THE CORRECTOR IN STOCHASTIC HOMOGENIZATION: NEAR-OPTIMAL RATES WITH OPTIMAL STOCHASTIC INTEGRABILITY

ANTOINE GLORIA & FELIX OTTO

Abstract. We consider uniformly elliptic coefficient fields that are randomly distributed according to a stationary ensemble of a finite range of dependence. We show that the gradient $\nabla\phi$ of the corrector ϕ, when spatially averaged over a scale $R \gg 1$ decays like $R^{-\alpha}$ for any $\alpha < \frac{d}{2}$. We establish these rates on the level of Gaussian bounds in terms of the stochastic integrability.

Keywords: stochastic homogenization, corrector, quantitative estimate.

2010 Mathematics Subject Classification: 35J15, 35K10, 35B27, 60H25, 60F99.

Contents

Date: November 24, 2015.

1. Introduction and context

We are interested in uniformly elliptic coefficient fields a that are randomly distributed according to a stationary ensemble $\langle\cdot\rangle$ of a finite range of dependence. By the qualitative stochastic homogenization theory of Kozlov [19] and Papanicolaou & Varadhan [24], the behavior of the inverse operator $(-\nabla\cdot a\nabla)^{-1}$ at large scales is described by that of the constant-coefficient operator $(-\nabla\cdot a_{\mathrm{hom}}\nabla)^{-1}$, where a_{hom} are the homogenized coefficients, characterized in direction e by the corrector ϕ, unique (up to additive constant) sublinear solution at infinity of

$$-\nabla\cdot a(\nabla\phi+e)=0 \quad \text{in } \mathbb{R}^d,$$

via the formula

$$a_{\mathrm{hom}}e \;=\; \langle a(\nabla\phi+e)\rangle.$$

The corrector ϕ is the key to quantitative homogenization properties, and the main goal of this contribution is to understand how quantitative ergodic properties of the coefficient field a are transmitted to ϕ. We shall show that the gradient $\nabla\phi$ of the corrector ϕ displays, when spatially averaged, stochastic cancellations at almost the rate as if $\nabla\phi$ was a local function of the coefficient field. More precisely, the spatial average of $\nabla\phi$ on scale R decays like $R^{-\alpha}$ for any $\alpha<\frac{d}{2}$, which amounts almost to Central Limit Theorem (CLT) cancellations. Crucially, we establish these rates on the level of *Gaussian* bounds in terms of the stochastic integrability.

Optimal rates like these in stochastic homogenization have been captured, at first for small ellipticity contrast in the pioneering work by Naddaf and Spencer [23], and later on by Conlon and Naddaf [7] and Conlon and Spencer [8], and more recently in the general case by the authors in [17, 18, 16], and by Neukamm and the authors in [14]. The above cancellations in $\nabla\phi$, even for the exact rate $\alpha=\frac{d}{2}$ have been established in [13]. However, these rates have been only captured with a *suboptimal* integrability, the best being the exponential moments in [13, Theorem 2]. On the other hand, suboptimal rates but with the Gaussian bounds have been established by Armstrong and Smart in [3, Theorem 3.1], and recently by Armstrong, Kuusi, and Mourrat in [1, Theorem 1.1] in a much more quantified way. Hence the merit of the present work is to capture (near-) optimal rates and optimal stochastic integrability at the same time.

This work borrows both in philosophy and tools quite a bit from earlier works. On the one hand, we work under the assumption of *finite range* like Armstrong and Smart in [3] (which was extended to mixing conditions in [2] by Armstrong and

Mourrat in the spirit of the early work [26] by Yurinskii) and just use stochastic cancellations which come from summing independent random variables (rather: spatially averaging stationary random fields that are local). In particular we do not appeal to a more general concentration of measure property, which captures CLT cancellations in random variables of a more complicated structure. Typically these arguments build on an underlying product or Gaussian structure, and pass via assumptions on the ensemble like the Spectral Gap (SG) introduced into stochastic homogenization by Naddaf and Spencer in [23] and extensively used by the authors in [17, 18, 16] and Neukamm and the authors in [14] or the Logarithmic Sobolev Inequality (LSI) introduced by Marahrens and the second author in [20] and refined by Neukamm and the authors in [13], Fischer and the second author in [11], and Duerinckx and the first author in [10].

On the other hand, we completely bypass the variational and subadditive arguments used in [3] (and adapted to a non-symmetric situation in [2]), and thus the related part of the strategy laid out in [1]. Instead we use the parabolic approach as in [14], which yields a convenient disintegration of scales. However, like the pioneering works by Avellaneda & Lin [4] in the periodic and [3] in the random case, we compare the actual solution to the homogenized solution on all scales. When it comes to estimating the homogenization error, we use the tools introduced in [13], namely the vector potential σ for the harmonic coordinates. Here, in order to be able to buckle, we use the "modified" (our language for introducing a massive term as an infra-red regularization) version (ϕ_T, σ_T) of this augmented corrector (ϕ, σ), which necessitates to introduce a third field g_T. Another important technical element in order to get a small *relative* homogenization error is a parabolic version of a novel (deterministic) inner regularity estimate introduced by Bella and the second author in [6, Proof of Theorem 2, Step 6] and refined by Bella, Giunti, and the second author in [5, Lemma 4].

We refer the reader to the introduction of [13] for a more thorough discussion of the literature on quantitative stochastic homogenization of linear elliptic equations in divergence form (and in particular the Lipschitz regularity theory on large scales first introduced in [3]).

The present contribution significantly differs from and completes [13], and we shall conclude this introduction with the relation between these two works. The latter contains three main results. The first result is the development of an intrinsic $C^{1,\alpha}$-regularity theory for random elliptic operators in divergence form at large scales (that is, from a random minimal radius r_* onwards, which is finite almost surely under mere ergodicity of the coefficients), cf. [13, Lemma 2]. This implies in particular a new Liouville theorem for a-harmonic functions that are strictly sub-quadratic at infinity, under the qualitative assumption of ergodicity, cf. [13, Corollary 1]. The range of validity of this intrinsic regularity theory depends on the stochastic integrability of r_*. The second result is the establishment of the optimal stochastic integrability of r_* for a class of correlated Gaussian coefficient fields that display arbitrarily slow algebraically-decaying correlations, based on (a coarsened variant of) LSI, cf. [13, Theorem 1]. The third result of concerns quantitative stochastic homogenization, and more precisely yields quantitative control of the decay of spatial averages of the gradient of the corrector [13, Theorem 2], of the spatial growth of the corrector [13, Theorem 3], and of the quantitative two-scale expansions [13, Corollary 3] for these Gaussian coefficient fields. As mentioned above, under LSI, the rates are optimal, but the stochastic integrability is only exponential whereas we rather expect Gaussian bounds. The proof is based on a sensitivity calculus as in [17, 18, 14, 16] combined with the $C^{1,\alpha}$-regularity theory and the optimal stochastic integrability of r_* (the results are suboptimal in terms of stochastic integrability because the proof does not exploit the potential spatial decorrelations of the stationary field r_*).

Whereas the decay of spatial averages of the gradient of the corrector is obtained as a consequence of the combination of a sensitivity calculus, the regularity theory (in form of the optimal stochastic integrability of the minimal radius r_*), and a LSI in [13], it is considered here as the primary object of interest, which is a twist of point of view. In turn, the stochastic integrability of the minimal radius r_* and the various quantitative estimates obtained in [13] (and also estimates of systematic errors in the spirit of [14]) can be infered from the optimal decay and stochastic integrability of spatial averages of the gradient of the (extended) corrector, which is the driving quantity. The present contribution thus completes the intrinsic regularity theory of [13] and shows that the objects (like the extended corrector), tools, and results introduced and proved there are robust enough

- to treat the case of coefficient fields with a finite range of dependence, and therefore bypass the use of functional inequalities like SG and LSI;
- to prove near-optimal estimates with optimal stochastic integrability.

In the rest of this article, we focus on the control of spatial averages of the extended corrector.

