Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

A Liouville theorem for stationary and ergodic ensembles of parabolic systems

by

Peter Bella, Alberto Chiarini, and Benjamin Fehrman

Preprint no.: 37 2017

A LIOUVILLE THEOREM FOR STATIONARY AND ERGODIC ENSEMBLES OF PARABOLIC SYSTEMS

PETER BELLA, ALBERTO CHIARINI, AND BENJAMIN FEHRMAN

Abstract. A first-order Liouville theorem is obtained for random ensembles of uniformly parabolic systems under the mere qualitative assumptions of stationarity and ergodicity. Furthermore, the paper establishes, almost surely, an intrinsic large-scale $\mathcal{C}^{1,\alpha}$-regularity estimate for caloric functions.

1. Introduction and main results.

This paper considers random ensembles of uniformly parabolic systems

$$u_t = \nabla \cdot a \nabla u, \tag{1}$$

where the law of the coefficient field a is assumed to be *stationary* with respect to space-time translations and *ergodic*. Precisely, for a probability space of coefficient fields $(\Omega, \mathcal{F}, \langle\cdot\rangle)$, where $\langle\cdot\rangle$ is used simultaneously to denote the law and expectation of the ensemble, the stationarity asserts that the coefficients are statistically homogeneous in time and space in the sense that

$$\forall x \in \mathbb{R}^d, \forall t \in \mathbb{R}: \qquad a(\cdot,\cdot) \text{ and } a(\cdot + x, \cdot + t) \text{ have the same law under } \langle\cdot\rangle. \tag{2}$$

The ergodicity asserts that every translationally invariant function of the coefficient field is constant. That is, for every bounded random variable F:

$$\text{if } \forall x \in \mathbb{R}^d, \forall t \in \mathbb{R}, \text{ and for } \langle\cdot\rangle\text{-a.e. } a: F(a) = F(a(\cdot + x, \cdot + t)), \text{ then } F = c \ \langle\cdot\rangle\text{-a.s.} \tag{3}$$

Finally, the ensemble is bounded and uniformly elliptic in the sense that there exists a deterministic $\lambda \in (0, 1]$ such that

$$|a\xi| \le |\xi| \text{ and } \lambda |\xi|^2 \le \xi \cdot a\xi \qquad \forall \xi \in \mathbb{R}^d, \text{ and for } \langle\cdot\rangle\text{-a.e. } a. \tag{4}$$

Assumptions (2) and (3) are the minimal statistical requirements on the ensemble $\langle\cdot\rangle$ which guarantee the qualitative homogenization of equations like (1), see (12). Their role in this paper, and in homogenization theory generally, appears most essentially through applications of the ergodic theorem. See, for instance, the foundational work of Papanicolaou and Varadhan [23], who worked in the elliptic setting.

However, conditions (2) and (3) are merely qualitative and contain no quantitative information about the mixing properties of the ensemble. Therefore, while the results of this paper apply to a very general class of environments, the corresponding homogenization may occur at an arbitrarily slow rate. In order to obtain more quantitative statements, such as in the recent work Armstrong, Bordas and Mourrat [3], it would be necessary to quantify the ergodicity in the way, for example, of a spectral gap inequality or a finite-range of dependence.

The qualitative theory of homogenization for systems like (1) aims to characterize, for $\langle\cdot\rangle$-a.e. a, the limiting behavior, as $\epsilon \to 0$, of solutions to the rescaled equation

$$\begin{cases} u_t^\epsilon = \nabla \cdot a^\epsilon \nabla u^\epsilon & \text{in } \mathbb{R}^d \times (0, \infty) \\ u^\epsilon = u_0 & \text{on } \mathbb{R}^d \times \{0\}, \end{cases} \tag{5}$$

The first author is supported by the German Science Foundation DFG in context of the Emmy Noether junior research group BE 5922/1-1.

The second author is supported by LabEx Archiméde.

The third author is supported by the National Science Foundation Mathematical Sciences Postdoctoral Research Fellowship under Grant Number 1502731.

where

$$a^\epsilon(\cdot,\cdot) := a\left(\frac{\cdot}{\epsilon}, \frac{\cdot}{\epsilon^2}\right)$$

is a parabolic rescaling of the coefficient field. This is understood classically through the introduction of a space-time corrector $\phi = \{\phi_i\}_{i\in\{1,\dots,d\}}$ satisfying, for each $i \in \{1,\dots,d\}$,

$$\phi_{i,t} = \nabla \cdot a(\nabla\phi_i + e_i) \text{ in } \mathbb{R}^{d+1}. \tag{6}$$

Then, in view of the linearity, for each $\xi \in \mathbb{R}^d$ the corresponding corrector ϕ_ξ is defined by the sum

$$\phi_\xi := \xi_i\phi_i, \tag{7}$$

where here, and throughout the paper, the notation employs Einstein's summation convention over repeated indices.

The gradient of the corrector $\nabla\phi$ is a random field which is stationary with finite energy. That is, for each $x \in \mathbb{R}^d$, $t \in \mathbb{R}$ and $a \in \Omega$,

$$\nabla\phi(x,t;a) = \nabla\phi(0,0;a(\cdot+x,\cdot+t)),$$

and, for each $i \in \{1,\dots,d\}$,

$$\left\langle |\nabla\phi_i|^2 \right\rangle < \infty.$$

These facts are used to prove the strict sublinearity of the large-scale L^2-averages of ϕ on parabolic cylinders. Namely, for each $R > 0$, let B_R denote the ball of radius R centered at the origin and let $\mathcal{C}_R$ denote the parabolic cylinder

$$\mathcal{C}_R := B_R \times (-R^2, 0].$$

The corrector satisfies, for $\langle\cdot\rangle$-a.e. a, for each $i \in \{1,\dots,d\}$,

$$\lim_{R\to\infty} \frac{1}{R}\left(\fint_{\mathcal{C}_R} |\phi_i|^2\right)^{\frac{1}{2}} = 0, \tag{8}$$

where here, and throughout the paper, the integration variables will be omitted unless there is a possibility of confusion. This sublinearity is essentially equivalent to homogenization, see (12) below, and is crucial for the arguments of this paper.

The corrector is used to identify the homogenized coefficient field a_{hom} as the expectation of the components of the flux according to the rules, for $i \in \{1,\dots,d\}$,

$$a_{\mathrm{hom}}e_i := \langle a(\nabla\phi_i + e_i)\rangle, \tag{9}$$

where the flux $q = \{q_i\}_{i\in\{1,\dots,d\}}$ is defined, for each $i \in \{1,\dots,d\}$, by

$$q_i := a(\nabla\phi_i + e_i). \tag{10}$$

It is a classical fact that the homogenized coefficient field a_{hom} is uniformly elliptic and bounded, as is shown in Lemma 2. The solution of the corresponding constant-coefficient parabolic equation

$$\begin{cases} v_t = \nabla \cdot a_{\mathrm{hom}}\nabla v & \text{in } \mathbb{R}^d \times (0,\infty) \\ v = u_0 & \text{on } \mathbb{R}^d \times \{0\}, \end{cases} \tag{11}$$

then characterizes the limiting behavior, for $\langle\cdot\rangle$-a.e. a and as $\epsilon \to 0$, of the solutions to (5). Indeed, by obtaining an energy estimate for the error in the asymptotic expansion

$$u^\epsilon \simeq v + \epsilon\phi_i\left(\frac{\cdot}{\epsilon}, \frac{\cdot}{\epsilon^2}\right)\partial_i v,$$

which relies upon the sublinearity (8), it follows that, for $\langle\cdot\rangle$-a.e. a, for every $u_0 \in L^2(\mathbb{R}^d)$ and $T > 0$, as $\epsilon \to 0$,

$$u^\epsilon \to v \text{ strongly in } L^2(\mathbb{R}^d \times [0,T]). \tag{12}$$

This almost sure convergence is the qualitative homogenization of the original ensemble.

Looking ahead, observe that the behavior of the solution u^ϵ to (5) on a unit scale, for $\epsilon > 0$ small, corresponds to a characterization of the large-scale behavior of the solution u satisfying (1). Namely, the behavior of the solution u^ϵ on a unit scale corresponds to the behavior of u on scale ϵ^{-1} in space and ϵ^{-2} in time. The purpose of this paper will be to characterize the extent

to which solutions of (1) inherit, on large-scales and for $\langle\cdot\rangle$-a.e. a, the regularity of solutions to constant-coefficient parabolic equations.

A concise statement of this large-scale regularity is contained in the following first-order Liouville theorem, which is the main theorem of the paper.

Theorem 1. *Suppose that $\langle\cdot\rangle$ is stationary (2), ergodic (3) and bounded and uniformly elliptic (4). Then, $\langle\cdot\rangle$-a.e. a satisfies the following first-order Liouville property: if u is an ancient whole-space a-caloric function, that is if u is a distributional solution of*

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

which is strictly subquadratic on parabolic cylinders in the sense that, for some $\alpha \in (0,1)$,

$$\lim_{R\to\infty} \frac{1}{R^{1+\alpha}} \left(\fint_{\mathcal{C}_R} |u|^2 \right)^{\frac{1}{2}} = 0,$$

then there exists $c \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$ such that

$$u(x,t) = c + x \cdot \xi + \phi_\xi(x,t) \ \text{ in } \ \mathbb{R}^d \times (-\infty, 0),$$

for the corrector ϕ_ξ defined in (7).

The proof of Theorem 1 is strongly motivated by the work of Gloria, Neukamm and Otto [17], who considered precisely these questions for stationary and ergodic ensembles of elliptic equations. It is based on controlling the large-scale L^2-deviation of the gradient of an a-caloric function from the span of the a-caloric gradients $\{\xi + \nabla\phi_\xi\}_{\xi\in\mathbb{R}^d}$. The excess of an a-caloric function measures this deviation, and is defined, for each $R > 0$ and a-caloric function u on $\mathcal{C}_R$, by

$$\mathrm{Exc}(u;R) := \inf_{\xi\in\mathbb{R}^d} \fint_{\mathcal{C}_R} (\nabla u - \xi - \nabla\phi_\xi) \cdot a(\nabla u - \xi - \nabla\phi_\xi). \tag{13}$$

In Proposition 4 below, for $\langle\cdot\rangle$-a.e. a, the excess of an a-caloric function will be shown to decay like a power law in the radius. However, before the statement, it is useful to observe some essential differences between the parabolic and elliptic settings. In what follows, the superscript "ell" will be used to differentiate elliptic objects from their parabolic counterparts.

In the elliptic case, for a stationary and ergodic ensemble $\langle\cdot\rangle^{\mathrm{ell}}$ of bounded, uniformly elliptic coefficient fields a^{ell}, the corrector $\phi^{\mathrm{ell}} = \{\phi_i^{\mathrm{ell}}\}_{i\in\{1,\dots,d\}}$ is defined by the equations, for $i \in \{1,\dots,d\}$,

$$-\nabla \cdot a^{\mathrm{ell}}(\nabla\phi_i^{\mathrm{ell}} + e_i) = 0 \ \text{ in } \ \mathbb{R}^d, \tag{14}$$

and, for each $\xi \in \mathbb{R}^d$,

$$\phi_\xi^{\mathrm{ell}} := \phi_i^{\mathrm{ell}}\xi_i.$$

These correctors play a virtually identical role to the parabolic correctors (6) in elliptic homogenization theory.

Excess for uniformly elliptic ensembles was first defined in [17, Lemma 2], although they worked not with the intrinsic energy defined by a but with the equivalent L^2-energy. This differs, for instance, from the definition used in the work of the first author, third author and Otto [11], which considered degenerate elliptic ensembles for which it was essential to incorporate the environment a. These notions motivated definition (13), and measured the deviation of the gradient of an a^{ell}-harmonic function u on B_R, by which is meant a solution

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

from the span of a^{ell}-harmonic gradients $\{\xi + \nabla\phi_\xi^{\mathrm{ell}}\}_{\xi\in\mathbb{R}^d}$. Precisely, for each $R > 0$ and a^{ell}-harmonic function u on B_R,

$$\mathrm{Exc}^{\mathrm{ell}}(u;R) := \inf_{\xi\in\mathbb{R}^d} \fint_{B_R} (\nabla u - \xi - \nabla\phi_\xi^{\mathrm{ell}}) \cdot a^{\mathrm{ell}}(\nabla u - \xi - \nabla\phi_\xi^{\mathrm{ell}}). \tag{15}$$

The decay of the excess was controlled in [17, Lemma 2] through the introduction of a flux correction $\sigma^{\mathrm{ell}} = \{\sigma_i^{\mathrm{ell}}\}_{i\in\{1,\dots,d\}}$. Namely, the flux $q^{\mathrm{ell}} = \{q_i^{\mathrm{ell}}\}_{i\in\{1,\dots,d\}}$ is defined, for each $i \in \{1,\dots,d\}$, by

$$q_i^{\mathrm{ell}} := a^{\mathrm{ell}}(\nabla\phi_i^{\mathrm{ell}} + e_i),$$

where, in analogy with the parabolic setting, the homogenized coefficient field $a_{\mathrm{hom}}^{\mathrm{ell}}$ is defined by the expectation of the components of the flux, for $i \in \{1,\dots,d\}$,

$$a_{\mathrm{hom}}^{\mathrm{ell}} e_i := \left\langle a^{\mathrm{ell}}(\nabla\phi_i^{\mathrm{ell}} + e_i)\right\rangle^{\mathrm{ell}}.$$

Therefore, strictly in the elliptic case, the corrector equation (14) asserts that the components of the flux are divergence-free and may be viewed as closed $(d-1)$-forms on the whole space. Hence, for each $i \in \{1,\dots,d\}$, there exists a $(d-2)$-form, which is represented by a skew-symmetric matrix $\sigma_i^{\mathrm{ell}} = (\sigma_{ijk})_{j,k\in\{1,\dots,d\}}^{\mathrm{ell}}$, satisfying

$$\nabla\cdot\sigma_i^{\mathrm{ell}} = q_i^{\mathrm{ell}} - a_{\mathrm{hom}}^{\mathrm{ell}} e_i, \tag{16}$$

where the divergence of the tensor-field σ_i^{ell} is defined, for each $i,j \in \{1,\dots,d\}$, by

$$(\nabla\cdot\sigma_i^{\mathrm{ell}})_j = \sum_{k=1}^{d} \partial_k \sigma_{ijk}^{\mathrm{ell}}.$$

Furthermore, the flux correction σ^{ell} is fixed according to the choice of gauge, for each $i,j,k \in \{1,\dots,d\}$,

$$\Delta\sigma_{ijk}^{\mathrm{ell}} = \partial_k q_{ij}^{\mathrm{ell}} - \partial_j q_{ik}^{\mathrm{ell}}.$$

In [17, Lemma 2], the sublinearity of the large-scale L^2-averages of the extended corrector $(\phi^{\mathrm{ell}}, \sigma^{\mathrm{ell}})$ on large balls is shown to imply that the elliptic excess (15) decays as a power law in the radius.

Precisely, for each $\alpha \in (0,1)$, there exists $C_0^{\mathrm{ell}} = C_0^{\mathrm{ell}}(\alpha,d,\lambda) > 0$ and $C_1^{\mathrm{ell}} = C_1^{\mathrm{ell}}(\alpha,d,\lambda) > 0$ for which, whenever a pair of radii $0 < r < R < \infty$ satisfy, for every $\rho \in [r,R]$, for each $i \in \{1,\dots,d\}$,

$$\frac{1}{\rho}\left(\fint_{B_\rho} \left|\phi_i^{\mathrm{ell}}\right|^2 + \left|\sigma_i^{\mathrm{ell}}\right|^2\right)^{\frac{1}{2}} \le \frac{1}{C_0^{\mathrm{ell}}},$$

then, for every a^{ell}-harmonic function u in B_R,

$$\mathrm{Exc}^{\mathrm{ell}}(u;r) \le C_1^{\mathrm{ell}}\left(\frac{r}{R}\right)^{2\alpha}\mathrm{Exc}^{\mathrm{ell}}(u;R). \tag{17}$$

Observe, in particular, that this is a deterministic result. Indeed, the stochastic properties of the extended corrector $(\phi^{\mathrm{ell}}, \sigma^{\mathrm{ell}})$ are necessary to prove that, for $\langle\cdot\rangle$-a.e. a, the large-scale L^2-averages are sublinear. But, by taking this fact as an input, it follows from a Campanato iteration that the excess of an arbitrary a^{ell}-harmonic function decays according to (17). In this paper, the analogous result will also be obtained for the parabolic excess, as shown in Proposition 4 below.

The first essential difference is that, unlike in the elliptic case, the fluxes $\{q_i\}_{i\in\{1,\dots,d\}}$ defined in (10) are not divergence-free, and so an immediate analogue of the flux correction $\sigma^{\mathrm{ell}} = \{\sigma_i^{\mathrm{ell}}\}_{i\in\{1,\dots,d\}}$ cannot be defined. Instead, the flux is essentially decomposed according to the Weyl decomposition, where the parabolic σ is constructed to correct the divergence-free component. Precisely, for each $i \in \{1,\dots,d\}$,

$$q_i = q_{i,\mathrm{sol}} + \nabla\psi_i,$$

where the solenoidal part $q_{i,\mathrm{sol}}$ is divergence-free and $\nabla\psi_i$ is a stationary, finite-energy gradient chosen to satisfy

$$\Delta\psi_i = \nabla\cdot q_i. \tag{18}$$

Indeed, for each $i \in \{1,\dots,d\}$, one first defines $\nabla\psi_i$ according to (18) and then observes that $q_i - \nabla\psi_i$ is divergence-free.

The flux correction $\sigma = \{\sigma_i\}_{i\in\{1,\dots,d\}}$ is then defined, for each $i \in \{1,\dots,d\}$, by the equation

$$\nabla \cdot \sigma_i = (q_i - \nabla\psi_i) - \langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle, \tag{19}$$

where $\langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle$ denotes the conditional expectation of q_i with respect to the sub-sigma-algebra $\mathcal{F}_{\mathbb{R}^d} \subset \mathcal{F}$ of subsets of Ω which are invariant with respect to spatial translations of the coefficient fields. They are fixed following the choice of gauge, for each $i,j,k \in \{1,\dots,d\}$,

$$\Delta\sigma_{ijk} = \partial_k(q_i - \nabla\psi_i)_j - \partial_j(q_i - \nabla\psi_i)_k. \tag{20}$$

Finally, for each $i \in \{1,\dots,d\}$, it is necessary to correct the oscillations of the conditional expectation $\langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle$ about its mean. The corrector $\zeta = \{\zeta_{ij}\}_{i,j\in\{1,\dots,d\}}$ is constructed explicitly for this purpose and satisfies, for each $i \in \{1,\dots,d\}$,

$$\partial_t\zeta_i = \langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle - \langle q_i\rangle = \langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle - a_{\mathrm{hom}}e_i. \tag{21}$$

In particular, this final correction ζ is constant in space, as a $\mathcal{F}_{\mathbb{R}^d}$-measurable field, and depends only on time.

In comparison with the elliptic setting, where the decay of the excess was determined by the sublinearity of the large-scale L^2-averages of $(\phi^{\mathrm{ell}}, \sigma^{\mathrm{ell}})$, the decay of the parabolic excess will be determined by the sublinearity of the large-scale L^2-averages of the corrector (ϕ,ψ,σ), measured with respect to the scaling in space, and the sublinearity of the large-scale L^2-averages of ζ, measured with respect to the scaling in time. The first lemma of the paper establishes the existence of the extended corrector (ϕ,ψ,σ,ζ).

