homogenized coefficients has been considered in [40]. Variance reduction methods have been developed in [41, 42]. Multiscale finite element methods have been developed in [43, 44] and extended in [45], see [46] for a review. Heterogeneous multiscale methods have been developed and studied in [47, 48] and many other works, see [49, 50] for reviews. In [51] they proposed a method that aims at recovering local oscillations by solving a local problem using the approximate solution of the homogenized equation as its boundary condition. Localized orthogonal decomposition approaches have been studied in [52, 53]. Relationship between numerical homogenization and Bayesian inference have been investigated in [54, 55].

## 2. Ideas behind the algorithm

### 2.1. Correctors, homogenized coefficients and fluxes, effective multipoles

Let us recall the standard first-order correctors  $\phi_i$ , which play a central role in homogenization theory [56, 57]. For each direction  $i = 1, \dots, d$ , the first-order corrector  $\phi_i$  is defined such that  $x_i + \phi_i$  is  $a$ -harmonic<sup>3</sup>:

$$-\nabla \cdot a(e_i + \nabla \phi_i) = 0. \quad (17)$$

Provided the ensemble is stationary and ergodic, the operator  $-\nabla \cdot a \nabla$  homogenizes on large scale to  $-\nabla \cdot a_h \nabla$ , with the constant and deterministic homogenized coefficient  $a_h$  given by

$$a_h e_i := \langle q_i \rangle = \lim_{L \uparrow \infty} \fint_{B_L} q_i \quad \text{where} \quad q_i := a(e_i + \nabla \phi_i). \quad (18)$$

---

<sup>3</sup>See below for a uniqueness argument.

Given  $a_h$ , we define  $\tilde{u}_h$  to be the solution of the homogenized equation

$$-\nabla \cdot a_h \nabla \tilde{u}_h = \nabla \cdot g. \quad (19)$$

The flux corrector  $\sigma$ , introduced in [12] in the setting of stochastic homogenization (see earlier ideas in periodic homogenization in [58, Proposition 7.2]), plays a convenient role in quantitative analysis. Since  $q_i - a_h e_i$ , which can be viewed as a  $(d-1)$ -form, has zero expectation and is divergence free in view of (17) and (18), there exists a  $(d-2)$ -form, which can be expressed as a skew-symmetric tensor field  $\sigma_i$ , such that

$$q_i - a_h e_i = \nabla \cdot \sigma_i, \quad (20)$$

where we use the convention  $(\nabla \cdot \sigma_i)_j = \partial_k \sigma_{ijk}$ . Clearly  $\sigma_i$  is only determined up to a  $(d-3)$ -form, so that a gauge has to be chosen to make its construction unique. It is convenient to choose

$$-\Delta \sigma_{ijk} = \partial_j q_{ik} - \partial_k q_{ij} = \nabla \cdot (q_{ik} e_j - q_{ij} e_k). \quad (21)$$

With the help of the  $\sigma_i$ 's one can express the homogenization error in divergence form: for any  $a_h$ -harmonic function  $u_h$ , the two-scale expansion  $(1 + \phi_i \partial_i) u_h$  is close to being  $a$ -harmonic in the sense of

$$-\nabla \cdot a \nabla (1 + \phi_i \partial_i) u_h = \nabla \cdot (\phi_i a - \sigma_i) \nabla \partial_i u_h. \quad (22)$$

The functions  $(\phi, \sigma)$  are uniquely determined up to a random constant by requiring  $(\nabla \phi, \nabla \sigma)$  to be stationary fields (i.e.  $\nabla \phi(a(\cdot + z), x) = \nabla \phi(a, x + z)$  for any  $x, z \in \mathbb{R}^d$ ) with finite second moments and zero mean [12, Lemma 1]. In dimension 3 under the unit-range assumption,  $(\phi, \sigma)$  themselves are stationary, and have finite stretched exponential moments [16, Theorem 4.1, Proposition 6.2], [3, Corollary 2]. Hence  $(\phi, \sigma)$  are uniquely determined by requiring  $\langle \phi \rangle = \langle \sigma \rangle = 0$  (see discussions in [12]).

In dimension 3, it is well-known that enriching the two-scale expansion by second-order correctors  $\psi_{ij}$  leads to a better approximation [23]. More precisely, given an  $a_h$ -harmonic function  $u_h$ , we may improve our two-scale expansion to  $(1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h$ , which is a second-order approximation of an  $a$ -harmonic function. The characterizing property of second-order correctors is the following equation<sup>4</sup>

$$-\nabla \cdot a \nabla \psi_{ij} = \nabla \cdot (\phi_i a - \sigma_i) e_j. \quad (23)$$

Equation (23) encodes the following property of the  $\psi_{ij}$ 's: for an  $a_h$ -harmonic quadratic polynomial  $u_h$ ,  $(1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h$  is exactly  $a$ -harmonic. In practice we only need  $\{\psi_{ij}\}_{i,j}$  in form of linear combinations  $E_{ij} \psi_{ij}$  with coefficients  $\{E_{ij}\}_{i,j}$  that are symmetric and satisfy the trace condition  $a_{hij} E_{ij} = 0$ , which is a five dimensional space for  $d = 3$ . Hence, it suffices to restrict to  $\{\psi_{ij}\}_{(i,j) \in \mathcal{J}}$ , where  $\mathcal{J}$  is the index set defined in Algorithm 1.

Analogously to first-order corrector case, we need second-order flux correctors, which were first introduced in [23] for quantitative analysis. Since by (23),

$$p_{ij} := a \nabla \psi_{ij} + (\phi_i a - \sigma_i) e_j \quad (24)$$

is divergence free, there exist  $\Psi_{ij} = \{\Psi_{ijkn}\}_{k,n=1,\dots,d}$ , skew-symmetric with respect to  $k$  and  $n$ , such that

$$p_{ij} = \nabla \cdot \Psi_{ij}. \quad (25)$$

---

<sup>4</sup>It might be possible to construct  $\psi$  via  $-\nabla \cdot a(\nabla \psi_{ij} + \phi_i e_j) = e_j \cdot (q_i - a_h e_i)$  in order to completely avoid computing  $\sigma_i$ . We will not discuss the details here.

Since  $a$  is symmetric, the second-order homogenized coefficient vanishes (see discussions in [23, calculations on page 5] or [1, Corollary 1]). Similar to the argument for  $\sigma_i$ ,  $\Psi_{ij}$  can be viewed as a  $(d-2)$ -form and is determined only up to a  $(d-3)$ -form, and the following choice of gauge is convenient:

$$-\Delta \Psi_{ijkn} = \partial_k p_{ijn} - \partial_n p_{ijk}. \quad (26)$$

We learn from (23) that  $\nabla \psi_{ij}$ , and thus  $p_{ij}$  by (24), and  $\nabla \Psi_{ijkn}$  by (26), can be constructed to be stationary if and only if  $\phi_i$  and  $\sigma_i$  are, which is the case only for  $d > 2$ . Although  $\Psi$  is not needed to formulate our algorithm, it will be used in our analysis to upgrade the homogenization error (22) from first to second-order: for any  $a_h$ -harmonic function  $u_h$ ,

$$-\nabla \cdot a \nabla (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h = -\nabla \cdot (\psi_{ij} a - \Psi_{ij}) \nabla \partial_{ij} u_h.$$

The second-order correctors  $\psi$  and  $\Psi$  are not expected to be stationary, but rather to grow at a rate a bit worse than  $2 - \frac{d}{2} = \frac{1}{2}$  away from the origin, which we capture through an exponent  $\beta < \frac{d}{2} = \frac{3}{2}$  that measures the amount the growth rate stays below quadratic. The first-order correctors  $\phi$  and  $\sigma$  are stationary for  $d > 2$ , but they are of course not bounded on  $\mathbb{R}^d$ ; we capture this through an exponent  $\alpha \in (\frac{1}{2}, 1)$  that measures the amount the growth rate below linear. Following [1, 23], we introduce the random radius  $r_{**} \geq 1$  starting from which we have the desired growth rate:

$$\frac{1}{r^2} \left( \int_{B_r} |(\psi, \Psi) - \int_{B_r} (\psi, \Psi)|^2 \right)^{\frac{1}{2}} \leq \left( \frac{r_{**}}{r} \right)^\beta \quad \text{for all } r \geq r_{**}, \quad (27)$$

and

$$\frac{1}{r} \left( \int_{B_r} |(\phi, \sigma)|^2 \right)^{\frac{1}{2}} \leq \left( \frac{r_{**}}{r} \right)^\alpha \quad \text{for all } r \geq r_{**}. \quad (28)$$

For convenience we take  $\alpha = 1 - \varepsilon$  and  $\beta = \frac{3}{2} - \varepsilon$  for the  $\varepsilon > 0$  fixed in Algorithm 1. This  $r_{**}$  is the one that appears in the statement of Theorem 1.2.

As observed in [1, 59], if we solve the boundary value problem

$$-\nabla \cdot a \nabla u_1 = \nabla \cdot g \text{ in } Q_L, \quad u_1 = (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) \tilde{u}_h \text{ on } \partial Q_L,$$

the solution has error  $O((\frac{\ell}{L})^d)$ , as it fails to capture the correct multipole behavior, which is the far-field behavior generated from the intrinsic moments of the localized r.h.s.  $\nabla \cdot g$ . We now recall the discussions in [1] and describe the far-field behavior of  $u$  in order to design the correct boundary condition on  $\partial Q_L$  to solve (4).

Let  $X_k$  be the space of  $a$ -harmonic functions of growth rate  $\leq k$ , more precisely<sup>5</sup>

$$X_k := \left\{ u : -\nabla \cdot a \nabla u = 0, \quad \limsup_{R \rightarrow \infty} R^{1-k} \left( \int_{B_R} |\nabla u|^2 \right)^{\frac{1}{2}} < \infty \right\},$$

and  $Y_k(r)$  be  $a$ -harmonic functions outside  $B_r$  with decay rate  $\geq k + d - 2$ , more precisely,

$$Y_k(r) := \left\{ u : -\nabla \cdot a \nabla u = 0 \text{ in } B_r^c, \quad \limsup_{R \rightarrow \infty} R^{k+(d-2)+1} \left( \frac{1}{R^d} \int_{B_R^c} |\nabla u|^2 \right)^{\frac{1}{2}} < \infty \right\}.$$

---

<sup>5</sup>Here  $B_r^c := \mathbb{R}^d \setminus B_r$ .



We also use  $X_k^h$  and  $Y_k^h(r)$  to denote similar spaces with  $a_h$  in place of  $a$  in the definition of  $X_k$  and  $Y_k(r)$ , respectively. The spaces  $X_k$  and  $Y_k(r)$  are related through the bilinear form: for  $u \in Y_k(r)$  and  $v \in X_k$ ,

$$(u, v)_a := - \int \nabla \eta \cdot (va \nabla u - ua \nabla v), \quad (29)$$

where the cutoff function  $\eta \equiv 1$  in  $B_r$  and  $\eta \equiv 0$  in  $B_{2r}^c$ . Note that the value of the integral does not depend on the choice of  $\eta$  since  $va \nabla u - ua \nabla v$  is divergence-free outside of  $B_r$ . We use  $(u_h, v_h)_{a_h}$  to denote the bilinear form similarly defined as (29) with  $a_h$  in place of  $a$ . In the case of a constant coefficient  $a_h$ , the spaces  $X_k^h$  and  $Y_k^h(r)$  are well-understood:  $X_k^h$  is the finite-dimensional linear space spanned by  $a_h$ -harmonic polynomials of degree at most  $k$ , while  $Y_k^h(r)$  is the infinite-dimensional linear space spanned by  $k$ -th or higher derivatives of  $G_h$ , the Green's function of  $-\nabla \cdot a_h \nabla$ . In the language of electrostatics, the quotient spaces  $Y_1^h(r)/Y_2^h(r)$  and  $Y_2^h(r)/Y_3^h(r)$  are spanned by first and second derivatives of  $G_h$  respectively, and thus represent dipoles and quadrupoles. As shown in [1, Proposition 5], similar to the pairing  $(\cdot, \cdot)_{a_h}$  providing an isomorphism between  $Y_1^h(r)/Y_3^h(r)$  and  $(X_2^h/X_0^h)^*$ ,  $(\cdot, \cdot)_a$  provides an isomorphism between  $Y_1(r)/Y_3(r)$  and  $(X_2/X_0)^*$ . For Liouville principles which describe the equivalence between spaces  $X_k^h$  and  $X_k$ , see also [14, 22].

By [1, Lemma 4], we know that  $u \in Y_1(\ell)$ . Therefore, by [1, Theorem 2], under the assumptions (27) and (28), for  $\ell \geq r_{**}$ , there exists a  $u_h \in Y_1^h(\ell)$ , uniquely defined up to an element in  $Y_3^h(\ell)$  by

$$(u, v)_a = (u_h, v_h)_{a_h}, \quad \text{for any } v_h \in X_2^h \text{ where } v := (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) v_h \in X_2, \quad (30)$$

that captures the effective multipole behavior of  $u$ . The condition  $\ell \geq r_{**}$ , which corresponds to  $\ell \geq r_{**}^{(L)}$  in Theorem 1.2, can be understood as a condition ensuring that the right-hand side  $g$  (whose characteristic scale is  $\ell$ ) is smooth, in the sense that it does not vary on too small scales, as is usually required for right-hand sides in homogenization. We state a modification of this result in the following Proposition 2.1.

**Proposition 2.1.** *Consider a coefficient field  $a$  on  $\mathbb{R}^d$  satisfying (1). Suppose that there exists a tensor  $a_h$  and, for  $i = 1, 2, 3$ , a scalar field  $\phi_i$  and a skew-symmetric tensor field  $\sigma_i$  such that (17), (18), and (20) hold, and for  $i, j = 1, 2, 3$ , a scalar field  $\psi_{ij}$  and a skew-symmetric tensor field  $\Psi_{ij}$  such that (23), (24), and (25) hold. For fixed  $\varepsilon > 0$  and  $\alpha = 1 - \varepsilon$ ,  $\beta = \frac{3}{2} - \varepsilon$ , suppose that there exists a radius  $r_{**}$  such that (27) and (28) hold. Moreover, let us assume the following convergences in distribution as  $R \rightarrow \infty$ :*

$$\begin{aligned} \phi_i(R \cdot) &\rightharpoonup 0 \\ \sigma_i(R \cdot) &\rightharpoonup 0 \\ (e_j \cdot \sigma_i \nabla \phi_k - e_j \cdot \sigma_k \nabla \phi_i)(R \cdot) &\rightharpoonup 0. \end{aligned} \quad (31)$$

Let  $g$  be of the form (2) for some  $\ell \geq r_{**}$  and  $\hat{g}$ , let  $u$  be the solution of (3), and let  $u_h \in Y_1^h(\ell)$  satisfy (30). Then we have for any  $R \geq r_{**}$ ,

$$\left( \frac{1}{R^d} \int_{B_R^c} \left| \nabla \left( u - (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h \right) \right|^2 \right)^{\frac{1}{2}} \leq C \left( \frac{\ell}{R} \right)^d \left( \frac{r_{**}}{R} \right)^\beta,$$

where  $C$  is a constant depending only on  $d$ ,  $\lambda$ ,  $\varepsilon$ , and  $\hat{g}$ .

Solving (4) with boundary condition  $u_L = (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij})u_h$  will indeed provide the desired level of approximation, which resembles Step 8 of [Algorithm 1](#) and is the statement of [Corollary 2.2](#). The reason we could not directly use  $(1 + \phi_i \partial_i + \psi_{ij} \partial_{ij})u_h$  for approximation is that [Proposition 2.1](#) does not hold inside  $B_{r_{**}}$ .

**Corollary 2.2.** *Under the same assumptions as in [Proposition 2.1](#), for  $L \geq \ell$  let  $\hat{u}$  be the solution of the equation*

$$-\nabla \cdot a \nabla \hat{u} = \nabla \cdot g \text{ in } Q_L, \quad \hat{u} = (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij})u_h \text{ on } \partial Q_L, \quad (32)$$

where  $u_h \in Y_1^h(\ell)$  satisfies (30). Then we have for any  $R \in [r_{**}, L]$ ,

$$\left( \int_{B_R} |\nabla(\hat{u} - u)|^2 \right)^{\frac{1}{2}} \leq C \left( \frac{\ell}{L} \right)^d \left( \frac{r_{**}}{L} \right)^\beta,$$

where  $C$  is a constant depending only on  $d, \lambda, \varepsilon$ , and  $\hat{g}$ .

For the purpose of [Algorithm 1](#) we now derive the explicit expression for  $u_h$  from (30). We start with  $\tilde{u}_h$  defined in (19), which is an element of  $Y_1^h(\ell)$ . We make the ansatz

$$u_h = \tilde{u}_h + \xi_i \partial_i G_h + \sum_{(i,j) \in \mathcal{J}} c_{ij} \partial_{ij} G_h \quad \text{mod } Y_3^h(\ell), \quad (33)$$

where we recall the index set  $\mathcal{J}$  is defined in (12). This ansatz is motivated by the fact that  $\{\partial_i G_h\}_{i=1,2,3}$  is a basis of the 3-dimensional space  $Y_1^h(\ell)/Y_2^h(\ell)$  and  $\{\partial_{ij} G_h\}_{(i,j) \in \mathcal{J}}$  a basis of the 5-dimensional space  $Y_2^h(\ell)/Y_3^h(\ell)$ , so that  $\xi_i \partial_i G_h + \sum_{(i,j) \in \mathcal{J}} c_{ij} \partial_{ij} G_h$  is a general element of  $Y_1^h(\ell)/Y_3^h(\ell)$ .

By equations (3) and (19) and integration by parts, we have for all  $v_h \in X_2^h$ ,

$$(u, v)_a = \int g \cdot \nabla (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) v_h, \quad \text{and} \quad (\tilde{u}_h, v_h)_{a_h} = \int g \cdot \nabla v_h. \quad (34)$$

If  $v_h \in X_0^h$  (i.e.  $v_h$  is a constant) then both expressions vanish. Moreover, comparing (34) with (30), we obtain the following identity for the multipole correction of  $\tilde{u}_h$ :

$$(\xi_i \partial_i G_h + \sum_{(i,j) \in \mathcal{J}} c_{ij} \partial_{ij} G_h, v_h)_{a_h} = \int g \cdot \nabla (\phi_i \partial_i + \psi_{ij} \partial_{ij}) v_h \quad \text{for } v_h \in X_2^h/X_0^h. \quad (35)$$

Note that for  $v_h \in X_2^h$ ,

$$(\partial_i G_h, v_h)_{a_h} = - \int \nabla \eta \cdot (v_h a_h \nabla \partial_i G_h - \partial_i G_h a_h \nabla v_h) = \partial_i v_h(0),$$

and similarly  $(\partial_{ij} G_h, v_h)_{a_h} = -\partial_{ij} v_h$ . Therefore, substituting this into (35), we obtain

$$\xi_i \partial_i v_h(0) - \sum_{(i,j) \in \mathcal{J}} c_{ij} \partial_{ij} v_h = \int g \cdot \nabla (\phi_i \partial_i + \psi_{ij} \partial_{ij}) v_h. \quad (36)$$

Thus, choosing  $v_h = x_i$  for  $i = 1, 2, 3$  in (36) we obtain

$$\xi_i = \int g \cdot \nabla \phi_i,$$

which is consistent with the expression obtained in [2]. To determine the coefficients  $c_{ij}$ , we test (36) with 5 linearly independent  $a_h$ -harmonic polynomials  $v_{h,ij}$  that are homogeneous of degree 2, and obtain linear equations of  $c_{ij}$ . Choosing the basis

$$v_{h,ij} = (1 - \frac{1}{2}\delta_{ij})(x_i x_j - \frac{a_{hij}}{a_{h11}} x_1^2) \text{ for } (i,j) \in \mathcal{J}, \quad (37)$$

leads to the explicit formula which highly resembles (13) in Algorithm 1

$$c_{ij} = - \int g \cdot \nabla \left( \phi_k \partial_k v_{h,ij} + (2 - \delta_{ij})(\psi_{ij} - \frac{a_{hij}}{a_{h11}} \psi_{11}) \right). \quad (38)$$

Summing up, the function  $u_h$  that captures the correct multipole behavior is given by

$$u_h = \tilde{u}_h + \left( \int g \cdot \nabla \phi_i \right) \partial_i G_h + c_{ij} \partial_{ij} G_h, \quad (39)$$

with  $c_{ij}$  given by (38). This motivates Algorithm 2 to obtain an approximation  $\hat{u}$  to the solution of (3).

---

**Algorithm 2** Idealized algorithm

---

- 1: Solve (17) for first-order correctors  $\phi_i$ .
  - 2: Determine the homogenized coefficients  $a_h$  via (18).
  - 3: Evaluate (19) for  $\tilde{u}_h$  on  $\partial Q_L$  by  $\tilde{u}_h = G_h * (\nabla \cdot g)$ .
  - 4: Solve (21) for first-order flux correctors  $\sigma_{ijk}$  and (23) for second-order correctors  $\psi_{ij}$ .
  - 5: Obtain  $u_h$  via (39).
  - 6: Solve (32) for  $\hat{u}$ , which is the approximation we desire.
- 

Algorithm 2 is, however, not computationally practical since several quantities like  $\phi_i, \sigma_i, \psi_{ij}$  still require solving a whole-space problem like (17) and thus knowledge of a realization of  $a$  outside of  $Q_{2L}$ . Fortunately, we can replace each of these quantities in Algorithm 2 by a computable surrogate with a small approximation error. This leads to Algorithm 1 and the error is only affected by a multiplicative constant. This is a consequence of the following Proposition 2.3, which allows us to pass from Corollary 2.2 to the error estimates of Algorithm 1 in Theorem 1.2:

**Proposition 2.3.** *Under the same assumptions as in Proposition 2.1, for  $\sqrt{T} = L^{1-\varepsilon} \geq \ell \geq r_{**}$ , define  $\phi_T^{(L)}, \sigma_T^{(L)}$  and  $\psi_T^{(L)}$  as in (6), (10), and (11). We assume they are good approximations of  $\phi, \sigma, \psi$ , in the sense of<sup>6</sup>*

$$\left( \int_{B_R} |(\sqrt{T} \nabla(\phi_T^{(L)} - \phi), \phi_T^{(L)} - \phi, \nabla(\psi_T^{(L)} - \psi))|^2 \right)^{\frac{1}{2}} \leq \sqrt{T} \left( \frac{r_{**}}{L} \right)^\beta \text{ for } R \in \left\{ \ell, \frac{5}{4}L \right\}, \quad (40)$$

and

$$\left( \int_{B_L} |\psi_T^{(L)} - \psi|^2 \right)^{\frac{1}{2}} \leq L^2 \left( \frac{r_{**}}{L} \right)^\beta. \quad (41)$$

Define  $a_h^{(L)}$  as in (7), and we assume it is a good approximation of  $a_h$  in the sense of

$$|a_h^{(L)} - a_h| \leq \left( \frac{r_{**}}{L} \right)^\beta. \quad (42)$$

---

<sup>6</sup>Here and for the rest of the paper the notation “ $(f_1, f_2) \lesssim X$ ” (when obviously  $X \geq 0$ ) means  $f_1 \lesssim X$  and  $f_2 \lesssim X$ .

Given  $\hat{g}$ , let  $u$  be defined as in (3) and  $u^{(L)}$  in (16). Then we have for any  $R \in [r_{**}, L]$ ,

$$\left( \int_{B_R} |\nabla(u^{(L)} - u)|^2 \right)^{\frac{1}{2}} \leq C \left( \frac{\ell}{L} \right)^d \left( \frac{r_{**}}{L} \right)^\beta,$$

with a constant  $C$  of the same type as in [Theorem 1.2](#).

## 2.2. Massive approximation of correctors

We start with introducing the notations for this section and beyond. Given a length scale  $R > 0$ , we define the exponential averaging function

$$\eta_R(x) := \frac{c_d}{R^d} \exp\left(-\frac{|x|}{R}\right)$$

with the constant  $c_d$  such that  $\int_{\mathbb{R}^d} \eta_R(x) dx = 1$ . We also define the Gaussian

$$G_R(x) := \frac{1}{(2\pi R^2)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{2R^2}\right).$$

For any function  $f$ , we use

$$f_R := f * G_R \tag{43}$$

to denote the convolution of  $f$  with  $G_R$ . For any  $s > 0$ , we define the following norm for a random variable  $F$  that quantifies its tail:

$$\|F\|_s := \inf \left\{ M \geq 0 : \left\langle \exp\left(\left(\frac{|F|}{M} + c\right)^s\right) \right\rangle - \exp(c^s) \leq 1 \right\} \quad \text{with } c = \begin{cases} \left(\frac{1-s}{s}\right)^{\frac{1}{s}} & s \in (0, 1), \\ 0 & \text{otherwise.} \end{cases} \tag{44}$$

Here the constant  $c$  is chosen such that the function  $[0, \infty) \ni x \mapsto \exp((x+c)^s) - \exp(c^s)$  is convex, which by Jensen's inequality makes  $\|\cdot\|_s$  a norm. Therefore, if  $\|F\|_s < \infty$ , then  $\langle \exp(r|F|^{\tilde{s}}) \rangle < \infty$  for  $r < \frac{1}{\|F\|_s^{\tilde{s}}}$  and  $\tilde{s} \in (0, s]$ .

We now illustrate why we may expect  $\phi_T^{(L)}$  and  $\psi_T^{(L)}$  to be good approximations of  $\phi$  and  $\psi$  in the sense of (40). Let  $\phi_T$  satisfy the equation

$$\frac{1}{T} \phi_T - \nabla \cdot a(e + \nabla \phi_T) = 0. \tag{45}$$

The massive corrector  $\phi_T$ , which was first considered in the early works of [4, 57] and then [7, 8, 60], is an approximation of  $\phi$  that has the advantage of being defined deterministically and it is automatically stationary. Indeed, in the class

$$\sup_x \int_{B_1(x)} (\phi_T^2 + |\nabla \phi_T|^2) < \infty$$

there exists a unique solution to (45). The massive corrector serves as the bridge between  $\phi$  and  $\phi_T^{(L)}$ , and we will show its closeness to both of them. The following [Proposition 2.4](#) shows that if we choose the length scale  $\sqrt{T}$  to be close to  $L$ , then the estimates (40) on  $\phi - \phi_T$  hold with high probability.

**Proposition 2.4.** For  $d \geq 3$  and  $\sqrt{T} \geq 1$ , we have<sup>7</sup>

$$\left\| \left( \int \eta_{\sqrt{T}} |\sqrt{T} \nabla(\phi_T - \phi)|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim \sqrt{T}^{-\frac{1}{2}}. \quad (46)$$

Moreover, there exists a random radius  $r_*$  with

$$\|r_*\|_d \lesssim 1, \quad (47)$$

and such that for all  $R > 0$ , we have<sup>8</sup>

$$\left\| I(R \geq r_*) \left( \int \eta_R |\sqrt{T} \nabla(\phi_T - \phi), \phi_T - \phi|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim \sqrt{T}^{-\frac{1}{2}}. \quad (48)$$

**Remark 2.5.** For  $d > 4$ , the approximation error (48) saturates at  $\sqrt{T}^{-1}$ . A better approximation which has error  $\sqrt{T}^{1-\frac{d}{2}}$  is given in [3, Theorem 3] and arises from iterated Richardson extrapolation of  $\phi_T$ . We expect a similar strategy to work in the optimal approximation of higher order correctors.

**Remark 2.6.** The quantity  $r_*$ , also known as the minimal radius, is the smallest scale on which the elliptic and parabolic  $C^{0,1}$ -estimates (See Lemma D.1) hold. Large scale regularity was first considered in [61] for periodic homogenization (see also the monograph [62]), and [12, 13] introduced the random variable  $r_*$  in the stochastic setting. For any fixed  $\delta \in (0, 1)$  we define  $r_*$  as in [3]<sup>9</sup>:

$$r_* = r_*(0) := \inf \left\{ r \geq 1 \mid \forall R \geq r, \frac{1}{R} \left( \int_{B_R} |(\phi, \sigma) - \int_{B_R} (\phi, \sigma)|^2 \right)^{\frac{1}{2}} \leq \delta \right\}. \quad (49)$$

Comparing the above (49) with (28), we observe  $r_{**} \geq r_*$ . The stochastic estimate (47) is proved in [3, Corollary 6]. For slightly different definitions of  $r_*$  under various probabilistic settings, we refer to [13, Theorem 1.1] and [12, Theorem 1] for corresponding stochastic estimates.

On the other hand,  $\phi_T$  near the origin can be well-approximated by a function that only depends on  $a$  through its restriction to the finite domain  $Q_L$ , which is exactly achieved by the function  $\phi_T^{(L)}$  defined in (6). Though the Dirichlet boundary conditions break the stationarity of  $\phi_T^{(L)}$ , deterministic methods are sufficient to prove that the difference  $\phi_T - \phi_T^{(L)}$  is sub-algebraically small in  $\frac{\sqrt{T}}{L}$  for  $\sqrt{T} \ll L$ . The exact statement of this is deferred to Proposition 2.9.

As a comparison, if we consider the Dirichlet approximation as [2, (18)]

$$-\nabla \cdot a \nabla \phi_i^{(L)} = \nabla \cdot a e_i \text{ in } Q_{2L}, \quad \phi_i^{(L)} = 0 \text{ on } \partial Q_{2L},$$

then it is unclear if one could prove anything stronger than what is proved in [2]

$$\left( \int_{Q_L} |\nabla(\phi^{(L)} - \phi)|^2 \right)^{\frac{1}{2}} \lesssim L^{-1},$$

<sup>7</sup>We use " $\|F\|_{s-} \lesssim X$ " hereafter to denote for any  $s' \in (0, s)$ ,  $\|F\|_{s'} \lesssim_{s'} X$ .

<sup>8</sup>Hereafter we use  $I$  to denote indicator (characteristic) functions.

<sup>9</sup>Note that the definitions in [12, 13] are slightly different.

which is the desired scaling for  $d = 2$  but insufficient for the CLT scaling required in (40) for  $d = 3$ . The bottleneck is that  $\phi^{(L)}$  is not stationary and it is thus unclear how probabilistic tools can be applied here. While one may use appropriate oversampling techniques to obtain desirable approximations using  $\phi^{(L)}$ , in this work we opt for our approximation  $\phi_T^{(L)}$  in (6) since we could prove the required approximation bounds and it is not more numerically difficult to compute than  $\phi^{(L)}$ .