2. Notation, objects, and statement of results

We say that a coefficient field $a = a(x)$ on d-dimensional Euclidean space $\mathbb{R}^d$ is uniformly λ-elliptic provided

$$\xi \cdot a(x)\xi \geq \lambda|\xi|^2 \quad \text{and} \quad \xi \cdot a(x)\xi \geq |a(x)\xi|^2 \tag{1}$$

for all points $x \in \mathbb{R}^d$ and tangent vectors ξ, where the ellipticity ratio $\lambda > 0$ is fixed once for all. In the case of symmetric a, (1) is equivalent to $\lambda|\xi|^2 \leq \xi \cdot a(x)\xi \leq |\xi|^2$; in general, the second condition in (1), which is equivalent to $\xi \cdot a(x)^{-1}\xi \geq |\xi|^2$ and thus also invariant under transposition, yields the more standard upper bound $|a(x)\xi| \leq |\xi|$; but it is in the form of (1) that the constant in the upper bound is preserved under homogenization, see (25) in Lemma 2. While we use scalar notation and language like above, a is allowed to be the coefficient field of an elliptic system. We stress that a needs not be symmetric, in fact, no iota in the proof would change for asymmetric a.

We consider an ensemble of (that is, a probability measure on the space of) λ-uniformly elliptic coefficient fields a and use the physicists notation $\langle\cdot\rangle$ to address both the probability measure and the expectation. We assume that $\langle\cdot\rangle$ is *stationary* in the sense that for any integrable random variable $F = F(a)$ and every shift vector $z \in \mathbb{R}^d$, which acts on coefficient fields via $a(\cdot+z)(x) = a(x+z)$ and thus on random variables via $F^z(a) = F(a(\cdot+z))$, we have $\langle F\rangle = \langle F^z\rangle$. Moreover, we assume that $\langle\cdot\rangle$ has finite range, in fact, has a *unity range of dependence* by which we mean that for two square integrable random variables $F_i = F_i(a)$, $i = 1, 2$, that have the property that F_i depends on a only through $a_{|D_i}$ for two open sets D_i of distance larger than unity, we have $\langle F_1F_2\rangle = \langle F_1\rangle\langle F_2\rangle$. We will be cavalier about measurability issues; however, in case of the qualitative anchoring of our quantitative estimate stated in Lemma 8, it is advantageous to be precise: Taking inspiration from the work of Dal Maso & Modica [9], where the space Ω of λ-uniformly coefficient fields was endowed with the topology coming from Spagnolo's G-convergence (which in their nonlinear context is seen as Γ-convergence), we take the one coming from Murat & Tartar's H- convergence [22] instead, which extends the ideas from [9] to our non-symmetric context and also makes Ω compact. We consider probability measures $\langle\cdot\rangle$ which respect this topology. It turns out that the qualitative homogenization result, even uniformly in stationary ensembles $\langle\cdot\rangle$ of range unity, follows handily thanks to this natural choice of topology, cf Lemma 8.

It is a classical result in qualitative stochastic homogenization that under these conditions (in fact: finite range might be replaced by mere qualitative ergodicity, cf. [19, 24]), for a direction e (a unit vector in $\mathbb{R}^d$), there exists a stationary gradient field $\nabla\phi = \nabla\phi(a, x)$ (where stationary means shift-invariant in the sense of $\nabla\phi(a, x+z) = \nabla\phi(a(\cdot+z), x)$) that is $\langle\cdot\rangle$-square integrable, of vanishing expectation, and such that for $\langle\cdot\rangle$-a. e. realization a we have

$$-\nabla\cdot a(\nabla\phi + e) = 0. \tag{2}$$

Its associated current will play a crucial role

$$q := a(\nabla\phi + e).$$

In view of (2), ϕ itself, which can be constructed as a non-stationary random field that is non-unique up to a random additive constant, might be seen as the correction to the scalar potential of the curl-free harmonic vector field $\nabla\phi + e$, a closed and thus exact 1-form. In view of (2), also q is a closed and thus exact $(d-1)$-form and hence admits a "vector potential", that is, a $(d-2)$-form σ, which can be represented by a skew-symmetric tensor field $\sigma = \{\sigma_{jk}\}_{j,k=1,\cdots,d}$. Clearly, this $(d-2)$ form is non-unique up to a random $(d-3)$-form. The natural choice of gauge is given by

$$-\triangle\sigma_{jk} = \partial_j q_k - \partial_k q_j. \tag{3}$$

It has been recently established in [13, Lemma 1] that under the conditions of stationarity and ergodicity there indeed exists a stationary gradient field $\nabla\sigma$ that is square integrable, of vanishing expectation, such that we have the representation $q-\langle q\rangle = \nabla\cdot\sigma$, and such that (3) holds almost wrt $\langle\cdot\rangle$. Since within the framework of this work, we may recover $(\nabla\phi, q, \nabla\sigma)$ by an approximation via a massive term in (2) and (3), see (20) and (23) (and [16, Lemma 2.7]), which is in fact how we establish the bounds, we ask the reader not to worry about how to construct $(\nabla\phi, q, \nabla\sigma)$ as random objects and how they are uniquely characterized. For simplicity, we don't indicate the — linear — dependence of $(\nabla\phi, \nabla\sigma, q)$ on e by our notation.

The main result is about stochastic cancellations in spatial averages of the triplet $(\nabla\phi, \nabla\sigma, q)$ formed by the gradient of scalar potential (that is, the field), the gradient of vector potential, and the current. It is convenient to define spatial averages f_R on scale R of a field $f = f(x)$ by convolution with the Gaussian $\frac{1}{R^d}\sqrt{2\pi}^{-d}\exp(-\frac{1}{2}|\frac{x}{R}|^2)$. This is convenient mostly because of its connection to the constant-coefficient heat kernel and in particular its semi-group property $(f_R)_r = f_{\sqrt{R^2+r^2}}$. If not specified otherwise, we don't distinguish between $f_R(0)$ and f_R in our notation, which is an acceptable abuse of language since the convolution will only be applied to random fields $f(a, x)$ that are stationary in the above sense, which implies that the distribution of $f_R(y)$ under the stationary $\langle\cdot\rangle$ does not depend on y.

Theorem 1. *Suppose $\langle\cdot\rangle$ is an ensemble of λ-uniformly elliptic coefficient fields which is stationary and of unity range of dependence. Then $(\nabla\phi, \nabla\sigma, q)$, when spatially averaged over a scale R, displays almost CLT-cancellations in terms of Gaussian moments in the sense that for all exponents $\alpha < \frac{d}{2}$ we have*

$$\sup_{\nu} \frac{1}{\nu^2}\Big\langle \exp\big(\nu R^{\alpha}(\nabla\phi, \nabla\sigma, q - \langle q\rangle)_R\big)\Big\rangle \lesssim 1 \quad \textit{for all } R \geq 1, \tag{4}$$

where here and in the proof, $\lesssim$ means $\leq C$ where C is a constant only depending on the dimension d, the ellipticity ratio $\lambda > 0$, and $\alpha < \frac{d}{2}$. Moreover, here and in the sequel $(\nabla\phi, \nabla\sigma, q - \langle q\rangle)$ stands for one of the components, and $|(\nabla\phi, \nabla\sigma, q - \langle q\rangle)|$ for the Euclidean norm of this triplet.

Note that because of $\langle\nabla\phi\rangle = \langle\nabla\sigma\rangle = 0$, these two fields do have not to be recentered. While from the point of view of applications, we are interested in the stochastic cancellations of the augmented corrector $(\nabla\phi, \nabla\sigma)$, the proof rather passes via the field/current pair $(\nabla\phi, q)$; in particular the current q plays a crucial role. However, the proof crucially relies on a version σ_T of the vector potential modified by a massive term.

