The Choice of Representative Volumes in the Approximation of Effective Properties of Random Materials

JULIAN FISCHER

Communicated by S. MÜLLER

Abstract

The effective large-scale properties of materials with random heterogeneities on a small scale are typically determined by the method of representative volumes: a sample of the random material is chosen—the representative volume—and its effective properties are computed by the cell formula. Intuitively, for a fixed sample size it should be possible to increase the accuracy of the method by choosing a material sample which captures the statistical properties of the material particularly well; for example, for a composite material consisting of two constituents, one would select a representative volume in which the volume fraction of the constituents matches closely with their volume fraction in the overall material. Inspired by similar attempts in materials science, Le Bris, Legoll and Minvielle have designed a selection approach for representative volumes which performs remarkably well in numerical examples of linear materials with moderate contrast. In the present work, we provide a rigorous analysis of this selection approach for representative volumes in the context of stochastic homogenization of linear elliptic equations. In particular, we prove that the method essentially never performs worse than a random selection of the material sample and may perform much better if the selection criterion for the material samples is chosen suitably.

1. Introduction

The most widely employed method for determining the effective large-scale properties of a material with random heterogeneities on a small scale is the method

of representative volumes. It basically proceeds by taking a small sample of the material—a "representative volume element" (RVE)—and determining the properties of the sample by the cell formula. The criteria for the choice of the representative volume have been the subject of an ongoing debate; while in principle increasing the size of the material sample increases the accuracy of the approximation of the material properties, this comes at a correspondingly larger computational cost. It has been conjectured that for a fixed size of the material sample, selecting a material sample which captures certain statistical properties of the material in a particularly good way may be beneficial; for example, for a composite material consisting of two constituent materials, one would try to select a material sample for which the volume fraction of each constituent material within the sample matches the overall volume fraction of this constituent in the composite as closely as possible (see Fig. 1). Alternatively, for linear materials one might try to match the averaged material coefficient in the sample with the average taken over the full material. There have been efforts in materials science and mechanics towards replicating further statistical properties of the material in a representative volume, an approach called "special quasirandom structures" [82, 83, 86] or "statistically similar representative volume elements" [15–18, 28, 81]. A particularly successful approach in this direction has been developed for linear materials by LE BRIS et al. [64]; their method proceeds by considering a large number of material samples, evaluating one or more cheaply computable statistical quantities of the samples (like, for example, the spatial average of the coefficient), and then choosing the sample as the representative volume that is most representative for the material as measured by these quantities. In the present work, in the context of stochastic homogenization of linear elliptic PDEs we provide the first rigorous justification of these approaches.[1]

For materials with random heterogeneities on small scales, the approximation of the effective material coefficient by the method of representative volumes is a random quantity itself, as the outcome depends on the sample of the material. In the setting of linear elliptic PDEs with random coefficient fields—which corresponds to the setting of heat conduction, electrical currents, or electrostatics in a material with random microstructure—GLORIA and OTTO [48, 53, 54] have investigated the structure of the error of the approximation of the effective material coefficient by the method of representative volumes: the leading-order contribution to the error (with respect to the size of the RVE) consists of random fluctuations; in expectation the approximation of effective coefficients by the method of representative volumes is accurate to higher order, that is the systematic error of the RVE method is of higher order.[2] For a given size of the RVE—which corresponds to a fixed computational effort—the accuracy of the RVE method may therefore be increased significantly by reducing the variance of the approximations of the effective coefficient. It is

---

[1] Note that for one-dimensional linear elliptic PDEs—a case in which homogenization is linear in the inverse of the coefficient and thus independent of the geometry of the material— an analysis has directly been provided in [64].

[2] At least if a suitable *periodization* of the probability distribution of the coefficient field is available, see below for an explanation of this concept.
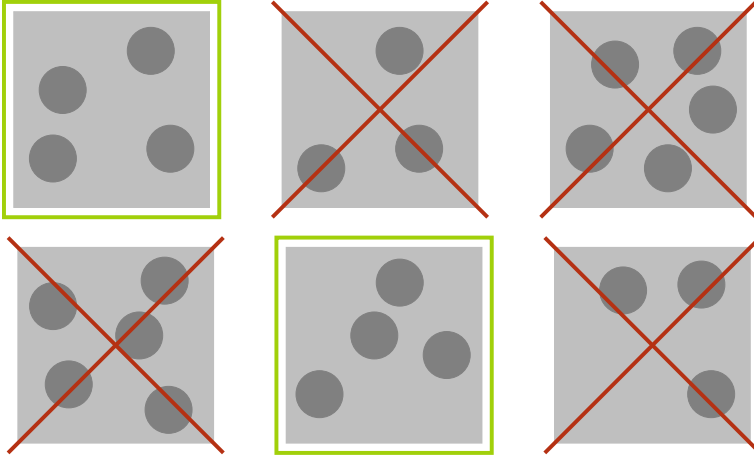
Fig. 1. Among the six depicted material samples, the method of Le Bris, Legoll, and Minvielle in its simplest realization would choose either the first sample or the fifth sample as the representative volume element and discard the others, as the volume fraction of the inclusions in the first and the fifth sample is closest to the overall material average. Note that in the depicted material samples the volume fraction of the inclusions is proportional to the number of inclusions, as all inclusions are of equal size. For a better illustration of the method, both the size and the number of the depicted samples have been chosen much smaller than in actual computations

precisely such a reduction of the variance by which the selection approach for representative volumes of Le Bris et al. [64] achieves its gain in accuracy.

For linear elliptic PDEs with random coefficients and moderate ellipticity contrast, the reduction of the variance by the ansatz of Le Bris et al. [64] is particularly remarkable; by selecting the representative volume according to the criterion that the averaged coefficient in the RVE should be particularly close to the averaged coefficient in the overall material, in numerical examples with ellipticity contrast ~ 5 they observed a variance reduction by a factor of ~ 10. Going beyond this simple selection criterion, they devised a criterion based on an expansion of the effective coefficient in the regime of small ellipticity contrast, which numerically achieves a remarkable variance reduction factor of ~ 60 even for a moderate ellipticity contrast ~ 5. Note that this basically corresponds to the gain of about one order of magnitude in accuracy for a negligible additional computational cost and implementation effort.

However, the analysis of the selection approach for representative volumes has been restricted to the one-dimensional setting [64], in which the homogenization of linear elliptic PDEs is linear in the inverse coefficient and therefore independent of the geometry of the material. Besides the highly nonlinear dependence of the effective coefficient on the heterogeneous coefficient field in dimensions $d \geqq 2$, one of the main challenges in the analysis of the selection method for representative volumes is the fact that it is only expected to increase the accuracy by a (though often very large) constant factor, at least for a fixed set of statistical quantities by which the selection is performed. At the same time, the available error estimates for

the representative volume element method in stochastic homogenization are only optimal up to constant factors. For this reason, the analysis of the selection approach for representative volumes necessitates a fine-grained analysis of the structure of fluctuations in stochastic homogenization.

1.1. Stochastic Homogenization of Linear Elliptic PDEs: A Brief Outline

The subject of the present contribution is the rigorous justification of the selection method for representative volumes by LE BRIS et al. [64] in the context of linear elliptic equations

$$-\nabla \cdot (a\nabla u) = f \tag{1}$$

with random coefficient fields a on $\mathbb{R}^d$ for arbitrary spatial dimension d. Note that this setting describes, for example, heat conduction or electrostatics in a random material. Our assumptions on the probability distribution of the coefficient field a are standard in the theory of stochastic homogenization; we assume just uniform ellipticity and boundedness, stationarity, and finite range of dependence (see conditions (A1)–(A3) below). In particular, our analysis includes the case of a two-material composite with random non-overlapping inclusions as depicted in Fig. 1.

The theory of stochastic homogenization of linear elliptic PDEs predicts that for coefficient fields with only short-range correlations on a scale $\varepsilon \ll 1$ the solution u to the equation with random coefficient field (1) may be approximated by the solution u_{hom} of an effective equation of the form

$$-\nabla \cdot (a_{\mathsf{hom}}\nabla u_{\mathsf{hom}}) = f, \tag{2}$$

where $a_{\mathsf{hom}} \in \mathbb{R}^{d\times d}$ is a constant effective coefficient which describes the effective behavior of the material. In this context of linear materials, the method of representative volumes is employed to compute the effective coefficient a_{hom}.

Let us describe the method of representative volumes for the approximation of the effective material coefficient a_{hom} in more detail. It proceeds by choosing a sample of the material, say, a cube with side length $L\varepsilon$ for some $L \gg 1$, uniformly at random. Roughly speaking—for the moment passing silently over the question of boundary conditions—by solving the equation for the homogenization corrector ϕ_i associated with the i-th coordinate direction on the representative volume

$$-\nabla \cdot (a(e_i + \nabla\phi_i)) = 0 \qquad \text{on } [0, L\varepsilon]^d \tag{3}$$

($e_i \in \mathbb{R}^d$ denoting the i-th vector of the standard basis) one may obtain an approximation a^{RVE} for the effective coefficient a_{hom} in terms of the averaged fluxes

$$a^{\mathsf{RVE}}e_i := \fint_{[0, L\varepsilon]^d} a(e_i + \nabla\phi_i) \, \mathrm{d}x. \tag{4}$$

This expression is also known in homogenization as the *cell formula*. As already mentioned before, the approximation a^{RVE} for the effective material coefficient

a_{hom} is a random variable itself, as it depends on the realization of the random coefficient field a on the sample volume $[0, L\varepsilon]^d$. It has been proven by GLORIA and OTTO [54,55] and also observed in numerical computations that the main contribution to the error of the RVE method is caused by the random fluctuations of the approximation a^{RVE}, while the systematic error is of higher order: For spatial dimensions $d \geqq 1$ one has

$$\sqrt{\operatorname{Var} a^{\mathsf{RVE}}} \lesssim L^{-d/2}, \tag{5}$$

but

$$\left| \mathbb{E}[a^{\mathsf{RVE}}] - a_{\mathsf{hom}} \right| \lesssim L^{-d} |\log L|^d. \tag{6}$$

As a consequence, a reduction of the fluctuations of the approximations a^{RVE} would lead to an increase in accuracy of the approximation for the effective coefficient a_{hom}. It has been observed numerically by LE BRIS et al. [64], and shall be proven below rigorously, that the selection approach for representative volumes achieves its gain in accuracy precisely by reducing the fluctuations of the approximations for the effective coefficients.

1.2. Informal Summary of Our Main Results

In the present work, we prove that in the setting of stochastic homogenization of linear elliptic equations the selection approach for representative volumes by LE BRIS et al. [64]

- essentially never performs worse than a completely random selection of the representative volume element, but may perform much better for suitable selection criteria,
- basically maintains the order of the systematic error of the approximation for the effective coefficient, and
- reduces also the error in the approximation for the effective coefficient that may occur with a given low probability, that is reduces also the "outliers" of the approximation for the effective coefficient.

As mentioned before, in the setting of linear elliptic PDEs the method of representative volumes is employed to obtain an approximation a^{RVE} for the effective (homogenized) coefficient a_{hom}. The role of "material samples" is assumed by realizations of the random coefficient field $a : [0, L\varepsilon]^d \to \mathbb{R}^{d \times d}$, on which the computation of the approximations a^{RVE} is based.

The selection approach for representative volumes proposed in [64] then proceeds as follows: at first, one or more statistical quantities $\mathcal{F}$ are chosen which assign a real number $\mathcal{F}(a) \in \mathbb{R}$ to any realization $a : [0, L\varepsilon]^d \to \mathbb{R}^{d \times d}$. Note that the simplest statistical quantity proposed in [64] is the spatial average $\mathcal{F}(a) := \fint_{[0,L\varepsilon]^d} a \, dx$. Next, one considers a sequence of independent samples of the random coefficient field until a sample meets the selection criterion

$$\left| \mathcal{F}(a) - \mathbb{E}[\mathcal{F}(a)] \right| \leqq \delta \sqrt{\operatorname{Var} \mathcal{F}(a)} \tag{7}$$

for some chosen parameter δ with $CL^{-d/2}|\log L|^C \leq \delta \leq 1$. Finally, the approximation for the effective coefficient is computed by solving the equation for the homogenization corrector (3) and using the cell formula (4) for this sample of the random coefficient field.

To give a flavor of our main result, let us formulate it informally in the case of a single statistical quantity $\mathcal{F}(a)$. We denote the approximation for the effective coefficient by the standard representative volume element method (without selection of material samples) by a^{RVE} and the approximation for the effective coefficient by the selection approach for representative volumes by $a^{\text{sel-RVE}}$. In this case, our main theorems Theorems 2 and 3 may be summarized as follows:

- The systematic error of the approximation $a^{\text{sel-RVE}}$ is essentially (up to powers of $\log L$ and some prefactors) of the same order as the systematic error of the standard representative volume element method a^{RVE}: We have

$$\left| \mathbb{E}\left[a^{\text{sel-RVE}} \right] - a_{\text{hom}} \right| \leqq \frac{C\kappa^{3/2}}{\delta} L^{-d} |\log L|^C.$$

  The quantity κ will be discussed below.
- The fluctuations of the approximation $a^{\text{sel-RVE}}$ are reduced by the fraction of the variance of a^{RVE} that is explained by $\mathcal{F}(a)$. More precisely, we derive the estimate

$$\frac{\text{Var } a^{\text{sel-RVE}}}{\text{Var } a^{\text{RVE}}} \leqq 1 - (1 - \delta^2)|\rho_{\mathcal{F}(a),a^{\text{RVE}}}|^2 + \frac{C\kappa^{3/2} r_{\text{Var}}}{\delta} L^{-d/2} |\log L|^C$$

  where $\rho_{\mathcal{F}(a),a^{\text{RVE}}} \in [-1, 1]$ denotes the correlation coefficient of $\mathcal{F}(a)$ and a^{RVE}, given by

$$\rho_{\mathcal{F}(a),a^{\text{RVE}}} := \frac{\text{Cov}[a^{\text{RVE}}, \mathcal{F}(a)]}{\sqrt{\text{Var } \mathcal{F}(a) \text{Var } a^{\text{RVE}}}},$$

  and where $r_{\text{Var}} := \frac{L^{-d}}{\text{Var } a^{\text{RVE}}}$ denotes the ratio between the expected order of fluctuations of a^{RVE} and the actual magnitude of fluctuations. Note that the last term in the estimate on Var $a^{\text{sel-RVE}}$ converges to zero as the size L of the representative volume increases.
- The probability of "outliers" is reduced by the selection method just as suggested by the variance reduction, at least in an "intermediate" region between the "bulk" and the "outer tail" of the probability distribution: One has a moderate-deviations-type estimate of the form

$$\mathbb{P}\left[\frac{|a_{ij}^{\text{sel-RVE}} - a_{\text{hom},ij}|}{\sqrt{\left(1 - |\rho_{\mathcal{F}(a),a^{\text{RVE}}}|^2 + \delta^2\right)\text{Var } a_{ij}^{\text{RVE}}} + L^{-d/2-\beta}} \geqq s \right]$$

$$\leqq \left(1 + \frac{C\delta}{\sqrt{1 - |\rho|^2} s} + \frac{C}{\delta L^\beta} \right) \mathbb{P}[|\mathcal{N}_1| \geqq s] + \frac{C}{\delta} \exp(-L^{2\beta})$$

  for any $s \geq C \max\{(1-|\rho|^2)^{1/2}\delta^{-1}, \delta(1-|\rho|^2)^{-1/2}\}$ and some $\beta = \beta(d) > 0$, where $\mathcal{N}_1$ denotes the centered normal distribution with unit variance.

- In the above bounds, $\kappa := (1 - |\rho_{\mathcal{F}(a),a^{\mathrm{RVE}}}|^2)^{-1}$ denotes (essentially) the condition number of the covariance matrix Var $(a^{\mathrm{RVE}}, \mathcal{F}(a))$. For the case that the correlation $|\rho_{\mathcal{F}(a),a^{\mathrm{RVE}}}|$ is close to one, we derive bounds which are independent of κ but come at the cost of a lower rate of convergence in L, namely

$$\left|\mathbb{E}\big[a^{\mathrm{sel\text{-}RVE}}\big] - a_{\mathsf{hom}}\right| \leqq \frac{C}{\delta} L^{-d/2 - d/8} |\log L|^C$$

and

$$\frac{\mathrm{Var}\ a^{\mathrm{sel\text{-}RVE}}}{\mathrm{Var}\ a^{\mathrm{RVE}}} \leqq 1 - (1 - \delta^2)\big|\rho_{\mathcal{F}(a),a^{\mathrm{RVE}}}\big|^2 + \frac{Cr_{\mathrm{Var}}}{\delta} L^{-d/8} |\log L|^C.$$

Our estimate on the variance reduction achieved by the selection approach for representative volumes is implicit in the sense that it is determined by the correlation coefficient

$$\rho_{\mathcal{F}(a),a^{\mathrm{RVE}}} := \frac{\mathrm{Cov}[a^{\mathrm{RVE}}, \mathcal{F}(a)]}{\sqrt{\mathrm{Var}\ \mathcal{F}(a)\mathrm{Var}\ a^{\mathrm{RVE}}}}.$$

In fact, the failure of the correlation coefficient $\rho_{\mathcal{F}(a),a^{\mathrm{RVE}}}$ to be nonzero also implies the failure of gaining accuracy by the selection approach for the representative volumes (see Theorem 4): In such a case of vanishing correlation, the method of LE BRIS et al. [64] is not superior (but essentially also not inferior) to the standard method of choosing a representative volume randomly.

This raises the question whether such a degeneracy of the correlation coefficient can occur for "natural" choices of the statistical quantity $\mathcal{F}(a)$. In Theorem 4, we shall prove that even for a "natural" choice like $\mathcal{F}(a) := \fint_{[0,\varepsilon L]^d} a\,\mathrm{d}x$ there is *a priori* no guarantee that there is a nonzero correlation between a^{RVE} and $\mathcal{F}(a)$: We construct an example of a probability distribution of a for which the covariance of a^{RVE} and the average of the coefficient field $\fint a$ in fact vanishes, while the variances Var $\fint_{[0,\varepsilon L]^d} a\,\mathrm{d}x$ and Var a^{RVE} are nondegenerate.

However, the failure of the variance reduction approaches to effectively reduce the variance is presumably limited to rather artificial examples: we prove that the covariance of a^{RVE} and the average of the coefficient field $\fint a$ is positive for coefficient fields which are obtained from iid random variables by applying a "monotone" function, see Proposition 5.

1.3. Outline of Our Strategy

The basic idea underlying our analysis of the selection approach for representative volumes is the observation that the joint probability distribution of the approximation for the effective coefficient a^{RVE} and one or more statistical quantities $\mathcal{F}(a)$ like the average of the coefficient field $\mathcal{F}(a) := \fint_{[0,L\varepsilon]^d} a$ is close to a multivariate Gaussian, up to an error of the order $L^{-d}|\log L|^C$ in a suitable notion of distance between probability measures. The selection of representative volumes by the criterion (7)—which amounts to conditioning on the event $|\mathcal{F}(a) - \mathbb{E}[\mathcal{F}(a)]| \leqq \delta\sqrt{\mathrm{Var}\ \mathcal{F}(a)}$—then reduces the variance of the probability
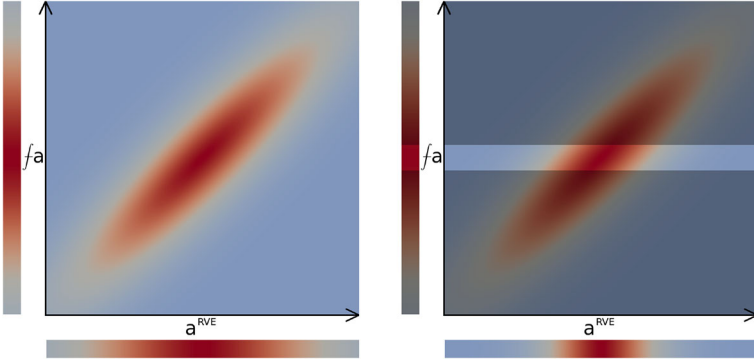
Fig. 2. For a multivariate Gaussian probability distribution, conditioning on the event of one variable being close to its expectation reduces the variance of the other variable, provided that the two random variables are nontrivially correlated. In our setting, conditioning on the event "spatial average of coefficient field is close to its expectation" reduces the variance of the random variable "approximation for the effective conductivity" a^{RVE}, as their joint probability distribution is close to a multivariate Gaussian

distribution of a^{RVE} by the variance explained by the statistical quantity $\mathcal{F}(a)$, up to error terms due to the deviation of the probability distribution from a multivariate Gaussian and the non-perfectness of the conditioning $\delta > 0$, see Fig. 2. Note that for an ideal multivariate Gaussian distribution, the expected value of the approximation a^{RVE} would be left unchanged under conditioning since the criterion (7) is symmetric around $\mathbb{E}[\mathcal{F}(a)]$, that is the conditioning would not introduce a bias. As a consequence, for our approximate multivariate Gaussian $(a^{\mathrm{RVE}}, \mathcal{F}(a))$ the expectation of a^{RVE} is changed under conditioning only by the distance of our probability distribution to a multivariate Gaussian, which is a higher-order term. Note that both the reduction of the variance by conditioning and the estimate on the bias introduced by the conditioning rely crucially on the fact that our probability distribution is close to a multivariate Gaussian (and not another probability distribution); it is obvious from the picture in Fig. 2 that a probability distribution other than a multivariate Gaussian could introduce a large bias under conditioning and even an *increase* in variance. Our analysis of the selection approach for representative volumes by LE BRIS et al. [64] is a first practical application of the beautiful theory of fluctuations in stochastic homogenization, which has been developed in recent years and which our work both draws ideas from and contributes to.

The underlying reason for the convergence of the joint probability distribution of a^{RVE} and one or more functionals $\mathcal{F}(a)$ towards a multivariate Gaussian is a central limit theorem for suitable collections of vector-valued random variables. We show that the approximation a^{RVE} for the effective coefficient a_{hom}—and also the functionals $\mathcal{F}(a)$ that are used in the work of LE BRIS et al. [64]—may be written as a sum of random variables with a local dependence structure with multiple levels, see Definition 6 and Proposition 7. For such sums of vector-valued random variables with multilevel local dependence, a proof of quantitative normal approximation is provided in the companion article [43] (see also Theorem 9 below). To the best of our

knowledge such quantitative normal approximation results were previously known only for sums of random variables with local dependence structure [33,34,80] (corresponding more or less to just the lowest level of random variables in Fig. 4 below), a framework into which the approximation for the effective coefficient a^{RVE} does not fit. Note that the sharp boundaries of the region defined by the selection criterion (7) (see also the sharp boundaries in Fig. 2) necessitate the use of a rather strong (though standard) distance between probability measures for our quantitative normal approximation result (see Definition 8); in particular, a stronger notion of distance between probability measures than the 1-Wasserstein distance must be used.

As a by-product, our work also provides a proof of quantitative normal approximation for a^{RVE} in a different setting than available in the literature so far. To the best of our knowledge, the results on quantitative normal approximation for a^{RVE} in the literature always rely on an assumption that the coefficient field a is obtained as a function of iid random variables [39,52,77] or that the probability distribution of a is subject to a second-order Poincaré inequality like in [38]. In contrast, our result holds under the assumption of *finite range of dependence*, in which to the best of our knowledge only a qualitative normal approximation result had been known [6].

The companion article [43] also provides a result on moderate deviations in the sense of Kramers for sums of random variables with multilevel local dependence structure, see Theorem 10. Our result on the reduction of the error by the selection approach for representative volumes in the case of unlikely events (Theorem 3) is based on this moderate deviations theorem.

Our counterexample for the variance reduction—which shows that even "natural" statistical quantities like the spatial average $\mathcal{F}(a) := \fint_{[0,L\varepsilon]^d} a \, dx$ do not necessarily explain a positive fraction of the variance of a^{RVE}—is based on the nonlinear dependence of the effective coefficient in periodic homogenization on the underlying coefficient field. More precisely, our counterexample consists of an interpolation between a standard random checkerboard and a random checkerboard with two types of tiles, one tile type being a constant coefficient field and one tile type being a second-order laminate microstructure; see Section 6 for details of the construction.

1.4. Computation of Effective Properties of Random Materials: A More Detailed Look

In the homogenization of periodic linear materials—that is in the homogenization of the linear elliptic PDE (1) with periodic coefficient field a in the sense $a(x) = a(x + \varepsilon k)$ for all $k \in \mathbb{Z}^d$—it is possible to compute the effective coefficient a_{hom} by exploiting the periodicity of the coefficient field, basically reducing the problem to solving a PDE—the PDE for the homogenization corrector—on a single periodicity cell: for a period of length ε, the effective coefficient is given by the cell formula

$$a_{\text{hom}} e_i \cdot e_j := \fint_{[0,\varepsilon]^d} a(e_i + \nabla \phi_i) \cdot e_j \, dx,$$

with the homogenization corrector ϕ_i defined as the unique ε-periodic solution with zero average to the PDE

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

As a consequence, in periodic homogenization the numerical computation of the effective coefficient a_{hom} typically requires only modest effort.

In contrast, in stochastic homogenization this simplification is no longer possible due to the absence of a periodic structure in the random coefficient field $a^{\mathbb{R}^d} : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ and the computation of the effective coefficient becomes a computationally costly problem. The effective coefficient in stochastic homogenization is given by the infinite volume limit cell formula[3]

$$a_{\mathsf{hom}} e_i \cdot e_j := \lim_{L \to \infty} \fint_{[0, L\varepsilon]^d} a^{\mathbb{R}^d} (e_i + \nabla \phi_i^{\mathrm{L,Dir}}) \cdot e_j \, \mathrm{d}x$$

with $\phi_i^{\mathrm{L,Dir}}$ denoting the solution to the corrector problem with Dirichlet boundary conditions

$$-\nabla \cdot (a^{\mathbb{R}^d} (e_i + \nabla \phi_i^{\mathrm{L,Dir}})) = 0 \qquad \text{in } [0, L\varepsilon]^d,$$
$$\phi_i^{\mathrm{L,Dir}} \equiv 0 \qquad \text{on } \partial[0, L\varepsilon]^d.$$

In practice, in order to approximate the effective coefficient a_{hom} a representative volume $[0, L\varepsilon]^d$ of finite size must be chosen. However, the approximation of the effective coefficient by the standard cell formula with Dirichlet boundary conditions for the corrector

$$a_{\mathsf{hom}} e_i \cdot e_j \approx a_{\mathrm{Dir}}^{\mathrm{RVE}} e_i \cdot e_j := \fint_{[0, L\varepsilon]^d} a^{\mathbb{R}^d} (e_i + \nabla \phi_i^{\mathrm{L,Dir}}) \cdot e_j \, \mathrm{d}x$$

is only of first-order accuracy $\mathbb{E}[|a_{\mathrm{Dir}}^{\mathrm{RVE}} - a_{\mathsf{hom}}|^2]^{1/2} \lesssim L^{-1}$ due to the presence of a boundary layer: the artificial Dirichlet boundary condition leads to the creation of a boundary layer in an $O(\varepsilon)$-neighborhood of the boundary $\partial[0, L\varepsilon]^d$. The limitation to first-order accuracy is present even in the systematic error $\mathbb{E}[a^{\mathrm{RVE}}] - a_{\mathsf{hom}}$. Note that while replacing the volume average in the cell formula by an average taken strictly in the interior of the representative volume typically increases the accuracy [84], for general probability distributions it does not increase the order of convergence due to global effects of the boundary layer. To achieve the convergence rates $|\mathbb{E}[a^{\mathrm{RVE}}] - a_{\mathsf{hom}}| \lesssim L^{-d} |\log L|^d$ and $\mathbb{E}[|a^{\mathrm{RVE}} - a_{\mathsf{hom}}|^2]^{1/2} \lesssim L^{-d/2}$ stated in (6) and (5), the boundary layer phenomenon must necessarily be addressed by the use of a more careful approximation technique than the method of correctors with Dirichlet boundary data.

One possibility of avoiding the creation of boundary layers is the use of a so-called "periodization" of the probability distribution: Given a probability distribution of coefficient fields $a^{\mathbb{R}^d}$, one first fixes the size $L\varepsilon$ of the desired representative volume and then attempts to construct a probability distribution of $L\varepsilon$-periodic coefficient fields a such that the law of $a|_{x+[0, \frac{1}{2}L\varepsilon]^d}$ (i.e the law of a restricted

---

[3] This limit is to be read in an almost sure sense: By ergodicity, for almost every realization of a this limit exists and is equal to a matrix which is independent of the realization.

to some box of half the size of the representative volume) coincides with the law of $a^{\mathbb{R}^d}|_{x+[0,\frac{1}{2}L\varepsilon]^d}$ for any $x \in \mathbb{R}^d$. For one realization of the periodized probability distribution of coefficient fields a one may then solve the corrector equation $-\nabla \cdot (a(e_i + \nabla\phi_i)) = 0$ with periodic boundary conditions on $\partial[0, L\varepsilon]^d$ and define the approximation a^{RVE} for the effective coefficient a_{hom} as

$$a^{\text{RVE}}e_i := \fint_{[0,L\varepsilon]^d} a(e_i + \nabla\phi_i) \, dx. \tag{8}$$

This approximation a^{RVE} then has the desired approximation properties (5) and (6). Note that this construction requires the knowledge of the probability distribution of $a^{\mathbb{R}^d}$ and must be done in a case-by-case basis; it is therefore not feasible in all practical situations.

To give an example, random non-overlapping inclusions like in Fig. 1 may be constructed by considering a Poisson point process on $\mathbb{R}^d \times [0, 1]$, ordering the points $(x_k, y_k) \in \mathbb{R}^d \times [0, 1]$ with respect to their last coordinate y_k, and then successively placing inclusions in $\mathbb{R}^d$ centered at the x_k and with diameter ε if the "previous" points $x_l, l < k$, have a distance of at least ε from x_k (that is $|x_l - x_k| \geqq \varepsilon$). The result of such a construction is shown in Fig. 3a. For this probability distribution, one may define a periodization in a natural way by considering a Poisson point process on $[0, L\varepsilon)^d \times [0, 1]$ and defining an $L\varepsilon$-periodic coefficient field with non-overlapping inclusions in the obvious way, replacing the Euclidean distance $|x_l - x_k|$ by the periodicity-adjusted distance $|x_l - x_k|_{\text{per}} := \inf_{z \in \mathbb{Z}^d} |x_l - x_k + L\varepsilon z|$. A sample from the periodized probability distribution is shown in Fig. 3b.

If no periodization of the probability distribution is available—for example if only samples from the probability distribution are available and the underlying probability distribution is not known, like in applications where one has access to samples of the materials—, one has to resort to an alternative means of increasing the rate of convergence of the method of representative volumes. One feasible option is to "screen" the effect of the boundary by introducing a "massive" term in
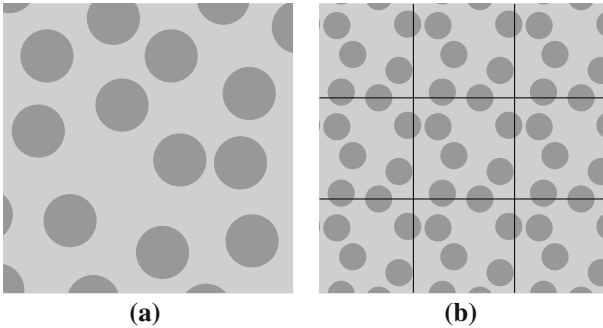

(a)   **(b)**

Fig. 3. a An example of random spherical inclusions distributed according to a Poisson point process, with overlapping inclusions removed. **b** A sample from the corresponding periodization of the probability distribution (rescaled); the periodicity cell is indicated by black lines

the PDE for the homogenization corrector [24,47,54]: Fixing a scale $\sqrt{T} \sim \frac{L}{\log L}$, one replaces the equation for the homogenization corrector by the PDE

$$-\nabla \cdot (a^{\mathbb{R}^d}(e_i + \nabla \phi_i^{L,T})) + \frac{1}{T}\phi_i^{L,T} = 0 \qquad \text{in } [0, L\varepsilon]^d,$$

$$\phi_i^{L,T} \equiv 0 \qquad \text{on } \partial[0, L\varepsilon]^d$$

and approximates the effective coefficient a_{hom} by

$$a_{\text{hom}}e_i \approx a^{\text{RVE}}e_i := \frac{1}{\int_{[0,L\varepsilon]^d} \eta \, dx} \int_{[0,L\varepsilon]^d} \eta \, a^{\mathbb{R}^d}(e_i + \nabla \phi_i^{L,T}) \, dx,$$

where η is a smooth nonnegative weight supported in the slightly smaller box $[\frac{1}{8}L\varepsilon, (1-\frac{1}{8})L\varepsilon]^d$. In up to four spatial dimensions $d \leqq 4$, this approximation also admits error estimates of the form

$$\sqrt{\text{Var } a^{\text{RVE}}} \lesssim L^{-d/2}$$

and

$$\left| \mathbb{E}[a^{\text{RVE}}] - a_{\text{hom}} \right| \lesssim L^{-d}(\log L)^C.$$

Due to the already substantial length of the present paper, we shall limit ourselves to the analysis of the selection approach for representative volumes in the context of periodizations of the probability distribution and defer the analysis of the screening approach to a future work.

Generally speaking, in the method of representative volumes the equation for the homogenization corrector may be solved by any numerical algorithm that is feasible for the given size of the representative volume; for example, standard finite element methods may be employed for representative volumes of moderate size, while for very large representative volumes one may use appropriate instances of modern computational homogenization methods like the multiscale finite element method, heterogeneous multiscale methods, and related approaches (see for example [1,14, 29,40,60,61,71]) or the local orthogonal decomposition method by MÅLQVIST and PETERSEIM [70].

Note that besides the modern numerical homogenization methods—which are in principle applicable to any elliptic PDE involving a heterogeneous coefficient field—there have been numerous numerical works on the more specific problem of the approximation of effective coefficients in stochastic homogenization, see for example [13,32,41,42,62,72,79].

1.5. The Selection Approach for Representative Volumes by Le Bris, Legoll and Minvielle

Let us describe the selection approach for representative volumes by LE BRIS et al. [64] in more detail. The selection approach for representative volumes achieves its gain in accuracy of approximations a^{RVE} for the effective coefficient a_{hom} (as compared to the standard representative volume element method with completely

random choice of the material sample) by selecting only those realizations of the random coefficient field $a|_{[0,L\varepsilon]^d}$ which capture some important statistical properties of the coefficient field a in an exceptionally good way. For example, in the simplest setting LE BRIS et al. [64] propose to restrict one's attention to realizations of the coefficient field a for which the average on $[0, L\varepsilon]^d$ is exceptionally close to its expected value in the sense that

$$\left| \fint_{[0,L\varepsilon]^d} a \, dx - \mathbb{E}\left[\fint_{[0,L\varepsilon]^d} a \, dx\right]\right| \leq \delta L^{-d/2} \qquad (9)$$

for some $\delta \ll 1$. Note that for generic realizations of a only

$$\left| \fint_{[0,L\varepsilon]^d} a \, dx - \mathbb{E}\left[\fint_{[0,L\varepsilon]^d} a \, dx\right]\right| \sim L^{-d/2}$$

is true by the central limit theorem for the averages $\fint_{[0,L\varepsilon]^d} a \, dx$ and the finite range of dependence ε.

On a numerical level, such a selection approach typically provides an increase in computational efficiency if the accuracy is indeed increased by conditioning on the event (9): usually, the most expensive step in the computation of the approximations a^{RVE} is the computation of the homogenization corrector as the solution to the PDE (3). In contrast, the generation of random coefficient fields a and the evaluation of the average of a is typically cheap. Therefore it is often worth generating about $\frac{1}{\delta}$ independent realizations of a to obtain on average one realization of a which satisfies (9); for this single realization, the corrector equation (3) is solved numerically and the approximation a^{RVE} for the effective coefficient is computed. This strategy is also applicable to situations in which the probability distribution of the coefficient field is not known, but one has only access to a large number of samples of the coefficient field, like in applications in which one has access to data from actual material samples.

The selection criterion (9) based on the average of the coefficient field in the material sample is the first out of two selection criteria proposed by LE BRIS et al. [64]. In order to reduce the variance of a^{RVE} further, they propose to consider several such statistical quantities at the same time, for example in addition to the spatial average

$$\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a(x) \, dx,$$

the quantities

$$(\mathcal{F}_{2-\text{point}})_{i,j}(a) := \fint_{[0,L\varepsilon]^d} a \nabla v_i \cdot e_j \, dx \, dy \qquad (10)$$

for some (approximation of the) solution v_i to the constant-coefficient equation

$$-\Delta v_i = \nabla \cdot (a e_i),$$

require that all of these statistical quantities be close to their expectation at the same time. The quantities (10) arise as a second-order correction to the effective conductivity a^{RVE} in the expansion in the regime of small ellipticity contrast: Expanding the homogenization corrector ϕ_i and the approximate effective conductivity a^{RVE} as a power series in ν for the family of coefficient fields

$$a = \mathrm{Id} + \nu \hat{a},$$

we deduce

$$\phi_i = \phi_i^0 + \nu \phi_i^1 + \nu^2 \phi_i^2 + O(\nu^3)$$

with $\phi_i^0 \equiv 0$, $\phi_i^1 = v_i$, and ϕ_i^2 defined as the solution to another PDE. As a consequence, for the approximation of the effective conductivity we obtain

$$a^{\mathrm{RVE}} e_i = \fint_{[0, L\varepsilon]^d} a e_i + \nu a \nabla v_i + \nu^2 \, \mathrm{Id} \, \nabla \phi_i^2 + O(\nu^3) \, \mathrm{d}x$$

$$= \fint_{[0, L\varepsilon]^d} a e_i + \nu a \nabla v_i \, \mathrm{d}x + O(\nu^3)$$

where in the last step we have used the periodicity of ϕ_i^2. To see that the contribution of v_i is actually of second order in ν, one uses again $a = \mathrm{Id} + \nu \hat{a}$ and the periodicity of v_i.

By selecting the representative volumes by the *two* criteria (9) and

$$\left| \mathcal{F}_{2-\mathrm{point}} - \mathbb{E}\big[\mathcal{F}_{2-\mathrm{point}} \big] \right| \lessapprox \tilde{\delta} L^{-d/2} \tag{11}$$

at the same time, in the model problem of the random checkerboard with an ellipticity ratio of 5 Le Bris, Legoll, and Minvielle were able to reduce the variance of the approximations $a^{\mathrm{sel\text{-}RVE}}$ for the effective conductivity by a factor of 60, compared to the approximations a^{RVE} by the standard representative volume element method.