Lemma 2. *Suppose that the ensemble $\langle\cdot\rangle$ satisfies (2), (3) and (4). There exist $C = C(d,\lambda) > 0$ and four random fields $\phi = \{\phi_i\}_{i\in\{1,\dots,d\}}$, $\psi = \{\psi_i\}_{i\in\{1,\dots,d\}}$, $\sigma = \{\sigma_{ijk}\}_{i,j,k\in\{1,\dots,d\}}$ and $\zeta = \{\zeta_{ij}\}_{i,j\in\{1,\dots,d\}}$ on $\mathbb{R}^{d+1}$ with the following properties:*

The gradient fields are stationary, finite energy random processes with vanishing expectation: for each $i,j,k \in \{1,\dots,d\}$,

$$\left\langle |\nabla\phi_i|^2 \right\rangle + \left\langle |\nabla\psi_i|^2 \right\rangle + \left\langle |\nabla\sigma_{ijk}|^2 \right\rangle + \left\langle |\partial_t\zeta_{ij}|^2 \right\rangle \le C,$$

and

$$\langle\nabla\phi_i\rangle = \langle\nabla\psi_i\rangle = \langle\nabla\sigma_{ijk}\rangle = \langle\partial_t\zeta_{ij}\rangle = 0.$$

For each $i \in \{1,\dots,d\}$, the field $\sigma_i = (\sigma_{ijk})_{j,k\in\{1,\dots,d\}}$ is skew-symmetric in its last two indices: for each $i,j,k \in \{1,\dots,d\}$,

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

The fields ψ and σ are stationary in time: for each $x \in \mathbb{R}^d$, $t \in \mathbb{R}$ and $a \in \Omega$,

$$\psi(x,t;a) = \psi(x,0;a(\cdot,\cdot+t)),$$

and

$$\sigma(x,t;a) = \sigma(x,0;a(\cdot,\cdot+t)).$$

Furthermore, for $\langle\cdot\rangle$-a.e. a, the following equations are satisfied in the sense of distributions on $\mathbb{R}^{d+1}$. The field ϕ satisfies (6): for each $i \in \{1,\dots,d\}$,

$$\phi_{i,t} = \nabla\cdot a(\nabla\phi_i + e_i).$$

The potential part of the flux is corrected by ψ according to (18): for each $i \in \{1,\dots,d\}$,

$$\Delta\psi_i = \nabla\cdot q_i.$$

The field σ corrects the divergence-free part of the flux according to (19): for each $i \in \{1,\dots,d\}$,

$$\nabla\cdot\sigma_i = q_i - \nabla\psi_i - \langle q_i \mid \mathcal{F}_{\mathbb{R}^d}\rangle,$$

where $\langle\cdot \mid \mathcal{F}_{\mathbb{R}^d}\rangle$ denotes the conditional expectation with respect to the sub-sigma-algebra $\mathcal{F}_{\mathbb{R}^d} \subset \mathcal{F}$ of subsets of Ω which are invariant with respect to spatial translations of the coefficient field. Furthermore, σ is constructed according to the choice of gauge, for each $i,j,k \in \{1,\dots,d\}$,

$$\Delta\sigma_{ijk} = \partial_k(q_i - \nabla\psi_i)_j - \partial_j(q_i - \nabla\psi_i)_k.$$

The field ζ corrects the oscillation of the conditional expectation about its mean: for each $i \in \{1, \dots, d\}$, the random vector field ζ_i is constant in space and satisfies

$$\partial_t \zeta_i = \langle q_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle - \langle q_i \rangle = \langle q_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle - a_{\mathrm{hom}} e_i.$$

Finally, the homogenized coefficient field a_{hom} defined in (9) is bounded and uniformly elliptic: for each $\xi \in \mathbb{R}^d$,

$$\lambda |\xi|^2 \leq \xi \cdot a_{\mathrm{hom}} \xi \quad \text{and} \quad |a_{\mathrm{hom}} \xi| \leq \frac{1}{\lambda} |\xi| .$$

The following two propositions effectively split the probabilistic and deterministic aspects of the paper. Proposition 3 contains the probabilistic parts, and uses the stationarity and ergodicity of the ensemble to prove that the large-scale L^2-averages of (ϕ, ψ, σ) are sublinear with respect to the spatial scaling and that those of ζ are sublinear with respect to the time scaling. This fact is essentially classical for the case of the correctors ϕ and ζ, although a new argument for the sublinearity of ϕ is presented which may be of independent interest. A new argument is required to prove the sublinearity of σ and ψ.

The difference is the following. The corrector ϕ is, in general, not stationary in either space or time but equation (6) yields some control over both its spatial and temporal derivatives. Similarly, the corrector ζ has an explicit, stationary time derivative but is itself not stationary. In the second case, since equations (18) and (19) yield only the spatial regularity for ψ and σ, it is necessary to use the fact that both fields are stationary in time in order to obtain the convergence.

In fact, the following proposition will prove the sublinearity of the normalized corrector where, in the case of ϕ, the components are normalized by their large-scale averages on a parabolic cylinder, in the case of ζ, using the fact that the Sobolev embedding implies that ζ is continuous, the components are normalized by their value at time zero and, in the case of ψ and σ, the functions are normalized, for each fixed time, by their large-scale averages on a ball. This is in fact equivalent to the sublinearity of the corrector $(\phi, \psi, \sigma, \zeta)$ without a normalization, see for instance [11, Lemma 2], but since this observation is not necessary for the arguments of the paper it is omitted.

For an arbitrary function $\varphi : \mathbb{R}^{d+1} \to \mathbb{R}$, define, for each $R > 0$ and $t \in \mathbb{R}$,

$$(\varphi)_R := \fint_{\mathcal{C}_R} \varphi \ \text{ and } \ (\varphi)_{t,R} := \fint_{B_R} \varphi(\cdot, t).$$

The precise normalization considered and the corresponding sublinearity are contained in the following proposition.

Proposition 3. *Suppose that the ensemble $\langle \cdot \rangle$ satisfies (2), (3) and (4). Then, for $\langle \cdot \rangle$-a.e a, the corrector (ϕ, ψ, σ) is strictly sublinear with respect to the spatial scaling and the corrector ζ is strictly sublinear with respect to the time scaling in the sense that, for each $i, j, k \in \{1, \dots, d\}$,*

$$\lim_{R \to \infty} \frac{1}{R} \left(\fint_{C_R} |\phi_i - (\phi_i)_R|^2 + |\psi_i - (\psi_i)_{t,R}|^2 + |\sigma_{ijk} - (\sigma_{ijk})_{t,R}|^2 \right)^{\frac{1}{2}} = 0, \tag{22}$$

and

$$\lim_{R \to \infty} \frac{1}{R^2} \left(\fint_{C_R} |\zeta_{ij} - \zeta_{ij}(0)|^2 \right)^{\frac{1}{2}} = 0. \tag{23}$$

Furthermore, for $\langle \cdot \rangle$-a.e. a, for each $i \in \{1, \dots, d\}$, the large-scale L^2-averages of the components of the flux satisfy

$$\lim_{R \to \infty} \left(\fint_{\mathcal{C}_R} |q_i|^2 \right)^{\frac{1}{2}} = \left\langle |q_i|^2 \right\rangle^{\frac{1}{2}}. \tag{24}$$

It is important to observe at this point that equations (18) and (19) defining ψ and σ are invariant if either ψ or σ is altered by a time stationary constant. This explains why in (22), for each $R > 0$, it remains useful to allow for a time-dependent normalization. The equations (6) and (21) defining ϕ and ζ are not likewise invariant, and therefore the corresponding normalizations appearing in (22) and (23) are necessarily achieved by subtracting a true constant.

The deterministic aspect of the paper uses a Campanato iteration, which takes the conclusion of Proposition 3 as input. Namely, it will be shown that, for any $\alpha \in (0,1)$, the parabolic excess decays like a power law in the radius as soon as the quantities appearing in (22) and (23) are sufficiently small and as soon as (24) is sufficiently close to its expectation. This is to say that there exists a random but $\langle\cdot\rangle$-a.e. a finite radius $r_*(a)$ such that, whenever $r_* < r < R < \infty$, for every a-caloric function u in $\mathcal{C}_R$, the parabolic excess satisfies, for $C_1 = C_1(\alpha, d, \lambda) > 0$,

$$\operatorname{Exc}(u; r) \leq C_1 \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R).$$

This is the content of the following proposition.

Proposition 4. *Suppose that the ensemble $\langle\cdot\rangle$ satisfies (2), (3) and (4). Fix a Hölder exponent $\alpha \in (0,1)$. Then, there exist constants $C_0 = C_0(\alpha, d, \lambda) > 0$ and $C_1(\alpha, d, \lambda) > 0$ with the following property:*

If $R_1 < R_2$ are two radii such that, for each $R \in [R_1, R_2]$ and for each $i, j, k \in \{1, \dots, d\}$,

$$\frac{1}{R}\left(\fint_{\mathcal{C}_R} |\phi_i - (\phi_i)_R|^2 + |\psi_i - (\psi_i)_{t,R}|^2 + |\sigma_{ijk} - (\sigma_{ijk})_{t,R}|^2\right)^{\frac{1}{2}} \leq \frac{1}{C_0},$$

and

$$\frac{1}{R^2}\left(\fint_{\mathcal{C}_R} |\zeta_{ij} - \zeta_{ij}(0)|^2\right)^{\frac{1}{2}} \leq \frac{1}{C_0},$$

and such that, for each $i \in \{1, \dots, d\}$ and $R \in [R_1, R_2]$,

$$\left(\fint_{\mathcal{C}_R} |q_i|^2\right)^{\frac{1}{2}} \leq 2\left\langle |q_i|^2 \right\rangle^{\frac{1}{2}},$$

then any distributional solution u to the parabolic equation

$$u_t = \nabla \cdot a \nabla u \ \text{ in } \ \mathcal{C}_{R_2}$$

satisfies

$$\operatorname{Exc}(u; R_1) \leq C_1 \left(\frac{R_1}{R_2}\right)^{2\alpha} \operatorname{Exc}(u; R_2).$$

The proof of Proposition 4 is motivated by the proof of [17, Lemma 2] from the elliptic setting. There, the flux correction σ^{ell} was used to express the residuum of the homogenization error in a useful divergence form. That is, for $R > 0$, given an a^{ell}-harmonic function u in B_R, define its $a^{\mathrm{ell}}_{\mathrm{hom}}$-harmonic extension v into B_R to be the solution

$$\begin{cases} -\nabla \cdot a^{\mathrm{ell}}_{\mathrm{hom}} \nabla v &= 0 \quad \text{in } B_R \\ v &= u \quad \text{on } \partial B_R. \end{cases}$$

Then, for a smooth cutoff function η vanishing along the boundary ∂B_R, define the augmented homogenization error w^{ell} to be the following modification of the classical two-scale expansion

$$w^{\mathrm{ell}} := u - (1 + \eta \phi_i^{\mathrm{ell}} \partial_i) v,$$

where the cutoff is used in order to guarantee the difference w^{ell} vanishes on the boundary.

It was proven that the augmented homogenization error w^{ell} satisfies

$$\begin{cases} -\nabla \cdot a^{\mathrm{ell}} \nabla w^{\mathrm{ell}} = \ \nabla \cdot \left((1-\eta)(a^{\mathrm{ell}} - a^{\mathrm{ell}}_{\mathrm{hom}})\nabla v + (\phi_i^{\mathrm{ell}} a^{\mathrm{ell}} - \sigma_i^{\mathrm{ell}})\nabla(\eta \partial_i v)\right) & \text{in } B_R \\ w^{\mathrm{ell}} = \ 0 & \text{on } \partial B_R, \end{cases} \tag{25}$$

which, by testing the equation with w^{ell}, yields a useful energy estimate that provides the starting point for a Campanato iteration.

In particular, by analyzing the right hand side of (25), the energy of w^{ell} can be controlled by the growth of the extended corrector $(\phi^{\mathrm{ell}}, \sigma^{\mathrm{ell}})$, the choice of the cutoff function η and the interior and boundary regularity of the $a^{\mathrm{ell}}_{\mathrm{hom}}$-harmonic function v. The argument is completed by observing that, owing to the regularity of $a^{\mathrm{ell}}_{\mathrm{hom}}$-harmonic functions, the energy of the homogenization error is a good approximation for the excess.

The methods of this paper apply the same philosophy to the parabolic setting. However, similarly to what was done in the proof of [11, Theorem 2], it is furthermore necessary to introduce a spatial regularization of the a-caloric function u. The purpose of this is to obtain more regularity in time, since such functions are already sufficiently regular in space. Precisely, if u is an a-caloric function then, in general, its time derivative $u_t \in H^{-1}$ and no better, where H^{-1} denotes the dual space of the Sobolev space H^1. However, for every $\epsilon > 0$, if u^ϵ denotes the spatial convolution of u on scale $\epsilon > 0$, then it is possible to show that $u_t^\epsilon \in L^2$, see Section 5.1 below. This additional approximation is necessary in order to apply the boundary estimate of Section 5.3.

For $\epsilon > 0$, the a_{hom}-caloric function v^ϵ will then be the a_{hom}-caloric extension of the spatial regularization u^ϵ into $\mathcal{C}_R$. Namely, for $R > 0$ and $\epsilon > 0$, given an a-caloric function u in $\mathcal{C}_{R+\epsilon}$, define the a_{hom}-caloric extension v^ϵ of u^ϵ into $\mathcal{C}_R$ to be the solution

$$\begin{cases} v_t^\epsilon &= \nabla \cdot a_{\mathrm{hom}} \nabla v^\epsilon \quad \text{in } \mathcal{C}_R \\ v^\epsilon &= u^\epsilon \qquad\qquad\qquad \text{on } \partial_p \mathcal{C}_R, \end{cases}$$

where $\partial_p \mathcal{C}_R$ denotes the parabolic boundary

$$\partial_p \mathcal{C}_R := \left(B_R \times \{-R^2\}\right) \cup \left(\partial B_R \times [-R^2, 0]\right).$$

Then, again motivated by the classical two-scale expansion, for $\epsilon > 0$ and a smooth cutoff function η vanishing on the parabolic boundary $\partial_p \mathcal{C}_R$, the augmented homogenization error w will be defined as

$$w := u - (1 + \eta \phi_i \partial_i) v^\epsilon.$$

It will be shown in the proof of Proposition 4 that the augmented homogenization error satisfies

$$\begin{cases} w_t - \nabla \cdot a \nabla w = & \nabla \cdot ((1-\eta)(a - a_{\mathrm{hom}}) \nabla v^\epsilon) + \nabla \cdot ((\phi_i a + \psi_i - \sigma_i) \nabla (\eta \partial_i v^\epsilon)) \quad \text{in } \mathcal{C}_R \\ & + \partial_t \zeta_i \cdot \nabla (\eta \partial_i v^\epsilon) - \phi_i (\eta \partial_i v^\epsilon)_t - \psi_i \Delta (\eta \partial_i v^\epsilon) \\ w = & u - u^\epsilon \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial_p \mathcal{C}_R. \end{cases} \tag{26}$$

As in the elliptic setting, the energy estimate obtained by testing this equation with w, for an appropriately chosen cutoff η, will be the starting point of the Campanato iteration used to control the decay of the excess. In this case, there is a contribution from the boundary, which will be controlled first by fixing $\epsilon > 0$ small. From the right hand side of (26), the energy of the homogenization error will then be controlled by the growth of the extended parabolic corrector (ϕ, ψ, σ) and, after integrating in parts by time, the growth of ζ and q. It is furthermore necessary to make a good choice for the cutoff function η and to use the interior and boundary regularity of the a_{hom}-caloric function v^ϵ. The argument is completed by observing that, owing to the interior regularity of a_{hom}-caloric functions, the homogenization error w provides a good approximation for the excess.

Finally, the following parabolic Caccioppoli inequality will be used in the proofs of Theorem 1 and Proposition 4. The proof is classical, and is included for the convenience of the reader.

Lemma 5. *Suppose that $\langle \cdot \rangle$ satisfies (4). There exists $C = C(\lambda) > 0$ such that, for $\langle \cdot \rangle$-a.e. a, for every $R > 0$ and distributional solution u of the equation*

$$u_t = \nabla \cdot a \nabla u \ \text{ in } \mathcal{C}_R,$$

and for every $c \in \mathbb{R}$ and $\rho \in (0, \frac{R}{2})$,

$$\int_{\mathcal{C}_{R-\rho}} |\nabla u|^2 \leq \frac{C}{\rho^2} \int_{\mathcal{C}_R \setminus \mathcal{C}_{R-\rho}} |u - c|^2. \tag{27}$$

In comparison with the elliptic setting, the qualitative homogenization theory of divergence-form operators with coefficients depending on time and space is relatively under studied. While the case of periodic coefficients has long been understood, and a full explanation can be found in the classic reference Bensoussan, Lions and Papanicolaou [12, Chapter 3], the qualitative stochastic homogenization of stationary and ergodic ensembles like (1) was obtained only more recently by Rhodes [25, 26]. However, related problems were earlier handled, such as the case of a Brownian motion in the presence of a divergence-free drift, by Komorowski and Olla [18],

Landim, Olla and Yau [21] and Oelschläger [22]. In the discrete setting, related questions have been considered, for instance, by Andres [1], Bandyopadhyay and Zeitouni [10] and Rassoul-Agha and Seppäläinen [24] in the uniformly elliptic setting and, for degenerate environments, by Andres, the second author, Deuschel and Slowik [2].

The quantitative homogenization of such ensembles has only recently been considered, and the preprint [3] contains, to our knowledge, the first results in this direction. In particular, in [3, Theorem 1.2], a full hierarchy of Liouville theorems is obtained for ensembles satisfying a finite-range dependence in space and time. Their method is motivated by the work of Armstrong and Smart [6] from the elliptic setting, which adapted the approach of Avellaneda and Lin [7] from the context of periodic homogenization.

In [7], a full hierarchy of Liouville properties was established for uniformly elliptic and periodic coefficient fields based upon the previous works Avellaneda and Lin [8, 9], which developed a large-scale regularity theory in Hölder and L^p-spaces. In [6], the approach of [7] was adapted to stationary and ergodic ensembles satisfying a finite-range dependence. Their proof, which obtained a large-scale $C^{0,1}$-regularity theory, was based upon a variational approach and the quantification of the convergence of certain sub-additive and super-additive energies. Their work was later extended by Armstrong and Mourrat [5] to more general mixing conditions, and subsequently gave rise to a significant literature on the subject. The interested reader is pointed to the recent monograph Armstrong, Kuusi and Mourrat [4], and the references therein.

The approach of this paper follows closely the work [17], which derived, for uniformly elliptic ensembles, a large-scale $C^{1,\alpha}$-regularity estimate and first-order Liouville property under the qualitative assumptions of stationarity and ergodicity. The method was based upon the introduction of an intrinsic notion of excess, as defined in (15), as well as the construction of the flux correction σ^{ell} defined in (16). The introduction of σ^{ell} was used to prove that the homogenization error solves the divergence-form equation (25), which provided the starting point for a Campanato iteration as explained above.

Subsequently, Fischer and Otto [15] obtained a full hierarchy of Liouville properties under a mild quantification of the ergodicity. In Fischer and Otto [16], the necessary quantification of ergodicity from [15] was shown to be satisfied by a general class of Gaussian environments. However, absent some mild quantification of ergodicity in the sense of either [5] or [15], the existence of higher order Liouville and large-scale regularity statements remains an open question.

Finally, motivated by the work of the second author and Deuschel [14], the first author, third author and Otto [11] derived a large-scale $C^{1,\alpha}$-regularity theory and first-order Liouville theorem for degenerate elliptic equations, where the boundedness and uniform ellipticity (4) was replaced by certain moment conditions. It is expected that the results of this paper can be similarly extended to degenerate environments, and the setting of [2] will serve as the starting point for future work.