Let us mention some consequences of Theorem 1 in relation to [13] (the proof of which are left to the reader). First it implies by integration, that the minimal radius r_* introduced in [13, Corallary 2] for ergodic coefficients satisfies for all $\varepsilon > 0$,

$$\langle \exp(r_*^{d(1-\varepsilon)})\rangle \;<\; \infty$$

if the ensemble $\langle\cdot\rangle$ is of finite range of dependence (recovering the corresponding result of [3, Theorem 1.2]), and therefore extends [13, Theorem 1] to this class of coefficients. Second it implies the following control of the growth of the extended corrector: for all $\varepsilon > 0$ there exists $C_\varepsilon < \infty$ and there exists $C < \infty$ such that

$$\left.\begin{array}{ll} d=2 \ : & \left\langle \exp\left(\frac{1}{C_\varepsilon}(1+|x|)^{-\varepsilon}\big((\phi,\sigma)_1(x)-(\phi,\sigma)_1(0)\big)^2\right)\right\rangle \\ d>2 \ : & \left\langle \exp\left(\frac{1}{C}\big((\phi,\sigma)_1(x)-(\phi,\sigma)_1(0)\big)^2\right)\right\rangle \end{array}\right| \le 1,$$

which is new, optimal both in scaling and in stochastic integrability for $d > 2$ (dimension $d = 2$ is critical), and extends [13, Theorem 3] to the case of finite range of dependence; it yields the existence of stationary correctors for $d > 2$ and compares favourably to the suboptimal estimates proved in [1, Theorem 1.2]. These bounds on the extended corrector directly yield a quantitative control of the error in the two-scale expansion, cf. [13, Corollary 3].

We conclude with a consequence of interest of the proof of Theorem 1: nearly-optimal control of the so-called systematic errors. These extend the bounds obtained in [14, Corollary 1 and Lemma 8] in the case of discrete elliptic equations with i.i.d. conductances to the continuum setting of elliptic systems and finite range of dependence, cf. also [16, Proposition 2 and Corollary 2] for similar results up to dimension $d = 4$ (albeit for scalar equations and under a spectral gap assumption).

Theorem 2. *For all $T > 0$ we define the massive approximation ϕ_{Ti} of the corrector ϕ_i in direction e_i as the unique stationary solution with finite second moment of* (20)*. Likewise, we denote by ϕ'_{Tj} the massive approximation of the adjoint corrector ϕ'_j, associated with the pointwise transpose coefficient field a' of a, in direction e_j. For all $\kappa \in \mathbb{N}$ we define the Richardson extrapolation of ϕ_{Ti} wrt T by*

$$\phi^1_{Ti} := \phi_{Ti}, \quad \phi^{\kappa+1}_{Ti} := \frac{1}{2^\kappa - 1}(2^\kappa \phi^\kappa_{2Ti} - \phi^\kappa_{Ti}),$$

and likewise for ϕ'^κ_{Tj}, and we define the approximations a^κ_{hT} and $\tilde{a}^\kappa_{hT}$ of the homogenized coefficients a_{hom} by

$$e_j \cdot a^\kappa_{hT} e_i \ := \ \langle (\nabla \phi'^\kappa_{Tj} + e_j) \cdot a(\nabla \phi^\kappa_{Ti} + e_i)\rangle. \tag{5}$$

If $\langle\cdot\rangle$ has range of dependence unity, then the following estimates of the systematic errors hold true: for all $d \ge 2$, $\kappa > \frac{d}{4}$, and all $\alpha < \frac{d}{2}$,

$$\langle |\nabla \phi^\kappa_{Ti} - \nabla \phi_i|^2\rangle^{\frac12} \ \lesssim \ T^{-\frac{\alpha}{2}}, \tag{6}$$

$$|a^\kappa_{hT} - a_{\mathrm{hom}}| \ \lesssim \ T^{-\alpha}, \tag{7}$$

where the multiplicative constant depends on κ and α next to d and λ.

Recall that stationarity allows one to define a differential calculus in probability through the correspondance for stationary fields ψ (cf. [24, Section 2]):

$$D_i\psi(0) = \lim_{h\downarrow 0} \frac{\psi(a(\cdot + he_i), 0) - \psi(a,0)}{h} = \lim_{h\downarrow 0} \frac{\psi(a, he_i) - \psi(a,0)}{h} = \nabla_i \psi(a,0).$$

This defines a Hilbert space: $\mathcal{H}^1 = \{\psi \in L^2(\langle\cdot\rangle) \,|\, \langle |D\psi|^2\rangle < \infty\}$. In the case when the coefficients a are symmetric, the operator $\mathcal{L} = -D \cdot a(0) D$ defines a quadratic form on $\mathcal{H}^1$. We denote by $\mathcal{L}$ its Friedrichs extension on $L^2(\langle\cdot\rangle)$. Since $\mathcal{L}$ is a self-adjoint non-negative operator, by the spectral theorem, it admits the spectral resolution

$$\mathcal{L} = \int_0^\infty \mu P(d\mu). \tag{8}$$

We obtain as a corollary of Theorem 2 the following bounds on the bottom of the spectrum of $\mathcal{L}$ projected on $D \cdot a(0)e \in (\mathcal{H}^1)'$:

Corollary 1. *Let $\langle\cdot\rangle$ be an ensemble with range of dependence unity that takes values into the set of symmetric coefficient fields, and $\mathfrak{d} := D \cdot a(0)e_i$. Then the spectral resolution P of $\mathcal{L}$ satisfies: for all $\alpha < \frac{d}{2}$ and $\tilde\mu > 0$,*

$$\langle \mathfrak{d} P(d\mu)\mathfrak{d}\rangle([0,\tilde\mu]) \lesssim \tilde\mu^{\alpha+1}, \tag{9}$$

where the multiplicative constant depends on α next to d and λ.

This corollary is a direct extension of [14, Corollary 1] to the continuum setting with finite range of dependence.

3. Strategy of proof and statement of lemmas

The gradient of the corrector $\nabla\phi$ can be (formally) recovered from the parabolic initial value problem

$$\partial_\tau u - \nabla\cdot a\nabla u = 0, \quad u(\tau = 0) = \nabla \cdot ae, \tag{10}$$

as

$$\nabla\phi = \int_0^\infty \nabla u d\tau. \tag{11}$$

Indeed, integration of (10) over $\tau \in (0,\infty)$ with the understanding that $\lim_{\tau\uparrow\infty} \nabla u = 0$ gives (2). Our estimate on $(\nabla\phi, q)$ passes via this representation; for every λ-uniformly elliptic a, there exists a solution of (10) that is unique under mild growth conditions and thus stationary. This parabolic equation, acting as here on stationary fields, can be lifted to probability space and then is the Fokker-Planck equation of the process of the "environment as seen from the particle". It has been the starting point of a quantitative approach to homogenization in [21, 12, 14], where optimal decay rates for algebraic moments have been established in [12, Theorem 7] for low dimensions and in [14, Theorem 1] for all dimensions. The proof of Theorem 1 yields the following by-product, which is at the core of Theorem 2, and quantifies the decay in time of the semi-group in the spirit of [14, Theorem 1].

Lemma 1. *For all $\alpha < \frac{d}{4}$, there exists $C_\alpha < \infty$ such that*

$$\Big\langle \exp\Big(\frac{1}{C_\alpha} T^{1+\alpha} (|\nabla u(T)|^2)^{\frac12}_{\sqrt{2T}}\Big)\Big\rangle \lesssim 1 \quad \textit{for all}\, T \geq 1. \tag{12}$$

In particular,

$$\langle |\nabla u(T)|^2\rangle^{\frac12} \lesssim T^{-1-\alpha} \quad \textit{for all}\, T \geq 1. \tag{13}$$

In this paper, (11) provides a useful disintegration of $\nabla\phi$ into centered stationary random variables $\nabla u(t)$. It is useful because on the one hand, $\nabla u(t)$ is (approximately) local on the scale $\sqrt{t}$ in the sense that it depends on a only through $a_{|B_{M\sqrt{t}}}$ up to exponentially small terms in $M \gg 1$, cf. Definition 2 and Lemma 10, which yields cancellations when spatially averaged, see Lemma 9. On the other hand, for $T \gg t$, $\nabla u(T)$ is close to a spatial average of $\nabla u(t)$ over scales $\sqrt{T}$ provided (10) homogenizes well for sufficiently large t.