Another remarkable feature of the selection approach for representative volumes by Le Bris, Legoll, and Minvielle is its compatibility with the vast majority of numerical homogenization methods: As the selection approach for representative volumes operates at the level of the choice of the coefficient field a, it may be combined with essentially any numerical discretization method for the corrector problem (59). Note that there exist many numerical homogenization methods that are particularly well-adapted to certain geometries of the microstructure; the selection approach for representative volumes may be employed in most of these methods to achieve a further speedup.

The selection approach for representative volumes is only one out of several variance reduction concepts in the context of stochastic homogenization: BLANC et al. [22,23,25] have succeeded in reducing the variance by the method of antithetic variables; note that however for this approach the achievable variance reduction factor is much more limited. The method of control variates has also been demonstrated to be successful in the context of the computation of effective coefficients in stochastic homogenization [25,65].

1.6. A Brief Overview of Quantitative Stochastic Homogenization

For the sake of completeness, let us give a short overview of the tremendous progress that has been achieved in the quantitative theory of stochastic homogenization in recent years. The earliest (non-optimal) quantitative homogenization results for linear elliptic equations are due to YURINSKI [85]. A decade later, NADDAF and SPENCER [76] introduced the use of spectral gap inequalities in stochastic homogenization and derived optimal fluctuation estimates in the regime of small ellipticity contrast $||a - \text{Id}||_{L^\infty} \ll 1$, that is in a perturbative setting. Another decade later, Caffarelli and Souganidis derived the first—though only logarithmic—rates of convergence for nonlinear stochastic homogenization problems [31]. GLORIA and OTTO [53,54] and GLORIA et al. [49] succeeded in the derivation of optimal homogenization rates for discrete linear elliptic equations with i.i.d. random conductances. Subsequently, these results were generalized to elliptic equations on $\mathbb{R}^d$ and correlated probability distributions by GLORIA et al. [50,51]. For coefficient fields a whose correlations decay quickly on scales larger than $\varepsilon > 0$, these quantitative estimates for the homogenization error—that is, for the difference between the solutions to the PDE with the random coefficient field (1) and its homogenized approximation (2)—read

$$||u - u_{\text{hom}}||_{L^p} \leqq \begin{cases} \mathcal{C}(a)||f||_{L^2}\varepsilon\sqrt{|\log \varepsilon|} & \text{for } d = 2, \\ \mathcal{C}(a)||f||_{L^2}\varepsilon & \text{for } d \geqq 3, \end{cases} \tag{12}$$

with $\mathcal{C}(a)$ satisfying stretched exponential moment bounds and for suitable $p = p(d)$. ARMSTRONG and SMART [9] were the first to obtain power-law rates of convergence for nonlinear equations, deriving and employing an Avellanda–Lin type regularity estimate [12]; see also ARMSTRONG and MOURRAT [8]. Their estimates also come with optimal—almost Gaussian—stochastic moment bounds. Recently, the progress in stochastic homogenization culminated in the derivation of the optimal homogenization rates with optimal stochastic moment bounds by ARMSTRONG, et al. [5] and GLORIA and OTTO [55]: For finite range of dependence ε, a quantitative error bound for the homogenization error of the form (12) holds true with a random constant $\mathcal{C}(a)$ with almost Gaussian moments $\mathbb{E}[\exp(\mathcal{C}(a)^{2-\delta}/C(\delta))] \leqq 2$ for any $\delta > 0$.

Higher-order approximation results in terms of homogenized problems have been derived in [19–21,56,69], relying on the concept of higher-order correctors which was first used in the stochastic homogenization context in [44] to establish Liouville principles of arbitrary order in the spirit of Avellaneda and Lin's result in periodic homogenization [11]. Further works in quantitative stochastic homogenization include the analysis of nondivergence form equations [7], a regularity theory up to the boundary [45], degenerate elliptic equations [2,46], and the homogenization of parabolic equations [3,66]. Recently, ARMSTRONG and DARIO [4] and DARIO [36] succeeded in establishing quantitative homogenization for supercritical Bernoulli bond percolation on the standard lattice.

The fluctuations of the mathematical objects arising in the stochastic homogenization of linear elliptic PDEs have been the subject of a beautiful series of

works, starting with the work of NOLEN [77] and a subsequent work of GLORIA and NOLEN [52] on quantitative normal approximation for (a single component of) the approximation of the effective conductivity a^{RVE} and a work of MOURRAT and OTTO [74] on the correlation structure of fluctuations in the homogenization corrector ϕ_i. MOURRAT and NOLEN [73] have shown a quantitative normal approximation result for the fluctuations of the corrector. GU and MOURRAT [57] have derived a description of fluctuations in the solutions to the equation with random coefficient field (1). Recently, a pathwise description of fluctuations of the solutions to the equation with random coefficient field (1)—namely, in terms of deterministic linear functionals of the so-called *homogenization commutator* $\Xi := (a - a_{\text{hom}})(\text{Id} + \nabla \phi)$, a random field converging (for $\varepsilon \to 0$) towards white noise—was developed by DUERINCKX et al. [39]. The scaling limit of certain energetic quantities—related to the homogenization commutator—as well as the scaling limit of the homogenization corrector has been identified in the setting of finite range of dependence by ARMSTRONG et al. [5]. As far as quantitative normal approximation results are concerned, all of these works work under the assumption of i.i.d. coefficients (in the discrete setting) or second-order Poincaré inequalities. To the best of our knowledge, the present work provides the first quantitative description of fluctuations (though so far limited to the approximation of the effective conductivity a^{RVE}) when the decorrelation in the coefficient field is quantified by the assumption of finite range of dependence instead of functional inequalities.

Note that despite its long history [35,63,67,78], the qualitative theory of stochastic homogenization has also been a very active area of research in the past years, see for example [10,27,58,59]; however, due to the substantial length of the present manuscript we shall not provide a more detailed discussion and refer the reader to these references instead.

Notation Throughout the paper, we shall use standard notation for Sobolev spaces and weak derivatives; for a space-time function $v(x, s)$, we denote by ∇v its spatial gradient (in the weak sense) and by $\partial_s v$ its (weak) time derivative. The notation $\fint_B f \, dx := \frac{\int_B f \, dx}{\int_B 1 \, dx}$ is used for the average integral over a set B of positive but finite Lebesgue measure. The space of measurable functions f with $||f||_{L^p} := (\int_{\mathbb{R}^d} |f|^p \, dx)^{1/p} < \infty$ will be denoted by L^p. By L^p_{loc} we denote the space of functions f with $f \chi_{\{|x| \leq R\}} \in L^p$ for all $R < \infty$. We shall also use the weighted space L^p_h of functions with $||f||_{L^p_h} := (\int_{\mathbb{R}^d} |f(x)|^p h(x) \, dx)^{1/p} < \infty$ for a nonnegative measurable weight function h. By $H^1(\mathbb{R}^d)$ we denote as usual the Sobolev space of functions $v \in L^2(\mathbb{R}^d)$ with $\nabla v \in L^2(\mathbb{R}^d)$; similarly, $H^1_{loc}(\mathbb{R}^d)$ is the space of functions v with $v \in L^2_{loc}(\mathbb{R}^d)$ and $\nabla v \in L^2_{loc}(\mathbb{R}^d)$. For a Banach space X we denote by $L^p([0, T]; X)$ the usual Lebesgue–Bochner space.

As usual, we shall denote by C and c constants whose value may change from occurrence to occurrence. We are going to use the notation $\mathcal{C}(a)$ and similar expressions to denote a random constant subject to suitable moment bounds; again, the precise value of $\mathcal{C}(a)$ may change from occurrence to occurrence.

For a vector $v \in \mathbb{R}^m$ we denote by $|v|$ its Euclidean norm. We denote the identity matrix in $\mathbb{R}^{N \times N}$ by Id or Id_N. For a matrix $A \in \mathbb{R}^{m \times m}$ we shall denote by $|A|$ its natural norm $|A| := \max_{v, w \in \mathbb{R}^m, |v|=|w|=1} |v \cdot Aw|$ and by A^* its transpose

(as all our matrices are real). For $x \in \mathbb{R}^d$ we denote by $|x|_\infty = \max_i |x_i|$ its supremum norm. By $|x - y|_{\mathrm{per}}$ respectively (for sets) $\mathrm{dist}_{\mathrm{per}}(U, V)$, we denote the periodicity-adjusted distance (in the context of the torus $[0, L\varepsilon]^d$). By $|x - y|_\infty^{\mathrm{per}}$ and $\mathrm{dist}_\infty^{\mathrm{per}}(x, y)$, we denote the corresponding distances associated with the maximum norm. For a positive definite matrix A, we denote by $\kappa(A)$ its condition number.

Given a positive definite symmetric matrix $\Lambda \in \mathbb{R}^{N \times N}$, we denote the Gaussian with covariance matrix Λ by

$$\mathcal{N}_\Lambda(x) := \frac{1}{(2\pi)^{N/2}\sqrt{\det \Lambda}} \exp\left(-\frac{1}{2}\Lambda^{-1}x \cdot x\right).$$

For $\gamma > 0$, we equip the space of random variables X with stretched exponential moment $\mathbb{E}[\exp(|X|^\gamma/a)] < \infty$ for some $a = a(X) > 0$ with the norm $||X||_{\exp^\gamma} := \sup_{p \geq 1} p^{-1/\gamma}\mathbb{E}[|X|^p]^{1/p}$. For a discussion of this choice of norm, see Appendix B.

For a map $f : \mathbb{R}^N \to V$ into a normed vector space V, we denote for any $r > 0$ by $\mathrm{osc}_r f(x_0) := \sup_{x, y \in \{|x - x_0| \leq r\}} |f(x) - f(y)|_V$ its oscillation in the ball of radius r around x_0.

The conditional expectation of a random variable X given Y is denoted by $\mathbb{E}[X|Y]$.

2. Main Results

In the present work, we establish a rigorous justification of the selection approach for representative volumes by LE BRIS et al. [64] in the context of stochastic homogenization of linear elliptic PDEs for quite general probability distributions of the coefficient field $a^{\mathbb{R}^d}$. Our only assumptions on the probability distribution of the coefficient field $a^{\mathbb{R}^d} : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ are uniform ellipticity and boundedness, stationarity, and finite range of dependence, which is a standard set of assumptions in stochastic homogenization [9,55] (note that we equip the space of uniformly elliptic and bounded coefficient fields with the topology of Murat and Tartar's H-convergence [75]). Let us remark that all of our results and proofs are also valid in the case of strongly elliptic systems, upon adapting the notation in the obvious way:

(A1) *Uniform ellipticity* of a coefficient field a as usual means that there exists a positive real number $\lambda > 0$ such that almost surely we have $a(x)v \cdot v \geq \lambda|v|^2$ for almost every $x \in \mathbb{R}^d$ and every $v \in \mathbb{R}^d$. Furthermore we assume *uniform boundedness* in the sense that almost surely $|a(x)v| \leq \frac{1}{\lambda}|v|$ holds for almost every $x \in \mathbb{R}^d$ and every $v \in \mathbb{R}^d$.

(A2) *Stationarity* means that the law of the shifted coefficient field $a(\cdot + x)$ must coincide with the law of $a(\cdot)$ for every $x \in \mathbb{R}^d$. On a heuristic level, this means that "the probability distribution of a is everywhere the same" or, in other words, that the material is spatially statistically homogeneous.

(A3) *Finite range of dependence* ε means that for any two Borel sets $A, B \subset \mathbb{R}^d$ with $\mathrm{dist}(A, B) \geq \varepsilon$ the restrictions $a|_A$ and $a|_B$ must be stochastically independent. In particular, this assumption restricts the correlations in the coefficient field to the scale $\varepsilon \ll 1$.

Note that these assumptions include for example the case of a two-material composite with random (either overlapping or non-overlapping) inclusions of diameter ε, the centers distributed according to a Poisson point process (up to removal in case of overlap); see Fig. 3a. Further examples include coefficient fields $a^{\mathbb{R}^d}(x) := \xi(\tilde{a}(x))$ that arise by pointwise application of a nonlinear function $\xi : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ to a (tensor-valued) stationary Gaussian random field $\tilde{a}$ with finite range of dependence ε and integrable correlations, provided that the function ξ is Lipschitz and takes values in the set of uniformly elliptic and bounded matrices.

For the approximation of the effective coefficient a_{hom}, it is of advantage to work with a so-called *periodization* of the stationary ensemble of random coefficient fields $a^{\mathbb{R}^d}$ (employing terminology from statistical mechanics, a probability measure on the space of coefficient fields shall also be called an *ensemble* of coefficient fields). By a periodization of an ensemble of coefficient fields $a^{\mathbb{R}^d}$ we understand an ensemble of coefficient fields a which are almost surely $L\varepsilon\mathbb{Z}^d$-periodic for some $L \gg 1$ and for which the probability distribution of a on each cube of size of half the period $\frac{L\varepsilon}{2}$ coincides with the probability distribution of the original coefficient field $a^{\mathbb{R}^d}$, that is for which the probability distribution of $a|_{x+[0,L\varepsilon/2]^d}$ coincides with the distribution of $a^{\mathbb{R}^d}|_{x+[0,L\varepsilon/2]^d}$ for all $x \in \mathbb{R}^d$. For such a periodization, the condition (A3) is replaced by the following conditions $(A3_a)$, $(A3_b)$, $(A3_c)$:

$(A3_a)$ The coefficient field a is almost surely $L\varepsilon\mathbb{Z}^d$-periodic.

$(A3_b)$ There exists a *finite range of dependence* $\varepsilon > 0$ such that for any two measurable $L\varepsilon\mathbb{Z}^d$-periodic sets $A, B \subset \mathbb{R}^d$ with $\mathrm{dist}(A, B) \geqq \varepsilon$ the restrictions $a|_A$ and $a|_B$ are stochastically independent.

$(A3_c)$ For any $x_0 \in \mathbb{R}^d$ the law of the restriction $a|_{x_0+[-\frac{L\varepsilon}{4},\frac{L\varepsilon}{4}]^d}$ coincides with the corresponding law for some (non-periodic) ensemble of coefficient fields $a^{\mathbb{R}^d}$ satisfying (A1)–(A3).

Furthermore, to include examples like the random checkerboard in our analysis, we need the following notion of discrete stationarity:

(A2') We say that our probability distribution of coefficient fields a satisfies *discrete stationarity* if the law of the shifted coefficient field $a(\cdot + x)$ coincides with the law of $a(\cdot)$ for every shift $x \in \varepsilon\mathbb{Z}^d$.

Our main assumptions stated in Assumption 1 below consist of two parts. First, we assume that the probability distribution of coefficient fields $a^{\mathbb{R}^d}$ satisfies the standard assumptions from stochastic homogenization and that there exists a suitable periodization a of the probability distribution. Second, we require the statistical quantities $\mathcal{F}(a)$ to admit a "multilevel local dependence structure decomposition" as introduced in Definition 6 below. Let us remark that both the spatial average

$$\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a \, \mathrm{d}x$$

and the higher-order quantity $\mathcal{F}_{2-\mathrm{point}}(a)$ considered by LE BRIS et al. [64] as defined in (10) satisfy the conditions in Definition 6; a proof of this fact is provided in Proposition 7 below. As a consequence, both the spatial average $\mathcal{F}_{avg}(a)$ and

the higher-order quantity $\mathcal{F}_{2-\text{point}}(a)$ may be chosen as the statistical quantities by which the selection of representative volumes is performed in our main theorems Theorem 2 and Theorem 3.

Assumption 1. (*Assumptions and Notation*) Consider a probability distribution of random coefficient fields $a^{\mathbb{R}^d}$ on $\mathbb{R}^d$, $d \geq 1$, which satisfies the conditions of ellipticity, stationarity, and finite range of dependence (A1)–(A3). Let $L \geq 2$ and suppose that there exists an $L\varepsilon$-periodization a of the probability distribution of $a^{\mathbb{R}^d}$ subject to (A1), (A2), (A3$_a$)–(A3$_c$). Denote by a^{RVE} the approximation for the effective coefficient a_{hom} by the standard representative volume element method with a material sample of size $[0, L\varepsilon]^d$, that is set

$$a^{\text{RVE}} e_i := \fint_{[0,L\varepsilon]^d} a(e_i + \nabla\phi_i)\,\mathrm{d}x$$

with ϕ_i being the unique $L\varepsilon$-periodic solution with vanishing average to the corrector equation

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

Let $\mathcal{F}(a) = (\mathcal{F}_1(a), \ldots, \mathcal{F}_N(a))$ be a collection of statistical quantities of the coefficient field a which are subject to the conditions of Definition 6 with $K \leq C_0$, $B \leq C_0|\log L|^{C_0}$, and $\gamma \geq c_0$ for some $0 < c_0, C_0 < \infty$. Suppose that the covariance matrix of $\mathcal{F}(a)$ is nondegenerate and bounded in the natural scaling in the sense

$$L^{-d}\,\text{Id} \leq \text{Var}\ \mathcal{F}(a) \leq C_0 L^{-d}\,\text{Id}. \tag{13}$$

For any $1 \leq i, j \leq d$ introduce the condition number κ_{ij} of the covariance matrix of $(a_{ij}^{\text{RVE}}, \mathcal{F}(a))$

$$\kappa_{ij} := \kappa\left(\text{Var}\ (a_{ij}^{\text{RVE}}, \mathcal{F}(a))\right)$$

and the ratio $r_{\text{Var},ij}$ between the expected order of fluctuations and the actual fluctuations of the approximation a_{ij}^{RVE}

$$r_{\text{Var},ij} := \frac{L^{-d}}{\text{Var}\ a_{ij}^{\text{RVE}}}.$$

Denote by C a constant depending on d, λ, γ, N, and C_0.

Under the above assumptions, the selection approach for representative volumes to capture certain statistical properties of the material in the representative volume particularly well—as proposed by Le Bris et al. [64]—leads to the following increase in accuracy of the computed material coefficients:

Theorem 2. (Justification of the Selection Approach for Representative Volumes) *Let the assumptions and notations of Assumption 1 be in place. Denote by $a^{\text{sel-RVE}}$ the approximation for the effective coefficient a_{hom} by the selection approach for representative volumes introduced by Le Bris et al.* [64] *in the case of a representative volume of size $L\varepsilon$. Suppose that the representative volumes $a|_{[0,L\varepsilon]^d}$ are selected from the periodized probability distribution according to the criterion*

$$\left|\mathcal{F}(a) - \mathbb{E}\big[\mathcal{F}(a)\big]\right| \leqq \delta L^{-d/2} \tag{14}$$

for some $\delta \in (0,1]$. Let the selection criterion be chosen not too strict in the sense that $\delta^N \geqq CL^{-d/2}|\log L|^{C(d,\gamma,C_0)}$. Then the selection approach for representative volumes is subject to the following error analysis:

(a) *The systematic error of the approximation $a^{\text{sel-RVE}}$ satisfies the estimate*

$$\left|\mathbb{E}\big[a^{\text{sel-RVE}}\big] - a_{\text{hom}}\right| \leqq \frac{C\kappa_{ij}^{3/2}}{\delta^N}L^{-d}|\log L|^{C(d,\gamma)}. \tag{15}$$

(b) *The variance of the approximation $a^{\text{sel-RVE}}$ is estimated from above by*

$$\frac{\text{Var } a_{ij}^{\text{sel-RVE}}}{\text{Var } a_{ij}^{\text{RVE}}} \leqq 1 - (1-\delta^2)|\rho|^2 + \frac{C\kappa_{ij}^{3/2}r_{\text{Var},ij}}{\delta^N}L^{-d/2}|\log L|^{C(d,\gamma)}, \tag{16}$$

where $|\rho|^2$ is the fraction of the variance of a_{ij}^{RVE} explained by the $\mathcal{F}(a)$, that is, $|\rho|^2$ is the maximum of the squared correlation coefficient between a_{ij}^{RVE} and any linear combination of the $\mathcal{F}_n(a)$. The explained fraction of the variance is given by the formula

$$|\rho|^2 := \frac{\text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)] \cdot (\text{Var } \mathcal{F}(a))^{-1} \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}]}{\text{Var } a_{ij}^{\text{RVE}}}. \tag{17}$$

(c) *The probability that a randomly chosen coefficient field a satisfies the selection criterion* (14) *is at least*

$$\mathbb{P}\big[|\mathcal{F}(a) - \mathbb{E}\big[\mathcal{F}(a)\big]| \leqq \delta L^{-d/2}\big] \geqq c(N)\delta^N. \tag{18}$$

(d) *The systematic error and the variance of $a^{\text{sel-RVE}}$ may be estimated independently of κ_{ij} at the price of lower rate of convergence in L*

$$\left|\mathbb{E}\big[a^{\text{sel-RVE}}\big] - a_{\text{hom}}\right| \leqq \frac{C}{\delta^N}L^{-d/2-d/8}|\log L|^{C(d,\gamma)} \tag{19}$$

and

$$\frac{\text{Var } a_{ij}^{\text{sel-RVE}}}{\text{Var } a_{ij}^{\text{RVE}}} \leqq 1 - (1-\delta^2)|\rho|^2 + \frac{Cr_{\text{Var},ij}}{\delta^N}L^{-d/8}|\log L|^{C(d,\gamma)}. \tag{20}$$

The previous theorem states that the approximation of effective coefficients by the selection approach for representative volumes is essentially at least as accurate as a random selection of samples (except for a possible additional relative error of the order $CL^{-d/2}|\log L|^C$, which however converges to zero quickly as L increases), at least when measuring the mean-square error. If the selection is based on a statistical quantity $\mathcal{F}(a)$ which is capable of explaining a large part of the variance of a_{ij}^{RVE}, the selection approach achieves a much better accuracy than a random selection of samples (namely, by a factor of about $\sqrt{1 - |\rho|^2}$).

However, the previous theorem only provides a statement about the reduction of the mean-square error by the selection approach for representative volumes. A natural question is whether this reduction of the error also applies to rare events: More precisely, if we fix a small probability $p > 0$, is the bound on the error $|a_{ij}^{\text{sel-RVE}} - a_{\text{hom},ij}|$ which holds with probability $1 - p$ also improved as suggested by the variance reduction estimate (16)? The following theorem shows that this is in fact true for "moderate deviations", that is basically for probabilities $p \gtrsim \exp(-L^\beta)$ for some $\beta > 0$. More precisely, the theorem is to be read as follows: up to error terms that converge to zero as $L \to \infty$ and $s \to \infty$, the probability of $a_{ij}^{\text{sel-RVE}}$ deviating from $a_{\text{hom},ij}$ by more than s times the ideally reduced standard deviation $\sqrt{(1 - |\rho|^2)\text{Var } a_{ij}^{\text{RVE}}}$ behaves like the probability of a normal distribution deviating from its mean by more than s standard deviations, at least in some regime $s \lesssim L^{\beta/3}$.

Theorem 3. *Let the assumptions and notations of Theorem 2 be in place. Suppose in addition $L \geq C$. Then the selection approach for representative volumes leads to a reduction of the "outliers" of the probability distribution of $a^{\text{sel-RVE}}$ in the sense of the moderate-deviations-type bound*

$$
\mathbb{P}\left[\frac{|a_{ij}^{\text{sel-RVE}} - a_{\text{hom},ij}|}{\sqrt{(1 + \frac{C\delta}{\sqrt{1 - |\rho|^2}s})(1 - |\rho|^2)\text{Var } a_{ij}^{\text{RVE}} + CL^{-d-\beta}}} \geq s \right] \tag{21}
$$

$$
\leq \left(1 + \frac{C}{\delta^N L^\beta} + \frac{C\delta}{\sqrt{1 - |\rho|^2}s}\right)\mathbb{P}[|\mathcal{N}_1| \geq s] + \frac{C}{\delta^N}\exp\left(-L^\beta\right)
$$

for any $s \geq \max\left\{1, \frac{\delta}{\sqrt{1 - |\rho|^2}}\right\}$ and some $\beta = \beta(d) > 0$.

We have shown in the preceding two theorems that the selection approach for representative volumes by Le Bris et al. essentially does not increase the error; it succeeds in reducing the fluctuations of the approximations as soon as the functionals $\mathcal{F}(a)$ and the approximation a^{RVE} have a nonzero covariance.

However, as we shall show in the next theorem there exist cases in which the selection approach for representative volumes in fact fails to reduce the variance significantly, even for a "natural" statistical quantity like the average of the coefficient field

$$
\mathcal{F}(a) := \fint_{[0, L\varepsilon]^d} a \, dx.
$$

Theorem 4. (Possible Failure of the Reduction of the Variance) *Suppose that the assumptions of Theorem* 2 *hold. Then the estimate* (16) *on the reduction of the variance is sharp in the sense*

$$\frac{\operatorname{Var} a_{ij}^{\text{sel-RVE}}}{\operatorname{Var} a_{ij}^{\text{RVE}}} \geqq 1 - |\rho|^2 - \frac{C\kappa_{ij}^{3/2} r_{\text{Var},ij}}{\delta^N} L^{-d/2} |\log L|^{C(d,\gamma)}. \qquad (22)$$

Furthermore, for $d \geqq 2$ there exist $L\varepsilon$-periodic probability distributions of coefficient fields a which satisfy the conditions of ellipticity, discrete stationarity, and finite range of dependence (A1), (A2'), (A3$_a$)–(A3$_c$) with the following property: the covariance of a^{RVE} and the spatial average $\fint a$ vanishes

$$\operatorname{Cov}\left[a^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a \, \mathrm{d}x \right] = 0, \qquad (23)$$

while the fluctuations of a^{RVE} and $\fint_{[0,L\varepsilon]^d} a$ are nondegenerate in the sense that

$$\operatorname{Var} a^{\text{RVE}} \geqq cL^{-d} \operatorname{Id} \otimes \operatorname{Id},$$

$$\operatorname{Var} \fint_{[0,L\varepsilon]^d} a \, \mathrm{d}x \geqq cL^{-d} \operatorname{Id} \otimes \operatorname{Id},$$

for some universal constant c. These coefficient fields may be chosen to be of the form $a(x) = \tilde{a}(x) \operatorname{Id}$ for some scalar random field $\tilde{a}$.

As a consequence, for these probability distributions of coefficient fields the selection approach for representative volumes based on the spatial average $\fint a$ fails to efficiently reduce the variance in the sense that

$$\frac{\operatorname{Var} a_{ij}^{\text{sel-RVE}}}{\operatorname{Var} a_{ij}^{\text{RVE}}} \geqq 1 - \frac{C\kappa_{ij}^{3/2} r_{\text{Var},ij}}{\delta^N} L^{-d/2} |\log L|^{C(d,\gamma)}. \qquad (24)$$

Let us note that it is presumably not too difficult to replace the random checkerboard in our construction of the counterexample featuring (23) by random spherical inclusions distributed according to a Poisson point process (with overlaps of the inclusions). This would yield a counterexample subject to the continuous stationarity (A2).

The next theorem suggests that the failure of effective variance reduction is atypical and may be limited to rather artificial examples. For a large class of random coefficient fields—namely for coefficient fields that are obtained from a collection of iid random variables ξ_k, $k \in \varepsilon \mathbb{Z}^d$, by applying a stationary monotone map with finite range of dependence—the correlation coefficient between a^{RVE} and the average $\mathcal{F}(a) := \fint a$ is bounded from below by a positive number. Therefore, for such (ensembles of) coefficient fields both the method of special quasirandom structures and the method of control variates in fact reduce the variance by some factor $\tau < 1$ when applied with the choice $\mathcal{F}(a) := \fint a$.

Proposition 5. (Reduction of the Variance for a Large Class of Coefficient Fields)
Let $\varepsilon > 0$ and let $L \geq 2$ be an integer and let V denote some measure space. Let (Γ_k), $k \in \varepsilon\mathbb{Z}^d \cap [0, L\varepsilon)^d$, be a collection of independent identically distributed V-valued random variables, and denote by $(\tilde{\Gamma}_k)$ an independent copy. Extend Γ_k to $k \in \varepsilon\mathbb{Z}^d$ by $L\varepsilon$-periodicity. For $k \in \varepsilon\mathbb{Z}^d$ and $z \in V$, denote by $\Delta_{k,z}\Gamma$ the collection $(\tilde{\Gamma}_k)$ obtained by setting $\tilde{\Gamma}_k := z$ and $\tilde{\Gamma}_j = \Gamma_j$ for all $j \neq k$.

Let $a = a(x, \Gamma)$ be a measurable map into the uniformly elliptic $L\varepsilon$-periodic symmetric coefficient fields with the property that $a(x, \Gamma)$ depends only on the Γ_k with $|x - k|_{\mathrm{per}} \leq K\varepsilon$ for some $K \geq 1$ (in a measurable way). Suppose that the map is stationary in the sense that $a(x + y, \Gamma) = a(x, \Gamma_{\cdot + y})$ for any $y \in \varepsilon\mathbb{Z}^d$.

Suppose that the dependence of a on Γ is monotone in the sense that for every $k \in \varepsilon\mathbb{Z}^d$ and every pair $z_1, z_2 \in V$, either for all x the inequality

$$a(x, \Delta_{k,z_1}\Gamma) \geqq a(x, \Delta_{k,z_2}\Gamma)$$

holds, or for all x the reverse inequality

$$a(x, \Delta_{k,z_1}\Gamma) \leqq a(x, \Delta_{k,z_2}\Gamma)$$

holds. Suppose furthermore that there exists $\nu > 0$ such that we have the quantified monotonicity

$$\mathbb{E}\left[\sum_{k \in \varepsilon\mathbb{Z}^d \cap [0, L\varepsilon)^d} \sqrt{\varepsilon^{-d} \int_{[0, L\varepsilon]^d} |(a - a(\Delta_{k,\tilde{\Gamma}_k}\Gamma))\xi \cdot \xi|\, \mathrm{d}x} \left(a(x, \Gamma) - a(x, \Delta_{k,\tilde{\Gamma}_k}\Gamma)\right)_+^{1/2} \,\Big|\, \Gamma \right]$$

$$\geqq \nu \,\mathrm{Id} \tag{25}$$

for all $x \in [0, L\varepsilon)^d$ and all Γ, where $\left(a(x, \Gamma) - a(x, \Delta_{k,\tilde{\Gamma}_k}\Gamma)\right)_+^{1/2}$ denotes the matrix square root and where $\tilde{\Gamma}$ denotes an independent copy of Γ.

Then the probability distribution of $a = a(x, \Gamma)$ satisfies the conditions of ellipticity, periodicity, and finite range of dependence (A1), (A3$_a$), and (A3$_b$) (with ε replaced by $4K\varepsilon$), as well as the discrete stationarity (A2'). Furthermore, for such coefficient fields a the correlation between $\xi \cdot a^{\mathrm{RVE}}\xi$ (where $\xi \in \mathbb{R}^d$ is any nonzero vector) and the average

$$\mathcal{F}_{avg}(a) := \fint_{[0, L\varepsilon]^d} \xi \cdot a\xi \, \mathrm{d}x$$

is bounded from below by a positive number in the sense

$$\rho = \frac{\mathrm{Cov}[a_{ij}^{\mathrm{RVE}}, \mathcal{F}_{avg}(a)]}{\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}\ \mathrm{Var}\ \mathcal{F}_{avg}(a)}} \geqq \frac{\nu^2}{C(d, \lambda, K)}.$$

In the statements of our main theorems, we have made use of the following notion of "multilevel local dependence decomposition"; this structure will also be at the heart of the proof of our main results (an illustration of this decomposition is provided in Fig. 4):
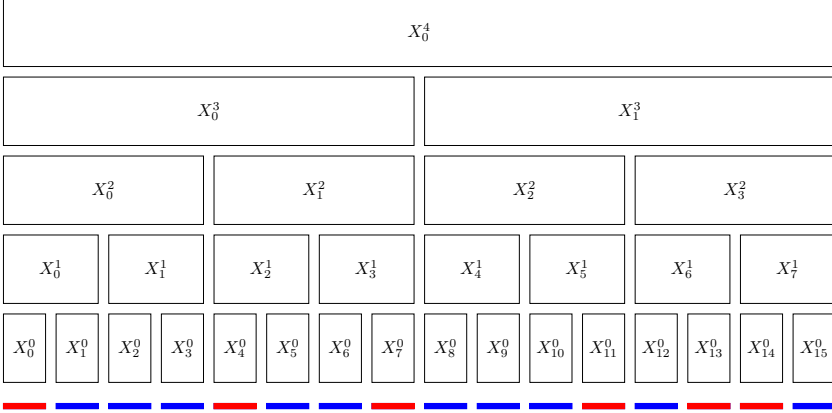
Fig. 4. An illustration of the "multilevel local dependence structure" introduced in Definition 6 (in a one-dimensional setting). At the bottom, a sample of the random coefficient field a is depicted; the X_y^k may depend not only on the values of the coefficient field directly below their box, but on the coefficient field in a region that is wider by a factor of $K \log L$

Definition 6. (*Sums of Random Variables with Multilevel Local Dependence Structure*) Let $d \geqq 1$, $N \in \mathbb{N}$, $\varepsilon > 0$, and $L \geqq 2$. Consider a probability distribution of coefficient fields a on $\mathbb{R}^d$ subject to the assumptions of ellipticity and boundedness, stationarity, and finite range of dependence ε (A1), (A2), and (A3), or the periodization of such an ensemble subject to the conditions (A1), (A2), and (A3$_a$) - (A3$_c$). Let $X = X(a)$ be an $\mathbb{R}^N$-valued random variable.

We then say that X is a sum of random variables with multilevel local dependence if there exist random variables $X_y^m = X_y^m(a)$, $0 \leqq m \leqq 1 + \log_2 L$ and $y \in 2^m \varepsilon \mathbb{Z}^d \cap [0, L\varepsilon)^d$, and constants $K \geqq 1$, $\gamma \in (0, 2]$, and $B \geqq 1$ with the following properties:

- The random variable $X_y^m(a)$ only depends on $a|_{y + K \log L \, [-2^m \varepsilon, 2^m \varepsilon]^d}$. More precisely, $X_y^m(a)$ is a measurable function of $a|_{y + K \log L \, [-2^m \varepsilon, 2^m \varepsilon]^d}$ equipped with the topology of H-convergence.
- We have

$$X = \sum_{m=0}^{1+\log_2 L} \sum_{y \in 2^m \varepsilon \mathbb{Z}^d \cap [0, L\varepsilon)^d} X_y^m.$$

- The random variables X_y^m satisfy the bound

$$||X_y^m||_{\exp^\gamma} \leqq B L^{-d}. \tag{26}$$

The next proposition shows that the approximation a^{RVE} of the effective coefficient by the method of representative volumes may indeed be rewritten as a sum of random variables with a multilevel local dependence structure. We establish the same result for the spatial average of the coefficient field $\mathcal{F}_{avg}(a) := \fint_{[0, L\varepsilon]^d} a \, \mathrm{d}x$ and the second-order term $\mathcal{F}_{2-\text{point}}(a)$ in the low ellipticity contrast expansion of a^{RVE} given by (10).

Furthermore, the last result of the next proposition shows that the fraction of the variance of a^{RVE} that is explained by the statistical quantities $\mathcal{F}_{avg}(a)$ and $\mathcal{F}_{2-\mathrm{point}}(a)$—that is, the gain in accuracy achieved by the selection approach for representative volumes when employing these statistical quantities—stabilizes as the size L of the representative volume increases; more precisely, it converges to some limit with rate $L^{-d/2}|\log L|^C$.

Proposition 7. *Let the assumptions (A1), (A2), (A3$_a$)–(A3$_c$) be satisfied, that is consider the periodization of a stationary ensemble of random coefficient fields. For any coefficient field a, denote by ϕ_i the unique (up to additions of constants) periodic solution to the corrector equation*

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

Then the approximation a^{RVE} of the effective coefficient a_{hom} by the representative volume element method, given by

$$a^{\mathrm{RVE}}e_i := \fint_{[0,L\varepsilon]^d} a(e_i + \nabla\phi_i)\,\mathrm{d}x,$$

is a sum of a family of random variables with multilevel local dependence. More precisely, a^{RVE} satisfies the criteria of Definition 6 for any $\gamma < 1$ with $K := C(d,\lambda)$ and $B := C(d,\gamma,\lambda)|\log L|^{C(d,\gamma)}$.

Furthermore, the spatial average

$$\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a\,\mathrm{d}x$$

is also a sum of a family of random variables with multilevel local dependence. The criteria of Definition 6 are satisfied by $\mathcal{F}_{avg}(a)$ for any $\gamma < \infty$ with $K := C(d)$ and $B := C(d,\gamma)$.

Additionally, the second-order correction to the effective conductivity in the setting of small ellipticity contrast $\mathcal{F}_{2-\mathrm{point}}$, given by

$$\mathcal{F}_{2-\mathrm{point}}(a) := -\fint_{[0,L\varepsilon]^d} a\nabla v_i \cdot e_j\,\mathrm{d}x, \tag{27}$$

with v_i denoting the solution to

$$-\Delta v_i = \nabla \cdot (a\nabla e_i), \tag{28}$$

is a sum of random variables with multilevel local dependence structure: the random variable $\mathcal{F}_{2-\mathrm{point}}(a)$ satisfies the criteria of Definition 6 for any $\gamma < 1$ with $K := C(d,\lambda)$ and $B := C(d,\gamma,\lambda)|\log L|^{C(d,\gamma)}$.

Finally, the rescaled variances and covariances of a^{RVE} and the statistical quantities $\mathcal{F}_{avg}(a)$ and $\mathcal{F}_{2-\mathrm{point}}(a)$ converge as $L \to \infty$. There exist positive

semidefinite matrices V_{RVE}, V_{avg}, $V_{2-\text{point}}$ *and matrices* $V_{c,\text{RVE},avg}$, $V_{c,\text{RVE},2-\text{point}}$, $V_{c,avg,2-\text{point}}$ *independent of L such that the estimates*

$$|L^d \text{Var } a^{\text{RVE}} - V_{\text{RVE}}| \leqq CL^{-d/2}(\log L)^C,$$

$$|L^d \text{Var } \mathcal{F}_{avg}(a) - V_{avg}| \leqq CL^{-d/2}(\log L)^C,$$

$$|L^d \text{Var } \mathcal{F}_{2-\text{point}}(a) - V_{2-\text{point}}| \leqq CL^{-d/2}(\log L)^C,$$

and

$$|L^d \text{Cov}[a^{\text{RVE}}, \mathcal{F}_{avg}(a)] - V_{c,\text{RVE},avg}| \leqq CL^{-d/2}(\log L)^C,$$

$$|L^d \text{Cov}[a^{\text{RVE}}, \mathcal{F}_{2-\text{point}}(a)] - V_{c,\text{RVE},2-\text{point}}| \leqq CL^{-d/2}(\log L)^C,$$

$$|L^d \text{Cov}[\mathcal{F}_{avg}(a), \mathcal{F}_{2-\text{point}}(a)] - V_{c,avg,2-\text{point}}| \leqq CL^{-d/2}(\log L)^C,$$

hold true.