In principle, one could also hope to combine the methods of this paper with those of [17], in the presence of a logarithmic Sobolev inequality like that used in [17, Theorem 1], to obtain more quantitative information. For example, the minimal radius $r_*(a) > 0$ defined, for $C_0 > 0$ from Proposition 4, after suppressing the normalizations in the notation, by

$$r_*(a) = \inf \left\{ r \geq 1 \;\middle|\; \text{For all } R \geq r, \text{ for each } i \in \{1, \ldots, d\}, \; \left(\fint_{\mathcal{C}_R} |q_i|^2 \right)^{\frac{1}{2}} \leq 2 \left\langle |q_i|^2 \right\rangle^{\frac{1}{2}} \right.$$
$$\left. \text{and } \frac{1}{R} \left(\fint_{B_R} |\phi_i|^2 + |\psi_i|^2 + |\sigma_i|^2 \right)^{\frac{1}{2}} + \frac{1}{R^2} \left(\fint_{\mathcal{C}_R} |\zeta_i|^2 \right)^{\frac{1}{2}} \leq \frac{1}{C_0} \right\},$$

which effectively defines the initial scale on which the $C^{1,\alpha}$-regularity of Proposition 4 begins to take effect, is expected to have stretched exponential moments in the sense of [17, Theorem 1]. Furthermore, again assuming a logarithmic Sobolev inequality, it should be possible to obtain a quantitative two-scale expansion for a-caloric functions like [17, Corollary 3]. Lastly, following the methods of [15], it may be possible to prove higher order Liouville statements under a mild quantification of the ergodicity.

The paper is organized as follows. The proofs are presented in the order of their appearance: Theorem 1, Lemma 2, Proposition 3, Proposition 4 and Lemma 5. In order to simplify the notation, the statements and proofs are written for the non-symmetric scalar setting. However, at the cost of increasing some constants, all of the arguments carry through unchanged for non-symmetric systems. Throughout, the notation $\lesssim$ is used to denote a constant whose dependencies are specified in every case by the statement of the respective theorem, proposition or lemma.

2. The proof of Theorem 1

Fix a coefficient field a satisfying the conclusions of Lemma 2, Proposition 3 and Proposition 4, and suppose that u is a distributional solution of

$$u_t = \nabla \cdot a\nabla u \ \text{ on } \ \mathbb{R}^d \times (-\infty, 0),$$

which is strictly subquadratic in the sense that, for some $\alpha \in (0,1)$,

$$\lim_{R\to\infty} \frac{1}{R^{1+\alpha}} \left(\fint_{\mathcal{C}_R} |u|^2 \right)^{\frac{1}{2}} = 0.$$

Fix $C_0 = C_0(\alpha, d, \lambda) > 0$ satisfying the hypothesis of Proposition 4. Then, using Proposition 3, fix $R_0 > 0$ such that, for every $R > R_0$, for each $i, j, k \in \{1, \ldots, d\}$,

$$\frac{1}{R} \left(\fint_{\mathcal{C}_R} |\phi_i - (\phi_i)_R|^2 + |\psi_i - (\psi_i)_{t,R}|^2 + |\sigma_{ijk} - (\sigma_{ijk})_{t,R}|^2 \right)^{\frac{1}{2}} \leq \frac{1}{C_0},$$

and

$$\frac{1}{R^2} \left(\fint_{\mathcal{C}_R} |\zeta_{ij} - \zeta_{ij}(0)|^2 \right)^{\frac{1}{2}} < \frac{1}{C_0},$$

and such that, for each $i \in \{1, \ldots, d\}$ and $R > R_0$,

$$\left(\fint_{\mathcal{C}_R} |q_i|^2 \right)^{\frac{1}{2}} \leq 2 \left\langle |q_i|^2 \right\rangle^{\frac{1}{2}}.$$

The definition of the excess, Proposition 4 and the Caccioppoli inequality (27) imply that, for each $R_0 < \rho < R$,

$$\operatorname{Exc}(u;\rho) \lesssim \left(\frac{\rho}{R}\right)^{2\alpha} \operatorname{Exc}(u;R) \leq \left(\frac{\rho}{R}\right)^{2\alpha} \fint_{\mathcal{C}_R} |\nabla u|^2 \lesssim \frac{\rho^{2\alpha}}{(2R)^{2+2\alpha}} \fint_{\mathcal{C}_{2R}} |u|^2 .$$

Therefore, since u is strictly subquadratic,

$$\operatorname{Exc}(u;\rho) \lesssim \rho^{2\alpha} \limsup_{R\to\infty} \frac{1}{(2R)^{2+2\alpha}} \fint_{\mathcal{C}_{2R}} |u|^2 = 0.$$

This implies that, for every $\rho > 0$,

$$\operatorname{Exc}(u;\rho) = 0.$$

It is then immediate from the definition of parabolic excess (13), since $\mathcal{C}_{R_1} \subset \mathcal{C}_{R_2}$ whenever $R_1 < R_2$, that there exists $\xi \in \mathbb{R}^d$ for which the difference

$$z(x,t) := u(x,t) - \xi \cdot x - \phi_\xi(x,t)$$

satisfies

$$\nabla z = 0 \ \text{ in } \ \mathbb{R}^d \times (-\infty, 0).$$

However, because z is a distributional solution of

$$z_t = \nabla \cdot a\nabla z \ \text{ in } \ \mathbb{R}^d \times (-\infty, 0),$$

it follows that z is necessarily constant in time as well. Therefore, there exists $c \in \mathbb{R}$ such that $u = c + \xi \cdot x + \phi_\xi$, which completes the argument.

3. The proof of Lemma 2

The construction of the corrector $(\phi, \psi, \sigma, \zeta)$ will be achieved by lifting the relevant equations (6), (18), (19) and (21) to the probability space Ω, and thereby identifying ϕ by its stationary, finite energy gradient and time derivative, ψ and σ by their stationary, finite energy gradients, and ζ by its stationary time derivative. For this, it is necessary to define the horizontal derivative of a random variable as induced by shifts of the coefficient field in space and time. Then, these will be used to define an analogue of the Sobolev space H^1 on the probability space.

Given an $L^2(\Omega)$ random variable f, define, for each $i \in \{1, \dots, d\}$, the horizontal derivative

$$D_i f(a) := \lim_{h \to 0} \frac{f(a(\cdot + he_i, \cdot)) - f(a)}{h}, \tag{28}$$

where the above limit is understood in the sense of strong $L^2(\Omega)$-convergence. Of course, it is not true in general that the above limit exists for every $f \in L^2(\Omega)$, but the horizontal derivatives are closed, densely defined operators on $L^2(\Omega)$, see [23], with domains, for $i \in \{1, \dots, d\}$,

$$\mathcal{D}(D_i) := \{\ f \in L^2(\Omega) \mid D_i f \ \text{ exists as an element of } \ L^2(\Omega)\ \}.$$

Similarly, for each $f \in L^2(\Omega)$, define the horizontal time derivative

$$D_0 f(a) := \lim_{h \to 0} \frac{f(a(\cdot, \cdot + h)) - f(a)}{h}, \tag{29}$$

which is a closed, densely defined operator on $L^2(\Omega)$ with domain

$$\mathcal{D}(D_0) := \{\ f \in L^2(\Omega) \mid D_0 f \ \text{ exists as an element of } \ L^2(\Omega)\ \}.$$

The analogue of the Hilbert space H^1 is then defined on the probability space as the intersection

$$\mathcal{H}^1 := \cap_{i=0}^d \mathcal{D}(D_i),$$

equipped with the inner product, for each $f, g \in \mathcal{H}^1$,

$$(f, g)_{\mathcal{H}^1} = \langle fg \rangle + \langle D_0 f D_0 g \rangle + \langle Df \cdot Dg \rangle,$$

for the horizontal spatial gradient

$$Df := (D_1 f, \dots, D_d f).$$

Finally, define the space of spatial potentials

$$L^2_{\text{pot}}(\Omega; \mathbb{R}^d) := \overline{\{\ Df \mid f \in \mathcal{H}^1\ \}}^{\ L^2(\Omega)-\text{weak}}, \tag{30}$$

as the $L^2(\Omega)$-weak closure of spatial gradients arising from $\mathcal{H}^1$ functions. Indeed, it is immediate from the weak convergence that elements of $L^2_{\text{pot}}(\Omega; \mathbb{R}^d)$ are potentials in the sense that every $F = (F_1, \dots, F_d) \in L^2_{\text{pot}}(\Omega; \mathbb{R}^d)$ satisfies the distributional equality, for each $i, j \in \{1, \dots, d\}$,

$$D_i F_j = D_j F_i.$$

In other words, every $F \in L^2_{\text{pot}}(\Omega; \mathbb{R}^d)$ is curl-free.

The following general fact about potential vector fields will be used in the construction of σ and to prove the sublinearity for the corrector $(\phi, \psi, \sigma, \zeta)$. It will be shown that, with respect to the sub-sigma-algebra of subsets that are invariant with respect to spatial translations of the coefficient fields, the conditional expectation of a potential vector field vanishes as a random variable.

Lemma 6. *For every $F = (F_1, \dots, F_d) \in L^2_{pot}(\Omega; \mathbb{R}^d)$, for each $i \in \{1, \dots, d\}$,*

$$\langle F_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle = 0 \ \text{ in } \ L^2(\Omega), \tag{31}$$

where $\langle \cdot \mid \mathcal{F}_{\mathbb{R}^d} \rangle$ denotes the conditional expectation with respect to the sub-sigma-algebra $\mathcal{F}_{\mathbb{R}^d} \subset \mathcal{F}$ of subsets which are invariant with respect to spatial translations of the coefficient field.

In particular, for every $F = (F_1, \dots, F_d) \in L^2_{pot}(\Omega; \mathbb{R}^d)$, for every $i \in \{1, \dots, d\}$,

$$\langle F_i \rangle = 0.$$

The proof of Lemma 6 now follows. Let $F = (F_1, \dots, F_d) \in L^2_{\mathrm{pot}}(\Omega; \mathbb{R}^d)$ be arbitrary. Owing to the definition of the conditional expectation, it is sufficient to show that, for every $\mathcal{F}_{\mathbb{R}^d}$-measurable function $g \in L^2(\Omega)$, for each $i \in \{1, \dots, d\}$,

$$\langle F_i g \rangle = 0. \tag{32}$$

To prove (32), owing to definition (30) there exists a sequence of functions $\{\varphi_n\}_{n=1}^\infty \subset \mathcal{H}^1$ such that, as $n \to \infty$,

$$D\varphi_n \rightharpoonup F \ \text{ weakly in } \ L^2(\Omega; \mathbb{R}^d). \tag{33}$$

The weak convergence (33) implies that, for each $i \in \{1, \dots, d\}$,

$$\lim_{n\to\infty} \langle (D_i \varphi_n)\, g \rangle = \langle F_i g \rangle . \tag{34}$$

Then, for each $n \geq 1$, since $\varphi_n \in \mathcal{H}^1$, it follows that, since spatial translations of the coefficient field preserve the measure, see (2), for each $i \in \{1, \dots, d\}$,

$$\langle (D_i \varphi_n)\, g \rangle = \lim_{h\to 0} \frac{1}{h} \langle (\varphi_n(a(\cdot + he_i, \cdot)) - \varphi_n(a))\, g \rangle = \lim_{h\to 0} \frac{1}{h} \langle \varphi_n \left(g(a(\cdot - he_i, \cdot)) - g(a) \right) \rangle = 0, \tag{35}$$

where the final equality follows from the fact that $g \in L^2(\Omega)$ and the fact that g is invariant with respect to spatial shifts of the coefficient field as an $\mathcal{F}_{\mathbb{R}^d}$-measurable function. In combination, (34) and (35) imply (32). Since the $\mathcal{F}_{\mathbb{R}^d}$-measurable $g \in L^2(\Omega)$ and $F \in L^2_{\mathrm{pot}}(\Omega; \mathbb{R}^d)$ were arbitrary, this completes the proof of (31).

The final statement is then immediate since, for every $F \in L^2_{\mathrm{pot}}(\Omega; \mathbb{R}^d)$ and $i \in \{1, \dots, d\}$, the conditional expectation satisfies

$$\langle F_i \rangle = \langle\langle\ F_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle\rangle = 0,$$

where the final equality follows from (31) and completes the proof of Lemma 6.

3.1. **The construction of ϕ.** For the construction of the corrector ϕ, it is sufficient to construct, for each $k \in \{1, \dots, d\}$, a stationary gradient $D\phi_k \in L^2_{\mathrm{pot}}(\Omega; \mathbb{R}^d)$ and a stationary time-derivative $D_0\phi_k \in \mathcal{H}^{-1}$, where $\mathcal{H}^{-1}$ denotes the dual-space of $\mathcal{H}^1$, satisfying

$$\langle D_0\phi_k f \rangle + \langle Df \cdot a(D\phi_k + e_k) \rangle = 0 \ \text{ for every } \ f \in \mathcal{H}^1. \tag{36}$$

The corrector ϕ will then be defined on $\mathbb{R}^{d+1}$, for $\langle\cdot\rangle$-a.e. a, by integration.

The first step is to introduce an approximation of (36) which is coercive with respect to the $\mathcal{H}^1$-norm. The Riesz representation theorem and the uniform ellipticity of the ensemble (4) guarantee that, for each $k \in \{1, \dots, d\}$ and $\beta \in (0, 1)$, there exists a unique element $\phi_k^\beta \in \mathcal{H}^1$ satisfying

$$\beta \left\langle \phi_k^\beta f \right\rangle + \beta \left\langle D_0\phi_k^\beta D_0 f \right\rangle + \left\langle D_0\phi_k^\beta f \right\rangle + \left\langle Df \cdot a(D\phi_k^\beta + e_k) \right\rangle = 0 \ \text{ for every } \ f \in \mathcal{H}^1. \tag{37}$$

Therefore, for each $k \in \{1, \dots, d\}$ and $\beta \in (0, 1)$, after testing (37) with ϕ_k^β, the uniform ellipticity of the ensemble and Hölder's inequality yield the estimate

$$\left\langle \left| D\phi_k^\beta \right|^2 \right\rangle + \beta \left\langle \left| \phi_k^\beta \right|^2 + \left| D_0\phi_k^\beta \right|^2 \right\rangle \lesssim 1, \tag{38}$$

where the fact that

$$\left\langle D_0\phi_k^\beta \phi_k^\beta \right\rangle = \frac{1}{2} \left\langle D_0 (\phi_k^\beta)^2 \right\rangle = \lim_{h\to 0} \frac{1}{2h} \left\langle (\phi_k^\beta)^2(a(\cdot, \cdot + h)) - (\phi_k^\beta)^2(a) \right\rangle = 0 \tag{39}$$

is also used to obtain (38), and follows from the fact that shifts of the coefficient field in time and space preserve the underlying measure of the ensemble, see (2), and $\phi_k^\beta \in L^2(\Omega)$.

Then, for each $k \in \{1, \dots, d\}$ and $\beta \in (0, 1)$, equation (37), Hölder's inequality and (38) imply that, for each $f \in \mathcal{H}^1$,

$$\left| \left\langle D_0\phi_k^\beta f \right\rangle \right| \lesssim \|f\|_{\mathcal{H}^1} \ \text{ and, therefore, } \ \|D_0\phi_k^\beta\|_{\mathcal{H}^{-1}} \lesssim 1. \tag{40}$$

Hence, for each $k \in \{1,\dots,d\}$, the definition of the potential space (30) with estimates (38) and (40) imply that there exist $\Psi_k \in L^2_{\rm pot}(\Omega;\mathbb{R}^d)$ and $\xi_k \in \mathcal{H}^{-1}$ such that, after passing to a subsequence $\{\beta_j \to 0\}_{j=1}^{\infty}$, as $j \to \infty$,

$$D\phi_k^{\beta_j} \rightharpoonup \Psi_k \text{ weakly in } L^2(\Omega;\mathbb{R}^d) \text{ and } D_0\phi_k^{\beta_j} \rightharpoonup \xi_k \text{ weakly in } \mathcal{H}^{-1}. \tag{41}$$

The convergence (41) combined with equation (37) and estimate (38) prove that, for each $k \in \{1,\dots,d\}$,

$$\langle \xi_k f\rangle + \langle Df \cdot a(\Psi_k + e_k)\rangle = 0 \text{ for every } f \in \mathcal{H}^1.$$

Finally, for each $k \in \{1,\dots,d\}$, as weak limits of functions $\phi_k^{\beta} \in \mathcal{H}^1$, the pair (Ψ_k, ξ_k) are curl-free in the sense that, for each $i \in \{1,\dots,d\}$, distributionally

$$D_0\Psi_{ki} = D_i\xi_k.$$

Therefore, for each $k \in \{1,\dots,d\}$, the argument is completed by defining $D\phi_k := \Psi_k$ and $D_0\phi_k := \xi_k$.

3.2. **The construction of ψ.** For each $k \in \{1,\dots,d\}$, let $D\phi_k \in L^2_{\rm pot}(\Omega;\mathbb{R}^d)$ denote the stationary gradient corresponding to ϕ_k, which was constructed in the previous step. Then, for each $k \in \{1,\dots,d\}$, define the lift of the flux q_k to the probability space according to the rule

$$Q_k := a(D\phi_k + e_k). \tag{42}$$

The existence of ψ follows from the following general fact.

Lemma 7. *For every $F \in L^2(\Omega;\mathbb{R}^d)$, there exists $\Psi \in L^2_{pot}(\Omega;\mathbb{R}^d)$ satisfying*

$$\langle \Psi \cdot Df\rangle = \langle F \cdot Df\rangle \quad \textit{for every } f \in \mathcal{H}^1. \tag{43}$$

The existence of ψ follows from Lemma 7 in the following way. For each $k \in \{1,\dots,d\}$, choose $F = Q_k$ and define $D\psi_k := \Psi$, which defines ψ_k, for $\langle\cdot\rangle$-a.e. a, as a function on $\mathbb{R}^d$ via integration. Then, for $\langle\cdot\rangle$-a.e. a, for each $k \in \{1,\dots,d\}$, the function ψ_k is extended to $\mathbb{R}^{d+1}$ as a stationary function in time.

In order to prove Lemma 7, the Riesz representation theorem asserts that, for each $\beta \in (0,1)$, there exists a unique $\psi^\beta \in \mathcal{H}^1$ satisfying

$$\beta\left\langle \psi^\beta f\right\rangle + \beta\left\langle D_0\psi^\beta D_0 f\right\rangle + \left\langle D\psi^\beta \cdot Df\right\rangle = \langle F \cdot Df\rangle \text{ for every } f \in \mathcal{H}^1. \tag{44}$$

For each $\beta \in (0,1)$, after testing (44) with ψ^β, Hölder's inequality and Young's inequality yield the estimate

$$\left\langle \left|D\psi^\beta\right|^2\right\rangle + \beta\left\langle \left|\psi^\beta\right|^2 + \left|D_0\psi^\beta\right|^2\right\rangle \lesssim 1. \tag{45}$$

Therefore, the definition of the potential space (30) and estimate (45) imply that there exists $\Psi \in L^2_{\rm pot}(\Omega;\mathbb{R}^d)$ such that, after passing to a subsequence $\{\beta_j \to 0\}_{j=1}^{\infty}$, as $j \to \infty$,

$$D\psi^{\beta_j} \rightharpoonup \Psi \text{ weakly in } L^2(\Omega;\mathbb{R}^d). \tag{46}$$

In combination with equation (44), estimates (45) and the convergence (46) imply that Ψ satisfies (43), which completes the proof of Lemma 7.