Integrating (10) over $\tau \in (0,T)$ we obtain the central formula

$$u(T) = \nabla \cdot q(T) \quad \text{where} \quad q(T) := a(\int_0^T \nabla u d\tau + e). \tag{14}$$

Our strategy will be to propagate in time T stochastic estimates on the field/current pair $(\int_0^T \nabla u d\tau, q(T))$. By stochastic estimate we mean that we capture cancellations of spatial averages $(\int_0^T \nabla u d\tau, q(T))_R$ on the level of Gaussian bounds, see Definition 1. We will propagate these estimates in Lemma 6 on the basis of the decomposition

$$(\int_0^T \nabla u d\tau, q(T))_R = F_0 + F_1, \tag{15}$$

with

$$F_0 := (\int_0^t \nabla u d\tau, q(t))_R + (\mathrm{id}, a_{hT})(\nabla \int_t^T v d\tau)_R, \tag{16}$$

$$F_1 := (\int_t^T \nabla u - \nabla v d\tau, \int_t^T a\nabla u - a_{hT}\nabla v d\tau)_R, \tag{17}$$

where v solves the initial value problem for the constant-coefficient equation

$$\partial_\tau v - \nabla \cdot a_{hT}\nabla v = 0 \ \text{ for } \tau \geq t - r^2. \quad v(t-r^2) = u(t). \tag{18}$$

The time increment $r^2 \leq t$ provides an additional smoothing on scale r. Here the constant coefficient a_{hT} is defined via

$$a_{hT}e_i = \langle q_{Ti}\rangle, \quad i = 1, \cdots, d, \tag{19}$$

(where e_i denotes the i-th unit vector of $\mathbb{R}^d$) by what we call the modified corrector ϕ_{Ti} and its current q_{Ti} which themselves are characterized by the elliptic equation with massive term

$$\frac{1}{T}\phi_{Ti} - \nabla \cdot q_{Ti} = 0 \quad \text{where} \quad q_{Ti} := a(\nabla\phi_{Ti} + e_i). \tag{20}$$

Note that there is a slight inconsistency of notation in (19) and (5) because of the massive term (in particular, $\langle \nabla\phi_{Ti} \cdot a(\nabla\phi_{Ti} + e_i)\rangle = -T^{-1}\langle|\phi_{Ti}|^2\rangle \neq 0$; the approximation defined in (5) is more accurate). Again, for every λ-uniformly elliptic coefficient field a, there exists a solution of (20) that is unique under mild growth conditions and thus stationary. In Lemma 2 we will argue that the tensor a_{hT} defined through (19) is elliptic. For later use, we note that also the *modified* corrector may be recovered from u via $\phi_{Ti} = \int_0^\infty \exp(-\frac{\tau}{T})u d\tau$, provided the direction e in the definition of u in (10) agrees with e_i: Multiplying (10) with $\exp(-\frac{\tau}{T})$ and integrating over $\tau \in (0,\infty)$ yields (20) after an integration by parts. A further integration by parts gives the formula $\phi_{Ti} = \int_0^\infty \frac{1}{T}\exp(-\frac{t}{T})\int_0^t u d\tau dt$. From this together with (14) and (20) we read off the relation between the currents

$$q_T = \int_0^\infty \frac{1}{T}\exp(-\frac{t}{T})q(t)dt. \tag{21}$$

Formula (21) identifies q_T as a convex combination of the $\{q(t)\}_t$, which will be crucial when buckling in Lemma 7.

We will call F_0 the dominant term and F_1 the homogenization error, a deterministic estimate of F_0 is provided by Lemma 3, the deterministic estimate of F_1 is provided by Lemma 4 and the corresponding stochastic estimate in Corollary 2. However, the estimate of F_1 is conditioned on the smallness of the augmented modified corrector $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$, see (23) and (24) below for a definition of $(\frac{\sigma_T}{\sqrt{T}}, g_T)$. The latter is deterministically estimated in Lemma 5 and then stochastically in Corollary 3. All stochastic cancellations follow from taking spatial averages of local fields wrt to a finite range ensemble, see Lemma 9. However, the fields of interest are only approximately local, see Definition 2 and Lemma 10. This only approximate locality is the source of a logarithm in the estimates and ultimately leads to the only near-CLT scaling.

The next element of our strategy consists in a suitable representation of the error in the two-scale expansion based on the modified corrector

$$w := u - (v + \phi_{Ti}\partial_i v), \tag{22}$$

where we use Einstein's convention of summation over repeated indices, where we recall that v is defined in (18). In Lemma 2 we characterize the residuum $\partial_\tau w - \nabla\cdot a\nabla w$. In order to write this residuum (almost) in divergence form, it is necessary to also appeal to a suitable choice of a vector potential σ_{Ti} of the current q_{Ti}, which in the case of general d is written as a skew-symmetric tensor field $\{\sigma_{Tijk}\}_{j,k=1,\cdots,d}$, see our initial discussion of σ_{jk}. The appropriate gauge is given by

$$\frac{1}{T}\sigma_{Tijk} - \triangle\sigma_{Tijk} = \partial_j q_{Tik} - \partial_k q_{Tij} \tag{23}$$

which when compared to (3) also contains a massive term. Because of the massive terms in both (20) and (23), we no longer have $\nabla\cdot\sigma_{Ti} = q_{Ti} - \langle q_{Ti}\rangle$, where the divergence $\nabla\cdot\sigma$ of a tensor field is defined via $(\nabla\cdot\sigma)_j = \partial_k\sigma_{jk}$. To capture the defect in this relation we introduce another auxiliary vector field:

$$g_{Ti} - T\triangle g_{Ti} = q_{Ti} - \langle q_{Ti}\rangle - \nabla\phi_{Ti}. \tag{24}$$

Again, (24) is suitably well-posed in the whole space so that it defines a stationary field g_T. Equipped with these notations, we have

Lemma 2. *The tensor a_{hT} defined through (19) is elliptic in the sense of*

$$\xi\cdot a_{hT}\xi \ge \lambda|\xi|^2 \quad \textit{and} \quad \xi\cdot a_{hT}\xi \ge |a_{hT}\xi|^2. \tag{25}$$

The auxiliary field g_T defined through (24) satisfies

$$q_{Ti} = a_{hT} e_i + \nabla \cdot \sigma_{Ti} + g_{Ti}. \tag{26}$$

The homogenization error F_1 may be expressed in terms of the error w in the two-scale expansion as

$$\nabla u - \nabla v \;=\; \nabla w + \nabla(\phi_{Ti}\partial_i v), \tag{27}$$

$$a\nabla u - a_{hT}\nabla v \;=\; a\nabla w + \nabla\cdot(\partial_i v \sigma_{Ti}) + \partial_i v g_{Ti} + (\phi_{Ti} a - \sigma_{Ti})\nabla\partial_i v. \tag{28}$$

The residuum in the two-scale expansion is given by

$$\begin{aligned}\partial_\tau w - \nabla\cdot a\nabla w = \;& \phi_{Ti}(\frac{1}{T} - \partial_\tau)\partial_i v + g_{Ti}\cdot\nabla\partial_i v \\ & + \nabla\cdot((\phi_{Ti} a - \sigma_{Ti})\nabla\partial_i v),\end{aligned} \tag{29}$$

$$w(t) = u(t) - v(t) - \phi_{Ti}\partial_i v(t). \tag{30}$$

Our strategy relies on good control of the solution v to the constant coefficient initial value value problem (18). On the one hand, in order to capture the stochastic cancellations in the leading order term F_0, cf. (16), we use that its contribution $\int_t^T (\mathrm{id}, a_{hT})(\nabla v)_R d\tau$ is a convolution of the initial data $v(t-r^2) \overset{(18)}{=} u(t) \overset{(14)}{=} \nabla\cdot(q(t) - \langle q(t)\rangle)$. On the other hand, in order to estimate the homogenization error F_1, cf. (17), in view of its representation (29) in Lemma 2 we also need control of higher derivatives $(\nabla^2 v, \nabla\partial_\tau v)$. The following lemma provides both by a pointwise estimate of spatial averages of $(\nabla v, \nabla^2 v, \nabla\partial_\tau v)(t+\tau)$ in terms of spatial averages of $q(t) - \langle q(t)\rangle$.