It is interesting to compare our approach on quantitative normal approximation of a^{RVE} with concepts employed in the derivation of optimal error estimates in stochastic homogenization [5,6,55]. A central theme in [5] is the approximate additivity of certain energetic quantities: the energy quantity on a certain scale may approximately be written as a sum of the energy quantities on smaller scales, allowing for an application of the central limit theorem. In [55], the application of the central limit theorem is facilitated by the homogenization of the flux propagation in the parabolic semigroup associated with the random elliptic operator. In our context, while we also introduce an additive decomposition of a^{RVE}, we do not require the summands to be of the same structure as a^{RVE} and allow for a multilevel structure. This enables us to derive an optimal-order normal approximation result for the fluctuations.

Note that in [5,6] a certain localization property of the considered energy quantity has been established. In principle, sufficiently strong localization properties of a random field allow for a multilevel decomposition of (linear functionals of) the random field in the sense of Definition 6 and therefore for an application of our quantitative normal approximation result in Theorem 9; see, in particular, the proof of [43, Theorem 2] for such a construction. However, the locality of the energy quantity established in [5,6] is non-optimal and in general not sufficient for our purposes. In the forthcoming work [37], an optimal-order localization result for (linear functionals of) the homogenization commutator $\Xi := (a - a_{\text{hom}})(\text{Id} + \nabla\phi)$ will be provided, implying an optimal-order normal approximation result.

3. Strategy of the Proof and Intermediate Results

Our main result relies on a quantitative normal approximation result for the joint probability distribution of the approximation of the effective conductivity a^{RVE} and auxiliary random variables $\mathcal{F}(a)$ like the spatial average $\fint_{[0,L\varepsilon]^d} a \, dx$. The distance of the probability distribution to a multivariate Gaussian will be quantified through the following notion of distance between probability measures. Note that

this distance is a standard choice in the theory of multivariate normal approximation, see for example [33] and the references therein.

Definition 8. Given a symmetric positive definite matrix $\Lambda \in \mathbb{R}^{N \times N}$ and some $\bar{L} < \infty$, we consider the classes $\Phi_\Lambda^{\bar{L}}$ of functions $\phi : \mathbb{R}^N \to \mathbb{R}$ subject to the following properties:

- ϕ is smooth and its first derivative is bounded in the sense $|\nabla \phi(x)| \leqq \bar{L}$ for all $x \in \mathbb{R}^N$.
- For any $r > 0$ and any $x_0 \in \mathbb{R}^N$, we have

$$\int_{\mathbb{R}^N} \mathrm{osc}_r \phi(x) \, \mathcal{N}_\Lambda(x - x_0) \, \mathrm{d}x \leqq r, \tag{29}$$

where $\mathrm{osc}_r \phi(x)$ is the oscillation of ϕ defined as

$$\mathrm{osc}_r \phi(x) := \sup_{|z| \leqq r} \phi(x + z) - \inf_{|z| \leqq r} \phi(x + z)$$

and where

$$\mathcal{N}_\Lambda(x) := \frac{1}{(2\pi)^{N/2} \sqrt{\det \Lambda}} \exp\left(-\frac{1}{2} \Lambda^{-1} x \cdot x \right).$$

The class Φ_Λ is defined as

$$\Phi_\Lambda := \bigcup_{\bar{L} > 0} \Phi_\Lambda^{\bar{L}}.$$

Furthermore, we introduce the distance $\mathcal{D}$ between the law of an $\mathbb{R}^N$-valued random variable X and the N-variate Gaussian $\mathcal{N}_\Lambda$ as

$$\mathcal{D}(X, \mathcal{N}_\Lambda) := \sup_{\phi \in \Phi_\Lambda} \left(\mathbb{E}[\phi(X)] - \int_{\mathbb{R}^N} \phi(x) \mathcal{N}_\Lambda(x) \, \mathrm{d}x \right). \tag{30}$$

Note that defining the distance $\mathcal{D}$ with the class of functions Φ_Λ^1 instead of Φ_Λ would lead to the 1-Wasserstein distance. The distance $\mathcal{D}$ is a stronger distance than the 1-Wasserstein distance. The 1-Wasserstein distance is defined by taking the supremum in (30) only over all functions ϕ which are 1-Lipschitz. In contrast, the condition (29) corresponds more or less to a slightly stronger condition than an L_{loc}^1-type bound for $\nabla \phi$: It in particular implies by letting $r \to 0$

$$\int_{\mathbb{R}^N} |\nabla \phi|(x) \mathcal{N}_\Lambda(x - x_0) \, \mathrm{d}x \leqq 1 \tag{31}$$

for any $x_0 \in \mathbb{R}^N$.

It is well-known that Stein's method of normal approximation allows one to establish a quantitative result on normal approximation for sums of random variables with local dependence structure, see for example [33,34,80] and the references therein. However, the approximation of the effective coefficient a^{RVE}—that is, the random variable a^{RVE} as defined by (4)—features global dependencies. It is shown

in Proposition 7 that a^{RVE} may nevertheless be approximated by a sum of random variables with a *multilevel local dependence* structure. We then employ the following quantitative central limit theorem for sums of vector-valued random variables with a multilevel local dependence structure, which is not covered by the normal approximation results for sums of random variables with a given dependency graph in the literature and which is established in the companion article [43]:

Theorem 9. ([43, Theorem 4]) *Consider a probability distribution of uniformly elliptic and bounded coefficient fields a on $\mathbb{R}^d$ or a periodization of such a probability distribution, and suppose that assumptions (A1)–(A3) respectively (A1), (A2), $(A3_a)$–$(A3_c)$ are satisfied. Let $X = X(a)$ be a random variable that is a sum of random variables with multilevel local dependence in the sense of Definition 6. Then the law of the random variable X is close to a multivariate Gaussian in the sense that*

$$\mathcal{D}(X - \mathbb{E}[X], \mathcal{N}_\Lambda) \leqq C(d, \gamma, N, K) B^3 (\log L)^{C(d,\gamma)} \big(L^{-d} |\Lambda^{1/2}| |\Lambda^{-1/2}|^3 \big) L^{-d}, \tag{32}$$

where $\Lambda := \mathrm{Var}\ X$ and where the constant $C(d, \gamma, N, K)$ depends in a polynomial way on d, N, and K.

Furthermore, we have, for any symmetric positive definite $\Lambda \in \mathbb{R}^{N \times N}$ with $\Lambda \geqq \mathrm{Var}\ X$ and $|\Lambda - \mathrm{Var}\ X| \leqq L^{-d}$,

$$\mathcal{D}(X - \mathbb{E}[X], \mathcal{N}_\Lambda) \leqq C(d, \gamma, N, K) B^3 (\log L)^{C(d,\gamma)} \big(L^{-d} |\Lambda^{1/2}| |\Lambda^{-1/2}|^3 \big) L^{-d} \tag{33}$$

$$+ C(d, N)(\log L)^{C(d,\gamma)} |\Lambda - \mathrm{Var}\ X|^{1/2},$$

providing a better bound in the case of degenerate covariance matrices $\mathrm{Var}\ X$.

Our result on moderate deviations of the probability distribution of $a^{\mathrm{sel\text{-}RVE}}$ is based on the following simple general moderate deviations result for sums of random variables with multilevel local dependence structure:

Theorem 10. ([43, Theorem 5]) *Consider an ensemble of coefficient fields a on $\mathbb{R}^d$, $d \geqq 1$, or its periodization for some $L \geqq 1$, subject to the conditions (A1)–(A3) respectively (A1), (A2), and $(A3_a)$–$(A3_c)$. Let $X = X(a)$ be a random variable that may be written as a sum of random variables with multilevel local dependence structure $X = \sum_{m=0}^{1+\log_2 L} \sum_{i \in 2^m \varepsilon \mathbb{Z}^d \cap [0, L\varepsilon)^d} X_i^m$ in the sense of Definition 6.*

Then there exists $\beta = \beta(d, \gamma) > 0$ and a positive definite symmetric matrix $\Lambda \in \mathbb{R}^{N \times N}$ with $|\Lambda - \mathrm{Var}\ X| \leqq C(d, \gamma, N, K) B^2 L^{-2\beta} L^{-d}$ such that for any measurable $A \subset \mathbb{R}^N$ we have the estimate

$$\mathbb{P}\big[X \in A\big] \leqq \int_{\{x \in \mathbb{R}^N : \mathrm{dist}(x,A) \leqq L^{-\beta} L^{-d/2}\}} \mathcal{N}_\Lambda(x)\, \mathrm{d}x$$

$$+ C(d, \gamma, N, K) \exp\Big(- \frac{c}{B^C} L^{2\beta} \Big).$$

4. Justification of the Selection Approach for Representative Volumes

We now provide the proof of our main result—the error estimates for the selection approach for representative volumes by Le Bris et al. [64]—which is stated in Theorems 2 and 3.

The idea for the proof of all statements of Theorem 2 is that Theorem 9 enables us in conjunction with Proposition 7 to approximate the joint probability distribution of a^{RVE} and $\mathcal{F}(a)$ by a multivariate Gaussian with the same covariance matrix. The probability distribution of $a^{\text{sel-RVE}}$ arises as the probability distribution of a^{RVE} conditioned on the event (14). As a consequence, the probability distribution of $a^{\text{sel-RVE}}$ may be approximated by the marginal of the conditional probability distribution of an ideal multivariate Gaussian. The results of Theorem 2 on the probability distribution of $a^{\text{sel-RVE}}$ are then a consequence of corresponding properties of multivariate normal distributions.

Proof of Theorem 2. For the proof of the theorem we may assume without loss of generality that $\mathbb{E}[\mathcal{F}(a)] = 0$. Throughout the proof, the constants c and C may depend on d, λ, N, γ, c_0, and C_0, if not otherwise stated.

Recall that the probability distribution of $a^{\text{sel-RVE}}$ is given by the probability distribution of a^{RVE} conditioned on the event (14). Theorem 9 and Proposition 7 entail that the joint probability distribution of any component a_{ij}^{RVE} of a^{RVE} and $\mathcal{F}(a)$ is close to a multivariate Gaussian $\mathcal{N}_{\text{Var } (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(\cdot - \mathbb{E}[a_{ij}^{\text{RVE}}], \cdot)$. As a consequence of this result, the probability distribution of $a_{ij}^{\text{sel-RVE}}$ may be approximated in a quantitative sense by the first-variable marginal of the conditional distribution of $\mathcal{N}_{\text{Var } (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(\cdot - \mathbb{E}[a_{ij}^{\text{RVE}}], \cdot)$ given the event $|\mathcal{F}(a)| \leqq \delta L^{-d/2}$. As we shall show below, the latter marginal probability distribution has the density

$$
\mathcal{M}^\delta(x) := \frac{1}{p} \int_{\mathbb{R}^N} \mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}}} \left(x - \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1} y - \mathbb{E}[a_{ij}^{\text{RVE}}] \right)
\tag{34}
$$

$$
\times \, \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y) \, \mathrm{d}y,
$$

where the renormalization factor p is given by

$$
p = \int_{\mathbb{R}} \int_{\mathbb{R}^N} \chi_{\{|y| \leqq \delta L^{-d/2}\}}(y) \mathcal{N}_{\text{Var } (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(x, y) \, \mathrm{d}y \, \mathrm{d}x,
$$

and where the unexplained variance Var $a_{ij}^{\text{RVE}}|_{\text{unexpl}}$ (that is the variance of a_{ij}^{RVE} which is not explained by the $\mathcal{F}_n(a)$) is given by

$$
\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}} = \text{Var } a_{ij}^{\text{RVE}} - \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1} \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}].
$$

The assertions (15) and (16) on the systematic error and the variance reduction in Theorem 2 will be a consequence of the lower bound (18) on the probability of a random coefficient field satisfying the selection criterion, the related lower bound

$$
\int_{\mathbb{R}^N} \int_{\mathbb{R}} \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(x, y) \, \mathrm{d}x \, \mathrm{d}y \geqq c(N) C_0^{-N/2} \delta^N,
\tag{35}
$$

the stretched exponential moment bounds for any $\gamma < 1/2$,

$$||a^{\text{RVE}} - \mathbb{E}[a^{\text{RVE}}]||_{\exp^\gamma} \leqq C(d, \lambda, \gamma) L^{-d/2} |\log L|^C, \tag{36a}$$

$$||\mathcal{N}_{\text{Var }(a^{\text{RVE}}, \mathcal{F}(a))}||_{\exp^\gamma} \leqq C(d, \lambda, \gamma, C_0) L^{-d/2} |\log L|^C, \tag{36b}$$

and the approximation result of the distribution of $a_{ij}^{\text{sel-RVE}}$ by $\mathcal{M}^\delta$

$$\left| \mathbb{E}[\tilde{\phi}(a_{ij}^{\text{sel-RVE}})] - \int_{\mathbb{R}} \tilde{\phi}(x) \mathcal{M}^\delta(x) \, dx \right| \leqq \frac{C \kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)} \tag{37}$$

for any continuous $\tilde{\phi} : \mathbb{R} \to \mathbb{R}$ satisfying

$$|\tilde{\phi}| \leqq L^{-d/2} \tag{38a}$$

and

$$\int_{\mathbb{R}} \text{osc}_r \tilde{\phi}(x) \mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}}}(x - x_0) \, dx \leqq r \tag{38b}$$

for all $r > 0$ and all $x_0 \in \mathbb{R}$. To obtain the κ-independent estimates (19) and (20), the bound (37) is replaced by

$$\left| \mathbb{E}[\tilde{\phi}(a_{ij}^{\text{sel-RVE}})] - \int_{\mathbb{R}} \tilde{\phi}(x) \mathcal{M}^\delta(x) \, dx \right| \leq \frac{C}{\delta^N} L^{-d/2-d/8} |\log L|^{C(d,\gamma)}. \tag{39}$$

We defer the proof of (18) and (37) (as well as (39)) to the last step and first demonstrate that these estimates entail the assertions (15) and (16) of our theorem.

Step 1: Estimate on the systematic error. In order to derive the estimate on the systematic error (15), we first use the formula (34) and Fubini's theorem to see that

$$\int x \, \mathcal{M}^\delta(x) \, dx \tag{40}$$
$$= \frac{1}{p} \int_{\mathbb{R}^N} \left(\mathbb{E}[a_{ij}^{\text{RVE}}] + \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1} y \right)$$
$$\times \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y) \, dy$$
$$= \mathbb{E}[a_{ij}^{\text{RVE}}],$$

where in the second step we have used the symmetry of the Gaussian $\mathcal{N}_{\text{Var } \mathcal{F}(a)}$. In other words, if the probability distribution of $(a^{\text{RVE}}, \mathcal{F}(a))$ were an ideal multivariate Gaussian, we would have the perfect equality $\mathbb{E}[a^{\text{sel-RVE}}] = \mathbb{E}[a^{\text{RVE}}]$.

We would now like to transfer the property (40) (up to an error) from $\mathcal{M}^\delta$ to our actual probability distribution $a^{\text{sel-RVE}}$ by choosing $\tilde{\phi}(x) := x$ in the estimate (37). However, this choice is not possible due to the upper bound on $\tilde{\phi}$ in (38a). Instead, for some cutoff factor $B_c \geqq 1$ we consider the function $\tilde{\phi}(x) = \min\{\max\{x - \mathbb{E}[a_{ij}^{\text{RVE}}], -B_c L^{-d/2}\}, B_c L^{-d/2}\}$. Note that for this choice of $\tilde{\phi}$ we have $|\nabla \tilde{\phi}| \leqq 1$

and $|\tilde{\phi}| \leqq B_c L^{-d/2}$. As a consequence, $\frac{1}{B_c}\tilde{\phi}$ satisfies (38) and hence is an admissible choice in (37), which gives, by (40),

$$
\begin{aligned}
&\left| \mathbb{E}[a_{ij}^{\text{sel-RVE}}] - \mathbb{E}[a_{ij}^{\text{RVE}}] \right| \\
&= \left| \mathbb{E}[a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]] - \int_{\mathbb{R}} (x - \mathbb{E}[a_{ij}^{\text{RVE}}]) \, \mathcal{M}^\delta(x) \, dx \right| \\
&\leqq \mathbb{E}\left[\left| (a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]) - \tilde{\phi}(a_{ij}^{\text{sel-RVE}}) \right| \right] \\
&\quad + \int_{\mathbb{R}} |(x - \mathbb{E}[a_{ij}^{\text{RVE}}]) - \tilde{\phi}(x)| \, \mathcal{M}^\delta(x) \, dx \\
&\quad + \left| \mathbb{E}[\tilde{\phi}(a_{ij}^{\text{sel-RVE}})] - \int_{\mathbb{R}} \tilde{\phi}(x) \, \mathcal{M}^\delta(x) \, dx \right| \\
&\overset{(37)}{\leqq} \mathbb{E}\left[\left(|a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]| - B_c L^{-d/2} \right)_+ \right] \\
&\quad + \int_{\mathbb{R}} \left(|x - \mathbb{E}[a_{ij}^{\text{RVE}}]| - B_c L^{-d/2} \right)_+ \mathcal{M}^\delta(x) \, dx \\
&\quad + B_c \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)}.
\end{aligned}
$$

Using first the lower bounds (18) and (35) and the representation (44) and then in the next step Hölder's inequality, the previous estimate implies

$$
\begin{aligned}
&\left| \mathbb{E}[a_{ij}^{\text{sel-RVE}}] - \mathbb{E}[a_{ij}^{\text{RVE}}] \right| \\
&\leqq \frac{C}{\delta^N} \mathbb{E}\left[\left(|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]| - B_c L^{-d/2} \right)_+ \right] \\
&\quad + \frac{C(N)}{\delta^N} \int_{\mathbb{R}} \int_{\mathbb{R}^N} \left(|x - \mathbb{E}[a_{ij}^{\text{RVE}}]| - B_c L^{-d/2} \right)_+ \\
&\qquad\qquad \times \mathcal{N}_{\text{Var } (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(x - \mathbb{E}[a_{ij}^{\text{RVE}}], y) \, dy \, dx \\
&\quad + B_c \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)} \\
&\leqq \frac{C}{\delta^N} \mathbb{E}\left[|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]|^2 \right]^{1/2} \mathbb{P}\left[|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]| \geqq B_c L^{-d/2} \right]^{1/2} \\
&\quad + \frac{C(N)}{\delta^N} \mathbb{E}\left[|\mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}}|^2 \right]^{1/2} \mathbb{P}\left[|\mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}}| \geqq B_c L^{-d/2} \right]^{1/2} \\
&\quad + B_c \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)}.
\end{aligned}
$$

This yields, by Lemma 19b and the bounds (36a) and (36b),

$$
\begin{aligned}
&\left| \mathbb{E}[a_{ij}^{\text{sel-RVE}}] - \mathbb{E}[a_{ij}^{\text{RVE}}] \right| \\
&\leqq \frac{C(N)}{\delta^N} \exp\left(-c \left(\frac{B_c}{|\log L|^C} \right)^\gamma \right) + B_c \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)}.
\end{aligned}
$$

Choosing $B_c := C|\log L|^{C(\gamma)}$, we deduce

$$|\mathbb{E}[a_{ij}^{\text{sel-RVE}}] - \mathbb{E}[a_{ij}^{\text{RVE}}]| \leq \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)}. \tag{41}$$

Plugging in the bound for the systematic error of the standard representative volume element method $|\mathbb{E}[a^{\text{RVE}}] - a_{\text{hom}}| \leq CL^{-d}|\log L|^C$ from [55] (note that this estimate for the systematic error of the standard representative volume element method may also be derived by slightly modifying the proof of our Proposition 7), we obtain (15). Repeating the previous proof but replacing the use of the estimate (37) by (39), we obtain (19).

Step 2: Proof of the variance reduction estimate. To prove the variance estimate (16), we proceed similarly and define for a cutoff factor $B_c \geq 1$ the function $\phi(x) := \min\{(x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2, B_c^2 L^{-d}\}$. Note that this function satisfies the global bounds $|\nabla\phi| \leq 2B_c L^{-d/2}$ and $|\phi| \leq B_c^2 L^{-d}$. Thus, $\frac{1}{2B_c^2 L^{-d/2}}\phi$ satisfies (38) and is therefore an admissible choice in (37), yielding

$$\left| \mathbb{E}\big[\min\{(a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}])^2, B_c^2 L^{-d}\}\big] \right. \tag{42}$$

$$\left. - \int_{\mathbb{R}} \min\{(x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2, B_c^2 L^{-d}\} \mathcal{M}^\delta(x)\, dx \right|$$

$$\leq 2B_c^2 L^{-d/2} \cdot \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-d} |\log L|^{C(d,\gamma)}.$$

The tails (subject to truncation in our choice of ϕ) can be estimated by

$$\mathbb{E}\big[|(a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 - \phi(a_{ij}^{\text{sel-RVE}})|\big]$$

$$+ \int_{\mathbb{R}} |(x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 - \phi(x)|\, \mathcal{M}^\delta(x)\, dx$$

$$\leq \mathbb{E}\big[\big(|a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]|^2 - B_c^2 L^{-d}\big)_+\big]$$

$$+ \int_{\mathbb{R}} \big(|x - \mathbb{E}[a_{ij}^{\text{RVE}}]|^2 - B_c^2 L^{-d}\big)_+ \mathcal{M}^\delta(x)\, dx$$

$$\leq \frac{C}{\delta^N} \mathbb{E}\big[\big(|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]|^2 - B_c^2 L^{-d}\big)_+\big]$$

$$+ \frac{C}{\delta^N} \int_{\mathbb{R}} \int_{\mathbb{R}^N} \big(|x - \mathbb{E}[a_{ij}^{\text{RVE}}]|^2 - B_c^2 L^{-d}\big)_+ \mathcal{N}_{\text{Var}\ (a_{ij}^{\text{RVE}}, \mathcal{F}(a))}(x, y)\, dy\, dx,$$

where in the last step we have used (18), (35), and (44). Applying Hölder's inequality, we obtain

$$\mathbb{E}\big[|(a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 - \phi(a_{ij}^{\text{sel-RVE}})|\big]$$

$$+ \int_{\mathbb{R}} |(x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 - \phi(x)|\, \mathcal{M}^\delta(x)\, dx$$

$$\leq \frac{C}{\delta^N} \mathbb{E}\big[|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]|^4\big]^{1/2} \mathbb{P}\big[|a_{ij}^{\text{RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}]| \geq B_c L^{-d/2}\big]^{1/2}$$

$$+ \frac{C}{\delta^N} \mathbb{E}[|\mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}}|^4]^{1/2} \cdot \mathbb{P}[|\mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}}| \geqq B_c L^{-d/2}]^{1/2}$$

$$\leqq \frac{C}{\delta^N} \exp\left(-c\left(\frac{B_c}{|\log L|^C}\right)^\gamma\right),$$

where in the last step we have used Lemma 19b and the bounds (36a) and (36b).

Combining this estimate with (42) and choosing $B_c := C|\log L|^{C(d,\gamma)}$, we infer

$$\left| \mathbb{E}[(a_{ij}^{\text{sel-RVE}} - \mathbb{E}[a_{ij}^{\text{RVE}}])^2] - \int (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 \mathcal{M}^\delta(x)\, \mathrm{d}x \right| \tag{43}$$

$$\leqq \frac{C\kappa_{ij}^{3/2}}{\delta^N} L^{-3d/2} |\log L|^{C(d,\gamma)}.$$

In other words, the variance of $a_{ij}^{\text{sel-RVE}}$ is determined up to an error by the variance of the probability distribution $\mathcal{M}^\delta$. To estimate the latter, a straightforward computation yields

$$\int (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 \mathcal{M}^\delta(x)\, \mathrm{d}x$$

$$\overset{(34)}{=} \frac{1}{p} \int_{\mathbb{R}^N} \int_{\mathbb{R}} \mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}}} \left(x - \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1}y - \mathbb{E}[a_{ij}^{\text{RVE}}]\right)$$

$$\times \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y) \cdot (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2\, \mathrm{d}x\, \mathrm{d}y$$

$$= \frac{1}{p} \int_{\mathbb{R}^N} \int_{\mathbb{R}} \left(\tilde{x} + \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1}y\right)^2 \mathcal{N}_{\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}}}(\tilde{x})\, \mathrm{d}\tilde{x}$$

$$\times \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y)\, \mathrm{d}y.$$

By the symmetry of the set $\{|y| \leqq \delta L^{-d/2}\}$ and the probability density $\mathcal{N}_{\text{Var } \mathcal{F}(a)}(y)$ we have $\int_{\mathbb{R}^N} y \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y)\, \mathrm{d}y = 0$. As a consequence, we get

$$\int (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 \mathcal{M}^\delta(x)\, \mathrm{d}x$$

$$= \frac{1}{p} \int_{\mathbb{R}^N} \left(\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}} + \left(\text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1}y\right)^2\right)$$

$$\times \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y)\, \mathrm{d}y$$

$$\overset{(13)}{\leqq} \left(\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}} + \delta^2 \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1} \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}]\right)$$

$$\times \frac{1}{p} \int_{\mathbb{R}^N} \chi_{\{|y| \leqq \delta L^{-d/2}\}} \mathcal{N}_{\text{Var } \mathcal{F}(a)}(y)\, \mathrm{d}y$$

$$= \left(\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}} + \delta^2 \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var } \mathcal{F}(a))^{-1} \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}]\right)$$

$$= \left(1 - (1 - \delta^2)|\rho|^2\right) \text{Var } a_{ij}^{\text{RVE}}.$$

Together with (43), this entails (16). To prove (20), we repeat the proof of (43) and just replace the use of (37) in the proof of (43) by (39).

Note that the lower bound (22) on the variance given in Theorem 4 follows also from the estimates (43) and (15) and the lower bound $\int (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 \mathcal{M}^\delta(x)\, dx \geqq (1 - |\rho|^2)\text{Var}\ a_{ij}^{\text{RVE}}$, the latter of which is derived analogously to the upper bound $\int (x - \mathbb{E}[a_{ij}^{\text{RVE}}])^2 \mathcal{M}^\delta(x)\, dx \leqq (1 - (1 - \delta^2)|\rho|^2)\text{Var}\ a_{ij}^{\text{RVE}}$.

Step 3: The probability density of the reference distribution. For the purpose of this subsection, introduce the abbreviation for the covariance matrix

$$\Lambda := \text{Var}\ (a_{ij}^{\text{RVE}}, \mathcal{F}(a)) = \begin{pmatrix} \text{Var}\ a_{ij}^{\text{RVE}} & \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)] \\ \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}] & \text{Var}\ \mathcal{F}(a) \end{pmatrix}.$$

The probability density $\mathcal{M}^\delta$ of the first-variable marginal of the corresponding multivariate Gaussian conditioned on $|\mathcal{F}(a)| \leq \delta L^{-d/2}$, which is the probability distribution by which we approximate the distribution of $a_{ij}^{\text{sel-RVE}}$, is given by

$$\mathcal{M}^\delta(x) = \frac{1}{\int_{\mathbb{R}} \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_\Lambda(\tilde{x}, y)\, dy\, d\tilde{x}} \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \tag{44}$$
$$\times\ \mathcal{N}_\Lambda(x - \mathbb{E}[a_{ij}^{\text{RVE}}], y)\, dy.$$

Our goal is to show that this probability density $\mathcal{M}^\delta$ may be rewritten in the form (34). To this end, we recall some basic linear algebra. The Schur complement of the symmetric block matrix

$$M := \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$$

(with $A^T = A$ and $D^T = D$) is given by $T := A - BD^{-1}B^T$, and the inverse of the matrix may be written as

$$\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}B^T T^{-1} & D^{-1} + D^{-1}B^T T^{-1}BD^{-1} \end{pmatrix}.$$

The determinant may be expressed as $\det M = \det T \cdot \det D$. The Schur complement allows us to rewrite the quadratic form defined by M^{-1} as

$$M^{-1}(x, y) \cdot (x, y) = T^{-1}(x - BD^{-1}y) \cdot (x - BD^{-1}y) + D^{-1}y \cdot y.$$

As a consequence, we get for $M := \Lambda$ that

$$T = \text{Var}\ a_{ij}^{\text{RVE}} - \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var}\ \mathcal{F}(a))^{-1} \text{Cov}[\mathcal{F}(a), a_{ij}^{\text{RVE}}]$$
$$= \text{Var}\ a_{ij}^{\text{RVE}}|_{\text{unexpl}}$$

and

$$\mathcal{N}_\Lambda(x, y) = \frac{1}{(2\pi)^{(N+1)/2}\sqrt{\det \Lambda}} \exp\left(-\frac{1}{2}\Lambda^{-1}(x, y) \cdot (x, y)\right) \tag{45}$$
$$= \mathcal{N}_{\text{Var}\ a_{ij}^{\text{RVE}}|_{\text{unexpl}}}(x - \text{Cov}[a_{ij}^{\text{RVE}}, \mathcal{F}(a)](\text{Var}\ \mathcal{F}(a))^{-1}y)\mathcal{N}_{\text{Var}\ \mathcal{F}(a)}(y).$$

Now, (34) and (44) are seen to be equivalent.

Step 4: Proof of the normal approximation estimate and the lower bound on the probability of the event $|\mathcal{F}(a)| \leq \delta L^{-d/2}$. First, let us show the lower bound (35). We have

$$\int_{\mathbb{R}^N} \int_{\mathbb{R}} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_{\mathrm{Var}\ (a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a))}(x, y) \ \mathrm{d}x \ \mathrm{d}y$$

$$= \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y) \ \mathrm{d}y$$

$$\overset{(13)}{\geq} \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_{C_0 L^{-d}}(y) \ \mathrm{d}y$$

$$\geq \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \frac{1}{(2\pi C_0 L^{-d})^{N/2}} \exp(-\delta^2) \ \mathrm{d}y$$

$$\geq c(N) C_0^{-N/2} \delta^N,$$

establishing (35).

The estimate (36b) is a consequence of the estimate on $\mathrm{Var}\ (a^{\mathrm{RVE}}, \mathcal{F}(a))$ which follows from (36a), (13), and the exponential moment bounds for Gaussians. The bound (36a) is a consequence of Lemma 12 (note that by Proposition 7, Lemma 12 is indeed applicable).

Our next goal is to show (37) and (39). Let $\tilde{\phi} : \mathbb{R} \rightarrow \mathbb{R}$ satisfy (38) and suppose that we would like to estimate the error

$$\mathbb{E}\big[\tilde{\phi}\big(a_{ij}^{\mathrm{sel\text{-}RVE}}\big)\big] - \int_{\mathbb{R}} \tilde{\phi}(x) \mathcal{M}^\delta(x) \ \mathrm{d}x.$$

As the distribution of $a_{ij}^{\mathrm{sel\text{-}RVE}}$ is obtained from the distribution of a_{ij}^{RVE} by conditioning on the event $|\mathcal{F}(a)| \leq \delta L^{-d/2}$, by (34) and (44) this error expression is equal to

$$\mathbb{E}\big[\tilde{\phi}\big(a_{ij}^{\mathrm{sel\text{-}RVE}}\big)\big] - \int_{\mathbb{R}} \tilde{\phi}(x) \mathcal{M}^\delta(x) \ \mathrm{d}x$$

$$= \frac{1}{\mathbb{P}[|\mathcal{F}(a)| \leq \delta L^{-d/2}]} \bigg(\mathbb{E}\big[\chi_{\{|\mathcal{F}(a)| \leq \delta L^{-d/2}\}} \tilde{\phi}\big(a_{ij}^{\mathrm{RVE}}\big)\big]$$

$$- \int_{\mathbb{R} \times \mathbb{R}^N} \tilde{\phi}(x) \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_\Lambda(x - \mathbb{E}[a_{ij}^{\mathrm{RVE}}], y) \ \mathrm{d}(x, y) \bigg)$$

$$+ \int_{\mathbb{R}} \tilde{\phi}(x) \mathcal{M}^\delta(x) \ \mathrm{d}x \bigg(\frac{\int_{\mathbb{R} \times \mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_\Lambda(x, y) \ \mathrm{d}(x, y)}{\mathbb{P}[|\mathcal{F}(a)| \leq \delta L^{-d/2}]} - 1 \bigg). \qquad (46)$$

Up to the normalizing factor $1/\mathbb{P}\big[|\mathcal{F}(a)| \leq \delta L^{-d/2}\big]$, the first term on the right-hand side is given by

$$\mathbb{E}\big[\phi\big(a^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]\big)\big] - \int_{\mathbb{R} \times \mathbb{R}^N} \phi(x, y) \mathcal{N}_\Lambda(x, y) \ \mathrm{d}(x, y),$$

where $\phi : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$ is defined as

$$\phi(x, y) := \begin{cases} \tilde{\phi}(x + \mathbb{E}[a_{ij}^{\mathrm{RVE}}]) & \text{for } |y| \leq \delta L^{-d/2}, \\ 0 & \text{for } |y| > \delta L^{-d/2}. \end{cases} \qquad (47)$$

We would now like to show that (a suitable multiple of) the function ϕ is admissible in the error bound (33). By the estimate

$$
\mathrm{osc}_r\phi(x, y) \leqq \chi_{\{|y|\leqq\delta L^{-d/2}+r\}}\mathrm{osc}_r\tilde{\phi}(x + \mathbb{E}[a_{ij}^{\mathrm{RVE}}])
$$
$$
+ \chi_{|y|\in[\delta L^{-d/2}-r,\delta L^{-d/2}+r]}|\tilde{\phi}(x + \mathbb{E}[a_{ij}^{\mathrm{RVE}}])|,
$$

we obtain, for any $z_0 = (x_0 - \mathbb{E}[a_{ij}^{\mathrm{RVE}}], y_0) \in \mathbb{R} \times \mathbb{R}^N$, also making use of the abbreviation $Q := \mathrm{Cov}[a^{\mathrm{RVE}}, \mathcal{F}(a)](\mathrm{Var}\ \mathcal{F}(a))^{-1}$,

$$
\int_{\mathbb{R}\times\mathbb{R}^N}\mathrm{osc}_r\phi(z)\mathcal{N}_\Lambda(z - z_0)\,\mathrm{d}z
$$
$$
\leqq \int_{\mathbb{R}^N}\int_{\mathbb{R}}(\mathrm{osc}_r\tilde{\phi})(x)\chi_{\{|y|\leqq\delta L^{-d/2}+r\}}\mathcal{N}_\Lambda(x - x_0, y - y_0)\,\mathrm{d}x\,\mathrm{d}y
$$
$$
+ \int_{\mathbb{R}^N}\int_{\mathbb{R}}|\tilde{\phi}|(x)\chi_{|y|\in[\delta L^{-d/2}-r,\delta L^{-d/2}+r]}\mathcal{N}_\Lambda(x - x_0, y - y_0)\,\mathrm{d}x\,\mathrm{d}y
$$
$$
\overset{(45)}{\leqq} \int_{\mathbb{R}^N}\chi_{\{|y|\leqq\delta L^{-d/2}+r\}}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y - y_0)
$$
$$
\times \int_{\mathbb{R}}(\mathrm{osc}_r\tilde{\phi})(x)\mathcal{N}_{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}|_{\mathrm{unexpl}}}(x - x_0 - Q(y - y_0))\,\mathrm{d}x\,\mathrm{d}y
$$
$$
+ \int_{\mathbb{R}^N}\chi_{|y|\in[\delta L^{-d/2}-r,\delta L^{-d/2}+r]}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y - y_0)
$$
$$
\times \int_{\mathbb{R}}|\tilde{\phi}|(x)\mathcal{N}_{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}|_{\mathrm{unexpl}}}(x - x_0 - Q(y - y_0))\,\mathrm{d}x\,\mathrm{d}y
$$
$$
\overset{(38)}{\leqq} \int_{\mathbb{R}^N}\chi_{\{|y|\leqq\delta L^{-d/2}+r\}}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y - y_0) \cdot r\,\mathrm{d}y
$$
$$
+ \int_{\mathbb{R}^N}\chi_{|y|\in[\delta L^{-d/2}-r,\delta L^{-d/2}+r]}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y - y_0) \cdot L^{-d/2}\,\mathrm{d}y,
$$

and therefore

$$
\int_{\mathbb{R}\times\mathbb{R}^N}\mathrm{osc}_r\phi(z)\mathcal{N}_\Lambda(z - z_0)\,\mathrm{d}z
$$
$$
\leqq r + L^{-d/2}\int_{\mathbb{R}^N}\chi_{|y+y_0|\in[\delta L^{-d/2}-r,\delta L^{-d/2}+r]}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\,\mathrm{d}y
$$
$$
\leqq r + L^{-d/2}\sup_{|W|\leqq|B_1^N|\left((\delta L^{-d/2}+r)^N-(\delta L^{-d/2}-r)_+^N\right)}\int_W\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\,\mathrm{d}y
$$
$$
\leqq r + L^{-d/2}\min\left\{1, |B_1^N| \cdot \frac{(\delta L^{-d/2}+r)^N - (\delta L^{-d/2}-r)_+^N}{(2\pi)^{N/2}\sqrt{\det \mathrm{Var}\ \mathcal{F}(a)}}\right\}
$$
$$
\leqq r + L^{-d/2}\min\left\{1, |B_1^N|\frac{N(\delta L^{-d/2}+r)^{N-1} \cdot 2r}{(2\pi)^{N/2}\sqrt{\det \mathrm{Var}\ \mathcal{F}(a)}}\right\}
$$
$$
\leqq r + L^{-d/2}\min\left\{1, \frac{C(N)r^N + C(N)(\delta L^{-d/2})^{N-1}r}{\sqrt{\det \mathrm{Var}\ \mathcal{F}(a)}}\right\}
$$

$$\overset{(13)}{\leqq} r + \frac{C(N)L^{-d/2}}{\sqrt{\det \mathrm{Var}\ \mathcal{F}(a)}^{1/N}}r + \frac{C(N)\delta^{N-1}L^{-d/2}}{\sqrt{\det \mathrm{Var}\ \mathcal{F}(a)}^{1/N}}r.$$

By our assumption (13), this yields, for any $z_0 \in \mathbb{R} \times \mathbb{R}^N$,

$$\fint_{\mathbb{R}\times\mathbb{R}^N} \mathrm{osc}_r\phi(z)\mathcal{N}_\Lambda(z-z_0)\,\mathrm{d}z \leqq Cr.$$

Looking at Definition 8, we would have $\frac{1}{C}\phi \in \Phi_\Lambda$ if it were not for the *qualitative* Lipschitz continuity condition for functions in Φ_Λ. However, for a standard family of mollifiers ρ_ε supported in $\{|x|^2 + |y|^2 \leqq \varepsilon\}$ the approximations $\phi_\varepsilon(x,y) := (\rho_\varepsilon * \phi)(x, (1-2\delta^{-1}L^{d/2}\varepsilon)y)$ satisfy $\frac{1}{C}\phi_\varepsilon \in \Phi_\Lambda$ for any $\varepsilon \in (0, \frac{1}{4}\delta L^{-d/2}]$ (see Definition 8) for some constant C. Furthermore, the ϕ_ε converge poinwise to ϕ for $\varepsilon \to 0$ (by (47) and the continuity assumption on $\tilde\phi$; it is here that we need the dilation factor $(1-2\delta^{-1}L^{d/2}\varepsilon)$ in the second variable due to the discontinuity in the definition (47)) and satisfy a uniform bound of the form $|\phi_\varepsilon(x,y)| \leqq L^{-d/2}$ (by (47) and (38a)). Choosing the functions $\frac{1}{C}\phi_\varepsilon$ in the definition of the distance $\mathcal{D}$ and passing to the limit $\varepsilon \to 0$, we infer

$$\left| \mathbb{E}\big[\chi_{\{|\mathcal{F}(a)|\leqq\delta L^{-d/2}\}}\tilde\phi\big(a_{ij}^{\mathrm{RVE}}\big)\big] \right.$$
$$\left. - \int_{\mathbb{R}\times\mathbb{R}^N} \tilde\phi(x)\chi_{\{|y|\leqq\delta L^{-d/2}\}}\mathcal{N}_\Lambda(x-\mathbb{E}[a_{ij}^{\mathrm{RVE}}], y)\,\mathrm{d}(x,y) \right|$$
$$\overset{(47),(30)}{\leqq} C\mathcal{D}((a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}], \mathcal{F}(a)), \mathcal{N}_\Lambda).$$

Theorem 9 is applicable to the random variable $X := (a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a))$ by our assumptions on $\mathcal{F}(a)$ (see Assumption 1) and by the multilevel decomposition of a_{ij}^{RVE} provided by Proposition 7. In total, with the notation $\Lambda := \mathrm{Var}\ (a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a))$ the application of Theorem 9 to $(a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a))$ yields

$$\left| \mathbb{E}\big[\chi_{\{|\mathcal{F}(a)|\leqq\delta L^{-d/2}\}}\tilde\phi\big(a_{ij}^{\mathrm{RVE}}\big)\big] \right.$$
$$\left. - \int_{\mathbb{R}\times\mathbb{R}^N} \tilde\phi(x)\chi_{\{|y|\leqq\delta L^{-d/2}\}}\mathcal{N}_\Lambda(x-\mathbb{E}[a_{ij}^{\mathrm{RVE}}], y)\,\mathrm{d}(x,y) \right| \qquad (48)$$
$$\leqq C(d,\gamma,N,K)B^3|\log L|^{C(d,\gamma)}(L^{-d}|\Lambda|^{1/2}|\Lambda^{-1/2}|^3)L^{-d}$$
$$\leqq C(d,\lambda,\gamma,N,C_0)\kappa_{ij}^{3/2}L^{-d}|\log L|^{C(d,\gamma)},$$

where in the last step we have used (13) (which entails $L^{-d} \leqq |\Lambda^{1/2}|^2$) and the definition of κ_{ij}.