Equipped with estimates on  $\phi$  we may derive estimates for the approximation of homogenized coefficients  $a_h$ . In analogy to (18) we introduce the modified flux

$$q_T := a(e + \nabla \phi_T). \quad (50)$$

For any smooth weight function  $\omega(x)$  supported in the unit ball satisfying  $\int_{\mathbb{R}^d} \omega = 1$ , and rescaled as  $\omega_L(x) = \frac{1}{L^d} \omega(\frac{x}{L})$ ,  $\int \omega_L q_T$  is a good approximation of  $a_h$ , which establishes that the assumption (42) of Proposition 2.3 holds with high probability:

**Lemma 2.7.** *Let  $L \geq \sqrt{T} \geq 1$ . Then for  $d \geq 3$  we have*

$$\|a_h e_i - \int \omega_L q_{i,T}\|_{2-} \lesssim \sqrt{T}^{-\frac{3}{2}}. \quad (51)$$

Using the same procedure, we can also approximate the first-order flux corrector  $\sigma$  and the second-order corrector  $\psi$  using their massive counterparts, denoted as  $\sigma_T$  and  $\psi_T$  respectively, which are defined by the equations

$$\frac{1}{T} \sigma_{ijk,T} - \Delta \sigma_{ijk,T} = \nabla \cdot (q_{ik,T} e_j - q_{ij,T} e_k), \quad (52)$$

$$\frac{1}{T} \psi_{ij,T} - \nabla \cdot a \nabla \psi_{ij,T} = \nabla \cdot (\phi_{i,T} a - \sigma_{i,T}) e_j. \quad (53)$$

Similar to  $\phi_T$ , both  $\sigma_T$  and  $\psi_T$  are well-defined stationary fields, and can be approximated by functions that only depend on  $a$  through  $a|_{Q_{2L}}$ , which are  $\sigma_T^{(L)}$  and  $\psi_T^{(L)}$  defined in (10) and (11), respectively. The following proposition gives estimates on  $\psi_T$ , establishing that the assumptions (40) and (41) hold with high probability.

**Proposition 2.8.** *Let  $r_*$  be the same random radius as defined in Proposition 2.4. Then for  $d \geq 3$  and  $\sqrt{T} \geq 1$ ,  $R > 0$ ,*

$$\left\| I(R \geq r_*) \left( \int \eta_R |\psi_T|^2 \right)^{\frac{1}{2}} \right\|_{1-} \lesssim \sqrt{T}^{\frac{1}{2}}, \quad (54)$$

$$\left\| I(R \geq r_*) \left( \int \eta_R |\nabla(\psi - \psi_T)|^2 \right)^{\frac{1}{2}} \right\|_{1-} \lesssim \sqrt{T}^{-\frac{1}{2}}. \quad (55)$$

The massive correctors  $\phi_T$  and  $\psi_T$  can both be approximated by functions that depend on  $a$  only through its restriction to a finite domain, with an error smaller than any power of  $\frac{\sqrt{T}}{L}$  for  $\sqrt{T} \ll L$ , which is the result of the following Proposition 2.9:

**Proposition 2.9.** *Let  $L \geq \sqrt{T} \geq 1$ . Let  $\phi_T^{(L)}, \sigma_T^{(L)}$  and  $\psi_T^{(L)}$  be defined through (6), (10) and (11), then for any  $p < \infty$ ,*

$$\left( \int_{Q_{\frac{7}{4}L}} |(\sqrt{T} \nabla(\phi_T - \phi_T^{(L)}), \phi_T - \phi_T^{(L)})|^2 \right)^{\frac{1}{2}} \lesssim_p \left( \frac{\sqrt{T}}{L} \right)^p, \quad (56)$$

$$\left( \int_{Q_{\frac{5}{4}L}} |(\sqrt{T} \nabla(\psi_T - \psi_T^{(L)}), \psi_T - \psi_T^{(L)})|^2 \right)^{\frac{1}{2}} \lesssim_p \sqrt{T} \left( \frac{\sqrt{T}}{L} \right)^p. \quad (57)$$

The reason we cannot prove (56) with  $Q_{2L}$  is that the Dirichlet boundary conditions generate a boundary layer where  $\phi_T$  may be dramatically different from  $\phi_T^{(L)}$ . Therefore near  $\partial Q_{2L}$ ,  $\phi_T^{(L)}$  is not trustworthy and should not be used for the computation of  $\sigma_T^{(L)}$ , which is the reason why in Algorithm 1 the domain of computation shrinks when computing  $\sigma_T^{(L)}$  and, for the same reason, shrinks further for  $\psi_T^{(L)}$ .

We finally present Proposition 2.10, which is the main ingredient of the proof for the stochastic estimate on  $r_{**}$  in (5), and, together with the estimate (47) on  $r_*$ , bounds the probability of the event  $r_{**} \geq R$  for any large  $R$ :

**Proposition 2.10.** *Let  $r_*$  be the same random radius as defined in Proposition 2.4, and denote  $(\nabla\psi, \nabla\psi)_R$  as the convolution of the two functions  $(\nabla\psi, \nabla\psi)$  with the Gaussian kernel  $G_R$  as in (43), then for  $d \geq 3$  and  $R \geq 1$ ,*

$$\|I(R \geq r_*) \left( \int_{B_R} |\nabla(\psi, \psi)|^2 \right)^{\frac{1}{2}}\|_{1-} \lesssim 1, \quad (58)$$

$$\|I(R \geq r_*) (\nabla\psi, \nabla\psi)_R\|_{1-} \lesssim R^{1-\frac{d}{2}}. \quad (59)$$

### 3. Numerical example

For the numerical test, we will consider a discrete elliptic equation<sup>10</sup> on  $\mathbb{Z}^3$ , so that we do not need to worry about error due to discretization. To set up the elliptic problem, we say the points  $x, y \in \mathbb{Z}^3$  are neighbors if  $\|x - y\|_{\ell_1} = 1$ , and draw an edge between  $x$  and  $y$  if they are neighbors. Denote  $\mathbb{B}$  as the set of (undirected) edges,  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  as the canonical basis in  $\mathbb{Z}^3$ , and  $(a_e)_{e \in \mathbb{B}}$  as the random field. The discrete gradient is defined as

$$\nabla f(x) = (f(x + \mathbf{e}_1) - f(x), f(x + \mathbf{e}_2) - f(x), f(x + \mathbf{e}_3) - f(x))$$

<sup>10</sup>We remark here why we consider the discrete equation in our numerical example. Our theoretical result holds without spatial discretization, and we do not assume any smoothness on  $a$ , so we may without loss of generality assume correlation length equal to 1. Unfortunately, computations in 3D are difficult, since to apply the finite difference (or any standard numerical) scheme, if we divide each side of the cube into  $N$  smaller intervals, then we have  $O(N^3)$  grid points, thus solving the full equation would require prohibitive memory. In practice, this can be resolved by using multigrid algorithms or using a supercomputer, but we do not plan to do that as neither is the purpose of this work. Such limitation in computational resources significantly restrict our choice of discretization and correlation lengths in our numerical tests. As a compromise, we solve discrete problems in our numerical example and hence both correlation and discretization lengths equal to unity.

and the divergence of  $F = (F_1, F_2, F_3)$  is defined as

$$\nabla \cdot F(x) = \sum_{i=1}^3 (F_i(x) - F_i(x - \mathbf{e}_i)).$$

The discrete equation we consider, with the above notations, is still in the form

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

The coefficient field  $a = a_e$  on edges  $e \in \mathbb{B}$  are i.i.d. random matrices with values Id and 9 Id with probability 1/2 each. For the right hand side, we take some function  $f(x)$  compactly supported in the box  $\{-1, 0, 1\}^3$  with average zero, so that there exists some vector valued function  $g(x)$  such that  $f = \nabla \cdot g$  for a function  $g$  supported in the slightly larger box  $Q_2$ .

We compare Algorithm 1 with three other algorithms:

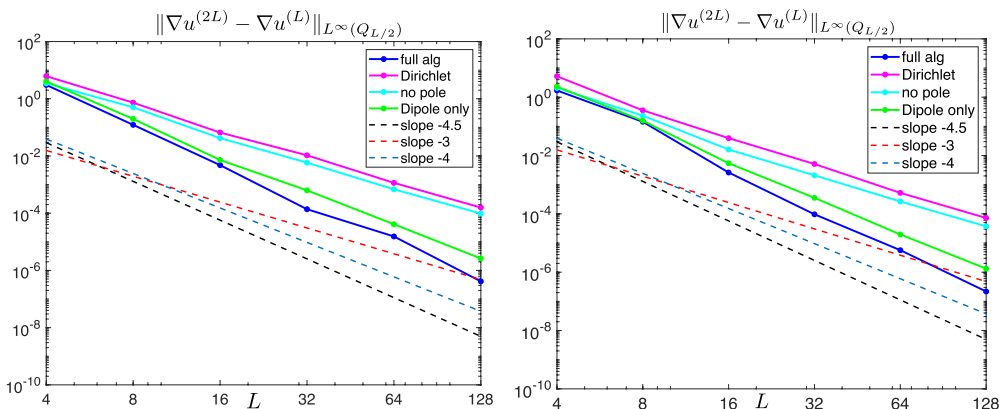
1. Solving the equation (4) with zero Dirichlet boundary condition.
2. Solving the equation (4) with modified correctors but without dipole or quadruple corrections, i.e., the boundary condition given by

$$u_{nc}^{(L)} = (1 + \phi_{i,T}^{(L)} \partial_i + \psi_{ij,T}^{(L)} \partial_{ij}) \tilde{u}_h^{(L)} \quad \text{on } \partial Q_L. \quad (60)$$

3. Solving the equation (4) with boundary condition corrected up to first-order correctors and dipoles, which is the algorithm proposed in [2]:

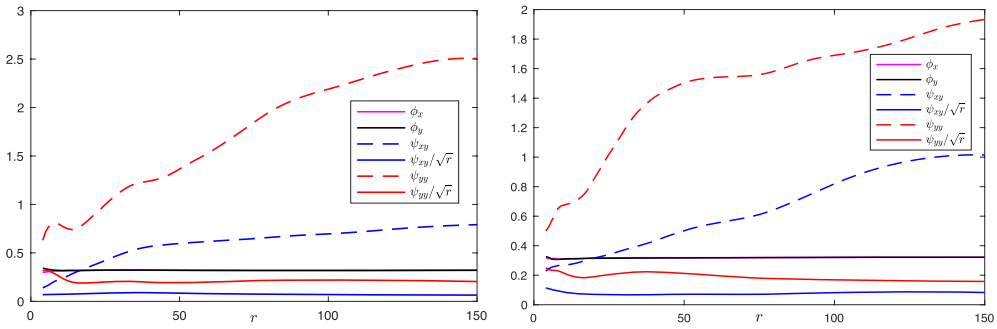
$$u_{dp}^{(L)} = (1 + \phi_{k,T}^{(L)} \partial_k) (\tilde{u}_h^{(L)} + (\int g \cdot \nabla \phi_{i,T}^{(L)}) \partial_i G_h^{(L)}) \quad \text{on } \partial Q_L. \quad (61)$$

We compare the numerical rate of  $\sup_{Q_{L/2}} |\nabla(u^{(2L)} - u^{(L)})|$  and plot it for various  $L$  for all four algorithms. We would comment here that while our analysis is for the gradient averaged over a region, for simplicity we only compare the gradient at a single point. From Figure 1 we can observe that the Dirichlet algorithm and the no multipole algorithm both have convergence rates of  $O(L^{-3})$ , the dipole algorithm has a convergence rate of  $O(L^{-4})$  while the proposed algorithm achieves  $O(L^{-4.5})$  convergence rate, which are consistent with our theoretical results.



**Figure 1.** Numerical convergence rate of  $\sup_{Q_{L/2}} |\nabla u^{(2L)} - \nabla u^{(L)}|$  for the proposed Algorithm 1, an algorithm with zero Dirichlet boundary condition, an algorithm without multipole corrections (defined in (60)), and an algorithm with only dipole corrections (defined in (61)). The two figures correspond to two independent realizations of the random media and the same r.h.s.





**Figure 2.**  $L^2$ -averages of  $\phi$  and fluctuations of  $\psi$ . The two figures correspond to independent realizations of the random media.

We also numerically compute  $(f_{Q_r} |\phi_T^{(L)}|^2)^{\frac{1}{2}}$  and  $(f_{Q_r} |\psi_T^{(L)} - f_{Q_r} \psi_T^{(L)}|^2)^{\frac{1}{2}}$  for a variety of  $r$ . Figure 2 indicates that the quantities  $(f_{Q_r} |\phi_T^{(L)}|^2)^{\frac{1}{2}}$  and  $\frac{1}{\sqrt{r}}(f_{Q_r} |\psi_T^{(L)} - f_{Q_r} \psi_T^{(L)}|^2)^{\frac{1}{2}}$  are almost constants for all  $r$ , which is consistent with their growth estimates (28) and (27). The figure also indicates numerically that  $r_{**}$  should be of order 1.

## 4. Strategy of the proof

### 4.1. Parabolic semigroup representation of correctors and fluxes

Our approach is based on the semigroup framework developed in [3], see also [16, Chapter 9]. The central objects we study are the operators  $S(t)$  and  $\mathcal{S}(t)$  related to  $-\nabla \cdot a \nabla$ .

**Definition 4.1.** The linear operators  $S(t)$  and  $\mathcal{S}(t)$  are defined as follows: for an arbitrary vector field  $g$ , let  $v$  solve the parabolic equation

$$\partial_t v - \nabla \cdot a \nabla v = 0 \text{ for } t > 0, \quad v(t=0) = \nabla \cdot g,$$

then

$$S(t)g := v(t), \text{ and } \mathcal{S}(t)g := g + a \int_0^t d\tau \nabla v(\tau). \quad (62)$$

We also define the operator  $\bar{S}(t)$  similarly to  $S(t)$  with  $a$  replaced by constant coefficient  $\text{Id}$ , and  $\mathcal{S}^h(t)$  similarly to  $\mathcal{S}(t)$  with  $a$  replaced by  $a_h$ .

Roughly speaking,  $\mathcal{S}(t)g$  is the flux accumulated from time 0 to  $t$  with initial condition  $\nabla \cdot g$ . As opposed to  $S(t)$ ,  $\mathcal{S}(t)$ , which is the same as the operator  $S_{0 \rightarrow t}$  defined in [3], does satisfy the semigroup property [3, Lemma 2]

$$\mathcal{S}(t_1)\mathcal{S}(t_2)g = \mathcal{S}(t_1 + t_2)g. \quad (63)$$

Since the solution of the standard heat equation is the convolution of the initial condition with the heat kernel, which coincides with  $G_{\sqrt{t}}$ ,  $\bar{S}$  can be rewritten as

$$\bar{S}(t)g = (\nabla \cdot g)_{\sqrt{t}} \stackrel{(43)}{=} G_{\sqrt{t}} * (\nabla \cdot g). \quad (64)$$

The semigroup operators  $S(t)$  and  $\bar{S}(t)$  are essential since all correctors and their massive approximations can be represented using these operators, which is the building block for the

estimates of correctors and fluxes in [Propositions 2.4–2.9](#). More precisely, these operators provide a resolution of the correctors in terms of quantities of controlled locality and amplitude. Below in [Remark 4.2](#) we collect the semigroup representation formulas for all correctors and fluxes. Here for simplicity of notation, we omit the indices, and introduce the vector product notation  $\times$  so that  $\sigma = \int_0^\infty dt_0 \bar{S}(t_0) \times q$  stands for  $\sigma_{ijk} = \int_0^\infty dt_0 \bar{S}(t_0)(q_{ik}e_j - q_{ij}e_k)$ .

**Remark 4.2.** The correctors  $\phi, \sigma, \psi, \Psi$  and their massive approximations  $\phi_T, \psi_T$  can be represented as follows

$$\phi = \int_0^\infty dt S(t)ae, \quad (65)$$

$$\phi_T = \int_0^\infty dt \exp(-\frac{t}{T})S(t)ae, \quad (66)$$

$$\sigma = \int_0^\infty dt_0 \bar{S}(t_0) \times q, \quad (67)$$

$$\begin{aligned} \nabla \psi = & \int_0^\infty dt_0 \int_0^\infty dt_1 \nabla S(t_0)(aS(t_1)ae - \bar{S}(t_1) \times ae) \\ & - \int_0^\infty dt_0 \int_0^\infty dt_1 \int_0^\infty dt_2 \nabla S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae, \end{aligned} \quad (68)$$

$$\begin{aligned} \psi_T = & \int_0^\infty dt_0 \exp(-\frac{t_0}{T}) \int_0^\infty dt_1 \exp(-\frac{t_1}{T})S(t_0)(aS(t_1)ae - \bar{S}(t_1) \times ae) \\ & - \int_0^\infty dt_0 \exp(-\frac{t_0}{T}) \int_0^\infty dt_1 \exp(-\frac{t_1}{T}) \int_0^\infty dt_2 \exp(-\frac{t_2}{T})S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae, \end{aligned} \quad (69)$$

$$\nabla \Psi = \int_0^\infty dt \nabla \bar{S}(t) \times (a\nabla \psi + (a\phi - \sigma)e). \quad (70)$$

We would like to comment here that (68) and (70) are formulated on the level of gradients, since only  $\nabla(\psi, \Psi)$  are well-defined stationary random fields. To show the algebraic equivalence of the quantities, it suffices to show the r.h.s. of the equations satisfy the same elliptic equations as their counterparts on the l.h.s. The core argument we need is, suppose a function  $w_T$  has the representation

$$w_T = \int_0^\infty dt \exp(-\frac{t}{T})S(t)g, \quad (71)$$

then it satisfies the massive equation  $\frac{1}{T}w_T - \nabla \cdot a\nabla w_T = \nabla \cdot g$ . Indeed,

$$\begin{aligned} -\nabla \cdot a\nabla w_T &= -\int_0^\infty dt \exp(-\frac{t}{T})\nabla \cdot a\nabla S(t)g = -\int_0^\infty dt \exp(-\frac{t}{T})\partial_t S(t)g \\ &= S(0)g + \int_0^\infty dt (\partial_t \exp(-\frac{t}{T}))S(t)g = \nabla \cdot g - \frac{1}{T}w_T. \end{aligned}$$

Similarly, if  $w$  has the representation

$$w = \int_0^\infty S(t)g dt, \quad (72)$$

then  $w$  satisfies  $-\nabla \cdot a \nabla w = \nabla \cdot g$ . In particular, we obtain for  $a = \text{Id}$ , the function

$$w = \int_0^\infty dt (\nabla \cdot g)_{\sqrt{t}}. \quad (73)$$

satisfies  $-\Delta w = \nabla \cdot g$ . Choosing  $g = ae$  in (71) and (72) and we obtain the representations (66) for  $\phi_T$  and (65) for  $\phi$ , respectively, and choosing  $g = q$  in (73) yields (67) for  $\sigma$ . The representations for  $(\nabla \psi, \psi_T, \nabla \Psi)$  can be derived by applying the above arguments iteratively and we will not write the details here. To rigorously obtain the equivalences, in view of uniqueness of stationary correctors, it suffices to show that the r.h.s. integrals converge and therefore represent stationary functions, which is done in the proofs of (86) for  $(\phi, \sigma)$  and (58), (54) for  $(\nabla \psi, \nabla \Psi, \psi_T)$  (such result for  $\phi_T$  is classic).

#### 4.2. CLT-cancellations and propagation of locality

The next important notion is the so-called “relative approximate locality”.

**Definition 4.3.** Let  $p > \frac{d}{2}$  be fixed, and let  $g$  and  $\bar{g}$  denote two stationary random fields. We say  $g$  is approximately local on scale  $r \geq 1$  relative to  $\bar{g}$  if, for two realizations  $a$  and  $\tilde{a}$  satisfying  $a = \tilde{a}$  in  $B_{2R}$  for some  $R \geq r$ ,

$$\left( \int_{B_R} |g(a) - g(\tilde{a})|^2 \right)^{\frac{1}{2}} \leq \left( \frac{r}{R} \right)^p \int \eta_R (\bar{g}(a) + \bar{g}(\tilde{a})). \quad (74)$$

Roughly speaking, the random field  $g|_{B_R}$  “essentially depends on  $a$  only through  $a|_{B_r}$  up to an error term  $(\frac{r}{R})^p \bar{g}$ ”. This locality is at the basis of a CLT scaling, see Lemma C.1.

In order to estimate higher-order correctors and their fluxes, it is important to estimate the operator  $S(T)$  acting on arbitrary  $g$ . More precisely, we extend the results in [3] on approximate locality from the special and obviously local initial data  $ae$  to more general initial data  $g$  that are only approximately local. Our goal is to derive stochastic bounds and approximate locality properties for  $(S(T)g, \nabla S(T)g, \mathcal{S}(T)g)$ , given  $g$  approximately local on scale  $r_0 \geq 1$  relative to  $\bar{g}$ . In addition, while  $ae$  is obviously bounded, we only assume stochastic integrability of  $g$  and  $\bar{g}$ , in the sense of controlling  $\|g\|_s$  and  $\|\bar{g}\|_s$  for some  $0 < s \leq 2$ .

In the range  $\sqrt{T} \leq r_0$ ,  $S(T)g$  does not benefit from stochastic cancellations through parabolic propagation of  $g$ . Therefore we can do no better than the following consequence of a deterministic estimate (see Lemma B.2)

$$\left\| \left( \int \eta_{r_0} |(T \nabla S(T)g, \sqrt{T} S(T)g, \mathcal{S}(T)g)|^2 \right)^{\frac{1}{2}} \right\|_s \lesssim \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s. \quad (75)$$

The scenario is more subtle in the regime  $\sqrt{T} \geq r_0$ . While  $S(T)g$  now benefits from stochastic cancellations, for which the ratio of the locality scale  $r_0$  of  $g$  and the parabolic scale  $\sqrt{T}$  matters, it also suffers an increase of the locality scale to  $\sqrt{T}$ , as well as a loss of stochastic integrability to  $\frac{2s}{s+2}$ . The first and third aspects are captured by Proposition 4.4, the second by Proposition 4.6.

**Proposition 4.4.** Suppose  $g$  and  $\bar{g}$  satisfy (74) for  $r_0 \geq 1$ . Then for all  $R \geq 1$ ,  $T \geq 0$ , we have<sup>11</sup>

$$\left\| \left( S(T)g - \langle S(T)g \rangle, \int_0^T dt \nabla S(t)g \right)_R \right\|_{\frac{2s}{s+2}-} \lesssim \left( \frac{r_0}{R} \right)^{\frac{d}{2}} \left( \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s + \|\bar{g}\|_s \right) \quad (76)$$

and

$$\left\| \left( \int \eta_{\sqrt{T}} |T \nabla S(T)g, \sqrt{T} S(T)g|^2 \right)^{\frac{1}{2}} \right\|_{\frac{2s}{s+2}-} \lesssim (1 \wedge \frac{r_0}{\sqrt{T}})^{\frac{d}{2}} \left( \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s + \|\bar{g}\|_s \right). \quad (77)$$

**Remark 4.5.** The same estimates hold with  $S$  replaced by  $\bar{S}$ , with stochastic integrability improved from  $\frac{2s}{s+2}-$  to  $s$ , see [3, Lemma 16].

**Proposition 4.6.** Suppose  $g$  and  $\bar{g}$  satisfy (74) for  $r_0 \geq 1$ . For all  $T \geq 0$ ,  $(T \nabla S(T)g, \sqrt{T} S(T)g, S(T)g)$  is approximately local on scale  $r_0 \vee \sqrt{T}$  relative to some stationary  $\bar{G}$  with

$$\|\bar{G}\|_{s_0} \lesssim (1 \wedge \frac{r_0}{\sqrt{T}})^{\frac{d}{2}} \left( \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s + \|\bar{g}\|_s \right). \quad (78)$$

Here  $s_0 = \frac{2s}{s+2}-$  when  $\sqrt{T} > r_0$  and  $s_0 = s$  when  $\sqrt{T} \leq r_0$ .

Propositions 4.4 and 4.6 tell us that if initially  $\left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s$  and  $\|\bar{g}\|_s$  have the same upper bound, then so will  $\left\| \left( \int \eta_{(r_0 \vee \sqrt{T})} |T \nabla S(T)g, \sqrt{T} S(T)g, S(T)g|^2 \right)^{\frac{1}{2}} \right\|_{\frac{2s}{s+2}-}$  and  $\|\bar{G}\|_{\frac{2s}{s+2}-}$ , so that we can essentially “sweep under the rug” the quantity quantifying the locality since it has the same size as the original quantity. This is convenient for estimating multiple time integrals.

In particular, for  $g = ae$ , in which case we may set  $r_0 = 1$ ,  $\bar{g} = 0$  and  $s = \infty$ , we recover the following stochastic estimates on  $S(t)ae$  proved in [3], which are special cases of Propositions 4.4 and 4.6.

**Lemma 4.7.** [3, Corollary 4] For all  $T \geq 0$ ,

$$\left\| \left( \int \eta_{\sqrt{T}} |T \nabla S(T)ae, \sqrt{T} S(T)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim (1 \wedge \frac{1}{\sqrt{T}})^{\frac{d}{2}}. \quad (79)$$

Moreover,  $(T \nabla S(T)ae, T a \nabla S(T)ae, \sqrt{T} S(T)ae, S(T)ae)$  is approximately local on scale  $1 \vee \sqrt{T}$  relative to some stationary  $\bar{g}$  with

$$\|\bar{g}\|_{2-} \lesssim (1 \wedge \frac{1}{\sqrt{T}})^{\frac{d}{2}}. \quad (80)$$

With Lemma 4.7 and the representations (65), (66) of  $\phi$  and  $\phi_T$ , we see why  $\phi_T$  is a good approximation of  $\phi$ : in the range  $t \gg T$ , the contribution of  $S(t)ae$  to  $\phi_T$  is exponentially small, and the locality of  $S(t)ae$  in conjunction with the finite range condition on  $a$  make the contribution to  $\phi$  small as well, which (almost) matches the bound (40); in the range  $t \lesssim T$ , we can use  $1 - \exp(-\frac{t}{T}) \lesssim \frac{t}{T}$  and the stochastic decay estimates of  $S(t)ae$  to control  $\phi_T - \phi$ , which matches (40) in spatial dimension 3 as well. This is the main intuition behind the proof of Proposition 2.4. We use the same strategy to prove that  $\psi_T$  is a good approximation of  $\psi$ ,

<sup>11</sup>Here we use  $r_1 \vee r_2$  and  $r_1 \wedge r_2$  to denote  $\max\{r_1, r_2\}$  and  $\min\{r_1, r_2\}$ , respectively.

where we again divide the representation of  $\psi_T - \psi$  into several regimes depending on the relationship between the  $t_i$ 's.

Another motivation of introducing the semigroup framework comes from deriving stochastic estimates for  $\psi$ ,  $\Psi$ , and  $r_{**}$  in our finite range setting. For ensembles that satisfy a logarithmic Sobolev inequality, these estimates are already established in [23], while similar estimates are absent in our setting of finite range dependence (as a side note, contrary to intuition, ensembles that satisfy finite range dependence may not satisfy an LSI, see [63, example after Theorem 6]). The semigroup framework will be our main tool for establishing stochastic bounds on  $\psi$ ,  $\Psi$  (see Proposition 2.8 for its precise statement), which eventually leads to the stochastic estimates on  $r_{**}$ , c.f. (5).

We would also comment here that the representation of  $(\psi, \Psi)$  involves two instances of the random  $S(t)$ , losing stochastic integrability twice, which is the reason why we only obtain the stochastic integrability of  $\frac{2s}{s+2}|_{s=2-} = 1-$ . For the same reason, we would expect the  $k$ -th order corrector (when dimension  $d \geq 2k - 1$ ) to have stochastic integrability  $\frac{2}{k}-$ .

### 4.3. Other technical lemmas

We continue with presenting auxiliary lemmas for the proof of Proposition 4.4. In order to capitalize in stochastic cancellations as in the proof of [3, Theorem 1], for which we divide into dyadic scales and apply the CLT-estimate Lemma C.1 to  $(S(T) - S^h(\frac{T}{2})S(\frac{T}{2}))g$ , we need the following approximate locality result.

**Lemma 4.8.** *For  $\sqrt{T} \geq r_0$ ,  $(S(T) - S^h(\frac{T}{2})S(\frac{T}{2}))g$  is approximately local on scale  $\sqrt{T}$  relative to  $\bar{F} + (\frac{r_0}{\sqrt{T}})^{\frac{d}{2}}\bar{g}$ , where*

$$\bar{F} := \int_0^T dt \left(\frac{\sqrt{t}}{\sqrt{T}}\right)^{\frac{d}{2}} \int_0^{\sqrt{t}} dr \left(\frac{r}{\sqrt{t}}\right)^{\frac{d}{2}} |(S(t)g - \langle S(t)g \rangle_r)|. \quad (81)$$

We also need some auxiliary estimates of  $(\phi, \sigma, q)$  to prove Proposition 4.4, which are listed in the following Lemma 4.9. The first two results are proven in [3] while the latter three are not explicitly stated in [3] since they involve stationary  $(\phi, \sigma, q)$ , so we provide a proof for them.

**Lemma 4.9.** *Let  $d \geq 3$ ,  $r_*$  be the random radius defined in Proposition 2.4, then for any  $R \geq 1$ ,*

$$[3, \text{Corollary1}] \quad \|(\nabla\phi, \nabla\sigma, q - \langle q \rangle, S(t)ae - \langle S(t)ae \rangle)_R\|_{2-} \lesssim R^{-\frac{d}{2}}, \quad (82)$$

$$[3, \text{Corollary4}] \quad \|(\int \eta_R |\nabla\phi|^2)^{\frac{1}{2}}\|_{2-} \lesssim 1, \quad (83)$$

$$\|(\phi, \sigma)_R\|_{2-} \lesssim R^{1-\frac{d}{2}}, \quad (84)$$

$$\|I(R \geq r_*)(a\phi - \langle a\phi \rangle)_R\|_{2-} \lesssim R^{1-\frac{d}{2}}. \quad (85)$$

$$\|(\int \eta_R |(\phi, \sigma)|^2)^{\frac{1}{2}}\|_{2-} \lesssim 1, \quad (86)$$

We conclude with two technical lemmas which facilitate the proof of [Proposition 2.8](#). The first lemma acts like a combination of [Lemma B.2](#) (with  $g$  replaced by  $aS(t_1)g$ ) and [Corollary D.2](#).