Lemma 3. *We have for all $\tau \geq 0$ and $R < \infty$*

$$\begin{aligned}&\sum_{k=2}^{3} \big|(\sqrt{\tau + r^2 + R^2}^{k}(\nabla^{k-1} v(t+\tau))_R\big| \\ &\quad + \big|(\tau + r^2 + R^2)^2(\nabla\partial_\tau v(t+\tau))_R\big| \lesssim \big|(q(t) - \langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+r^2+R^2}}\big|_{C\sqrt{\tau+r^2+R^2}},\end{aligned} \tag{31}$$

where here and in the proof C denotes a generic constant only depending on d and $\lambda > 0$.

The upcoming, purely deterministic, Lemma 4 is at the core of our result; through Corollary 2 it establishes that the *relative* homogenization error is small provided the augmented modified corrector $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$ is small. More precisely, the input is that up to time t, the field/current pair displays stochastic cancellations of Gaussian moments when spatially averaged over scales r as encoded in (33). Then the output is that the homogenization error in the field/current pair, when propagating from t to T, is under control in the same way, cf. (34). However, this only holds *conditioned* on the event (35) that the linear growth of the modified augmented corrector $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$ is below a threshold δ. As we learn from (34), in order for the estimate to propagate from t to T, this threshold δ has to compensate a high power of the ratio $\frac{\sqrt{T}}{\sqrt{t}}$ of the time scales $t \ll T$. The form of measuring the stochastic cancellations in terms of the averaging scale (r' in (33) and R in (34)), namely the exponent $\frac{d}{2}+1$ instead of the CLT exponent $\frac{d}{2}$, is dictated by the deterministic estimate of Lemma 4 and has no further deep meaning.

Lemma 4. *For $t \ll T$ and $R \leq \sqrt{T}$ we have for the homogenization error*

$$\begin{aligned}
&(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T |(\nabla u-\nabla v, a\nabla u - a_{hT}\nabla v)_R| d\tau \qquad (32)\\
&\lesssim \ (\frac{\sqrt{T}}{\sqrt{t}})^{\frac{d}{2}+1}\frac{r}{\sqrt{t}}\fint_{\frac{t}{2}}^{t} dt' \fint_0^{\sqrt{t'}} dr' (\frac{r'}{\sqrt{t'}})^{\frac{d}{2}+1}\int \eta_{\sqrt{T}}|(q(t')-\langle q(t')\rangle)_{r'}|\\
&+ \ \Big(\int \eta_{2\sqrt{T}}|(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},g_T)|^2\Big)^{\frac{1}{2}}(\frac{\sqrt{T}}{r})^3\int \eta_{2\sqrt{T}}|(q(t)-\langle q(t)\rangle)_r|,
\end{aligned}$$

where $\eta_R = \frac{1}{R^d}\exp(-\frac{|x|}{R})$ denotes the exponential localization function on scale R and $\lesssim$ means $\leq C$ where C is a generic constant that only depends on λ and d.

The first rhs term of (32) comes from the initial data in (30), the second one from the rhs in (29). Next to Lemma 2 for the representation of the homogenization error in terms of the error w in the two-scale expansion and the representation of the residuum of the latter, and Lemma 3 for the higher-order estimates on the constant-coefficient solution v, Lemma 4 crucially relies on the novel local regularity result Lemma 11 on solutions of $\partial_\tau u - \nabla\cdot a\nabla u = 0$ for the first rhs term.

Corollary 2. *Suppose for some constant $\Lambda < \infty$ we have*

$$\begin{aligned}
\Big\langle \exp\Big(\nu(\frac{r'}{\sqrt{t'}})^{\frac{d}{2}+1}\Big(\int_0^{t'}\nabla u d\tau, q(t')-\langle q(t')\rangle\Big)_{r'}\Big)\Big\rangle \leq \exp(\Lambda\nu^2)\\
\text{for all } \nu\in\mathbb{R},\ r'\leq\sqrt{t'} \text{ and } \frac{t}{2}\leq t'\leq t. \qquad (33)
\end{aligned}$$

Then for any threshold $\delta \ll 1$ we may choose the regularization scale r in (18) such that

$$\begin{aligned}
&\Big\langle I(\mathcal{G}_\delta)\exp\Big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T(\nabla u-\nabla v, a\nabla u - a_{hT}\nabla v)_R d\tau\Big|\Big\rangle\\
&\qquad \lesssim \exp(C\Lambda(\frac{\sqrt{T}}{\sqrt{t}})^{d+4}\delta^{\frac{4}{d+10}}\nu^2) \quad \text{for all } \nu\in\mathbb{R} \text{ and } R\leq\sqrt{T}, \qquad (34)
\end{aligned}$$

where $\mathcal{G}_\delta$ denotes the "good event" of

$$\mathcal{G}_\delta := \{a|\Big(\int \eta_{2\sqrt{T}}|(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},g_T)|^2\Big)^{\frac{1}{2}}\leq\delta\}, \qquad (35)$$

and $I(\mathcal{G}_\delta)$ the associated indicator function.

Corollary 2, through the definition of good event, displays the necessity to control the modified augmented corrector $(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},g_T)$. The next purely deterministic lemma shows that for any pair of times $t \ll T$, *weak* smallness of the current fluctuations $q_t - \langle q_t\rangle$ implies *strong* smallness of $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$. Here, "weak smallness" is quantified by the control of spatial averages on scales slightly smaller than $\sqrt{t}$ (and the expectation thereof), and "strong smallness" refers to the square of $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$ itself, as needed in Lemma 4. It will be crucial for the application in Corollary 3 that the quantified weak smallness enters only once averaged over the large scale $\sqrt{T}$.

Lemma 5. *For $\sqrt{t} \le \delta\sqrt{T}$ and $0 < \delta \le 1$ we have*

$$\begin{aligned}&\Big(\int \eta_{\sqrt{T}r_0}|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2\Big)^{\frac{1}{2}}\\ &\lesssim \frac{1}{\delta^{\frac{d}{2}+7}}\Big(\int \eta_{\sqrt{T}r_0}|(q_t - \langle q_t\rangle)_{\delta\sqrt{t}}| + \langle|(q_t - \langle q_t\rangle)_{\delta\sqrt{t}}|\rangle\Big) + \delta, \end{aligned} \tag{36}$$

where $r_0 \sim 1$ is the radius from the localized elliptic energy estimate (174). Here $A \sim B$ means $A \lesssim B$ and $B \lesssim A$.

While we need Lemma 5 to be quantitative in the sense that we need an algebraic estimate, we don't need to be optimal in terms of the exponents of δ, and make no attempt to be. It makes use of arguments introduced in [13, Section 2.3].

Corollary 2 shows the need to estimate the probability of the "bad event" $\mathcal{G}_\delta^c = \{a|(\int \eta_{2\sqrt{T}}|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2)^{\frac{1}{2}} \ge \delta\}$, cf. (35). Combining the deterministic estimate of the augmented modified corrector $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$ in Lemma 5 in terms of spatial averages $|(q_t - \langle q_t\rangle)_r|$ with stochastic cancellations, cf. Lemma 9, in the centered random variable $|(q_t - \langle q_t\rangle)_r| - \langle|(q_t - \langle q_t\rangle)_r|\rangle$, which is approximately local on scale $\sqrt{t}$ according to Lemma 10, we obtain an estimate on the probability of $\mathcal{G}_\delta^c$: It is exponentially small in the ratio of the time scales $t \ll T$ with a scaling that is, up to a logarithm, in agreement with CLT scaling. However, it only holds provided t is so large that some stochastic cancellations did occur in the sense that the flux q_t of the modified corrector is spatially weakly close to its expectation $\langle q_t\rangle$, cf. (38). The latter is mildly quantified in the sense that the spatial averages monitoring the weak closeness are only slightly smaller than $\sqrt{t}$, namely $\delta\sqrt{t}$ and that closeness is requires only to first moment in probability. Corollary 3 captures the first genuine stochastic effect in the flow of lemmas.