3.3. **The construction of σ.** For each $k \in \{1,\dots,d\}$, let $D\psi_k \in L^2_{\rm pot}(\Omega;\mathbb{R}^d)$ denote the stationary gradient corresponding to ψ_k constructed in the previous step and let Q_k denote the lift of the flux from (42). Lemma 7 applies directly to this situation, and proves that, for each $i,j,k \in \{1,\dots,d\}$, there exists $\Sigma_{ijk} \in L^2_{\rm pot}(\Omega;\mathbb{R}^d)$ satisfying

$$\begin{aligned}\langle \Sigma_{ijk} \cdot Df\rangle &= \langle (Q_i - D\psi_i)_k D_j f\rangle - \langle (Q_i - D\psi_i)_j D_k f\rangle \\ &= \langle ((Q_i - D\psi_i)_k e_j - (Q_i - D\psi_i)_j e_k) \cdot Df\rangle \quad \text{for every } f \in \mathcal{H}^1.\end{aligned} \tag{47}$$

Then, for each $i,j,k \in \{1,\dots,d\}$, the definition $D\sigma_{ijk} := \Sigma_{ijk}$ defines σ_{ijk}, for $\langle\cdot\rangle$-a.e. a, on $\mathbb{R}^d$ via integration. These functions are then extended to $\mathbb{R}^{d+1}$, for each $i,j,k \in \{1,\dots,d\}$, for $\langle\cdot\rangle$-a.e. a, as stationary functions in time.

Since it is clear from the proof of existence that, for each $i,j,k \in \{1,\dots,d\}$, the gradients can be constructed to satisfy

$$\Sigma_{ijk} = -\Sigma_{ikj},$$

after integrating it follows that, for each $i,j,k \in \{1,\dots,d\}$,

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

Or, perhaps more simply, for each $i \in \{1,\dots,d\}$, one may first construct σ_{ijk} for every $j > k \in \{1,\dots,d\}$ and then simply define $\sigma_{ikj} := -\sigma_{ijk}$.

It remains to prove that, for each $i \in \{1,\dots,d\}$,

$$D \cdot \sigma_i = (Q_i - D\psi_i) - \langle Q_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle , \tag{48}$$

where, for each $i,j \in \{1,\dots,d\}$,

$$(D \cdot \sigma_i)_j = \sum_{k=1}^{d} D_k \sigma_{ijk}. \tag{49}$$

To simplify the notation in what follows, define the vector, for each $i \in \{1,\dots,d\}$,

$$\Psi_i := (Q_i - D\psi_i), \tag{50}$$

where the construction of ψ_i guarantees that the vector Ψ_i is divergence-free in the sense of the distributional equality

$$D_l \Psi_{il} = 0. \tag{51}$$

Equation (48) now follows from the following distributional equalities. For each $i,j \in \{1,\dots,d\}$, thanks to equation (47), in the sense of distributions

$$D_l(D_l(D \cdot \sigma_i)_j) = D_l(D_l(D_k \sigma_{ijk}) = D_k(D_l(D_l \sigma_{ijk})) = D_k(D_k \Psi_{ij} - D_j \Psi_{ik}).$$

Therefore, for each $i,j \in \{1,\dots,d\}$, in view of (51) and after relabeling the final integral,

$$D_l(D_l(D \cdot \sigma_i)_j) = D_k(D_k \Psi_{ij} - D_j \Psi_{ik}) = D_k(D_k \Psi_{ij}) - D_j(D_k \Psi_{ik}) = D_l(D_l \Psi_{ij}).$$

Hence, in the sense of distributions, for each $i,j \in \{1,\dots,d\}$,

$$D_l \left(D_l \left((D \cdot \sigma_i)_j - \Psi_{ij} \right) \right) = 0. \tag{52}$$

Equation (52) implies that, for each $i,j \in \{1,\dots,d\}$, the difference $((D \cdot \sigma_i)_j - \Psi_{ij})$ is invariant with respect to spatial translations of the coefficient field. That is, for each $i,j \in \{1,\dots,d\}$,

$$((D \cdot \sigma_i)_j - \Psi_{ij}) = \langle ((D \cdot \sigma_i)_j - \Psi_{ij}) \mid \mathcal{F}_{\mathbb{R}^d} \rangle . \tag{53}$$

The fact that (53) implies (48) follows from Lemma 6. Indeed, for each $i \in \{1,\dots,d\}$,

$$D\psi_i \in L^2_{\mathrm{pot}}(\Omega;\mathbb{R}^d),$$

and by a straightforward repetition of the arguments leading to Lemma 6, for each $i \in \{1,\dots,d\}$,

$$\langle\, (D \cdot \sigma_i) \mid \mathcal{F}_{\mathbb{R}^d} \,\rangle = 0 \ \text{ in } \ L^2(\Omega;\mathbb{R}^d).$$

Therefore, for each $i,j \in \{1,\dots,d\}$,

$$((D \cdot \sigma_i)_j - \Psi_{ij}) = \langle ((D \cdot \sigma_i)_j - \Psi_{ij}) \mid \mathcal{F}_{\mathbb{R}^d} \rangle = - \langle Q_{ij} \mid \mathcal{F}_{\mathbb{R}^d} \rangle ,$$

which is (48). This completes the argument.

3.4. **The construction of ζ.** The construction of ζ is explicit. Namely, for each $i \in \{1,\dots,d\}$, for the lift of the flux Q_i defined in (42), define the stationary derivative in time according to the rule

$$D_0 \zeta_i = \langle Q_i \mid \mathcal{F}_{\mathbb{R}^d} \rangle - \langle Q_i \rangle .$$

For each $i \in \{1,\dots,d\}$, the function ζ_i is defined on $\mathbb{R}$, for $\langle\cdot\rangle$-a.e a, by integration in time and extended to $\mathbb{R}^{d+1}$ as a spatially constant function.

3.5. **The boundedness and uniform ellipticity of $a_{\rm hom}$.** For the reader's convenience, the argument of [17, Lemma 2] is repeated here. For each $\xi \in \mathbb{R}^d$, the linearity and (9) assert that the homogenized coefficients are defined according to the rule

$$a_{\rm hom}\xi := \langle a(\nabla\phi_\xi + \xi)\rangle .$$

It is first shown that, for each $\xi \in \mathbb{R}^d$,

$$|a_{\rm hom}\xi| \le \frac{1}{\lambda}|\xi| . \tag{54}$$

For each $\xi \in \mathbb{R}^d$, since ϕ_ξ satisfies (6), the uniform ellipticity of the ensemble (4) and Jensen's inequality imply that

$$|a_{\rm hom}\xi|^2 = |\langle a(\nabla\phi_\xi + \xi)\rangle|^2 \le \left\langle |a(\nabla\phi_\xi + \xi)|^2 \right\rangle \le \left\langle |(\nabla\phi_\xi + \xi)|^2 \right\rangle \le \frac{1}{\lambda}\langle (\nabla\phi_\xi + \xi)\cdot a(\nabla\phi_\xi + \xi)\rangle .$$

Then, using the corrector equation (6), (39) and the Cauchy-Schwarz inequality,

$$|a_{\rm hom}\xi|^2 \le \frac{1}{\lambda}\langle (\nabla\phi_\xi + \xi)\cdot a(\nabla\phi_\xi + \xi)\rangle = \frac{1}{\lambda}\xi\cdot\langle a(\nabla\phi_\xi + \xi)\rangle \le \frac{1}{\lambda}|\xi|\,|a_{\rm hom}\xi| .$$

Dividing by $|a_{\rm hom}\xi|$ yields (54) and completes the proof.

It remains only to prove that, for each $\xi \in \mathbb{R}^d$,

$$\lambda|\xi|^2 \le \xi\cdot a_{\rm hom}\xi. \tag{55}$$

This follows from the convexity of the map

$$(X, v) \in \mathcal{S}(d)_{>0}\times\mathbb{R}^d \to v\cdot X^{-1}v, \tag{56}$$

where $\mathcal{S}(d)_{>0}$ denotes the space of positive, $d\times d$ symmetric matrices. Indeed, if $X \in \mathcal{S}(d)_{>0}$ and $v \in \mathbb{R}^d$,

$$\frac{1}{2}\left(v\cdot X^{-1}v\right) = \sup_{w\in\mathbb{R}^d}\left\{w\cdot v - \frac{1}{2}w\cdot Xw\right\},$$

is a supremum over linear functions in (X, v), and therefore the map (56) is convex. Hence, for each $\xi \in \mathbb{R}^d$, using the corrector equation (6) and Jensen's inequality,

$$\begin{aligned}\xi\cdot a_{\rm hom}\xi =& \langle(\nabla\varphi_\xi + \xi)\cdot a(\nabla\varphi_\xi + \xi)\rangle = \langle(\nabla\phi_\xi + \xi)\cdot a_{\rm sym}(\nabla\phi_\xi + \xi)\rangle\\ =& \left\langle(\nabla\phi_\xi + \xi)\cdot(a_{\rm sym}^{-1})^{-1}(\nabla\phi_\xi + \xi)\right\rangle \ge \langle\nabla\phi_\xi + \xi\rangle\cdot\left\langle(a_{\rm sym}^{-1})\right\rangle^{-1}\langle\nabla\phi_\xi + \xi\rangle\\ \ge& \lambda|\xi|^2 ,\end{aligned}$$

where $a_{\rm sym}$ denotes the symmetric part of a, and where the final inequality is obtained using the boundedness of the ensemble (4) and the vanishing expectation of the gradient from Lemma 6. This completes the proof of (55).

4. The proof of Proposition 3

4.1. **The sublinearity of σ and ψ.** The sublinearity of σ and ψ will follow from the following general fact. Recall that, for a function $\varphi : \mathbb{R}^{d+1} \to \mathbb{R}$, for each $R > 0$ and $t \in \mathbb{R}$,

$$(\varphi)_R = \fint_{\mathcal{C}_R}\varphi \ \text{ and } \ (\varphi)_{t,R} = \fint_{B_R}\varphi(\cdot, t).$$

Lemma 8. *Suppose that φ is a scalar random field on $\mathbb{R}^{d+1}$ which is stationary in time and has a stationary, finite energy gradient $\nabla\varphi$ in the potential space $L^2_{pot}(\Omega;\mathbb{R}^d)$. That is, assume that*

$$\left\langle|\nabla\varphi|^2\right\rangle < \infty \ \text{ with } \ \nabla\varphi \in L^2_{pot}(\Omega;\mathbb{R}^d). \tag{57}$$

Then, for $\langle\cdot\rangle$-a.e. a, the normalized large-scale L^2-averages of φ are strictly sublinear in the sense that

$$\lim_{R\to\infty}\frac{1}{R^2}\fint_{\mathcal{C}_R}|\varphi - (\varphi)_{t,R}|^2 = 0. \tag{58}$$

To prove Lemma 8, since φ is stationary in time, it will first be shown that, for $\langle\cdot\rangle$-a.e. a, the normalized large-scale L^2-averages of $\varphi(\cdot,0)$ on large balls are sublinear. Define, for each $\epsilon>0$,

$$\varphi^\epsilon(\cdot) := \epsilon\varphi(\frac{\cdot}{\epsilon},0).$$

Then, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to\infty}\frac{1}{R^2}\fint_{B_R}|\varphi(\cdot,0)-(\varphi)_{0,R}|^2 = \lim_{\epsilon\to0}\fint_{B_1}|\varphi^\epsilon-(\varphi^\epsilon)|^2 = 0, \tag{59}$$

where the first equality is immediate by scaling if, for each $\epsilon>0$,

$$(\varphi^\epsilon) := \fint_{B_1}\varphi^\epsilon.$$

To obtain (59), the Poincaré inequality in space, together with the ergodic theorem, imply that, for $\langle\cdot\rangle$-a.e. a,

$$\limsup_{\epsilon\to0}\fint_{B_1}|\varphi^\epsilon-(\varphi^\epsilon)|^2 \le \limsup_{\epsilon\to0}\fint_{B_1}|\nabla\varphi^\epsilon|^2 = \left\langle |\nabla\varphi|^2 \mid \mathcal{F}_{\mathbb{R}^d}\right\rangle < \infty,$$

where the final term is finite, for $\langle\cdot\rangle$-a.e. a, thanks to (57). Therefore, for $\langle\cdot\rangle$-a.e. a, the sequence $\{(\varphi^\epsilon-(\varphi^\epsilon))\}_{\epsilon\in(0,1)}$ is bounded in $H^1(B_1)$, and hence compact in $L^2(B_1)$ by the Rellich-Kondrachov embedding theorem. Since the ergodic theorem, see Krengel [19, Theorem 2.3], implies that, for $\langle\cdot\rangle$-a.e. a, as $\epsilon\to0$, the gradient

$$\nabla\varphi^\epsilon \rightharpoonup \langle\nabla\varphi\mid\mathcal{F}_{\mathbb{R}^d}\rangle = 0 \text{ weakly in } L^2(B_1;\mathbb{R}^d),$$

where the vanishing of the conditional expectation follows from Lemma 6 and (57), the compact embedding implies that, for $\langle\cdot\rangle$-a.e a,

$$(\varphi^\epsilon-(\varphi^\epsilon))\to0 \text{ strongly in } L^2(B_1),$$

which proves (59).

The sublinearity of the normalized large-scale averages at time zero will now be upgraded to the sublinearity of the normalized large-scale averages on parabolic cylinders using Egorov's theorem. Precisely, using (59), it will be shown that, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to\infty}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac12} = 0. \tag{60}$$

To prove (60), for $\epsilon\in(0,1)$ use Egorov's theorem and (59) to find a measurable subset $A_\epsilon\subset\Omega$ and $R_\epsilon>0$ such that, for every $a\in A_\epsilon$ and $R>R_\epsilon$,

$$\frac{1}{R}\left(\fint_{B_R}|\varphi(\cdot,0)-(\varphi)_{0,R}|^2\right)^{\frac12} < \epsilon \text{ with } \langle\chi_{A_\epsilon}\rangle > 1-\epsilon, \tag{61}$$

where $\chi_{A_\epsilon}\in L^\infty(\Omega)$ denotes the indicator function of A_ϵ. The large-scale averages on cylinders appearing in (60) will be decomposed according to the event A_ϵ.

For each $R>0$ and $\epsilon\in(0,1)$, use the triangle inequality to form the decomposition

$$\begin{aligned}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac12} \le{}& \frac{1}{R}\left(\fint_{-R^2}^0\chi_{A_\epsilon}(a(\cdot,\cdot+t))\fint_{B_R}|(\varphi-(\varphi)_{t,R})|^2\right)^{\frac12}\\ &+\frac{1}{R}\left(\fint_{-R^2}^0(1-\chi_{A_\epsilon}(a(\cdot,\cdot+t)))\fint_{B_R}|(\varphi-(\varphi)_{t,R})|^2\right)^{\frac12}.\end{aligned} \tag{62}$$

For the second term of (62), for each $R > 0$ and $\epsilon \in (0,1)$, Hölder's inequality and the Sobolev embedding theorem imply that

$$\frac{1}{R}\left(\fint_{-R^2}^{0}(1-\chi_{A_\epsilon}(a(\cdot,\cdot+t)))\fint_{B_R}|(\varphi-(\varphi)_{t,R})|^2\right)^{\frac{1}{2}}\lesssim$$
$$\frac{1}{R}\left(\fint_{\mathcal{C}_R}(1-\chi_{A_\epsilon}(a(\cdot,\cdot+t)))\right)^{\frac{1}{d}}\left(\fint_{\mathcal{C}_R}|(\varphi-(\varphi)_{t,R})|^{\frac{2d}{d-2}}\right)^{\frac{d-2}{2d}}\lesssim$$
$$\left(\fint_{\mathcal{C}_R}(1-\chi_{A_\epsilon}(a(\cdot,\cdot+t)))\right)^{\frac{1}{d}}\left(\fint_{\mathcal{C}_R}|\nabla\varphi|^2\right)^{\frac{1}{2}}, \quad (63)$$

where the argument is written only for the case $d \geq 3$, since the modifications necessary to handle the cases $d = 1$ and $d = 2$ are straightforward and follow from the Sobolev embedding theorem.

After combining (62) and (63), it is then immediate from the ergodic theorem and the definition of A_ϵ from (61) that, for each $\epsilon > 0$, for $\langle\cdot\rangle$-a.e. a,

$$\limsup_{R\to\infty}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac{1}{2}}\leq\langle(1-\chi_{A_\epsilon})\mid\mathcal{F}_t\rangle^{\frac{1}{d}}\left\langle|\nabla\varphi|^2\right\rangle^{\frac{1}{2}}+\epsilon, \quad (64)$$

where $\langle\cdot\mid\mathcal{F}_t\rangle$ denotes the conditional expectation with respect to the sub-sigma-algebra $\mathcal{F}_t\subset\mathcal{F}$ of subsets of Ω which are invariant with respect to translations of the coefficient fields in time.

Since, as $\epsilon\to 0$, the choice (61) implies that

$$\langle(1-\chi_{A_\epsilon})\mid\mathcal{F}_t\rangle\to 0 \text{ strongly in } L^1(\Omega),$$

using (57), after choosing a countable sequence $\{\epsilon_n\to 0\}_{n=1}^{\infty}$, it follows from (62) and (64) that, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to\infty}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac{1}{2}}=0, \quad (65)$$

which proves (60). This completes the proof of Lemma 8, and thereby proves the sublinearity of σ and ψ thanks to Lemma 2.

4.2. **The sublinearity of ϕ.** The sublinearity of ϕ will follow from the following general fact.

Lemma 9. *Suppose that φ is a scalar random field on $\mathbb{R}^{d+1}$ which has a stationary, finite energy gradient $\nabla\varphi$ in the potential space $L^2_{pot}(\Omega;\mathbb{R}^d)$. That is, assume that*

$$\left\langle|\nabla\varphi|^2\right\rangle<\infty \text{ with } \nabla\varphi\in L^2_{pot}(\Omega;\mathbb{R}^d). \quad (66)$$

Furthermore, assume that, for $\langle\cdot\rangle$-a.e. a, the field φ satisfies

$$\varphi_t=\nabla\cdot F \text{ distributionally in } \mathbb{R}^{d+1}, \quad (67)$$

for the stationary extension of a finite energy field $F\in L^2(\Omega;\mathbb{R}^d)$. Then, for $\langle\cdot\rangle$-a.e. a, the normalized large-scale L^2-averages of φ are strictly sublinear in the sense that

$$\lim_{R\to\infty}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_R|^2\right)^{\frac{1}{2}}=0. \quad (68)$$

To prove Lemma 9, for each $R > 0$, use the triangle inequality to obtain

$$\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_R|^2\right)^{\frac{1}{2}}\leq\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac{1}{2}}+\frac{1}{R}\left(\fint_{\mathcal{C}_R}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac{1}{2}}. \quad (69)$$

Since φ has a stationary spatial gradient, for each $R > 0$ and $t\in\mathbb{R}$, the scalar random field

$$(x,t)\to(\varphi-(\varphi)_{t,R})(x,t)=\fint_{B_R}\varphi(x,t)-\varphi(y,t)\,\mathrm{d}y$$

is stationary in time, with a stationary, finite-energy gradient in the potential space. Therefore, Lemma 8 applies to this random field, and asserts that, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to\infty}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|\varphi-(\varphi)_{t,R}|^2\right)^{\frac{1}{2}}=0. \tag{70}$$

It remains to prove that, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to 0}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac{1}{2}}=\lim_{R\to 0}\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac{1}{2}}=0.$$

Let $\rho\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ be a smooth, symmetric convolution kernel satisfying $\operatorname{supp}(\rho)\subset B_1$ and, for each $\epsilon>0$, define the rescaling $\rho^\epsilon(\cdot)=\epsilon^{-d}\rho(\frac{\cdot}{\epsilon})$. Then, for each $\epsilon>0$, define the spatial convolution, for each $x\in\mathbb{R}^d$ and $t\in\mathbb{R}$,

$$\varphi^\epsilon(x,t):=\int_{\mathbb{R}^d}\rho^\epsilon(y-x)\varphi(y,t)\,\mathrm{d}y.$$

The introduction of the convolution kernel provides a test function which will be used to apply the equation (67) satisfied by φ.