**Lemma 4.10.** *For any  $R > 0$ ,  $t_1 \geq t_0 > 0$ ,*

$$\begin{aligned} & \left\| I(R \geq r_*) \left( \int_{\eta_R} \left| \int_0^{t_0} d\tau \nabla S(\tau) aS(t_1)g \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \lesssim \left\| I(R \vee \sqrt{t_1} \geq r_*) \left( \int_{\eta_{R \vee \sqrt{t_1}}} (|\sqrt{t_1} \nabla S(\frac{t_1}{2})g|^2 + |S(\frac{t_1}{2})g|^2) \right)^{\frac{1}{2}} \right\|_{2-}. \end{aligned} \quad (87)$$

The second lemma deals with estimating the triple integral term of  $\psi_T$ , see (69), in the regime  $t_2 < t_1$ . The goal is to utilize the bounds on  $S(t)ae$ , i.e. [Lemma 4.9](#).

**Lemma 4.11.** *For  $T \geq 1$ ,  $t_3 \geq t_1 > 0$ ,*

$$\begin{aligned} & \left\| \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3 + t_2}{T})) (\sqrt{t_1} \nabla \bar{S}(t_1) \times a \nabla S(t_2) ae, \bar{S}(t_1) \times a \nabla S(t_2) ae) \right\|_{2-} \\ & \lesssim (1 \wedge \frac{t_3}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}, \end{aligned} \quad (88)$$

$$\left\| \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) (\sqrt{t_1} \nabla \bar{S}(t_1) \times a \nabla S(t_2) ae, \bar{S}(t_1) \times a \nabla S(t_2) ae) \right\|_{2-} \lesssim \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}. \quad (89)$$

Moreover,  $\int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3 + t_2}{T})) \bar{S}(t_1) \times a \nabla S(t_2) ae$  is approximately local on scale  $1 \vee \sqrt{t_1}$  relative to some stationary  $\bar{g}$  with

$$\|\bar{g}\|_{2-} \lesssim (1 \wedge \frac{t_3}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}. \quad (90)$$

## 5. Proofs

### 5.1. Proof of Theorem 1.2

*Step 1:* Stochastic estimates of  $r_{**}$ . The idea of the proof is based on [64, Theorem 1 (ii)]. Our goal is to estimate the “failure probability”  $\langle I(r_{**} \geq R_0) \rangle$  for an arbitrary  $R_0$ . We separate the event  $r_{**} \geq R_0$  into three possible scenarios. The first scenario to leave out is  $r_* \geq \sqrt{R_0}$  which has probability at most  $\exp(-\frac{1}{C} R_0^{\frac{3}{2}})$ , as  $r_*$  has stochastic integrability  $d$  (47). For the rest of the proof we assume  $r_* \leq \sqrt{R_0} \leq R_0$ .

Next we look at the failure probability due to first-order correctors  $(\phi, \sigma)$ . We recall (28) which defines the constraint on  $r_{**}$  coming from  $(\phi, \sigma)$ . By (86) in [Lemma 4.9](#), which we reformulate in terms of boxes instead of balls, we have for all  $R \geq 1$ ,

$$\left\| \left( \int_{Q_R} |(\phi, \sigma)|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim 1,$$

so that by Chebyshev's inequality, we have<sup>12</sup>

$$\begin{aligned} \langle I\left(\left(\oint_{Q_R} |(\phi, \sigma)|^2\right)^{\frac{1}{2}} \geq R^{1-\alpha} R_0^\alpha\right) \rangle &= \langle I\left(\exp\left(\left(\oint_{Q_R} |(\phi, \sigma)|^2\right)^{\frac{2-}{2}}\right) \geq \exp\left((R^{1-\alpha} R_0^\alpha)^{2-}\right)\right) \rangle \\ &\lesssim \exp\left(-\frac{1}{C}(R^{1-\alpha} R_0^\alpha)^{2-}\right). \end{aligned}$$

Here the second line uses the third expression of the norm  $\|\cdot\|_s$  in (A1). Therefore if  $r_{**} \geq R_0$  because of the “failure” of first-order correctors, which means that there must be some  $R \geq R_0$  (which we may assume to be dyadic) such that

$$\left(\oint_{Q_R} |(\phi, \sigma)|^2\right)^{\frac{1}{2}} \geq R^{1-\alpha} R_0^\alpha,$$

the probability is dominated by<sup>13</sup>

$$\sum_{\substack{R \geq R_0 \\ \text{dyadic}}} \exp\left(-\frac{1}{C}(R^{1-\alpha} R_0^\alpha)^{2-}\right) \sim \exp\left(-\frac{1}{C} R_0^{2-}\right),$$

where we used  $\alpha < 1$ .

We now look at the failure probability due to second-order correctors. We first argue that it is enough to prove for any  $R \geq 1$ ,

$$\left\| I(\sqrt{R} \geq r_*) \left( \oint_{B_R} |(\psi, \Psi) - \oint_{B_R} (\psi, \Psi)|^2 \right)^{\frac{1}{2}} \right\|_{1-} \lesssim R^{\frac{1}{2}}. \quad (91)$$

Indeed, if (91) holds true, then we can bound the failure probability using Chebyshev inequality:

$$\begin{aligned} &\sum_{\substack{R \geq R_0 \\ \text{dyadic}}} \langle I\left(\left(\oint_{B_R} |(\psi, \Psi) - \oint_{B_R} (\psi, \Psi)|^2\right)^{\frac{1}{2}} \geq R_0^\beta R^{2-\beta}\right) I(\sqrt{R_0} \geq r_*) \rangle \\ &\lesssim \sum_{\substack{R \geq R_0 \\ \text{dyadic}}} \exp\left(-\left(\frac{R_0^\beta R^{2-\beta}}{cR^{\frac{1}{2}}}\right)^{1-}\right) \sim \exp\left(-\frac{1}{C} R_0^{\frac{3}{2}-}\right), \end{aligned}$$

where we used  $\beta < \frac{3}{2}$ . To prove (91), we again divide into dyadic series and use (58) and (59) in Proposition 2.8. We will abuse notation and use  $f_r$  to denote  $f_{Q_r} f$ , as [3, Lemma 13, Step 4] shows its equivalence to  $f * G_r$  under stochastic norm  $\|\cdot\|$ .

$$\begin{aligned} &\left\| I(\sqrt{R} \geq r_*) \left( \oint_{B_R} |(\psi, \Psi) - \oint_{B_R} (\psi, \Psi)|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\ &\leq \left\| I(\sqrt{R} \geq r_*) \left( \left( \oint_{B_R} |(\psi, \Psi) - (\psi, \Psi)_{r_*}|^2 \right)^{\frac{1}{2}} + \sum_{\substack{R/2 \geq r \geq r_* \\ \text{dyadic}}} \left( \oint_{B_R} |(\psi, \Psi)_r - (\psi, \Psi)_{2r}|^2 \right)^{\frac{1}{2}} \right) \right\|_{1-} \\ &\leq \left\| I(\sqrt{R} \geq r_*) \left( \oint_{B_R} |(\psi, \Psi) - (\psi, \Psi)_{r_*}|^2 \right)^{\frac{1}{2}} \right\|_{1-} \end{aligned}$$

<sup>12</sup>Here and for the rest of the paper  $X \lesssim \exp(-R^{s-})$  denotes for any  $s_0 \in (0, s)$ ,  $X \lesssim_{s_0} \exp(-R^{s_0})$ .

<sup>13</sup>Here  $\sim$  means both  $\lesssim$  and  $\gtrsim$  hold.

$$\begin{aligned}
& + \sum_{\substack{R/2 \geq r \geq 1 \\ \text{dyadic}}} \left\| I(r \geq r_*) \left( \oint_{B_R} |(\psi, \Psi)_r - (\psi, \Psi)_{2r}|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \lesssim \|I(\sqrt{R} \geq r_*) r_* \left( \int \eta_R |\nabla(\psi, \Psi)|^2 \right)^{\frac{1}{2}}\|_{1-} + \sum_{\substack{R/2 \geq r \geq 1 \\ \text{dyadic}}} \left\| I(r \geq r_*) r \left( \int \eta_R |\nabla(\psi, \Psi)_r|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \leq \sqrt{R} \|I(\sqrt{R} \geq r_*) \left( \int \eta_R |\nabla(\psi, \Psi)|^2 \right)^{\frac{1}{2}}\|_{1-} + \sum_{\substack{R/2 \geq r \geq 1 \\ \text{dyadic}}} r \|I(r \geq r_*) \nabla(\psi, \Psi)_r\|_{1-} \\
& \stackrel{(58), (59)}{\lesssim} \sqrt{R} + \sum_{\substack{R/2 \geq r \geq 1 \\ \text{dyadic}}} \sqrt{r} \lesssim \sqrt{R}.
\end{aligned}$$

Here in the third inequality we used the Poincaré inequality in convolution (see [3, (201)]): for any  $R \gtrsim r_0$ ,

$$\int \eta_R (f - f_r)^2 \lesssim r^2 \int \eta_R |\nabla f|^2. \quad (92)$$

This establishes  $\langle I(r_{**} \geq R_0) \rangle \lesssim \exp(-\frac{1}{C} R_0^{\frac{3}{2}-})$  for any  $R_0 \geq 1$ , which is equivalent to (5).

*Step 2: Estimation of failure probabilities.* The plan is to pass from the deterministic estimate [Proposition 2.3](#) to the probabilistic statement [Theorem 1.2](#), and estimate the probability for the assumptions (40)–(42) in the Proposition to hold. We would like to comment that the rest of the assumptions in [Proposition 2.3](#) are standard, and, thanks to our finite range assumption, hold with probability 1 (see [Section 2.1](#) for the standard properties of correctors and  $r_{**}$ , and [1, Corollary 1] for an argument of (31)).

The starting point is the stochastic bounds that are established on stationary approximations of these quantities (i.e., the quantities without  $L$  in superscripts), namely [Proposition 2.4](#), [Lemma 2.7](#), and [Proposition 2.8](#). Therefore, to estimate the “failure probability” of  $|a_h e_i - \int \omega q_{i,T}|$ , we use a Chebyshev inequality as well as (51) (notice that  $r_{**} \geq 1$ ):

$$\begin{aligned}
\langle I(|a_h e_i - \int \omega q_{i,T}| \geq (\frac{r_{**}}{L})^\beta) \rangle & \leq \langle I(|a_h e_i - \int \omega q_{i,T}| \geq L^{-\beta}) \rangle \\
& \lesssim \exp(-\frac{1}{C} (L^\beta T^{-\frac{3}{4}})^{2-}) = \exp(-\frac{1}{C} L^{\varepsilon-}).
\end{aligned} \quad (93)$$

We may also replace  $q_{i,T}$  with  $q_{i,T}^{(L)}$  in the above (93), since by the deterministic [Proposition 2.9](#), we can estimate

$$\begin{aligned}
|\int \omega(q_{i,T} - q_{i,T}^{(L)})| & \stackrel{(8), (18)}{=} |\int \omega a \nabla(\phi_{i,T} - \phi_{i,T}^{(L)})| \\
& \lesssim (\int \omega^2)^{\frac{1}{2}} (\int_{Q_L} |\nabla(\phi_{i,T} - \phi_{i,T}^{(L)})|^2)^{\frac{1}{2}} \stackrel{(56)}{\lesssim} (\frac{\sqrt{T}}{L})^p,
\end{aligned}$$

which is much smaller than  $L^{-\beta}$  for  $\sqrt{T} = L^{1-\varepsilon}$  and  $p$  sufficiently large. This shows

$$\langle I(|a_h e_i - \int \omega q_{i,T}^{(L)}| \geq (\frac{r_{**}}{L})^\beta) \rangle \lesssim \exp(-\frac{1}{C} L^{\varepsilon-}).$$



The failure probability for the terms in (40) will be estimated slightly differently as they involve  $r_*$ . We take  $(\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}}$  as an example since it is among the terms that have the worst stochastic integrability. Again by Chebyshev inequality, we obtain

$$\langle I((\int_{B_\ell} |\nabla(\psi_T - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{1}{L})^\beta) I(\ell \geq r_*) \rangle \stackrel{(55)}{\lesssim} \exp(-\frac{1}{C}(L^\beta T^{-\frac{3}{4}})^{1-}) = \exp(-\frac{1}{C}L^{\frac{\epsilon}{2}-}),$$

which, in view of (57) and  $r_{**} \geq r_*$ , can be changed to

$$\langle I((\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{1}{L})^\beta) I(\ell \geq r_*) \rangle \lesssim \exp(-\frac{1}{C}L^{\frac{\epsilon}{2}-}).$$

We now remove the constraint  $\ell \geq r_*$  using Bayes' formula

$$\begin{aligned} & \langle I((\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{r_{**}}{L})^\beta) \Big| \ell \geq r_{**} \rangle \\ & \leq \langle I((\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{1}{L})^\beta) \Big| \ell \geq r_{**} \rangle \\ & = \frac{\langle I((\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{1}{L})^\beta) I(\ell \geq r_{**}) \rangle}{\langle I(\ell \geq r_{**}) \rangle} \\ & \leq \frac{\langle I((\int_{B_\ell} |\nabla(\psi_T^{(L)} - \psi)|^2)^{\frac{1}{2}} \geq \sqrt{T}(\frac{1}{L})^\beta) I(\ell \geq r_*) \rangle}{\langle I(\ell \geq r_{**}) \rangle} \\ & \stackrel{(5)}{\lesssim} \frac{\exp(-\frac{1}{C}L^{\frac{\epsilon}{2}-})}{1 - \exp(-\frac{1}{C}\ell^{\frac{3}{2}-})}. \end{aligned}$$

We take  $\ell$  to be reasonably large and independent of  $L, T, \beta$  so that  $1 - \exp(-\frac{1}{C}\ell^{\frac{3}{2}-}) \geq \frac{1}{2}$ . Since the probability of all failure events can be estimated as such, we derive the total failure probability estimate as desired.

Finally for the failure probability for (41), we first obtain by Chebyshev inequality

$$\langle I((\int_{B_L} \psi_T^2)^{\frac{1}{2}} \geq T(\frac{1}{L})^\beta) I(L \geq r_*) \rangle \stackrel{(54)}{\lesssim} \exp(-\frac{1}{C}L^{\frac{\epsilon}{2}-}),$$

which, in view of (99) and (57), as well as  $r_{**} \geq r_*$ , can be changed to

$$\langle I((\int_{B_L} |\psi_T^{(L)} - \psi|^2)^{\frac{1}{2}} \geq L^2(\frac{r_{**}}{L})^\beta) I(L \geq r_{**}) \rangle \lesssim \exp(-\frac{1}{C}L^{\frac{\epsilon}{2}-}),$$

and we again finish the proof by dropping the constraint  $L \geq r_{**}$  using Bayes' formula.  $\square$

## 5.2. Proof of Proposition 2.1

By [1, Theorem 2], for  $u_h$  that satisfies (30), we have

$$\sup_{R \geq r_{**}} (\frac{R}{r_{**}})^{d+\beta} \left( \frac{1}{R^d} \int_{B_R^c} |\nabla(u - (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij})u_h)|^2 \right)^{\frac{1}{2}} \lesssim \sup_{R \geq r_{**}} (\frac{R}{r_{**}})^d \left( \frac{1}{R^d} \int_{B_R^c} |\nabla u|^2 \right)^{\frac{1}{2}}.$$

Thus, it suffices to prove that, for every  $R \geq r_{**}$ ,

$$R^{\frac{d}{2}} \left( \int_{B_R^c} |\nabla u|^2 \right)^{\frac{1}{2}} \lesssim \ell^d.$$

For  $r_{**} \leq R \leq 2\ell$ , the proof is a standard energy estimate:

$$R^{\frac{d}{2}} \left( \int_{B_R^c} |\nabla u|^2 \right)^{\frac{1}{2}} \leq R^{\frac{d}{2}} \left( \int |\nabla u|^2 \right)^{\frac{1}{2}} \stackrel{(3)}{\lesssim} R^{\frac{d}{2}} \left( \int_{B_\ell} |g|^2 \right)^{\frac{1}{2}} \leq R^{\frac{d}{2}} \ell^{\frac{d}{2}} \lesssim \ell^d.$$

The proof for  $R \geq 2\ell$  uses [1, Lemma 4 (c)]: since  $u \in Y_1(L)$ ,

$$\left( \frac{1}{R^d} \int_{B_R^c} |\nabla u|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{\ell}{R} \right)^d \left( \frac{1}{\ell^d} \int_{B_\ell^c} |\nabla u|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{\ell}{R} \right)^d \ell^{-\frac{d}{2}} \left( \int |\nabla u|^2 \right)^{\frac{1}{2}} \stackrel{(3)}{\lesssim} \frac{\ell^{\frac{d}{2}}}{R^d} \left( \int |g|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{\ell}{R} \right)^d.$$

□

### 5.3. Proof of Corollary 2.2

Define  $w = \hat{u} - u$  and  $w_D = (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h - u$ , then

$$-\nabla \cdot a \nabla w = 0 \quad \text{in } Q_L, \quad w = w_D \quad \text{on } \partial Q_L. \quad (94)$$

Since  $w$  is  $a$ -harmonic, by (D3) (note that  $r_{**} \geq r_*$ ),

$$\int_{Q_R} |\nabla w|^2 \lesssim \int_{Q_L} |\nabla w|^2 \quad \text{for } r_{**} \leq R \leq L. \quad (95)$$

Thus, it suffices to prove

$$\left( \int_{Q_L} |\nabla w|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{Q_{2L}-Q_L} |\nabla w_D|^2 \right)^{\frac{1}{2}}, \quad (96)$$

as (95) and (96) together with Proposition 2.1 yield Corollary 2.2.

By rescaling, we may without loss of generality assume  $L = 1$ , and we can further assume  $\int_{Q_2-Q_1} w_D = 0$  as the expressions in both sides of (96) are invariant through the subtraction of constants. This allows us to define an extension  $\bar{w}_D$  of  $w_D$  on  $Q_2$  such that  $\bar{w}_D = w_D$  on  $Q_1$  and (using Poincaré inequality)

$$\left( \int_{Q_2} |\nabla \bar{w}_D|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{Q_2-Q_1} |\nabla w_D|^2 \right)^{\frac{1}{2}}. \quad (97)$$

Hence (94) can be reformulated as

$$-\nabla \cdot a \nabla (w - \bar{w}_D) = \nabla \cdot a \nabla \bar{w}_D \quad \text{in } Q_1, \quad w - \bar{w}_D = 0 \quad \text{on } \partial Q_1.$$

Now (96) follows from the standard energy estimate  $\left( \int_{Q_1} |\nabla (w - \bar{w}_D)|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{Q_1} |\nabla \bar{w}_D|^2 \right)^{\frac{1}{2}}$ , triangle inequality and (97). □

### 5.4. Proof of Proposition 2.3

The proof largely resembles that of [2, Proposition 1]. We divide the proof into four steps.

*Step 1:* We upgrade (27) and show a seemingly stronger condition

$$\frac{1}{r^2} \left( \int_{B_r} |(\psi, \Psi) - \int_{B_{r_{**}}} (\psi, \Psi)|^2 \right)^{\frac{1}{2}} \leq \left( \frac{r_{**}}{r} \right)^\beta \quad \text{for all } r \geq r_{**}. \quad (98)$$

What separates (98) from (27) is

$$\left| \int_{B_r} (\psi, \Psi) - \int_{B_{r_{**}}} (\psi, \Psi) \right| \lesssim r^2 \left( \frac{r_{**}}{r} \right)^\beta \quad \text{for } r \geq r_{**}.$$

To prove this, we apply dyadic decomposition, and (since  $\beta < 2$ ) reduce this to

$$|\oint_{B_r} (\psi, \Psi) - \oint_{B_{r'}} (\psi, \Psi)| \lesssim r^2 \left(\frac{r_{**}}{r}\right)^\beta \quad \text{for } 2r' \geq r \geq r' \geq r_{**},$$

which, by triangle inequality, is a consequence of (27). Hence, as  $(\psi, \Psi)$  are defined up to a constant, we may assume

$$\frac{1}{r^2} \left( \oint_{B_r} |(\psi, \Psi)|^2 \right)^{\frac{1}{2}} \lesssim \left(\frac{r_{**}}{r}\right)^\beta \quad \text{for all } r \geq r_{**}. \quad (99)$$

For the rest of this proof, by a scaling argument we may without loss of generality assume  $\ell = 1$ .

*Step 2:* We compare  $\tilde{u}_h^{(L)}$ , defined in (9) which satisfies  $-\nabla \cdot a_h^{(L)} \nabla \tilde{u}_h^{(L)} = \nabla \cdot g$ , with  $\tilde{u}_h$ , the solution of (19), and claim that

$$|\nabla(\tilde{u}_h^{(L)} - \tilde{u}_h)| + L|\nabla^2(\tilde{u}_h^{(L)} - \tilde{u}_h)| + L^2|\nabla^3(\tilde{u}_h^{(L)} - \tilde{u}_h)| \lesssim \left(\frac{1}{L}\right)^d \left(\frac{r_{**}}{L}\right)^\beta \quad \text{on } Q_L^c. \quad (100)$$

To prove (100), we use that the support of  $g$  is contained in  $B_1$ . It is well known that  $|\nabla^n G_h(x)| \lesssim |x|^{2-d-n}$  for any multi-index  $n$ . Using the representation formula,

$$\tilde{u}_h = G_h * \nabla \cdot g = \nabla G_h * g$$

for  $|x| \gg 1$ , and standard Schauder theory for  $|x| \lesssim 1$  (here we use the regularity of  $g \in C^{2,\gamma}$ ), we obtain for all  $x$ ,

$$\begin{aligned} |\tilde{u}_h(x)| &\lesssim \frac{1}{(1+|x|)^{d-1}}, & |\nabla \tilde{u}_h(x)| &\lesssim \frac{1}{(1+|x|)^d}, \\ |\nabla^2 \tilde{u}_h(x)| &\lesssim \frac{1}{(1+|x|)^{d+1}}, & |\nabla^3 \tilde{u}_h(x)| &\lesssim \frac{1}{(1+|x|)^{d+2}}, \end{aligned} \quad (101)$$

and

$$\sup_{0 < |y| \leq 1} \frac{|\nabla^3 \tilde{u}_h(x+y) - \nabla^3 \tilde{u}_h(x)|}{|y|^\gamma} \lesssim \frac{1}{(1+|x|)^{d+2}}. \quad (102)$$

Next we consider  $w = \tilde{u}_h^{(L)} - \tilde{u}_h$ , which satisfies the equation

$$-\nabla \cdot a_h^{(L)} \nabla w = \nabla \cdot (a_h^{(L)} - a_h) \nabla \tilde{u}_h.$$

Similarly we have the representation

$$w = G_h^{(L)} * \nabla \cdot (a_h^{(L)} - a_h) \nabla \tilde{u}_h = \nabla G_h^{(L)} * (a_h^{(L)} - a_h) \nabla \tilde{u}_h.$$

Therefore, using (101) and (102), again we appeal to representation formula and standard Schauder theory and derive

$$\begin{aligned} |\nabla w(x)| &\lesssim |a_h^{(L)} - a_h| \frac{1}{(1+|x|)^d} \stackrel{(42)}{\lesssim} \left(\frac{r_{**}}{L}\right)^\beta \frac{1}{(1+|x|)^d}, \\ |\nabla^2 w(x)| &\lesssim |a_h^{(L)} - a_h| \frac{1}{(1+|x|)^{d+1}} \stackrel{(42)}{\lesssim} \left(\frac{r_{**}}{L}\right)^\beta \frac{1}{(1+|x|)^{d+1}}, \\ |\nabla^3 w(x)| &\lesssim |a_h^{(L)} - a_h| \frac{1}{(1+|x|)^{d+2}} \stackrel{(42)}{\lesssim} \left(\frac{r_{**}}{L}\right)^\beta \frac{1}{(1+|x|)^{d+2}}, \end{aligned}$$

which is exactly (100).

Step 3: We compare  $u_h^{(L)}$  and  $u_h$  defined in (15) and (39) and claim

$$|\nabla(u_h^{(L)} - u_h)| + L|\nabla^2(u_h^{(L)} - u_h)| + L^2|\nabla^3(u_h^{(L)} - u_h)| \lesssim \left(\frac{1}{L}\right)^d \left(\frac{r^{**}}{L}\right)^\beta \text{ on } Q_L^c \quad (103)$$

and

$$|\nabla u_h^{(L)}| + L|\nabla^2 u_h^{(L)}| + L^2|\nabla^3 u_h^{(L)}| \lesssim \left(\frac{1}{L}\right)^d \text{ on } Q_L^c. \quad (104)$$

To this purpose, we define

$$\xi_i := \int \nabla g \cdot \phi_i \text{ and } \xi_{i,T}^{(L)} := \int g \cdot \nabla \phi_{i,T}^{(L)}. \quad (105)$$

In view of (100), it suffices to control the dipole and quadrupole terms:

$$\begin{aligned} & |\nabla(\xi_{i,T}^{(L)} \partial_i G_h^{(L)} - \xi_i \partial_i G_h)| + L|\nabla^2(\xi_{i,T}^{(L)} \partial_i G_h^{(L)} - \xi_i \partial_i G_h)| \\ & + L^2|\nabla^3(\xi_{i,T}^{(L)} \partial_i G_h^{(L)} - \xi_i \partial_i G_h)| \lesssim \left(\frac{1}{L}\right)^d \left(\frac{r^{**}}{L}\right)^\beta \end{aligned} \quad (106)$$

and

$$\begin{aligned} & |\nabla(c_{ij,T}^{(L)} \partial_{ij} G_h^{(L)} - c_{ij} \partial_{ij} G_h)| + L|\nabla^2(c_{ij,T}^{(L)} \partial_{ij} G_h^{(L)} - c_{ij} \partial_{ij} G_h)| \\ & + L^2|\nabla^3(c_{ij,T}^{(L)} \partial_{ij} G_h^{(L)} - c_{ij} \partial_{ij} G_h)| \lesssim \left(\frac{1}{L}\right)^d \left(\frac{r^{**}}{L}\right)^\beta, \text{ both on } Q_L^c. \end{aligned} \quad (107)$$

We have some obvious estimates for constant coefficient Green's functions:

$$|\nabla^2 G_h^{(L)}| + L|\nabla^3 G_h^{(L)}| + L^2|\nabla^4 G_h^{(L)}| + L^3|\nabla^5 G_h^{(L)}| \lesssim \left(\frac{1}{L}\right)^d \text{ on } Q_L^c,$$

and

$$\begin{aligned} & |\nabla^2(G_h^{(L)} - G_h)| + L|\nabla^3(G_h^{(L)} - G_h)| + L^2|\nabla^4(G_h^{(L)} - G_h)| + L^3|\nabla^5(G_h^{(L)} - G_h)| \\ & \lesssim |a_h^{(L)} - a_h| \left(\frac{1}{L}\right)^d \stackrel{(42)}{\lesssim} \left(\frac{1}{L}\right)^d \left(\frac{r^{**}}{L}\right)^\beta \text{ on } Q_L^c. \end{aligned}$$

Hence, recalling the definitions of  $c_{ij}$  and  $c_{ij,T}^{(L)}$ , c.f. (38) and (13), it suffices to show

$$|\xi| \lesssim 1 \text{ and } |c_{ij}| \lesssim 1 \quad (108)$$

and

$$|\xi_T^{(L)} - \xi| \lesssim \left(\frac{r^{**}}{L}\right)^\beta \text{ and } \frac{1}{L}|c_{ij,T}^{(L)} - c_{ij}| \lesssim \left(\frac{r^{**}}{L}\right)^\beta. \quad (109)$$

The arguments of (108) and (109) for  $\xi$  and  $\xi_T^{(L)}$  are straightforward:

$$|\xi| = \left| \int \phi \nabla \cdot g \right| \lesssim \left( \int_{B_1} \phi^2 \right)^{\frac{1}{2}} \stackrel{(28)}{\lesssim} 1,$$

and

$$|\xi_T^{(L)} - \xi| = \left| \int g \cdot \nabla(\phi_T^{(L)} - \phi) \right| \lesssim \left( \int_{B_1} |\nabla(\phi_T^{(L)} - \phi)|^2 \right)^{\frac{1}{2}} \stackrel{(40)}{\lesssim} \left(\frac{r^{**}}{L}\right)^\beta.$$

We now prove (108) for  $c_{ij}$ . By the growth condition (27) of  $\psi$  for  $r = 1$ , we have

$$\left| \int \psi \nabla \cdot g \right| = \left| \int (\psi - \int_{B_1} \psi) \nabla \cdot g \right| \lesssim \left( \int_{B_1} |\psi - \int_{B_1} \psi|^2 \right)^{\frac{1}{2}} \lesssim 1. \quad (110)$$

Moreover, since  $\partial_k v_{h,ij}$ , defined in (37), is a polynomial homogeneous of order 1, we have

$$|\partial_k v_{h,ij}(x)| \lesssim 1 \text{ and } |\nabla \partial_k v_{h,ij}(x)| \lesssim 1 \text{ for } x \in B_1, \quad (111)$$

thus

$$|\int \phi_k \partial_k v_{h,ij} \nabla \cdot g| \lesssim (\int_{B_1} (\phi_k \partial_k v_{h,ij})^2)^{\frac{1}{2}} \lesssim (\int_{B_1} \phi_k^2)^{\frac{1}{2}} \stackrel{(28)}{\lesssim} 1. \quad (112)$$

Combining (110) and (112) we established (108) for  $c_{ij}$ .