Applying a similar line of argument to the random variable $\mathcal{F}(a)$ and the function

$$\phi(y) := \begin{cases} 1 & \text{for } |y| \leqq \delta L^{-d/2}, \\ 0 & \text{for } |y| > \delta L^{-d/2}, \end{cases}$$

we obtain

$$
\left| \mathbb{E}\big[\chi_{\{|\mathcal{F}(a)| \leq \delta L^{-d/2}\}} \big] - \int_{\mathbb{R}^N} \chi_{\{|y| \leq \delta L^{-d/2}\}} \mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\, \mathrm{d}y \right| \tag{49}
$$
$$
\leq C(d, \gamma, N, C_0) C_0^{3N/2} L^{-d/2} |\log L|^{C(d,\gamma,C_0)},
$$

where we have estimated $\kappa(\mathrm{Var}\ \mathcal{F}(a))$ by (13). Together with the lower bound (35) and our assumption $\delta^N \geq C L^{-d/2} |\log L|^{C(d,\gamma,C_0)}$, this estimate implies (18).

Plugging in the estimate (48), the lower bound (18), and the estimate (49) as well as the assumption (38a) into (46), we deduce (37). The estimate (39) follows by repeating the above steps, but appealing in the proof of (48) to the bound (33) instead of (32) and choosing $\Lambda := \mathrm{Var}\ (a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a)) + L^{-d/2-d/8}\,\mathrm{Id}$ (which ensures by (13) that $\kappa(\Lambda) \leq C L^{d/8}$). $\square$

We now turn to the proof of the moderate-deviations-type result for the selection approach for representative volumes stated in Theorem 3.

Proof of Theorem 3. Fix $\tilde{S} \geq C L^{-d/2-\beta/2}$. Our goal is to estimate the probability

$$
\mathbb{P}\big[|a_{ij}^{\mathrm{sel\text{-}RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]| \geq \tilde{S} \big]
$$
$$
= \mathbb{P}\Big[|a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]| \geq \tilde{S} \ \Big| \ |\mathcal{F}(a)| \leq \delta L^{-d/2} \Big]
$$
$$
= \frac{\mathbb{P}\big[|a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]| \geq \tilde{S} \text{ and } |\mathcal{F}(a)| \leq \delta L^{-d/2} \big]}{\mathbb{P}[|\mathcal{F}(a)| \leq \delta L^{-d/2}]}. \tag{50}
$$

The main task is the derivation of a suitable estimate for the numerator. To this aim, we apply the moderate deviations estimate from Theorem 10 to the random variable $(a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}], \mathcal{F}(a))$ and the set $A := A_1 \times A_2$ with

$$
A_1 := \big\{ x \in \mathbb{R} : |x| \geq \tilde{S} + C L^{-d/2-\beta} \big\},
$$
$$
A_2 := \big\{ y \in \mathbb{R}^N : |y| \leq \delta L^{-d/2} \big\}.
$$

By Proposition 7 and our assumptions, the application of Theorem 10 is possible, resulting in the estimate

$$
\mathbb{P}\big[(a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}], \mathcal{F}(a)) \in A \big]
$$
$$
\leq \int_{\{(x,y) \in \mathbb{R} \times \mathbb{R}^N : \mathrm{dist}((x,y),A) \leq C L^{-\beta} L^{-d/2}\}} \mathcal{N}_{\tilde{\Lambda}}(x, y)\, \mathrm{d}(x, y)
$$
$$
+ C \exp(-c L^{2\beta} |\log L|^{-C})
$$
$$
\leq \int_{\mathbb{R}^N} \chi_{\{|y| \leq (\delta + C L^{-\beta}) L^{-d/2}\}} \int_{\mathbb{R} \setminus [-\tilde{S}, \tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y)\, \mathrm{d}x\, \mathrm{d}y + C \exp(-L^\beta) \tag{51}
$$

for some positive definite matrix $\tilde{\Lambda}$ with

$$
|\tilde{\Lambda} - \mathrm{Var}\ (a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a))| \leq C(d, \gamma, N, K) |\log L|^C L^{-2\beta} L^{-d}. \tag{52}
$$

We intend to apply the factorization property (45) to the matrix $\tilde{\Lambda}$ with the notation

$$\tilde{\Lambda} = \begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{B}^T & \tilde{D} \end{pmatrix}.$$

By (52) and the bounds $L^{-d} \operatorname{Id} \leqq \operatorname{Var} \mathcal{F}(a) \leqq CL^{-d} \operatorname{Id}$ (see (13)) and $\operatorname{Var} a_{ij}^{\mathrm{RVE}} \leqq CL^{-d} |\log L|^d$ (see (36a)), we deduce

$$|\tilde{D}^{-1} - (\operatorname{Var} \mathcal{F}(a))^{-1}| \leqq CL^d L^{-2\beta} \tag{53}$$

and

$$|\tilde{B}\tilde{D}^{-1} - \operatorname{Cov}[a_{ij}^{\mathrm{RVE}}, \mathcal{F}(a)](\operatorname{Var} \mathcal{F}(a))^{-1}| \leqq CL^{-2\beta} |\log L|^d. \tag{54}$$

As a consequence of these estimates and (52), the formula (17) for $|\rho|^2$ implies for $\tilde{T} := \tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{B}^T$ that

$$\left| \tilde{T} - (1 - |\rho|^2) \operatorname{Var} a_{ij}^{\mathrm{RVE}} \right| \leqq CL^{-d-2\beta} |\log L|^d. \tag{55}$$

Using the bounds $\operatorname{Var} a_{ij}^{\mathrm{RVE}} \leqq CL^{-d} |\log L|^d$ and $|\rho| \leqq 1$ as well as (54), (17), and (13), we obtain for any $|y| \leqq (\delta + CL^{-\beta})L^{-d/2}$ that

$$|\tilde{B}\tilde{D}^{-1}y| \leqq C\delta|\rho|\sqrt{\operatorname{Var} a_{ij}^{\mathrm{RVE}}} + CL^{-d/2-\beta}. \tag{56}$$

Applying the factorization property (45) to the first term on the right-hand side of (51), we obtain

$$\int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{s}, \tilde{s}]} \mathcal{N}_{\tilde{\Lambda}}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

$$= \int_{\mathbb{R}^N} \int_{\mathbb{R}\setminus[-\tilde{s}, \tilde{s}]} \mathcal{N}_{\tilde{T}}(x - \tilde{B}\tilde{D}^{-1}y) \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\tilde{D}}(y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\leqq \int_{\mathbb{R}^N} \int_{\mathbb{R}\setminus[-\tilde{s}, \tilde{s}]} \mathcal{N}_{\tilde{T}}(x) \cdot \exp\left(\tilde{T}^{-1}x \cdot \tilde{B}\tilde{D}^{-1}y\right)$$

$$\times \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\tilde{D}}(y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\overset{(56)}{\leqq} \int_{\mathbb{R}^N} \int_{\mathbb{R}\setminus[-\tilde{s}, \tilde{s}]} \frac{1}{\sqrt{2\pi\tilde{T}}} \exp\left(\frac{-x^2 + C\delta|\rho|\sqrt{\operatorname{Var} a_{ij}^{\mathrm{RVE}}}|x| + C|x|L^{-d/2-\beta}}{2\tilde{T}}\right)$$

$$\times \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\tilde{D}}(y) \, \mathrm{d}x \, \mathrm{d}y.$$

Assuming that $\tilde{S} \geqq CL^{-d/2-\beta/2}$, we deduce

$$\int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\leqq \int_{\mathbb{R}^N} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \frac{1}{\sqrt{2\pi\tilde{T}}} \exp\left(\frac{-\left(1 - \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} - L^{-\beta/2}\right)x^2}{2\tilde{T}}\right)$$

$$\times \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\tilde{D}}(y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\leqq \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \frac{1}{1 - \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} - L^{-\beta/2}} \mathcal{N}_V(x) \, \mathrm{d}x$$

$$\times \int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\tilde{D}}(y) \, \mathrm{d}y,$$

with

$$V := \frac{\tilde{T}}{1 - C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}\tilde{S}^{-1} - L^{-\beta/2}}. \tag{57}$$

Using (53) to estimate the last factor in this estimate and assuming for the moment $\tilde{S} \geqq C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}$ as well as $L \geqq C(\beta)$ to estimate the quotient in the first factor, we get

$$\int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\leqq \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_V(x) \, \mathrm{d}x$$

$$\times \int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a) - CL^{-d-2\beta}\,\mathrm{Id}}(y) \, \mathrm{d}y.$$

Using the bound $L^{-d}\,\mathrm{Id} \leqq \mathrm{Var}\ \mathcal{F}(a)$ from (13) and assuming $L^{-2\beta} \leqq c$, we get

$$\int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\leqq \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_V(x) \, \mathrm{d}x$$

$$\times \int_{\mathbb{R}^N} \chi_{\{|y| \leqq (\delta + 2CL^{-\beta})L^{-d/2}\}} \mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y) \, \mathrm{d}y,$$

and therefore by the upper bound $|\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}| \leq C(L^{-d/2})^{-d}$ and the estimate on the volume $|\{\delta L^{-d/2} \leq |y| \leq (\delta + 2CL^{-\beta})L^{-d/2}\}| \leq C(L^{-d/2})^{d-1}L^{-d/2-\beta}$,

$$
\int_{\mathbb{R}^N} \chi_{\{|y|\leq(\delta+CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y)\, \mathrm{d}x\, \mathrm{d}y
$$

$$
\leq \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_V(x)\, \mathrm{d}x
$$

$$
\times \left(\int_{\mathbb{R}^N} \chi_{\{|y|\leq\delta L^{-d/2}\}}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\, \mathrm{d}y + CL^{-\beta}\right).
$$

By $\tilde{T} \leq (1 - |\rho|^2)\mathrm{Var}\ a_{ij}^{\mathrm{RVE}} + CL^{-d-\beta}\,\mathrm{Id}$ (which follows from (55)) and $\mathrm{Var}\ a_{ij}^{\mathrm{RVE}} \leq CL^{-d}\,\mathrm{Id}$, we deduce from (57) under the assumptions $\tilde{S} \geq C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}$ and $L \geq C(\beta)$

$$
V \leq \tilde{V} := \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}}\right)(1 - |\rho|^2)\mathrm{Var}\ a_{ij}^{\mathrm{RVE}} + CL^{-d-\beta/2}. \quad (58)
$$

As a consequence, we obtain

$$
\int_{\mathbb{R}^N} \chi_{\{|y|\leq(\delta+CL^{-\beta})L^{-d/2}\}} \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{\Lambda}}(x, y)\, \mathrm{d}x\, \mathrm{d}y
$$

$$
\leq \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{V}}(x)\, \mathrm{d}x
$$

$$
\times \left(\int_{\mathbb{R}^N} \chi_{\{|y|\leq\delta L^{-d/2}\}}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\, \mathrm{d}y + CL^{-\beta}\right).
$$

Plugging this bound into (51), we obtain

$$
\mathbb{P}\big[|a_{ij}^{\mathrm{RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]| \geq \tilde{S} + CL^{-d/2-\beta} \text{ and } |\mathcal{F}(a)| \leq \delta L^{-d/2}\big]
$$

$$
\leq \left(1 + \frac{C\delta|\rho|\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{V}}(x)\, \mathrm{d}x
$$

$$
\times \left(\int_{\mathbb{R}^N} \chi_{\{|y|\leq\delta L^{-d/2}\}}\mathcal{N}_{\mathrm{Var}\ \mathcal{F}(a)}(y)\, \mathrm{d}y + CL^{-\beta}\right)
$$

$$
+ C\exp(-L^\beta).
$$

Inserting the previous estimate into (50) and using (49), (35), and (18) as well as the assumption $\delta^N \geq CL^{-d/2}$ to estimate the denominator, we get

$$
\mathbb{P}\big[|a_{ij}^{\mathrm{sel\text{-}RVE}} - \mathbb{E}[a_{ij}^{\mathrm{RVE}}]| \geq \tilde{S} + CL^{-d/2-\beta}\big]
$$

$$
\leq \left(1 + \frac{C\delta\sqrt{\mathrm{Var}\ a_{ij}^{\mathrm{RVE}}}}{\tilde{S}} + \frac{C}{\delta^N}L^{-\beta/2}\right) \int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]} \mathcal{N}_{\tilde{V}}(x)\, \mathrm{d}x + \frac{C}{\delta^N}\exp(-L^\beta).
$$

Note that we have the estimate $|\mathbb{E}[a_{ij}^{\text{RVE}}] - a_{\text{hom},ij}| \leq CL^{-d}|\log L|^C$. By redefining $\tilde{S}$ (and possibly increasing the constant in (58); recall that $\tilde{S} \geq L^{-d/2-\beta/2}$), we obtain

$$\mathbb{P}\left[|a_{ij}^{\text{sel-RVE}} - a_{\text{hom},ij}| \geq \tilde{S}\right]$$
$$\leq \left(1 + \frac{C\delta\sqrt{\text{Var } a_{ij}^{\text{RVE}}}}{\tilde{S}} + \frac{C}{\delta^N}L^{-\beta/2}\right)\int_{\mathbb{R}\setminus[-\tilde{S},\tilde{S}]}\mathcal{N}_{\tilde{V}}(x)\,dx + \frac{C}{\delta^N}\exp(-L^\beta).$$

Finally, we set $\tilde{S} := \sqrt{(1 + \frac{C\delta}{\sqrt{1-|\rho|^2}s})(1 - |\rho|^2)\text{Var } a_{ij}^{\text{RVE}} + L^{-d-\beta/2}} \cdot s$. Upon redefining β, this yields the desired estimate (21). $\square$

5. The Multilevel Local Dependence Structure of the Approximation for the Effective Conductivity

We now prove that the approximation a^{RVE} for the effective conductivity obtained by the representative volume element method may indeed be written as a sum of a family of random variables with multilevel local dependence structure in the sense of Definition 6. Furthermore, we show that the same is true for the spatial average of the coefficient field $\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a\,dx$ and also for the second-order correction $\mathcal{F}_{2-\text{point}}(a)$ to a^{RVE} in the setting of small ellipticity contrast.

Proof of Proposition 7. *Part 1: The spatial average of the coefficient.* First, let us show that the average $\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a\,dx$ is approximately the sum of a family of random variables with multilevel local dependence structure. Decomposing

$$\mathcal{F}_{avg}(a) = \fint_{[0,L\varepsilon]^d} a\,dx = \sum_{y\in\varepsilon\mathbb{Z}^d\cap[0,L\varepsilon)^d}\underbrace{\frac{1}{L^d}\fint_{y+[0,\varepsilon]^d} a\,dx}_{=:X_y^0},$$

defining the X_y^0 as indicated in this formula, and setting $X_y^m := 0$ for $m \geq 1$, we immediately observe that the average $\mathcal{F}_{avg}(a)$ is the sum of a family of random variables with multilevel local dependence structure with $K := 1$. The bound (26) follows immediately from the uniform bound on a (with $B := ||a||_{L^\infty}$ and arbitrary $\gamma > 0$).

Part 2: The approximation a^{RVE} for the effective coefficient. Next, let us show that a^{RVE} is approximately the sum of a family of random variables with multilevel local dependence structure. For simplicity of notation, let us assume that $\varepsilon = 1$.

Recall that the corrector ϕ_i associated with the periodized ensemble is the unique L-periodic solution to the equation

$$\nabla \cdot (a(e_i + \nabla\phi_i)) = 0 \tag{59}$$

with vanishing average $\fint_{[0,L]^d} \phi_i \, dx = 0$. We shall use the decomposition of the (L-periodic) corrector ϕ_i according to

$$\phi_i(\cdot) = \int_0^\infty u_i(\cdot, s) \, ds, \tag{60}$$

where $u_i = u_i(x, s)$ is the (L-periodic) solution to the parabolic PDE

$$\frac{d}{ds} u_i = \nabla \cdot (a \nabla u_i) \qquad \text{in } [0, L]^d \times [0, \infty),$$
$$u_i(\cdot, 0) = \nabla \cdot (a e_i) \qquad \text{in } [0, L]^d.$$

Observe that the parabolic PDE directly entails

$$\nabla \cdot \left(a \left(e_i + \nabla \int_0^t u_i(\cdot, s) \, ds \right) \right) = u_i(\cdot, t). \tag{61}$$

Thus, decay of u_i for $t \to \infty$ implies that ϕ_i may indeed be decomposed as $\int_0^\infty u_i(\cdot, s) \, ds$. Note that exponential decay of u_i (with an L-dependent constant) is immediate by the standard energy estimate, the vanishing average of $u_i(\cdot, s)$ for any $s \geq 0$ (as the average of the initial conditions on $[0, L]^d$ vanishes), and the Poincaré inequality.

Recall the key result from [55] which states that under the assumptions of ellipticity, stationarity, and finite range of dependence (A1)–(A3) the full-space variant $u_i^{\mathbb{R}^d}(\cdot, s)$—that is, the solution to the equation

$$\frac{d}{ds} u_i^{\mathbb{R}^d} = \nabla \cdot (a^{\mathbb{R}^d} \nabla u_i^{\mathbb{R}^d}) \qquad \text{in } \mathbb{R}^d \times [0, \infty),$$
$$u_i^{\mathbb{R}^d}(\cdot, 0) = \nabla \cdot (a^{\mathbb{R}^d} e_i) \qquad \text{in } \mathbb{R}^d,$$

with $a^{\mathbb{R}^d}$ denoting a coefficient field from the original (non-periodic) ensemble of coefficient fields—actually decays like $s^{-(1+d/2)/2}$ in suitable norms.

Theorem 11. ([55], Corollary 4) *Consider an ensemble of random coefficient fields* $a^{\mathbb{R}^d}$ *subject to the assumptions (A1)–(A3) with range of dependence $\varepsilon := 1$. Then for any $T > 0$ we have the estimate*

$$\left(\fint_{\{|x| \leq \sqrt{T}\}} |\nabla u_i^{\mathbb{R}^d}(\cdot, T)|^2 \, dx \right)^{1/2} \leq \mathcal{C}(a^{\mathbb{R}^d}, T) \, T^{-1-d/4}, \tag{62a}$$

$$\left(\fint_{\{|x| \leq \sqrt{T}\}} |u_i^{\mathbb{R}^d}(\cdot, T)|^2 \, dx \right)^{1/2} \leq \mathcal{C}(a^{\mathbb{R}^d}, T) \, T^{-1/2-d/4}, \tag{62b}$$

where the random constant $\mathcal{C}(a^{\mathbb{R}^d}, T)$ satisfies for any $\delta > 0$ a bound of the form

$$\mathbb{E} \left[\exp \left(\frac{\mathcal{C}(a^{\mathbb{R}^d}, T)^{2-\delta}}{C(d, \lambda, \delta)} \right) \right] \leq 2.$$

Note that the second inequality (62b) is actually not contained in [55, Corollary 4]. However, it is an easy consequence of (62a) (the proof is provided below).

By ϕ_j^* and u_j^* we shall denote the corresponding quantities for the adjoint coefficient field a^*, that is $\phi_j^*(\cdot) := \int_0^\infty u_j^*(\cdot, s)\, ds$ with u_j^* being the L-periodic solution to

$$\frac{d}{ds} u_j^* = \nabla \cdot (a^* \nabla u_j^*) \qquad \text{in } [0, L]^d \times [0, \infty),$$

$$u_j^*(\cdot, 0) = \nabla \cdot (a^* e_j) \qquad \text{in } [0, L]^d.$$

The full space variants $u_j^{*, \mathbb{R}^d}$ satisfy also estimates of the form (62a)–(62b), as the conditions (A1)–(A3) are invariant under passing to the adjoint coefficient fields.

We introduce a "cutoff scale" L_K as the largest integer power of 2 not larger than $\frac{L}{16K \log L}$ for some constant $K \geq 1$ that remains to be chosen. Defining $T_L := (L_K)^2$, we now compute, using the properties (59), (60) and (61), that

$$
\begin{aligned}
a^{\text{RVE}} e_i \cdot e_j &= \fint_{[0,L]^d} a(e_i + \nabla \phi_i) \cdot e_j \, dx \\
&\stackrel{(59)}{=} \fint_{[0,L]^d} a(e_i + \nabla \phi_i) \cdot (e_j + \nabla \phi_j^*) \, dx \\
&\stackrel{(60)}{=} \fint_{[0,L]^d} a\left(e_i + \nabla \int_0^1 u_i(\cdot, s)\, ds\right) \cdot \left(e_j + \nabla \int_0^1 u_j^*(\cdot, s)\, ds\right) dx \\
&\quad + \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} a\nabla \int_{4^k}^{4^{k+1}} u_i(\cdot, s)\, ds \cdot \left(e_j + \nabla \int_0^{4^k} u_j^*(\cdot, s)\, ds\right) dx \\
&\quad + \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} a\left(e_i + \nabla \int_0^{4^k} u_i(\cdot, s)\, ds\right) \cdot \nabla \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s)\, ds\, dx \\
&\quad + \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} a\nabla \int_{4^k}^{4^{k+1}} u_i(\cdot, s)\, ds \cdot \nabla \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s)\, ds\, dx \\
&\quad + \fint_{[0,L]^d} a\nabla \int_{4T_L}^{\infty} u_i(\cdot, s)\, ds \cdot \left(e_j + \nabla \int_0^{4T_L} u_j^*(\cdot, s)\, ds\right) dx \\
&\quad + \fint_{[0,L]^d} a\left(e_i + \nabla \int_0^{4T_L} u_i(\cdot, s)\, ds\right) \cdot \nabla \int_{4T_L}^{\infty} u_j^*(\cdot, s)\, ds\, dx \\
&\quad + \fint_{[0,L]^d} a\nabla \int_{4T_L}^{\infty} u_i(\cdot, s)\, ds \cdot \nabla \int_{4T_L}^{\infty} u_j^*(\cdot, s)\, ds \\
&\stackrel{(61)}{=} \fint_{[0,L]^d} a\left(e_i + \nabla \int_0^1 u_i(\cdot, s)\, ds\right) \cdot \left(e_j + \nabla \int_0^1 u_j^*(\cdot, s)\, ds\right) dx \\
&\quad - \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} \int_{4^k}^{4^{k+1}} u_i(\cdot, s)\, ds\, u_j^*(\cdot, 4^k)\, dx
\end{aligned}
$$

$$- \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} u_i(\cdot, 4^k) \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s) \, ds \, dx$$

$$+ \sum_{k=0}^{\log_2 L_K} \fint_{[0,L]^d} a\nabla \int_{4^k}^{4^{k+1}} u_i(\cdot, s) \, ds \cdot \nabla \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s) \, ds \, dx$$

$$- \fint_{[0,L]^d} \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \, u_j^*(\cdot, 4T_L) \, dx$$

$$- \fint_{[0,L]^d} u_i(\cdot, 4T_L) \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \, dx$$

$$+ \fint_{[0,L]^d} a\nabla \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \cdot \nabla \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \, dx.$$

We now decompose the integrals into integrals over cubes with side length $\sim 2^k$, resulting in

$$a^{\text{RVE}} e_i \cdot e_j \tag{63}$$

$$= \sum_{x_0 \in \mathbb{Z}^d} \frac{1}{L^d} \fint_{(x_0 + [0,1]^d) \cap [0,L]^d} a\left(e_i + \nabla \int_0^1 u_i(\cdot, s) \, ds\right)$$

$$\cdot \left(e_j + \nabla \int_0^1 u_j^*(\cdot, s) \, ds\right) dx$$

$$- \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d} \frac{1}{L^d} \int_{(x_0 + [0,2^k]^d) \cap [0,L]^d} \int_{4^k}^{4^{k+1}} u_i(\cdot, s) \, ds \, u_j^*(\cdot, 4^k) \, dx$$

$$- \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d} \frac{1}{L^d} \int_{(x_0 + [0,2^k]^d) \cap [0,L]^d} u_i(\cdot, 4^k) \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s) \, ds \, dx$$

$$+ \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d} \frac{1}{L^d} \int_{(x_0 + [0,2^k]^d) \cap [0,L]^d} a\nabla \int_{4^k}^{4^{k+1}} u_i(\cdot, s) \, ds$$

$$\cdot \nabla \int_{4^k}^{4^{k+1}} u_j^*(\cdot, s) \, ds \, dx$$

$$- \fint_{[0,L]^d} \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \, u_j^*(\cdot, 4T_L) \, dx$$

$$- \fint_{[0,L]^d} u_i(\cdot, 4T_L) \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \, dx$$

$$+ \fint_{[0,L]^d} a\nabla \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \cdot \nabla \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \, dx.$$

We now intend to replace u_i and u_j^* in each of these expressions by a proxy with localized dependence. To this end, for any $k \in \mathbb{N}_0$ and any $x_0 \in 2^k \mathbb{Z}^d$, define the coefficient field a_{k,x_0} on the full space $\mathbb{R}^d$ as

$$a_{k,x_0}(x) := \begin{cases} a(x) & \text{for } |x - x_0| \leq \sqrt{K|\log L|}\, 2^{k-1}, \\ \text{Id} & \text{otherwise.} \end{cases} \qquad (64)$$

Define a corresponding u_{i,k,x_0} as the solution to the equation

$$\frac{d}{dt} u_{i,k,x_0} = \nabla \cdot (a_{k,x_0} \nabla u_{i,k,x_0}), \qquad (65a)$$

$$u_{i,k,x_0}(\cdot, 0) = \nabla \cdot (a_{k,x_0} e_i), \qquad (65b)$$

and introduce, analogously, the function u^*_{i,k,x_0} as the solution to the equation with a_{k,x_0} replaced by a^*_{k,x_0}. Note that while u_i and a are defined on $[0, L]^d$ and extended to $\mathbb{R}^d$ by periodicity, both a_{k,x_0} and u_{i,k,x_0} are defined on $\mathbb{R}^d$ and lack any periodicity.

By Lemma 15—applied with $M := \frac{1}{2}\sqrt{K|\log L|}$ and $r := 2^k$—we have

$$\fint_{\{|x-x_0| \leq 2d \cdot 2^k\}} |u_i(\cdot, t) - u_{i,k,x_0}(\cdot, t)|^2 \, dx \qquad (66)$$
$$\leq C\sqrt{K \log L}^{d/2} \exp(-cK|\log L|) \leq C(d, \lambda, K) L^{-cK}$$

for any $t \leq 4^{k+1}$ and

$$\int_0^{4^{k+1}} \fint_{\{|x-x_0| \leq d \cdot 2^k\}} |\nabla u_i - \nabla u_{i,k,x_0}|^2 \, dx \, dt \qquad (67)$$
$$\leq C \exp(-cK|\log L|) \leq C(d, \lambda, K) L^{-cK},$$

and analogous estimates for the difference $u^*_j - u^*_{j,k,x_9}$.

As our probability distribution of coefficient fields a on $[0, L]^d$ is the periodization of a probability distribution of coefficient fields $a^{\mathbb{R}^d}$ on $\mathbb{R}^d$, by definition of a periodization (see (A3$_c$)) for each $x_0 \in [0, L)^d$ and any $k \leq \log_2 L_K$ the law of $a|_{x_0 + K \log L[-2^k, 2^k]^d}$ coincides with the law of $a^{\mathbb{R}^d}|_{x_0 + K \log L[-2^k, 2^k]^d}$. As a consequence, the law of u_{i,k,x_0} coincides with the law of $u^{\mathbb{R}^d}_{i,k,x_0}$, where $u^{\mathbb{R}^d}_{i,k,x_0}$ is defined analogously to u_{i,k,x_0} (replacing a in the definition by $a^{\mathbb{R}^d}$). Therefore, any moment bound on $u^{\mathbb{R}^d}_{i,k,x_0}$ carries over to u_{i,k,x_0}. Applying Lemma 15 to $u^{\mathbb{R}^d}_{i,k,x_0}$, we obtain estimates analogous to (66) and (67). The estimates from Theorem 11 therefore carry over to $u^{\mathbb{R}^d}_{i,k,x_0}$, provided that we choose $K \geq C$; we have for $t \in [4^k, 4^{k+1}]$ and $T = 4^k$ with $2^k \leq L$ that

$$\left(\fint_{\{|x-x_0| \leq d \cdot 2^k\}} |u^{\mathbb{R}^d}_{i,k,x_0}(t)|^2 \, dx \right)^{1/2} \leq C(a^{\mathbb{R}^d}, t)\, t^{-1/2 - d/4},$$

$$\left(\fint_T^{4T} \fint_{\{|x-x_0| \leq d \cdot 2^k\}} |\nabla u^{\mathbb{R}^d}_{i,k,x_0}(T)|^2 \, dx \, dt \right)^{1/2} \leq C(a^{\mathbb{R}^d}, T)\, T^{-1-d/4},$$

for some random constants $\mathcal{C}(a^{\mathbb{R}^d}, t), \mathcal{C}(a^{\mathbb{R}^d}, T)$, with

$$||\mathcal{C}(a^{\mathbb{R}^d}, t)||_{\exp^{2-\delta}} \leqq C(d, \lambda, K, \delta),$$
$$||\mathcal{C}(a^{\mathbb{R}^d}, T)||_{\exp^{2-\delta}} \leqq C(d, \lambda, K, \delta)$$

for any $\delta > 0$. By the coincidence of laws, we get, for $t \in [4^k, 4^{k+1}]$ and $T = 4^k$,

$$\left(\fint_{\{|x-x_0| \leqq d \cdot 2^k\}} |u_{i,k,x_0}(t)|^2 \, \mathrm{d}x \right)^{1/2} \leqq \mathcal{C}(a, t) \, t^{-1/2 - d/4}, \qquad (68a)$$

$$\left(\fint_T^{4T} \fint_{\{|x-x_0| \leqq d \cdot 2^k\}} |\nabla u_{i,k,x_0}(T)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1/2} \leqq \mathcal{C}(a, T) \, T^{-1 - d/4}, \qquad (68b)$$

for random constants $\mathcal{C}$ satisfying

$$||\mathcal{C}(a, t)||_{\exp^{2-\delta}} \leqq C(d, \lambda, K, \delta),$$
$$||\mathcal{C}(a, T)||_{\exp^{2-\delta}} \leqq C(d, \lambda, K, \delta)$$

for any $\delta > 0$. Furthermore, the bound (102) yields an estimate of the form

$$\left(\fint_{\{|x-x_0| \leqq d\}} \left| e_i + \nabla \int_0^1 u_{i,0,x_0} \, \mathrm{d}s \right|^2 \mathrm{d}x \right)^{1/2} \leqq C(d, \lambda). \qquad (69)$$