First, for each $R>0$ and $\epsilon\in(0,R)$, it follows from the support of the convolution kernel, Fubini's theorem and Jensen's inequality that

$$\begin{aligned}
|(\varphi)_R-(\varphi^\epsilon)_R|^2=&\left|\fint_{\mathcal{C}_R}\int_{\mathbb{R}^d}\rho^\epsilon(x)\left(\varphi(x+y,t)-\varphi(y,t)\right)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\right|^2\\
\leq&\epsilon^2\left|\int_0^1\int_{\mathbb{R}^d}\rho^\epsilon(x)\fint_{\mathcal{C}_R}|\nabla\varphi(y+rx,t)|\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}r\right|^2\\
\leq&\epsilon^2\int_0^1\int_{\mathbb{R}^d}\rho^\epsilon(x)\fint_{\mathcal{C}_R}|\nabla\varphi(y+rx,t)|^2\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}r\\
\lesssim&\epsilon^2\fint_{\mathcal{C}_{2R}}|\nabla\varphi|^2.
\end{aligned}\tag{71}$$

Similarly, the identical argument yields, for each $R>0$, $t\in\mathbb{R}$ and $\epsilon\in(0,R)$,

$$|(\varphi)_{t,R}-(\varphi^\epsilon)_{t,R}|^2\lesssim\epsilon^2\fint_{B_{2R}}|\nabla\varphi(\cdot,t)|^2. \tag{72}$$

Therefore, for each $R>0$ and $\epsilon\in(0,R)$, the triangle inequality, (71) and (72) imply that, after adding and subtracting $((\varphi^\epsilon)_{t,R}-(\varphi^\epsilon)_R)$,

$$\frac{1}{R}\left(\fint_{-R^2}^0|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac{1}{2}}\lesssim\frac{1}{R}\left(\fint_{-R^2}^0|(\varphi^\epsilon)_{t,R}-(\varphi^\epsilon)_R|^2\right)^{\frac{1}{2}}+\frac{\epsilon}{R}\left(\fint_{\mathcal{C}_{2R}}|\nabla\varphi|^2\right)^{\frac{1}{2}}. \tag{73}$$

For the first term on the right hand side of (73), for each $R\geq 1$ and $\epsilon\in(0,R)$, the equation (67) satisfied by φ and the Poincaré inequality in time yields

$$\begin{aligned}
\frac{1}{R}\left(\fint_{-R^2}^0|(\varphi^\epsilon)_{t,R}-(\varphi^\epsilon)_R|^2\right)^{\frac{1}{2}}\lesssim&\left(\fint_{-R^2}^0|\partial_t(\varphi^\epsilon)_{t,R}|^2\right)^{\frac{1}{2}}\\
\lesssim&\left(\fint_{-R^2}^0\left|\fint_{B_R}\int_{\mathbb{R}^d}\nabla\rho^\epsilon(x-y)\cdot F(x,t)\,\mathrm{d}x\,\mathrm{d}y\right|^2\,\mathrm{d}t\right)^{\frac{1}{2}}\\
\lesssim&\left(\fint_{-R^2}^0\left|\int_{\mathbb{R}^d}|\nabla\rho^\epsilon(x)|\fint_{B_R(x)}|F(y,t)|\,\mathrm{d}y\,\mathrm{d}x\right|^2\,\mathrm{d}t\right)^{\frac{1}{2}}\\
\lesssim&\left(\fint_{-R^2}^0\left|\left(\int_{\mathbb{R}^d}|\nabla\rho^\epsilon(x)|\,\mathrm{d}x\right)\left(\fint_{B_{2R}}|F(y,t)|\,\mathrm{d}y\right)\right|^2\,\mathrm{d}t\right)^{\frac{1}{2}}.
\end{aligned}\tag{74}$$

Then, continuing with (74), the definition of the convolution kernel and Jensen's inequality yield

$$\begin{aligned}\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi^\epsilon)_{t,R}-(\varphi^\epsilon)_R|^2\right)^{\frac12}\lesssim&\frac{1}{\epsilon}\left(\fint_{-R^2}^{0}\left(\fint_{B_{2R}}|F(y,t)|\ \mathrm{d}y\right)^2\mathrm{d}t\right)^{\frac12}\\ \lesssim&\frac{1}{\epsilon}\left(\fint_{\mathcal{C}_{2R}}|F|^2\right)^{\frac12}.\end{aligned}\tag{75}$$

Therefore, combining (73) with (75), for each $R>0$ and $\epsilon\in(0,R)$,

$$\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac12}\lesssim\frac{1}{\epsilon}\left(\fint_{\mathcal{C}_{2R}}|F|^2\right)^{\frac12}+\frac{\epsilon}{R}\left(\fint_{\mathcal{C}_{2R}}|\nabla\varphi|^2\right)^{\frac12}.$$

Let $\delta\in(0,1)$ be arbitrary and, for each $R>0$, fix $\epsilon(R):=\delta R$. For this choice, for each $R>0$ and $\delta\in(0,1)$,

$$\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac12}\lesssim\frac{1}{\delta R}\left(\fint_{\mathcal{C}_{2R}}|F|^2\right)^{\frac12}+\delta\left(\fint_{\mathcal{C}_{2R}}|\nabla\varphi|^2\right)^{\frac12}.\tag{76}$$

Since the ergodic theorem implies that, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{R\to\infty}\left(\fint_{\mathcal{C}_{2R}}|F|^2\right)^{\frac12}=\left\langle|F|^2\right\rangle^{\frac12}\quad\text{and}\quad\lim_{R\to\infty}\left(\fint_{\mathcal{C}_{2R}}|\nabla\varphi|^2\right)^{\frac12}=\left\langle|\nabla\varphi|^2\right\rangle^{\frac12},$$

it follows from (76) that, for $\langle\cdot\rangle$-a.e. a, for every $\delta\in(0,1)$,

$$\limsup_{R\to\infty}\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac12}\lesssim\delta\left\langle|\nabla\varphi|^2\right\rangle^{\frac12}.$$

Therefore, since $\delta\in(0,1)$ is arbitrary,

$$\lim_{R\to0}\frac{1}{R}\left(\fint_{\mathcal{C}_R}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac12}=\lim_{R\to0}\frac{1}{R}\left(\fint_{-R^2}^{0}|(\varphi)_{t,R}-(\varphi)_R|^2\right)^{\frac12}=0.\tag{77}$$

In combination, (69), (70) and (77) combine to prove (68), and thereby complete the proof of Lemma 9. The sublinearity of the corrector ϕ, for $\langle\cdot\rangle$-a.e. a, is then immediate from Lemma 2.

4.3. **The sublinearity of ζ.** The fact that, for $\langle\cdot\rangle$-a.e. a, for each $i,j\in\{1,\dots,d\}$,

$$\lim_{R\to0}\frac{1}{R^2}\left(\fint_{\mathcal{C}_R}|\zeta_{ij}-(\zeta_{ij})_R|^2\right)^{\frac12}=\lim_{R\to\infty}\frac{1}{R^2}\left(\fint_{-R^2}^{0}|\zeta_{ij}-(\zeta_{ij})_R|^2\right)^{\frac12}=0,$$

follows similarly to (59). That is, for each $i\in\{1,\dots,d\}$, the Poincaré inequality and the ergodic theorem together with the Rellich-Kondrachov embedding theorem imply that the family

$$\left\{\epsilon^2\zeta_{ij}(\frac{\cdot}{\epsilon^2})-\epsilon^2\fint_{-\frac{1}{\epsilon^2}}^{0}\zeta_{ij}\right\}_{\epsilon\in(0,1)}$$

is compact in $L^2([0,1])$ and converges weakly to zero, as $\epsilon\to0$, in $H^1([0,1])$. Therefore, for $\langle\cdot\rangle$-a.e. a, for each $i,j\in\{1,\dots,d\}$, as $\epsilon\to0$,

$$\left(\epsilon^2\zeta_{ij}(\frac{\cdot}{\epsilon^2})-\epsilon^2\fint_{-\frac{1}{\epsilon^2}}^{0}\zeta_{ij}\right)\to0\ \text{ in }\ L^2([0,1]).\tag{78}$$

Furthermore, now exploiting the fact that ζ is a one-dimensional function, it follows from the Sobolev embedding theorem and the Arzelà-Ascoli theorem that, for $\langle\cdot\rangle$-a.e. a, for each $i,j\in\{1,\dots,d\}$, the family

$$\left\{\epsilon^2\zeta_{ij}(\frac{\cdot}{\epsilon^2})-\epsilon^2\fint_{-\frac{1}{\epsilon^2}}^{0}\zeta_{ij}\right\}_{\epsilon\in(0,1)}$$

is compact in $C^{0,\frac{1}{2}}([0,1])$ and, by repeating the argument of (59), converges weakly to zero, as $\epsilon \to 0$, in $H^1([0,1])$. Therefore, for $\langle\cdot\rangle$-a.e. a, for each $i,j \in \{1,\dots,d\}$, as $\epsilon \to 0$,

$$\left(\epsilon^2 \zeta_{ij}(\frac{\cdot}{\epsilon^2}) - \epsilon^2 \fint_{-\frac{1}{\epsilon^2}}^0 \zeta_{ij} \right) \to 0 \text{ in } \mathcal{C}^{0,\frac{1}{2}}([0,1]).$$

In particular, for each $i,j \in \{1,\dots,d\}$, as $\epsilon \to 0$,

$$\left| \epsilon^2 \zeta_{ij}(0) - \epsilon^2 \fint_{-\frac{1}{\epsilon^2}}^0 \zeta_{ij} \right| \to 0. \tag{79}$$

Hence, in combination, (78) and (79) prove after rescaling that, for $\langle\cdot\rangle$-a.e. a, for each $i,j \in \{1,\dots,d\}$,

$$\limsup_{R\to 0} \frac{1}{R^2} \left(\fint_{-R^2}^0 |\zeta_{ij} - \zeta_{ij}(0)| \right)^{\frac{1}{2}} \leq \limsup_{R\to\infty} \frac{1}{R^2} \left(\fint_{-R^2}^0 |\zeta_{ij} - (\zeta_{ij})_R| \right)^{\frac{1}{2}} + \limsup_{R\to\infty} \frac{1}{R^2} |(\zeta_{ij})_R - \zeta_{ij}(0)| = 0,$$

which completes the argument since ζ is constant in space.

4.4. **The large-scale averages of q.** It is an immediately consequence of the ergodic theorem and the fact that the flux q is stationary with finite energy that, for $\langle\cdot\rangle$-a.e. a, for each $i \in \{1,\dots,d\}$,

$$\lim_{R\to\infty} \left(\fint_{\mathcal{C}_R} |q_i|^2 \right)^{\frac{1}{2}} = \left\langle |Q_i|^2 \right\rangle^{\frac{1}{2}},$$

which completes the argument, and the proof of Proposition 3.

5. The proof of Proposition 4

The proof of Proposition 4 is split into five steps. The first defines the augmented homogenization error. The second proves that the augmented homogenization error satisfies a parabolic equation. The third recalls some classical estimates governing the interior and boundary regularity of a_{hom}-caloric functions. The fourth uses the equation satisfied by the augmented homogenization error to derive an energy estimate. And, finally, the fifth uses the energy estimate to complete the proof of excess decay.

In what follows, to simplify the notation, observe that it may be assumed without loss of generality that, for each $R > 0$, $t \in \mathbb{R}$ and $i,j,k \in \{1,\dots,d\}$,

$$(\phi_i)_R = (\psi_i)_{t,R} = (\sigma_{ijk})_{t,R} = \zeta_{ij}(0) = 0.$$

Indeed, otherwise in the arguments to follow, at each step replace the components of the corrector, for each $R > 0$, $t \in \mathbb{R}$ and $i,j,k \in \{1,\dots,d\}$, by the normalizations defined by

$$\tilde{\phi}_i := \phi_i(x,t) - (\phi_i)_R, \quad \tilde{\psi}_i := \psi_i - (\psi_i)_{t,R}, \quad \tilde{\sigma}_{ijk} := \sigma_{ijk} - (\sigma_{ijk})_{t,R} \text{ and } \tilde{\zeta}_{ij} := \zeta_{ij} - \zeta_{ij}(0). \tag{80}$$

The argument now begins with the definition of the augmented homogenization error.

5.1. **The augmented homogenization error.** The analysis augmented homogenization error and its corresponding energy estimate will first be obtained on scale $R = 1$. The general results will then follow by scaling. Suppose that u is an a-caloric function in $\mathcal{C}_1$. That is, in the sense of distributions, suppose that u satisfies

$$u_t = \nabla \cdot a \nabla u \text{ in } \mathcal{C}_1. \tag{81}$$

Then, let $\rho \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ be a standard convolution kernel satisfying $\mathrm{supp}(\rho) \subset B_1$. For each $\epsilon \in (0,\frac{1}{4})$, let $\rho^\epsilon(\cdot) := \epsilon^{-d}\rho(\frac{\cdot}{\epsilon})$ and define the spatial convolution

$$u^\epsilon(x,t) = \int_{\mathbb{R}^d} \rho^\epsilon(y-x) u(y,t)\,\mathrm{d}y \text{ on } \mathcal{C}_{1-\epsilon}.$$

It is necessary to observe some useful energy estimates for u and its convolution.

First, it is immediate from (81) and the uniform ellipticity of a from (4) that

$$\|u_t\|_{L^2([-1,0];H^{-1}(B_1))} \lesssim \int_{\mathcal{C}_1} |\nabla u|^2 .$$

Therefore, since the convolution preserves this estimate, for each $\epsilon \in (0, \frac{1}{4})$,

$$\|u_t^\epsilon\|_{L^2([-1,0];H^{-1}(B_{1-\epsilon}))} \lesssim \int_{\mathcal{C}_1} |\nabla u|^2 . \tag{82}$$

It is important to keep these estimates in mind when considering the application of the constant-coefficient regularity estimates (101) and (105) below.

Next, although there is no convolution in time, the spatial convolution nevertheless provides some temporal regularity in the sense that, for each $\epsilon \in (0, \frac{1}{4})$,

$$u_t^\epsilon(x,t) = -\int_{\mathbb{R}^d} \nabla \rho^\epsilon(y-x) \cdot a \nabla u(y,t) \, \mathrm{d}y \ \text{ in } \ \mathcal{C}_{1-\epsilon}.$$

Therefore, the time-derivative has a uniformly bounded energy. That is, for each $\epsilon \in (0, \frac{1}{4})$, the Minkowski integral inequality, Hölder's inequality, the definition of the convolution kernel and the uniform ellipticity of a imply that

$$\begin{aligned}
\left(\int_{\mathcal{C}_{\frac{3}{4}}} |u_t^\epsilon|^2 \right)^{\frac{1}{2}} &\leq \left(\int_{\mathcal{C}_{1-\epsilon}} \left| \int_{\mathbb{R}^d} \nabla \rho^\epsilon(y-x) \cdot a \nabla u(y) \, \mathrm{d}y \right|^2 \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \\
&\lesssim \int_{\mathbb{R}^d} \left(\int_{\mathcal{C}_{1-\epsilon}} |\nabla \rho^\epsilon(y)|^2 \, |\nabla u(y+x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \mathrm{d}y \\
&= \int_{\mathbb{R}^d} |\nabla \rho^\epsilon(y)| \left(\int_{\mathcal{C}_{1-\epsilon}} |\nabla u(y+x,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \\
&\lesssim \frac{1}{\epsilon} \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}} .
\end{aligned}$$

The convolution error is also well controlled by the energy of u. Precisely, for each $\epsilon \in (0, \frac{1}{4})$, it follows from Jensen's inequality and the definition of the convolution kernel that

$$\begin{aligned}
\left(\int_{\mathcal{C}_{\frac{3}{4}}} |u^\epsilon - u|^2 \right)^{\frac{1}{2}} &= \left(\int_{\mathcal{C}_{\frac{3}{4}}} \left| \int_0^1 \int_{\mathbb{R}} \rho^\epsilon(y) \nabla u(x+sy,t) \cdot y \, \mathrm{d}y \, \mathrm{d}s \right|^2 \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}} \\
&\leq \epsilon \left(\int_{\mathbb{R}^d} \rho^\epsilon(y) \int_0^1 \int_{\mathcal{C}_{\frac{3}{4}}} |\nabla u(x+sy,t)|^2 \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \, \mathrm{d}y \right)^{\frac{1}{2}} \\
&\leq \epsilon \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}} .
\end{aligned}$$

Lastly, it is immediate from Jensen's inequality that the convolution preserves the energy in the sense that, for each $\epsilon \in (0, \frac{1}{4})$,

$$\left(\int_{\mathcal{C}_{\frac{3}{4}}} |\nabla u^\epsilon|^2 \right)^{\frac{1}{2}} \leq \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}} .$$

Fubini's theorem therefore implies that, for each $\epsilon \in (0, \frac{1}{4})$, there exists $r_\epsilon \in (\frac{1}{2}, \frac{3}{4})$ such that

$$\left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |u^\epsilon - u|^2 \right)^{\frac{1}{2}} \lesssim \epsilon \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}}, \tag{83}$$

and

$$\left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |u_t^\epsilon|^2 \right)^{\frac{1}{2}} \lesssim \frac{1}{\epsilon} \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}}, \tag{84}$$

and, finally, such that

$$\left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |\nabla u|^2 \right)^{\frac{1}{2}} + \left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |\nabla u^\epsilon|^2 \right)^{\frac{1}{2}} \lesssim \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}}. \tag{85}$$

It will be for this radius that the a_{hom}-caloric extension of u^ϵ is constructed.

Namely, for each $\epsilon \in (0, \frac{1}{4})$, let v^ϵ denote the solution

$$\begin{cases} v_t^\epsilon &= \nabla \cdot a \nabla v^\epsilon \quad \text{in } C_{r_\epsilon} \\ v^\epsilon &= u^\epsilon \qquad\quad \text{on } \partial_p C_{r_\epsilon}. \end{cases} \tag{86}$$

These functions will now come to define the augmented homogenization error after the introduction of a cutoff function.

For each $\epsilon \in (0, \frac{1}{4})$ and $\rho \in (0, \frac{1}{8})$, let $\eta_\rho^\epsilon \in \mathcal{C}_c^\infty(\mathbb{R}^{d+1})$ be smooth cutoff function satisfying $0 \le \eta_\rho^\epsilon \le 1$ with, for each $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$,

$$\eta_\rho^\epsilon(x,t) = \begin{cases} 1 & \text{if } (x,t) \in \overline{\mathcal{C}}_{r_\epsilon - 2\rho} \\ 0 & \text{if } (x,t) \in \big(\mathbb{R}^d \times (-\infty, 0)\big) \setminus \mathcal{C}_{r_\epsilon - \rho}. \end{cases} \tag{87}$$

Furthermore, for each $\epsilon \in (0, \frac{1}{4})$ and $\rho \in (0, \frac{1}{8})$, for each $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$,

$$\big|\nabla \eta_\rho^\epsilon(x,t)\big| + \big|\partial_t \eta_\rho^\epsilon(x,t)\big| \lesssim \frac{1}{\rho} \text{ and } \big|\nabla^2 \eta_\rho^\epsilon(x,t)\big| \lesssim \frac{1}{\rho^2}. \tag{88}$$

Then, for $\rho \in (0, \frac{1}{4})$ and $\epsilon \in (0, \frac{1}{4})$ to be specified later, define the augmented homogenization error w according to the rule

$$w = u - (1 + \eta_\rho^\epsilon \phi_i \partial_i) v^\epsilon. \tag{89}$$

The augmented homogenization error (89) will now be shown to satisfy a useful parabolic equation. The computation is motivated by the analogous computation in [17, Lemma 2], but there are significant differences owing to the parabolic setting and the use of the parabolic extended corrector $(\phi, \psi, \sigma, \zeta)$.