The proof of (109) for  $|c_T^{(L)} - c|$  is more complicated. Recalling the definitions of  $c_{ij}$  and  $c_{ij,T}^{(L)}$ , c.f. (38) and (13), we may write  $|c_T^{(L)} - c| := |\int \Xi \cdot g|$  and decompose  $\Xi$  into different terms:

$$\begin{aligned} \Xi_{ij} = & (\partial_k v_{h,ij} - \partial_k v_{h,ij}^{(L)}) \nabla \phi_k + \partial_k v_{h,ij}^{(L)} \nabla (\phi_k - \phi_{k,T}^{(L)}) + \phi_k \nabla (\partial_k v_{h,ij} - \partial_k v_{h,ij}^{(L)}) \\ & + (\phi_k - \phi_{k,T}^{(L)}) \nabla \partial_k v_{h,ij}^{(L)} + (2 - \delta_{ij}) \nabla (\psi_{ij} - \psi_{ij,T}^{(L)}) + (2 - \delta_{ij}) \left( \frac{a_{hij}}{a_{h11}} - \frac{a_{hij}^{(L)}}{a_{h11}^{(L)}} \right) \nabla \psi_{11}. \end{aligned}$$

Since  $x_i + \phi_i$  is  $a$ -harmonic, we can use Caccioppoli's estimate and the growth condition (28) on  $\phi$  to derive

$$(\int_{B_R} |\nabla \phi|^2)^{\frac{1}{2}} \lesssim 1 + \frac{1}{R} (\int_{B_{2R}} \phi^2)^{\frac{1}{2}} \lesssim 1 + \left(\frac{r_{**}}{R}\right)^\alpha \lesssim 1, \text{ for } R \geq r_{**}. \quad (113)$$

Using the same argument as (113) for  $\psi$ , we obtain

$$\begin{aligned} (\int_{B_R} |\nabla \psi|^2)^{\frac{1}{2}} & \stackrel{(23)}{\lesssim} \frac{1}{R} (\int_{B_{2R}} |\psi - \int_{B_{2R}} \psi|^2)^{\frac{1}{2}} + (\int_{B_{2R}} |(\phi, \sigma)|^2)^{\frac{1}{2}} \\ & \stackrel{(27),(28)}{\lesssim} R \left(\frac{r_{**}}{R}\right)^\beta + R \left(\frac{r_{**}}{R}\right)^\alpha \lesssim R \text{ for } R \geq r_{**}. \end{aligned} \quad (114)$$

Since  $\partial_k v_{h,ij}$  and  $\partial_k v_{h,ij}^{(L)}$  (defined in (14)) are two polynomials that are homogeneous of order 1, and their coefficients differ by  $|a_h^{(L)} - a_h|$ , which by (42) is bounded by  $(\frac{r_{**}}{L})^\beta$ , we have

$$|\partial_k v_{h,ij} - \partial_k v_{h,ij}^{(L)}| \lesssim \left(\frac{r_{**}}{L}\right)^\beta \text{ and } |\nabla (\partial_k v_{h,ij} - \partial_k v_{h,ij}^{(L)})| \lesssim \left(\frac{r_{**}}{L}\right)^\beta \text{ on } B_1. \quad (115)$$

Therefore combine (115) and (111) we obtain

$$|\partial_k v_{h,ij}^{(L)}(x)| \lesssim 1 \text{ and } |\nabla \partial_k v_{h,ij}^{(L)}(x)| \lesssim 1 \text{ on } B_1. \quad (116)$$

We are now ready to estimate

$$|c_T^{(L)} - c| = |\int \Xi \cdot g| \lesssim (\int_{B_1} |\Xi|^2)^{\frac{1}{2}},$$

For the first term of  $\Xi$  we use (115) and (113); the second term uses (116) and (40); the third term is controlled by (115) and (28); the fourth term can be estimated using (116) and (40); the fifth term can be bounded above by (40); finally the six term is controlled by (42) as well as (114). This finishes the proof of (109).

*Step 4: Conclusion.* We finally compare  $u^{(L)}$  defined through (16) with  $\hat{u}$  defined through Corollary 2.2 and claim that

$$(\int_{B_R} |\nabla (u^{(L)} - \hat{u})|^2)^{\frac{1}{2}} \lesssim \left(\frac{1}{L}\right)^d \left(\frac{r_{**}}{L}\right)^\beta \text{ for } L \geq R \geq r_{**}.$$

By [Corollary 2.2](#) this leads to the conclusion of the proposition. The difference  $w := u^{(L)} - \hat{u}$  satisfies

$$-\nabla \cdot a \nabla w = 0 \quad \text{in } Q_L, \quad w = w_D \text{ on } \partial Q_L$$

where  $w_D := (1 + \phi_{i,T}^{(L)} \partial_i + \psi_{ij,T}^{(L)} \partial_{ij}) u_h^{(L)} - (1 + \phi_i \partial_i + \psi_{ij} \partial_{ij}) u_h$ . By (a slight adaptation of) the argument [\(96\)](#) we have

$$\left( \int_{Q_L} |\nabla w|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{Q_{\frac{5}{4}L} - Q_L} |\nabla w_D|^2 \right)^{\frac{1}{2}}.$$

Hence by [\(95\)](#) it suffices to show

$$\left( \int_{Q_{\frac{5}{4}L} - Q_L} |\nabla w_D|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{1}{L} \right)^d \left( \frac{r_{**}}{L} \right)^\beta.$$

We break down  $w_D$  into eight parts, prove an estimate for each of these, and use a triangle inequality to get our desired result:

$$\begin{aligned} \nabla w_D &= \partial_i (u_h^{(L)} - u_h) (e_i + \nabla \phi_i) + \phi_i \nabla \partial_i (u_h^{(L)} - u_h) + \partial_i u_h^{(L)} \nabla (\phi_{i,T}^{(L)} - \phi_i) \\ &\quad + (\phi_{i,T}^{(L)} - \phi_i) \nabla \partial_i u_h^{(L)} + \partial_{ij} (u_h^{(L)} - u_h) \nabla \psi_{ij} + \psi_{ij} \nabla \partial_{ij} (u_h^{(L)} - u_h) \\ &\quad + \partial_{ij} u_h^{(L)} \nabla (\psi_{ij,T}^{(L)} - \psi_{ij}) + (\psi_{ij,T}^{(L)} - \psi_{ij}) \nabla \partial_{ij} u_h^{(L)}. \end{aligned}$$

The first estimate follows from [\(113\)](#) and [\(103\)](#). The second estimate follows from [\(103\)](#) and [\(28\)](#). The third and the fourth come from [\(40\)](#) and [\(104\)](#). The fifth term is bounded due to [\(103\)](#) and [\(114\)](#). The sixth term is controlled by [\(103\)](#) and [\(99\)](#). The seventh term is good thanks to [\(104\)](#) and [\(40\)](#). Finally the eighth term is controlled by [\(104\)](#) and [\(41\)](#).  $\square$

### 5.5. Proof of Proposition 2.4

*Step 1:* proof of [\(48\)](#). We will omit the proof for the second term since their proofs are identical. Thanks to the decomposition of  $\phi$  [\(65\)](#) and  $\phi_T$  [\(66\)](#), it suffices to prove

$$\left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^\infty dt (1 - \exp(-\frac{t}{T})) \nabla S(t) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim T^{-\frac{3}{4}}.$$

The proof is direct using [\(D11\)](#) and [Lemma 4.7](#):

$$\begin{aligned} &\int_0^\infty dt (1 - \exp(-\frac{t}{T})) \left\| I(R \geq r_*) \left( \int \eta_R |\nabla S(t) a e|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{\text{(D11)}}{\lesssim} \int_0^\infty dt (1 \wedge \frac{t}{T}) \left\| \left( \int \eta_{\sqrt{t}} |\nabla S(\frac{t}{2}) a e|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(79)}{\lesssim} \int_0^\infty dt (1 \wedge \frac{t}{T}) \frac{1}{t} (1 \wedge \frac{1}{\sqrt{t}})^{\frac{d}{2}} \stackrel{d \geq 3}{\approx} T^{-\frac{3}{4}}. \end{aligned}$$

*Step 2:* proof of [\(46\)](#). Similar to Step 1, we need to show

$$\left\| \left( \int \eta_{\sqrt{T}} \left| \int_0^\infty dt (1 - \exp(-\frac{t}{T})) \nabla S(t) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim T^{-\frac{3}{4}}.$$

For the range  $t \in (0, T)$  we have as above

$$\left\| \left( \int \eta_{\sqrt{T}} \left| \int_0^T dt (1 - \exp(-\frac{t}{T})) \nabla S(t) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-}$$

$$\lesssim \int_0^T dt (1 - \exp(-\frac{t}{T})) \left\| \left( \int \eta_{\sqrt{T}} |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \stackrel{(79), d \geq 3}{\lesssim} T^{-\frac{3}{4}}.$$

For the range  $t \in (T, \infty)$ , we divide the integral into dyadic intervals  $(2^k T, 2^{k+1} T)$  and use the Meyers' inequality [3, Lemma 7] in form of

$$\left( \int \eta_{\sqrt{T}} \int_{2^k T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{\sqrt{2^{k+1} T}}{\sqrt{T}} \right)^{\frac{d}{2}-\gamma} \left( \int \eta_{\sqrt{2^{k+1} T}} \int_{2^{k-1} T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right)^{\frac{1}{2}}, \quad (117)$$

where  $\gamma = \gamma(d, \lambda) > 0$ . We then use the third characterization of the norm  $\|\cdot\|_s$  in (A1) and the stochastic estimate (79) to obtain

$$\begin{aligned} \left\| \left( \int \eta_{\sqrt{2^{k+1} T}} \int_{2^{k-1} T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} &\stackrel{(A1)}{\sim} \left\| \int \eta_{\sqrt{2^{k+1} T}} \int_{2^{k-1} T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right\|_{1-}^{\frac{1}{2}} \\ &\leq \left( \int_{2^{k-1} T}^{2^{k+1} T} dt \left\| \int \eta_{\sqrt{2^{k+1} T}} |\nabla S(t)ae|^2 \right\|_{1-} \right)^{\frac{1}{2}} \\ &\stackrel{(A1)}{\sim} \left( \int_{2^{k-1} T}^{2^{k+1} T} dt \left\| \left( \int \eta_{\sqrt{2^{k+1} T}} |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-}^2 \right)^{\frac{1}{2}} \\ &\stackrel{(79)}{\lesssim} \left( \int_{2^{k-1} T}^{2^{k+1} T} dt t^{-2-\frac{d}{2}} \right)^{\frac{1}{2}} \sim (2^{k-1} T)^{-\frac{1}{2}-\frac{d}{4}}. \end{aligned}$$

With the above two estimates we are ready to finish the proof:

$$\begin{aligned} &\left\| \left( \int \eta_{\sqrt{T}} \int_T^\infty dt (1 - \exp(-\frac{t}{T})) |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\lesssim \sum_{k=0}^\infty \left\| \left( \int \eta_{\sqrt{T}} \int_{2^k T}^{2^{k+1} T} dt (1 - \exp(-\frac{t}{T})) |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\lesssim \sum_{k=0}^\infty \left( \int_{2^k T}^{2^{k+1} T} dt (1 - \exp(-\frac{t}{T}))^2 \right)^{\frac{1}{2}} \left\| \left( \int \eta_{\sqrt{T}} \int_{2^k T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(117)}{\lesssim} \sum_{k=0}^\infty (2^k T)^{\frac{1}{2}} 2^{\frac{k+1}{2}(\frac{d}{2}-\gamma)} \left\| \left( \int \eta_{\sqrt{2^{k+1} T}} \int_{2^{k-1} T}^{2^{k+1} T} dt |\nabla S(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\lesssim \sum_{k=0}^\infty (2^k T)^{\frac{1}{2}} 2^{\frac{k+1}{2}(\frac{d}{2}-\gamma)} (2^{k-1} T)^{-\frac{1}{2}-\frac{d}{4}} \lesssim T^{-\frac{d}{4}}. \end{aligned}$$

□

## 5.6. Proof of Lemma 2.7

Substituting the definition (18) into (82), where we may change the convolution kernel from Gaussian to any Schwartz function with the same scale and preserve the CLT-scaling (see [3, Lemma 13, Step 4]), and choosing  $R = L$ , we obtain

$$\|a_h e_i - \int \omega_L q_i\|_{2-} \lesssim L^{-\frac{3}{2}}.$$

The rest of the proof follows from a Cauchy-Schwarz,  $q - q_T = a\nabla(\phi - \phi_T)$ , and [Proposition 2.4](#), (46), as well as (A7):

$$\|\int \omega_L(q - q_T)\|_{2-} \leq (\int \omega_L)^{\frac{1}{2}} \|(\int \omega_L |q - q_T|^2)^{\frac{1}{2}}\|_{2-} \lesssim \|(\int \eta_L |\nabla(\phi - \phi_T)|^2)^{\frac{1}{2}}\|_{2-} \lesssim \sqrt{T}^{-\frac{3}{2}}.$$

With a triangle inequality we arrive at  $\|a_{he_i} - \int \omega_L q_{i,T}\|_{2-} \lesssim \sqrt{T}^{-\frac{3}{2}}$ .  $\square$

### 5.7. Proof of Propositions 2.8 and 2.10

We will focus on the proof of (55) with full details and discuss the other estimates afterwards. We recall the decomposition formulas (68), (69) for  $\psi$  and  $\psi_T$  and write at least formally

$$\begin{aligned} \psi - \psi_T = & \underbrace{\int_0^\infty dt_0 \int_0^\infty dt_1 (1 - \exp(-\frac{t_0 + t_1}{T})) S(t_0) (aS(t_1)ae - \bar{S}(t_1) \times ae)}_{\text{double integral term}} \\ & - \underbrace{\int_0^\infty dt_0 \int_0^\infty dt_1 \int_0^\infty dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) S(t_0) \bar{S}(t_1) \times a\nabla S(t_2)ae}_{\text{triple integral term}}. \end{aligned} \quad (118)$$

The strategy is to divide these integrals into several regimes depending on the ordering of  $t_0, t_1, t_2$ , then estimate the layers one-by-one, using either the deterministic [Lemma B.2](#) when the time variable in question is not the largest, or the stochastic [Propositions 4.4](#) and [4.6](#) and [Lemma 4.7](#) when the time variable being estimated is the largest of all remaining variable, with [Corollary D.2](#) being used to adjust the averaging scales whenever necessary. As we mentioned in [Remark 4.2](#), our proof actually first shows that the integrals of (68)–(70) satisfy the desired bounds (58), (54), which guarantee that the integrals are stationary functions, and hence must correspond to correctors thanks to their algebraic equivalence and uniqueness of stationary correctors.

*Proof of (55).* We divide the estimates into two parts, one for the double integral term, the other for the triple integral term.

*Step 1:* estimates for the double integral term. We only estimate the term with  $S(t_0)aS(t_1)ae$  since the estimates for the other term are identical. We divide the double integral into two regimes:  $t_0 > t_1$  and  $t_1 > t_0$ .

*Case 1.1:*  $t_0 > t_1$ . The main ingredient we use in this case is (77) with  $g = \sqrt{t_1}aS(t_1)ae$  which by [Lemma 4.7](#) is approximately local on scale  $1 \vee \sqrt{t_1}$  with stochastic integrability  $2-$ , so applying the variable-coefficient semigroup  $S(\frac{t_0}{2})$  will lead to the CLT factor  $(\frac{1 \vee \sqrt{t_1}}{1 \vee \sqrt{t_0}})^{\frac{d}{2}}$  with a loss of stochastic integrability to  $1-$ :

$$\begin{aligned} & \|I(R \geq r_*) (\int \eta_R |\int_0^\infty dt_0 \int_0^{t_0} dt_1 (1 - \exp(-\frac{t_0 + t_1}{T})) \nabla S(t_0) aS(t_1)ae|^2)^{\frac{1}{2}}\|_{1-} \\ & \leq \int_0^\infty dt_0 \int_0^{t_0} dt_1 (1 - \exp(-\frac{t_0 + t_1}{T})) \|I(R \geq r_*) (\int \eta_R |\nabla S(t_0) aS(t_1)ae|^2)^{\frac{1}{2}}\|_{1-} \\ & \stackrel{(D11)}{\lesssim} \int_0^\infty dt_0 (1 \wedge \frac{t_0}{T}) \int_0^{t_0} dt_1 \|(\int \eta_{R \vee \sqrt{t_0}} |\nabla S(\frac{t_0}{2}) aS(t_1)ae|^2)^{\frac{1}{2}}\|_{1-} \\ & \stackrel{(77)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} (1 \wedge \frac{t_0}{T}) \int_0^{t_0} dt_1 (\frac{1 \vee \sqrt{t_1}}{1 \vee \sqrt{t_0}})^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} (\|(\int \eta_{\sqrt{t_1}} |\sqrt{t_1} S(t_1)ae|^2)^{\frac{1}{2}}\|_{2-} + \|\bar{g}\|_{2-}) \end{aligned}$$



$$\stackrel{(79),(80)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} (1 \wedge \frac{1}{\sqrt{t_0}})^{\frac{d}{2}} (1 \wedge \frac{t_0}{T}) \int_0^{t_0} dt_1 (1 \vee \sqrt{t_1})^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} T^{-\frac{1}{4}}.$$

Case 1.2:  $t_1 > t_0$ . Before we start the proof, we present here a general integration by parts formula in time which will be used repeatedly hereafter: for any  $t, t_3, T \geq 0$ ,

$$\begin{aligned} \int_0^t d\tau (1 - \exp(-\frac{\tau + t_3}{T})) S(\tau) g &= (1 - \exp(-\frac{t + t_3}{T})) \int_0^t d\tau S(\tau) g \\ &\quad - \frac{1}{T} \exp(-\frac{t_3}{T}) \int_0^t dt' \exp(-\frac{t'}{T}) \int_0^{t'} d\tau S(\tau) g. \end{aligned} \quad (119)$$

To estimate this case, we first switch the order of integration of  $t_0$  and  $t_1$  and use (119) with  $\tau = t_0$ ,  $t = t_3 = t_1$  and  $g = aS(t_1)ae$ , then appeal to Lemma 4.10 with  $ae$  playing the role of  $g$ :

$$\begin{aligned} &\left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^\infty dt_1 \int_0^{t_1} dt_0 (1 - \exp(-\frac{t_0 + t_1}{T})) \nabla S(t_0) aS(t_1) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(119)}{\leq} \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^\infty dt_1 (1 - \exp(-\frac{2t_1}{T})) \int_0^{t_1} dt_0 \nabla S(t_0) aS(t_1) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\quad + \left\| I(R \geq r_*) \left( \int \eta_R \left| \frac{1}{T} \int_0^\infty dt_1 \exp(-\frac{t_1}{T}) \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \int_0^{t_0} d\tau \nabla S(\tau) aS(t_1) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\leq \int_0^\infty dt_1 (1 - \exp(-\frac{2t_1}{T})) \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 \nabla S(t_0) aS(t_1) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\quad + \frac{1}{T} \int_0^\infty dt_1 \exp(-\frac{t_1}{T}) \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_0} d\tau \nabla S(\tau) aS(t_1) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(87)}{\lesssim} \int_0^\infty dt_1 (1 \wedge \frac{t_1}{T}) \left\| \left( \int \eta_{\sqrt{t_1}} |(\sqrt{t_1} \nabla S(\frac{t_1}{2}) ae, S(\frac{t_1}{2}) ae)|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(79)}{\lesssim} \int_0^\infty dt_1 (1 \wedge \frac{t_1}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} T^{-\frac{1}{4}}. \end{aligned}$$

Step 2: estimates for the triple integral term. We divide the triple integral into five regimes. In particular, we use Lemma 4.10 whenever  $t_0 < t_1$ , and Lemma 4.11 whenever  $t_1 > t_2$  (in which case the  $t_2$  integral always stays inside the stochastic norm  $\|\cdot\|$ ).

Case 2.1:  $t_0 > t_1 > t_2$ . In this case, notice that  $\int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \bar{S}(t_1) \times a \nabla S(t_2) ae$  has locality scale  $1 \vee \sqrt{t_1}$  by Lemma 4.11 and stochastic integrability 2−, therefore, by Proposition 4.4, applying the semigroup  $S(t_0)$  will induce stochastic cancellations by a CLT-factor  $(\frac{1 \vee \sqrt{t_1}}{1 \vee \sqrt{t_0}})^d$  with a loss of stochastic integrability to 1−:

$$\begin{aligned} &\int_0^\infty dt_0 \int_0^{t_0} dt_1 \left\| I(R \geq r_*) \left( \int \eta_R \left| \nabla S(t_0) \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \bar{S}(t_1) \times a \nabla S(t_2) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\ &\stackrel{(D11)}{\lesssim} \int_0^\infty dt_0 \int_0^{t_0} dt_1 \left\| \left( \int \eta_{R \vee \sqrt{t_0}} \left| \nabla S(\frac{t_0}{2}) \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \bar{S}(t_1) \times a \nabla S(t_2) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\ &\stackrel{(77)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{1 \vee \sqrt{t_0}} \right)^{\frac{d}{2}} \left\| \left( \int \eta_{\sqrt{t_1}} \left| \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \bar{S}(t_1) \times a \nabla S(t_2) ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ &\stackrel{(88)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} (1 \wedge \frac{1}{\sqrt{t_0}})^{\frac{d}{2}} \int_0^{t_0} dt_1 (1 \vee \sqrt{t_1})^{\frac{d}{2}} (1 \wedge \frac{t_0}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} T^{-\frac{1}{4}}. \end{aligned}$$

Here the  $\bar{g}$  term is not necessary when using (77), thanks to Lemma 4.11, which we used for  $t_3 = t_0 + t_1 \sim t_0$ .

Case 2.2:  $t_0 > t_2 > t_1$ . Now  $\sqrt{t_1}\bar{S}(t_1) \times a\nabla S(t_2)ae$  has locality scale  $1 \vee \sqrt{t_2}$  (using Proposition 4.6 for  $t_1$  and Lemma 4.7 for  $t_2$ ) and stochastic integrability 2− (using Lemma 4.7 and Remark 4.5), so Proposition 4.4 tells us that applying  $S(t_0)$  on it will result in stochastic cancellations by a CLT-factor  $(\frac{1 \vee \sqrt{t_2}}{1 \vee \sqrt{t_0}})^d$  and a loss of stochastic integrability to 1−:

$$\begin{aligned}
& \int_0^\infty dt_0 \int_0^{t_0} dt_1 \int_{t_1}^{t_0} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \left\| I(R \geq r_*) \left( \int \eta_R |\nabla S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \stackrel{(D11)}{\lesssim} \int_0^\infty dt_0 (1 \wedge \frac{t_0}{T}) \int_0^{t_0} dt_1 \int_{t_1}^{t_0} dt_2 \left\| \left( \int \eta_{\sqrt{t_0}} |\nabla S(\frac{t_0}{2})\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \stackrel{(77)}{\lesssim} \int_0^\infty dt_0 (1 \wedge \frac{t_0}{T}) \frac{1}{t_0} \int_0^{t_0} dt_1 \int_{t_1}^{t_0} dt_2 \left( \frac{1 \vee \sqrt{t_2}}{1 \vee \sqrt{t_0}} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \\
& \quad \times \left( \left\| \left( \int \eta_{\sqrt{t_2}} |\sqrt{t_1}\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{G}\|_{2-} \right) \\
& \stackrel{(B3),(78),(139)}{\lesssim} \int_0^\infty dt_0 (1 \wedge \frac{t_0}{T}) \frac{1}{t_0} \int_0^{t_0} dt_1 \frac{1}{\sqrt{t_1}} \\
& \quad \times \int_{t_1}^{t_0} dt_2 \left( \frac{1 \vee \sqrt{t_2}}{1 \vee \sqrt{t_0}} \right)^{\frac{d}{2}} \frac{1}{t_2} \left( \left\| \left( \int \eta_{\sqrt{t_2}} |t_2 \nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{g}\|_{2-} \right) \\
& \stackrel{(79),(80)}{\lesssim} \int_0^\infty dt_0 (1 \wedge \frac{t_0}{T}) \frac{1}{t_0} (1 \wedge \frac{1}{\sqrt{t_0}})^{\frac{d}{2}} \int_0^{t_0} dt_1 \frac{1}{\sqrt{t_1}} \int_{t_1}^{t_0} dt_2 (1 \vee \sqrt{t_2})^{\frac{d}{2}} \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}} \lesssim T^{-\frac{1}{4}}.
\end{aligned}$$

Case 2.3:  $t_2 > t_0 > t_1$ . In this case, since both  $t_0$  and  $t_1$  are small compared to the innermost  $t_2$ , we use the deterministic Lemma B.2 on both  $t_0$  and  $t_1$ , and the Lipschitz estimate (D11) and finally Lemma 4.7 on  $t_2$ :

$$\begin{aligned}
& \int_0^\infty dt_0 \int_0^{t_0} dt_1 \int_{t_0}^\infty dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \left\| I(R \geq r_*) \left( \int \eta_R |\nabla S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(D11)}{\lesssim} \int_0^\infty dt_0 \int_0^{t_0} dt_1 \int_{t_0}^\infty dt_2 (1 \wedge \frac{t_2}{T}) \left\| I(R \vee \sqrt{t_0} \geq r_*) \left( \int \eta_{R \vee \sqrt{t_0}} |\nabla S(\frac{t_0}{2})\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(B3)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \int_{t_0}^\infty dt_2 (1 \wedge \frac{t_2}{T}) \left\| I(R \vee \sqrt{t_0} \geq r_*) \left( \int \eta_{R \vee \sqrt{t_0}} |\bar{S}(t_1) \times a\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(B3)}{\lesssim} \int_0^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \frac{1}{\sqrt{t_1}} \int_{t_0}^\infty dt_2 (1 \wedge \frac{t_2}{T}) \left\| I(R \vee \sqrt{t_0} \geq r_*) \left( \int \eta_{R \vee \sqrt{t_0}} |\nabla S(t_2)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(D11)}{\lesssim} \int_0^\infty dt_0 \frac{1}{\sqrt{t_0}} \int_{t_0}^\infty dt_2 (1 \wedge \frac{t_2}{T}) \left\| \left( \int \eta_{\sqrt{t_2}} |\nabla S(\frac{t_2}{2})ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(79)}{\lesssim} \int_0^\infty dt_0 \frac{1}{\sqrt{t_0}} \int_{t_0}^\infty dt_2 (1 \wedge \frac{t_2}{T}) \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}} \lesssim T^{-\frac{1}{4}}.
\end{aligned}$$

Case 2.4:  $t_2 > t_1 > t_0$ . We start the proof by using (119) on  $t_0$  with  $t = t_1$ ,  $t_3 = t_1 + t_2$ , and  $g = \bar{S}(t_1) \times a\nabla S(t_2)ae$ ,

$$\begin{aligned}
& \int_0^\infty dt_2 \int_0^{t_2} dt_1 \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \nabla S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(119)}{\leq} \int_0^\infty dt_2 \int_0^{t_2} dt_1 (1 - \exp(-\frac{2t_1 + t_2}{T})) \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 \nabla S(t_0)\bar{S}(t_1) \times a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \quad + \frac{1}{T} \int_0^\infty dt_2 \exp(-\frac{t_2}{T}) \int_0^{t_2} dt_1 \exp(-\frac{t_1}{T}) \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \\
& \quad \times \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_0} d\tau \nabla S(\tau)\bar{S}(t_1) \times a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-}.
\end{aligned}$$

We now state a slight modification of [Lemma 4.10](#) in the sense of

$$\begin{aligned} & \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_0} d\tau \nabla S(\tau) \bar{S}(t_1) g \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \lesssim \left\| I(R \vee \sqrt{t_1} \geq r_*) \left( \int \eta_{R \vee \sqrt{t_1}} (|\sqrt{t_1} \nabla \bar{S}(\frac{t_1}{2}) g|^2 + |\bar{S}(\frac{t_1}{2}) g|^2) \right)^{\frac{1}{2}} \right\|_{2-}. \end{aligned} \quad (120)$$

In this regime we apply (120) to  $g = a \nabla S(t_2) a e$ , so applying [Lemma B.2](#) on  $t_1$  and finally the Lipschitz estimate (D11) and [Lemma 4.7](#) on  $t_2$  yields

$$\begin{aligned} & \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_0} d\tau \nabla S(\tau) \bar{S}(t_1) \times a \nabla S(t_2) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \stackrel{(120)}{\lesssim} \left\| I(R \vee \sqrt{t_1} \geq r_*) \left( \int \eta_{R \vee \sqrt{t_1}} (|\sqrt{t_1} \nabla \bar{S}(\frac{t_1}{2}) \times a \nabla S(t_2) a e|^2 \right. \right. \\ & \quad \left. \left. + |\bar{S}(\frac{t_1}{2}) \times a \nabla S(t_2) a e|^2) \right)^{\frac{1}{2}} \right\|_{2-} \\ & \stackrel{(B3)}{\lesssim} \frac{1}{\sqrt{t_1}} \left\| \left( \int \eta_{R \vee \sqrt{t_1}} |\nabla S(t_2) a e|^2 \right)^{\frac{1}{2}} \right\|_{2-} \stackrel{(D11)}{\lesssim} \frac{1}{\sqrt{t_1}} \left\| \left( \int \eta_{\sqrt{t_2}} |\nabla S(\frac{t_2}{2}) a e|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \stackrel{(79)}{\lesssim} \frac{1}{\sqrt{t_1}} \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}}. \end{aligned}$$

Combining all the arguments, we are able to conclude

$$\begin{aligned} & \int_0^\infty dt_2 \int_0^{t_2} dt_1 \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \nabla S(t_0) \bar{S}(t_1) \right. \right. \right. \\ & \quad \left. \left. \times a \nabla S(t_2) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \lesssim \int_0^\infty dt_2 (1 \wedge \frac{t_2}{T}) \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}} \int_0^{t_2} dt_1 \frac{1}{\sqrt{t_1}} \\ & \quad + \frac{1}{T} \int_0^\infty dt_2 \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}} \int_0^{t_2} dt_1 \frac{1}{\sqrt{t_1}} \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \\ & \stackrel{d \geq 3}{\lesssim} T^{-\frac{1}{4}}. \end{aligned}$$

*Case 2.5:*  $t_1 > t_0 \vee t_2$ . We start by using (119) on  $t_0$  as in Case 2.4:

$$\begin{aligned} & \int_0^\infty dt_1 \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \nabla S(t_0) \bar{S}(t_1) \right. \right. \right. \\ & \quad \left. \left. \times a \nabla S(t_2) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \stackrel{(119)}{\leq} \int_0^\infty dt_1 \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_0 \int_0^{t_1} dt_2 (1 - \exp(-\frac{2t_1 + t_2}{T})) \nabla S(t_0) \bar{S}(t_1) \right. \right. \right. \\ & \quad \left. \left. \times a \nabla S(t_2) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\ & \quad + \frac{1}{T} \int_0^\infty dt_1 \int_0^{t_1} dt_0 \exp(-\frac{t_1 + t_0}{T}) \\ & \quad \times \left\| I(R \geq r_*) \left( \int \eta_R \left| \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \int_0^{t_0} \nabla S(\tau) \bar{S}(t_1) \times a \nabla S(t_2) a e \right|^2 \right)^{\frac{1}{2}} \right\|_{2-}. \end{aligned}$$

We again continue by using (120) with  $g = a\nabla S(t_2)ae$ , though this time we use Lemma 4.11 with  $t_3 = 2t_1$  on  $t_1, t_2$ :

$$\begin{aligned}
& \int_0^\infty dt_1 \left\| I(R \geq r_*) \left( \int_0^{t_1} \eta_R \left| \int_0^{t_1} dt_0 \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_0 + t_1 + t_2}{T})) \nabla S(t_0) \bar{S}(t_1) \right. \right. \right. \\
& \quad \times \left. \left. \left. a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(120)}{\lesssim} \int_0^\infty dt_1 \left\| \left( \int_0^{t_1} \eta_{\sqrt{t_1}} \left| \int_0^{t_1} dt_2 (1 - \exp(-\frac{2t_1 + t_2}{T})) (|\sqrt{t_1} \nabla \bar{S}(\frac{t_1}{2}) \right. \right. \right. \\
& \quad \times \left. \left. \left. a\nabla S(t_2)ae|^2 + |\bar{S}(\frac{t_1}{2}) \times a\nabla S(t_2)ae|^2 \right) \right)^{\frac{1}{2}} \right\|_{2-} \\
& \quad + \frac{1}{T} \int_0^\infty dt_1 \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \\
& \quad \times \left\| \left( \int_0^{t_1} \eta_{\sqrt{t_1}} \left| \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) (|\sqrt{t_1} \nabla \bar{S}(\frac{t_1}{2}) \times a\nabla S(t_2)ae|^2 + |\bar{S}(\frac{t_1}{2}) \right. \right. \right. \\
& \quad \times \left. \left. \left. a\nabla S(t_2)ae|^2) \right)^{\frac{1}{2}} \right\|_{2-} \\
& \stackrel{(88),(89)}{\lesssim} \int_0^\infty (1 \wedge \frac{t_1}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \\
& \quad + \frac{1}{T} \int_0^\infty dt_1 \int_0^{t_1} dt_0 \exp(-\frac{t_0}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \lesssim^{d \geq 3} T^{-\frac{1}{4}}.
\end{aligned}$$

This finishes the proof of (55). The estimates (58) and (54) follow by simple modifications of the above argument: Indeed, notice that when any of the  $t_i$ 's is larger than  $T$ , the weight  $1 - \exp(-\frac{t_0+t_1+t_2}{T})$  is  $\sim 1$ . If we replace the weights  $1 - \exp(-\frac{t_0+t_1+t_2}{T})$  by 1 and repeat the integral estimates we get (58) for  $\psi$  (the calculations are actually easier since integration by parts on the weights are no longer needed). If we replace the weights by  $\exp(-\frac{t_0+t_1+t_2}{T})$  and remove the gradient in front of  $S(t_0)$  (which will make the bound worse by  $\sqrt{T}$ ) we obtain (54).