By (61), its analogue for $u_{i,0,x_0}$, and the definition of a_{0,x_0}, we have in $\{|x - x_0| \leqq 2d\}$ that $-\nabla \cdot (a\nabla(\int_0^1 u_i(\cdot, s) - u_{i,0,x_0}(\cdot, s) \, \mathrm{d}s)) = u_i(\cdot, 1) - u_{i,0,x_0}(\cdot, 1)$, which implies, by the Caccioppoli inequality,

$$\fint_{\{|x-x_0| \leqq d\}} \left| e_i + \nabla \int_0^1 u_i \, \mathrm{d}s - \left(e_i + \nabla \int_0^1 u_{i,0,x_0} \, \mathrm{d}s \right) \right|^2 \mathrm{d}x$$

$$\leqq C \fint_{\{|x-x_0| \leqq 2d\}} |u_i(\cdot, 1) - u_{i,0,x_0}(\cdot, 1)|^2 \, \mathrm{d}x$$

$$+ C \fint_{\{|x-x_0| \leqq 2d\}} \left| \int_0^1 u_i(\cdot, s) - u_{i,0,x_0}(\cdot, s) \, \mathrm{d}s \right|^2 \mathrm{d}x$$

$$\leqq C \fint_{\{|x-x_0| \leqq 2d\}} |u_i(\cdot, 1) - u_{i,0,x_0}(\cdot, 1)|^2 \, \mathrm{d}x$$

$$+ C \fint_{\{|x-x_0| \leqq 2d\}} \int_0^1 |u_i(\cdot, s) - u_{i,0,x_0}(\cdot, s)|^2 \, \mathrm{d}s \, \mathrm{d}x$$

$$\overset{(66)}{\leqq} C(K) L^{-cK}. \qquad (70)$$

As a consequence of our definition of u_{i,k,x_0}, for the choice

$$X_{x_0}^0 := \frac{1}{L^d} \int_{(x_0+[0,1]^d)\cap[0,L]^d} a\left(e_i + \nabla \int_0^1 u_{i,0,x_0}(\cdot,s)\,ds\right) \tag{71a}$$
$$\cdot \left(e_j + \nabla \int_0^1 u_{j,0,x_0}^*(\cdot,s)\,ds\right) dx,$$

$$X_{x_0}^{k+1} := -\frac{1}{L^d} \int_{(x_0+[0,2^k]^d)\cap[0,L]^d} \int_{4^k}^{4^{k+1}} u_{i,k,x_0}(\cdot,s)\,ds\, u_{j,k,x_0}^*(\cdot,4^k)\,dx \tag{71b}$$
$$-\frac{1}{L^d} \int_{(x_0+[0,2^k]^d)\cap[0,L]^d} u_{i,k,x_0}(\cdot,4^k) \int_{4^k}^{4^{k+1}} u_{j,k,x_0}^*(\cdot,s)\,ds\,dx$$
$$+\frac{1}{L^d} \int_{(x_0+[0,2^k]^d)\cap[0,L]^d} a\nabla \int_{4^k}^{4^{k+1}} u_{i,k,x_0}(\cdot,s)\,ds$$
$$\cdot \nabla \int_{4^k}^{4^{k+1}} u_{j,k,x_0}^*(\cdot,s)\,ds\,dx$$

for $0 \leqq k \leqq \log_2 L_K$, we see, by (64) and (65) and $\sqrt{K\log L} \geqq 1$, that $X_{x_0}^k$ is a random variable which depends only on $a|_{x_0+K\log L[-2^k,2^k]^d}$, that is the first condition of Definition 6 is satisfied. Furthermore, by (68) and (69), we obtain, for any $0 < \gamma < 1$, an estimate of the form

$$||X_y^k||_{\exp^\gamma} \leqq C(d,\lambda,\gamma,K)L^{-d}. \tag{72}$$

We now intend to replace the terms in the first five terms on the right-hand side of (63) by the $X_{x_0}^k$ with $0 \leqq k \leqq \log_2 L_K + 1$, using the estimates (66), (67), (70), and Hölder's inequality to bound the arising error. For example, we may estimate

$$\left| -\frac{1}{L^d} \int_{(x_0+[0,2^k]^d)\cap[0,L]^d} u_i(\cdot,4^k) \int_{4^k}^{4^{k+1}} u_j^*(\cdot,s)\,ds\,dx \right.$$
$$\left. -\left(-\frac{1}{L^d} \int_{(x_0+[0,2^k]^d)\cap[0,L]^d} \int_{4^k}^{4^{k+1}} u_{i,k,x_0}(\cdot,s)\,ds\, u_{j,k,x_0}^*(\cdot,4^k)\,dx\right) \right|$$
$$\leqq \frac{4^{(k+1)/2}}{L^d} \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} |u_i(\cdot,4^k)|^2\,dx\right)^{1/2}$$
$$\times \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} \int_{4^k}^{4^{k+1}} |u_j^*(\cdot,s) - u_{j,k,x_0}^*(\cdot,s)|^2\,ds\,dx\right)^{1/2}$$
$$+ \frac{4^{(k+1)/2}}{L^d} \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} |u_i(\cdot,4^k) - u_{i,k,x_0}(\cdot,4^k)|^2\,dx\right)^{1/2}$$
$$\times \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} \int_{4^k}^{4^{k+1}} |u_{j,k,x_0}^*(\cdot,s)|^2\,ds\,dx\right)^{1/2}$$
$$\overset{(66)}{\leqq} \frac{C(d,\lambda,K)}{L^d} \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} |u_{i,k,x_0}(\cdot,4^k)|^2\,dx + L^{-cK}\right)^{1/2}$$

$$\times \, (2^k)^{d/2} L^{-cK}$$

$$+ \frac{C(d, \lambda, K)}{L^d} \cdot (2^k)^{d/2} L^{-cK}$$

$$\times \left(\int_{(x_0+[0,2^k]^d)\cap[0,L]^d} \int_{4^k}^{4^{k+1}} |u_{j,k,x_0}^*(\cdot, s)|^2 \, ds \, dx \right)^{1/2},$$

where in the last step we have used $4^k \leq CL^2$ and $(2^k)^{d/2} \leq CL^{d/2}$, absorbing these factors in the factor L^{-cK} (possible for $cK \geq 4+2d$). Proceeding analogously for the other terms in (63), we deduce

$$\left| a^{\text{RVE}} e_i \cdot e_j - \sum_{x_0 \in \mathbb{Z}^d \cap [0,L)^d} X_{x_0}^0 - \sum_{k=1}^{1+\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} X_{x_0}^k \right|$$

$$\leq C \sum_{x_0 \in \mathbb{Z}^d \cap [0,L)^d} \frac{1}{L^d} \left(\int_{x_0+[0,1]^d} \left| e_i + \nabla \int_0^1 u_{i,0,x_0}(\cdot, s) \, ds \right|^2 \right.$$

$$\left. + \left| e_j + \nabla \int_0^1 u_{j,0,x_0}^*(\cdot, s) \, ds \right|^2 dx + L^{-cK} \right)^{1/2} L^{-cK}$$

$$+ C \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} \frac{1}{L^d} \left(\int_{x_0+[0,2^k]^d} |u_{i,k,x_0}(\cdot, 4^k)|^2 \right.$$

$$\left. + |u_{j,k,x_0}^*(\cdot, 4^k)|^2 \, dx + L^{-cK} \right)^{1/2}$$

$$\times \, (2^k)^{d/2} L^{-cK}$$

$$+ C \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} \frac{1}{L^d} \left(\int_{x_0+[0,2^k]^d} \int_{4^k}^{4^{k+1}} |u_{i,k,x_0}(\cdot, s)|^2 \right.$$

$$\left. + |u_{j,k,x_0}^*(\cdot, s)|^2 \, ds \, dx \right)^{1/2}$$

$$\times \, (2^k)^{d/2} L^{-cK}$$

$$+ C \sum_{k=0}^{\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} \frac{1}{L^d} \left(\int_{x_0+[0,2^k]^d} \int_{4^k}^{4^{k+1}} |\nabla u_{i,k,x_0}(\cdot, s)|^2 \right.$$

$$\left. + |\nabla u_{j,k,x_0}^*(\cdot, s)|^2 \, ds \, dx + L^{-cK} \right)^{1/2} L^{-cK}$$

$$+ \fint_{[0,L]^d} \left| \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \right| |u_j^*(\cdot, 4T_L)| \, dx$$

$$+ \fint_{[0,L]^d} |u_i(\cdot, 4T_L)| \left| \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \right| dx$$

$$+ C \fint_{[0,L]^d} \left| \nabla \int_{4T_L}^{\infty} u_i(\cdot, s) \, ds \right| \left| \nabla \int_{4T_L}^{\infty} u_j^*(\cdot, s) \, ds \right| dx.$$

Inserting the estimates (68) and (69), we get, for some $\mathcal{C}(a)$ with $||\mathcal{C}(a)||_{\exp^\gamma} \leqq C(d, \lambda, K, \gamma)$ for any $\gamma \in (0, 1)$,

$$\left| a^{\text{RVE}} e_i \cdot e_j - \sum_{x_0 \in \mathbb{Z}^d \cap [0,L]^d} X_{x_0}^0 - \sum_{k=1}^{1+\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L]^d} X_{x_0}^k \right| \tag{73}$$

$$\leqq C L^{-cK} + \mathcal{C}(a) \sum_{k=0}^{\log_2 L_K} L^{-cK} + \mathcal{C}(a) \sum_{k=0}^{\log_2 L_K} \sqrt{4^k} L^{-cK} + \mathcal{C}(a) \sum_{k=0}^{\log_2 L_K} L^{-cK}$$

$$+ \fint_{[0,L]^d} \left| \int_{4T_L}^\infty u_i(\cdot, s) \, ds \right| |u_j^*(\cdot, 4T_L)| \, dx$$

$$+ \fint_{[0,L]^d} |u_i(\cdot, 4T_L)| \left| \int_{4T_L}^\infty u_j^*(\cdot, s) \, ds \right| dx$$

$$+ C \fint_{[0,L]^d} \left| \nabla \int_{4T_L}^\infty u_i(\cdot, s) \, ds \right| \left| \nabla \int_{4T_L}^\infty u_j^*(\cdot, s) \, ds \right| dx.$$

The bound (66) and its equivalent for $u_i^{\mathbb{R}^d}$ and $u_{i,k,x_0}^{\mathbb{R}^d}$ enable us to transfer the bounds in Theorem 11 from $u_i^{\mathbb{R}^d}$ to u_i. Recalling that $T_L = (L_K)^2$, we obtain

$$\fint_{[0,L]^d} |u_i(\cdot, T_L)|^2 \, dx = \sum_{x_0 \in L_K \mathbb{Z}^d} L^{-d} \int_{x_0 + [0,L_K]^d \cap [0,L]^d} |u_i(\cdot, T_L)|^2 \, dx \tag{74}$$

$$\overset{(66)}{\leqq} C \sum_{x_0 \in L_K \mathbb{Z}^d} L^{-d} \left(L^{-cK} + \int_{y + [0,L_K]^d \cap [0,L]^d} |u_{i,\log_2 L_K, x_0}(\cdot, T_L)|^2 \, dx \right)$$

and

$$\int_{y + [0,L_K]^d \cap [0,L]^d} |u_{i,\log_2 L_K, x_0}^{\mathbb{R}^d}(\cdot, T_L)|^2 \, dx$$

$$\leqq 2 \int_{y + [0,L_K]^d \cap [0,L]^d} |u_i^{\mathbb{R}^d}(\cdot, T_L)|^2 \, dx + 2C L^{-cK/2}.$$

The latter estimate entails, in view of Theorem 11 (choosing $K \geqq C$ and recalling that $\sqrt{T_L} = L_K \leqq \frac{L}{4K \log L}$), that

$$\left(\fint_{y + [0,L_K]^d \cap [0,L]^d} |u_{i,\log_2 L_K, x_0}^{\mathbb{R}^d}(\cdot, T_L)|^2 \, dx \right)^{1/2}$$

$$\leqq \mathcal{C}(a^{\mathbb{R}^d}, y, T_L) T_L^{-1/2 - d/4},$$

where again $||\mathcal{C}(a^{\mathbb{R}^d}, y, T_L)||_{\exp^{2-\delta}} \leqq C(d, \lambda, K, \delta)$. By coincidence of the laws of $a|_{x_0 + K \log L[-L_K, L_K]^d}$ and $a^{\mathbb{R}^d}|_{x_0 + K \log L[-L_K, L_K]^d}$, we get, for $K \geqq C$ from the previous estimate and (74),

$$\fint_{[0,L]^d} |u_i(\cdot, T_L)|^2 \, dx \leqq \mathcal{C}(a, T_L) T_L^{-1 - d/2} \tag{75}$$

where $\|\mathcal{C}(a, T_L)\|_{\exp^\gamma} \leqq C(d, \lambda, K, \gamma)$ for any $\gamma < 1$. An analogous bound holds for u_j^*. Finally, the energy estimate for u_i implies

$$\frac{d}{dt}\fint_{[0,L]^d} |u_i|^2 \, dx \leqq -c\fint_{[0,L]^d} |\nabla u_i|^2 \, dx.$$

As the average of u_i over $[0, L]^d$ vanishes, the Poincaré inequality implies, for $T \geqq T_L$,

$$\frac{d}{dt}\fint_{[0,L]^d} |u_i|^2 \, dx \leqq -\frac{c}{2}\fint_{[0,L]^d} |\nabla u_i|^2 \, dx - \frac{c}{2CL^2}\fint_{[0,L]^d} |u_i|^2 \, dx,$$

and as a consequence,

$$\fint_{[0,L]^d} |u_i(\cdot, T)|^2 \, dx + \int_{\max\{T_L, T/2\}}^T \fint_{[0,L]^d} |\nabla u_i|^2 \, dx \, dt$$
$$\leqq C(d, \lambda) \exp\left(-\frac{T - T_L}{C(d, \lambda)L^2}\right)\fint_{[0,L]^d} |u_i(\cdot, T_L)|^2 \, dx.$$

Note that this estimate yields, in particular, that

$$\fint_{[0,L]^d} \left|\int_{T_L}^\infty \nabla u_i \, dt\right|^2 dx \leqq 2 \sum_{l=-\log_2 \frac{L^2}{T_L}}^\infty \fint_{[0,L]^d} \left|2^l \int_{2^l L^2}^{2^{l+1} L^2} \nabla u_i \, dt\right|^2 dx$$

$$\leqq 2 \sum_{l=-\log_2 \frac{L^2}{T_L}}^\infty \fint_{[0,L]^d} 2^l (2^l L^2) \int_{2^l L^2}^{2^{l+1} L^2} |\nabla u_i|^2 \, dt \, dx$$

$$\leqq C \sum_{l=-\log_2 \frac{L^2}{T_L}}^\infty 2^{2l} L^2 \exp(-c2^l)\fint_{[0,L]^d} |u_i(\cdot, T_L)|^2 \, dx$$

$$\overset{(75)}{\leqq} \mathcal{C}(a, T_L)\frac{L^2}{T_L}T_L^{-d/2} \leqq \mathcal{C}(a, T_L)(K|\log L|)^{d+2}L^{-d/2},$$

where in the last step we have used that $\sqrt{T_L} = L_K$ is the largest power of 2 with $L_K \leqq \frac{L}{4K \log L}$.

Plugging these bounds and (75) into (73), we get, for $K \geqq C$,

$$\left|a^{\mathrm{RVE}} e_i \cdot e_j - \sum_{k=0}^{1+\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} X_{x_0}^k\right| \tag{76}$$
$$\leqq CL^{-2d} + \mathcal{C}(a, T_L)(K|\log L|)^{2d+4}L^{-d},$$

with $\|\mathcal{C}(a, T_L)\|_{\exp^\gamma} \leqq C(d, \lambda, K, \gamma)$ for any $\gamma < 1$. Choosing $\gamma \in (0, 1)$ and $B := C(d, \lambda, K, \gamma)(4K \log L)^{2+d}$ in Definition 6, defining the variable $X_0^{\log_2 L+1}$ (which may depend on a on the full volume $[0, L]^d$) to account for the remaining difference $a^{\mathrm{RVE}} e_i \cdot e_j - \sum_{k=0}^{1+\log_2 L_K} \sum_{x_0 \in 2^k \mathbb{Z}^d \cap [0,L)^d} X_{x_0}^k$, and setting the remaining

$X_i^k := 0$ for $\log_2 L_K + 1 < k < \log_2 L + 1$, establishes that a^{RVE} may be rewritten as a sum of a family of random variables with multilevel local dependence.

Part 3: The higher-order statistical quantity Next, we derive the multilevel decomposition of the higher-order quantity in the small ellipticity contrast setting $\mathcal{F}_{2-\text{point}}$. To do this, we decompose the solution v_i to (28) as

$$v_i(\cdot) = \int_0^\infty w_i(\cdot, s) \, \mathrm{d}s, \tag{77}$$

where w_i is defined as the solution to the parabolic PDE

$$\frac{\mathrm{d}}{\mathrm{d}t} w_i = \Delta w_i,$$
$$w_i(\cdot, 0) = \nabla \cdot (a e_i).$$

As before, the representation (77) follows from the exponential decay of w_i, as we have $-\Delta \int_0^T w_i(\cdot, t) \, \mathrm{d}t = \nabla \cdot (a e_i) - w_i(\cdot, T)$.

We introduce analogous definitions for v_j^*. Again, we may assume without loss of generality that $\varepsilon = 1$. We then observe, following an argument of Mourrat [72], that by formula (78) below, we have

$$
\begin{aligned}
\mathcal{F}_{2-\text{point}}(a) &= \fint_{[0,L]^d} a \nabla v_i \cdot e_j \, \mathrm{d}x \\
&= \fint_{[0,L]^d} -\nabla v_i \cdot \nabla v_j^* \, \mathrm{d}x \\
&= \fint_{[0,L]^d} \int_0^\infty \int_0^\infty -\nabla w_i(\cdot, s_1) \cdot \nabla w_j^*(\cdot, s_2) \, \mathrm{d}s_1 \, \mathrm{d}s_2 \, \mathrm{d}x \\
&\overset{(78)}{=} \fint_{[0,L]^d} \int_0^\infty \int_0^\infty -\nabla w_i\left(\cdot, \frac{s_1 + s_2}{2}\right) \cdot \nabla w_j^*\left(\cdot, \frac{s_1 + s_2}{2}\right) \, \mathrm{d}s_1 \, \mathrm{d}s_2 \, \mathrm{d}x.
\end{aligned}
$$

Next, we deduce

$$
\begin{aligned}
\mathcal{F}_{2-\text{point}}(a) &= -\fint_{[0,L]^d} \int_0^\infty 4s \, \nabla w_i(\cdot, s) \cdot \nabla w_j^*(\cdot, s) \, \mathrm{d}s \, \mathrm{d}x \\
&= -\sum_{x_0 \in \mathbb{Z}^d} \frac{1}{L^d} \int_0^1 \int_{(x_0 + [0,1]^d) \cap [0,L]^d} 4s \, \nabla w_i(\cdot, s) \cdot \nabla w_j^*(\cdot, s) \, \mathrm{d}x \, \mathrm{d}s \\
&\quad - \sum_{k=1}^{\log_2 L_K} \sum_{x_0 \in \mathbb{Z}^d} \frac{1}{L^d} \int_{4^{k-1}}^{4^k} \fint_{[0,L]^d} 4s \, \nabla w_i(\cdot, s) \cdot \nabla w_j^*(\cdot, s) \, \mathrm{d}x \, \mathrm{d}s \\
&\quad - \frac{1}{L^d} \int_{T_L}^\infty \fint_{[0,L]^d} 4s \, \nabla w_i(\cdot, s) \cdot \nabla w_j^*(\cdot, s) \, \mathrm{d}x \, \mathrm{d}s.
\end{aligned}
$$

We may now proceed to argue just as in the case of a^{RVE}. The required decay estimates for the semigroup of the form

$$\left(\fint_{\{|x - x_0| \leq \sqrt{T}\}} |\nabla w_i(\cdot, T)|^2 \, \mathrm{d}x \right)^{1/2} \leq \mathcal{C}(a, T, x_0) T^{-1 - d/4}$$

(with $||\mathcal{C}(a, T, x_0)||_{\exp 2} \leqq C(d, \lambda)$) are now a consequence of the explicit heat kernel representation of the solution w_i (as we are now dealing with a constant-coefficient parabolic equation), the finite range of dependence $\varepsilon = 1$ of the initial data $w_i(\cdot, 0) = \nabla \cdot (ae_i)$, and standard Gaussian concentration estimates (or, alternatively—though then with a less strong stretched exponential bound—the concentration estimates of Lemma 20).

In the computation above we have used the simple fact that

$$
\begin{aligned}
&\fint_{[0,L]^d} \nabla w_i(\cdot, s_1) \otimes \nabla w_j^*(\cdot, s_2) \, dx \\
&= \fint_{[0,L]^d} \nabla w_i\left(\cdot, \frac{s_1 + s_2}{2}\right) \otimes \nabla w_j^*\left(\cdot, \frac{s_1 + s_2}{2}\right) dx \\
&\quad - \int_0^1 \frac{d}{d\rho} \fint_{[0,L]^d} \nabla w_i\left(\cdot, \frac{(2-\rho)s_1 + \rho s_2}{2}\right) \otimes \nabla w_j^*\left(\cdot, \frac{\rho s_1 + (2-\rho)s_2}{2}\right) dx \, d\rho \\
&= \fint_{[0,L]^d} \nabla w_i\left(\cdot, \frac{s_1 + s_2}{2}\right) \otimes \nabla w_j^*\left(\cdot, \frac{s_1 + s_2}{2}\right) dx.
\end{aligned}
\tag{78}
$$

Part 4: Convergence of the variance Finally, we prove that the rescaled variances $L^d \mathrm{Var}\, a^{\mathrm{RVE}}$, $L^d \mathrm{Var}\, \mathcal{F}_{avg}(a)$, and $L^d \mathrm{Var}\, \mathcal{F}_{2-\mathrm{point}}(a)$ and the covariances $L^d \mathrm{Cov}[a^{\mathrm{RVE}}, \mathcal{F}_{avg}(a)]$, $L^d \mathrm{Cov}[a^{\mathrm{RVE}}, \mathcal{F}_{2-\mathrm{point}}(a)]$, and $L^d \mathrm{Cov}[\mathcal{F}_{avg}(a), \mathcal{F}_{2-\mathrm{point}}(a)]$ converge for $L \to \infty$. We limit ourselves to proving the convergence of the rescaled variance $L^d \mathrm{Var}\, a^{\mathrm{RVE}}$; the proofs for the convergence of the other variances and the covariances are analogous. Furthermore, to simplify notation, we limit ourselves to proving the convergence of the variance for $L = 2^n$ for some $n \in \mathbb{N}$; the proof in the general case is similar.

By Lemma 12, we obtain $\mathrm{Var}\, a^{\mathrm{RVE}} \leqq C(d, \lambda, K) L^{-d} |\log L|^{C(d)}$. Using (76) and this estimate, we deduce

$$
\left| \mathrm{Var}\, a^{\mathrm{RVE}} - \sum_{k=0}^{1+\log_2 L_K} \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{y \in 2^k \mathbb{Z}^d \cap [0,L)^d} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [0,L)^d} \mathrm{Cov}[X_y^k, X_{\tilde{y}}^{\tilde{k}}] \right|
$$
$$
\leqq C(d, \lambda, K) |\log L|^C L^{-3d/2}.
$$

Expanding the sum and using stochastic independence of many of these terms, we may write

$$
\left| \mathrm{Var}\, a^{\mathrm{RVE}} - \sum_{k=0}^{1+\log_2 L_K} \sum_{y \in 2^k \mathbb{Z}^d \cap [0,L)^d} \sum_{\substack{\tilde{y} \in 2^k \mathbb{Z}^d \cap [0,L)^d: \\ |y-\tilde{y}|_{\mathrm{per}} \leqq CK \log L \cdot 2^k}} \mathrm{Cov}[X_y^k, X_{\tilde{y}}^k] \right.
$$
$$
\left. - \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}+1}^{1+\log_2 L_K} \sum_{y \in 2^k \mathbb{Z}^d \cap [0,L)^d} \sum_{\substack{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [0,L)^d: \\ |y-\tilde{y}|_{\mathrm{per}} \leqq CK \log L \cdot 2^k}} (\mathrm{Cov}[X_y^k, X_{\tilde{y}}^{\tilde{k}}] + \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k}}, X_y^k]) \right|
$$
$$
\leqq C(d, \lambda, K) |\log L|^C L^{-3d/2}.
$$

Denote by $X_y^{k,\mathbb{R}^d}$ the quantities defined as in (71) but with u_{i,k,x_0} and u_{j,k,x_0}^* replaced by $u_i^{\mathbb{R}^d}$ and $u_j^{*,\mathbb{R}^d}$, that is for example, for $k \geqq 0$ and $y \in 2^k \mathbb{Z}^d$,

$$
\begin{aligned}
X_y^{k,\mathbb{R}^d} := & -\frac{1}{L^d} \int_{(y+[0,2^k]^d)} \int_{4^k}^{4^{k+1}} u_i^{\mathbb{R}^d}(\cdot, s) \, ds \, u_j^{*,\mathbb{R}^d}(\cdot, 4^k) \, dx \\
& -\frac{1}{L^d} \int_{(y+[0,2^k]^d)} u_i^{\mathbb{R}^d}(\cdot, 4^k) \int_{4^k}^{4^{k+1}} u_j^{*,\mathbb{R}^d}(\cdot, s) \, ds \, dx \\
& +\frac{1}{L^d} \int_{(y+[0,2^k]^d)} a\nabla \int_{4^k}^{4^{k+1}} u_i^{\mathbb{R}^d}(\cdot, s) \, ds \cdot \nabla \int_{4^k}^{4^{k+1}} u_j^{*,\mathbb{R}^d}(\cdot, s) \, ds \, dx.
\end{aligned}
$$

Set $X_y^{k,\infty} := L^d X_y^{k,\mathbb{R}^d}$. Note that $\mathrm{Cov}[X_y^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}]$ does not depend on L (by definition of $X_y^{k,\mathbb{R}^d}$). By the full-space variants of the estimates (66), (67), and (70) (that is the estimates for the differences $u_i^{\mathbb{R}^d} - u_{i,k,x_0}^{\mathbb{R}^d}$ etc., which are derived in exactly the same way) and (72) as well as the equality of laws of (products of the) u_{i,k,x_0} etc. and (products of the) $u_{i,k,x_0}^{\mathbb{R}^d}$ etc., we get for $k, \tilde{k} \leqq 1 + \log_2 L_K$ that

$$
\left| \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k}}, X_y^k] - \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\mathbb{R}^d}, X_y^{k,\mathbb{R}^d}] \right| \leqq C(d, \lambda, K) L^{-cK}. \tag{79}
$$

By the definition of the X_y^k (see (71)), the definition of the u_{i,k,x_0}, and the stationarity of the probability distribution of $a^{\mathbb{R}^d}$, the covariance $\mathrm{Cov}[X_y^{k,\mathbb{R}^d}, X_{\tilde{y}}^{\tilde{k},\mathbb{R}^d}]$ depends only on $k, \tilde{k}, y - \tilde{y}, L$, and the law of $a^{\mathbb{R}^d}$ (but not on y for fixed $y - \tilde{y}$). Furthermore, by (72) we have $|\mathrm{Cov}[X_{\tilde{y}}^{\tilde{k}}, X_y^k]| \leqq C L^{-2d}$. This implies, by (79),

$$
\begin{aligned}
& \left| \mathrm{Var}\, a^{\mathrm{RVE}} - \sum_{k=0}^{1+\log_2 L_K} \left(\frac{L}{2^k}\right)^d \sum_{\substack{\tilde{y} \in 2^k \mathbb{Z}^d \cap [-L/2, L/2]^d: \\ |\tilde{y}-0| \leqq CK \log L \cdot 2^k}} \mathrm{Cov}[X_0^{k,\mathbb{R}^d}, X_{\tilde{y}}^{k,\mathbb{R}^d}] \right. \\
& \left. - \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}+1}^{1+\log_2 L_K} \left(\frac{L}{2^k}\right)^d \right. \\
& \left. \quad\quad \times \sum_{\substack{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-L/2, L/2]^d: \\ |\tilde{y}-0| \leqq CK \log L \cdot 2^k}} (\mathrm{Cov}[X_0^{k,\mathbb{R}^d}, X_{\tilde{y}}^{\tilde{k},\mathbb{R}^d}] + \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\mathbb{R}^d}, X_0^{k,\mathbb{R}^d}]) \right| \\
& \leqq C(d, \lambda, K) |\log L|^C L^{-3d/2} + \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}}^{1+\log_2 L_K} \left(\frac{L}{2^k}\right)^d \cdot C(d, \lambda, K) L^{-cK} \\
& \leqq C(d, \lambda, K) |\log L|^C L^{-3d/2}
\end{aligned}
$$

for K chosen large enough.

The fact that (by stochastic independence) we have $\mathrm{Cov}[L^d X_{\tilde{y}}^{\tilde{k}}, L^d X_y^k] = 0$ for $|y - \tilde{y}|_{\mathrm{per}} \geqq C(d) 2^k K \log L$ and $k \geqq \tilde{k}$ implies together with (79) and the

definition of $X_y^{k,\infty}$ that (by selecting K large enough and by choosing L to be just small enough for $|y - \tilde{y}| \geq C(d)2^k K \log L$ to hold in case $|y - \tilde{y}| \geq C(d)K2^k$, and otherwise—that is for $|y - \tilde{y}| \leq C(d)K2^k$—appealing to the upper bound (72))

$$\left| \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\infty}, X_y^{k,\infty}] \right| \leq C(d, \lambda, K) \exp\left(-\frac{|y - \tilde{y}|}{C(d, \lambda)2^k} \right). \tag{80}$$

As a consequence, we obtain

$$\left| L^d \mathrm{Var}\ a^{\mathrm{RVE}} - \sum_{k=0}^{1+\log_2 L_K} (2^k)^{-d} \sum_{\tilde{y} \in 2^k \mathbb{Z}^d} \mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{k,\infty}] \right.$$

$$\left. - \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}+1}^{1+\log_2 L_K} (2^k)^{-d} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} (\mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}] + \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\infty}, X_0^{k,\infty}]) \right|$$

$$\leq C(d, \lambda, K)|\log L|^C L^{-d/2}$$

$$+ \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}}^{1+\log_2 L_K} (2^k)^{-d} \sum_{\substack{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d: \\ |\tilde{y}-0| > CK \log L \cdot 2^k}} C(d, \lambda, K) \exp\left(-\frac{|\tilde{y} - 0|}{C2^k} \right)$$

$$\leq C(d, \lambda, K)|\log L|^C L^{-d/2}$$

$$+ \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=\tilde{k}}^{1+\log_2 L_K} (2^k)^{-d} \cdot \left(\frac{2^k}{2^{\tilde{k}}} \right)^d C(d, \lambda, K) \exp(-cK \log L)$$

$$\leq C(d, \lambda, K)|\log L|^C L^{-d/2}.$$

This implies

$$\left| L^d \mathrm{Var}\ a^{\mathrm{RVE}} - \sum_{k=0}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^k \mathbb{Z}^d} \mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{k,\infty}] \right.$$

$$\left. - \sum_{\tilde{k}=0}^{\infty} \sum_{k=\tilde{k}+1}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} (\mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}] + \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\infty}, X_0^{k,\infty}]) \right|$$

$$\leq C(d, \lambda, K)|\log L|^C L^{-d/2}$$

$$+ 2 \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=2+\log_2 L_K}^{\infty} (2^k)^{-d} \left| \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} \mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}] \right|$$

$$+ 2 \sum_{\tilde{k}=2+\log_2 L_K}^{\infty} \sum_{k=\tilde{k}}^{\infty} (2^k)^{-d} \left| \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} \mathrm{Cov}[X_0^{k,\infty}, X_y^{\tilde{k},\infty}] \right|.$$

We now distinguish the cases $\tilde{y} \in [-R_k 2^k, R_k 2^k]^d$ and $\tilde{y} \notin [-R_k 2^k, R_k 2^k]^d$ for some R_k to be chosen. Using (80) in the latter case, we get

$$
\left| L^d \mathrm{Var}\ a^{\mathrm{RVE}} - \sum_{k=0}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^k \mathbb{Z}^d} \mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{k,\infty}] \right.
$$

$$
\left. - \sum_{\tilde{k}=0}^{\infty} \sum_{k=\tilde{k}+1}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} (\mathrm{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}] + \mathrm{Cov}[X_{\tilde{y}}^{\tilde{k},\infty}, X_0^{k,\infty}]) \right|
$$

$$
\leqq C(d, \lambda, K) |\log L|^C L^{-d/2}
$$

$$
+ 2 \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=2+\log_2 L_K}^{\infty} (2^k)^{-d} \left| \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-R_k 2^k, R_k 2^k]^d} \mathrm{Cov}[X_0^{k,\infty}, X_y^{\tilde{k},\infty}] \right|
$$

$$
+ 2 \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=2+\log_2 L_K}^{\infty} (2^k)^{-d} \cdot C(d, \lambda, K) \left(\frac{2^k}{2^{\tilde{k}}} \right)^d \exp\left(- \frac{R_k}{C} \right)
$$

$$
+ 2 \sum_{\tilde{k}=2+\log_2 L_K}^{\infty} \sum_{k=\tilde{k}}^{\infty} (2^k)^{-d} \left| \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-R_k 2^k, R_k 2^k]^d} \mathrm{Cov}[X_0^{k,\infty}, X_y^{\tilde{k},\infty}] \right|
$$

$$
+ 2 \sum_{\tilde{k}=2+\log_2 L_K}^{\infty} \sum_{k=\tilde{k}}^{\infty} (2^k)^{-d} \cdot C(d, \lambda, K) \left(\frac{2^k}{2^{\tilde{k}}} \right)^d \exp\left(- \frac{R_k}{C} \right).
$$

For $\tilde{k} \leqq k$ and $R2^k \leqq L_K$ we have, by Lemma 12 and (72),

$$
\left| \mathrm{Cov}\left[X_y^k, \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-R2^k, R2^k]^d} X_{\tilde{y}}^{\tilde{k}} \right] \right|
$$

$$
\leqq \sqrt{|\mathrm{Var}\ X_y^k|} \sqrt{\left| \mathrm{Var} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-R2^k, R2^k]^d} X_{\tilde{y}}^{\tilde{k}} \right|}
$$

$$
\leqq C(d, \lambda, K) L^{-2d} \left(\frac{R2^k}{2^{\tilde{k}}} \right)^{d/2} \left| \log \frac{R2^k}{2^{\tilde{k}}} \right|^{d/2},
$$

which entails, by (79), upon choosing $L^{1/2} = R2^k$,

$$
\left| \mathrm{Cov}\left[X_y^{k,\infty}, \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d \cap [-R2^k, R2^k]^d} X_{\tilde{y}}^{\tilde{k},\infty} \right] \right|
$$

$$
\leqq C(d, \lambda, K) \left(\frac{R2^k}{2^{\tilde{k}}} \right)^{d/2} |\log(R2^k)|^d.
$$

As a consequence, choosing $R_k = Sk$ for $S \geqq 1$ large enough, we get

$$\left| L^d \operatorname{Var} a^{\mathrm{RVE}} - \sum_{k=0}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^k \mathbb{Z}^d} \operatorname{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{k,\infty}] \right.$$

$$\left. - \sum_{\tilde{k}=0}^{\infty} \sum_{k=\tilde{k}+1}^{\infty} (2^k)^{-d} \sum_{\tilde{y} \in 2^{\tilde{k}} \mathbb{Z}^d} (\operatorname{Cov}[X_0^{k,\infty}, X_{\tilde{y}}^{\tilde{k},\infty}] + \operatorname{Cov}[X_{\tilde{y}}^{\tilde{k},\infty}, X_0^{k,\infty}]) \right|$$

$$\leqq C(d, \lambda, K) |\log L|^C L^{-d/2}$$

$$+ 2 \sum_{\tilde{k}=0}^{1+\log_2 L_K} \sum_{k=2+\log_2 L_K}^{\infty} (2^k)^{-d} \cdot C(d, \lambda, K) \left(\frac{R_k 2^k}{2^{\tilde{k}}} \right)^{d/2} |\log(R_k 2^k)|^d$$

$$+ C(d, \lambda, K) \sum_{k=2+\log_2 L_K}^{\infty} \exp\left(-\frac{R_k}{C} \right)$$

$$+ 2 \sum_{\tilde{k}=2+\log_2 L_K}^{\infty} \sum_{k=\tilde{k}}^{\infty} (2^k)^{-d} \cdot C(d, \lambda, K) \left(\frac{R_k 2^k}{2^{\tilde{k}}} \right)^{d/2} |\log(R_k 2^k)|^d$$

$$+ C(d, \lambda, K) (L_K)^{-d} \sum_{k=0}^{\infty} \exp\left(-\frac{R_k}{C} \right)$$

$$\leqq C(d, \lambda, K) |\log L|^C L^{-d/2}$$

$$+ C(d, \lambda, K) (L_K)^{-d/2} |\log L_K|^C$$

$$+ C(d, \lambda, K) \exp\left(-\frac{S \log L_K}{C} \right)$$

$$+ C(d, \lambda, K) (L_K)^{-d} |\log L_K|^C$$

$$+ C(d, \lambda, K) (L_K)^{-d}$$

$$\leqq C(d, \lambda, K) |\log L|^C L^{-d/2}.$$

In total, we have shown convergence of the rescaled variance $L^d \operatorname{Var} a^{\mathrm{RVE}}$ towards a limit independent of L with the desired rate.

The proof of the other cases is analogous. $\quad\square$

Proof of Theorem 11. The estimate (62a) is contained in [55, Corollary 4]. In view of the Poincaré inequality the bound (62b) is a consequence of (62a) and an estimate on a (weighted) average of $u_i^{\mathbb{R}^d}$. Hence, we only need to derive a bound on

$$\int u_i^{\mathbb{R}^d}(\cdot, T) \frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}} \right) \mathrm{d}x$$

for a suitably chosen smooth function ψ supported in $\{|x| \leqq 1\}$. To this end, we compute

$$\int u_i^{\mathbb{R}^d}(\cdot, T) \frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}}\right) dx$$

$$= \int u_i^{\mathbb{R}^d}(\cdot, 2T) \frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}}\right) dx - \int_T^{2T} \int \frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}}\right) \frac{d}{dt} u_i^{\mathbb{R}^d} \, dx \, dt$$

$$= \int u_i^{\mathbb{R}^d}(\cdot, 2T) \frac{1}{\sqrt{2T}^d} \psi\left(\frac{x}{\sqrt{2T}}\right) dx$$

$$+ \int u_i^{\mathbb{R}^d}(\cdot, 2T) \left(\frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}}\right) - \frac{1}{\sqrt{2T}^d} \psi\left(\frac{x}{\sqrt{2T}}\right)\right) dx$$

$$- \int_T^{2T} \int \frac{1}{\sqrt{T}^{d+1}} \nabla\psi\left(\frac{x}{\sqrt{T}}\right) \cdot a \nabla u_i^{\mathbb{R}^d} \, dx \, dt,$$

which yields upon applying the Poincaré inequality to the second term (note that the second factor in the integral has vanishing average) and using the bound (62a)

$$\left| \int u_i^{\mathbb{R}^d}(\cdot, T) \frac{1}{\sqrt{T}^d} \psi\left(\frac{x}{\sqrt{T}}\right) dx - \int u_i^{\mathbb{R}^d}(\cdot, 2T) \frac{1}{\sqrt{2T}^d} \psi\left(\frac{x}{\sqrt{2T}}\right) dx \right|$$

$$\leq C(d)\mathcal{C}(a, 2T)(2T)^{-1/2-d/4} + C(d, \lambda) \int_T^{2T} \mathcal{C}(a, t) t^{-1-d/4} \sqrt{T}^{-1} \, dt$$

$$\leq C(d)\mathcal{C}(a, T) T^{-1/2-d/4}.$$

Summing over a dyadic sequence of times $2^k T$ and using the fact that almost surely

$$\lim_{T \to \infty} \int u_i^{\mathbb{R}^d}(\cdot, T) \sqrt{T}^{-d} \psi(x/\sqrt{T}) \, dx = 0,$$

we infer (62b) (upon redefining the constant $\mathcal{C}(a, T)$). $\square$

In the previous proofs, we have made use of the following elementary concentration estimate for sums of random variables with multilevel local dependence:

Lemma 12. ([43], Lemma 9) *Consider a probability distribution of uniformly elliptic and bounded coefficient fields a on $\mathbb{R}^d$ or a periodization of such a probability distribution, and suppose that assumptions (A1)–(A3) respectively (A1), (A2), $(A3_a)$–$(A3_c)$ are satisfied. Let $X = X(a)$ be a random variable that is approximately a sum of random variables with multilevel local dependence in the sense of Definition 6. Then for $\tilde{\gamma} := \gamma/(\gamma + 1)$ the concentration estimate*

$$||X - \mathbb{E}[X]||_{\exp^{\tilde{\gamma}}} \leq C(d, \gamma, K) B |\log L|^{d/2} L^{-d/2}$$

holds true.

6. Failure and Success of the Variance Reduction Approaches

We now establish our theorems on the failure and the success of the variance reduction approaches in stochastic homogenization. We start with the counterexample that shows that in general there is no guarantee that the variance reduction techniques provide an effective reduction of the variance, even for "natural" choices of the statistical quantity $\mathcal{F}(a)$ like the spatial average $\mathcal{F}_{avg}(a) := \fint_{[0,L\varepsilon]^d} a \, dx$.

Proof of Theorem 4. Before turning to the main result of Theorem 4, the failure of the spatial average $\mathcal{F}_{avg}(a)$ to explain a fraction of the variance of a^{RVE} (inequality (23)), let us first show (22). The estimate (22) is in fact a consequence of the estimate (43) in the proof of Theorem 2 in combination with (41) and the lower bound for the variance of $\mathcal{M}^\delta$ which is a straightforward consequence of the formula (34) and the definition of $\text{Var } a_{ij}^{\text{RVE}}|_{\text{unexpl}} = (1 - |\rho|^2)\text{Var } a_{ij}^{\text{RVE}}$.

Note that the derivation of (24) from (23) requires the estimate (22) under the assumption (A2') instead of (A2). However, the only place where the assumption (A2) entered in our analysis is in Proposition 7, where it was used to apply the result of [55] on the decay of the semigroup. However, the arguments of [55] may be modified to yield the corresponding estimate under the assumption of discrete stationarity (A2').