Corollary 3. *For any threshold $\delta \le 1$ and time T we have*

$$\log\Big\langle I\Big(\big(\int \eta_{2\sqrt{T}}|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2\big)^{\frac{1}{2}} \ge \delta\Big)\Big\rangle \lesssim -(\frac{\sqrt{T}}{\sqrt{t}}\frac{1}{\log(e + \frac{\sqrt{T}}{\sqrt{t}})})^d\delta^{d+16}, \tag{37}$$

provided $t \le T$ and δ are such that

$$\langle|(q_t - \langle q_t\rangle)_{\delta\sqrt{t}}|\rangle \ll \delta^{\frac{d}{2}+8}. \tag{38}$$

Equipped with Corollaries 2 & 3 we are now in the position to propagate forward in time Gaussian bounds on spatial averages $(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_r$, $r \le \sqrt{t}$, of the field/current pair $(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)$. The form of Corollary 2 indicates that we cannot hope at this stage to capture CLT cancellations in terms of the averaging scale r. Following (33) in the statement of Corollary 2, we instead monitor non-CLT cancellations, however on the level of Gaussian bounds, through

Definition 1.

$$\Lambda(t) := \sup_{r\le\sqrt{t},\nu} \frac{1}{\nu^2} \log \big\langle \exp\big(\nu(\frac{r}{\sqrt{t}})^{\frac{d}{2}+1}(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_r\big)\big\rangle. \tag{39}$$

The central ingredient of this paper is the following propagation estimate which involves three times, t_0 next to $t \le T$.

Lemma 6. *For any triplet of times $t_0, t \le T$ and any threshold $\delta \ll 1$ we have*

$$\Lambda(T) \lesssim (\frac{\sqrt{t}}{\sqrt{T}})^d \Lambda(t) \log^d(e + \frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d) + \delta^{\frac{4}{d+10}}(\frac{\sqrt{T}}{\sqrt{t}})^{d+4} \sup_{t'\in[\frac{t}{2},t]} \Lambda(t') + \frac{1}{\delta^{d+16}}(\frac{\sqrt{t_0}}{\sqrt{T}})^d \log^d(e + \frac{\sqrt{T}}{\sqrt{t_0}}), \tag{40}$$

provided that we are in the regime where t_0 is so large that some stochastic cancellations have set in in the sense of

$$\langle|(q_{t_0} - \langle q_{t_0}\rangle)_{\delta\sqrt{t_0}}|\rangle \ll \delta^{\frac{d}{2}+8}. \tag{41}$$

The first rhs term in (40) comes from the dominant term F_0, cf. (16), via Lemma 3 and stochastic cancellations; the second from the homogenization error F_1, cf. (17), via Corollary 2; and the third one from the modified augmented corrector $(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)$ via Corollary 3.

The following lemma establishes near-optimal bounds on $\{\Lambda(T)\}_T$ introduced in Definition 1, building on Lemma 6. It involves a time t_0 which has to be so large that stochastic cancellations have already set in on the associated scale $\sqrt{t_0}$; hence it is conditional on the anchoring provided by Lemma 8 below. The core of the proof of Lemma 7 is a Campanato iteration based on the propagation rule (40).

Lemma 7. *For any exponent $\alpha < \frac{d}{2}$ we have for any time $t_0 \gg 1$ the implication*

$$\sup_{t\ge t_0} \Lambda(t) \ll 1 \quad \Longrightarrow \quad \forall\, t \ge t_0 \;\; \Lambda(t) \lesssim (\frac{t_0}{t})^\alpha, \tag{42}$$

where like in Theorem 1, $A \lesssim B$ means $A \le C(d,\lambda,\alpha)B$ and $A \ll 1$ means that there exists a $C(d,\lambda,\alpha)$ such that $A \le \frac{1}{C}$.

In the following lemma, we show that qualitative homogenization yields that $\Lambda(T)$ converges to zero as $T \uparrow \infty$, which we will provides the "anchoring" for Lemma 7.

Lemma 8. *We have*

$$\Lambda(T) \lesssim 1 \quad \textit{and} \quad \lim_{T\uparrow\infty} \Lambda(T) = 0, \tag{43}$$

where the limit is uniform up to a dependence on the dimension d and the ellipticity ratio $\lambda > 0$.

Based on the anchoring in Lemma 8 and the buckling in Lemma 7, we may now upgrade the ensuing near-optimal decay of $\Lambda(T)$ in T, which however encodes only non-CLT decay of averages of $(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_R$ in terms of the averaging scale R, cf. Definition 1, to the near optimal decay of $(\nabla\phi, q - \langle q\rangle)_R$ in the averaging scale R in a first step, and then also of $(\nabla\sigma)_R$ in a second step. The main ingredient in this upgrade is deterministic and consists of the two local regularity results on solutions of $\partial_\tau u - \nabla\cdot a\nabla u = 0$ provided by the novel Lemma 11 and the more standard Lemma 12. The second of these lemmas provides a quantified equi-integrability of $|\nabla u|^2$, while the first one controls this quantity in terms of (the square of) a very weak norm of u.

We now come to the auxiliary lemmas needed in the above flow of the proof. The following lemma provides the only genuine stochastic ingredient, namely CLT cancellations for spatial averages of centered (ie of vanishing expectation) stationary fields q that are local. It is the only place where we use that the ensemble $\langle\cdot\rangle$ has finite (in fact, unity) range. The CLT cancellations degrade by a logarithm due to the only approximate locality, cf. Definition 2, which is the form of locality satisfied by the random variables of interest, cf. Lemma 10.

Definition 2. *We say that a random variable g is approximately local on scale $\sqrt{T}$ if and only if there exists a constant $C = C(d, \lambda)$ such that we have*

$$|g(a) - g(\tilde a)| \lesssim \exp(-\frac{1}{C}\frac{R}{\sqrt{T}}) \quad \textit{provided } a = \tilde a \textit{ on } B_R \quad \textit{for all } R \geq \sqrt{T}, \tag{44}$$

When applied to a stationary field $g = g(a, x)$, we mean $g(a) = g(a, x = 0)$ and $B_R = B_R(0)$ in (49).

Lemma 9. *Let g denote a centered stationary field. We assume that its convolution $g_{\sqrt{t}}$ on scale $\sqrt{t} \geq 1$ is approximately local on scale $\sqrt{t}$ in the sense of Definition 2, that is,*

$$|g_{\sqrt{t}}(a) - g_{\sqrt{t}}(\tilde a)| \leq \exp(-\frac{r}{\sqrt{t}}) \quad \textit{provided } a = \tilde a \textit{ in } B_r \;\; \textit{for all } r \geq \sqrt{t}, \tag{45}$$

and that it has Gaussian bounds of the form

$$\log\langle\exp(\nu g_{\sqrt{t}})\rangle \leq \Lambda\nu^2 \quad \textit{for all } \nu \tag{46}$$

for some $\Lambda \leq 1$. Then the spatial averages g_R on larger scales $R \geq \sqrt{t}$ display CLT cancellations (up to a logarithm) on the level of Gaussian bounds in the sense of

$$\log\langle\exp(\nu g_R)\rangle \lesssim \frac{\Lambda\sqrt{t}^d}{R^d}(\log(e + \frac{R^d}{\Lambda\sqrt{t}^d}))^d\nu^2 \quad \textit{for all } \nu \textit{ and } R \geq \sqrt{t}. \tag{47}$$

We need two locality statements: 1) for the solution u of the parabolic equation (10), or rather its field/current pair $(\int_0^T \nabla u d\tau, q(T))$, cf. (14), which we need for the dominant term F_0 and the homogenization error F_1 in the proof of Lemma 7, and 2) for the modified corrector ϕ_T, or rather its current q_T, cf. (20), which we need for the estimate of the augmented modified corrector in the proof of Corollary 3. The upcoming lemma also contains a statement on uniform bounds which come up in its proof and which are needed at several places.