*Proof of (59) for  $\nabla \psi$ .* When any one of the  $t_i$ 's in representation (68) is larger than  $R^2$ , we can use  $|f_R| \lesssim (\int \eta_R |f|^2)^{\frac{1}{2}}$  to derive the bound  $R^{1-\frac{d}{2}}$ . For example, in the case  $t_0 > t_1 > t_2$  of the triple integral in (68) (which means  $t_0 > R^2$ ), we can estimate as

$$\begin{aligned}
& \left\| I(R \geq r_*) \left( \int_{R^2}^\infty dt_0 \int_0^{t_0} dt_1 \int_0^{t_1} dt_2 \nabla S(t_0) \bar{S}(t_1) \times a\nabla S(t_2)ae \right)_R \right\|_{1-} \\
& \lesssim \int_{R^2}^\infty dt_0 \int_0^{t_0} dt_1 \left\| I(R \geq r_*) \left( \int_0^{t_1} \eta_R \left| \int_0^{t_1} dt_2 \nabla S(t_0) \bar{S}(t_1) \times a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \stackrel{(D11)}{\lesssim} \int_{R^2}^\infty dt_0 \int_0^{t_0} dt_1 \left\| \left( \int_0^{t_1} \eta_{\sqrt{t_0}} \left| \int_0^{t_1} dt_2 \nabla S(\frac{t_0}{2}) \bar{S}(t_1) \times a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{1-} \\
& \stackrel{(77)}{\lesssim} \int_{R^2}^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{\sqrt{t_0}} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \left( \left\| \left( \int_0^{t_1} \eta_{\sqrt{t_1}} \left| \int_0^{t_1} dt_2 \sqrt{t_1} \bar{S}(t_1) \right. \right. \right. \right. \\
& \quad \times \left. \left. \left. a\nabla S(t_2)ae \right|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{g}\|_{2-} \right)
\end{aligned}$$

$$\begin{aligned}
&\lesssim \int_{R^2}^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{\sqrt{t_0}} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \left( \|(\sqrt{t_1} \nabla \times (S(t_1)ae - ae))_{\sqrt{t_1}}\|_{2-} + \|\bar{g}\|_{2-} \right) \\
&\stackrel{(82),(B3),(80)}{\lesssim} \int_{R^2}^\infty dt_0 \frac{1}{t_0} \int_0^{t_0} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{\sqrt{t_0}} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} R^{1-\frac{d}{2}}.
\end{aligned}$$

We may apply the same modifications for all other cases and obtain the same bound. Hence we only need to consider the case when all  $t_i$ 's are in  $(0, R^2)$ , which will be done now.

For the double integral term of  $\psi$  in (68), notice that since  $aS(t_1)ae$  has locality scale  $1 \vee \sqrt{t_1}$  by Lemma 4.7, we can apply (76) of Proposition 4.4 with  $aS(t_1)ae$  playing the role of  $g$  and estimate

$$\begin{aligned}
&\left\| \left( \int_0^{R^2} dt_0 \int_0^{R^2} dt_1 \nabla S(t_0) aS(t_1) ae \right)_R \right\|_{1-} \\
&\leq \int_0^{R^2} dt_1 \left\| \left( \int_0^{R^2} dt_0 \nabla S(t_0) aS(t_1) ae \right)_R \right\|_{1-} \\
&\stackrel{(76)}{\lesssim} \int_0^{R^2} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{R} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \left( \left\| \left( \int \eta_{\sqrt{t_1}} |\sqrt{t_1} aS(t_1) ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{g}\|_{2-} \right) \\
&\stackrel{(79),(80)}{\lesssim} R^{-\frac{d}{2}} \int_0^{R^2} dt_1 (1 \vee \sqrt{t_1})^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \lesssim R^{1-\frac{d}{2}}.
\end{aligned}$$

For triple integral in (68) we divide the integral into the two ranges  $t_1 > t_2$  and  $t_1 < t_2$ . In the regime  $t_1 > t_2$ , we appeal to (76) of Proposition 4.4 with  $g = (\nabla \times (S(t_1)ae - ae))_{\sqrt{t_1}}$  which has locality scale  $1 \vee \sqrt{t_1}$  (which follows from a combination of Proposition 4.6 and Lemma 4.7). In the regime  $t_1 < t_2$ , we also apply (76), this time with  $g = \bar{S}(t_1) \times a\nabla S(t_2)ae$ , which has locality scale  $1 \vee \sqrt{t_2}$  (again using Proposition 4.6 and Lemma 4.7).

$$\begin{aligned}
&\left\| \left( \int_0^{R^2} dt_0 \int_0^{R^2} dt_1 \int_0^{R^2} dt_2 \nabla S(t_0) \bar{S}(t_1) \times a\nabla S(t_2) ae \right)_R \right\|_{1-} \\
&\stackrel{(62)}{\leq} \int_0^{R^2} dt_1 \left\| \left( \int_0^{R^2} dt_0 \nabla S(t_0) (\nabla \times (S(t_1)ae - ae))_{\sqrt{t_1}} \right)_R \right\|_{1-} \\
&\quad + \int_0^{R^2} dt_1 \int_{t_1}^{R^2} dt_2 \left\| \left( \int_0^{R^2} dt_0 \nabla S(t_0) \bar{S}(t_1) \times a\nabla S(t_2) ae \right)_R \right\|_{1-} \\
&\stackrel{(76)}{\lesssim} \int_0^{R^2} dt_1 \left( \frac{1 \vee \sqrt{t_1}}{R} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \left( \left\| \sqrt{t_1} (\nabla \times (S(t_1)ae - ae))_{\sqrt{t_1}} \right\|_{2-} + \|\bar{g}\|_{2-} \right) \\
&\quad + \int_0^{R^2} dt_1 \int_{t_1}^{R^2} dt_2 \left( \frac{1 \vee \sqrt{t_2}}{R} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} \left( \left\| \left( \int \eta_{\sqrt{t_2}} |\sqrt{t_1} \bar{S}(t_1) \right. \right. \right. \\
&\quad \left. \left. \left. \times a\nabla S(t_2) ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{G}\|_{2-} \right) \\
&\stackrel{(82),(B3),(80)}{\lesssim} R^{-\frac{d}{2}} \int_0^{R^2} (1 \vee \sqrt{t_1})^{\frac{d}{2}} \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} dt_1 \\
&\quad + R^{-\frac{d}{2}} \int_0^{R^2} dt_1 \frac{1}{\sqrt{t_1}} \int_{t_1}^{R^2} dt_2 (1 \vee \sqrt{t_2})^{\frac{d}{2}} \\
&\quad \times \frac{1}{t_2} \left( \left\| \left( \int \eta_{\sqrt{t_2}} |t_2 \nabla S(t_2) ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{g}\|_{2-} \right)
\end{aligned}$$

$$\stackrel{(79)}{\lesssim} R^{1-\frac{d}{2}} + R^{-\frac{d}{2}} \int_0^{R^2} dt_1 \frac{1}{\sqrt{t_1}} \int_{t_1}^{R^2} dt_2 (1 \vee \sqrt{t_2})^{\frac{d}{2}} \frac{1}{t_2} (1 \wedge \frac{1}{\sqrt{t_2}})^{\frac{d}{2}} \lesssim R^{1-\frac{d}{2}}.$$

Using almost identical arguments we can also establish

$$\|I(R \geq r_*)(a\nabla\psi - \langle a\nabla\psi \rangle)_R\|_{1-} \lesssim R^{1-\frac{d}{2}}, \quad (121)$$

the only difference is that whenever we use (76), we use the result for  $\mathcal{S}(t)g - \langle \mathcal{S}(t)g \rangle$ . We thus do not write the argument here.

*Proof of (58) and (59) for  $\nabla\Psi$ .* The proof is based on the representation (70). For (58) we can estimate

$$\begin{aligned} & \|I(R \geq r_*)(\int \eta_R |\nabla\Psi|^2)^{\frac{1}{2}}\|_{1-} \\ & \leq \|I(R \geq r_*)(\int \eta_R |\int_0^{R^2} dt \nabla\bar{S}(t) \times (a\nabla\psi, a\phi, \sigma)|^2)^{\frac{1}{2}}\|_{1-} \\ & \quad + \int_{R^2}^\infty dt \|I(R \geq r_*)(\int \eta_R |\nabla(\nabla \times (a\nabla\psi, a\phi, \sigma))_{\sqrt{t}}|^2)^{\frac{1}{2}}\|_{1-} \\ & \stackrel{(B3)}{\lesssim} \|I(R \geq r_*)(\int \eta_R |(a\nabla\psi, a\phi, \sigma)|^2)^{\frac{1}{2}}\|_{1-} \\ & \quad + \int_{R^2}^\infty dt \|I(R \geq r_*)\nabla(\nabla \times (a\nabla\psi, a\phi, \sigma))_{\sqrt{t}}\|_{1-} \\ & \stackrel{(58),(121),(86),(84),(85)}{\lesssim} 1 + \int_{R^2}^\infty \sqrt{t}^{-2+1-\frac{d}{2}} dt \stackrel{d \geq 3}{\lesssim} 1. \end{aligned}$$

The strategy for (59) is similar,

$$\begin{aligned} \|I(R \geq r_*)\nabla\Psi_R\|_{1-} & \lesssim \int_0^\infty dt \|I(R \geq r_*)\nabla(\nabla \times (a\nabla\psi, a\phi, \sigma))_{\sqrt{R^2+t}}\|_{1-} \\ & \stackrel{(85),(84),(121)}{\lesssim} \int_0^\infty dt \sqrt{R^2+t}^{-2+1-\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} R^{1-\frac{d}{2}}. \end{aligned}$$

□

### 5.8. Proof of Proposition 2.9

*Step 1:* proof of (56). In fact, in this step we prove a more general statement: for any  $\frac{1}{2} \leq r' < r$  to be specified later, if  $F_T$  satisfies the whole-space equation

$$\frac{1}{T}F_T - \nabla \cdot a\nabla F_T = \nabla \cdot g,$$

and  $F_T^{(L)}$  satisfies the following equation in  $Q_{rL}$ :

$$\begin{cases} \frac{1}{T}F_T^{(L)} - \nabla \cdot a\nabla F_T^{(L)} = \nabla \cdot g & \text{in } Q_{rL}, \\ F_T^{(L)} = 0 & \text{on } \partial Q_{rL}, \end{cases} \quad (122)$$

then for any  $p < \infty$ ,

$$\left( \int_{Q_{r'L}} |(\sqrt{T}\nabla(F_T - F_T^{(L)}), F_T - F_T^{(L)})|^2 \right)^{\frac{1}{2}} \lesssim_{r,r',p} \left( \frac{\sqrt{T}}{L} \right)^p \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}}. \quad (123)$$

Choosing  $g = ae$  so that  $F_T = \phi_T$ ,  $F_T^{(L)} = \phi_T^{(L)}$ , then pick  $r = 2$ ,  $r' = \frac{7}{4}$  and (123) turns into (56).

To prove (123), we define a solution operator  $S^{(L)}$  for finite domain: given  $r > 0$ , let  $v(x, t)$  be the solution for

$$\begin{cases} \partial_t v - \nabla \cdot a \nabla v = 0 & \text{in } Q_{rL} \times (0, \infty), \\ v = 0 & \text{on } \partial Q_{rL} \times (0, \infty), \\ v(t = 0) = \nabla \cdot g & \text{in } Q_{rL}, \end{cases} \quad (124)$$

and we define

$$S^{(L)}(t)g := v(t).$$

Following the discussions which leads to (71),

$$F_T^{(L)} = \int_0^\infty \exp\left(-\frac{t}{T}\right) S^{(L)}(t)g \, dt.$$

Following the steps of [3, Lemma 1], we can show

$$\left( \int_{Q_{rL}} |(t \nabla S^{(L)}(t)g, \sqrt{t} S^{(L)}(t)g)|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{Q_{rL}} |g|^2 \right)^{\frac{1}{2}}. \quad (125)$$

Now we introduce the intermediate length scale  $\tilde{T} = \sqrt{TL}$ . The idea is to divide the  $t$ -interval into  $[0, \tilde{T}]$  and  $[\tilde{T}, \infty)$ . In  $[0, \tilde{T}]$ ,  $\exp(-\frac{t}{T}) \lesssim 1$  and we use Lemma B.3 for small  $t$ , while in the large  $t$  regime we directly use the small factor  $\exp(-\frac{t}{T})$  to derive subalgebraic bound:

$$\begin{aligned} & \left( \int_{Q_{r'L}} |\nabla(F_T - F_T^{(L)})|^2 \right)^{\frac{1}{2}} \\ & \stackrel{(71), (122)}{=} \left( \int_{Q_{r'L}} \left| \int_0^\infty dt \exp\left(-\frac{t}{T}\right) \nabla(S(t) - S^{(L)}(t))g \right|^2 \right)^{\frac{1}{2}} \\ & \leq \left( \int_{Q_{r'L}} \left| \int_0^{\tilde{T}} dt \exp\left(-\frac{t}{T}\right) \nabla(S(t) - S^{(L)}(t))g \right|^2 \right)^{\frac{1}{2}} \\ & \quad + \left( \int_{Q_{r'L}} \left| \int_{\tilde{T}}^\infty dt \exp\left(-\frac{t}{T}\right) \nabla(S(t) - S^{(L)}(t))g \right|^2 \right)^{\frac{1}{2}} \\ & \leq \left( \int_0^{\tilde{T}} dt \exp\left(-\frac{2t}{T}\right) \right)^{\frac{1}{2}} \left( \int_{Q_{r'L}} \int_0^{\tilde{T}} dt |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\ & \quad + \int_{\tilde{T}}^\infty dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\ & \lesssim \sqrt{T} \left( \int_{Q_{r'L}} \int_0^{\tilde{T}} dt |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\ & \quad + \int_{\tilde{T}}^\infty dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

For the first term, note that  $(S(t) - S^{(L)}(t))g$  satisfies

$$\begin{cases} \partial_t(S(t) - S^{(L)}(t))g - \nabla \cdot a \nabla(S(t) - S^{(L)}(t))g = 0 & \text{in } Q_{rL} \times (0, \infty), \\ (S(0) - S^{(L)}(0))g = 0 & \text{in } Q_{rL}. \end{cases} \quad (126)$$

Thus, by [Lemma B.2](#) and a slight modification of [Lemma B.3](#) with  $\tilde{T}$  playing the role of  $T$  and  $rL, r'L$  replacing  $2L, L$ , for any  $p < \infty$ ,

$$\begin{aligned}
 & \sqrt{T} \left( \int_{Q_{r'L}} \int_0^{\tilde{T}} dt |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\
 & \stackrel{(B4)}{\lesssim} \frac{\sqrt{T}}{\tilde{T}} \left( \frac{\sqrt{\tilde{T}}}{L} \right)^{2p-1} \int_0^{\tilde{T}} dt \left( \frac{\sqrt{t}}{\sqrt{\tilde{T}}} \right)^p \left( \int_{Q_{rL} \setminus Q_{r'L}} |(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\
 & \lesssim \frac{1}{\sqrt{\tilde{T}}} \left( \frac{\sqrt{T}}{L} \right)^p \int_0^{\tilde{T}} dt \left( \frac{\sqrt{t}}{\sqrt{\tilde{T}}} \right)^p \left( \int_{Q_{rL}} |S(t)g|^2 \right)^{\frac{1}{2}} + \left( \int_{Q_{rL}} |S^{(L)}(t)g|^2 \right)^{\frac{1}{2}} \\
 & \stackrel{(B3), (125)}{\lesssim} \frac{1}{\sqrt{\tilde{T}}} \left( \frac{\sqrt{T}}{L} \right)^p \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}} \int_0^{\tilde{T}} dt \frac{\sqrt{t}^{p-1}}{\sqrt{\tilde{T}}^p} \lesssim \left( \frac{\sqrt{T}}{L} \right)^p \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}}.
 \end{aligned}$$

For the second term

$$\int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}},$$

since  $\exp(-\frac{t}{T}) \leq \exp(-\frac{\tilde{T}}{T}) \lesssim_p \left(\frac{\sqrt{\tilde{T}}}{L}\right)^p$  decays faster than any algebraic power of  $\frac{\sqrt{\tilde{T}}}{L}$ , we only need to show that  $\left( \int_{Q_{r'L}} |\nabla(S(t) - S^{(L)}(t))g|^2 dt \right)^{\frac{1}{2}}$  grows at most algebraically fast in  $\frac{L}{\sqrt{\tilde{T}}}$ . This can be easily achieved by energy estimates:

$$\begin{aligned}
 & \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla(S(t) - S^{(L)}(t))g|^2 \right)^{\frac{1}{2}} \\
 & \lesssim \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla S(t)g|^2 dt \right)^{\frac{1}{2}} + \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) \left( \int_{Q_{r'L}} |\nabla S^{(L)}(t)g|^2 dt \right)^{\frac{1}{2}} \\
 & \stackrel{(125)}{\lesssim} L^{-\frac{d}{2}} \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) \sqrt{t}^{\frac{d}{2}} \left( \int_{Q_{\sqrt{t}}} |\nabla S(t)g|^2 dt \right)^{\frac{1}{2}} + \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) t^{-1} \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}} \\
 & \stackrel{(B3)}{\lesssim} \left( L^{-\frac{d}{2}} \int_{\tilde{T}}^{\infty} dt \exp\left(-\frac{t}{T}\right) t^{\frac{d}{4}-1} + \frac{T}{\tilde{T}} \exp\left(-\frac{\tilde{T}}{T}\right) \right) \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}} \lesssim \left( \frac{\sqrt{T}}{L} \right)^p \left( \int \eta_{rL} |g|^2 \right)^{\frac{1}{2}}.
 \end{aligned}$$

The proof for  $F_T - F_T^{(L)}$  without gradient is identical, and now we finish the proof of [\(123\)](#).

*Step 2:* proof of [\(57\)](#). As  $\psi$ 's are defined through  $\sigma$ 's, we need to prove the intermediate estimate

$$\left( \int_{Q_{\frac{3}{2}L}} |(\sqrt{T}\nabla(\sigma_T - \sigma_T^{(L)}), \sigma_T - \sigma_T^{(L)})|^2 \right)^{\frac{1}{2}} \lesssim_p \left( \frac{\sqrt{T}}{L} \right)^p. \quad (127)$$

Let  $\tilde{\sigma}_T$  be the solution of

$$\begin{cases} \frac{1}{T} \tilde{\sigma}_T - \Delta \tilde{\sigma}_T = \nabla \cdot (q_{k,T} e_j - q_{j,T} e_k) & \text{in } Q_{\frac{7}{4}L}, \\ \tilde{\sigma}_T = 0 & \text{on } \partial Q_{\frac{7}{4}L}. \end{cases} \quad (128)$$

Now substitute  $g$  with  $q_{k,T} e_j - q_{j,T} e_k$  and pick  $r = \frac{7}{4}, r' = \frac{3}{2}$  in [\(123\)](#), we obtain

$$\left( \int_{Q_{\frac{3}{2}L}} |(\sqrt{T}\nabla(\sigma_T - \tilde{\sigma}_T), \sigma_T - \tilde{\sigma}_T)|^2 \right)^{\frac{1}{2}} \stackrel{(123)}{\lesssim} \left( \frac{\sqrt{T}}{L} \right)^p \left( \int \eta_{\frac{7}{4}L} |q_T|^2 \right)^{\frac{1}{2}} \stackrel{(B1)}{\lesssim} \left( \frac{\sqrt{T}}{L} \right)^p.$$



To finish the proof of (127) we need to pass from  $\tilde{\sigma}_T$  to  $\sigma_T^{(L)}$ . We subtract (128) by (10), and appeal to the standard elliptic energy estimate

$$\begin{aligned} \left( \int_{Q_{\frac{3}{2}L}} |(\sqrt{T}\nabla(\sigma_T^{(L)} - \tilde{\sigma}_T), \sigma_T^{(L)} - \tilde{\sigma}_T)|^2 \right)^{\frac{1}{2}} &\lesssim \sqrt{T} \left( \int_{Q_{\frac{7}{4}L}} |q_T - q_T^{(L)}|^2 \right)^{\frac{1}{2}} \\ &\lesssim \sqrt{T} \left( \int_{Q_{\frac{7}{4}L}} |\nabla(\phi_T - \phi_T^{(L)})|^2 \right)^{\frac{1}{2}} \stackrel{(56)}{\lesssim} \left( \frac{\sqrt{T}}{L} \right)^p. \end{aligned}$$

The proof of (57) now follows from an identical argument as that of (127): pick  $r = \frac{3}{2}$  and  $r' = \frac{5}{4}$  in (123) and we have

$$\begin{aligned} &\left( \int_{Q_{\frac{5}{4}L}} |(\sqrt{T}\nabla(\psi_T - \psi_T^{(L)}), \psi_T - \psi_T^{(L)})|^2 \right)^{\frac{1}{2}} \\ &\lesssim \left( \frac{\sqrt{T}}{L} \right)^p \left( \int_{Q_{\frac{3}{2}L}} (\phi_T^2 + |\sigma_T|^2) \right)^{\frac{1}{2}} + \sqrt{T} \left( \int_{Q_{\frac{3}{2}L}} ((\phi_T - \phi_T^{(L)})^2 + |\sigma_T - \sigma_T^{(L)}|^2) \right)^{\frac{1}{2}} \\ &\stackrel{(B1), (56), (127)}{\lesssim} \sqrt{T} \left( \frac{\sqrt{T}}{L} \right)^p. \end{aligned}$$

□

### 5.9. Proof of Proposition 4.4

We only include the proof for  $\mathcal{S}(T)g$  since the proof for  $\int_0^T dt \nabla S(t)g$  is identical. The proof of (76) is divided into four steps:

*Step 1:* Proof of (76) when  $\sqrt{T} \leq r_0$ . When  $r \leq r_0$ , using the fact that Gaussian is dominated by exponential averaging functions, we have

$$\begin{aligned} r^{\frac{d}{2}} \|(\mathcal{S}(T)g - \langle \mathcal{S}(T)g \rangle)_r\|_s &\lesssim r^{\frac{d}{2}} \left\| \left( \int G_r |\mathcal{S}(T)g|^2 \right)^{\frac{1}{2}} \right\|_s \\ &\lesssim r_0^{\frac{d}{2}} \left\| \left( \int \eta_{r_0} |\mathcal{S}(T)g|^2 \right)^{\frac{1}{2}} \right\|_s \stackrel{(B3)}{\lesssim} r_0^{\frac{d}{2}} \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s. \end{aligned} \quad (129)$$

When  $r \geq r_0$ , in view of Proposition 4.6 (in this regime the bounds on  $\bar{G}$  only uses Lemma B.2), we may apply Lemma C.1 to  $F = \mathcal{S}(T)g - \langle \mathcal{S}(T)g \rangle$  and derive with the help of (129)

$$\sup_{r \geq r_0} r^{\frac{d}{2}} \|(\mathcal{S}(T)g - \langle \mathcal{S}(T)g \rangle)_r\|_s \lesssim r_0^{\frac{d}{2}} \left( \left\| \left( \int \eta_{r_0} |g|^2 \right)^{\frac{1}{2}} \right\|_s + \|\bar{g}\|_s \right). \quad (130)$$

*Step 2:* We claim that there exists a constant  $C = C(\lambda, d)$  such that for any  $\delta \ll 1$ ,

$$F_0 := \frac{1}{\sqrt{T}} \left( \int \eta_{\sqrt{T}} |(\phi, \sigma)_{\frac{\delta\sqrt{T}}{C}}|^2 \right)^{\frac{1}{2}} \leq \delta \quad (131)$$

implies

$$\frac{1}{\sqrt{T}} \left( \int \eta_{\sqrt{T}} |(\phi, \sigma)|^2 \right)^{\frac{1}{2}} \leq 2\delta.$$

Here  $\phi_{\frac{\delta\sqrt{T}}{C}}$  denotes the convolution of  $\phi$  and Gaussian  $G_{\frac{\delta\sqrt{T}}{C}}$  (instead of the massive corrector).

We first establish the simple Caccioppoli-type estimate

$$(\int \eta_{\sqrt{T}} |\nabla(\phi, \sigma)|^2)^{\frac{1}{2}} \lesssim 1 + \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} |(\phi, \sigma)|^2)^{\frac{1}{2}}. \quad (132)$$

We test (17) with  $\eta_{\sqrt{T}}\phi$  and obtain

$$\int \nabla(\eta_{\sqrt{T}}\phi) \cdot a \nabla \phi = - \int \nabla(\eta_{\sqrt{T}}\phi) \cdot a e = - \int \phi \nabla \eta_{\sqrt{T}} \cdot a e - \int \eta_{\sqrt{T}} \nabla \phi \cdot a e.$$

We then use  $|\nabla \eta_{\sqrt{T}}| \leq \frac{1}{\sqrt{T}} \eta_{\sqrt{T}}$ , ellipticity (1) and Young's inequality to obtain for any  $\varepsilon \in (0, 1)$ ,

$$\begin{aligned} \int \eta_{\sqrt{T}} |\nabla \phi|^2 &\lesssim \int \eta_{\sqrt{T}} \nabla \phi \cdot a \nabla \phi = \int \nabla(\eta_{\sqrt{T}}\phi) \cdot a \nabla \phi - \int \phi \nabla \eta_{\sqrt{T}} \cdot a \nabla \phi \\ &\lesssim \int |\phi| |\nabla \eta_{\sqrt{T}}| + \int \eta_{\sqrt{T}} |\nabla \phi| + \int |\phi| |\nabla \eta_{\sqrt{T}}| |\nabla \phi| \\ &\lesssim \frac{1}{\sqrt{T}} \int \eta_{\sqrt{T}} |\phi| + \int \eta_{\sqrt{T}} |\nabla \phi| + \frac{1}{\sqrt{T}} \int \eta_{\sqrt{T}} |\phi| |\nabla \phi| \\ &\lesssim \varepsilon \int \eta_{\sqrt{T}} |\nabla \phi|^2 + \frac{1}{\varepsilon} (1 + \frac{1}{T} \int \eta_{\sqrt{T}} \phi^2), \end{aligned}$$

and we finish the proof of (132) for  $\phi$  by choosing a small  $\varepsilon > 0$  to absorb the first r.h.s. term into l.h.s. Similarly for  $\sigma$  we obtain from (21)

$$\begin{aligned} \int \eta_{\sqrt{T}} |\nabla \sigma|^2 &\lesssim \varepsilon \int \eta_{\sqrt{T}} |\nabla \sigma|^2 + \frac{1}{\varepsilon} (1 + \int \eta_{\sqrt{T}} |\nabla \phi|^2 + \frac{1}{T} \int \eta_{\sqrt{T}} \sigma^2) \\ &\lesssim \varepsilon \int \eta_{\sqrt{T}} |\nabla \sigma|^2 + \frac{1}{\varepsilon} (1 + \frac{1}{T} \int \eta_{\sqrt{T}} \phi^2 + \frac{1}{T} \int \eta_{\sqrt{T}} \sigma^2), \end{aligned}$$

and we therefore finish the proof of (132) after absorption.