Let us now turn to the construction of our counterexample featuring the degenerate covariance (23). The construction is based on the following ideas:

- The approximation a^{RVE} for the effective coefficient depends in a uniformly continuous way on a as a map $L^\infty([0, L\varepsilon]^d; \mathbb{R}^{d \times d}) \to \mathbb{R}^{d \times d}$, as long as a is uniformly elliptic and bounded.
- Consider a probability distribution of coefficient fields a for which a is almost surely almost everywhere a multiple of the identity matrix. If in addition the law of a is invariant under reflections of coordinate axes and invariant under exchange of coordinate axes (that is, invariant under diagonal reflections), the covariance

$$\text{Cov}\left[a^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a \, dx\right]$$

is a multiple of $\text{Id} \otimes \text{Id}$. For a proof of this fact, see Lemma 13, below.
- Consider the "periodized random checkerboard" with the set of tiles $\mathcal{T} := \{x_0 + [0, \varepsilon)^d : x_0 \in \varepsilon\mathbb{Z}^d \cap [0, L\varepsilon)^d\}$. On each tile $T \in \mathcal{T}$, choose at random (and independently from the other tiles) $a(x) = \text{Id}$ with probability 0.5 and $a(x) = \frac{1}{2} \text{Id}$ with probability 0.5. By Proposition 5 and the preceding considerations, for this probability distribution the covariance

$$\text{Cov}\left[a^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a \, dx\right]$$

is a positive multiple of $\text{Id} \otimes \text{Id}$; in fact, one has a lower bound of the form $\gtrsim L^{-d} \text{Id} \otimes \text{Id}$.
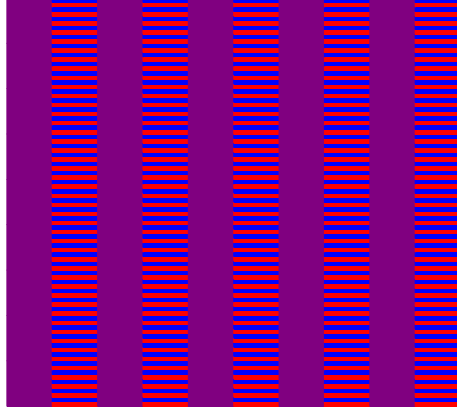
Fig. 5. A single tile with (second-order laminate) microstructure, as used in the proof of Theorem 4. Blue corresponds to the regions with $a(x) = \lambda\,\mathrm{Id}$, red to the regions with $a(x) = \mathrm{Id}$, and violet to the regions with $a(x) = \mu\,\mathrm{Id}$

- We now consider a "periodized random checkerboard with microstructure" with the set of tiles $\mathcal{T} := [0, \varepsilon)^d + (\varepsilon\mathbb{Z}^d \cap [0, L\varepsilon)^d)$: Fix some $\tau \ll 1$ with $1/\tau \in 2\mathbb{N}$. On each tile $T = \varepsilon k + [0, \varepsilon)^d \in \mathcal{T}$, choose at random (and independently from the other tiles) $a_\tau(x) = \sigma\,\mathrm{Id}$ with probability 0.5 (where $\sigma > 0$ is to be chosen below) and $a_\tau(x) = A_\tau((x - \varepsilon k)/\varepsilon)$ with probability 0.5, where $A_\tau : [0, 1]^2 \to \mathbb{R}^{2\times 2}$ is the tile described in Fig. 5, rotated and reflected at random (with equal probability for all 8 orientations and independently on all such tiles; see Fig. 6 for an illustration).

  The probability distribution of a satisfies the same isotropy properties as in the case of the periodized random checkerboard. Thus, by Lemma 13 the covariance

  $$\mathrm{Cov}\left[a_\tau^{\mathrm{RVE}}, \fint_{[0, L\varepsilon]^d} a_\tau\,\mathrm{d}x\right]$$

  is a multiple of $\mathrm{Id} \otimes \mathrm{Id}$.

- We shall argue below that for suitable $\sigma, \lambda, \mu > 0$ and for $\tau \ll 1$ small enough the covariance

  $$\mathrm{Cov}\left[a_\tau^{\mathrm{RVE}}, \fint_{[0, L\varepsilon]^d} a_\tau\,\mathrm{d}x\right]$$

  is negative; in fact, one has an upper bound of the form $\lesssim -L^{-d}\,\mathrm{Id} \otimes \mathrm{Id}$.

- Linearly interpolating between a_τ and a—that is, considering for $\kappa \in [0, 1]$ the coefficient field

  $$a_{\tau,\kappa} := (1 - \kappa)a + \kappa a_\tau$$

  defined on the product probability space, that is for independent a_τ and a—we find a probability distribution of coefficient fields $\tilde{a}$ for which the covariance

  $$\mathrm{Cov}\left[\tilde{a}^{\mathrm{RVE}}, \fint_{[0, L\varepsilon]^d} \tilde{a}\,\mathrm{d}x\right]$$
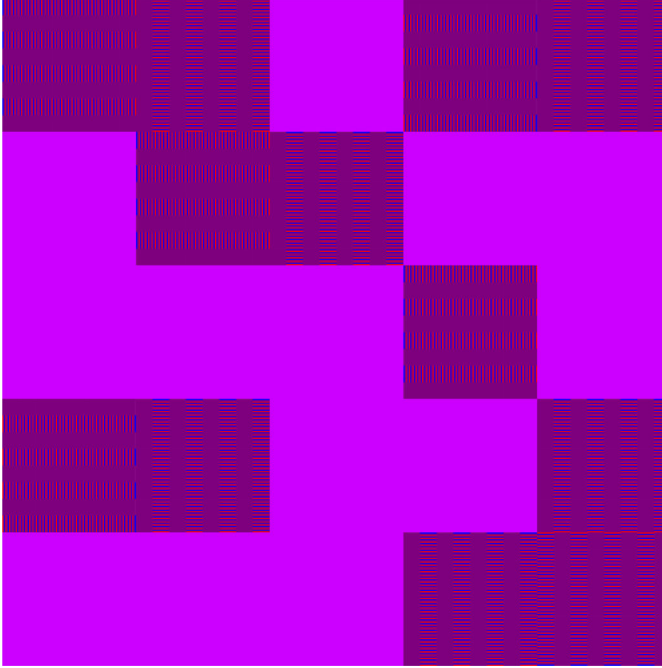
Fig. 6. A single realization of the probability distribution of our counterexample (with an exaggerated size of the microstructure in the tiles with microstructure). The tiles with microstructure behave almost like a homogeneous tile with an effective conductivity. Note that the tiles with microstructure are oriented randomly in order to enforce exact isotropy of the (co-)variances Var a^{RVE} and Cov$[a^{\mathrm{RVE}}, \fint a \, dx]$

vanishes. This is possible by the continuous dependence of a^{RVE} and $\fint_{[0, L\varepsilon]^d} a \, dx$ on a (and hence the continuous dependence on $\kappa \in [0, 1]$ in the case of the family $a_{\tau, \kappa}$) and by the fact that for all $\kappa \in [0, 1]$ the covariance

$$\mathrm{Cov}\left[a_{\tau, \kappa}^{\mathrm{RVE}}, \fint_{[0, L\varepsilon]^d} a_{\tau, \kappa} \, dx\right]$$

is a multiple of $\mathrm{Id} \otimes \mathrm{Id}$ (this latter property holds again by the isotropy properties of the probability distribution and Lemma 13, below).

- For any $\kappa \in (0, 1)$ the variances Var $\fint_{[0, L\varepsilon]^d} a_{\tau, \kappa} \, dx$ and Var $a_{\tau, \kappa}^{\mathrm{RVE}}$ are non-degenerate in the sense $\gtrsim L^{-d} \mathrm{Id} \otimes \mathrm{Id}$. For the spatial average $\fint_{[0, L\varepsilon]^d} a_{\tau, \kappa} \, dx$ this non-degeneracy is an easy consequence of the formula

$$\mathrm{Var} \fint_{[0, L\varepsilon]^d} a_{\tau, \kappa} \, dx = (1 - \kappa)^2 \mathrm{Var} \fint_{[0, L\varepsilon]^d} a \, dx + \kappa^2 \mathrm{Var} \fint_{[0, L\varepsilon]^d} a_\tau \, dx$$

(which follows from the definition of $a_{\tau, \kappa}$ and the independence of a and a_τ) and the fact that the latter two variances satisfy such a lower bound (note that the spatial average of the coefficient field on a tile with microstructure A_τ does not equal $\sigma \mathrm{Id}$). The non-degeneracy of Var $a_{\tau, \kappa}^{\mathrm{RVE}}$ is shown as follows:

first, a new coefficient field $a_{\tau,\kappa,\text{eff}}$ is introduced by letting $a_{\tau,\kappa,\text{eff}} = a_{\tau,\kappa}$ on each tile without microstructure but replacing the values of $a_{\tau,\kappa}$ by the effective coefficient from periodic homogenization on each tile with microstructure. Note that $a_{\tau,\kappa,\text{eff}}$ corresponds to a standard random checkerboard. Denote by $a_{\tau,\kappa,\text{eff}}^{\text{RVE}}$ the approximation for the effective coefficient associated with the coefficient field $a_{\tau,\kappa,\text{eff}}$ (that is the result of formula (8) for the coefficient field $a_{\tau,\kappa,\text{eff}}$). The nondegeneracy of Var $a_{\tau,\kappa}^{\text{RVE}}$ now follows from the nondegeneracy Var $a_{\tau,\kappa,\text{eff},ii}^{\text{RVE}} \gtrsim L^{-d}$ and the convergence $|a_{\tau,\kappa}^{\text{RVE}} - a_{\tau,\kappa,\text{eff}}^{\text{RVE}}| \to 0$ for $\tau \to 0$ (uniformly in κ, see below).

Note that $a_{\tau,\kappa,\text{eff}}^{\text{RVE}}$ corresponds to a random checkerboard with tiles $(\kappa\sigma + (1 - \kappa))$ Id, $\kappa\sigma + (1 - \kappa) \cdot \frac{1}{2}$ Id, $\kappa A_\tau + (1 - \kappa)$ Id, and $\kappa A_\tau + (1 - \kappa) \cdot \frac{1}{2}$ Id, each tile chosen with probability $\frac{1}{4}$ (and the microscopic tiles rotated and reflected at random). Thus the nondegeneracy of Var $a_{\tau,\kappa,\text{eff},ii}^{\text{RVE}}$ for $1 \leq i \leq d$ follows from the covariance estimate of Proposition 5 and the quantitative upper bound Var $\fint_{[0,L\varepsilon]^d} a_{\tau,\kappa,\text{eff}} \, \mathrm{d}x \leq CL^{-d}$.

To complete the proof, it only remains to establish the negativity of the covariance

$$\text{Cov}\left[a_\tau^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a_\tau \, \mathrm{d}x\right]$$

for $\tau \ll 1$ small enough and suitable σ, μ, λ, as well as the convergence $a_{\tau,\kappa}^{\text{RVE}} \to a_{\tau,\kappa,\text{eff}}^{\text{RVE}}$ for $\tau \to 0$, uniformly in κ. The underlying idea for our choice of the tiles in Fig. 5 is that we intend to exploit the nonlinear dependence of the effective coefficients in periodic homogenization on the coefficient field, equipping such a tile with an effective coefficient that is unrelated to the spatial average of the coefficient field. Heuristically, by classical results in periodic homogenization we expect the following to happen:

- Consider our (sub)pattern of periodic horizontal stripes of equal height (that is the red-and-blue subpattern in Fig. 5), in which the coefficient field a alternatingly takes the values Id and λ Id. Then the (large-scale) effective coefficient for this pattern is given by

$$\begin{pmatrix} \frac{1+\lambda}{2} & 0 \\ 0 & \frac{2\lambda}{1+\lambda} \end{pmatrix},$$

that is by the arithmetic mean in the horizontal direction and by the harmonic mean in the vertical direction.

- Consider now the pattern of periodic vertical stripes of equal width, in which the coefficient alternatingly takes the value μ Id respectively is given by the pattern of horizontal stripes from the previous step. The effective coefficient for this (second-order laminate) pattern is (at least in the limit of an infinitesimally fine horizontal pattern) given by the arithmetic mean of the effective coefficients in the vertical direction and the harmonic mean of the effective coefficients in the horizontal direction, that is by

$$\begin{pmatrix} \frac{2\mu(1+\lambda)}{2\mu+1+\lambda} & 0 \\ 0 & \frac{\lambda}{1+\lambda} + \frac{\mu}{2}. \end{pmatrix}.$$

Choosing $\mu := \frac{3\lambda^2+(1-\lambda)\sqrt{9\lambda^2+14\lambda+9}+2\lambda+3}{4(\lambda+1)}$—which is positive for any $\lambda \in (0, 1]$—, the effective coefficient becomes a multiple of the identity matrix. Note that the spatial average of the coefficient field on a tile is given by

$$\frac{\mu + \frac{\lambda+1}{2}}{2} \text{ Id}.$$

- Consider the coefficient field $a_{\tau,\text{eff}}$ that is obtained from our random checkerboard with microstructure a_τ by replacing a_τ on the tiles with microstructure with the effective coefficient $(\frac{\lambda}{1+\lambda} + \frac{\mu}{2})$ Id. The coefficient field $a_{\tau,\text{eff}}$ is now just a usual random checkerboard; by Lemma 13 and Proposition 5, the covariance

$$\text{Cov}\left[a_{\tau,\text{eff}}^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a_{\tau,\text{eff}} \, dx\right]$$

is a positive multiple of $\text{Id} \otimes \text{Id}$, and we have a lower bound of the form $\geqq cL^{-d} \text{ Id} \otimes \text{Id}$ for the choice of λ, μ, and τ to be made below. Note that $a_{\tau,\text{eff}}$—and hence also the preceding covariance—is actually independent of τ (we just keep the τ to emphasize that $a_{\tau,\text{eff}}$ is the coefficient field obtained from a_τ in the homogenization limit $\tau \to 0$). We shall prove below that a_τ^{RVE} is (quantitatively) close to $a_{\tau,\text{eff}}^{\text{RVE}}$ for $\tau \ll 1$ small enough, which implies that

$$\text{Cov}\left[a_\tau^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a_{\tau,\text{eff}} \, dx\right]$$

is close to a positive multiple of $\text{Id} \otimes \text{Id}$ (again with a lower bound of the form $\geqq cL^{-d} \text{ Id} \otimes \text{Id}$).

- The average $\fint_{[0,L\varepsilon]^d} a_\tau \, dx$ is an affine function of $\fint_{[0,L\varepsilon]^d} a_{\tau,\text{eff}} \, dx$: The coefficient field $a_{\tau,\text{eff}}$ is constant on each tile and may only take the values σ Id or $(\frac{\lambda}{1+\lambda} + \frac{\mu}{2})$ Id. On the tiles on which the value of $a_{\tau,\text{eff}}$ is σ Id, a_τ also takes the constant value σ Id. However, on the tiles on which $a_{\tau,\text{eff}}$ is given by $(\frac{\lambda}{1+\lambda} + \frac{\mu}{2})$ Id (that is on the tiles on which a_τ features a microstructure), the average of a_τ is $\frac{2\mu+\lambda+1}{4}$ Id. We thus have

$$\fint_{[0,L\varepsilon]^d} a_\tau \, dx = \frac{\sharp \text{microtiles}}{L^d} \cdot \frac{2\mu + \lambda + 1}{4} \text{ Id} + \frac{L^d - \sharp \text{microtiles}}{L^d} \cdot \sigma \text{ Id}$$

$$= \sigma \text{ Id} + \frac{\sharp \text{microtiles}}{L^d} \cdot \left(\frac{2\mu + \lambda + 1}{4} - \sigma\right) \text{ Id}$$

and

$$\fint_{[0,L\varepsilon]^d} a_{\tau,\text{eff}} \, dx = \frac{\sharp \text{microtiles}}{L^d} \cdot \left(\frac{\lambda}{1 + \lambda} + \frac{\mu}{2}\right) \text{ Id} + \frac{L^d - \sharp \text{microtiles}}{L^d} \cdot \sigma \text{ Id}$$

$$= \sigma \text{ Id} + \frac{\sharp \text{microtiles}}{L^d} \cdot \left(\frac{\lambda}{1 + \lambda} + \frac{\mu}{2} - \sigma\right).$$

Choosing σ such that $\sigma > \frac{\lambda}{1+\lambda} + \frac{\mu}{2}$ but $\sigma < \frac{2\mu+\lambda+1}{4}$—which is possible for $\lambda > 0$ small enough—, we obtain a relation of the form

$$\fint_{[0,L\varepsilon]^d} a_\tau \, dx = A \operatorname{Id} - B \fint_{[0,L\varepsilon]^d} a_{\tau,\mathrm{eff}} \, dx$$

for suitable positive constants A and B. Thus, the sign of the covariance flips upon replacing the $a_{\tau,\mathrm{eff}}$ by a_τ in the spatial average, that is

$$\operatorname{Cov}\left[a_\tau^{\mathrm{RVE}}, \fint_{[0,L\varepsilon]^d} a_\tau \, dx\right]$$

must be a negative multiple of $\operatorname{Id} \otimes \operatorname{Id}$, with an upper bound of the form $\leqq -cL^{-d} \operatorname{Id} \otimes \operatorname{Id}$.

It now only remains to prove two things: We need to show that a_τ^{RVE} is quantitatively close to $a_{\tau,\mathrm{eff}}$ if we choose the width τ of the vertical stripes and the height τ^2 of the horizontal stripes in the pattern in Fig. 5 small enough and we need to establish the corresponding assertion for the interpolated coefficient field $a_{\tau,\kappa,\mathrm{eff}}$. As the latter result is shown similarly—though with two different microscopic tiles $\kappa A_\tau + (1-\kappa)\frac{1}{2}\operatorname{Id}$ and $\kappa A_\tau + (1-\kappa)\frac{1}{2}\operatorname{Id}$, depending on whether the random checkerboard a equals Id or $\frac{1}{2}\operatorname{Id}$ on the tile (and correspondingly, with two sets of homogenization correctors and two characteristic functions $\chi_{microtile1}$ and $\chi_{microtile2}$, see below for this notation)—we only provide the proof of the latter result.

For the remainder of the proof, we shall fix without loss of generality $\varepsilon := 1$ to avoid even more cumbersome notation. Again, to avoid even more cumbersome notation, we only give the proof in the case that all tiles with microstructure have the same orientation as in Fig. 5.

To see this quantitative closeness, we construct an approximate homogenization corrector $\phi_{i,\mathrm{appr}}$ for a_τ^{RVE}. To this end, let $\phi_{i,\mathrm{eff}}$ be the homogenization corrector associated with the coefficient field $a_{\tau,\mathrm{eff}}$, that is let $\phi_{i,\mathrm{eff}}$ solve

$$-\nabla \cdot (a_{\tau,\mathrm{eff}}(e_i + \nabla\phi_{i,\mathrm{eff}})) = 0$$

on $[0, L]^2$ with periodic boundary conditions. We now intend to build the approximate homogenization corrector $\phi_{i,\mathrm{appr}}$ for a_τ^{RVE} by a nested two-scale expansion, using the homogenization correctors for the periodic laminate microstructures.

By Meyer's estimate, there exists $p > 2$ with

$$\fint_{[0,L]^2} |\nabla\phi_{i,\mathrm{eff}}|^p \, dx \leqq C(d,\lambda). \tag{81}$$

Furthermore, $a_{\tau,\mathrm{eff}}$ is constant on each tile $k + [0,1)^2$, which implies on each tile $T = k + [0,1)^2$ (with $k \in \mathbb{Z}^2$) for each $x \in T$ by regularity theory for constant coefficient equations

$$|\nabla^2\phi_{i,\mathrm{eff}}(x)| \leqq \frac{C}{\mathrm{dist}(x,\partial T)}\left(\fint_{\{|y-x|\leqq\mathrm{dist}(x,\partial T)/2\}} |e_i + \nabla\phi_{i,\mathrm{eff}}|^2 \, dy\right)^{1/2}$$

$$\leqq \frac{C}{\mathrm{dist}(x,\partial T)^{1+d/2}}\left(\fint_T |e_i + \nabla\phi_{i,\mathrm{eff}}|^2 \, dy\right)^{1/2}. \tag{82}$$

Let ρ_δ denote a standard mollifier. The L^p estimate and the estimate on $\nabla^2 \phi_{i,\text{eff}}$ imply (for notational convenience we extend $\phi_{i,\text{eff}}$ by periodicity)

$$\fint_{[0,L]^2} \left| \nabla \phi_{i,\text{eff}} - \nabla(\rho_\delta * \phi_{i,\text{eff}}) \right|^{(p+2)/2} dx \le C\delta^\alpha \tag{83}$$

for some $\alpha > 0$ (for a proof of this estimate, split the domain into a neighborhood of size $\delta^{1/5}$ of the tile boundaries ∂T, on which one uses the Hölder inequality and the L^p bound on $\nabla \phi_{i,\text{eff}}$ in (81), and the interior $\{x \in T : \text{dist}(x, \partial T) \ge \delta^{1/5}\}$, where one applies the regularity estimate (82)).

Let $\phi_{i,h}$ denote the 2-periodic homogenization corrector for the coefficient field $a_h(x, y)$ associated with the pattern of horizontal stripes in Fig. 5 (that is let $a_h(x, y) = a_h(y)$ take alternatingly on intervals of length 1 the values Id and λ Id). Note that $\phi_{1,h} \equiv 0$ and that $\phi_{2,h}$ is explicitly given by

$$\phi_{2,h}(x, y) = \frac{1}{\fint_0^2 \frac{1}{e_2 \cdot a_h(x, \tilde{y})e_2} d\tilde{y}} \int_0^y \frac{1}{e_2 \cdot a_h(x, \tilde{y})e_2} d\tilde{y} - y.$$

We shall frequently use the uniform bound on the gradient $|\nabla \phi_{i,h}| \le C$ derived easily from this formula.

Let $\phi_{i,v}$ denote the 2-periodic homogenization corrector associated with the pattern of vertical stripes of width 1, in which the coefficient field $a_v(x, y) = a_v(x)$ alternatingly takes the values μ Id and

$$\begin{pmatrix} \frac{1+\lambda}{2} & 0 \\ 0 & \frac{2\lambda}{1+\lambda} \end{pmatrix}.$$

Note that we have $\phi_{2,v} \equiv 0$ and that $\phi_{1,v}$ is given explicitly by

$$\phi_{1,v}(x, y) = \frac{1}{\fint_0^2 \frac{1}{e_1 \cdot a_v(\tilde{x}, y)e_1} d\tilde{x}} \int_0^x \frac{1}{e_1 \cdot a_v(\tilde{x}, y)e_1} d\tilde{x} - x.$$

We shall again frequently use the uniform bound on the gradient $|\nabla \phi_{i,v}| \le C$.

We define the vector potential for the flux correction $\sigma_{h,ijk}$, skew-symmetric in its last two indices, as $\sigma_{h,212} := 0$ and

$$\sigma_{h,112} := \int_0^y (a_h(\tilde{y}) - a_{h,\text{eff}})e_1 \cdot e_1 \, d\tilde{y}. \tag{84}$$

Note that with this definition $\sigma_{h,ijk}$ satisfies $\nabla \cdot \sigma_{h,i} = a_h(e_i + \nabla \phi_{i,h}) - a_{h,\text{eff}}e_i$, as one checks by a case-by-case analysis.

Similarly, we define $\sigma_{v,ijk}$, skew-symmetric in its last two indices, as $\sigma_{v,121} := 0$ and

$$\sigma_{v,221} := \int_0^x (a_v(\tilde{x}) - a_{v,\text{eff}})e_2 \cdot e_2 \, d\tilde{x}, \tag{85}$$

which then satisfies $\nabla \cdot \sigma_{v,i} = a_v(e_i + \nabla \phi_{i,v}) - a_{v,\text{eff}}e_i$.

Let us denote the indicator function of the tiles with microstructure by $\chi_{\text{microtile}}$ (that is $\chi_{\text{microtile}}$ is 1 on all tiles $k + [0, 1)^d \subset [0, L)^d$ with microstructure and 0 on

the other tiles). Similarly, we denote by $\chi_{v\text{microstripe}}$ the indicator functions of all vertical stripes that according to Fig. 5 contain a micropattern of horizontal stripes. We then build our approximate correctors as

$$
\phi_{i,\text{appr},1} := \rho_{\delta_0} * \phi_{i,\text{eff}}
$$
$$
+ (\rho_{\tau\delta_1} * \chi_{\text{microtile}}) \sum_j (\delta_{ij} + \partial_j(\rho_{\delta_0} * \phi_{i,\text{eff}})) \left(\rho_{\delta_1\tau} * \tau\phi_{j,v}\left(\frac{\cdot}{\tau}\right) \right)
$$

and

$$
\phi_{i,\text{appr},2} := \phi_{i,\text{appr},1} + (\rho_{\tau^2\delta_2} * \chi_{v\text{microstripe}}) \sum_k (\partial_k\phi_{i,\text{appr},1} + \delta_{ik})\tau^2\phi_{k,h}\left(\frac{\cdot}{\tau^2}\right).
$$

We observe that $\phi_{i,\text{appr},1}$ satisfies the estimate

$$
|\nabla\phi_{i,\text{appr},1}| \leqq \left(\frac{C}{\min\{1,\delta_1\}} + \frac{C\tau}{\delta_0} \right)(\rho_{2\delta_0} * |\nabla\phi_{i,\text{eff}}| + 1). \tag{86}
$$

We also have the bound

$$
|\nabla\phi_{i,\text{appr},2}| \leqq \frac{C}{\min\{1,\delta_2\}}(|\nabla\phi_{i,\text{appr},1}| + 1) \tag{87}
$$
$$
+ \left(C + \frac{C\tau^2}{\delta_0} + \frac{C\tau}{\delta_1^2} + \frac{C\tau^3}{\delta_0^2} + \frac{C\tau}{\delta_1} \right)(\rho_{2\delta_0} * |\nabla\phi_{i,\text{eff}}| + 1).
$$

Furthermore, if we are at least $\tau\delta_1$ away from the tile boundaries and the boundaries of the vertical stripes (note that $\rho_{\delta_1\tau} * \nabla\phi_{j,v}(\cdot/\tau)$ is then equal to $\nabla\phi_{j,v}(\cdot/\tau)$ as the latter quantity is constant in each stripe; note also that then $\rho_{\tau\delta_1} * \chi_{\text{microtile}}$ is locally constant $= 0$ or $= 1$ and that we have a uniform bound on $\nabla\phi_{j,v}$), we have by (82) on each tile $T = k + [0,1)^2$, $k \in \mathbb{Z}^d \cap [0,L)^d$,

$$
\left| e_i + \nabla\phi_{i,\text{appr},1} - \sum_j (e_j + \chi_{\text{microtile}}\nabla\phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j\phi_{i,\text{eff}}) \right| \tag{88}
$$

$$
\leqq \frac{C}{\text{dist}(\cdot,\partial T)^2} \left(\fint_T |e_i + \nabla\phi_{i,\text{eff}}|^2 \, dx \right)^{1/2} (\delta_0 + \tau).
$$

If we are at least $\tau\delta_1$ away from the tile boundaries and the boundaries of the vertical stripes and at least $\tau^2\delta_2$ away from the boundary of the horizontal stripes, we get (note that $\rho_{\delta_2\tau^2} * \nabla\phi_{k,h}(\cdot/\tau^2)$ is then equal to $\nabla\phi_{k,h}(\cdot/\tau^2)$ as the latter quantity is constant in each small horizontal stripe; note also that then $\rho_{\tau^2\delta_2} * \chi_{h\text{microstripe}}$ is locally constant $= 0$ or $= 1$ and that we have a uniform bound on $\nabla\phi_{k,h}$)

$$
\left| e_i + \nabla\phi_{i,\text{appr},2} \right.
$$
$$
\left. - \sum_k (e_k + \chi_{v\text{microstripe}}\nabla\phi_{k,h}(\cdot/\tau^2)) \sum_j \left(\delta_{jk} + \chi_{\text{microtile}}\partial_k\phi_{j,v}(\cdot/\tau)\right)(\delta_{ij} + \partial_j\phi_{i,\text{eff}}) \right|
$$
$$
\overset{(88)}{\leqq} C \left| e_i + \nabla\phi_{i,\text{appr},1} - \sum_j (e_j + \chi_{\text{microtile}}\nabla\phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j\phi_{i,\text{eff}}) \right|
$$

$$+ C\tau^2 |\nabla^2 \phi_{i,\text{appr},1}| + \frac{C}{\text{dist}(\cdot, \partial T)^2} \left(\fint_T |e_i + \nabla\phi_{i,\text{eff}}|^2 \, dx \right)^{1/2} (\delta_0 + \tau)$$

$$\overset{(88),(82)}{\leqq} \frac{C}{\text{dist}(\cdot, \partial T)^2} \left(\fint_T |e_i + \nabla\phi_{i,\text{eff}}|^2 \, dx \right)^{1/2} \left(\delta_0 + \tau + \tau^2 + \frac{\tau^3}{\delta_0} \right).$$

Using the fact that by Meyers inequality we have for some $p = p(\lambda) > 2$ that

$$\fint_{[0,L]^2} |e_i + \nabla\phi_{i,\text{eff}}|^p \, dx \leqq C(d, \lambda),$$

we obtain, by choosing δ_0, δ_1, and δ_2 as appropriate powers of τ and using (87),

$$\fint_{[0,L\varepsilon]^d} \left| e_i + \nabla\phi_{i,\text{appr},2} - \sum_k (e_k + \chi_{v\text{microstripe}} \nabla\phi_{k,h}(\cdot/\tau^2)) \right.$$

$$\left. \times \sum_j (\delta_{jk} + \chi_{\text{microtile}} \partial_k \phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j \phi_{i,\text{eff}}) \right|^2 dx$$

$$\leqq C(d, \lambda)\tau^\eta \tag{89}$$

for some $\eta > 0$.

Having bounded the error in the gradient, we next estimate the error in the flux. In an analogous fashion to the definition of $a_{\tau,\text{eff}}$ as the effective coefficient from periodic homogenization on each tile, we define $a_{\tau,\text{veff}}$ as equal to $a_{\tau,\text{eff}} = a_\tau$ on the tiles without microstructure and equal to the effective coefficient from periodic homogenization on each vertical stripe of width τ on each tile with microstructure. Recalling the definitions (84) and (85), we may rewrite the error in the flux in a pointwise way as

$$a_\tau \sum_k (e_k + \chi_{v\text{microstripe}} \nabla\phi_{k,h}(\cdot/\tau^2)) \sum_j (\delta_{jk} + \chi_{\text{microtile}} \partial_k \phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j \phi_{i,\text{eff}})$$

$$-a_{\tau,\text{eff}}(e_i + \nabla\phi_{i,\text{eff}})$$

$$= \sum_j \left(a_\tau \sum_k (e_k + \chi_{v\text{microstripe}} \nabla\phi_{k,h}(\cdot/\tau^2)) - a_{\tau,\text{veff}} e_k \right)$$

$$\times (\delta_{jk} + \chi_{\text{microtile}} \partial_k \phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j \phi_{i,\text{eff}})$$

$$+ \sum_j \left(a_{\tau,\text{veff}}(e_j + \chi_{\text{microtile}} \nabla\phi_{j,v}(\cdot/\tau)) - a_{\tau,\text{eff}} e_j \right)(\delta_{ij} + \partial_j \phi_{i,\text{eff}})$$

$$= \chi_{v\text{microstripe}} \sum_k (\nabla \cdot (\tau^2 \sigma_{h,k}(\cdot/\tau^2))) \sum_j (\delta_{jk} + \chi_{\text{microtile}} \partial_k \phi_{j,v}(\cdot/\tau))(\delta_{ij} + \partial_j \phi_{i,\text{eff}})$$

$$+ \chi_{\text{microtile}} \sum_j (\nabla \cdot (\tau \sigma_{v,j}(\cdot/\tau)))(\delta_{ij} + \partial_j \phi_{i,\text{eff}}). \tag{90}$$

Thus, having chosen δ_0, δ_1, and δ_2 as suitable powers of τ, we obtain, by (89), (82) and (81),

$$\left| \fint_{[0,L]^2} a_\tau (e_i + \nabla\phi_{i,\text{appr},2}) \, dx - \fint_{[0,L]^2} a_{\tau,\text{eff}}(e_i + \nabla\phi_{i,\text{eff}}) \, dx \right|$$

$$\leqq C(d, \lambda)\tau^\eta.$$

It now only remains to show that $\nabla\phi_{i,\text{appr},2}$ is a good approximation for $\nabla\phi_i$. To do so, we consider the difference $\phi_i - \phi_{i,\text{appr},2}$ and observe that it satisfies the PDE

$$-\nabla \cdot (a_\tau(\nabla\phi_i - \nabla\phi_{i,\text{appr},2}))$$
$$= \nabla \cdot (a_\tau(e_i + \nabla\phi_{i,\text{appr},2}))$$
$$= \nabla \cdot (a_\tau(e_i + \nabla\phi_{i,\text{appr},2}) - a_{\tau,\text{eff}}(e_i + \nabla\phi_{i,\text{eff}})).$$

We now replace the divergence-form right-hand side using (89)

$$-\nabla \cdot (a_\tau(\nabla\phi_i - \nabla\phi_{i,\text{appr},2})$$
$$= \nabla \cdot g$$
$$+ \nabla \cdot \left(a_\tau \sum_k (e_k + \chi_{v\text{microstripe}}\nabla\phi_{k,h}(\cdot/\tau^2)) \right.$$
$$\times \sum_j \left(\delta_{jk} + \chi_{\text{microtile}}\partial_k\phi_{j,v}(\cdot/\tau)\right)(\delta_{ij} + \partial_j\phi_{i,\text{eff}})$$
$$\left. - a_{\tau,\text{eff}}(e_i + \nabla\phi_{i,\text{eff}}) \right)$$

for some g with $\int_{[0,L]^2} |g|^2 \leqq C\tau^\eta$ (recall that δ_1 and δ_2 have been chosen as a suitable small powers of τ and recall also the uniform L^p bound for $\nabla\phi_{i,\text{eff}}$ in (81)). This expression in turn may be rewritten by (83) and (90) for any $\beta > 0$ small enough as

$$-\nabla \cdot (a_\tau(\nabla\phi_i - \nabla\phi_{i,\text{appr},2}))$$
$$= \nabla \cdot \tilde{g} + \nabla \cdot \left((\rho_{\tau^{1+\beta}} * \chi_{v\text{microstripe}}) \sum_k (\nabla \cdot (\tau^2\sigma_{h,k}(\cdot/\tau^2))) \right.$$
$$\times \sum_j \left(\delta_{jk} + \rho_{\tau^{1+\beta}} * \chi_{\text{microtile}}\partial_k\phi_{j,v}(\cdot/\tau)\right)(\delta_{ij} + \rho_{\tau^\beta} * \partial_j\phi_{i,\text{eff}})$$
$$\left. + (\rho_{\tau^\beta} * \chi_{\text{microtile}}) \sum_j (\nabla \cdot (\tau\sigma_{v,j}(\cdot/\tau)))(\delta_{ij} + \rho_{\tau^\beta} * \partial_j\phi_{i,\text{eff}}) \right)$$

for some $\tilde{g}$ with $\int_{[0,L]^2} |\tilde{g}|^2 \leqq C\tau^{\text{delta}(\beta)}$.

Using the skew-symmetry of $\sigma_{v,i}$ and $\sigma_{h,i}$, we obtain

$$-\nabla \cdot (a_\tau(\nabla\phi_i - \nabla\phi_{i,\text{appr},2}))$$
$$= \nabla \cdot \tilde{g}$$
$$+ \sum_k (\nabla \cdot (\tau^2\sigma_{h,k}(\cdot/\tau^2)))$$
$$\cdot \nabla \sum_j (\rho_{\tau^{1+\beta}} * \chi_{v\text{microstripe}})\left(\delta_{jk} + \rho_{\tau^{1+\beta}} * \chi_{\text{microtile}}\partial_k\phi_{j,v}(\cdot/\tau)\right)$$
$$\times (\delta_{ij} + \rho_{\tau^\beta} * \partial_j\phi_{i,\text{eff}})$$
$$+ \sum_j (\nabla \cdot (\tau\sigma_{v,j}(\cdot/\tau))) \cdot \nabla\left((\rho_{\tau^\beta} * \chi_{\text{microtile}})(\delta_{ij} + \rho_{\tau^\beta} * \partial_j\phi_{i,\text{eff}})\right).$$

Using again the skew-symmetry of $\sigma_{v,i}$ and $\sigma_{h,i}$, we get

$$
\begin{aligned}
&- \nabla \cdot (a_\tau (\nabla \phi_i - \nabla \phi_{i,\text{appr},2})) \\
&= \nabla \cdot \tilde{g} \\
&\quad - \nabla \cdot \left(\sum_k \tau^2 \sigma_{h,k}(\cdot/\tau^2) \cdot \nabla \sum_j (\rho_{\tau^{1+\beta}} * \chi_{v\text{microstripe}}) \right. \\
&\qquad\qquad \times \left. \big(\delta_{jk} + \rho_{\tau^{1+\beta}} * \chi_{\text{microtile}} \partial_k \phi_{j,v}(\cdot/\tau) \big) \big(\delta_{ij} + \rho_{\tau^\beta} * \partial_j \phi_{i,\text{eff}} \big) \right) \\
&\quad - \nabla \cdot \left(\sum_j \tau \sigma_{v,j}(\cdot/\tau) \nabla \big((\rho_{\tau^\beta} * \chi_{\text{microtile}})(\delta_{ij} + \rho_{\tau^\beta} * \partial_j \phi_{i,\text{eff}}) \big) \right).
\end{aligned}
$$

Choosing $\beta > 0$ small enough, we finally end up with

$$
-\nabla \cdot (a_\tau (\nabla \phi_i - \nabla \phi_{i,\text{appr},2})) = \nabla \cdot \hat{g},
$$

with $\fint_{[0,L]^d} |\hat{g}|^2 \leqq C\tau^{\tilde{\nu}}$ for some $\tilde{\nu} > 0$. A standard energy estimate now implies

$$
\fint_{[0,L)^d} |\nabla \phi_i - \nabla \phi_{i,\text{appr},2}|^2 \, dx \leqq C\tau^{\tilde{\nu}}.
$$

$\square$

Lemma 13. *Consider a probability distribution of coefficient fields a subject to the conditions (A1), (A2), and (A3$_a$)–(A3$_c$). Suppose in addition that a is almost surely almost everywhere a multiple of the identity matrix. If, in addition, the law of a is invariant under reflections of coordinate axes (that is maps of the form $x \mapsto (x_1, \ldots, -x_i, \ldots, x_d)$) and invariant under exchange of coordinate axes (that is maps of the form $x \mapsto (x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_d)$), the covariance*

$$
\text{Cov}\left[a^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a \, dx \right]
$$

is a multiple of $\text{Id} \otimes \text{Id}$.