5.2. **The equation satisfied by the augmented homogenization error.** It is now shown that the augmented homogenization error (89) satisfies

$$\begin{cases} w_t - \nabla \cdot a \nabla w &= \nabla \cdot \big((1 - \eta_\rho^\epsilon)(a - a_{\mathrm{hom}}) \nabla v^\epsilon\big) + \nabla \cdot \big((\phi_i a + \psi_i - \sigma_i) \nabla (\eta_\rho^\epsilon \partial_i v^\epsilon)\big) \quad \text{in } \mathcal{C}_{r_\epsilon} \\ &\quad + \partial_t \zeta_i \cdot \nabla (\eta_\rho^\epsilon \partial_i v^\epsilon) - \phi_i (\eta_\rho^\epsilon \partial_i v^\epsilon)_t - \psi_i \Delta (\eta_\rho^\epsilon \partial_i v^\epsilon) \\ w &= u - u^\epsilon \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial_p \mathcal{C}_{r_\epsilon}. \end{cases} \tag{90}$$

Fix $\rho \in (0, \frac{1}{4})$ and $\epsilon \in (0, \frac{1}{4})$ and let w be defined by (89). Since the boundary condition is immediate from the definition, it remains only to compute the equation. First, using definition (89), the gradient is defined by

$$\nabla w = \nabla u - \nabla v^\epsilon - \nabla (\phi_i \eta_\rho^\epsilon \partial_i v^\epsilon).$$

Then, because u satisfies (81),

$$-\nabla \cdot a \nabla w = -u_t + \nabla \cdot a \nabla v^\epsilon + \nabla \cdot a (\nabla \phi_i \eta_\rho^\epsilon \partial_i v^\epsilon + \phi_i \nabla (\eta_\rho^\epsilon \partial_i v^\epsilon)). \tag{91}$$

It is necessary to further analyze the term

$$\nabla \cdot a \nabla v^\epsilon + \nabla \cdot a (\nabla \phi_i \eta_\rho^\epsilon \partial_i v^\epsilon),$$

which, after adding and subtracting the unit vectors $\{e_i\}_{i\in\{1,\dots,d\}}$, satisfies, for the fluxes $\{q_i\}_{i\in\{1,\dots,d\}}$ defined in (10),

$$\nabla\cdot a\nabla v^\epsilon+\nabla\cdot a(\nabla\phi_i\eta_\rho^\epsilon\partial_i v^\epsilon)=\nabla\cdot((1-\eta_\rho^\epsilon)a\nabla v^\epsilon)+\nabla\cdot(q_i\eta_\rho^\epsilon\partial_i v^\epsilon).$$

Then, after adding and subtracting the vectors $\{a_{\rm hom}e_i\}_{i\in\{1,\dots,d\}}$,

$$\begin{aligned}\nabla\cdot a\nabla v^\epsilon+\nabla\cdot a(\nabla\phi_i\eta_\rho^\epsilon\partial_i v^\epsilon)=&\\ \nabla\cdot((1-\eta_\rho^\epsilon)a\nabla v^\epsilon)+\nabla\cdot((q_i-a_{\rm hom}e_i)\eta_\rho^\epsilon\partial_i v^\epsilon)+\nabla\cdot(\eta_\rho^\epsilon a_{\rm hom}\nabla v^\epsilon).&\end{aligned}$$

Therefore, since v satisfies (86),

$$\nabla\cdot a\nabla v^\epsilon+\nabla\cdot a(\nabla\phi_i\eta_\rho^\epsilon\partial_i v^\epsilon)=v_t^\epsilon+\nabla\cdot((1-\eta_\rho^\epsilon)(a-a_{\rm hom})\nabla v^\epsilon)+\nabla\cdot((q_i-a_{\rm hom}e_i)\eta_\rho^\epsilon\partial_i v^\epsilon).\quad(92)$$

Returning to (91), in view of (92),

$$-\nabla\cdot a\nabla w=-u_t+v_t^\epsilon+\nabla\cdot((1-\eta_\rho^\epsilon)(a-a_{\rm hom})\nabla v^\epsilon)+\nabla\cdot((q_i-a_{\rm hom}e_i)\eta_\rho^\epsilon\partial_i v^\epsilon+\phi_i a\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)).\quad(93)$$

For the derivative in time, owing to definition (89),

$$w_t=u_t-v_t^\epsilon-\phi_{i,t}\eta_\rho^\epsilon\partial_i v^\epsilon-\phi_i(\eta_\rho^\epsilon\partial_i v^\epsilon)_t,$$

which, in combination with (93), yields the distributional equality

$$\begin{aligned}w_t-\nabla\cdot a\nabla w=&\nabla\cdot\big((1-\eta_\rho^\epsilon)(a-a_{\rm hom})\nabla v^\epsilon+\phi_i a\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\big)+(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\\ &+((\nabla\cdot q_i)-\phi_{i,t})\,\eta_\rho^\epsilon\partial_i v^\epsilon-\phi_i(\eta_\rho^\epsilon\partial_i v^\epsilon)_t.\end{aligned}$$

Therefore, since the correctors $\{\phi_i\}_{i\in\{1,\dots,d\}}$ satisfy (6), distributionally

$$\begin{aligned}w_t-\nabla\cdot a\nabla w=&\nabla\cdot\big((1-\eta_\rho^\epsilon)(a-a_{\rm hom})\nabla v^\epsilon+\phi_i a\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\big)\\ &+(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)-\phi_i(\eta_\rho^\epsilon\partial_i v^\epsilon)_t.\end{aligned}\quad(94)$$

It remains to analyze the term

$$(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon).$$

For the correctors $\{\psi_i\}_{i\in\{1\dots,d\}}$ satisfying (18), add and subtract the gradients $\{\nabla\psi_i\}_{i\in\{1,\dots,d\}}$ and add and subtract the conditional expectations $\{\langle q_i\mid\mathcal{F}_{\mathbb{R}^d}\rangle\}_{i\in\{1,\dots,d\}}$ to obtain

$$\begin{aligned}(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)=&(q_i-\nabla\psi_i-\langle q_i\mid\mathcal{F}_{\mathbb{R}^d}\rangle)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\\ &+(\langle q_i\mid\mathcal{F}_{\mathbb{R}^d}\rangle-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)+\nabla\psi_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon).\end{aligned}$$

Since the correctors $\{\sigma_i\}_{i\in\{1,\dots,d\}}$ satisfy (19) and the correctors $\{\zeta_i\}_{i\in\{1,\dots,d\}}$ satisfy (21),

$$(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)=(\nabla\cdot\sigma_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)+\partial_t\zeta_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)+\nabla\psi_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon).\quad(95)$$

Then, for each $i\in\{1,\dots,d\}$, the skew-symmetry of σ_i proven in Lemma 2 implies the distributional equality

$$\nabla\cdot(\sigma_i\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon))=-(\nabla\cdot\sigma_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon).\quad(96)$$

Indeed, for each $i\in\{1,\dots,d\}$, distributionally

$$\begin{aligned}\nabla\cdot(\sigma_i\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon))=\partial_j(\sigma_{ijk}\partial_k(\eta_\rho^\epsilon\partial_i v^\epsilon))=\partial_j\sigma_{ijk}\partial_k(\eta_\rho^\epsilon\partial_i v^\epsilon)+\sigma_{ijk}\partial_j\partial_k(\eta_\rho^\epsilon\partial_i v^\epsilon)=&\\ -\partial_j\sigma_{ikj}\partial_k(\eta_\rho^\epsilon\partial_i v^\epsilon)+\sigma_{ijk}\partial_j\partial_k(\eta_\rho^\epsilon\partial_i v^\epsilon)=-(\nabla\cdot\sigma_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon),&\end{aligned}$$

where the penultimate inequality follows from the skew-symmetry of σ and the final equality from the skew-symmetry of σ and the equality of mixed partial derivatives.

Therefore, returning to (95), the equality (96) and the distributional equality

$$\nabla\psi_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)=\nabla\cdot(\psi_i\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon))-\psi_i\Delta(\eta_\rho^\epsilon\partial_i v^\epsilon)$$

imply that

$$(q_i-a_{\rm hom}e_i)\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)=\nabla\cdot\big((\psi_i-\sigma_i)\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\big)+\partial_t\zeta_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)-\psi_i\Delta(\eta_\rho^\epsilon\partial_i v^\epsilon).\quad(97)$$

Therefore, returning to (94), in view of (97), it follows that

$$\begin{aligned}w_t-\nabla\cdot a\nabla w=&\nabla\cdot\big((1-\eta_\rho^\epsilon)(a-a_{\rm hom})\nabla v^\epsilon\big)+\nabla\cdot\big((\phi_i a+\psi_i-\sigma_i)\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)\big)\\ &+\partial_t\zeta_i\cdot\nabla(\eta_\rho^\epsilon\partial_i v^\epsilon)-\phi_i(\eta_\rho^\epsilon\partial_i v^\epsilon)_t-\psi_i\Delta(\eta_\rho^\epsilon\partial_i v^\epsilon),\end{aligned}\quad(98)$$

which completes the proof of (90). This equation will later be used to obtain an energy estimate for the augmented homogenization error. However, first, it is useful to recall three classical estimates concerning the boundary and interior regularity of a_{hom}-caloric functions.

5.3. **Interior and boundary estimates for a_{hom}-caloric functions.** In this subsection, three classical estimates are presented to control the interior and boundary regularity of a_{hom}-caloric functions.

In what follows, the boundary conditions will be assumed to be the trace of a function $\tilde{u} : \overline{\mathcal{C}}_1 \to \mathbb{R}$ satisfying

$$\tilde{u} \in L^2([-1,0]; H^1(B_1)) \text{ and } \tilde{u}_t \in L^2([-1,0]; H^{-1}(B_1)). \tag{99}$$

The first estimate is the a priori energy estimate for the a_{hom}-caloric extension of $\tilde{u}$ into $\mathcal{C}_1$. That is, if $\tilde{v}$ satisfies

$$\left\{ \begin{array}{ll} \tilde{v}_t &= \nabla \cdot a_{\text{hom}} \nabla \tilde{v} \quad \text{in } \mathcal{C}_1 \\ \tilde{v} &= \tilde{u} \qquad\qquad\quad \text{on } \partial_p \mathcal{C}_1, \end{array} \right. \tag{100}$$

then,

$$\int_{\mathcal{C}_1} |\nabla \tilde{v}|^2 \lesssim \int_{\mathcal{C}_1} |\nabla \tilde{u}|^2 + \|\tilde{u}_t\|^2_{L^2([-1,0];H^{-1}(B_1))}. \tag{101}$$

To prove (101), let $\tilde{z}$ denote the distributional solution of

$$\left\{ \begin{array}{ll} \tilde{z}_t &= \nabla \cdot a_{\text{hom}} \nabla \tilde{z} + \nabla \cdot a_{\text{hom}} \nabla \tilde{u} - \tilde{u}_t \quad \text{in } \mathcal{C}_1 \\ \tilde{z} &= 0 \qquad\qquad\qquad\qquad\quad \text{on } \partial \mathcal{C}_1. \end{array} \right. \tag{102}$$

Then, testing (102) with $\tilde{z}$ and, after applying the Poincaré inequality, Hölder's inequality and Young's inequality and using the uniform ellipticity of a_{hom} from Lemma 2, it follows that

$$\int_{\mathcal{C}_1} |\nabla \tilde{z}|^2 \lesssim \int_{\mathcal{C}_1} |\nabla \tilde{u}|^2 + \|\tilde{u}_t\|^2_{L^2([-1,0];H^{-1}(B_1))}. \tag{103}$$

However, thanks to (99) and (100), it is then immediate that

$$\tilde{v} = \tilde{z} + \tilde{u}.$$

Hence, with (103) and the triangle inequality,

$$\int_{\mathcal{C}_1} |\nabla \tilde{v}|^2 \lesssim \int_{\mathcal{C}_1} |\nabla \tilde{u}|^2 + \|\tilde{u}_t\|^2_{L^2([-1,0];H^{-1}(B_1))},$$

which proves (101).

An interior regularity estimate will now be obtained for a_{hom}-caloric functions. Suppose that $\tilde{v}$ satisfies (100) for $\tilde{u}$ satisfying (99). It then follows from a repeated application of the Caccioppoli inequality (27) that, for each $k \geq 0$, there exists $C(k) > 0$ such that

$$\int_{B_{1-\rho}} \left| \nabla^k \tilde{v} \right|^2 \leq \frac{C(k)}{(R\rho)^{2k}} \int_{B_1} |\nabla \tilde{v}|^2 .$$

Therefore, by choosing $k = \frac{d}{2}$, $k = \frac{d}{2} + 1$ and $k = \frac{d}{2} + 2$, the Sobolev embedding theorem implies that

$$\sup_{\mathcal{C}_{1-\rho}} \left(|\nabla \tilde{v}| + \rho \left| \nabla^2 \tilde{v} \right| + \rho^2 \left| \nabla^3 \tilde{v} \right| \right)^2 \lesssim \rho^{-(d+2)} \int_{\mathcal{C}_1} |\nabla \tilde{v}|^2 \lesssim \rho^{-(d+2)} \left(\int_{\mathcal{C}_1} |\nabla \tilde{u}|^2 + \|\tilde{u}_t\|^2_{L^2([-1,0];H^{-1}(B_1))} \right), \tag{104}$$

where the final inequality follows from (101).

The boundary regularity statement follows from a simplified version of Ladyzenskaja, Solonnikov and Uraltceva [20, Theorem 9.1] or, for the optimal statement, Weidemaier [27, Theorem 3.1]. This estimate will obtain H^2-regularity, and therefore requires more from the boundary

condition. In particular, this estimate explains the necessity of introducing the boundary regularization in the definition of the augmented homogenization error. Suppose that $\tilde{u}$ satisfies the trace estimates

$$\tilde{u} \in L^2([-1,0];H^1(\partial B_1)) \cap H^1(B_1 \times \{-1\}) \text{ and } \tilde{u}_t \in L^2([-1,0];L^2(\partial B_1)),$$

and that $\tilde{v}$ is the a_{hom}-caloric extension of $\tilde{u}$ into $\mathcal{C}_1$ in the sense of (100). Then, it follows from [20, Theorem 9.1] or [27, Theorem 3.1] that

$$\int_{\mathcal{C}_1} |\nabla \tilde{v}|^2 + \left|\nabla^2 \tilde{v}\right|^2 \lesssim \int_{\partial_p \mathcal{C}_1} \left|\nabla^{\tan} \tilde{u}\right|^2 + \int_{-1}^{0} \int_{\partial B_1} |\tilde{u}_t|^2 , \tag{105}$$

where $\nabla^{\tan}\tilde{u}$ denotes the tangential derivative of $\tilde{u}$ on the parabolic boundary. In particular, $\nabla^{\tan}\tilde{u}$ coincides with the full gradient on $B_1 \times \{-1\}$. Estimates (101), (104) and (105) will play an important role in the energy estimate to follow.

5.4. **The energy estimate for the augmented homogenization error.** Equation (90) will now be used to obtain an energy estimate for the augmented homogenization error w defined in (89). Precisely, it will be shown that

$$\begin{aligned}
\int_{\mathcal{C}_{r_\epsilon}} \nabla w \cdot a \nabla w \lesssim & \epsilon \int_{\mathcal{C}_1} |\nabla u|^2 + \frac{\rho^{\frac{2}{d}}}{\epsilon^2} \int_{\mathcal{C}_1} |\nabla u|^2 \\
& + \frac{1}{\rho^{d+4}} \int_{\mathcal{C}_1} \left(|\phi|^2 + |\psi|^2 + |\sigma|^2\right) \int_{\mathcal{C}_1} |\nabla u|^2 \\
& + \left(\frac{1}{\rho^{\frac{d}{2}+3}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} + \frac{1}{\rho^{d+6}} \int_{\mathcal{C}_1} |\zeta|^2 \right) \int_{\mathcal{C}_1} |\nabla u|^2 \\
& + \frac{1}{\rho^{d+4}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{C}_1} |q|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 .
\end{aligned} \tag{106}$$

The idea is to test equation (98) with w. However, for this it is necessary to introduce a cutoff to ensure that w vanishes along the upper boundary of the cylinder. For each $\delta \in (0,1)$, define a smooth cutoff function $\gamma_\delta : \mathbb{R} \to \mathbb{R}$ which is non-increasing and satisfies $0 \le \gamma_\delta \le 1$ with

$$\gamma_\delta(t) = \begin{cases} 1 & \text{if } t \le -\delta, \\ 0 & \text{if } t \ge 0. \end{cases} \tag{107}$$

Furthermore, for the Dirac mass δ_0 at zero, as $\delta \to 0$,

$$|(\gamma_\delta)_t| \rightharpoonup \delta_0 \text{ as distributions on } \mathbb{R}. \tag{108}$$

To begin, equation (98) is tested against $\gamma_\delta w$. Properties of the cutoff η_ρ^ϵ from (87) and (88), the uniform ellipticity of a from (4) and a_{hom} from Lemma 2 and Hölder's inequality imply

that, after bounding the time derivative of v^ϵ by its Hessian matrix,

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} |(\gamma_\delta)_t|\,|w|^2 + \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \nabla w \cdot a \nabla w \lesssim & \int_{-r_\epsilon^2}^0 \int_{\partial B_{r_\epsilon}} \gamma_\delta (u-u^\epsilon)\nu \cdot a(\nabla u - \nabla v^\epsilon) \\ & + \int_{B_{r_\epsilon}\times\{-r_\epsilon{}^2\}} \gamma_\delta\,|u-u^\epsilon|^2 \\ & + \int_{\mathcal{C}_{r_\epsilon}\setminus\mathcal{C}_{r_\epsilon-2\rho}} \gamma_\delta\,|\nabla v^\epsilon|\,|\nabla w| \\ & + \sup_{\mathcal{C}_{r_\epsilon-\rho}} \left(\left|\nabla^2 v^\epsilon\right| + \frac{1}{R\rho}\,|\nabla v^\epsilon|\right) \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta\,(|\phi|+|\psi|+|\sigma|)\,|\nabla w| \\ & + \sup_{\mathcal{C}_{r_\epsilon-\rho}} \left(\frac{1}{\rho^2}\,|\nabla v^\epsilon| + \frac{1}{\rho}\left|\nabla^2 v^\epsilon\right| + \left|\nabla^3 v^\epsilon\right|\right) \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta\,(|\phi|+|\psi|)\,|w| \\ & + \left| \int_{\mathcal{C}_{r_\epsilon}} \partial_t \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta w \right|, \end{aligned} \tag{109}$$

where ν denotes the interior normal and

$$|\phi| := \left(\sum_{i=1}^d |\phi_i|^2\right)^{\frac12}, \quad |\psi| := \left(\sum_{i=1}^d |\psi_i|^2\right)^{\frac12} \text{ and } \quad |\sigma| := \left(\sum_{i,j,k=1}^d |\sigma_{ijk}|^2\right)^{\frac12}.$$

For the first two boundary terms, it is immediate from the choice of $r_\epsilon \in (\frac12, \frac34)$ in (83) and (85), the uniform ellipticity of a, Hölder's inequality and the estimate for the Dirichlet to Neumann map, see [13], that

$$\begin{aligned}\int_{-r_\epsilon^2}^0 \int_{\partial B_{r_\epsilon}} (u-u^\epsilon)\nu\cdot a(\nabla u - \nabla v^\epsilon) + \int_{B_{r_\epsilon}\times\{-r_\epsilon^2\}} |u-u^\epsilon|^2 \lesssim \\ \left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |u-u^\epsilon|^2\right)^{\frac12} \left(\int_{\partial_p \mathcal{C}_{r_\epsilon}} |\nabla u|^2 + \left|\nabla^{\tan} u\right|^2\right)^{\frac12} + \int_{\partial_p \mathcal{C}_{r_\epsilon}} |u-u^\epsilon|^2 \lesssim \epsilon \int_{\mathcal{C}_1} |\nabla u|^2 .\end{aligned}$$

It is then necessary to analyze the final term on the right hand side of (109). Using the definition of w from (89) and the fact that ζ vanishes at $t=0$ owing to (80), it follows after integrating by parts variously in time and space that

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \partial_t \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta w = & \int_{\mathcal{C}_{r_\epsilon}} (\eta_\rho^\epsilon \partial_i v^\epsilon)_t \zeta_i \cdot \gamma_\delta \nabla w - \int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)(\gamma_\delta)_t w \\ & - \int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta (u_t - v_t) - \int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta (\eta_\rho^\epsilon \phi_j \partial_j v^\epsilon)_t, \end{aligned} \tag{110}$$

where this equality uses the fact that the corrector ζ and the cutoff γ_δ are constant in space.