By Poincaré inequality in convolution (92), we have for some constant  $C$  which may change from line to line,

$$\begin{aligned} \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} \phi^2)^{\frac{1}{2}} &\leq \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} |\phi - \phi_{\frac{\delta\sqrt{T}}{C}}|^2)^{\frac{1}{2}} + \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} |\phi_{\frac{\delta\sqrt{T}}{C}}|^2)^{\frac{1}{2}} \\ &\stackrel{(131)}{\leq} \frac{\delta}{C} (\int \eta_{\sqrt{T}} |\nabla \phi|^2)^{\frac{1}{2}} + \delta \\ &\stackrel{(132)}{\leq} \frac{\delta}{C} + \frac{\delta}{C\sqrt{T}} (\int \eta_{\sqrt{T}} \phi^2)^{\frac{1}{2}} + \delta. \end{aligned}$$

The term  $\frac{\delta}{C\sqrt{T}} (\int \eta_{\sqrt{T}} \phi^2)^{\frac{1}{2}}$  can be absorbed into the l.h.s., and for  $C$  sufficiently large, we obtain

$$\frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} \phi^2)^{\frac{1}{2}} \leq 2\delta.$$

Similarly,

$$\begin{aligned} \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} \sigma^2)^{\frac{1}{2}} &\leq \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} |\sigma - \sigma_{\frac{\delta\sqrt{T}}{C}}|^2)^{\frac{1}{2}} + \frac{1}{\sqrt{T}} (\int \eta_{\sqrt{T}} |\sigma_{\frac{\delta\sqrt{T}}{C}}|^2)^{\frac{1}{2}} \\ &\stackrel{(131)}{\leq} \frac{\delta}{C} (\int \eta_{\sqrt{T}} |\nabla \sigma|^2)^{\frac{1}{2}} + \delta \end{aligned}$$

$$\begin{aligned}
&\stackrel{(132)}{\leq} \frac{\delta}{C} \left(1 + \frac{1}{\sqrt{T}} \left(\int \eta_{\sqrt{T}} \phi^2\right)^{\frac{1}{2}} + \frac{1}{\sqrt{T}} \left(\int \eta_{\sqrt{T}} \sigma^2\right)^{\frac{1}{2}}\right) + \delta \\
&\stackrel{(131)}{\leq} \frac{\delta}{C} (1 + 2\delta) + \frac{\delta}{C\sqrt{T}} \left(\int \eta_{\sqrt{T}} \sigma^2\right)^{\frac{1}{2}} + \delta,
\end{aligned}$$

and we finish the proof for Step 2 after absorbing  $\frac{\delta}{C\sqrt{T}} \left(\int \eta_{\sqrt{T}} \sigma^2\right)^{\frac{1}{2}}$  into l.h.s. and using  $\delta \ll 1$ .

*Step 3:* Estimation of  $\left\| \left( \mathcal{S}(T) - \mathcal{S}^h\left(\frac{T}{2}\right) \mathcal{S}\left(\frac{T}{2}\right) \right) g \right\|_{\frac{2s}{s+2}-}$ . Here and for the rest of the paper we use  $\|\cdot\|$  to denote a norm on stationary random fields with a CLT-scaling built in:

$$\|f\|_s := \sup_{R \geq 1} R^{\frac{d}{2}} \|(f - \langle f \rangle)_R\|_s,$$

The strategy is to first estimate  $R^{\frac{d}{2}} \|(\mathcal{S}(T) - \mathcal{S}^h(\frac{T}{2}) \mathcal{S}(\frac{T}{2}))g\|_{\frac{2s}{s+2}-}$  for  $R \leq \sqrt{T}$  and then use [Lemma C.1](#) for larger  $R$ .

Following the proof of Lemmas 4–6 and Proposition 1 in [3], with  $(\phi, \sigma, a_h)$  playing the role of  $(\phi_T, \sigma_T, a_{hT})$  (and thus the  $g$  there vanishes), we obtain the following deterministic estimate of the homogenization error on the level of the flux semigroup  $\mathcal{S}(t)$  in a weak topology, i.e., there exists some possibly large  $p = p(\lambda, d)$  such that for all  $R \leq \sqrt{T}$ ,

$$\begin{aligned}
|F| &:= \left(\frac{R}{\sqrt{T}}\right)^{\frac{d}{2}} \left| \left( \mathcal{S}(T)g - \mathcal{S}^h\left(\frac{T}{2}\right) \mathcal{S}\left(\frac{T}{2}\right)g \right)_R \right| \\
&\lesssim \delta^{\frac{1}{p}} \int_{\frac{T}{4}}^{\frac{T}{2}} dt \int_0^{\sqrt{t}} dr \left(\frac{r}{\sqrt{t}}\right)^{\frac{d}{2}} \int \eta_{\sqrt{T}} |(\mathcal{S}(t)g - \langle \mathcal{S}(t)g \rangle)_r| =: \delta^{\frac{1}{p}} F_1,
\end{aligned} \tag{133}$$

provided

$$\frac{1}{\sqrt{T}} \left( \int \eta_{\sqrt{T}} |(\phi, \sigma)|^2 \right)^{\frac{1}{2}} \leq 2\delta. \tag{134}$$

The key feature of (133) is that the r.h.s.  $F_1$  is of the same nature as  $F$ : it is a weak norm of the flux.

Using Step 2, we may replace (134) by  $F_0 \leq 2\delta$  with the random variable  $F_0$  defined in (131) that has the desired cancellation bound  $\sqrt{T}^{-\frac{d}{2}}$ , as we may derive from (84),

$$\|F_0\|_{2-} \lesssim \delta (\delta \sqrt{T})^{-\frac{d}{2}}. \tag{135}$$

We also need the following pointwise bound on  $F$ :

$$\left(\frac{R}{\sqrt{T}}\right)^{\frac{d}{2}} \left| \left( \mathcal{S}(T) - \mathcal{S}^h\left(\frac{T}{2}\right) \mathcal{S}\left(\frac{T}{2}\right) \right) g \right|_R \lesssim \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}}. \tag{136}$$

The proof is identical to [3, Lemma 15, (3)], with the only difference that  $ae$  is replaced by  $(\int \eta_{\sqrt{T}} |g|^2)^{\frac{1}{2}}$ , and is thus omitted.

For any  $s_0 < \frac{2s}{s+2}$ , let  $s_1 < 2$  be such that  $\frac{1}{s_0} = \frac{1}{s_1} + \frac{1}{s}$ , by [Lemma A.3](#) we have for  $R \leq \sqrt{T}$  and  $\tilde{s}_1 \in (s_1, 2)$ , with  $2\delta$  playing the role of  $\delta$ ,

$$\begin{aligned}
\|F\|_{s_0} &\lesssim \delta^{\frac{1}{p}} \|F_1\|_{s_0} + \left(\frac{\|F_0\|_{\tilde{s}_1}}{\delta}\right)^{\frac{\tilde{s}_1}{s_1}} \|F\|_s \\
&\stackrel{(135), (136)}{\lesssim} \delta^{\frac{1}{p}} \sqrt{T}^{-\frac{d}{2}} \sup_{t \leq T} \|\mathcal{S}(t)g\|_{s_0} + (\delta \sqrt{T})^{-\frac{d}{2} \frac{\tilde{s}_1}{s_1}} \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s.
\end{aligned} \tag{137}$$

Multiply both sides by  $\sqrt{T}^{\frac{d}{2}}$  and we obtain for  $R \leq \sqrt{T}$ ,

$$R^{\frac{d}{2}} \left\| \left( \mathcal{S}(T)g - \mathcal{S}^h\left(\frac{T}{2}\right)\mathcal{S}\left(\frac{T}{2}\right)g \right)_R \right\|_{s_0} \lesssim \delta^{\frac{1}{p}} \sup_{0 \leq t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + (\delta\sqrt{T})^{-\frac{d}{2}(\frac{\tilde{s}_1}{s_1}-1)} \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s.$$

Since  $\tilde{s}_1 > s_1$ , by optimizing in  $\delta$ , it is clear that there exists an exponent  $\gamma > 0$  such that for all  $R \leq \sqrt{T}$ ,

$$R^{\frac{d}{2}} \left\| \left( \mathcal{S}(T)g - \mathcal{S}^h\left(\frac{T}{2}\right)\mathcal{S}\left(\frac{T}{2}\right)g \right)_R \right\|_{s_0} \lesssim \left( \frac{1}{\sqrt{T}} \right)^{\gamma} \left( \sup_{t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s \right).$$

In order to deal with the range  $R \geq \sqrt{T}$ , we appeal to [Lemma C.1](#) with  $F = (\mathcal{S}(T) - \mathcal{S}^h(\frac{T}{2})\mathcal{S}(\frac{T}{2}))g$  minus its expectation, and  $\bar{F}$  defined as in (81) of [Lemma 4.8](#), such that  $\sqrt{T}^{\frac{d}{2}} \|\bar{F}\|_{s_0} \lesssim \sup_{t \leq T} \|\mathcal{S}(t)g\|_{s_0}$ . Using Young's inequality for the last inequality, we get

$$\begin{aligned} & \left\| \left( \mathcal{S}(T) - \mathcal{S}^h\left(\frac{T}{2}\right)\mathcal{S}\left(\frac{T}{2}\right) \right)g \right\|_{s_0} \\ & \lesssim \left( \frac{1}{\sqrt{T}} \right)^{\gamma} \left( \sup_{t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s \right) \\ & \quad + \left( \left( \frac{1}{\sqrt{T}} \right)^{\gamma} \left( \sup_{t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s \right) \right)^{\frac{1}{2}} \left( \sup_{t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + r_0^{\frac{d}{2}} \|\bar{g}\|_{s_0} \right)^{\frac{1}{2}} \\ & \stackrel{T \geq 1}{\lesssim} \left( \frac{1}{\sqrt{T}} \right)^{\frac{\gamma}{2}} \left( \sup_{t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{T}} |g|^2 \right)^{\frac{1}{2}} \right\|_s + r_0^{\frac{d}{2}} \|\bar{g}\|_s \right). \end{aligned} \quad (138)$$

*Step 4:* Conclusion of the proof by decomposing the time interval into dyadic pieces and using the semigroup property (63) of  $\mathcal{S}(t)$ . The proof follows from [3, Theorem 1]. For any  $t_1 \geq r_0^2$ , using [3, Lemma 16],  $\left\| \mathcal{S}^h(t)g \right\|_s \lesssim \|g\|_s$  for all  $g$  and  $s$ ,

$$\begin{aligned} \left\| \mathcal{S}(T)g \right\|_{s_0} &= \left\| \mathcal{S}^h(T-t_1)\mathcal{S}(t_1)g + \sum_{t_1 < t \leq T} \mathcal{S}^h(T-t)(\mathcal{S}(t) - \mathcal{S}^h(\frac{t}{2})\mathcal{S}(\frac{t}{2}))g \right\|_{s_0} \\ &\leq \left\| \mathcal{S}^h(T-t_1)\mathcal{S}(t_1)g \right\|_{s_0} + \sum_{t_1 < t \leq T} \left\| \mathcal{S}^h(T-t)(\mathcal{S}(t) - \mathcal{S}^h(\frac{t}{2})\mathcal{S}(\frac{t}{2}))g \right\|_{s_0} \\ &\lesssim \left\| \mathcal{S}(t_1)g \right\|_{s_0} + \sum_{t_1 < t \leq T} \left\| \mathcal{S}(t)g - \mathcal{S}^h(\frac{t}{2})\mathcal{S}(\frac{t}{2})g \right\|_{s_0} \\ &\stackrel{(138)}{\lesssim} \left\| \mathcal{S}(t_1)g \right\|_{s_0} + \sum_{t_1 < t \leq T} \left( \frac{1}{\sqrt{t}} \right)^{\frac{\gamma}{2}} \left( \sup_{\tau \leq t} \left\| \mathcal{S}(\tau)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{t}} |g|^2 \right)^{\frac{1}{2}} \right\|_s + r_0^{\frac{d}{2}} \|\bar{g}\|_s \right) \\ &\stackrel{(A7)}{\lesssim} \left\| \mathcal{S}(t_1)g \right\|_{s_0} + \left( \frac{1}{\sqrt{t_1}} \right)^{\frac{\gamma}{2}} \left( \sup_{0 \leq t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0} + \left\| \left( \int \eta_{\sqrt{t_1}} |g|^2 \right)^{\frac{1}{2}} \right\|_s + r_0^{\frac{d}{2}} \|\bar{g}\|_s \right). \end{aligned}$$

Obviously, the estimate still holds if the l.h.s. is replaced by  $\sup_{t_1 \leq t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0}$ . Provided  $t_1 \gg 1$ , we may absorb  $\sup_{t_1 \leq t \leq T} \left\| \mathcal{S}(t)g \right\|_{s_0}$  into the l.h.s. Fixing  $t_1$  to be a large multiple of  $r_0^2 \geq 1$  we

may appeal (76) for  $\sup_{0 \leq t \leq t_1} \|\mathcal{S}(t)g\|_{s_0}$  (which we proved in Step 1, here we used  $s_0 \leq s$ ) and end up with

$$\begin{aligned} \sup_{0 \leq t \leq T} \|\mathcal{S}(t)g\|_{s_0} &\lesssim \sup_{0 \leq t \leq t_1} \|\mathcal{S}(t)g\|_{s_0} + \|(\int \eta_{\sqrt{t_1}} |g|^2)^{\frac{1}{2}}\|_s + r_0^{\frac{d}{2}} \|\bar{g}\|_s \\ &\stackrel{(76), (A7), r_0 \geq 1}{\lesssim} r_0^{\frac{d}{2}} (\|(\int \eta_{r_0} |g|^2)^{\frac{1}{2}}\|_s + \|\bar{g}\|_s). \end{aligned}$$

This finishes the proof of (76). The proof of (77) is now immediate: we use [3, Lemma 6] to obtain

$$\left( \int \eta_{\sqrt{T}} |(\sqrt{T} \nabla S(T)g, S(T)g)|^2 \right)^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{T}} \int_0^T dt \int_0^{\sqrt{t}} dr \left( \frac{r}{\sqrt{t}} \right)^{\frac{d}{2}} \int \eta_{\sqrt{t}} |(S(t)g - \langle S(t)g \rangle_r)|,$$

and then apply the stochastic bound (76).  $\square$

### 5.10. Proof of Proposition 4.6

Our goal is to prove, for any  $a = \tilde{a}$  in  $B_{2R}$  for  $R \geq r_0 \vee \sqrt{T}$ ,

$$\begin{aligned} &\left( \int_{B_R} |(T \nabla (S(T)g - \tilde{S}(T)\tilde{g}), \sqrt{T}(S(T)g - \tilde{S}(T)\tilde{g}), S(T)g - \tilde{S}(T)\tilde{g})|^2 \right)^{\frac{1}{2}} \\ &\lesssim \left( \frac{r_0 \vee \sqrt{T}}{R} \right)^p \int \eta_R (\bar{G} + \tilde{\bar{G}}), \end{aligned}$$

where  $\bar{G}$  is defined through

$$\bar{G}(0) := \frac{1}{\sqrt{T}} \int_0^T dt \left( \frac{\sqrt{t}}{\sqrt{T}} \right)^{\frac{d}{2}} \left( \int \eta_{(r_0 \vee \sqrt{t})} |S(t)g|^2 \right)^{\frac{1}{2}} + (1 \wedge \frac{r_0}{\sqrt{T}})^{\frac{d}{2}} \bar{g}. \quad (139)$$

Here  $\tilde{g}$ ,  $\tilde{S}$ ,  $\tilde{\bar{S}}$ , and  $\tilde{\bar{G}}$  are the quantities corresponding to  $a$  replaced by  $\tilde{a}$ . For simplicity we only present the estimates for  $S(T)g - \tilde{S}(T)\tilde{g}$  since the proofs for the other two quantities are identical.

We decompose  $S(T)g - \tilde{S}(T)\tilde{g}$  as  $S(T)g - \tilde{S}(T)\tilde{g} = w_1(T) + w_2(T)$ , where  $w_1, w_2$  satisfy the following equations:

$$\partial_t w_1 - \nabla \cdot a \nabla w_1 = \nabla \cdot ((\tilde{a} - a) \nabla \tilde{S}(T)\tilde{g}), \quad w_1(t=0) = \nabla \cdot (I(B_{2R}^c)(g - \tilde{g}))$$

and

$$\partial_t w_2 - \nabla \cdot a \nabla w_2 = 0, \quad w_2(t=0) = \nabla \cdot (I(B_{2R})(g - \tilde{g})).$$

Since  $a = \tilde{a}$  in  $B_{2R}$ ,  $w_1$  satisfies the assumptions of Lemma B.3, and hence

$$\left( \int_{B_R} |\sqrt{T} w_1(T)|^2 \right)^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{T}} \left( \frac{\sqrt{T}}{R} \right)^p \int_0^T dt \left( \frac{\sqrt{t}}{\sqrt{T}} \right)^p \left( \int_{B_{2R} \setminus B_R} w_1^2(t) \right)^{\frac{1}{2}}.$$

In order to estimate the r.h.s. we write  $w_1(T) = S(T)g - \tilde{S}(T)\tilde{g} - w_2(T)$  and apply the triangle inequality, the first contribution

$$\frac{1}{\sqrt{T}} \left( \frac{\sqrt{T}}{R} \right)^p \int_0^T dt \left( \frac{\sqrt{t}}{\sqrt{T}} \right)^p \left( \int_{B_{2R} \setminus B_R} |S(t)g|^2 \right)^{\frac{1}{2}}$$

is controlled by the first term of  $\bar{G}$  using [3, (204)]. The second contribution is controlled the same way. For the third term,  $\frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T (\frac{\sqrt{t}}{\sqrt{T}})^p (\int_{B_{2R} \setminus B_R} w_2^2(t))^{\frac{1}{2}} dt$ , we use Lemma B.2:

$$\frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T dt (\frac{\sqrt{t}}{\sqrt{T}})^p (\int_{B_{2R} \setminus B_R} w_2^2(t))^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T dt (\frac{\sqrt{t}}{\sqrt{T}})^p \frac{1}{\sqrt{t}} (\int_{B_{2R}} |g - \tilde{g}|^2)^{\frac{1}{2}}.$$

By assumption (74), the r.h.s.  $\lesssim (\frac{\sqrt{T}}{R})^p (\frac{r_0}{R})^p \int \eta_R(\bar{g} + \tilde{\bar{g}})$ , which is controlled by the second term of  $\bar{G}$ .

We are left with  $(\int_{B_R} |\sqrt{T} w_2(T)|^2)^{\frac{1}{2}}$ . By Lemma B.2 we have

$$(\int_{B_R} |\sqrt{T} w_2(T)|^2)^{\frac{1}{2}} \lesssim (\int_{B_{2R}} |g - \tilde{g}|^2)^{\frac{1}{2}}.$$

Again by locality assumption (74), the above is controlled by the second term of  $\bar{G}$ . The bound (78) is a direct consequence of Proposition 4.4 for  $\sqrt{T} \geq r_0$  and (75) (which only uses Lemma B.2) for  $\sqrt{T} \leq r_0$ .  $\square$

### 5.11. Proof of Lemma 4.8

We first reduce the problem to estimating  $(\int_{B_R} |S(T)g - \tilde{S}(T)\tilde{g}|^2)^{\frac{1}{2}}$  as the difference of the other term  $S^h(\frac{T}{2})(S(\frac{T}{2})g - \tilde{S}(\frac{T}{2})\tilde{g})$  can be estimated similarly. Here we recall that  $\tilde{g}$ ,  $\tilde{S}$  and  $\tilde{S}$  are the quantities correspond to  $\tilde{a}$  in place of  $a$ . Following the proof of Proposition 4.6, we can show for  $a = \tilde{a}$  in  $B_{2R}$  with  $R \geq r_0 \vee \sqrt{T}$ ,

$$\begin{aligned} (\int_{B_R} |S(T)g - \tilde{S}(T)\tilde{g}|^2)^{\frac{1}{2}} &\lesssim (\int_{B_R} |g - \tilde{g}|^2)^{\frac{1}{2}} + (\int_{B_R} |\int_0^T dt (a \nabla S(t)g - \tilde{a} \nabla \tilde{S}(t)\tilde{g})|^2)^{\frac{1}{2}} \\ &\stackrel{(74)}{\lesssim} (\frac{r_0}{R})^p \int \eta_R(\bar{g} + \tilde{\bar{g}}) + (\int_{B_R} |\int_0^T dt \nabla(S(t)g - \tilde{S}(t)\tilde{g})|^2)^{\frac{1}{2}} \\ &\lesssim \frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T dt (\frac{t}{T})^p (\int_{B_{2R} \setminus B_R} |S(t)g|^2)^{\frac{1}{2}} + (\frac{r_0}{R})^p \int \eta_R \bar{g} \\ &\quad + \frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T dt (\frac{t}{T})^p (\int_{B_{2R} \setminus B_R} |\tilde{S}(t)\tilde{g}|^2)^{\frac{1}{2}} + (\frac{r_0}{R})^p \int \eta_R \tilde{\bar{g}}. \end{aligned}$$

To finish the proof we have one more step

$$\frac{1}{\sqrt{T}}(\frac{\sqrt{T}}{R})^p \int_0^T dt (\frac{t}{T})^p (\int_{B_{2R} \setminus B_R} |S(t)g|^2)^{\frac{1}{2}} \lesssim (\frac{\sqrt{T}}{R})^p \int \eta_R \bar{F}(a),$$

which can be shown as in the proof of [3, Lemma 12].  $\square$

### 5.12. Proof of Lemma 4.9

We only need to establish (84), (85), and (86). Throughout the proof of this lemma we use  $\phi_R := \phi * G_R$  to denote the convolution of  $\phi$  and the Gaussian  $G_R$  (not to be confused with  $\phi_T$ ), and similarly  $\sigma_R$ . We first address (84). For the  $\phi$  part, we use (73) in form of

$$\phi = \int_0^\infty dt (-\Delta \phi)_{\sqrt{t}}.$$

Therefore

$$\begin{aligned}
 \|\phi_R\|_{2-} &= \left\| \left( \int_0^\infty dt (-\Delta\phi)_{\sqrt{t}} \right)_R \right\|_{2-} \\
 &\leq \int_0^\infty dt \|(-\Delta\phi)_{\sqrt{t+R^2}}\|_{2-} \leq \int_0^\infty dt \frac{1}{\sqrt{t+R^2}} \|(\nabla\phi)_{\frac{\sqrt{t+R^2}}{2}}\|_{2-} \\
 &\stackrel{(82)}{\lesssim} \int_0^\infty dt \sqrt{t+R^2}^{-1-\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} R^{1-\frac{d}{2}}.
 \end{aligned}$$

For  $\sigma$  part we use decomposition (67) and appeal to the bound (82) on  $q$ :

$$\begin{aligned}
 \|\sigma_R\| &= \left\| \left( \int_0^\infty dt (\nabla \times q)_{\sqrt{t}} \right)_R \right\|_{2-} \leq \int_0^\infty dt \|(\nabla \times q)_{\sqrt{t+R^2}}\|_{2-} \\
 &\stackrel{(82)}{\lesssim} \int_0^\infty dt \sqrt{t+R^2}^{-1-\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} R^{1-\frac{d}{2}}.
 \end{aligned}$$

We now turn to (85). It is slightly different from the above proof of (84) since the starting point is the semigroup decomposition (65):

$$\|I(R \geq r_*)(a\phi - \langle a\phi \rangle)_R\|_{2-} \leq \int_0^\infty dt \|I(R \geq r_*)(aS(t)ae - \langle aS(t)ae \rangle)_R\|_{2-}.$$

For the range  $\sqrt{t} \leq R$ , in view of Lemma 4.7 which ensures that  $aS(t)ae$  is local on scale  $1 \vee \sqrt{t}$ , we use Lemma C.1 on  $aS(t)ae - \langle aS(t)ae \rangle$ , then a Cauchy-Schwarz and the fact that Gaussian kernel is dominated by exponential kernel, and finally Lemma 4.7 to obtain

$$\begin{aligned}
 &\int_0^{R^2} dt \|aS(t)ae - \langle aS(t)ae \rangle\|_{2-} \\
 &\stackrel{(C1)}{\lesssim} \int_0^{R^2} dt \left( \frac{1 \vee \sqrt{t}}{R} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t}} \left( \sup_{r \leq 1 \vee \sqrt{t}} \left( \frac{r}{1 \vee \sqrt{t}} \right)^{\frac{d}{2}} \|\sqrt{t}aS(t)ae\|_{2-} + \|\bar{g}\|_{2-} \right) \\
 &\lesssim \int_0^{R^2} dt \left( \frac{1 \vee \sqrt{t}}{R} \right)^{\frac{d}{2}} \frac{1}{\sqrt{t}} \left( \left\| \left( \int \eta_{(1 \vee \sqrt{t})} |\sqrt{t}aS(t)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} + \|\bar{g}\|_{2-} \right) \\
 &\stackrel{(79), (80)}{\lesssim} R^{-\frac{d}{2}} \int_0^{R^2} dt (1 \vee \sqrt{t})^{\frac{d}{2}} \frac{1}{\sqrt{t}} (1 \wedge \frac{1}{\sqrt{t}})^{\frac{d}{2}} \sim R^{1-\frac{d}{2}}.
 \end{aligned}$$

For the range  $\sqrt{t} \geq R$ , we directly use the exponential kernel to control the Gaussian kernel, then appeal to (D12) and Lemma 4.7:

$$\begin{aligned}
 &\int_{R^2}^\infty dt \|I(R \geq r_*)(aS(t)ae - \langle aS(t)ae \rangle)_R\|_{2-} \\
 &\lesssim \int_{R^2}^\infty dt \|I(R \geq r_*) \left( \int \eta_R |S(t)ae|^2 \right)^{\frac{1}{2}}\|_{2-} \\
 &\stackrel{(D12)}{\lesssim} \int_{R^2}^\infty dt \left\| \left( \int \eta_{\sqrt{t}} |S(\frac{t}{2})ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \\
 &\stackrel{(79)}{\lesssim} \int_{R^2}^\infty dt \sqrt{t}^{-1-\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} R^{1-\frac{d}{2}}.
 \end{aligned}$$

We finally prove (86). The  $\phi$  part uses Poincaré inequality in convolution (92), and (84) for  $r = 1$ ,

$$\begin{aligned} \|(\int \eta_R |\phi|^2)^{\frac{1}{2}}\|_{2-} &\lesssim \|(\int \eta_R |\phi - \phi_1|^2)^{\frac{1}{2}}\|_{2-} + \|(\int \eta_R |\phi_1|^2)^{\frac{1}{2}}\|_{2-} \\ &\lesssim \|(\int \eta_R |\nabla \phi|^2)^{\frac{1}{2}}\|_{2-} + \|\phi_1\|_{2-} \stackrel{(83),(84)}{\lesssim} 1. \end{aligned}$$

For  $\sigma$  part, we again use decomposition (67)

$$\begin{aligned} \|(\int \eta_R |\sigma|^2)^{\frac{1}{2}}\|_{2-} &= \|(\int \eta_R |\int_0^\infty dt (\nabla \times q)_{\sqrt{t}}|^2)^{\frac{1}{2}}\|_{2-} \\ &\lesssim \|(\int \eta_R |\int_0^1 dt \bar{S}(t) \times q|^2)^{\frac{1}{2}}\|_{2-} + \int_1^\infty dt \|(\int_{B_R} |(\nabla \times q)_{\sqrt{t}}|^2)^{\frac{1}{2}}\|_{2-} \\ &\stackrel{(B3)}{\lesssim} \|(\int \eta_R |q|^2)^{\frac{1}{2}}\|_{2-} + \int_1^\infty dt \|(\nabla \times q)_{\sqrt{t}}\|_{2-} \\ &\stackrel{(82),(83)}{\lesssim} 1 + \int_1^\infty dt \sqrt{t}^{-1-\frac{d}{2}} \stackrel{d \geq 3}{\lesssim} 1. \end{aligned}$$

□

### 5.13. Proof of Lemma 4.10

If  $R \geq \sqrt{t_0}$ , the argument is direct:

$$\begin{aligned} &\|I(R \geq r_*)(\int \eta_R |\int_0^{t_0} d\tau \nabla S(\tau) a S(t_1) g|^2)^{\frac{1}{2}}\|_{2-} \\ &\stackrel{(B3)}{\lesssim} \|I(R \geq r_*)(\int \eta_R |S(t_1) g|^2)^{\frac{1}{2}}\|_{2-} \\ &\stackrel{(D12)}{\lesssim} \|I(R \vee \sqrt{t_1} \geq r_*)(\int \eta_{R \vee \sqrt{t_1}} |S(\frac{t_1}{2}) g|^2)^{\frac{1}{2}}\|_{2-}. \end{aligned}$$

We now assume  $\sqrt{t_0} \geq R$ . The starting point is the following: let  $u$  satisfy the equation

$$-\nabla \cdot a(\nabla u + g) = f,$$

then we have

$$\sup_{r \in [r_*, R]} \int \eta_r |\nabla u|^2 \lesssim \int \eta_R |\nabla u|^2 + \sup_{r \in [r_*, R]} \int \eta_r (|g|^2 + R^2 |\nabla g|^2 + R^2 f^2). \quad (140)$$

By a slight modification of [12, Corollary 3] as well as Poincaré inequality we have

$$\begin{aligned} \sup_{r \in [r_*, R]} \int_{B_r} |\nabla u + g|^2 &\lesssim \int_{B_R} |\nabla u + g|^2 + \sup_{r \in [r_*, R]} \frac{R^2}{r^2} \int_{B_r} |g - \int_{B_r} g|^2 + R^2 \sup_{r \in [r_*, R]} \int_{B_r} f^2 \\ &\lesssim \int_{B_R} |\nabla u + g|^2 + R^2 \sup_{r \in [r_*, R]} \int_{B_r} (|\nabla g|^2 + f^2). \end{aligned} \quad (141)$$