Proof. For such a probability distribution of coefficient fields a, the spatial average $\fint_{[0,L\varepsilon]^d} a \, dx$ is almost surely a multiple of the identity matrix, which entails that

$$
\text{Cov}\left[a^{\text{RVE}}, \fint_{[0,L\varepsilon]^d} a \, dx \right] = B \otimes \text{Id}
$$

for some $B \in \mathbb{R}^{d \times d}$.

The matrix B must also be a multiple of the identity matrix. Under reflection of the i-th coordinate, by the corrector equation (3) and the fact that a is pointwise a multiple of the identity matrix we have that the i-th corrector for the reflected coefficient field $\hat{a}(x) = a(x_1, \ldots, -x_i, \ldots, x_d)$ is given by $\hat{\phi}_i(x) = $

$-\phi_i(x_1, \ldots, -x_i, \ldots, x_d)$. Thus, the off-diagonal entries of a^{RVE} which are given by (for $i \neq j$, using also that $a(x) = a_{\text{scalar}}(x)\,\text{Id}$)

$$a^{\text{RVE}}e_i \cdot e_j = \fint_{[0,L\varepsilon]^d} a(e_i + \nabla\phi_i) \cdot e_j \, dx = \fint_{[0,L\varepsilon]^d} a_{scalar}(x)(e_j \cdot \nabla)\phi_i(x) \, dx$$

switch sign under such reflections, while the average $\fint_{[0,L\varepsilon]^d} a \, dx$ remains invariant. As our probability distribution is invariant under reflections, the off-diagonal entries of B must be zero. Similarly, as our probability distribution is invariant under exchange of coordinates, all diagonal entries of B must coincide; therefore the covariance must be a multiple of $\text{Id} \otimes \text{Id}$. □

We now turn to the proof of our theorem on successful variance reduction for random coefficient fields that are obtained by applying "monotone" functions to a collection of iid random variables.

Proof of Proposition 5. Without loss of generality (by rescaling), we may consider the case $\varepsilon = 1$.

Given any $\xi \in \mathbb{R}^d$, the L-periodic correctors associated with two L-periodic coefficient fields a and $\tilde{a}$ are given as the solutions to the PDEs

$$-\nabla \cdot (a\nabla\phi_\xi^{L,a}) = \nabla \cdot (a\xi) \tag{91}$$

and

$$-\nabla \cdot (\tilde{a}\nabla\phi_\xi^{L,\tilde{a}}) = \nabla \cdot (\tilde{a}\xi).$$

Define $\phi_\xi^{L,(1-s)a+s\tilde{a}}$ as the L-periodic solution to

$$-\nabla \cdot (((1-s)a + s\tilde{a})\nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) = \nabla \cdot (((1-s)a + s\tilde{a})\xi). \tag{92}$$

Setting

$$a^{\text{RVE},s}\xi \cdot \xi := \fint_{[0,L]^d} ((1-s)a + s\tilde{a})(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot \xi \, dx,$$

we then obtain

$$\frac{d}{ds}a^{\text{RVE},s}\xi \cdot \xi$$

$$= \frac{d}{ds}\fint_{[0,L]^d} ((1-s)a + s\tilde{a})(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot \xi \, dx$$

$$\stackrel{(92)}{=} \frac{d}{ds}\fint_{[0,L]^d} ((1-s)a + s\tilde{a})(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, dx$$

$$= \fint_{[0,L]^d} (\tilde{a} - a)(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, dx$$

$$+ 2\fint_{[0,L]^d} ((1-s)a + s\tilde{a})\nabla\frac{d}{ds}\phi_\xi^{L,(1-s)a+s\tilde{a}} \cdot (\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, dx$$

$$\stackrel{(92)}{=} \fint_{[0,L]^d} (\tilde{a} - a)(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, dx.$$

Given two coefficient fields a and $\tilde{a}$ with $a - \tilde{a} \geq 0$, we therefore have the estimate

$$a^{\mathrm{RVE},a}\xi \cdot \xi - a^{\mathrm{RVE},\tilde{a}}\xi \cdot \xi \tag{93}$$

$$\geqq \int_0^1 \fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, \mathrm{d}x \, \mathrm{d}s.$$

We now would like to derive a lower bound for the term on the right-hand side. We have, by (91) and (92),

$$-\nabla \cdot \left(((1-s)a + s\tilde{a})(\nabla\phi_\xi^{L,(1-s)a+s\tilde{a}} - \nabla\phi_\xi^{L,a})\right) = \nabla \cdot (s(\tilde{a} - a)(\xi + \nabla\phi_\xi^{L,a})).$$

Testing this PDE by the solution (note that $(1-s)a + s\tilde{a}$ is λ-uniformly elliptic) yields

$$\fint_{[0,L]^d} \lambda|\nabla\phi_\xi^{L,a} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}|^2 \, \mathrm{d}x$$

$$\leqq s\fint_{[0,L]^d} (\tilde{a} - a)(\xi + \nabla\phi_\xi^{L,a}) \cdot \left(\nabla\phi_\xi^{L,a} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}\right) \mathrm{d}x,$$

and therefore by Young's inequality (note that the matrix $a - \tilde{a}$ is symmetric and by (A1) bounded by $\frac{1}{\lambda}$ in the natural matrix norm), we have

$$\fint_{[0,L]^d} |\nabla\phi_\xi^{L,a} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}|^2 \, \mathrm{d}x$$

$$\leqq \frac{s^2}{\lambda^4}\fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,a}) \cdot (\xi + \nabla\phi_\xi^{L,a}) \, \mathrm{d}x.$$

In particular, we obtain, by (93) (and the analogous version of the previous estimate for $\phi_\xi^{L,\tilde{a}}$ instead of $\phi_\xi^{L,a}$) and $a \geq \tilde{a}$,

$$a^{\mathrm{RVE},a}\xi \cdot \xi - a^{\mathrm{RVE},\tilde{a}}\xi \cdot \xi$$

$$\geqq \int_0^{\lambda^2/2} \frac{1}{2}\fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,a}) \cdot (\xi + \nabla\phi_\xi^{L,a}) \, \mathrm{d}x$$

$$- 2\fint_{[0,L]^d} (a - \tilde{a})(\nabla\phi_\xi^{L,a} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\nabla\phi_\xi^{L,a} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, \mathrm{d}x \, \mathrm{d}s$$

$$+ \int_{1-\lambda^2/2}^1 \frac{1}{2}\fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,\tilde{a}}) \, \mathrm{d}x$$

$$- 2\fint_{[0,L]^d} (a - \tilde{a})(\nabla\phi_\xi^{L,\tilde{a}} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \cdot (\nabla\phi_\xi^{L,\tilde{a}} - \nabla\phi_\xi^{L,(1-s)a+s\tilde{a}}) \, \mathrm{d}x \, \mathrm{d}s$$

$$\geqq \int_0^{\lambda^2/2} \left(\frac{1}{2} - \frac{2s^2}{\lambda^4}\right)\fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,a}) \cdot (\xi + \nabla\phi_\xi^{L,a}) \, \mathrm{d}x \, \mathrm{d}s$$

$$+ \int_{1-\lambda^2/2}^1 \left(\frac{1}{2} - \frac{2(1-s)^2}{\lambda^4}\right)\fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla\phi_\xi^{L,\tilde{a}}) \cdot (\xi + \nabla\phi_\xi^{L,\tilde{a}}) \, \mathrm{d}x \, \mathrm{d}s$$

$$\geq \frac{\lambda^2}{8} \fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla \phi_\xi^{L,a}) \cdot (\xi + \nabla \phi_\xi^{L,a}) \, dx$$

$$+ \frac{\lambda^2}{8} \fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla \phi_\xi^{L,\tilde{a}}) \cdot (\xi + \nabla \phi_\xi^{L,a}) \, dx.$$

This entails

$$\left(a^{\text{RVE},a} \xi \cdot \xi - a^{\text{RVE},\tilde{a}} \xi \cdot \xi \right) \left(\fint_{[0,L]^d} a\xi \cdot \xi \, dx - \fint_{[0,L]^d} \tilde{a}\xi \cdot \xi \, dx \right)$$

$$\geq \frac{\lambda^2}{8} \left(\fint_{[0,L]^d} (a - \tilde{a})\xi \cdot \xi \, dx \right) \fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla \phi_\xi^{L,a}) \cdot (\xi + \nabla \phi_\xi^{L,a}) \, dx$$

$$+ \frac{\lambda^2}{8} \left(\fint_{[0,L]^d} (a - \tilde{a})\xi \cdot \xi \, dx \right) \fint_{[0,L]^d} (a - \tilde{a})(\xi + \nabla \phi_\xi^{L,\tilde{a}}) \cdot (\xi + \nabla \phi_\xi^{L,\tilde{a}}) \, dx.$$

The estimate (95) from Lemma 14 implies

$$\text{Cov}\left[a^{\text{RVE}} \xi \cdot \xi, \fint_{[0,L]^d} a\xi \cdot \xi \, dx \right]$$

$$\geq \frac{\lambda^2}{16} L^{-d} \mathbb{E}\Bigg[\sum_{k \in \mathbb{Z}^d \cap [0,L)^d} \sqrt{\fint_{[0,L]^d} |(a(\Gamma) - a(\Delta_{k,\tilde{\Gamma}_k}\Gamma))\xi \cdot \xi| \, dx}$$

$$\times \sqrt{\fint_{[0,L]^d} |(a(\Gamma) - a(\Delta_{k,\tilde{\Gamma}_k}\Gamma))(\xi + \nabla \phi_\xi^{L,a}) \cdot (\xi + \nabla \phi_\xi^{L,a})| \, dx}\,\Bigg]^2$$

$$\geq \frac{\lambda^2}{16} L^{-d} \mathbb{E}\Bigg[\sum_{k \in \mathbb{Z}^d \cap [0,L)^d} L^{-d} \sqrt{\int_{[0,L]^d} |(a - a(\Delta_{k,\tilde{\Gamma}_k}\Gamma))\xi \cdot \xi| \, dx}$$

$$\times (2K)^{-d/2} \int_{[0,L]^d} |(a(\Gamma) - a(\Delta_{k,\tilde{\Gamma}_k}\Gamma))^{1/2}(\xi + \nabla \phi_\xi^{L,a})| \, dx \Bigg]^2,$$

where in the last step we have used the Hölder inequality and the fact that $a(x,\Gamma) - a(x, \Delta_{k,\tilde{\Gamma}_k}\Gamma)$ is only nonzero for $|x - k| \leq K$.

By our assumption (25) we infer

$$\text{Cov}\left[a^{\text{RVE}} \xi \cdot \xi, \fint_{[0,L]^d} a\xi \cdot \xi \, dx \right]$$

$$\geq \frac{\lambda^2}{16} L^{-d} \mathbb{E}\left[L^{-d}(2K)^{-d/2} \int_{[0,L]^d} \nu |\xi + \nabla \phi_\xi^{L,a}| \, dx \right]^2$$

$$\geq \frac{\lambda^2}{16} L^{-d}(2K)^{-d} \nu^2 \mathbb{E}\left[\left| \fint_{[0,L]^d} \xi + \nabla \phi_\xi^{L,a} \, dx \right| \right]^2$$

$$\geq \frac{\lambda^2}{16} L^{-d}(2K)^{-d} \nu^2 |\xi|^2.$$

To conclude our proof, by

$$\rho_{a^{\mathrm{RVE}}\xi \cdot \xi, \mathcal{F}(a)} = \frac{\mathrm{Cov}[a^{\mathrm{RVE}}\xi \cdot \xi, \mathcal{F}(a)]}{\sqrt{\mathrm{Var}\ a^{\mathrm{RVE}}\xi \cdot \xi}\sqrt{\mathrm{Var}\ \mathcal{F}(a)}},$$

it suffices to bound $\mathrm{Var}\ a^{\mathrm{RVE}}\xi \cdot \xi$ and $\mathrm{Var}\ \mathcal{F}(a)$ by $C(d, \lambda, K)L^{-d}|\xi|^2$. A corresponding bound for $\mathrm{Var}\ a^{\mathrm{RVE}}\xi \cdot \xi$ is provided for example by the methods of GLORIA and OTTO [55]. To estimate $\mathrm{Var}\ \mathcal{F}(a)$, we simply apply (96), which yields

$$\mathrm{Var}\ \mathcal{F}(a)$$
$$\leqq \frac{1}{2} \sum_{k \in \mathbb{Z}^d \cap [0,L)^d} \mathbb{E}\left[\left(\fint_{[0,L)^d}(a(x, \Gamma) - a(x, \Delta_{k, \tilde{\Gamma}(k)}\Gamma))\xi \cdot \xi\, \mathrm{d}x\right)^2\right]$$
$$\leqq \frac{1}{2} \sum_{k \in \mathbb{Z}^d \cap [0,L)^d} \mathbb{E}\left[|\xi|^2 L^{-2d}(2K)^{2d}\right]$$
$$\leqq (2K)^{2d}|\xi|^2 L^{-d}.$$

$\square$

In the previous proof, we have used the following standard estimate for covariances of nonlinear functions of a finite number of independent random variables:

Lemma 14. *Let* $f : [0, 1]^N \to \mathbb{R}$, $g : [0, 1]^N \to \mathbb{R}$ *be two functions that are monotonous with respect to each of their arguments. Let* $X_i : \Omega \to [0, 1]$, $1 \leqq i \leqq N$, *and* $Y_i : \Omega \to [0, 1]$, $1 \leqq i \leqq N$, *be* $2N$ *independent identically distributed random variables. Define*

$$h_n(X, x, y)$$
$$:= |f(X_1, \ldots, X_{n-1}, x, X_{n+1}, \ldots, X_N) - f(X_1, \ldots, X_{n-1}, y, X_{n+1}, \ldots, X_N)|$$
$$\times |g(X_1, \ldots, X_{n-1}, x, X_{n+1}, \ldots, X_N) - g(X_1, \ldots, X_{n-1}, y, X_{n+1}, \ldots, X_N)|$$

and

$$H_n(X, x, y)$$
$$:= \frac{1}{2}|f(X_1, \ldots, X_{n-1}, x, X_{n+1}, \ldots, X_N) - f(X_1, \ldots, X_{n-1}, y, X_{n+1}, \ldots, X_N)|^2$$
$$+ \frac{1}{2}|g(X_1, \ldots, X_{n-1}, x, X_{n+1}, \ldots, X_N) - g(X_1, \ldots, X_{n-1}, y, X_{n+1}, \ldots, X_N)|^2.$$

Then

$$\mathrm{Cov}[f(X), g(X)] \geqq \frac{1}{2} \sum_{n=1}^{N} \mathbb{E}\left[\sqrt{h_n(X, X_n, Y_n)}\right]^2, \tag{94}$$

and, by Jensen's inequality,

$$\mathrm{Cov}[f(X), g(X)] \geqq \frac{1}{2} N^{-1} \mathbb{E}\left[\sum_{n=1}^{N} \sqrt{h_n(X, X_n, Y_n)}\right]^2. \tag{95}$$

Furthermore, we have

$$\text{Cov}[f(X), g(X)] \leqq \frac{1}{2} \sum_{n=1}^{N} \mathbb{E}\big[H_n(X, X_n, Y_n)\big]. \tag{96}$$

Proof. The proof proceeds similarly to the proof of the standard form of this lemma which provides the weaker assertion $\text{Cov}[f(X), g(X)] \geqq 0$; see for example [68, page 24] or [23, Lemma 2.1].

We have by the identity of the laws of $(X_1, \ldots, X_{n-1}, Y_n, Y_1, \ldots, Y_{n-1}, X_n)$ and $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$ (which allow us to swap X_n and Y_n in the expectations below),

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N) g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$= \frac{1}{2}\mathbb{E}\Big[\big(f(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots, X_N) - f(X_1, \ldots, X_{n-1}, Y_n, X_{n+1}, \ldots, X_N)\big)$$
$$\times \big(g(Y_1, \ldots, Y_{n-1}, X_n, X_{n+1}, \ldots, X_N) - g(Y_1, \ldots, Y_{n-1}, Y_n, X_{n+1}, \ldots, X_N)\big)\Big]$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N) g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big].$$

By the independence of the X_i and the Y_i, we infer

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N) g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$= \frac{1}{2}\mathbb{E}\Big[\int_{[0,1]^{n-1}} f(x, X_n, X_{n+1}, \ldots, X_N) - f(x, Y_n, X_{n+1}, \ldots, X_N)\, d\mathbb{P}_{(X_1, \ldots, X_{n-1})}(x)$$
$$\times \int_{[0,1]^{n-1}} g(y, X_n, X_{n+1}, \ldots, X_N) - g(y, Y_n, X_{n+1}, \ldots, X_N)\, d\mathbb{P}_{(Y_1, \ldots, Y_{n-1})}(y)\Big]$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N) g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big].$$

As both f and g are increasing functions in each of their arguments, the integrands in this formula are either nonnegative (for $X_n \geqq Y_n$) or nonpositive (for $X_n \leqq Y_n$). Thus, we have

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N) g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$= \frac{1}{2}\mathbb{E}\Big[\int_{[0,1]^{n-1}} |f(x, X_n, X_{n+1}, \ldots, X_N) - f(x, Y_n, X_{n+1}, \ldots, X_N)|\, d\mathbb{P}_{(X_1, \ldots, X_{n-1})}(x)$$
$$\times \int_{[0,1]^{n-1}} |g(y, X_n, X_{n+1}, \ldots, X_N) - g(y, Y_n, X_{n+1}, \ldots, X_N)|\, d\mathbb{P}_{(Y_1, \ldots, Y_{n-1})}(y)\Big]$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N) g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big] \tag{97}$$

and therefore by Hölder's inequality we have

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N) g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$\geqq \frac{1}{2}\mathbb{E}\Big[\sqrt{|f(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots, X_N) - f(X_1, \ldots, X_{n-1}, Y_n, X_{n+1}, \ldots, X_N)|}$$
$$\times \sqrt{|g(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots, X_N) - g(X_1, \ldots, X_{n-1}, Y_n, X_{n+1}, \ldots, X_N)|}\Big]^2$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N) g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big].$$

Taking the sum of these formulas for $n = 1, \ldots, N$, we infer

$$\mathbb{E}\big[f(X)g(X)\big] \geqq \frac{1}{2}\sum_{n=1}^{N}\mathbb{E}\Big[\sqrt{h_n(X, X_n, Y_n)}\Big]^2 + \mathbb{E}\big[f(X)g(Y)\big],$$

which establishes the desired lower bound (94) for the covariance.

To obtain (96), we apply Young's inequality and subsequently Jensen's inequality to (97), which yields

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N)g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$\leqq \frac{1}{2}\mathbb{E}\Big[\frac{1}{2}\int_{[0,1]^{n-1}}|f(x, X_n, X_{n+1}, \ldots, X_N) - f(x, Y_n, X_{n+1}, \ldots, X_N)|^2 \, d\mathbb{P}_{(X_1,\ldots,X_{n-1})}(x)$$
$$+ \frac{1}{2}\int_{[0,1]^{n-1}}|g(y, X_n, X_{n+1}, \ldots, X_N) - g(y, Y_n, X_{n+1}, \ldots, X_N)|^2 \, d\mathbb{P}_{(Y_1,\ldots,Y_{n-1})}(y)\Big]$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N)g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big].$$

This is equivalent to

$$\mathbb{E}\big[f(X_1, \ldots, X_{n-1}, X_n, \ldots, X_N)g(Y_1, \ldots, Y_{n-1}, X_n, \ldots, X_N)\big]$$
$$\leqq \frac{1}{2}\mathbb{E}\Big[\frac{1}{2}|f(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots, X_N) - f(X_1, \ldots, X_{n-1}, Y_n, X_{n+1}, \ldots, X_N)|^2$$
$$+ \frac{1}{2}|g(X_1, \ldots, X_{n-1}, X_n, X_{n+1}, \ldots, X_N) - g(X_1, \ldots, X_{n-1}, Y_n, X_{n+1}, \ldots, X_N)|^2\Big]$$
$$+ \mathbb{E}\big[f(X_1, \ldots, X_n, X_{n+1}, \ldots, X_N)g(Y_1, \ldots, Y_n, X_{n+1}, \ldots, X_N)\big].$$

Taking the sum with respect to n entails

$$\mathbb{E}\big[f(X)g(X)\big] \leqq \frac{1}{2}\sum_{n=1}^{N}\mathbb{E}[H_n(X, X_n, Y_n)] + \mathbb{E}\big[f(X)g(Y)\big],$$

which establishes the upper bound (96) for the covariance. □

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A. Gaussian Propagation Bounds for Parabolic PDEs

We now collect some elementary energy and propagation estimates for second-order linear parabolic equations. By a *nongrowing weak solution u* to the equation $\partial_t u = \nabla \cdot (a\nabla u)$ with initial data $u(\cdot, 0) = g$, we understand a function $u \in L^2_{loc}(\mathbb{R}^d \times [0, \infty))$ with $\nabla u \in L^2_{loc}(\mathbb{R}^d \times [0, \infty))$ satisfying the usual weak formulation of the PDE with test functions in $C^\infty_{cpt}(\mathbb{R}^d \times [0, \infty))$ and additionally the estimate

$$\sup_{r \geq 0} \int_0^T \fint_{\{|x| \leq r\}} |u|^2 \, dx \, dt < \infty$$

for any $T > 0$. Note that for initial data $u(\cdot, 0) = \nabla \cdot b$ for some vector field $b \in L^\infty(\mathbb{R}^d; \mathbb{R}^d)$, the initial data is incorporated into the weak formulation in a weak form, that is as

$$-\int_0^\infty \int_{\mathbb{R}^d} u \partial_t \eta \, dx \, dt = -\int_{\mathbb{R}^d} b \cdot \nabla \eta \, dx + \int_0^\infty \int_{\mathbb{R}^d} a\nabla u \cdot \nabla \eta \, dx \, dt.$$

Many of our computations in the next sections will be formal, but can be justified by the appropriate standard approximation arguments. Note also that the estimates which we shall prove ensure the *existence* of such nongrowing weak solutions for merely $b \in L^\infty(\mathbb{R}^d; \mathbb{R}^d)$, as they ensure that one may construct a solution by constructing solutions with the initial data b truncated outside of some large ball $\{|x| \leq R\}$ (in which case the standard existence theorems apply) and then passing to the limit $R \to \infty$.

Lemma 15. *Let a be a uniformly elliptic and bounded coefficient field on $\mathbb{R}^d$. For $r \geq 0$ and $M \geq 5d$, define the coefficient field*

$$a_{r,M}(x) := \begin{cases} a(x) & \text{for} |x| \leq Mr, \\ \text{Id} & \text{otherwise.} \end{cases}$$

Consider the unique nongrowing weak solutions u_i and $u_{i,r,M}$ to the equations

$$\frac{d}{dt} u_i = \nabla \cdot (a\nabla u_i),$$
$$u_i(\cdot, 0) = \nabla \cdot (ae_i),$$

and

$$\frac{d}{dt} u_{i,r,M} = \nabla \cdot (a_{r,M} \nabla u_{i,r,M}),$$
$$u_{i,r,M}(\cdot, 0) = \nabla \cdot (a_{r,M} e_i).$$

Then we have

$$\fint_{\{|x| \leq 2dr\}} |u_i(\cdot, t) - u_{i,r,M}(\cdot, t)|^2 \, dx \leq \frac{CM^{d/2}}{t} \exp\left(-c\frac{M^2 r^2}{t} \right) \qquad (98)$$

for any $t \leqq 16M^2r^2$ and

$$\int_0^{16r^2} \fint_{\{|x| \leqq dr\}} |\nabla u_i - \nabla u_{i,r,M}|^2 \, dx \, dt \leqq \frac{C}{r^2} \exp(-cM^2). \tag{99}$$

Proof. For an arbitrary function $\psi \in L^2(\mathbb{R}^d)$ supported in $\{|x| \leqq 2dr\}$ and any $T \in [0, 16M^2r^2]$, consider the solutions v_ψ and $v_{\psi,r,M}$ to the dual equations

$$-\frac{d}{dt}v_\psi = \nabla \cdot (a^* \nabla v_\psi),$$

$$v_\psi(\cdot, T) = \psi.$$

and

$$-\frac{d}{dt}v_{\psi,r,M} = \nabla \cdot (a^*_{r,M} \nabla v_{\psi,r,M}),$$

$$v_{\psi,r,M}(\cdot, T) = \psi.$$

We then have

$$\int_{\mathbb{R}^d} (u_i - u_{i,r,M})(\cdot, T)\psi \, dx = \int_{\mathbb{R}^d} u_i(\cdot, 0)v_\psi(\cdot, 0) - u_{i,r,M}(\cdot, 0)v_{\psi,r,M}(\cdot, 0) \, dx$$

$$+ \int_0^T \frac{d}{dt} \int_{\mathbb{R}^d} u_i v_\psi - u_{i,r,M} v_{\psi,r,M} \, dx \, dt$$

$$= -\int_{\mathbb{R}^d} a e_i \cdot \nabla v_\psi(\cdot, 0) - a_{r,M} e_i \cdot \nabla v_{\psi,r,M}(\cdot, 0) \, dx + 0$$

$$\leqq C \int_{\{|x| \geqq \frac{M}{2}r\}} |\nabla v_\psi(\cdot, 0)| + |\nabla v_{\psi,r,M}(\cdot, 0)| \, dx$$

$$+ C \int_{\{|x| \leqq \frac{M}{2}r\}} |\nabla (v_\psi(\cdot, 0) - v_{\psi,r,M}(\cdot, 0))| \, dx$$

$$\leqq C \left(\int_{\{|x| \geqq \frac{M}{2}r\}} \left(\frac{|x|}{Mr} \right)^{2d} \left(|\nabla v_\psi(\cdot, 0)|^2 + |\nabla v_{\psi,r,M}(\cdot, 0)|^2 \right) dx \right)^{1/2} (Mr)^{d/2}$$

$$+ C \left(\int_{\{|x| \leqq \frac{M}{2}r\}} |\nabla (v_\psi(\cdot, 0) - v_{\psi,r,M}(\cdot, 0))|^2 \, dx \right)^{1/2} (Mr)^{d/2}.$$

The penultimate term may be estimated by Lemma 17 (applied to the backward-in-time equations for v_ψ and $v_{\psi,r,M}$ and breaking up the "initial" condition ψ into pieces supported on scale $\sqrt{T}$ if necessary), resulting in the bound (note that $2dr \leqq \frac{Mr}{4}$)

$$\left| \int_{\mathbb{R}^d} (u_i - u_{i,r,M})(\cdot, T)\psi \, dx \right|$$

$$\leqq C \sum_{x_0 \in \frac{1}{d}\sqrt{T}\mathbb{Z}^d \cap \{|x| \leqq 2dr\}} \left(\int_{\{|x| \geqq \frac{M}{2}r\}} \left(\frac{|x|}{Mr} \right)^{2d} \cdot \sqrt{T}^d \exp\left(-\frac{|x - x_0|^2}{CT} \right) \right)$$

$$\times C(d, \lambda) \fint_{\{|x-x_0| \leqq \sqrt{T}\}} |\psi(\tilde{x})|^2 \, d\tilde{x} \cdot T^{-1} \, dx \right)^{1/2} (Mr)^{d/2}$$

$$+ C \left(\int_{\{|x| \leqq \frac{M}{2} r\}} |\nabla(v_\psi(\cdot, 0) - v_{\psi, r, M}(\cdot, 0))|^2 \, dx \right)^{1/2} (Mr)^{d/2},$$

and therefore by $\sqrt{T} \leqq 4Mr$,

$$\left| \int_{\mathbb{R}^d} (u_i - u_{i,r,M})(\cdot, T) \psi \, dx \right|$$

$$\leqq C \left(\frac{4Mr}{\sqrt{T}} \right)^d \times \exp\left(-\frac{M^2 r^2}{CT} \right) T^{-1/2} \left(\int |\psi|^2 \, dx \right)^{1/2} \times (Mr)^{d/2}$$

$$+ C \left(\int_{\{|x| \leqq \frac{M}{2} r\}} |\nabla(v_\psi(\cdot, 0) - v_{\psi, r, M}(\cdot, 0))|^2 \, dx \right)^{1/2} (Mr)^{d/2}. \qquad (100)$$

An estimate for the last term on the right-hand side of (100) can be obtained as follows: observe that

$$-\frac{d}{dt}(v_\psi - v_{\psi, r, M}) = \nabla \cdot (a^* \nabla (v_\psi - v_{\psi, r, M})) + \nabla \cdot ((a^* - a^*_{r,M}) \nabla v_{\psi, r, M}),$$

$$(v_\psi - v_{\psi, r, M})(\cdot, T) = 0.$$

We rewrite $(v_\psi - v_{\psi, r, M})(\cdot, 0)$ as $(v_\psi - v_{\psi, r, M})(\cdot, 0) = \int_0^T w_t(\cdot, 0) \, dt$ with w_{t_0} being the solution to the equation

$$-\frac{d}{dt} w_{t_0} = \nabla \cdot (a^* \nabla w_{t_0}),$$

$$w_{t_0}(\cdot, t_0) = \nabla \cdot ((a^* - a^*_{r,M}) \nabla v_{\psi, r, M}(\cdot, t_0)).$$

Considering the estimate (103) centered at x_0 (instead of 0) and integrating over the set $\{|x_0| \leqq \frac{M}{2} r\}$ and applying it to the backward-in-time equation for w_{t_0}, we obtain (also using the condition $t_0 \leqq T \leqq CM^2 r^2$)

$$\int_{\{|x| \leqq \frac{M}{2} r\}} |\nabla w_{t_0}(\cdot, 0)|^2 \, dx$$

$$\leqq C(d, \lambda) t_0^{-2} \int_{\{|x_0| \leqq \frac{M}{2} r\}} \int |(a^* - a^*_{r,M}) \nabla v_{\psi, r, M}|^2(x) \cdot t_0^{-d/2} \exp\left(-\frac{|x - x_0|^2}{C t_0} \right) dx \, dx_0$$

$$\leqq C(d, \lambda) t_0^{-2} \int_{\{|x_0| \leqq \frac{M}{2} r\}} \int_{\{|x| \geqq Mr\}} |\nabla v_{\psi, r, M}(x)|^2 \cdot t_0^{-d/2} \exp\left(-\frac{|x - x_0|^2}{C t_0} \right) dx \, dx_0$$

$$\leqq C(d, \lambda) t_0^{-2} \exp\left(-\frac{M^2 r^2}{2C t_0} \right) \int |\nabla v_{\psi, r, M}(x)|^2 \cdot \exp\left(-\frac{|x|^2}{4C t_0} \right) dx.$$

Lemma 17 (applied to $v_{\psi,r,M}$) implies by breaking up the "initial" condition ψ into contributions supported on balls of size $\sqrt{T - t_0}$

$$\int_{\mathbb{R}^d} |\nabla v_{\psi,r,M}(x)|^2 \, dx \leqq \frac{C}{T - t_0} \int_{\mathbb{R}^d} |\psi|^2 \, dx.$$

Combining the previous two estimates, we deduce

$$\int_{\{|x| \leq \frac{M}{2} r\}} |\nabla w_{t_0}(\cdot, 0)|^2 \, dx$$

$$\leqq C(d, \lambda) t_0^{-2} \exp\left(-\frac{M^2 r^2}{2C t_0} \right) (T - t_0)^{-1} \int |\psi(x)|^2 \, dx.$$

Taking the square root and integrating with respect to t_0, this entails

$$\left(\int_{\{|x| \leq \frac{M}{2} r\}} |\nabla(v_\psi - v_{\psi,r,M})|^2 \, dx \right)^{1/2}$$

$$\leqq C(d, \lambda) \int_0^T t_0^{-1} (T - t_0)^{-1/2} \exp\left(-\frac{M^2 r^2}{C t_0} \right) dt_0 \cdot \left(\int |\psi(x)|^2 \, dx \right)^{1/2}$$

$$\leqq C(d, \lambda) \left(\frac{T^{1/2}}{M^2 r^2} + \frac{1}{T^{1/2}} \right) \exp\left(-\frac{M^2 r^2}{CT} \right) \left(\int |\psi(x)|^2 \, dx \right)^{1/2}.$$

Using $T \leqq CM^2 r^2$ and plugging in this bound into (100), we get, by $M \geq 5d$,

$$\left| \int_{\mathbb{R}^d} (u_i - u_{i,r,M})(\cdot, T) \psi \, dx \right|$$

$$\leqq C(d, \lambda) T^{-1/2} \exp\left(-\frac{M^2 r^2}{CT} \right) \cdot \left(\int |\psi|^2 \, dx \right)^{1/2} (Mr)^{d/2}.$$

Passing to the supremum over all ψ supported in $\{|x| \leq 2dr\}$ with $\int |\psi|^2 \, dx \leq 1$, we deduce our bound (98).

Now choose a cutoff η with $\eta \equiv 1$ in $\{|x| \leq dr\}$ and $\eta \equiv 0$ outside of $\{|x| \leq 2dr\}$. For any $t \leq 16r^2$, we obtain by testing the equation for the difference $u_i - u_{i,r,M}$ with $(u_i - u_{i,r,M})\eta^2$

$$\int_t^{2t} \int \eta^2 \lambda |\nabla(u_i - u_{i,r,M})|^2 \, dx \, d\tilde{t}$$

$$\leqq \int_{\{|x| \leq 2dr\}} \eta^2 |u_i(\cdot, t) - u_{i,r,M}(\cdot, t)|^2 \, dx$$

$$+ C \int_t^{2t} \int |\nabla \eta|^2 |u_i(\cdot, t) - u_{i,r,M}(\cdot, t)|^2 \, dx \, d\tilde{t}$$

$$\leqq \int_{\{|x| \leq 2dr\}} |u_i(\cdot, t) - u_{i,r,M}(\cdot, t)|^2 \, dx$$

$$+ C \int_t^{2t} \int_{\{|x| \leq 2dr\}} \frac{C}{r^2} |u_i(\cdot, t) - u_{i,r,M}(\cdot, t)|^2 \, dx \, d\tilde{t}.$$

Using our bound (98) and $t \leq 16r^2$, we get

$$\int_t^{2t} \fint_{\{|x| \leq dr\}} |\nabla(u_i - u_{i,r,M})|^2 \, dx \, d\tilde{t} \leq C(d, \lambda) \frac{CM^{d/2}}{t} \exp\left(-c\frac{M^2 r^2}{t}\right).$$

Taking the sum over all $t = 2^k$ for $2^k \leq T$, we deduce our desired estimate (99). □

Lemma 16. *Let $a \in L^\infty(\mathbb{R}^d; \mathbb{R}^{d \times d})$ be a uniformly elliptic and bounded coefficient field in the sense of (A1). Let $b \in L^\infty(\mathbb{R}^d; \mathbb{R}^d)$ be a bounded vector field. Then the unique nongrowing weak solution w to the equation*

$$\frac{d}{dt} w = \nabla \cdot (a \nabla w),$$
$$w(\cdot, 0) = \nabla \cdot b,$$

satisfies for any $T > 0$ the estimate

$$\left(\fint_{\{|x| \leq \sqrt{T}\}} |w(\cdot, T)|^2 \, dx\right)^{1/2} \tag{101}$$

$$\leq C(d, \lambda) T^{-1/2} \left(\int |b(x)|^2 \cdot T^{-d/2} \exp\left(-\frac{|x|^2}{CT}\right) dx\right)^{1/2}.$$

Furthermore, we have the bounds

$$\fint_{\{|x| \leq 1\}} \left|\nabla \int_0^1 w(\cdot, t) \, dt\right|^2 dx \leq C(d, \lambda)\|b\|_{L^\infty}^2 \tag{102}$$

and

$$\left(\fint_{\{|x| \leq \sqrt{T}\}} |\nabla w(\cdot, T)|^2 \, dx\right)^{1/2} \tag{103}$$

$$\leq C(d, \lambda) T^{-1} \left(\int |b(x)|^2 \cdot T^{-d/2} \exp\left(-\frac{|x|^2}{CT}\right) dx\right)^{1/2}.$$

Proof. Let $T > 0$ and let $g \in L^2(\mathbb{R}^d)$ be a function supported in $\{|x| \leq \sqrt{T}\}$. Introducing the solution v to the dual (backward-in-time) equation

$$-\frac{d}{dt} v = \nabla \cdot (a^* \nabla v),$$
$$v(\cdot, T) = g,$$

we see that we have

$$\int w(\cdot, T) g \, dx$$

$$= -\int b \cdot \nabla v(\cdot, 0) \, dx + \int_0^T \frac{d}{dt} \int wv \, dx \, dt$$

$$= -\int b \cdot \nabla v(\cdot, 0) \, dx + \int_0^T \int -a\nabla w \cdot \nabla v + \nabla w \cdot a^* \nabla v \, dx \, dt$$

$$= -\int b \cdot \nabla v(\cdot, 0) \, dx.$$

Introducing

$$\Theta_T(x) := \exp\left(\frac{|x|^2}{CT}\right),$$

we obtain

$$\int w(\cdot, T)g\,dx \leqq \left(\int |b(x)|^2 \frac{1}{\Theta_T(x)}\,dx\right)^{1/2} \left(\int |\nabla v(\cdot, 0)|^2 \Theta_T(\cdot)\,dx\right)^{1/2}.$$

Lemma 17 (applied to v, which solves a parabolic PDE backward in time) provides the estimate

$$\int |\nabla v(\cdot, 0)|^2 \Theta_T(\cdot)\,dx \leqq C(d, \lambda)T^{-1} \int |g|^2\,dx.$$

Inserting this estimate in the previous inequality and passing to the supremum over all $g \in L^2$ supported in $\{|x| \leqq \sqrt{T}\}$ with $\int_{\{|x|\leqq\sqrt{T}\}} |g|^2\,dx \leqq 1$, we get

$$\left(\int_{\{|x|\leqq\sqrt{T}\}} |w(\cdot, T)|^2\,dx\right)^{1/2} \leqq CT^{-1/2}\left(\int |b(x)|^2 \frac{1}{\Theta_T(x)}\,dx\right)^{1/2}.$$

This establishes the estimate (101).