Lemma 10. *We have in the sense of Definition 2*

$$(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_{\sqrt{T}} \quad \textit{is approx. local on scale } \sqrt{T}, \tag{48}$$

$$\Big(\int \eta_{\sqrt{T}} |(\int_0^T \nabla u d\tau, q(T))|^2\Big)^{\frac{1}{2}} \lesssim 1, \tag{49}$$

$$\big(|(q_T - \langle q_T\rangle)_r| - \langle |(q_T - \langle q_T\rangle)_r|\rangle\big)_{\sqrt{T}} \quad \textit{is approx. local on scale } \sqrt{T}, \tag{50}$$

$$\Big(\int \eta_{\sqrt{T}} |q_T|^2\Big)^{\frac{1}{2}}, \big(|(q_T - \langle q_T\rangle)_r| - \langle |(q_T - \langle q_T\rangle)_r|\rangle\big)_{\sqrt{T}} \lesssim 1, \tag{51}$$

for arbitrary $r \le \sqrt{T}$.

We finally state the two inner regularity results for solutions of $\partial_t u - \nabla \cdot a \nabla u$ needed in the proof of Theorem 1, and in particular in the proof of Lemma 4.

Lemma 11. *For a solution of the homogeneous parabolic equation* $\partial_t u - \nabla \cdot a \nabla u = 0$ *for* $t \ge 0$ *we have for any exponent* $p < \infty$ *and time* T

$$\Big(\int \eta_{\sqrt{T}} |\sqrt{T} \nabla u(T)|^2\Big)^{\frac{1}{2}} \lesssim \fint_0^T dt \fint_0^{\sqrt{T}} dr (\frac{r}{\sqrt{T}})^p \int \eta_{2\sqrt{T}} |u_r(t)|, \tag{52}$$

where here and in the proof, $\lesssim$ *refers to* $\le C$, *where* C *denotes a generic constant that only depends on* d, $\lambda > 0$, *and on* $p < \infty$.

While it is well-known that the lhs of (52) is estimated by the (suitably localized) H^{-1}-norm of $u(t = 0)$, the rhs of (52) can be assimilated with much weaker norm (for p large) and moreover is an L^1-average in x instead of an L^2-average. Estimate (52) can be seen as a parabolic version of [5, Lemma 4]. The argument however is rather different since it is more about localization in time than in space.

The following quantified equi-integrability for $|\nabla u|^2$ in space is more standard and follows with help of the classical Meyers estimate, which has been extensively used in earlier work on quantitative stochastic homogenization (see [23] for its first use, in the elliptic case, and [14] for a use in the parabolic case).

Lemma 12. *There exists an exponent* $\varepsilon = \varepsilon(d, \lambda) > 0$ *(possibly very small) such that for any solution of* $\partial_\tau u - \nabla \cdot a \nabla u = 0$ *for* $\tau \ge 0$ *we have for all* $R \le \sqrt{T}$

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2} - \varepsilon} \Big(\int_{\frac{T}{2}}^T \int \eta_R |\nabla u|^2 d\tau\Big)^{\frac{1}{2}} \lesssim \Big(\int_0^T \int \eta_{\sqrt{T}} |\nabla u|^2 d\tau\Big)^{\frac{1}{2}}. \tag{53}$$

The proof uses many of the arguments from [14].

4. Proof of the main results

The proofs of the main results refer to some arguments in the proofs of the auxiliary results. These references are however kept minimal.

4.1. **Proof of Theorem 1: decay of averages of the extended corrector.** We fix an exponent $\alpha < \frac{d}{2}$ with the understanding that $\frac{d}{2} - \alpha \ll 1$. By Lemma 7 in conjunction with Lemma 8 we have

$$\sup_\nu \frac{1}{\nu^2} \log \big\langle \exp\big(\nu(\frac{r}{\sqrt{t}})^{\frac{d}{2}+1}(\int_0^t \nabla u d\tau, q(t) - \langle q(t) \rangle)_r\big)\big\rangle \lesssim \frac{1}{t^\alpha} \quad \text{for } r \leq \sqrt{t}. \tag{54}$$

Based on the representation $\nabla\phi = \int_0^\infty \nabla u d\tau$, cf. (11), and the definition $q(t) = a(\int_0^t \nabla u d\tau + e)$, cf. (14), and given an averaging radius R, we split the field/current pair of the corrector according to

$$(\nabla\phi, q)_R = \sum_{n=0}^{\infty} F_n, \quad \text{where} \tag{55}$$

$$F_0 := \big(\int_0^{R^2} \nabla u d\tau, q(R^2)\big)_R \quad \text{and} \quad F_n := \big((\mathrm{id}, a)\int_{2^{n-1}R^2}^{2^n R^2} \nabla u d\tau\big)_R \quad \text{for } n \geq 1.$$

The first term F_0 is directly estimated by (54), which we use for $t = R^2$ and $r = R = \sqrt{t}$ and then redefine ν:

$$\frac{1}{\nu^2} \log \big\langle \exp\big(\nu R^\alpha (F_0 - \langle F_0 \rangle)\big)\big\rangle \lesssim 1. \tag{56}$$

We now turn to F_n for an arbitrary yet fixed $n \in \mathbb{N}$; the main step is a deterministic estimate of F_n based on the two inner regularity results on solutions of the homogeneous parabolic equation $\partial_\tau u - \nabla \cdot a\nabla u = 0$ stated in Lemma 12 and Lemma 11, respectively. We set for abbreviation $T := 2^{n-1}R^2$, so that we have $R \leq \sqrt{T}$ and $F_n = \int_T^{2T}((\mathrm{id}, a)\nabla u)_R d\tau$. We start with Cauchy-Schwarz' inequality in t, Jensen's inequality in x, and the fact that the exponential average $\int \eta_R f$ dominates the Gaussian one f_R:

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}|F_n| \lesssim (\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}\big(T\int_T^{2T}\int \eta_R |\nabla u|^2 d\tau\big)^{\frac{1}{2}}.$$

We then appeal to the inner regularity estimate (53) in Lemma 12 (splitting our time interval $(T, 2T)$ into $(\frac{3}{2}T, 2T)$ and $(T, \frac{3}{2}T)$ and shifting them by T and $\frac{T}{2}$, respectively) to obtain

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}|F_n| \lesssim \big(T\int_{\frac{T}{2}}^{2T}\int \eta_{\sqrt{T}} |\nabla u|^2 d\tau\big)^{\frac{1}{2}}.$$

The localized higher-order parabolic energy estimate, see (**??**) in the proof of Lemma 10 with $f=0$, in its un-rescaled version

$$\frac{d}{dt}\int \eta_{\sqrt{T}}\nabla u\cdot a\nabla u \lesssim \frac{1}{T}\int \eta_{\sqrt{T}}\nabla u\cdot a\nabla u$$

shows that $t\mapsto \exp(-c\frac{t}{T})\int\eta_{\sqrt{T}}\nabla u\cdot a\nabla u$ is non-increasing, so that

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}|F_n|\lesssim \big(\int \eta_{\sqrt{T}}|T\nabla u(\tau=\frac{T}{2})|^2\big)^{\frac12}.$$

We then appeal to the inner regularity estimate (52) in Lemma 11 (with the time interval $(0,T)$ rescaled and shifted to $(\frac{T}{4},\frac{T}{2})$ and $p=\frac{d}{2}+2$) which now gives

$$\begin{aligned}
&\big(\int \eta_{\sqrt{T/4}}|T\nabla u(\tau=\frac{T}{2})|^2\big)^{\frac12}\\
&\lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}}dt\fint_0^{\sqrt{T/4}}dr(\frac{r}{\sqrt{T/4}})^{\frac{d}{2}+2}\int \eta_{2\sqrt{T/4}}|\sqrt{T}u_r(t)|\\
&\lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}}dt\fint_0^{\sqrt{t}}dr(\frac{r}{\sqrt{t}})^{\frac{d}{2}+2}\int \eta_{\sqrt{T}}|\sqrt{T}u_r(t)|.
\end{aligned}$$