The proof of (140) then follows from post-processing from  $\int_{B_r}$  to  $\int \eta_r$  (see the proof of Lemma D.1 for the argument) and a simple triangle inequality. Now, since the function  $w := \int_0^{t_0} d\tau S(\tau) a S(t_1) g$  satisfies the elliptic equation

$$-\nabla \cdot a(\nabla w + S(t_1)g) = -S(t_0)aS(t_1)g,$$



using (140) with  $\sqrt{t_0}$ ,  $S(t_1)g$ ,  $-S(t_0)aS(t_1)g$  playing the roles of  $R, g, f$  respectively, as well as Corollary D.2, we obtain

$$\begin{aligned}
& \|I(R \geq r_*)(\int \eta_R |\int_0^{t_0} d\tau \nabla S(\tau) aS(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \stackrel{(140)}{\lesssim} \|I(\sqrt{t_0} \geq r_*)(\int \eta_{\sqrt{t_0}} |\int_0^{t_0} d\tau \nabla S(\tau) aS(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \|\sup_{r \in [r_*, \sqrt{t_0}]} (\int \eta_r |S(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \sqrt{t_0} \|\sup_{r \in [r_*, \sqrt{t_0}]} (\int \eta_r |\nabla S(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \sqrt{t_0} \|I(\sqrt{t_0} \geq r_*) \sup_{r \in [r_*, \sqrt{t_0}]} (\int \eta_r |S(t_0)aS(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \stackrel{(D9), (D10)}{\lesssim} \|I(\sqrt{t_0} \geq r_*)(\int \eta_{\sqrt{t_0}} |\int_0^{t_0} d\tau \nabla S(\tau) aS(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \sqrt{t_0} \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |\nabla S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \sqrt{t_0} \|I(\sqrt{t_0} \geq r_*)(\int \eta_{\sqrt{t_0}} |S(\frac{t_0}{2})aS(t_1)g|^2)^{\frac{1}{2}}\|_{2-} \\
& \stackrel{(B3)}{\lesssim} \|I(\sqrt{t_0} \geq r_*)(\int \eta_{\sqrt{t_0}} |S(t_1)g|^2)^{\frac{1}{2}}\|_{2-} + \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-} \\
& \quad + \sqrt{t_0} \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |\nabla S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-} \\
& \stackrel{(D12)}{\lesssim} \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-} + \sqrt{t_0} \|I(\sqrt{t_1} \geq r_*)(\int \eta_{\sqrt{t_1}} |\nabla S(\frac{t_1}{2})g|^2)^{\frac{1}{2}}\|_{2-}.
\end{aligned}$$

□

#### 5.14. Proof of Lemma 4.11

We first prove (88). The starting point is to show for  $R \geq 1$

$$\|(ae - \langle ae \rangle)_R\|_{2-} \lesssim R^{-\frac{d}{2}}. \quad (142)$$

The proof uses (C2) for  $ae - \langle ae \rangle$  that is mean-zero and exactly local on scale  $r = 1$ :

$$\|(ae - \langle ae \rangle)_R\|_{2-} \stackrel{(C2)}{\lesssim} (\int G_R^2)^{\frac{1}{2}} \|ae - \langle ae \rangle\|_{2-} \lesssim R^{-\frac{d}{2}}.$$

The next step is to prove for  $t_2 \leq t_1$ ,

$$\|\nabla \times (\mathcal{S}(t_2)ae - ae)_{\sqrt{t_1}}\|_{2-} \lesssim \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}. \quad (143)$$

If  $t_1 \leq 1$  the proof is completely deterministic. Using a Cauchy-Schwarz and that the Gaussian kernel can be dominated by exponential kernel and [Lemma B.2](#) we obtain

$$\begin{aligned} \|\nabla \times (\mathcal{S}(t_2)ae - ae)_{\sqrt{t_1}}\|_{2-} &\lesssim \frac{1}{\sqrt{t_1}} \left\| \left( a \int_0^{t_2} d\tau \nabla \mathcal{S}(\tau)ae \right)_{\sqrt{t_1}/2} \right\|_{2-} \\ &\lesssim \frac{1}{\sqrt{t_1}} \left\| \left( \int \eta_{\frac{\sqrt{t_1}}{2}} |a \int_0^{t_2} d\tau \nabla \mathcal{S}(\tau)ae|^2 \right)^{\frac{1}{2}} \right\|_{2-} \lesssim \frac{1}{\sqrt{t_1}}. \end{aligned}$$

In the case  $t_1 \geq 1$  the proof follows from [\(82\)](#):

$$\|\nabla \times (\mathcal{S}(t_2)ae)_{\sqrt{t_1}}\|_{2-} = \|\nabla \times (\mathcal{S}(t_2)ae - \langle \mathcal{S}(t_2)ae \rangle)_{\sqrt{t_1}}\|_{2-} \stackrel{(82)}{\lesssim} \sqrt{t_1}^{-1-\frac{d}{2}},$$

and we prove [\(143\)](#) using a triangle inequality and [\(142\)](#). We are now ready to prove [\(88\)](#). For simplicity we only prove estimates for the second term as the proof for the other term is identical. Using integration by parts [\(119\)](#),

$$\begin{aligned} &\int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3 + t_2}{T})) \bar{\mathcal{S}}(t_1) \times a \nabla \mathcal{S}(t_2)ae \\ &= (1 - \exp(-\frac{t_3 + t_1}{T})) \int_0^{t_1} dt_2 \bar{\mathcal{S}}(t_1) \times a \nabla \mathcal{S}(t_2)ae \\ &\quad - \frac{1}{T} \exp(-\frac{t_3}{T}) \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \int_0^{t_2} d\tau \bar{\mathcal{S}}(t_1) \times a \nabla \mathcal{S}(\tau)ae \\ &\stackrel{(62),(64)}{=} (1 - \exp(-\frac{t_3 + t_1}{T})) \nabla \times (\mathcal{S}(t_1)ae - ae)_{\sqrt{t_1}} \\ &\quad - \frac{1}{T} \exp(-\frac{t_3}{T}) \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \nabla \times (\mathcal{S}(t_2)ae - ae)_{\sqrt{t_1}}. \end{aligned}$$

Therefore using [\(82\)](#) and  $t_3 \geq t_1$ ,

$$\begin{aligned} &\left\| \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3 + t_2}{T})) \bar{\mathcal{S}}(t_1) \times a \nabla \mathcal{S}(t_2)ae \right\|_{2-} \\ &\leq (1 - \exp(-\frac{t_3 + t_1}{T})) \|\nabla \times (\mathcal{S}(t_1)ae - ae)_{\sqrt{t_1}}\|_{2-} \\ &\quad + \frac{1}{T} \exp(-\frac{t_3}{T}) \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \|\nabla \times (\mathcal{S}(t_2)ae - ae)_{\sqrt{t_1}}\|_{2-} \\ &\stackrel{(143)}{\lesssim} (1 \wedge \frac{t_2}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} + \exp(-\frac{t_3}{T}) (1 \wedge \frac{t_2}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \\ &\lesssim (1 \wedge \frac{t_2}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}. \end{aligned}$$

The proof of [\(89\)](#) is similar, again we only write down the proof for the second term:

$$\begin{aligned} &\left\| \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \bar{\mathcal{S}}(t_1) \times a \nabla \mathcal{S}(t_2)ae \right\|_{2-} \\ &\leq \exp(-\frac{t_1}{T}) \|\nabla \times (\mathcal{S}(t_1)ae - ae)_{\sqrt{t_1}}\|_{2-} \\ &\quad + \frac{1}{T} \int_0^{t_1} dt_2 \exp(-\frac{t_2}{T}) \|\nabla \times (\mathcal{S}(t_2)ae - ae)_{\sqrt{t_1}}\|_{2-} \end{aligned}$$

$$\begin{aligned}
&\stackrel{(143)}{\lesssim} \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} + (1 \wedge \frac{t_2}{T}) \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}} \\
&\lesssim \frac{1}{\sqrt{t_1}} (1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}.
\end{aligned}$$

For the proof of approximate locality (90), in view of Proposition 4.6 applied to  $\tilde{S}(t_1)$ , it suffices to prove that  $\int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3+t_2}{T})) a \nabla S(t_2) a e$  is approximately local on scale  $1 \vee \sqrt{t_1}$  relative to some stationary  $\bar{g}_0$  with  $\|\bar{g}_0\|_{2-} \lesssim (1 \wedge \frac{t_3}{T})(1 \wedge \frac{1}{\sqrt{t_1}})^{\frac{d}{2}}$ . Now suppose  $a = \tilde{a}$  in  $B_{2R}$  for some  $R \geq 1 \vee \sqrt{t_1}$ , notice that  $S(t)ae - \tilde{S}(t)\tilde{a}e$  satisfies the conditions of Lemma B.3, we can estimate

$$\begin{aligned}
&\left( \int_{B_R} \left| \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3+t_2}{T})) \nabla (S(t_2)ae - \tilde{S}(t_2)\tilde{a}e) \right|^2 \right)^{\frac{1}{2}} \\
&\leq \int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3+t_2}{T})) \left( \int_{B_R} |\nabla (S(t_2)ae - \tilde{S}(t_2)\tilde{a}e)|^2 \right)^{\frac{1}{2}} \\
&\stackrel{(B4)}{\lesssim} \frac{1}{R} \int_0^{t_1} dt_2 (1 \wedge \frac{t_3}{T}) (\frac{\sqrt{t_2}}{R})^{p+2} \int_0^{t_2} d\tau (\frac{\sqrt{\tau}}{\sqrt{t_2}})^p \left( \int_{B_{2R} \setminus B_R} (S(\tau)ae - \tilde{S}(\tau)ae)^2 \right)^{\frac{1}{2}} \\
&\lesssim \frac{1}{R^{p+3}} (1 \wedge \frac{t_3}{T}) \int_0^{t_1} d\tau \tau^{\frac{p}{2}} \left( \left( \int_{B_{2R} \setminus B_R} |S(\tau)ae|^2 \right)^{\frac{1}{2}} + \left( \int_{B_{2R} \setminus B_R} |\tilde{S}(\tau)ae|^2 \right)^{\frac{1}{2}} \right) \int_{\tau}^{t_1} dt_2 \\
&\sim \frac{1}{R} (1 \wedge \frac{t_3}{T}) (\frac{\sqrt{t_1}}{R})^2 (\frac{1 \vee \sqrt{t_1}}{R})^p \int_0^{t_1} d\tau (\frac{\sqrt{\tau}}{1 \vee \sqrt{t_1}})^p \left( \left( \int_{B_{2R} \setminus B_R} |S(\tau)ae|^2 \right)^{\frac{1}{2}} \right. \\
&\quad \left. + \left( \int_{B_{2R} \setminus B_R} |\tilde{S}(\tau)ae|^2 \right)^{\frac{1}{2}} \right).
\end{aligned}$$

Now using the fact that  $R \geq 1 \vee \sqrt{t_1}$ , as well as [3, (204)], we derive that  $\int_0^{t_1} dt_2 (1 - \exp(-\frac{t_3+t_2}{T})) a \nabla S(t_2) a e$  is approximately local on scale  $1 \vee \sqrt{t_1}$  relative to some stationary  $\bar{g}_0$  with

$$\bar{g}_0(0) := \frac{1}{1 \vee \sqrt{t_1}} (1 \wedge \frac{t_3}{T}) \int_0^{t_1} d\tau (\frac{\sqrt{\tau}}{1 \vee \sqrt{t_1}})^p \left( \int \eta_{\sqrt{\tau}} |S(\tau)ae|^2 \right)^{\frac{1}{2}},$$

and we finish the proof after applying Lemma 4.7 on  $\bar{g}_0$ .  $\square$

## Acknowledgments

We would like to thank Marius Lemm for his suggestion on references. LW would also like to thank the Max Planck Institute for their generous support, and thank Peter Bella, Nicolas Clozeau, Antoine Gloria, Yu Gu and Jim Nolen for helping with the understanding of stochastic homogenization, and Yingzhou Li for assistance on numerical implementation.

## Funding

The work of JL and LW is supported in part by the National Science Foundation via grants DMS-1454939 and DMS-2012286.

## References

- [1] Bella, P., Giunti, A., Otto, F. (2020). Effective multipoles in random media. *Commun. Partial Differ. Equ.* 45(6):561–640.
- [2] Lu, J., Otto, F. (2021). Optimal artificial boundary condition for random elliptic media. *Found. Comput. Math.* 21:1643–1702.
- [3] Gloria, A., Otto, F. (2015). The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. arXiv preprint arXiv:1510.08290.
- [4] Yurinskii, V. V. (1986). Averaging of symmetric diffusion in random medium. *Siberian Math. J.* 27(4):603–613.
- [5] Naddaf, A., Spencer, T. (1997). On homogenization and scaling limit of some gradient perturbations of a massless free field. *Commun. Math. Phys.* 183(1):55–84.
- [6] Naddaf, A., Spencer, T. (1998). Estimates on the variance of some homogenization problems. preprint.
- [7] Gloria, A., Otto, F. (2011). An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Prob.* 39(3):779–856.
- [8] Gloria, A., Otto, F. (2012). An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Prob.* 22:1–28.
- [9] Gloria, A., Neukamm, S., Otto, F. (2015). Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on glauher dynamics. *Invent. Math.* 199(2):455–515.
- [10] Marahrens, D., Otto, F. (2015). Annealed estimates on the green function. *Probab. Theory Relat. Fields* 163(3–4):527–573.
- [11] Gloria, A., Neukamm, S., Otto, F. (2021). Quantitative estimates in stochastic homogenization for correlated coefficient fields. *Anal. PDE* 14(8):2497–2537.
- [12] Gloria, A., Neukamm, S., Otto, F. (2020). A regularity theory for random elliptic operators. *Milan J. Math.* 88(1):99–170.
- [13] Armstrong, S. N., Smart, C. K. (2016). Quantitative stochastic homogenization of convex integral functionals. *Ann. scientifiques de l'Ecole normale supérieure* 49:423–481.
- [14] Armstrong, S., Kuusi, T., Mourrat, J.-C. (2016). Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Commun. Math. Phys.* 347(2):315–361.
- [15] Armstrong, S., Kuusi, T., Mourrat, J.-C. (2017). The additive structure of elliptic homogenization. *Invent. Math.* 208(3):999–1154.
- [16] Armstrong, S., Kuusi, T., Mourrat, J.-C. (2019). *Quantitative Stochastic Homogenization and Large-Scale Regularity*, Vol. 352. cham: Springer.
- [17] Nolen, J. (2014). Normal approximation for a random elliptic equation. *Probab. Theory Relat. Fields* 159(3–4):661–700.
- [18] Gloria, A., Nolen, J. (2016). A quantitative central limit theorem for the effective conductance on the discrete torus. *Commun. Pure Appl. Math.* 69(12):2304–2348.
- [19] Mourrat, J.-C., Otto, F. (2016). Correlation structure of the corrector in stochastic homogenization. *Ann. Prob.* 44(5):3207–3233.
- [20] Gu, Y., Mourrat, J.-C. (2016). Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.* 14(1):452–481.
- [21] Mourrat, J.-C., Nolen, J. (2017). Scaling limit of the corrector in stochastic homogenization. *Ann. Appl. Prob.* 27(2):944–959.
- [22] Fischer, J., Otto, F. (2016). A higher-order large-scale regularity theory for random elliptic operators. *Commun. Partial Differ. Equ.* 41(7):1108–1148.
- [23] Bella, P., Fehrman, B., Fischer, J., Otto, F. (2017). 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.
- [24] Gu, Y. (2017). High order correctors and two-scale expansions in stochastic homogenization. *Probab. Theory Relat. Fields* 169(3–4):1221–1259.
- [25] Duerinckx, M., Otto, F. (2020). Higher-order pathwise theory of fluctuations in stochastic homogenization. *Stoch. Partial Differ. Equ. Anal. Comput.* 8(3):625–692.

- [26] Duerinckx, M. (2021). Non-perturbative approach to the Bourgain-Spencer conjecture in stochastic homogenization. arXiv preprint arXiv:2102.06319.
- [27] Duerinckx, M., Gloria, A., Lemm, M. (2019). A remark on a surprising result by bourgain in homogenization. *Commun. Partial Differ. Equ.* 44(12):1345–1357.
- [28] Bourgain, J. (2018). On a homogenization problem. *J. Stat. Phys.* 172(2):314–320.
- [29] Kim, J., Lemm, M. (2019). On the averaged Green's function of an elliptic equation with random coefficients. *Arch. Ration. Mech. Anal.* 234(3):1121–1166.
- [30] Anantharaman, A., Costaouec, R., Bris, C. L., Legoll, F., Thomines, F. (2012). Introduction to numerical stochastic homogenization and the related computational challenges: some recent developments. In: *Multiscale Modeling and Analysis for Materials Simulation*. Singapore: World Scientific, pp. 197–272.
- [31] Bourgeat, A., Piatnitski, A. (2004). Approximations of effective coefficients in stochastic homogenization. *Annales de l'IHP Probabilités et statistiques* 40:153–165.
- [32] Gloria, A. (2012). Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *ESAIM: Math. Model. Numer. Anal.* 46(1):1–38.
- [33] Gloria, A., Mourrat, J.-C. (2012). Spectral measure and approximation of homogenized coefficients. *Probab. Theory Relat. Fields* 154(1–2):287–326.
- [34] Egloffé, A.-C., Gloria, A., Mourrat, J.-C., Nguyen, T. N. (2015). Random walk in random environment, corrector equation and homogenized coefficients: from theory to numerics, back and forth. *IMA J. Numer. Anal.* 35(2):499–545.
- [35] Fischer, J. (2019). The choice of representative volumes in the approximation of effective properties of random materials. *Arch. Rational Mech. Anal.* 234(2):635–726.
- [36] Khoromskaia, V., Khoromskij, B. N., Otto, F. (2020). Numerical study in stochastic homogenization for elliptic partial differential equations: Convergence rate in the size of representative volume elements. *Numer. Linear Algebra Appl.* 27(3):e2296.
- [37] Mourrat, J.-C. (2019). Efficient methods for the estimation of homogenized coefficients. *Found. Comput. Math.* 19(2):435–483.
- [38] Hannukainen, A., Mourrat, J.-C., Stoppels, H. T. (2021). Computing homogenized coefficients via multiscale representation and hierarchical hybrid grids. *ESAIM: Math. Model. Numer. Anal.* 55:S149–S185.
- [39] Armstrong, S., Hannukainen, A., Kuusi, T., Mourrat, J.-C. (2021). An iterative method for elliptic problems with rapidly oscillating coefficients. *ESAIM: Math. Model. Numer. Anal.* 55(1):37–55.
- [40] Cances, E., Ehlacher, V., Legoll, F., Stamm, B. (2015). An embedded corrector problem to approximate the homogenized coefficients of an elliptic equation. *Comptes Rendus Math.* 353(9):801–806.
- [41] Le Bris, C., Legoll, F., Minvielle, W. (2016). Special quasirandom structures: a selection approach for stochastic homogenization. *Monte Carlo Methods Appl.* 22(1):25–54.
- [42] Blanc, X., Le Bris, C., Legoll, F. (2016). Some variance reduction methods for numerical stochastic homogenization. *Philos. Trans. Royal Soc. A: Math. Phys. Eng. Sci.* 374(2066):20150168.
- [43] Babuška, I., Osborn, J. E. (1983). Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.* 20(3):510–536.
- [44] Babuška, I., Caloz, G., Osborn, J. E. (1994). Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.* 31(4):945–981.
- [45] Hou, T. Y., Wu, X.-H. (1997). A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.* 134(1):169–189.
- [46] Efendiev, Y., Hou, T. Y. (2009). *Multiscale Finite Element Methods: Theory and Applications*, Vol. 4. New York: Springer.
- [47] Weinan, E., Engquist, B., Huang, Z. (2003). Heterogeneous multiscale method: a general methodology for multiscale modeling. *Phys. Rev. B* 67(9):092101.
- [48] Weinan, E., Ming, P., Zhang, P., et al. (2005). Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.* 18(1):121–156.
- [49] Weinan, E., Engquist, B., Li, X., Ren, W., Vanden-Eijnden, E. (2007). The heterogeneous multiscale method: a review. *Commun. Comput. Phys.* 2(3):367–450.

- [50] Abdulle, A., Weinan, E., Engquist, B., Vanden-Eijnden, E. (2012). The heterogeneous multiscale method. *Acta Numerica* 21:1–87.
- [51] Tinsley Oden, J., Vemaganti, K. S. (2000). Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials: I. error estimates and adaptive algorithms. *J. Comput. Phys.* 164(1):22–47.
- [52] Målqvist, A., Peterseim, D. (2014). Localization of elliptic multiscale problems. *Math. Comput.* 83(290):2583–2603.
- [53] Fischer, J., Gallistl, D., Peterseim, D. (2021). A priori error analysis of a numerical stochastic homogenization method. *SIAM J. Numer. Anal.* 59(2):660–674.
- [54] Owhadi, H. (2015). Bayesian numerical homogenization. *Multiscale Model. Simul.* 13(3):812–828.
- [55] Owhadi, H. (2017). Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Review* 59(1):99–149.
- [56] Kozlov, S. M. (1979). Averaging of random operators. *Matematicheskii Sbornik* 151(2):188–202.
- [57] Papanicolaou, G. C., Varadhan, S. R. S. (1979). Boundary value problems with rapidly oscillating random coefficients. *Colloquia Math. Soc., Janos Bolyai* 27:853–873.
- [58] Jikov, V. V., Kozlov, S. M., Oleinik, O. A. (1994). *Homogenization of Differential Operators and Integral Functionals*. Berlin: Springer.
- [59] Bella, P., Giunti, A., Otto, F. (2015). Quantitative stochastic homogenization: local control of homogenization error through corrector. arXiv preprint arXiv:1504.02487.
- [60] Gloria, A., Otto, F. (2017). Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc.* 19(11):3489–3548.
- [61] Avellaneda, M., Lin, F.-H. (1987). Compactness methods in the theory of homogenization. *Commun. Pure Appl. Math.* 40(6):803–847.
- [62] Shen, Z. (2018). *Periodic Homogenization of Elliptic Systems*, Vol. 269. Cham: Springer.
- [63] Bella, P., Otto, F. (2016). Corrector estimates for elliptic systems with random periodic coefficients. *Multiscale Model. Simul.* 14(4):1434–1462.
- [64] Fischer, J., Otto, F. (2017). Sublinear growth of the corrector in stochastic homogenization: optimal stochastic estimates for slowly decaying correlations. *Stoch. Partial Differ. Equ. Anal. Comput.* 5(2):220–255.
- [65] Ledoux, M., Talagrand, M. (2013). *Probability in Banach Spaces: Isoperimetry and Processes*. Berlin: Springer.
- [66] Su, Z. (1997). Central limit theorems for random processes with sample paths in exponential orlicz spaces. *Stoch. Process. Appl.* 66(1):1–20.

## A. Properties of stochastic norm $\|\cdot\|_s$

In this section we collect properties of the stochastic norm  $\|\cdot\|_s$  defined in (44). The first lemma presents two other equivalent formulations that are convenient for this work.

**Lemma A.1.** [65, Lemma 3.7] *For any  $s \in (0, 2]$  and the corresponding  $c$  defined in (44),*

$$\|F\|_s \sim_s \sup_{m \in \mathbb{N}} \frac{\langle |F|^m \rangle^{\frac{1}{m}}}{m^{\frac{1}{s}}} \sim_s \inf \{ M > 0 : \log \langle \exp((\frac{|F|}{M})^s) \rangle \leq 1 \}. \quad (\text{A1})$$

*Proof* We first state the following auxiliary result that will be used later: there exist constants  $c_0, c_1 > 0$  such that for all positive integers  $k$ ,

$$(c_0 k)^k \leq k! \leq (c_1 k)^k. \quad (\text{A2})$$

Our proof consists of three steps, with each one showing one expression of (A1) being dominated by another.

Step 1: middle  $\lesssim$  left. If  $\|F\|_s \leq 1$ , or equivalently

$$\langle \exp((|F| + c)^s) \rangle \leq 1 + \exp(c^s),$$

then we will show there exists a constant  $C$ , only depending on  $s$ , such that  $\langle |F|^m \rangle^{\frac{1}{m}} \leq Cm^{\frac{1}{s}}$  for all  $m \geq 1$ . By the Chebyshev inequality,

$$\langle I(|F| > \lambda) \rangle \leq \exp(-(\lambda + c)^s) \langle \exp((|F| + c)^s) \rangle \leq (1 + \exp(c^s)) \exp(-\lambda^s).$$

Therefore, using the formula for non-negative random variable  $F$

$$\langle \varphi(F) \rangle = \varphi(0) + \int_0^\infty \langle I(F \geq \lambda) \rangle \varphi'(\lambda) d\lambda \quad (\text{A3})$$

with  $\varphi(x) = x^m$  and a change of variables  $\tilde{\lambda} := \lambda^s$ ,

$$\begin{aligned} \langle |F|^m \rangle &\stackrel{(\text{A3})}{\leq} m(1 + \exp(c^s)) \int_0^\infty \lambda^{m-1} \exp(-\lambda^s) d\lambda \leq \frac{m(1 + \exp(c^s))}{s} \int_0^\infty \tilde{\lambda}^{\frac{m}{s}-1} e^{-\tilde{\lambda}} d\tilde{\lambda} \\ &= \frac{m(1 + \exp(c^s))}{s} \Gamma\left(\frac{m}{s}\right). \end{aligned}$$

Taking the  $m$ -th root and using the asymptotics of the Gamma function  $\Gamma(\frac{m}{s}) \lesssim \sqrt{m} (\frac{m}{se})^{\frac{m}{s}}$ , we obtain as desired  $\langle |F|^m \rangle^{\frac{1}{m}} \leq Cm^{\frac{1}{s}}$  using  $m^{\frac{1}{m}} \lesssim 1$ .

Step 2: left  $\lesssim$  right. Suppose  $\langle \exp(|F|^s) \rangle \leq e$ , then for any constant  $M$ ,

$$\langle I\left(\frac{|F|}{M} + c > \lambda\right) \rangle = \langle I(|F| > M(\lambda - c)) \rangle \leq \begin{cases} 1 & 0 < \lambda < c \\ \exp(1 - M^s(\lambda - c)^s) & \lambda \geq c, \end{cases}$$

and therefore using again (A3) with  $\varphi(x) = \exp(x^s)$ ,

$$\begin{aligned} \langle \exp\left(\left(\frac{|F|}{M} + c\right)^s\right) \rangle &\stackrel{(\text{A3})}{\leq} 1 + \int_0^c s\lambda^{s-1} \exp(\lambda^s) d\lambda + e \int_c^\infty s\lambda^{s-1} \exp(\lambda^s) \exp(-M^s(\lambda - c)^s) d\lambda \\ &\leq \exp(c^s) + e \int_c^\infty s\lambda^{s-1} \exp(\lambda^s) \exp(-M^s(\lambda - c)^s) d\lambda. \end{aligned}$$

The above integrand decreases to zero pointwise as  $M$  tends to infinity, and is obviously integrable for  $M > 1$ , thus the integral also converges to zero. Hence for some sufficiently large  $M > 1$  only depending on  $s$ ,  $\langle \exp((\frac{|F|}{M} + c)^s) \rangle \leq \exp(c^s) + 1$ , which is equivalent to  $\|F\|_s \leq 1$ .

Step 3: right  $\lesssim$  middle. This direction of the proof uses the algebraic moments of random variables to construct the stretched exponential norm (44). If

$$\frac{\langle |F|^m \rangle^{\frac{1}{m}}}{m^{1/s}} \leq 1, \quad \forall m \in \mathbb{N}, \quad (\text{A4})$$

then, let  $k_0$  be the largest integer such that  $k_0 s < 1$ . In the range  $k \geq k_0 + 1$ , which means  $ks \geq 1$ , for some large  $C$  to be chosen later,

$$\sum_{k=k_0+1}^\infty \frac{\langle |F|^{ks} \rangle}{C^{ks} k!} \stackrel{(\text{A4})}{\leq} \sum_{k=k_0+1}^\infty \frac{(ks)^k}{C^{ks} k!} \stackrel{(\text{A2})}{\leq} \sum_{k=k_0+1}^\infty \frac{s^k}{C^{ks} c_0^k} = \frac{(\frac{s}{C^s c_0})^{k_0+1}}{1 - \frac{s}{C^s c_0}} \leq \frac{s}{C^s c_0 - s}.$$

The lower moments are controlled by the higher moments: for  $1 \leq k \leq k_0$ , using Jensen's inequality on  $\langle \cdot \rangle$ , we have

$$\langle |F|^{ks} \rangle \leq \langle |F|^{(k+k_0)s} \rangle^{\frac{k}{k+k_0}} \leq 1 + \langle |F|^{(k+k_0)s} \rangle.$$

Therefore, using  $C \geq 1$ ,

$$\sum_{k=1}^{k_0} \frac{\langle |F|^{ks} \rangle}{C^{ks} k!} \leq C^{-s} (k_0 + (2k_0)! \sum_{k=k_0+1}^{2k_0} \frac{\langle |F|^{ks} \rangle}{k!}) \stackrel{(A4)}{\leq} C^{-s} (k_0 + (2k_0)! \sum_{k=k_0+1}^{2k_0} \frac{s^k}{c_0^k}).$$

Thus for fixed  $s, c_0$  we may choose  $C$  sufficiently large such that

$$\langle \exp(\frac{|F|^s}{C^s}) \rangle = 1 + \sum_{k=1}^{\infty} \frac{\langle |F|^{ks} \rangle}{C^{ks} k!} \leq 1 + C^{-s} (k_0 + (2k_0)! \sum_{k=k_0+1}^{2k_0} \frac{s^k}{c_0^k}) + \frac{s}{C^s c_0 - s} \leq e.$$

This establishes that the middle expression dominates the third one. Hence all three norms in (A1) are indeed equivalent.  $\square$

The second property of  $\|\cdot\|_s$  is a “stochastic Hölder inequality”.

**Lemma A.2.** For  $s, s_1, s_2 > 0$  with  $\frac{1}{s} = \frac{1}{s_1} + \frac{1}{s_2}$ ,

$$\|F_1 F_2\|_s \lesssim_{s, s_1, s_2} \|F_1\|_{s_1} \|F_2\|_{s_2}. \quad (\text{A5})$$

Taking  $F_2 = 1$  we obtain for any  $s_1 \leq s_2$ ,

$$\|F\|_{s_1} \lesssim \|F\|_{s_2}.$$

*Proof* The proof uses the third characterization of  $\|\cdot\|_s$  in (A1). Indeed, for any  $M_1, M_2 > 0$ , by Young’s inequality and the convexity of  $F \mapsto \log(\exp(F))$ ,

$$\begin{aligned} \log \left\langle \exp\left(\left|\frac{F_1 F_2}{M_1 M_2}\right|^s\right) \right\rangle &\leq \log \left\langle \exp\left(\frac{s}{s_1} \left|\frac{F_1}{M_1}\right|^{s_1} + \frac{s}{s_2} \left|\frac{F_2}{M_2}\right|^{s_2}\right) \right\rangle \\ &\leq \frac{s}{s_1} \log \left\langle \exp\left(\left|\frac{F_1}{M_1}\right|^{s_1}\right) \right\rangle + \frac{s}{s_2} \log \left\langle \exp\left(\left|\frac{F_2}{M_2}\right|^{s_2}\right) \right\rangle. \end{aligned}$$

Pick  $M_1 \geq \|F_1\|_{s_1}$  and  $M_2 \geq \|F_2\|_{s_2}$ , so that in view of the equivalence (A1), the r.h.s.  $\lesssim 1$ . This means  $\|F_1 F_2\|_s \lesssim M_1 M_2$ . Since  $M_1, M_2$  are arbitrary, this establishes (A5).  $\square$

The inequality (A5) allows us to prove the following lemma, which is a generalization of [3, Lemma 14]. This is at the origin in the loss of stochastic integrability from  $s = \infty$  for  $ae$ , to  $s = 2-$  for  $\nabla\phi$  and to  $s = 1-$  for  $\nabla\psi$ .