The first two terms of (110) are bounded immediately using the definition of the cutoff η_ρ^ϵ from (87) and (88), which yields

$$\left| \int_{\mathcal{C}_{r_\epsilon}} (\eta_\rho^\epsilon \partial_i v^\epsilon)_t \zeta_i \cdot \gamma_\delta \nabla w \right| \lesssim \sup_{\mathcal{C}_{r_\epsilon-\rho}} \left(\frac{1}{\rho^2}\,|\nabla v^\epsilon| + \left|\nabla^3 v^\epsilon\right|\right) \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta\,|\zeta|\,|\nabla w|, \tag{111}$$

where

$$|\zeta| := \left(\sum_{i,j=1}^d |\zeta_{ij}|^2\right)^{\frac12}.$$

Similarly,

$$\left|\int_{\mathcal{C}_{r_\epsilon}} (\zeta_i \cdot \nabla \left(\eta_\rho^\epsilon \partial_i v^\epsilon\right))(\gamma_\delta)_t w\right| \lesssim \sup_{\mathcal{C}_{r_\epsilon - \rho}} \left(\frac{1}{\rho}\left|\nabla v^\epsilon\right| + \left|\nabla^2 v^\epsilon\right|\right) \int_{\mathcal{C}_{r_\epsilon}} \left|(\gamma_\delta)_t\right| |\zeta| \, |w| \,. \tag{112}$$

It is necessary to analyze the final two terms of (110). For the first of these, using the equations (81) and (86) satisfied by u and v respectively,

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta(u_t - v_t) = &- \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \zeta_i \cdot \left(\nabla^2(\eta_\rho^\epsilon \partial_i v^\epsilon) \cdot a \nabla u\right) \\ &+ \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \zeta_i \cdot \left(\nabla^2(\eta_\rho^\epsilon \partial_i v^\epsilon) \cdot a_{\hom} \nabla v^\epsilon\right).\end{aligned}$$

Therefore, using the uniform ellipticity (4) of a and the uniform ellipticity of $a_{\hom}$ from Lemma 2, after bounding the time derivative of v by the norm of its Hessian matrix,

$$\begin{aligned}&\left|\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta(u_t - v_t^\epsilon)\right| \lesssim \\ &\qquad \sup_{\mathcal{C}_{r_\epsilon - \rho}} \left(\frac{1}{\rho^2}\left|\nabla v^\epsilon\right| + \frac{1}{\rho}\left|\nabla^2 v^\epsilon\right| + \left|\nabla^3 v^\epsilon\right|\right) \int_{\mathcal{C}_{r_\epsilon - \rho}} \gamma_\delta \left|\zeta\right| \left(\left|\nabla u\right| + \left|\nabla v^\epsilon\right|\right).\end{aligned} \tag{113}$$

For the final term of (110),

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta(\eta_\rho^\epsilon \phi_j \partial_j v^\epsilon)_t = &\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta \phi_j (\eta_\rho^\epsilon \partial_j v^\epsilon)_t \\ &+ \int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta \phi_{j,t} (\eta_\rho^\epsilon \partial_j v^\epsilon),\end{aligned}$$

and, therefore, using the equation (6) satisfied by the correctors $\{\phi_i\}_{i\in\{1,\dots,d\}}$,

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta(\eta_\rho^\epsilon \phi_j \partial_j v^\epsilon)_t = &\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta \phi_j (\eta_\rho^\epsilon \partial_j v^\epsilon)_t \\ &- \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \nabla \left(\zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\eta_\rho^\epsilon \partial_j v^\epsilon\right) \cdot q_j,\end{aligned}$$

for the fluxes $\{q_i\}_{i\in\{1,\dots,d\}}$ defined in (10). Hence, after bounding the time derivative of v by its Hessian matrix,

$$\begin{aligned}&\left|\int_{\mathcal{C}_{r_\epsilon}} \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon)\gamma_\delta(\eta_\rho^\epsilon \phi_j \partial_j v^\epsilon)_t\right| \lesssim \\ &\quad \sup_{\mathcal{C}_{r_\epsilon - \rho}} \left(\left|\nabla v^\epsilon\right| \left(\frac{1}{\rho^2}\left|\nabla v^\epsilon\right| + \frac{1}{\rho}\left|\nabla^2 v^\epsilon\right| + \left|\nabla^3 v^\epsilon\right|\right) + \left(\frac{1}{\rho}\left|\nabla v^\epsilon\right| + \left|\nabla^2 v^\epsilon\right|\right)^2\right) \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \left|\zeta\right| |q| \\ &\qquad + \sup_{\mathcal{C}_{r_\epsilon - \rho}} \left(\left(\frac{1}{\rho}\left|\nabla v^\epsilon\right| + \left|\nabla^2 v^\epsilon\right|\right)\left(\frac{1}{\rho^2}\left|\nabla v^\epsilon\right| + \left|\nabla^3 v^\epsilon\right|\right)\right) \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \left|\zeta\right| |\phi| \,,\end{aligned} \tag{114}$$

where

$$|q| := \left(\sum_{i=1}^d |q_i|^2\right)^{\frac{1}{2}}.$$

Therefore, in view of (82) and (109), it follows from the uniform ellipticity of a, the definition of γ_δ, the Poincaré inequality in space, Hölder's inequality and Young's inequality that

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} |(\gamma_\delta)_t|\, w^2 + \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \nabla w \cdot a \nabla w \lesssim & \epsilon \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \int_{\mathcal{C}_{r_\epsilon}\setminus\mathcal{C}_{r_\epsilon-2\rho}} |\nabla v^\epsilon|^2 \\ & + \frac{1}{\rho^{d+4}} \int_{\mathcal{C}_1} \left(|\phi|^2 + |\psi|^2 + |\sigma|^2\right) \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \left| \int_{\mathcal{C}_{r_\epsilon}} \partial_t \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon) \gamma_\delta w \right|. \end{aligned} \tag{115}$$

For the second term of (115), the choice of the radius $r_\epsilon \in (\frac{1}{2}, \frac{3}{4})$ satisfying (84) and estimate (101) for a_{hom}-caloric functions imply that

$$\int_{\mathcal{C}_{r_\epsilon}} |\nabla v^\epsilon|^2 + \left|\nabla^2 v^\epsilon\right|^2 \lesssim \int_{\partial_p \mathcal{C}_{r_\epsilon}} \left|\nabla^{\tan} u^\epsilon\right|^2 + \int_{-r_\epsilon^2}^0 \int_{\partial B_{r_\epsilon}} |u_t^\epsilon|^2 \lesssim \frac{1}{\epsilon^2} \int_{\mathcal{C}_{r_\epsilon}} |\nabla u^\epsilon|^2 \leq \frac{1}{\epsilon^2} \int_{\mathcal{C}_1} |\nabla u|^2 .$$

Therefore, it follows from the Sobolev embedding theorem that

$$\int_{\mathcal{C}_{r_\epsilon}\setminus\mathcal{C}_{r_\epsilon-2\rho}} |\nabla v^\epsilon|^2 \lesssim \left(\int_{\mathcal{C}_{r_\epsilon}} \chi_{\mathcal{C}_{r_\epsilon}\setminus\mathcal{C}_{r_\epsilon-2\rho}} \right)^{\frac{2}{d}} \left(\int_{\mathcal{C}_{r_\epsilon}} |\nabla v^\epsilon|^{\frac{2d}{d-2}} \right)^{\frac{d-2}{d}} \leq \frac{\rho^{\frac{2}{d}}}{\epsilon^2} \int_{\mathcal{C}_1} |\nabla u|^2 ,$$

where $\chi_{\mathcal{C}_{r_\epsilon}\setminus\mathcal{C}_{r_\epsilon-2\rho}}$ is the indicator function of the set $(\mathcal{C}_{r_\epsilon} \setminus \mathcal{C}_{r_\epsilon-2\rho})$, and where the argument is only written for the case $d \geq 3$, since the modifications necessary for the cases $d = 1$ and $d = 2$ are straightforward and rely only upon the Sobolev embedding theorem.

For the final term of (115), estimates (111), (112), (113) and (114) together with estimates (82) and (104), where Hölder's inequality is used for the final term, prove that, since removing γ_δ from the final three terms of the right hand side increases their magnitude,

$$\begin{aligned}\left| \int_{\mathcal{C}_{r_\epsilon}} \partial_t \zeta_i \cdot \nabla(\eta_\rho^\epsilon \partial_i v^\epsilon) \gamma_\delta w \right| \lesssim & \frac{1}{\rho^{\frac{d}{2}+3}} \int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta\, |\zeta|\, |\nabla w| \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}} \\ & + \frac{1}{\rho^{\frac{d}{2}+1}} \int_{\mathcal{C}_{r_\epsilon}} |(\gamma_\delta)_t|\, |\zeta|\, |w| \left(\int_{\mathcal{C}_1} |\nabla u|^2 \right)^{\frac{1}{2}} \\ & + \frac{1}{\rho^{\frac{d}{2}+3}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+4}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{C}_1} |q|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+5}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{C}_1} |\phi|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 . \end{aligned} \tag{116}$$

Therefore, following an application of Hölder's inequality and then Young's inequality, it follows from (115) and (116) that

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \gamma_\delta \nabla w \cdot a \nabla w \lesssim & \epsilon \int_{\mathcal{C}_1} |\nabla u|^2 + \frac{\rho^{\frac{2}{d}}}{\epsilon^2} \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+4}} \int_{\mathcal{C}_1} \left(|\phi|^2 + |\psi|^2 + |\sigma|^2 \right) \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \left(\frac{1}{\rho^{\frac{d}{2}+3}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} + \frac{1}{\rho^{d+6}} \int_{\mathcal{C}_1} |\zeta|^2 \right) \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+4}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{C}_1} |q|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+2}} \int_{\mathcal{C}_1} |(\gamma_\delta)_t| \, |\zeta|^2 \int_{\mathcal{C}_1} |\nabla u|^2 . \end{aligned} \tag{117}$$

In view of the construction of γ_δ from (107), and owing to the distributional convergence (108), the fact that ζ vanishes at $t = 0$ thanks to (80) implies, for $\langle\cdot\rangle$-a.e. a,

$$\lim_{\delta \to 0} \int_{\mathcal{C}_1} |(\gamma_\delta)_t| \, |\zeta|^2 = 0.$$

Therefore, after passing to the limit $\delta \to 0$ in (117), the construction of γ_δ in (107) implies that

$$\begin{aligned}\int_{\mathcal{C}_{r_\epsilon}} \nabla w \cdot a \nabla w \lesssim & \epsilon \int_{\mathcal{C}_1} |\nabla u|^2 + \frac{\rho^{\frac{2}{d}}}{\epsilon^2} \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+4}} \int_{\mathcal{C}_1} \left(|\phi|^2 + |\psi|^2 + |\sigma|^2 \right) \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \left(\frac{1}{\rho^{\frac{d}{2}+3}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} + \frac{1}{\rho^{d+6}} \int_{\mathcal{C}_1} |\zeta|^2 \right) \int_{\mathcal{C}_1} |\nabla u|^2 \\ & + \frac{1}{\rho^{d+4}} \left(\int_{\mathcal{C}_1} |\zeta|^2 \right)^{\frac{1}{2}} \left(\int_{\mathcal{C}_1} |q|^2 \right)^{\frac{1}{2}} \int_{\mathcal{C}_1} |\nabla u|^2 , \end{aligned} \tag{118}$$

which completes the proof of (106).

To obtain (118) for an arbitrary radius $R > 0$, suppose that u is an a-caloric function on B_R. Then, for each $\epsilon \in (0, \frac{R}{4})$, there exists a radius $R_\epsilon \in (\frac{R}{2}, \frac{3R}{4})$ and a cutoff function $\eta_\rho^{R_\epsilon}$ with $0 \leq \eta_\rho^{R_\epsilon} \leq 1$ and such that

$$\eta_\rho^{R_\epsilon}(x,t) = \begin{cases} 1 & \text{if } (x,t) \in \overline{\mathcal{C}}_{R_\epsilon - 2\rho R_\epsilon} \\ 0 & \text{if } (x,t) \in (\mathbb{R}^d \times (-\infty, 0)) \setminus \mathcal{C}_{R_\epsilon - \rho R_\epsilon}, \end{cases} \tag{119}$$

which, for the $a_{\hom}$-caloric extension v^ϵ of u^ϵ into $\mathcal{C}_{R_\epsilon}$, define the corresponding augmented homogenization error

$$w = u - (1 + \eta_\rho^{R_\epsilon} \phi_i \partial_i) v^\epsilon.$$

Then, for each $\epsilon \in (0, \frac{R}{4})$, after performing the rescalings

$$(\tilde{w}, \tilde{u}, \tilde{v^\epsilon}, \tilde{q})(\cdot, \cdot) = \frac{1}{R_\epsilon} (w, u, v^\epsilon, q)(R_\epsilon \cdot, R_\epsilon^2 \cdot),$$

it follows using equations (6), (18), (19) and (21) that the correctors rescale according to the rules

$$(\tilde{\phi}, \tilde{\psi}, \tilde{\sigma}) = \frac{1}{R_\epsilon} (\phi, \psi, \sigma)(R_\epsilon \cdot, R_\epsilon^2 \cdot) \text{ and } \tilde{\zeta}(\cdot) = \frac{1}{R_\epsilon^2} \zeta(R_\epsilon^2 \cdot).$$

Hence, after applying (118) and returning to the original scaling, it follows that, for each $\epsilon \in (0, \frac{R}{4})$ and $\rho \in (0, \frac{1}{8})$,

$$\begin{aligned}\fint_{\mathcal{C}_{R_\epsilon}} \nabla w \cdot a \nabla w \lesssim & \epsilon \fint_{\mathcal{C}_R} |\nabla u|^2 + \frac{\rho^{\frac{2}{d}}}{\epsilon^2} \fint_{\mathcal{C}_R} |\nabla u|^2 \\ & + \frac{1}{R^2 \rho^{d+4}} \fint_{\mathcal{C}_R} \left(|\phi|^2 + |\psi|^2 + |\sigma|^2 \right) \fint_{\mathcal{C}_R} |\nabla u|^2 \\ & + \left(\frac{1}{R^2 \rho^{\frac{d}{2}+3}} \left(\int_{\mathcal{C}_R} |\zeta|^2 \right)^{\frac{1}{2}} + \frac{1}{R^4 \rho^{d+6}} \int_{\mathcal{C}_R} |\zeta|^2 \right) \int_{\mathcal{C}_R} |\nabla u|^2 \\ & + \frac{1}{R^2 \rho^{d+4}} \left(\fint_{\mathcal{C}_R} |\zeta|^2 \right)^{\frac{1}{2}} \left(\fint_{\mathcal{C}_R} |q|^2 \right)^{\frac{1}{2}} \fint_{\mathcal{C}_R} |\nabla u|^2 ,\end{aligned} \tag{120}$$

which is the general form of the energy estimate that will be used in the proof of excess decay to follow.

5.5. **The proof of excess decay.** The energy estimate will now be used to prove the excess decay of Proposition 4. Fix $R > 0$ and suppose that u is an a-caloric function $\mathcal{C}_R$. Then, for each $\epsilon \in (0, \frac{R}{4})$ and $\rho \in (0, \frac{1}{8})$, choose a radius $R_\epsilon \in (\frac{R}{2}, \frac{3R}{4})$ and a cutoff $\eta_\rho^{R_\epsilon}$ such that, for the $a_{\hom}$-caloric extension v^ϵ of u^ϵ into $\mathcal{C}_{R_\epsilon}$, the conclusion of (120) is satisfied for the augmented homogenization error w defined by

$$w = u - (1 + \eta_\rho^{R_\epsilon} \phi_i \partial_i) v^\epsilon \ \text{ in } \ \mathcal{C}_{R_\epsilon}. \tag{121}$$

The proof of excess decay will now proceed in four steps.

Step 1. In the first step of the proof, it will be shown that, for any $\delta > 0$, there exists $C_2 = C_2(d, \lambda, \delta) > 0$ such that, whenever, for each $i \in \{1, \dots, d\}$,

$$\left(\fint_{\mathcal{C}_R} |q_i|^2 \right)^{\frac{1}{2}} \leq 2 \left\langle |q_i|^2 \right\rangle^{\frac{1}{2}}, \tag{122}$$

and

$$\frac{1}{R} \left(\fint_{\mathcal{C}_R} |\phi_i|^2 + |\psi_i|^2 + |\sigma_i|^2 \right)^{\frac{1}{2}} + \frac{1}{R^2} \left(\fint_{\mathcal{C}_R} |\zeta_i|^2 \right)^{\frac{1}{2}} \leq \frac{1}{C_2}, \tag{123}$$

then

$$\fint_{\mathcal{C}_{R_\epsilon}} \nabla w \cdot a \nabla w \lesssim \delta \fint_{\mathcal{C}_R} \nabla u \cdot a \nabla u. \tag{124}$$

The proof is a simple consequence of estimate (120) and the definition (121).

Fix $\delta > 0$. Then, assuming that (122) and (123) are satisfied for some $C_2 > 0$ to be fixed later, estimate (118), the choice (121) and the uniform ellipticity of a imply that

$$\begin{aligned}\fint_{\mathcal{C}_{R_1}} \nabla w \cdot a \nabla w \lesssim & \\ & \left(\epsilon + \frac{\rho^{\frac{2}{d}}}{\epsilon^2} + \frac{1}{C_2^2 \rho^{d+4}} + \frac{1}{C_2 \rho^{\frac{d}{2}+3}} + \frac{1}{C_2^2 \rho^{d+6}} + \frac{1}{C_2 \rho^{d+4}} \left\langle |q|^2 \right\rangle^{\frac{1}{2}} \right) \fint_{\mathcal{C}_{R_1}} \nabla u \cdot a \nabla u.\end{aligned}$$

Therefore, first choose $\epsilon_0 \in (0, \frac{1}{4})$ satisfying

$$\epsilon_0 < \frac{1}{3} \delta.$$

Then, choose $\rho_0 \in (0, \frac{1}{8})$ sufficiently small so as to guarantee that

$$\frac{\rho^{\frac{2}{d}}}{\epsilon_0^2} < \frac{1}{3} \delta.$$

Finally, fix $C_2 > 0$ large enough to ensure that

$$\left(\frac{1}{C_2^2 \rho_0^{d+4}} + \frac{1}{C_2 \rho_0^{\frac{d}{2}+3}} + \frac{1}{C_2^2 \rho_0^{d+6}} + \frac{1}{C_2 \rho_0^{d+4}} \left\langle |q^2| \right\rangle^{\frac{1}{2}} \right) < \frac{\delta}{3}.$$

Then, it follows that, for this choice of ϵ_0, ρ_0 and C_2,

$$\fint_{\mathcal{C}_{R_{\epsilon_0}}} \nabla w \cdot a \nabla w \lesssim \delta \fint_{\mathcal{C}_R} \nabla u \cdot a \nabla u,$$

which proves (124).