To prove the estimate (102), we first observe that we have

$$\nabla \cdot \left(a\nabla \int_0^1 w(\cdot, t)\,dt + b\right) = w(\cdot, 1).$$

Testing this PDE with $\eta^2 \int_0^1 w(\cdot, t)\,dt$ where η is a standard cutoff with $\eta \equiv 1$ in $\{|x| \leqq 1\}$ and $\eta \equiv 0$ outside of $\{|x| \leqq 2\}$, we obtain

$$\int \eta^2 \left|\nabla \int_0^1 w(\cdot, t)\,dt\right|^2\,dx$$

$$\leqq \int C|\eta|^2 |b|^2 + C(\eta^2 + |\nabla\eta|^2)\left|\int_0^1 w(\cdot, t)\,dt\right|^2 + C|\eta|^2 |w(\cdot, 1)|^2\,dx$$

$$\leqq \int C|\eta|^2 |b|^2 + C(\eta^2 + |\nabla\eta|^2)\int_0^1 \sqrt{t}|w(\cdot, t)|^2\,dt + C|\eta|^2 |w(\cdot, 1)|^2\,dx.$$

The estimate (101) entails

$$\left(\fint_{\{|x|\leqq\sqrt{t}\}} |w(\cdot, T)|^2\,dx\right)^{1/2} \leqq Ct^{-1/2}||b||_{L^\infty}.$$

The previous two estimates yield (102).

Finally, to prove (103), we first deduce from (101)

$$\left(\fint_{\{|x-x_0|\leqq\sqrt{T/2}\}} |w(\cdot, T/2)|^2\,dx\right)^{1/2}$$

$$\leqq C(d, \lambda)T^{-1/2}\left(\int |b(x)|^2 \cdot T^{-d/2}\exp\left(-\frac{|x-x_0|^2}{CT}\right)dx\right)^{1/2}.$$

Splitting the function $w(\cdot, T/2)$ into pieces each supported on a ball of size $\sqrt{T}/2$—that is, splitting $w(\cdot, T/2) = \sum_l \eta_l w(\cdot, T/2)$ with a partition of unity η_l subordinate to the set of balls $\{|x - x_0| \leqq \sqrt{T/2}\}$, $x_0 \in \frac{1}{d}\sqrt{T/2}\mathbb{Z}^d$—and applying Lemma 17 to the solutions of the parabolic equation with initial data $\eta_l w(\cdot, T/2)$ for all l (note that w is equal to the sum of all of these solutions), we obtain

$$\left(\fint_{\{|x| \leqq \sqrt{T/2}\}} |\nabla w(\cdot, T)|^2 \, dx \right)^{1/2}$$

$$\leqq \sum_{x_0 \in \frac{1}{d}\sqrt{T/2}\mathbb{Z}^d} \exp\left(-\frac{|x_0|^2}{CT/2} \right) \left(\fint_{\{|x - x_0| \leqq \sqrt{T/2}\}} |w(\cdot, T/2)|^2 \, dx \right)^{1/2} T^{-1/2}$$

$$\leqq \sum_{x_0 \in \frac{1}{d}\sqrt{T/2}\mathbb{Z}^d} T^{-d/2} \exp\left(-\frac{|x_0|^2}{CT/2} \right)$$

$$\times C(d, \lambda) T^{-1} \left(\int |b(x)|^2 \cdot T^{-d/2} \exp\left(-\frac{|x - x_0|^2}{CT} \right) dx \right)^{1/2}.$$

A straightforward estimate then entails (103) (with a different constant C). $\quad\square$

Lemma 17. *Let $a \in L^\infty(\mathbb{R}^d; \mathbb{R}^{d \times d})$ be a uniformly elliptic and bounded coefficient field in the sense of (A1) and let $T > 0$. Let $g \in L^2(\mathbb{R}^d)$ be a function supported in $\{|x| \leqq \sqrt{T}\}$. Then there exists $C = C(d, \lambda) > 0$ such that the unique nongrowing weak solution w to the equation*

$$\frac{d}{dt}w = \nabla \cdot (a\nabla w),$$

$$w(\cdot, 0) = g,$$

satisfies the estimate

$$\left(\int_{\mathbb{R}^d} |\nabla w(\cdot, T)|^2 \frac{1}{\sqrt{T}^d} \exp\left(\frac{|x|^2}{CT} \right) dx \right)^{1/2}$$

$$\leqq C(d, \lambda) \left(\fint_{\{|x| \leqq \sqrt{T}\}} |g|^2 \, dx \right)^{1/2} T^{-1/2}.$$

Proof. As

$$\Theta_T^m(x, t) := \exp\left(\frac{|x|^2}{4C_m(T + t)} \right)$$

satisfies

$$\frac{d}{dt}\Theta_T^m + C_m \frac{|\nabla \Theta_T^m|^2}{\Theta_T^m} = \left(-\frac{|x|^2}{4C_m(T + t)^2} + C_m \left| \frac{2x}{4C_m(T + t)} \right|^2 \right) \Theta_T^m \leqq 0,$$

$$\tag{104}$$

we have for $C_1 \geqq C(d, \lambda)$

$$
\frac{\mathrm{d}}{\mathrm{d}t} \int |w|^2 \Theta_T^1 \, \mathrm{d}x
$$

$$
= \int |w|^2 \frac{\mathrm{d}}{\mathrm{d}t} \Theta_T^1 + 2\Theta_T^1 a \nabla w \cdot \nabla w + 2a\nabla w \cdot w \nabla \Theta_T^1 \, \mathrm{d}x
$$

$$
\leqq -\lambda \int \Theta_T^1 |\nabla w|^2 \, \mathrm{d}x.
$$

This provides the bound

$$
\int_0^\infty \int |\nabla w|^2 \Theta_T^1 \, \mathrm{d}x \, \mathrm{d}t \leqq C \int |g|^2 \, \mathrm{d}x. \tag{105}
$$

We now would like to show (basically) $\nabla w \in C^\gamma ([\frac{1}{2}T, \frac{3}{2}T]; L^2_{\Theta_T^m})$ for some $\gamma > 0$ and some m. To this end, we abbreviate $\Theta_{T,t}^m := \Theta_T^m(\cdot, t)$ and compute

$$
\int |w(\cdot, t+h) - w(\cdot, t)|^2 \Theta_{T,t}^m \, \mathrm{d}x
$$

$$
= \int \int_t^{t+h} \frac{\mathrm{d}}{\mathrm{d}s} w(\cdot, s) \, \mathrm{d}s \, (w(\cdot, t+h) - w(\cdot, t)) \Theta_{T,t}^m \, \mathrm{d}x
$$

$$
= -\int \Theta_{T,t}^m \int_t^{t+h} a\nabla w \, \mathrm{d}s \cdot \nabla(w(\cdot, t+h) - w(\cdot, t)) \, \mathrm{d}x
$$

$$
- \int (w(\cdot, t+h) - w(\cdot, t)) \int_t^{t+h} a\nabla w \, \mathrm{d}s \cdot \nabla \Theta_{T,t}^m \, \mathrm{d}x.
$$

Applying the Hölder inequality to the first term and Young's inequality (and absorption) to the second term, we get

$$
\int |w(\cdot, t+h) - w(\cdot, t)|^2 \Theta_{T,t}^m \, \mathrm{d}x \tag{106}
$$

$$
\leqq C \left(\int \Theta_{T,t}^m \left| \int_t^{t+h} a\nabla w \, \mathrm{d}s \right|^2 \mathrm{d}x \right)^{1/2}
$$

$$
\times \left(\int \Theta_{T,t}^m |\nabla w(\cdot, t+h) - \nabla w(\cdot, t)|^2 \, \mathrm{d}x \right)^{1/2}
$$

$$
+ C \int \left| \int_t^{t+h} a\nabla w \, \mathrm{d}s \right|^2 \frac{|\nabla \Theta_{T,t}^m|^2}{\Theta_{T,t}^m} \, \mathrm{d}x
$$

$$
\leqq C\sqrt{h} \left(\int \int_t^{t+h} \Theta_{T,t}^m |\nabla w|^2 \, \mathrm{d}s \, \mathrm{d}x \right)^{1/2}
$$

$$
\times \left(\int \Theta_{T,t}^m |\nabla w(\cdot, t+h) - \nabla w(\cdot, t)|^2 \, \mathrm{d}x \right)^{1/2}
$$

$$
+ Ch \int \int_t^{t+h} |\nabla w|^2 \, \mathrm{d}s \frac{|\nabla \Theta_{T,t}^m|^2}{\Theta_{T,t}^m} \, \mathrm{d}x.
$$

Choosing a weight Θ_T^2 with slower growth than in (105)—for example, setting $C_2 := 4C_1$—, we may ensure that $\frac{|\nabla\Theta_T^2(\cdot,t)|^2}{\Theta_T^2(\cdot,t)} \leq \frac{C}{T}\Theta_T^1(\cdot,\tilde{t})$ and $\Theta_T^2(\cdot,t) \leq \Theta_T^1(\cdot,\tilde{t})$ for any $t, \tilde{t} \in [0, \frac{T}{3}]$. As a consequence, we may find for any $h \leq \frac{T}{10}$ a suitable $t \in [0, \frac{T}{10}]$ with

$$\int_t^{t+h} \int |\nabla w|^2 \left(\Theta_{T,t}^2 + T\frac{|\nabla\Theta_{T,t}^2|^2}{\Theta_{T,t}^2} \right) dx\, dt \leq C\frac{h}{T} \int_0^{\frac{T}{5}} \int |\nabla w|^2 \Theta_T^1 dx\, dt,$$

$$\int \Theta_{T,t}^2 \left(|\nabla w(\cdot, t+h)|^2 + |\nabla w(\cdot, t)|^2 \right) dx \leq \frac{C}{T} \int_0^{\frac{T}{5}} \int |\nabla w|^2 \Theta_T^1 dx\, dt.$$

Plugging in these bounds in the previous estimate and using (105), we obtain, for this t,

$$\int |w(\cdot, t+h) - w(\cdot, t)|^2 \Theta_T^2(\cdot, t)\, dx \leq C\left(\frac{h}{T} + \frac{h^2}{T^2} \right) \int |g|^2\, dx. \qquad (107)$$

Abbreviating $\Delta_h w(\cdot, t) := w(\cdot, t+h) - w(\cdot, t)$, we compute, for $C_m \geq C(d, \lambda)$,

$$\frac{d}{dt} \int |\Delta_h w|^2 \Theta_T^m\, dx \qquad (108)$$

$$= \int 2\Theta_T^m \Delta_h w \frac{d}{dt}\Delta_h w + |\Delta_h w|^2 \frac{d}{dt}\Theta_T^m\, dx$$

$$= \int -2\Theta_T^m a\nabla\Delta_h w \cdot \nabla\Delta_h w - 2\Delta_h w\, a\nabla\Delta_h w \cdot \nabla\Theta_T^m + |\Delta_h w|^2 \frac{d}{dt}\Theta_T^m\, dx$$

$$\overset{(104)}{\leq} -\lambda \int \Theta_T^m |\Delta_h\nabla w|^2\, dx.$$

Combining this with the existence of $t \in [0, \frac{T}{10}]$ for which the bound (107) holds, this entails, for any $h \leq \frac{T}{10}$,

$$\int_{\frac{T}{5}}^\infty \int \Theta_T^2 |\Delta_h\nabla w|^2\, dx\, dt \leq C\frac{h}{T} \int |g|^2\, dx.$$

We intend to plug back this estimate into (106). First, for any $h \in [0, \frac{T}{10}]$ we infer the existence of $t \in [\frac{T}{5}, \frac{T}{3}]$ which in addition to the bound

$$\int_t^{t+h} \int |\nabla w|^2 \left(\Theta_{T,t}^2 + T\frac{|\nabla\Theta_{T,t}^2|^2}{\Theta_{T,t}^2} \right) dx\, dt \leq C\frac{h}{T} \int_0^{\frac{T}{3}} \int |\nabla w|^2 \Theta_T^1 dx\, dt,$$

satisfies

$$\int \Theta_{T,t}^2 |\Delta_h\nabla w(\cdot, t)|^2\, dx \leq \frac{C}{T} \int_{\frac{T}{5}}^{\frac{T}{3}} \int \Theta_T^2 |\Delta_h\nabla w|^2\, dx\, dt.$$

Plugging these three estimates and (105) back into (106), we obtain for some $t \in [\frac{T}{5}, \frac{T}{3}]$ the improved bound

$$\int |w(\cdot, t+h) - w(\cdot, t)|^2 \Theta_{T,t}^2 \, dx \leqq C \left(\frac{h^{3/2}}{T^{3/2}} + \frac{h^2}{T^2} \right) \int |g|^2 \, dx. \qquad (109)$$

By (108) we obtain, for any $h \in [0, \frac{T}{10}]$,

$$\int_{\frac{T}{3}}^{\infty} \int \Theta_T^2 |\Delta_h \nabla w|^2 \, dx \, dt \leqq C \frac{h^{3/2}}{T^{3/2}} \int |g|^2 \, dx.$$

In other words, ∇w belongs to the Nikolskii space on the time interval $[\frac{T}{3}, 2T]$ with order of differentiability $\frac{3}{4}$, integrability 2, and values in $L^2_{\Theta_{T,2T}^2}(\mathbb{R}^d)$; furthermore, the Nikolskii seminorm is subject to a bound of the order $\frac{C}{T^{3/2}} \int |g|^2 \, dx$. By the embedding theorem for Nikolskii spaces, we deduce

$$\sup_{t \in [T/3, 2T]} \int \Theta_T^2(\cdot, 2T) |\nabla w(\cdot, t)|^2 \, dx$$

$$\leqq C \fint_{T/3}^{2T} \int \Theta_T^2(\cdot, 2T) |\nabla w|^2 \, dx \, dt$$

$$+ C T^{3/2} \sup_{h \in [0,T]} \fint_{T/3}^{2T-h} \int \Theta_T^2(\cdot, 2T) |h^{-3/4} \Delta_h \nabla w|^2 \, dx \, dt$$

$$\overset{(105)}{\leqq} \frac{C}{T} \int |g|^2 \, dx.$$

This establishes our lemma. $\square$

Appendix B. Calculus for Random Variables with Stretched Exponential Moments

On the space of random variables X with stretched exponential moments in the sense that

$$\mathbb{E}\left[\exp\left(\frac{|X|^\gamma}{C} \right) \right] \leqq 2,$$

for some $\gamma > 0$ and some $C > 0$, it is convenient to work with the norm

$$||X||_{\exp^\gamma} := \sup_{p \geqq 1} \frac{1}{p^{1/\gamma}} \mathbb{E}[|X|^p]^{1/p}.$$

For $\gamma \geqq 1$, this norm is equivalent to the Luxemburg norm associated with the convex function $\exp(x^\gamma) - 1$. However, this norm has two advantages: first, it simplifies calculus when considering the integrability of products of random variables or the concentration properties of independent random variables; secondly and more importantly, it is also a well-defined norm for $\gamma \in (0, 1)$, a parameter range which we shall employ heavily.

Lemma 18. *Let $\gamma > 0$. Consider a random variable X on some probability space. Define the quasinorm*

$$||X||_{\exp^\gamma, \text{quasi}} := \inf\left\{s > 0 : \mathbb{E}\left[\exp\left(\frac{|X|^\gamma}{s^\gamma}\right)\right] \leq 2\right\}.$$

Then we have $||X||_{\exp^\gamma, \text{quasi}} < \infty$ if and only if $||X||_{\exp^\gamma} < \infty$ and there exist constants $c(\gamma)$, $C(\gamma)$ such that the estimate

$$c(\gamma)||X||_{\exp^\gamma} \leq ||X||_{\exp^\gamma, \text{quasi}} \leq C(\gamma)||X||_{\exp^\gamma}$$

is satisfied.

Proof. The function

$$f_q(x) := x^q \exp(-x)$$

satisfies $f_q'(x) = (q - x)x^{q-1}\exp(-x)$ and attains its maximal value

$$\sup_{x \geq 0} f_q(x) = q^q \exp(-q)$$

at $x = q$. Applying the resulting estimate $x^q \leq q^q \exp(x)$ to $x := |X|^\gamma/||X||^\gamma_{\exp^\gamma, \text{quasi}}$ we deduce

$$\mathbb{E}\big[|X|^{\gamma q}\big] \leq q^q ||X||^{\gamma q}_{\exp^\gamma, \text{quasi}} \mathbb{E}[\exp(|X|^\gamma/||X||^\gamma_{\exp^\gamma, \text{quasi}})].$$

By definition we have $\mathbb{E}[\exp(|X|^\gamma/||X||^\gamma_{\exp^\gamma, \text{quasi}})] \leq 2$. Setting $p := \gamma q$ and taking the p-th root, we obtain, for any $p \geq 1$,

$$\mathbb{E}\big[|X|^p\big]^{1/p} \leq C(\gamma)p^{1/\gamma}||X||_{\exp^\gamma, \text{quasi}}.$$

This proves $||X||_{\exp^\gamma} \leq C(\gamma)||X||_{\exp^\gamma, \text{quasi}}$.

To establish the reverse inequality, observe that for $z \in [q, q + 1)$ we have $z^q e^{-z/2} \geq q^q e^{-(q+1)/2}$. This entails, for all $z \geq 0$,

$$\exp(z/2) \leq \sqrt{e} + \sum_{q=1}^{\infty} \sqrt{e}\frac{z^q}{(q/\sqrt{e})^q},$$

and therefore for any $b > 0$ (by setting $z = |X|^\gamma/b^\gamma$),

$$\mathbb{E}\big[\exp(|X|^\gamma/2b^\gamma)\big] \leq \sqrt{e} + \sum_{q=1}^{\infty} \sqrt{e}\frac{\mathbb{E}[|X|^{\gamma q}]}{b^{\gamma q}(q/\sqrt{e})^q}.$$

As a consequence, we obtain

$$\mathbb{E}\big[\exp(|X|^\gamma/2b^\gamma)\big] \leq \sqrt{e} + \sum_{q=1}^{\infty} \frac{C(\gamma)(\gamma q)^{q\gamma/\gamma}||X||^{\gamma q}_{\exp^\gamma}}{b^{\gamma q}(q/\sqrt{e})^q}.$$

Setting $b := a||X||_{\exp^\gamma}$, we get

$$\mathbb{E}\big[\exp(|X|^\gamma/2b^\gamma)\big] \leqq \sqrt{e} + \sum_{q=1}^{\infty} \frac{C(\gamma)\gamma^q \sqrt{e^q}}{a^{\gamma q}} \leqq \sqrt{e} + \frac{C(\gamma)a^{-\gamma}}{1 - \gamma\sqrt{e}/a^\gamma}.$$

Choosing $a := C(\gamma)$ large enough, we deduce

$$\mathbb{E}\left[\exp\left(\frac{|X|^\gamma}{C(\gamma)||X||_{\exp^\gamma}^\gamma}\right)\right] \leqq 2,$$

which entails $||X||_{\exp^\gamma,\text{quasi}} \leqq C(\gamma)||X||_{\exp^\gamma}$. $\quad\square$

Lemma 19. (Calculus for random variables with stretched exponential moments)
Let X, Y be random variables with stretched exponential moments in the sense that $||X||_{\exp^\gamma} < \infty$ and $||Y||_{\exp^\beta} < \infty$ for some $\gamma, \beta > 0$. Then:

(a) *The product XY has stretched exponential moments with exponent α given by $\frac{1}{\alpha} = \frac{1}{\gamma} + \frac{1}{\beta}$ and satisfies the bound*

$$||XY||_{\exp^\alpha} \leqq C(\beta,\gamma)||X||_{\exp^\gamma}||Y||_{\exp^\beta}.$$

(b) *There exists constants $c = c(\gamma) > 0$, $C = C(\gamma) < \infty$, with the following property: for any $K \geqq 0$, we have the estimate*

$$\mathbb{P}\big[|X| \geqq K||X||_{\exp^\gamma}\big] \leqq C \exp(-cK^\gamma).$$

Proof. For the first assertion, we estimate, for any $p \geqq 1$, by Hölder's inequality, that

$$\frac{1}{p^{1/\alpha}}\mathbb{E}\big[|XY|^p\big]^{1/p} = \frac{1}{p^{1/\gamma+1/\beta}}\mathbb{E}\big[|XY|^p\big]^{1/p}$$

$$\leqq \frac{1}{p^{1/\gamma+1/\beta}}\mathbb{E}\big[|X|^{2p}\big]^{1/2p}\mathbb{E}\big[|Y|^{2p}\big]^{1/2p}$$

$$\leqq 2^{1/\gamma+1/\beta}||X||_{\exp^\gamma}||Y||_{\exp^\beta}.$$

This establishes the first assertion.

For the second assertion, we estimate, for any $p \geqq 1$ and any $K > 0$,

$$\mathbb{P}\big[|X| \geqq K||X||_{\exp^\gamma,\text{quasi}}\big] \leqq \frac{\mathbb{E}\left[\exp\left(\left(\frac{|X|}{||X||_{\exp^\gamma,\text{quasi}}}\right)^\gamma\right)\right]}{\exp(K^\gamma)} \leqq 2\exp(-K^\gamma).$$

Using the fact that $||X||_{\exp^\gamma,\text{quasi}} \leqq C(\gamma)||X||_{\exp^\gamma}$, the second assertion follows upon redefining K. $\quad\square$

For independent random variables with stretched exponential moments, a standard argument via an inequality by Burkholder [30] provides a simple concentration estimate. Note that the estimate is not sharp and may be improved, see for example [26].

Lemma 20. *Let $X_1, \ldots, X_M$ be independent random variables with vanishing expectation and uniformly bounded stretched exponential moments*

$$||X_m||_{\exp^{\gamma_0}} \leqq b$$

for some $\gamma_0 > 0$ and some $b > 0$. Then the sum

$$X := \sum_{m=1}^{M} X_m$$

has uniformly bounded stretched exponential moments

$$||X||_{\exp^{\tilde{\gamma}}} \leqq C(\gamma_0)\sqrt{M}b$$

for $\tilde{\gamma} := \gamma_0/(\gamma_0 + 1)$.

Proof. The discrete-time stochastic process

$$m \mapsto \sum_{\tilde{m}=1}^{m} X_{\tilde{m}}$$

is a square-integrable martingale. An estimate by BURKHOLDER [30, Theorem 3.2]—applied for "timestep" $m := M$—yields for any $k \in \mathbb{N}$

$$\mathbb{E}\big[|X|^{2k}\big]^{1/2k} \leqq C \cdot 2k\mathbb{E}\bigg[\bigg|\sum_{m=1}^{M} |X_m|^2\bigg|^k\bigg]^{1/2k}.$$

This entails

$$\mathbb{E}\big[|X|^{2k}\big]^{1/2k} \leqq C \cdot 2k\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\big[|X_m|^{2k}\big]\right)^{1/2k}, \tag{110}$$

and therefore

$$\mathbb{E}\big[|X|^{2k}\big]^{1/2k} \leqq C \cdot 2k\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M}(2k)^{2k/\gamma_0}||X_m||_{\exp^{\gamma_0}}^{2k}\right)^{1/2k}$$

$$\leqq C\sqrt{M}(2k)^{1+1/\gamma_0}b.$$

We infer

$$(2k)^{-1/\gamma_0-1}\mathbb{E}\big[|X|^{2k}\big]^{1/2k} \leqq C\sqrt{M}b$$

for any $k \in \mathbb{N}$, which, by Hölder's inequality, entails

$$p^{-1/\gamma_0-1}\mathbb{E}\big[|X|^{p}\big]^{1/p} \leqq C(\gamma_0)\sqrt{M}b$$

for any $p \geqq 1$. $\quad\square$

References

1. ABDULLE, A.: On a priori error analysis of fully discrete heterogeneous multiscale FEM. *Multiscale Model. Simul.* **4**(2), 447–459, 2005

2. ANDRES, S., NEUKAMM, S.: Berry–Esseen theorem and quantitative homogenization for the random conductance model with degenerate conductances. Preprint arXiv:1706.09493, 2017

3. ARMSTRONG, S., BORDAS, A., MOURRAT, J.-C.: Quantitative stochastic homogenization and regularity theory of parabolic equations. Preprint arXiv:1705.07672, 2017

4. ARMSTRONG, S., DARIO, P.: Elliptic regularity and quantitative homogenization on percolation clusters. to appear in Commun. Pure Appl. Math. arXiv:1609.09431, 2018

5. ARMSTRONG, S., KUUSI, T., MOURRAT, J.-C.: The additive structure of elliptic homogenization. *Invent. Math.* **208**(3), 999–1154, 2017

6. ARMSTRONG, S., KUUSI, T., MOURRAT, J.-C.: Quantitative stochastic homogenization and large-scale regularity. Lecture Notes. Preprint arXiv:1705.05300, 2017

7. ARMSTRONG, S., LIN, J.: Optimal quantitative estimates in stochastic homogenization for elliptic equations in nondivergence form. *Arch. Ration. Mech. Anal.* **225**(2), 937–991, 2017

8. ARMSTRONG, S.N., MOURRAT, J.-C.: Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.* **219**(1), 255–348, 2016

9. ARMSTRONG, S.N., SMART, C.K.: Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)* **49**(2), 423–481, 2016

10. ARMSTRONG, S.N., SOUGANIDIS, P.E.: Stochastic homogenization of Hamilton–Jacobi and degenerate Bellman equations in unbounded environments. *J. Math. Pures Appl. (9)* **97**(5), 460–504, 2012

11. AVELLANEDA, M., LIN, F.: Une théorème de liouville pour des équations elliptiques à coefficients périodiques. *C. R. Acad. Sci. Paris Sér. I Math.* **309**, 245–250, 1989

12. AVELLANEDA, M., LIN, F.-H.: Compactness methods in the theory of homogenization. *Commun. Pure Appl. Math.* **40**(6), 803–847, 1987

13. AYOUL-GUILMARD, Q., NOUY, A., BINETRUY, C.: Tensor-based numerical method for stochastic homogenisation. Preprint arXiv:1805.00902, 2018

14. BABUŠKA, I., CALOZ, G., OSBORN, J.E.: Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.* **31**(4), 945–981, 1994

15. BALZANI, D., BRANDS, D., SCHRÖDER, J.: Construction of statistically similar representative volume elements. In: *Plasticity and Beyond*, Vol. 550 (Eds. Schröder J. and Hackl K.) Springer, Berlin, 355–412, 2014

16. BALZANI, D., BRANDS, D., SCHRÖDER, J., CARSTENSEN, C.: Sensitivity analysis of statistical measures for the reconstruction of microstructures based on the minimization of generalized least-square functionals. *Tech. Mech.* **30**, 297–315, 2010

17. BALZANI, D., SCHEUNEMANN, L., BRANDS, D., SCHRÖDER, J.: Construction of two- and three-dimensional statistically similar RVEs for coupled micro-macro simulations. *Comput. Mech.* **54**, 1269–1284, 2014

18. BALZANI, D., SCHRÖDER, J.: Some basic ideas for the reconstruction of statistically similar microstructures for multiscale simulations. *PAMM* **8**(1), 10533–10534, 2009

19. BELLA, P., FEHRMAN, B., FISCHER, J., OTTO, F.: Stochastic homogenization of linear elliptic equations: higher-order error estimates in weak norms via second-order correctors. *SIAM J. Math. Anal.* **49**(6), 4658–4703, 2017

20. BELLA, P., GIUNTI, A., OTTO, F.: Effective multipoles in random media. Preprint arXiv:1708.07672, 2017

21. BENOIT, A., GLORIA, A.: Long-time homogenization and asymptotic ballistic transport of classical waves. Preprint arXiv:1701.08600, 2017

22. BLANC, X., COSTAOUEC, R., LE BRIS, C., LEGOLL, F.: Variance reduction in stochastic homogenization: the technique of antithetic variables. In: Engquist, B., Runborg, O.,

Tsai, Y.-H. (eds.) Numerical Analysis of Multiscale Computations, Volume 82 of Lect. Notes Comput. Sci. Eng. Springer, Heidelberg, 47–70, 2012

23. Blanc, X., Costaouec, R., Le Bris, C., Legoll, F.: Variance reduction in stochastic homogenization using antithetic variables. *Markov Process. Relat. Fields* **18**(1), 31–66, 2012

24. Blanc, X., Le Bris, C.: Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings. *Netw. Heterog. Media* **5**, 1–29, 2010

25. Blanc, X., Le Bris, C., Legoll, F.: Some variance reduction methods for numerical stochastic homogenization. *Philos. Trans. A* **374**(2066), 20150168, 2016. 15

26. Boucheron, S., Lugosi, G., Massart, P.: Concentration inequalities using the entropy method. *Ann. Probab.* **31**(3), 1583–1614, 2003

27. Braides, A., Cicalese, M., Ruf, M.: Continuum limit and stochastic homogenization of discrete ferromagnetic thin films. *Anal. PDE* **11**(2), 499–553, 2018

28. Brands, D., Balzani, D., Scheunemann, L., Schröder, J., Richter, H., Raabe, D.: Computational modeling of dual-phase steels based on representative three-dimensional microstructures obtained from ebsd data. *Arch. Appl. Mech.* **86**(3), 575–598, 2016

29. Brezzi, F., Franca, L., Hughes, T., Russo, A.: $b = \int g$. *Comput. Methods Appl. Mech. Eng.* **145**, 329–339, 1997

30. Burkholder, D.L.: Distribution function inequalities for martingales. *Ann. Probab.* **1**(1), 19–42, 1973

31. Caffarelli, L.A., Souganidis, P.E.: Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media. *Invent. Math.* **180**(2), 301–360, 2010

32. Cancès, É., Ehrlacher, V., Legoll, F., Stamm, B.: An embedded corrector problem to approximate the homogenized coefficients of an elliptic equation. *C. R. Math.* **353**(9), 801–806, 2015

33. Chen, L.H.Y., Goldstein, L., Shao, Q.-M.: *Normal Approximation by Stein's Method. Probability and Its Applications (New York).* Springer, Heidelberg 2011

34. Chen, L.H.Y., Shao, Q.-M.: Normal approximation under local dependence. *Ann. Probab.* **32**(3A), 1985–2028, 2004

35. Dal Maso, G., Modica, L.: Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.* **368**, 28–42, 1986

36. Dario, P.: Optimal corrector estimates on percolation clusters. Preprint arXiv:1805.00902, 2018

37. Duerinckx, M., Fischer, J., Gloria, A., Otto, F.: The structure of fluctuations in stochastic homogenization: the case of finite range of dependence, 2019 **(in preparation)**

38. Duerinckx, M., Gloria, A.: Weighted second-order Poincaré inequalities: application to RSA models. Preprint arXiv:1711.03158, 2017

39. Duerinckx, M., Gloria, A., Otto, F.: The structure of fluctuations in stochastic homogenization. Preprint arXiv:1602.01717, 2016

40. E, W., Engquist, B.: The heterogeneous multiscale methods. *Commun. Math. Sci.* **1**(1), 87–132, 2003

41. Efendiev, Y., Kronsbein, C., Legoll, F.: Multilevel Monte Carlo approaches for numerical homogenization. *Multiscale Model. Simul.* **13**(4), 1107–1135, 2015

42. Eigel, M., Peterseim, D.: Simulation of composite materials by a network FEM with error control. *Comput. Methods Appl. Math. (online)* **15**(1), 21–37, 2015

43. Fischer, J.: Quantitative normal approximation for sums of random variables with multilevel local dependence. Preprint arXiv:1905.10273, 2018

44. Fischer, J., Otto, F.: A higher-order large-scale regularity theory for random elliptic operators. *Commun. Partial Differ. Equ.* **41**(7), 1108–1148, 2016

45. Fischer, J., Raithel, C.: Liouville principles and a large-scale regularity theory for random elliptic operators on the half-space. *SIAM J. Math. Anal.* **49**(1), 82–114, 2017

46. Giunti, A., Mourrat, J.-C.: Quantitative homogenization of degenerate random environments. *Ann. Inst. Henri Poincaré Probab. Stat.* **54**(1), 22–50, 2018

47. GLORIA, A.: Reduction of the resonance error. part 1: approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.* **21**(08), 1601–1630, 2011
48. GLORIA, A.: Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *ESAIM: M2AN* **46**(1), 1–38, 2012
49. GLORIA, A., NEUKAMM, S., OTTO, F.: An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *ESAIM Math. Model. Numer. Anal.* **48**(2), 325–346, 2014
50. Gloria, A., Neukamm, S., Otto, F.: A regularity theory for random elliptic operators. Preprint arXiv:1409.2678, 2014
51. GLORIA, A., NEUKAMM, S., OTTO, F.: Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199**(2), 455–515, 2015
52. GLORIA, A., NOLEN, J.: A quantitative central limit theorem for the effective conductance on the discrete torus. *Commun. Pure Appl. Math.* **69**(12), 2304–2348, 2016
53. GLORIA, A., OTTO, F.: An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.* **39**(3), 779–856, 2011
54. GLORIA, A., OTTO, F.: An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.* **22**(1), 1–28, 2012
55. GLORIA, A., OTTO, F.: The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. Preprint arXiv:1510.08290, 2015
56. GU, Y.: High order correctors and two-scale expansions in stochastic homogenization. *Probab. Theory Relat. Fields* **169**(3), 1221–1259, 2017
57. GU, Y., MOURRAT, J.-C.: Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.* **14**(1), 452–481, 2016
58. HEIDA, M., SCHWEIZER, B.: Stochastic homogenization of plasticity equations. *ESAIM Control Optim. Calc. Var.* **24**(1), 153–176, 2018
59. HORNUNG, P., PAWELCZYK, M., VELCIĆ, I.: Stochastic homogenization of the bending plate model. *J. Math. Anal. Appl.* **458**(2), 1236–1273, 2018
60. HOU, T.Y., WU, X.-H.: A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.* **134**(1), 169–189, 1997
61. HUGHES, T.J., FEIJÓO, G.R., MAZZEI, L., QUINCY, J.-B.: The variational multiscale method—a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Eng.* **166**(1), 3–24, 1998. **(Advances in Stabilized Methods in Computational Mechanics)**
62. KHOROMSKAIA, V., KHOROMSKIJ, B., OTTO, F.: A numerical primer in 2D stochastic homogenization: CLT scaling in the representative volume element, 2017. Preprint
63. KOZLOV, S.M.: The averaging of random operators. *Mat. Sb. (N.S.)* **109**(151), 188–202, 1979. 327
64. LE BRIS, C., LEGOLL, F., MINVIELLE, W.: Special quasirandom structures: a selection approach for stochastic homogenization. *Monte Carlo Methods Appl.* **22**(1), 25–54, 2016
65. LEGOLL, F., MINVIELLE, W.: A control variate approach based on a defect-type theory for variance reduction in stochastic homogenization. *Multiscale Model. Simul.* **13**(2), 519–550, 2015
66. LIN, J., SMART, C.K.: Algebraic error estimates for the stochastic homogenization of uniformly parabolic equations. *Anal. PDE* **8**(6), 1497–1539, 2015
67. LIONS, P.-L., SOUGANIDIS, P.E.: Correctors for the homogenization of Hamilton–Jacobi equations in the stationary ergodic setting. *Commun. Pure Appl. Math.* **56**(10), 1501–1524, 2003
68. LIU, J.S.: *Monte-Carlo Strategies in Scientific Computing. Springer Series in Statistics*. Springer, New York 2001
69. LU, J., OTTO, F.: Optimal artificial boundary condition for random elliptic media. Preprint arXiv:1803.09593, 2018
70. MÅLQVIST, A., PETERSEIM, D.: Localization of elliptic multiscale problems. *Math. Comput.* **83**, 2583–2603, 2014

71. MATACHE, A.-M., SCHWAB, C.: Two-scale FEM for homogenization problems. *M2AN Math. Model. Numer. Anal*. **36**(4), 537–572, 2002
72. MOURRAT, J.-C.: Efficient methods for the estimation of homogenized coefficients. Preprint arXiv:1609.06674, 2016
73. MOURRAT, J.-C., NOLEN, J.: Scaling limit of the corrector in stochastic homogenization. *Ann. Appl. Probab*. **27**(2), 944–959, 2017
74. MOURRAT, J.-C., OTTO, F.: Correlation structure of the corrector in stochastic homogenization. *Ann. Probab*. **44**(5), 3207–3233, 2016
75. MURAT, F., TARTAR, L.: *H-Convergence. Progress in Nonlinear Differential Equations and Their Applications*, vol. 31. Birkhäuser Boston Inc, Boston 1997
76. NADDAF, A., SPENCER, T.: Estimates on the variance of some homogenization problems, 1998. Unpublished preprint
77. NOLEN, J.: Normal approximation for the net flux through a random conductor. *Stoch. Partial Differ. Equ. Anal. Comput*. **4**(3), 439–476, 2016
78. PAPANICOLAOU, G.C., VARADHAN, S.R.S.: Boundary value problems with rapidly oscillating random coefficients. In: Random Fields, Vol. I, II (Esztergom, 1979), volume 27 of Colloquia Mathematica Societatis János Bolyai, pp. 835–873. North-Holland, Amsterdam, 1981
79. PETERSEIM, D., CARSTENSEN, C.: Finite element network approximation of conductivity in particle composites. *Numer. Math*. **124**(1), 73–97, 2013
80. RINOTT, Y., ROTAR, V.: A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal*. **56**(2), 333–350, 1996
81. SCHRÖDER, J., BALZANI, D., BRANDS, D.: Approximation of random microstructures by periodic statistically similar representative volume elements based on lineal-path functions. *Arch. Appl. Mech*. **81**(7), 975–997, 2011
82. von PEZOLD, J., DICK, A., FRIÁK, M., NEUGEBAUER, J.: Generation and performance of special quasirandom structures for studying the elastic properties of random alloys: application to al-ti. *Phys. Rev. B* **81**, 094203, 2010
83. WEI, S.-H., FERREIRA, L.G., BERNARD, J.E., ZUNGER, A.: Electronic properties of random alloys: special quasirandom structures. *Phys. Rev. B* **42**, 9622–9649, 1990
84. YUE, X., E, W.: The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys*. **222**(2), 556–572, 2007
85. YURINSKI, V.V.: Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh*. **27**, 167–180, 1986. 215
86. ZUNGER, A., WEI, S.-H., FERREIRA, L.G., BERNARD, J.E.: Special quasirandom structures. *Phys. Rev. Lett*. **65**, 353–356, 1990

JULIAN FISCHER
Institute of Science and Technology Austria (IST Austria),
Am Campus 1,
3400 Klosterneuburg
Austria.
e-mail: julian.fischer@ist.ac.at