**Lemma A.3.** Suppose for some  $\delta \in (0, 1]$  and  $p < \infty$  and random variables  $F, F_0, F_1$ , we have

$$F_0 \leq \delta \implies |F| \leq \delta^{\frac{1}{p}} F_1. \quad (\text{A6})$$

Then for  $s, s_0, s_1 > 0$  with  $\frac{1}{s_0} = \frac{1}{s} + \frac{1}{s_1}$ , and  $\tilde{s}_1 \geq s_1$  we have

$$\|F\|_{s_0} \lesssim \delta^{\frac{1}{p}} \|F_1\|_{s_0} + \left(\frac{\|F_0\|_{\tilde{s}_1}}{\delta}\right)^{\frac{\tilde{s}_1}{s_1}} \|F\|_s.$$

*Proof* By the triangle inequality,

$$\|F\|_{s_0} \leq \|I(F_0 \leq \delta)F\|_{s_0} + \|I(F_0 \geq \delta)F\|_{s_0}.$$

By our two assumptions (A6) and  $F_0 \leq \delta$ ,  $\|I(F_0 \leq \delta)F\|_{s_0} \leq \delta^{\frac{1}{p}} \|F_1\|_{s_0}$ . For the control of  $\|I(F_0 \geq \delta)F\|_{s_0}$ , we have by Lemma A.2,

$$\|I(F_0 \geq \delta)F\|_{s_0} \leq \|I(F_0 \geq \delta)\|_{s_1} \|F\|_s.$$



The final step is a Chebyshev inequality:

$$\|I(F_0 \geq \delta)\|_{s_1} = \|I(F_0 \geq \delta)\|_{s_1}^{\frac{s_1}{s_1-1}} \leq \left( \frac{\|F_0\|_{s_1}}{\delta} \right)^{\frac{s_1}{s_1-1}}.$$

□

The last property of  $\|\cdot\|_s$  is that for a stationary random field  $f$ ,  $t \mapsto \|(\int \eta_{\sqrt{t}} f^2)^{\frac{1}{2}}\|_s$  is decreasing in  $t$ :

**Lemma A.4.** *There exists a universal constant  $C$  such that for any  $t \leq T$  and stationary random field  $f$ ,*

$$\|(\int \eta_{\sqrt{T}} f^2)^{\frac{1}{2}}\|_s \leq C \|(\int \eta_{\sqrt{t}} f^2)^{\frac{1}{2}}\|_s. \quad (\text{A7})$$

*Proof* The proof uses [3, (211)] and the stationarity of  $f$ :

$$\begin{aligned} \|(\int \eta_{\sqrt{T}} f^2)^{\frac{1}{2}}\|_s &= \left( \left\| \int \eta_{\sqrt{T}} f^2 \right\|_{\frac{s}{2}} \right)^{\frac{1}{2}} \sim \left( \left\| \int \eta_{\sqrt{T}}(y) \int \eta_{\sqrt{t}}(x-y) f^2(x) \, dx \, dy \right\|_{\frac{s}{2}} \right)^{\frac{1}{2}} \\ &\leq \left( \int \eta_{\sqrt{T}}(y) \left\| \int \eta_{\sqrt{t}}(x-y) f^2(x) \, dx \right\|_{\frac{s}{2}} \, dy \right)^{\frac{1}{2}} \\ &= \|(\int \eta_{\sqrt{t}} f^2)^{\frac{1}{2}}\|_s. \end{aligned}$$

□

## B. Deterministic estimates

The next three lemmas are fundamental deterministic estimates. **Lemma B.1** is a localized energy estimate, **Lemma B.2** is a semigroup estimate that resolves the behavior in the time variable, and **Lemma B.3** is a crucial approximate locality estimate, which is the main tool for finding the approximate locality properties of several objects. These three lemmas are proved in [3] and we do not present their proofs here.

**Lemma B.1.** [3, Lemma 3] *Let  $T > 0$ , and let  $v, f$ , and  $g$  be related through the elliptic equation*

$$\frac{1}{T} v - \nabla \cdot a \nabla v = f + \nabla \cdot g,$$

*then we have for all  $R \geq \sqrt{T}$ ,*

$$\int \eta_R |(\nabla v, \frac{v}{\sqrt{T}})|^2 \lesssim \int \eta_R (|\sqrt{T} f|^2 + |g|^2). \quad (\text{B1})$$

*Let  $v, f, g$ , and  $v_0$  be related through the parabolic equation*

$$\begin{aligned} \partial_t v - \nabla \cdot a \nabla v &= f + \nabla \cdot g \quad \text{for } t > 0, \\ v &= v_0 \quad \text{for } t = 0, \end{aligned}$$

*then we have for all  $R \geq \sqrt{T}$ ,*

$$\sup_{t \leq T} \int \eta_R v^2(t) + \int_0^T dt \int \eta_R |(\nabla v(t), \frac{v(t)}{\sqrt{T}})|^2 \lesssim \int \eta_R v_0^2 + \int_0^T dt \int \eta_R (|\sqrt{T} f|^2 + |g|^2). \quad (\text{B2})$$

**Lemma B.2.** [3, Lemma 1] For all  $R \geq \sqrt{T} > 0$  we have

$$\int \eta_R |(T \nabla S(T)g, \sqrt{T} S(T)g, S(T)g)|^2 + \int \eta_R \left| \int_0^T dt (\nabla S(t)g, \frac{S(t)g}{\sqrt{T}}) \right|^2 \lesssim \int \eta_R |g|^2. \quad (\text{B3})$$

**Lemma B.3.** [3, Lemma 12] For some  $R \geq \sqrt{T}$  consider the following parabolic equation

$$\left. \begin{aligned} \partial_t w - \nabla \cdot a \nabla w &= 0 & \text{for } t > 0 \\ w &= 0 & \text{for } t = 0 \end{aligned} \right\} \text{ in } (0, T) \times B_{2R},$$

then we have for any exponent  $p < \infty$ ,

$$\begin{aligned} \sup_{t \leq T} \left( \int_{B_R} w^2(t) \right)^{\frac{1}{2}} + R \left( \int_{B_R} |\nabla w(T)|^2 \right)^{\frac{1}{2}} + \left( \int_{B_R} \int_0^T dt |\nabla w(t)|^2 \right)^{\frac{1}{2}} \\ \lesssim_p \left( \frac{\sqrt{T}}{R} \right)^p \int_0^T dt \left( \frac{\sqrt{t}}{\sqrt{T}} \right)^p \left( \int_{B_{2R} \setminus B_R} w^2(t) \right)^{\frac{1}{2}}. \quad (\text{B4}) \end{aligned}$$

**Lemma B.3** quantifies that  $w|_{B_R \times (0, T)}$  depends very little on  $w|_{B_{2R}^c \times (0, T)}$ , provided  $R \gg \sqrt{T}$ . In addition, the estimate can afford a time average that degenerates as  $t \ll T$ . The algebraic smallness is sufficient for our purposes.

## C. CLT estimates

The following **Lemma C.1**, which restates [3, Lemma 13], is the main probabilistic ingredient of this paper. It turns deterministic semigroup estimates on small scales and approximate locality properties into CLT-cancellations on large scales. The CLT-scaling typically arises from a combination of **Lemmas B.3** and **C.1**. We will give a proof here since the proof in [3] is only valid when  $s \in (1, 2]$ .

**Lemma C.1.** [3, Lemma 13] Let  $s \in (0, 2]$ . Let  $F$  and  $\bar{F}$  be stationary random fields such that  $\langle F \rangle = 0$  and suppose  $F$  is approximately local on scale  $\sqrt{T}$  relative to  $\bar{F}$  in the sense of (74). Then,

$$\sup_{R \geq \sqrt{T}} \left( \frac{R}{\sqrt{T}} \right)^{\frac{d}{2}} \|F_R\|_s \lesssim \sup_{r \leq \sqrt{T}} \left( \frac{r}{\sqrt{T}} \right)^{\frac{d}{2}} \|F_r\|_s + \left( \sup_{r \leq \sqrt{T}} \left( \frac{r}{\sqrt{T}} \right)^{\frac{d}{2}} \|F_r\|_s \right)^{1 - \frac{d}{2p}} \|\bar{F}\|_s^{\frac{d}{2p}}. \quad (\text{C1})$$

*Proof* By other parts of the proof of [3, Lemma 13] which are valid in the full range  $s \in (0, 2]$ , we only need to prove the following:

Suppose  $F$  is a mean-zero stationary random field that is exactly local on scale  $r \geq 1$ , in the sense that

$$F(a) = F(\tilde{a}) \text{ provided } a = \tilde{a} \text{ on } B_r,$$

then for any convolution kernel  $\tilde{G}$ ,

$$\|\tilde{G} * F\|_s \lesssim r^{\frac{d}{2}} \left( \int \tilde{G}^2 \right)^{\frac{1}{2}} \|F\|_s. \quad (\text{C2})$$

The proof is inspired by [66, Lemma 2.3]. We divide the proof into three steps:

*Step 1:* A CLT in iid symmetric random variables.

Suppose  $X_1, \dots, X_n$  are iid symmetric random variables (in other words, the law of  $X_1$  equals to the law of  $-X_1$ ), then for any real numbers  $w_1, \dots, w_n$ , we claim

$$\left\| \sum_{i=1}^n w_i X_i \right\|_s \lesssim \left( \sum_{i=1}^n w_i^2 \right)^{\frac{1}{2}} \|X_1\|_s. \quad (\text{C3})$$

Without loss of generality we assume  $\|X_1\|_s = 1$ . By Lemma A.1, it is sufficient to show for any integer  $m \geq 1$ ,

$$\left\langle \left( \sum_{i=1}^n w_i X_i \right)^{2m} \right\rangle^{\frac{1}{2m}} \leq C(2m)^{\frac{1}{s}} \left( \sum_{i=1}^n w_i^2 \right)^{\frac{1}{2}}.$$

Let  $c_0, c_1$  be constants as defined in (A2). Since the  $X_i$  are symmetric, if we expand  $\langle (\sum_{i=1}^n w_i X_i)^{2m} \rangle$ , then all odd powers of  $X_i$  will vanish, and therefore

$$\begin{aligned} \left\langle \left( \sum_{i=1}^n w_i X_i \right)^{2m} \right\rangle &= \sum_{m_1 + \dots + m_n = m} \binom{2m}{2m_1, \dots, 2m_n} \prod_{i=1}^n \langle (w_i X_i)^{2m_i} \rangle \\ &\stackrel{(\text{A1})}{\leq} \sum_{m_1 + \dots + m_n = m} \frac{(2m)!}{\prod_{i=1}^n (2m_i)!} \prod_{j=1}^n w_j^{2m_j} (C(2m_j)^{\frac{1}{s}})^{2m_j} \\ &\stackrel{(\text{A2})}{\leq} C^{2m} \sum_{m_1 + \dots + m_n = m} \frac{c_1^{2m} (2m)^{2m}}{\prod_{i=1}^n c_0^{2m_i} (2m_i)^{2m_i}} \prod_{j=1}^n w_j^{2m_j} (2m_j)^{\frac{2m_j}{s}} \\ &\leq \left( \frac{C c_1}{c_0} \right)^{2m} (2m)^{2m} \sum_{m_1 + \dots + m_n = m} \prod_{i=1}^n (2m_i)^{2m_i(\frac{1}{s}-1)} w_i^{2m_i}. \end{aligned}$$

Let  $\varphi(t) = (2m)^{2m} \sum_{m_1 + \dots + m_n = m} \prod_{i=1}^n (2m_i)^{2m_i(t-1)} w_i^{2m_i}$ . A direct calculation (only using  $\log(2m_i) \leq \log(2m)$ ) yields the differential inequality

$$\begin{aligned} \varphi'(t) &= (2m)^{2m} \sum_{m_1 + \dots + m_n = m} \left( \prod_{i=1}^n (2m_i)^{2m_i(t-1)} w_i^{2m_i} \right) \left( \sum_{j=1}^n 2m_j \log(2m_j) \right) \\ &\leq 2m \log(2m) \varphi(t). \end{aligned}$$

The value  $t = \frac{1}{2}$  is special as it corresponds to  $s = 2$ , since we have

$$\begin{aligned} \varphi\left(\frac{1}{2}\right) &= (2m)^{2m} \sum_{m_1 + \dots + m_n = m} \prod_{i=1}^n (2m_i)^{-m_i} w_i^{2m_i} \\ &= (2m)^m \sum_{m_1 + \dots + m_n = m} \frac{m^m}{\prod_{i=1}^n m_i^{m_i}} \prod_{j=1}^n w_j^{2m_j} \\ &\stackrel{(\text{A2})}{\leq} (2m)^m \sum_{m_1 + \dots + m_n = m} \frac{c_1^m m!}{\prod_{i=1}^n c_0^{m_i} m_i!} \prod_{j=1}^n w_j^{2m_j} \\ &= (2m)^m \left( \frac{c_1}{c_0} \right)^m \sum_{m_1 + \dots + m_n = m} \binom{m}{m_1, \dots, m_n} \prod_{i=1}^n w_i^{2m_i} = (2m)^m \left( \frac{c_1}{c_0} \right)^m \left( \sum_{i=1}^n w_i^2 \right)^m. \end{aligned}$$

Therefore we obtain by integration for  $t \geq \frac{1}{2}$  (or  $s \leq 2$ ),

$$\begin{aligned}\varphi(t) &\leq (2m)^m \left(\frac{c_1}{c_0}\right)^m \left(\sum_{i=1}^n w_i^2\right)^m \exp\left(2m \log(2m)\left(t - \frac{1}{2}\right)\right) \\ &= (2m)^m \left(\frac{c_1}{c_0}\right)^m \left(\sum_{i=1}^n w_i^2\right)^m (2m)^{2m(t-\frac{1}{2})} = (2m)^{2mt} \left(\frac{c_1}{c_0}\right)^m \left(\sum_{i=1}^n w_i^2\right)^m.\end{aligned}$$

Finally we take the  $2m$ -th root:

$$\left\langle \left(\sum_{i=1}^n w_i X_i\right)^{2m} \right\rangle^{\frac{1}{2m}} \leq \frac{C c_1}{c_0} (2m)^{\frac{1}{s}} \left(\frac{c_1}{c_0}\right)^{\frac{1}{2}} \left(\sum_{i=1}^n w_i^2\right)^{\frac{1}{2}}.$$

This finishes the proof.

*Step 2:* From symmetric random variables to mean-zero random variables; we claim that (C3) holds just under the assumption that  $\langle X_1 \rangle = 0$ .

For any  $X$  such that  $\langle X \rangle = 0$ , let  $Y$  be independent of  $X$  and equal to  $-X$  in distribution. Let  $\mu$  be the distribution of  $X$  so that  $f \mapsto \int \int f(x-y) d\mu(x) d\mu(y)$  denotes the distribution of  $X+Y$ , which is obviously symmetric. It remains to show

$$\|X+Y\|_s \sim \|X\|_s. \quad (\text{C4})$$

The “ $\lesssim$ ” direction is a direct consequence of the triangle inequality. For the other direction, notice that for any convex function  $f$ ,

$$\int \int f(x-y) d\mu(x) d\mu(y) \geq \int f\left(x - \int y d\mu(y)\right) d\mu(x) = \int f(x) d\mu(x). \quad (\text{C5})$$

We apply this to  $f(x) = \exp(|x|+c)^s - \exp(c^s)$ , which is convex for  $c$  defined in (44). Hence, (C5) turns into the estimate  $\|X+Y\|_s \gtrsim \|X\|_s$ . This suffices to prove equivalence (C4).

*Step 3: Conclusion.* Without loss of generality we assume  $\bar{G}$  is even, which can be removed in view of Step 4 of the proof of [3, Lemma 13]. We divide  $\mathbb{R}^d$  into cubes of size  $3r$ , so that for a fixed  $x \in [0, 3r)^d$ ,  $\{F(3z+x)\}_{z \in r\mathbb{Z}^d}$  are i.i.d. Therefore by stationarity, Step 2, passing to the limit  $n \rightarrow \infty$ ,

$$\begin{aligned}\|\bar{G} * F\|_s &= \|\bar{G} * F(0)\|_s = \left\| \int_{\mathbb{R}^d} \bar{G}(x) F(x) dx \right\|_s \\ &= \left\| r^d \sum_{z \in r\mathbb{Z}^d} \int_{[0, 3r)^d} \bar{G}(x+3z) F(x+3z) dx \right\|_s \\ &\lesssim r^d \int_{[0, 3r)^d} \left\| \sum_{z \in r\mathbb{Z}^d} \bar{G}(x+3z) F(x+3z) \right\|_s dx \\ &\lesssim r^d \int_{[0, 3r)^d} \left( \sum_{z \in r\mathbb{Z}^d} \bar{G}^2(x+3z) \right)^{\frac{1}{2}} dx \|F\|_s \\ &\lesssim r^d \left( \int_{[0, 3r)^d} \sum_{z \in r\mathbb{Z}^d} \bar{G}^2(x+3z) dx \right)^{\frac{1}{2}} \|F\|_s \\ &= r^{\frac{d}{2}} \left( \int_{\mathbb{R}^d} \bar{G}^2 \right)^{\frac{1}{2}} \|F\|_s. \quad \square\end{aligned}$$

The rest of the proof for this lemma follows from Steps 4-6 of [3, Lemma 13].

## D. Large scale regularity for stochastic homogenization

Below we present the crucial elliptic and parabolic large scale regularity estimates on scales larger than  $r_*$  defined in (49). This slightly differs from [16, Theorem 8.7], [3, Corollary 6] since we use exponential averaging  $\int \eta_R$  as opposed to the usual averaging  $\int_{B_R}$ . In Theorem 1.2, we use the growth properties (27) of second-order corrector and flux  $(\psi, \Psi)$  and the stochastic moment of  $r_*$  to derive stochastic bounds of  $r_{**}$ .

**Lemma D.1** ( $C^{0,1}$ -estimate). *The quantity  $r_*$  satisfies the following:*

(1) *For every  $R \geq r_*$  and  $u$  satisfying*

$$\partial_t u - \nabla \cdot a \nabla u = 0, \text{ in } (-R^2, 0) \times \mathbb{R}^d$$

*we have, for every  $r \in [r_*, R]$ ,*

$$\left( \int_{-r^2}^0 dt \int \eta_r |\nabla u(t)|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{-R^2}^0 dt \int \eta_R |\nabla u(t)|^2 \right)^{\frac{1}{2}}, \quad (\text{D1})$$

$$\left( \int_{-r^2}^0 dt \int \eta_r u^2(t) \right)^{\frac{1}{2}} \lesssim \left( \int_{-R^2}^0 dt \int \eta_R u^2(t) \right)^{\frac{1}{2}}. \quad (\text{D2})$$

(2) [3, Corollary 6], [13, Theorem 1.2] *For every  $R \geq r_*$ ,  $u$  satisfying*

$$-\nabla \cdot a \nabla u = 0 \text{ in } B_R,$$

*we have, for every  $r \in [r_*, R]$ ,*

$$\left( \int_{B_r} |\nabla u|^2 \right)^{\frac{1}{2}} \lesssim \left( \int_{B_R} |\nabla u|^2 \right)^{\frac{1}{2}}. \quad (\text{D3})$$

**Proof** *Step 1:* we prove (D1) from the usual version of  $C^{0,1}$  estimate [16, Remark 8.8]: for any  $R \geq r \geq r_*(0)$ ,

$$\left( \int_{B_r^{(P)}} |\nabla u|^2 \right)^{\frac{1}{2}} \leq C \left( \int_{B_R^{(P)}} |\nabla u|^2 \right)^{\frac{1}{2}} \text{ where } B_r^{(P)} := (-r^2, 0) \times Q_r. \quad (\text{D4})$$

Without loss of generality we assume  $r = 1$  and  $R \geq 2$ . We first claim that exponential average can be recovered from the ordinary one as follows:

$$\int \eta_R |\nabla u|^2 \sim \int_R^\infty d\rho \frac{\rho^d}{R^{d+1}} \exp\left(-\frac{\rho}{R}\right) \int_{B_\rho} |\nabla u|^2. \quad (\text{D5})$$

Indeed, let  $S(\rho) = \{x \in \mathbb{R}^d : |x| = \rho\}$  be the  $d$ -dimensional sphere of radius  $r$ . Then by Fubini's theorem

$$\begin{aligned} \int_R^\infty d\rho \frac{\rho^d}{R^{d+1}} \exp\left(-\frac{\rho}{R}\right) \int_{B_\rho} |\nabla u|^2 &\sim \int_R^\infty \frac{d\rho}{R^{d+1}} \exp\left(-\frac{\rho}{R}\right) \int_0^\rho d\rho' \int_{S(\rho')} |\nabla u|^2 \\ &\sim \int_0^R \frac{d\rho'}{R^{d+1}} \int_{S(\rho')} |\nabla u|^2 \int_R^\infty d\rho \exp\left(-\frac{\rho}{R}\right) \\ &\quad + \int_R^\infty \frac{d\rho'}{R^{d+1}} \int_{S(\rho')} |\nabla u|^2 \int_{\rho'}^\infty d\rho \exp\left(-\frac{\rho}{R}\right) \\ &\sim \int_0^R \frac{d\rho'}{R^d} \int_{S(\rho')} |\nabla u|^2 + \int_R^\infty \frac{d\rho'}{R^d} \exp\left(-\frac{\rho'}{R}\right) \int_{S(\rho')} |\nabla u|^2 \end{aligned}$$

$$\sim \int \eta_R |\nabla u|^2.$$

We now prove (D1). The strategy is to divide the  $r$  integral into  $(1, R)$  where we may appeal to (D4), and  $(R, \infty)$  where we use the faster decay of  $\eta_1$  compared to  $\eta_R$ :

$$\begin{aligned} & \int_{-1}^0 dt \int \eta_1 |\nabla u|^2 \\ & \stackrel{(D5)}{\lesssim} \int_{-1}^0 dt \int_1^R d\rho \rho^d \exp(-\rho) \int_{B_\rho} |\nabla u|^2 + \int_{-1}^0 dt \int_R^\infty d\rho \rho^d \exp(-\rho) \int_{B_\rho} |\nabla u|^2 \\ & \lesssim \int_1^R d\rho \rho^{d+2} \exp(-\rho) \int_{-\rho^2}^0 dt \int_{B_\rho} |\nabla u|^2 + R^2 \int_{-R^2}^0 dt \int_R^\infty d\rho \rho^d \exp(-\rho) \int_{B_\rho} |\nabla u|^2 \\ & \stackrel{(D4)}{\lesssim} \int_1^R d\rho \rho^{d+2} \exp(-\rho) \int_{-R^2}^0 dt \int_{B_R} |\nabla u|^2 + \int_{-R^2}^0 dt \int_R^\infty d\rho \frac{\rho^d}{R^{d+1}} \exp(-\frac{\rho}{R}) \int_{B_\rho} |\nabla u|^2 \\ & \stackrel{(D5)}{\lesssim} \int_{-R^2}^0 dt \int_{B_R} |\nabla u|^2 + \int_{-R^2}^0 dt \int \eta_R |\nabla u|^2 \\ & \lesssim \int_{-R^2}^0 dt \int \eta_R |\nabla u|^2. \end{aligned}$$

Here the second term of the third inequality uses that  $R^{d+3} \exp(-\rho) \lesssim \exp(-\frac{\rho}{R})$  when  $\rho \geq R \geq 2$ . This establishes (D1).

Step 2: proof of (D2). The proof starts from the usual version of  $C^{0,1}$  estimate without gradient [16, Theorem 8.7]: for any  $R \geq r \geq r_*(0)$ ,

$$(\int_{B_r^{(p)}} |u - \int_{B_r^{(p)}} u|^2)^{\frac{1}{2}} \lesssim \frac{r}{R} (\int_{B_R^{(p)}} |u - \int_{B_R^{(p)}} u|^2)^{\frac{1}{2}}. \quad (D6)$$

It suffices to prove

$$(\int_{B_r^{(p)}} |u|^2)^{\frac{1}{2}} \lesssim (\int_{B_R^{(p)}} |u|^2)^{\frac{1}{2}}. \quad (D7)$$

since we can then change spatial averaging from  $\int_{B_R}$  to  $\int \eta_R$  using the same post-processing argument as above.

Let  $u_r := \int_{B_r^{(p)}} u$ . W.l.o.g. we may assume  $R \geq 2r$ . Now we can replace  $r$  with  $2r$  in (D6) and get

$$(\int_{B_{2r}^{(p)}} |u - u_{2r}|^2)^{\frac{1}{2}} \lesssim \frac{r}{R} (\int_{B_R^{(p)}} |u - u_R|^2)^{\frac{1}{2}}.$$

Thus we can estimate

$$\begin{aligned} |u_r - u_{2r}| &= |\int_{B_r^{(p)}} (u - u_{2r})| \lesssim (\int_{B_r^{(p)}} |u - u_{2r}|^2)^{\frac{1}{2}} \\ &\lesssim (\int_{B_{2r}^{(p)}} |u - u_{2r}|^2)^{\frac{1}{2}} \lesssim \frac{r}{R} (\int_{B_R^{(p)}} |u - u_R|^2)^{\frac{1}{2}}. \end{aligned} \quad (D8)$$

We now obtain (D7) by telescoping

$$(\int_{B_r^{(p)}} |u|^2)^{\frac{1}{2}} \lesssim (\int_{B_r^{(p)}} |u - u_r|^2)^{\frac{1}{2}} + |u_r|$$

$$\begin{aligned}
&\stackrel{\text{(D6)}}{\lesssim} \frac{r}{R} \left( \int_{B_R^{(p)}} |u - u_R|^2 \right)^{\frac{1}{2}} + \sum_{\substack{\tilde{r} \in [r, R) \\ \text{dyadic}}} |u_{\tilde{r}} - u_{2\tilde{r}}| + |u_R| \\
&\stackrel{\text{(D8)}}{\lesssim} \frac{r}{R} \left( \int_{B_R^{(p)}} |u|^2 \right)^{\frac{1}{2}} + \sum_{\substack{\tilde{r} \in [r, R) \\ \text{dyadic}}} \frac{\tilde{r}}{R} \left( \int_{B_R^{(p)}} |u - u_R|^2 \right)^{\frac{1}{2}} + |u_R| \\
&\lesssim \left( \int_{B_R^{(p)}} |u|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

□

The following corollary is a convenient post-processing of [Lemma D.1](#) from time intervals to single time slices.

**Corollary D.2.** *For any  $\sqrt{t} \geq r \geq r_*$ , we have*

$$\left( \int \eta_r |\nabla S(t)g|^2 \right)^{\frac{1}{2}} \lesssim \left( \int \eta_{\sqrt{t}} |\nabla S(\frac{t}{2})g|^2 \right)^{\frac{1}{2}}, \quad (\text{D9})$$

$$\left( \int \eta_r |S(t)g|^2 \right)^{\frac{1}{2}} \lesssim \left( \int \eta_{\sqrt{t}} |S(\frac{t}{2})g|^2 \right)^{\frac{1}{2}}. \quad (\text{D10})$$

*Proof* We will only prove (D9) as the proof of (D10) is identical. The estimate uses (D1) and the fact that the function  $\tau \mapsto \int \eta_{\sqrt{t}} |\nabla S(\tau)g|^2$  is approximately non-increasing for any  $g$ , that is, for any  $t_1 \leq t_2 \leq r^2$ ,  $\int \eta_r |\nabla S(t_1)g|^2 \lesssim \int \eta_r |\nabla S(t_2)g|^2$  with the constant in  $\lesssim$  independent of  $r, t_1, t_2, g$ :

$$\begin{aligned}
\left( \int \eta_r |\nabla S(t)g|^2 \right)^{\frac{1}{2}} &\lesssim \left( \int \eta_r \int_{t-r^2}^t d\tau |\nabla S(\tau)g|^2 \right)^{\frac{1}{2}} \\
&\stackrel{\text{(D1)}}{\lesssim} \left( \int \eta_{\sqrt{t}} \int_{\frac{t}{2}}^t d\tau |\nabla S(\tau)g|^2 \right)^{\frac{1}{2}} \\
&\lesssim \left( \int \eta_{\sqrt{t}} |\nabla S(\frac{t}{2})g|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

□

**Remark D.3.** We further post-process [Corollary D.2](#) into the following: for any  $t, r \geq 0$ ,

$$\|I(r \geq r_*) \left( \int \eta_r |\nabla S(t)g|^2 \right)^{\frac{1}{2}}\|_s \lesssim \|I(r \vee \sqrt{t} \geq r_*) \left( \int \eta_{r \vee \sqrt{t}} |\nabla S(\frac{t}{2})g|^2 \right)^{\frac{1}{2}}\|_s, \quad (\text{D11})$$

$$\|I(r \geq r_*) \left( \int \eta_r |S(t)g|^2 \right)^{\frac{1}{2}}\|_s \lesssim \|I(r \vee \sqrt{t} \geq r_*) \left( \int \eta_{r \vee \sqrt{t}} |S(\frac{t}{2})g|^2 \right)^{\frac{1}{2}}\|_s. \quad (\text{D12})$$

The argument uses the fact that the function  $[0, r^2) \ni t \mapsto \int \eta_r |\nabla S(t)g|^2$  and  $[0, r^2) \ni t \mapsto \int \eta_r |S(t)g|^2$  are both approximately non-increasing for any  $g$  (see the proof of [[3](#), Lemma 3]).