In particular, since $\rho_0 \in (0, \frac{1}{8})$ and $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$, using the definition (119) of the cutoff $\eta_{\rho_0}^{R_{\epsilon_0}}$, it follows that

$$w = u - (1 + \phi_i \partial_i) v^{\epsilon_0} \text{ on } \mathcal{C}_{\frac{R}{4}}. \tag{125}$$

Therefore, since $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$, it follows from (124) and (125) that, for any $\delta > 0$ there exists $C_2 = C_2(d, \lambda) > 0$ such that, whenever (123) and (124) are satisfied, for each $r \in (0, \frac{R}{4}]$,

$$\fint_{\mathcal{C}_r} (\nabla u - \nabla v^{\epsilon_0} - \nabla(\phi_i \partial_i v^{\epsilon_0})) \cdot a (\nabla u - \nabla v^{\epsilon_0} - \nabla(\phi_i \partial_i v^{\epsilon_0})) \lesssim \delta \left(\frac{R}{r} \right)^{d+2} \fint_{\mathcal{C}_R} \nabla u \cdot a \nabla u. \tag{126}$$

This completes the first step of the proof.

Step 2. The second step will show that the left hand side of (126) is a good approximation for the excess by using the interior regularity of $a_{\hom}$-caloric functions and the Caccioppoli inequality. To simplify the notation in what follows, define

$$v := v^{\epsilon_0} \text{ in } \mathcal{C}_{R_{\epsilon_0}}.$$

Then, form the decomposition

$$\nabla u - \nabla v - \nabla(\phi_i \partial_i v) = \nabla u - \nabla v(0,0)(I_d + \nabla \phi) + (\nabla v(0,0) - \nabla v)(I_d + \nabla \phi) - \phi_i \nabla(\partial_i v), \tag{127}$$

where I_d denotes the $(d \times d)$-identity matrix and, for each $i, j \in \{1, \dots, d\}$,

$$(\nabla \phi)_{ij} := \partial_j \phi_i.$$

After fixing $\xi_0 = \nabla v(0,0)$, use (127), the triangle inequality and Young's inequality to prove that, in $\mathcal{C}_r$ for any $r \in (0, \frac{R}{4}]$,

$$|\nabla u - \xi_0 - \nabla \phi_{\xi_0}|^2 \lesssim |\nabla w|^2 + \sup_{\mathcal{C}_r} \left(|\nabla v - \nabla v(0,0)| \right)^2 |I + \nabla \phi|^2 + \sup_{\mathcal{C}_r} \left(|\nabla(\partial_i v)| \right)^2 |\phi_i|^2. \tag{128}$$

Estimate (104) implies that, after bounding the time derivative of v by the norm of its Hessian matrix, and using the uniform ellipticity of a and the choice $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$, for each $r \in (0, \frac{R}{4}]$,

$$\begin{aligned} \sup_{\mathcal{C}_r} \left(|\nabla v - \nabla v(0,0)| \right) &\lesssim \sup_{\mathcal{C}_r} \left(r \left| \nabla^2 v \right| + r^2 \left| \nabla^3 v \right| \right)^2 \\ &\lesssim \left(\frac{r}{R} \right)^2 \sup_{\mathcal{C}_{\frac{R}{4}}} \left(\frac{R}{4} \left| \nabla^2 v \right| + \left(\frac{R}{4} \right)^2 \left| \nabla^3 v \right| \right)^2 \lesssim \left(\frac{r}{R} \right)^2 \fint_{\mathcal{C}_R} \nabla u \cdot a \nabla u. \end{aligned} \tag{129}$$

Similarly, for each $i \in \{1, \dots, d\}$, using estimate (104), the uniform ellipticity of a and $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$, it follows that, for each $r \in (0, \frac{R}{4}]$,

$$\sup_{\mathcal{C}_r} \left(|\nabla(\partial_i v)| \right)^2 \lesssim \sup_{\mathcal{C}_{\frac{R}{4}}} \left(|\nabla(\partial_i v)| \right)^2 \lesssim \frac{1}{R^2} \fint_{\mathcal{C}_R} \nabla u \cdot a \nabla u. \tag{130}$$

Finally, since for each $i \in \{1, \dots, d\}$ the a-caloric coordinate $(x_i + \phi_i)$ satisfies

$$\partial_t (x_i + \phi_i) = \nabla \cdot a \nabla (x_i + \phi_i) \text{ in } \mathbb{R}^{d+1},$$

the Caccioppoli inequality (27) implies that, for each $r \in (0, \frac{R}{4}]$,

$$\fint_{\mathcal{C}_r} |e_i + \nabla \phi_i|^2 \lesssim \frac{1}{(2r)^2} \fint_{\mathcal{C}_{2r}} |\phi_i|^2 + \frac{1}{(2r)^2} \fint_{\mathcal{C}_{2r}} x_i^2 \lesssim \frac{1}{(2r)^2} \fint_{\mathcal{C}_{2r}} |\phi_i|^2 + 1. \tag{131}$$

Therefore, returning to (128), estimates (129), (130) and (131), with the uniform ellipticity of a and the choice $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$, imply that, for each $r \in (0, \frac{R}{4}]$,

$$\begin{aligned}\fint_{\mathcal{C}_r} (\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \cdot a(\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \lesssim \left(\frac{R}{r}\right)^{d+2} \fint_{\mathcal{C}_{R_{\epsilon_0}}} \nabla w \cdot a\nabla w \\ + \left(\frac{r}{R}\right)^2 \left(\frac{1}{(2r)^2} \fint_{\mathcal{C}_{2r}} |\phi_i|^2 + 1\right) \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u.\end{aligned} \tag{132}$$

This completes the second step.

Step 3. In the third step, inequality (132) will be combined with (124) to prove the excess decay along a subsequence. Namely, for every $\alpha \in (0,1)$, it will be shown that there exists $C_0 = C_0(d,\lambda,\alpha) > 0$ and $\theta_0 = \theta_0(\alpha, d, \lambda) \in (0, \frac{1}{4})$ such that, if $r_1 = \theta_0 R$ and if, for each $r \in [r_1, R]$,

$$\left(\fint_{\mathcal{C}_r} |q_i|^2\right)^{\frac{1}{2}} \leq 2 \left\langle |q_i|^2 \right\rangle^{\frac{1}{2}}, \tag{133}$$

and

$$\frac{1}{r}\left(\fint_{\mathcal{C}_r} |\phi_i|^2 + |\psi_i|^2 + |\sigma_i|^2\right)^{\frac{1}{2}} + \frac{1}{r^2}\left(\fint_{\mathcal{C}_r} |\zeta_i|^2\right)^{\frac{1}{2}} \leq \frac{1}{C_0}, \tag{134}$$

then

$$\operatorname{Exc}(u; r_1) \leq \left(\frac{r_1}{R}\right)^{2\alpha} \operatorname{Exc}(u; R). \tag{135}$$

Notice that the inequality appearing in (135) is exact.

Let $\delta > 0$ be arbitrary. In view of (124), there exists $C_2 = C_2(\delta, d, \lambda) \geq 1$ such that, whenever (133) and (134) are satisfied for the constant C_2, then, since $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$,

$$\fint_{\mathcal{C}_{R_{\epsilon_0}}} \nabla w \cdot a\nabla w \lesssim \delta \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u.$$

Therefore, it follows from inequality (132) and (134) that, for $C_3 = C_3(d,\lambda) > 0$, since $R_{\epsilon_0} \in (\frac{R}{2}, \frac{3R}{4})$,

$$\fint_{\mathcal{C}_{r_1}} (\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \cdot a(\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \leq C_3 \left(\delta \left(\frac{R}{r_1}\right)^{d+2} + \left(\frac{r_1}{R}\right)^2\right) \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u. \tag{136}$$

Choose $\theta_0 \in (0, \frac{1}{4})$ sufficiently small so as to guarantee

$$C_3\theta_0^2 \leq \frac{1}{2}\theta_0^{2\alpha}, \tag{137}$$

which is possible because $\alpha \in (0,1)$, and choose $\delta_0 > 0$ sufficiently small so as to guarantee

$$C_3\delta_0\theta_0^{-(d+2)} \leq \frac{1}{2}\theta_0^{2\alpha}. \tag{138}$$

It is then immediate from (136) that, by choosing θ_0 as in (137) and choosing $C_0 := C_2(\delta_0, d, \lambda)$ for δ_0 defined in (138), whenever (133) and (134) are satisfied for the constant C_0 and $r_1 = \theta_0 R$,

$$\fint_{\mathcal{C}_{r_1}} (\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \cdot a(\nabla u - \xi_0 - \nabla\phi_{\xi_0}) \leq \theta_0^{2\alpha} \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u = \left(\frac{r_1}{R}\right)^{2\alpha} \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u. \tag{139}$$

Since the excess is defined, for each $R > 0$, by

$$\operatorname{Exc}(u; R) = \inf_{\xi \in \mathbb{R}^d} \fint_{\mathcal{C}_R} (\nabla u - \xi - \nabla\phi_\xi) \cdot a(\nabla u - \xi - \nabla\phi_\xi),$$

inequality (139) implies that

$$\operatorname{Exc}(u; r_1) \leq \left(\frac{r_1}{R}\right)^{2\alpha} \fint_{\mathcal{C}_R} \nabla u \cdot a\nabla u. \tag{140}$$

However, because the left hand side of inequality (140) is invariant with respect to the addition of an arbitrary a-caloric gradient $(\xi + \nabla\phi_\xi)$ in the sense that, with (140), for every $\xi \in \mathbb{R}^d$,

$$\operatorname{Exc}(u; r_1) = \operatorname{Exc}(u - (\xi \cdot x + \phi_\xi); r_1) \leq \left(\frac{r_1}{R}\right)^{2\alpha} \fint_{\mathcal{C}_R} (\nabla u - \xi - \nabla\phi_\xi) \cdot a\nabla(u - \xi - \nabla\phi_\xi), \quad (141)$$

taking an infimum on the right hand side with respect to $\xi \in \mathbb{R}^d$ yields

$$\operatorname{Exc}(u; r_1) \leq \left(\frac{r_1}{R}\right)^{2\alpha} \operatorname{Exc}(u; R), \quad (142)$$

which completes the proof of (135), and the argument's third step.

Step 4. The final step completes the proof using (142) and an iteration argument. Fix $r_1 < R$ such that, for $C_0 > 0$ defined following (138), both (133) and (134) are satisfied for the constant C_0 for every $r \in [r_1, R]$. It will be shown that, in this case,

$$\operatorname{Exc}(u; r) \lesssim \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R). \quad (143)$$

Fix θ_0 as defined in (134). If $r \geq R\theta_0$, then using the definition of the excess, for $C = C(\theta_0) > 0$,

$$\begin{aligned} \operatorname{Exc}(u; r) \leq \left(\frac{R}{r}\right)^{d} \operatorname{Exc}(u; R) &= \left(\frac{R}{r}\right)^{d+2\alpha} \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R) \\ &\leq \theta_0^{-(d+2\alpha)} \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R) \leq C \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R). \quad (144) \end{aligned}$$

If $r < \theta_0 R$, then let n be the unique positive integer satisfying $\theta_0^{n-1} R \leq r < \theta_0^n R$. Proceeding inductively, and relying upon the fact that (142) obtains an exact inequality, for constants $C = C(\theta_0) > 0$ which can change between inequalities,

$$\begin{aligned} \operatorname{Exc}(u; r) \leq C\operatorname{Exc}(u; \theta_0^n R) \leq C(\theta_0^n)^{2\alpha}\operatorname{Exc}(u; R) = \\ C\theta_0^{2\alpha}(\theta_0^{n-1})^{2\alpha}\operatorname{Exc}(u; R) \leq C \left(\frac{r}{R}\right)^{2\alpha} \operatorname{Exc}(u; R). \quad (145) \end{aligned}$$

In combination, (144) and (145) prove (143) and complete the proof of Proposition 4.

6. The proof of Lemma 5

Fix a coefficient field a satisfying (4). Fix $R > 0$ and suppose that u is a distributional solution of

$$u_t = \nabla \cdot a\nabla u \text{ in } \mathcal{C}_R. \quad (146)$$

Let $c \in \mathbb{R}$ and $\rho \in (0, \frac{R}{2})$ be arbitrary. The Caccioppoli inequality is obtained by testing equation (146) with $\eta^2(u - c)$ for an appropriately chosen cutoff function η.

Precisely, fix $\eta \in \mathcal{C}_c^\infty(\mathbb{R}^{d+1})$ satisfying $0 \leq \eta \leq 1$ and, for $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$,

$$\eta(x,t) = \begin{cases} 1 & \text{if } (x,t) \in \overline{B}_{R-\rho} \times [\rho^2 - R^2, 0] \\ 0 & \text{if } (x,t) \in \mathbb{R}^{d+1} \setminus \mathcal{C}_R. \end{cases}$$

Furthermore, choose η satisfying

$$|\eta_t| \lesssim \frac{1}{\rho^2} \text{ and } |\nabla\eta| \lesssim \frac{1}{\rho} \text{ on } \mathbb{R}^{d+1}.$$

Test equation (146) against $\eta^2(u - c)$ and use the the definition of η and the identity

$$\nabla(\eta^2(u-c)) \cdot a\nabla u = \eta^2 \nabla u \cdot a\nabla u + 2\eta(u-c)\nabla\eta \cdot a\nabla u$$

to obtain

$$\int_{\mathcal{C}_R} \eta^2 \nabla u \cdot a\nabla u \lesssim \frac{1}{2} \int_{\mathcal{C}_R} (u-c)^2 \partial_t \eta + \int_{\mathcal{C}_R} \eta \left|(u-c)\right| \left|\nabla\eta\right| \left|\nabla u\right|.$$

Therefore, following applications of Hölder's inequality and Young's inequality, and after using definition of η and the uniform ellipticity of a, it follows that

$$\int_{\mathcal{C}_{R-\rho}} |\nabla u|^2 \lesssim \frac{1}{\rho^2} \int_{\mathcal{C}_R \setminus \mathcal{C}_{R-\rho}} (u-c)^2,$$

which completes the proof.

References

1. S. Andres, *Invariance principle for the random conductance model with dynamic bounded conductances*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 2, 352–374. MR 3189075
2. S. Andres, A. Chiarini, J.-D. Deuschel, and M. Slowik, *Quenched invariance principle for random walks with time-dependent ergodic degenerate weights*, ArXiv e-prints (2016).
3. S. Armstrong, A. Bordas, and J.-C. Mourrat, *Quantitative stochastic homogenization and regularity theory of parabolic equations*, ArXiv e-prints (2017).
4. S. Armstrong, T. Kuusi, and J.-C. Mourrat, *Quantitative stochastic homogenization and large-scale regularity*, ArXiv e-prints (2017).
5. S. N. Armstrong and J.-C. Mourrat, *Lipschitz regularity for elliptic equations with random coefficients*, ArXiv e-prints (2014).
6. S. N. Armstrong and C. K. Smart, *Quantitative stochastic homogenization of convex integral functionals*, Ann. Sci. Éc. Norm. Supér. (2015), To appear.
7. M. Avellaneda and F.-H. Lin, *Compactness methods in the theory of homogenization*, Comm. Pure Appl. Math. **40** (1987), no. 6, 803–847. MR 910954
8. ———, *Méthodes de compacité en homogénéisation*, C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), no. 4, 113–116. MR 901621
9. ———, *Un théorème de Liouville pour des équations elliptiques à coefficients périodiques*, C. R. Acad. Sci. Paris Sér. I Math. **309** (1989), no. 5, 245–250. MR 1010728
10. A. Bandyopadhyay and O. Zeitouni, *Random walk in dynamic Markovian random environment*, ALEA Lat. Am. J. Probab. Math. Stat. **1** (2006), 205–224. MR 2249655
11. P. Bella, B. Fehrman, and F. Otto, *A Liouville theorem for elliptic systems with degenerate ergodic coefficients*, ArXiv e-prints (2016).
12. A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, AMS Chelsea Publishing, Providence, RI, 2011, Corrected reprint of the 1978 original [MR0503330]. MR 2839402
13. Russell M. Brown, *The method of layer potentials for the heat equation in Lipschitz cylinders*, Amer. J. Math. **111** (1989), no. 2, 339–379. MR 987761
14. A. Chiarini and J.-D. Deuschel, *Invariance principle for symmetric diffusions in a degenerate and unbounded stationary and ergodic random medium*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 4, 1535–1563. MR 3573286
15. J. Fischer and F. Otto, *A higher-order large-scale regularity theory for random elliptic operators*, ArXiv e-prints (2015).
16. ———, *Sublinear growth of the corrector in stochastic homogenization: Optimal stochastic estimates for slowly decaying correlations*, ArXiv e-prints (2015).
17. A. Gloria, S. Neukamm, and F. Otto, *A regularity theory for random elliptic operators*, ArXiv e-prints (2014).
18. T. Komorowski and S. Olla, *On homogenization of time-dependent random flows*, Probab. Theory Related Fields **121** (2001), no. 1, 98–116. MR 1857110
19. U. Krengel, *Ergodic theorems*, De Gruyter Studies in Mathematics, vol. 6, Walter de Gruyter & Co., Berlin, 1985, With a supplement by Antoine Brunel. MR 797411
20. O. Ladyzenskaja, V. Solonnikov, and N. Uraltceva, *Linear and quasilinear equations of parabolic type*, Izdat. "Nauka", Moscow, 1967. MR 0241821
21. C. Landim, S. Olla, and H. T. Yau, *Convection-diffusion equation with space-time ergodic random flow*, Probab. Theory Related Fields **112** (1998), no. 2, 203–220. MR 1653837
22. K. Oelschläger, *Homogenization of a diffusion process in a divergence-free random field*, Ann. Probab. **16** (1988), no. 3, 1084–1126. MR 942757
23. G. C. Papanicolaou and S. R. S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, Random fields, Vol. I, II (Esztergom, 1979), Colloq. Math. Soc. János

Bolyai, vol. 27, North-Holland, Amsterdam-New York, 1981, pp. 835–873. MR 712714 (84k:58233)

24. F. Rassoul-Agha and T. Seppäläinen, *An almost sure invariance principle for random walks in a space-time random environment*, Probab. Theory Related Fields **133** (2005), no. 3, 299–314. MR 2198014
25. R. Rhodes, *On homogenization of space-time dependent and degenerate random flows*, Stochastic Process. Appl. **117** (2007), no. 10, 1561–1585. MR 2353040
26. ———, *On homogenization of space-time dependent and degenerate random flows. II*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 4, 673–692. MR 2446293
27. P. Weidemaier, *Maximal regularity for parabolic equations with inhomogeneous boundary conditions in Sobolev spaces with mixed L_p-norm*, Electron. Res. Announc. Amer. Math. Soc. **8** (2002), 47–51. MR 1945779

Universität Leipzig, Mathematisches Institut
E-mail address: bella@math.uni-leipzig.de

Université d'Aix-Marseille
E-mail address: alberto.chiarini@univ-amu.fr

Max Planck Institute for Mathematics in the Sciences
E-mail address: fehrman@mis.mpg.de