Based on (82) - (84) in the proof of Lemma 4 and the arguments there, this yields

$$\begin{aligned}
(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}|F_n|&\lesssim \big(\int \eta_{\sqrt{T}}|T\nabla u(\tau=\frac{T}{2})|^2\big)^{\frac12}\\
&\lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}}dt\fint_0^{\sqrt{t}}dr(\frac{r}{\sqrt{t}})^{\frac{d}{2}+2}\int \eta_{2\sqrt{T}}|\sqrt{T}u_r(t)|. \qquad (57)
\end{aligned}$$

We finally appeal to (14) in form of $u(t)=\nabla\cdot(q(t)-\langle q(t)\rangle)$ and to $|(\nabla f)_r| = |(\nabla f_{\frac{1}{\sqrt{2}}r})_{\frac{1}{\sqrt{2}}r}|\lesssim\frac{1}{r}|f_{\frac{1}{\sqrt{2}}r}|_r$ as well as $\int\eta_{\sqrt{T}}f_r\overset{(79)}{\lesssim}\int\eta_{\sqrt{T}}f$ for $r\lesssim\sqrt{T}$ to convert the above to

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}|F_n|\lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}}dt\fint_0^{\sqrt{t}}dr(\frac{r}{\sqrt{t}})^{\frac{d}{2}+1}\int \eta_{2\sqrt{T}}|(q(t)-\langle q(t)\rangle)_r|, \qquad (58)$$

where we replaced $\frac{1}{\sqrt{2}}r$ by r on the rhs by a change of variables at the expense of a further factor $\sqrt{2}^{\frac{d}{2}+2}$. Estimate (58) is our key deterministic estimate.

We now continue with the stochastic estimate: By convexity of $F\mapsto\langle\exp(F)\rangle$ and stationarity of $(q(t)-\langle q(t)\rangle)_r$, (58) yields

$$\begin{aligned}
&\big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}F_n\big|\big\rangle\\
&\lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}}dt\fint_0^{\sqrt{t}}dr\big\langle \exp\big(C|\nu(\frac{r}{\sqrt{t}})^{\frac{d}{2}+1}(q(t)-\langle q(t)\rangle)_r|\big)\big\rangle.
\end{aligned}$$

Appealing to $\exp|F| \le \exp(F)+\exp(-F)$, we now may insert (54)

$$\big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}-\varepsilon}F_n\big|\big\rangle \lesssim \fint_{\frac{T}{4}}^{\frac{T}{2}} dt\exp(C\frac{1}{t^\alpha}\nu^2) \le \exp(C\frac{1}{T^\alpha}\nu^2).$$

Redefining ν and using $\frac{\sqrt{T}}{R}=\sqrt{2}^{n-1}$, we rewrite this

$$\big\langle \exp\big|\nu(\sqrt{2}^{\alpha+\varepsilon-\frac{d}{2}})^n R^\alpha F_n\big|\big\rangle \lesssim \exp(C\nu^2).$$

With the same argument as at the end of the proof of Lemma 6 in order to pass from F_n to its centered (ie of vanishing expectation) version $F_n-\langle F_n\rangle$ we obtain from the above

$$\frac{1}{\nu^2}\log\big\langle \exp\big(\nu(\sqrt{2}^{\alpha+\varepsilon-\frac{d}{2}})^n R^\alpha(F_n-\langle F_n\rangle)\big)\big\rangle \lesssim 1. \tag{59}$$

Taking into account (56), we see that (59) holds for $n=0$ next to $n\in\mathbb{N}$. Wlog we may assume that α is so close to $\frac{d}{2}$ that $\alpha+\varepsilon-\frac{d}{2}>0$. Appealing to (55), we write

$$(\nabla\phi, q-\langle q\rangle)_R = \sum_{n=0}^{\infty}\omega_n(\sqrt{2}^{\alpha+\varepsilon-\frac{d}{2}})^n(F_n-\langle F_n\rangle), \tag{60}$$

where thanks to $\alpha+\varepsilon-\frac{d}{2}>0$, the discrete weight function $\{\omega_n := (\sqrt{2}^{\alpha+\varepsilon-\frac{d}{2}})^{-n}\}_n$ is summable $\sum_{n=0}^\infty \omega_n \sim 1$ and thus defines a convex combination after a modifying it by a multiplicative constant C. Hence (59) yields by convexity of $F\mapsto \log\langle\exp(F)\rangle$

$$\frac{1}{\nu^2}\log\big\langle \exp\big(\nu R^\alpha(\nabla\phi, q-\langle q\rangle)_R\big)\big\rangle \lesssim \exp(C\nu^2),$$

that is, (4) for the field/current pair $(\nabla\phi, q)$.

We finally turn to the gradient $\nabla\sigma$ of the vector potential σ. To avoid the use of indices we adopt the notation for $d=3$ and rewrite (3) as

$$-\triangle\sigma = \nabla\times q = \nabla\times(q-\langle q\rangle).$$

In order to leverage prior arguments as much as possible, we "disintegrate" $\nabla\sigma$ as we did for $\nabla\phi$, by formally writing $\nabla\sigma=\int_0^\infty \nabla v d\tau$, where v solves the constant-coefficient initial value problem

$$\partial_\tau v-\triangle v=0,\quad v(\tau=0)=\nabla\times(q-\langle q\rangle)$$

and thus has the following representation in terms of the heat kernel, that is, the convolution with our isotropic Gaussian:

$$\nabla v(\tau) = \nabla\nabla\times(q-\langle q\rangle)_{\sqrt{2\tau}}.$$

Proceeding now as in the proof of Lemma 3, we make use of the semi-group property to distribute the convolutions with Gaussians as follows

$$(\nabla v(\tau))_R = (\nabla\nabla\times(q-\langle q\rangle)_{\sqrt{\tau+\frac{R^2}{2}}})_{\sqrt{\tau+\frac{R^2}{2}}},$$

which yields the estimate

$$|(\nabla v(\tau))_R| \lesssim \frac{1}{\tau+R^2}|(q-\langle q\rangle)_{\sqrt{\tau+\frac{R}{2}}}|_{\sqrt{2\tau+R}}.$$

In preparation of stochastic estimate, following the proof of Lemma 6, we rewrite this as

$$R^\alpha|(\nabla v(\tau))_R| \lesssim \omega_R(\tau)|\sqrt{\tau+\tfrac{R^2}{2}}^{\alpha}(q-\langle q\rangle)_{\sqrt{\tau+\frac{R^2}{2}}}|_{\sqrt{2\tau+R^2}},$$

where $\omega_R(\tau) := \frac{R^\alpha}{\sqrt{\tau+R^2}^{\alpha+2}}$ is a weight function in the sense of $\int_0^\infty \omega_R d\tau \sim 1$ thanks to $\alpha > 0$. Based on $|(\nabla\sigma)_R| \leq \int_0^\infty |(\nabla v)_R| d\tau$ and by convexity of $F \mapsto \langle \exp|F|\rangle$ and stationarity of $q-\langle q\rangle$ this yields

$$\Big\langle \exp\big|\nu R^\alpha(\nabla\sigma)_R\big|\Big\rangle \lesssim \int_0^\infty d\tau\omega_R\Big\langle \exp\big|\nu\sqrt{\tau+\tfrac{R^2}{2}}^{\alpha}(q-\langle q\rangle)_{\sqrt{\tau+\frac{R^2}{2}}}\big|\Big\rangle.$$

With help of $\exp|F| \leq \exp(F)+\exp(-F)$ we may appeal to (4) for the current q with R replaced by $\sqrt{\tau+\frac{R}{2}}$:

$$\Big\langle \exp\big|\nu R^\alpha(\nabla\sigma)_R\big|\Big\rangle \lesssim \exp(C\nu^2).$$

With the same argument as at the end of the proof of Lemma 6 in order to pass from $|(\nabla\sigma)_R|$ to the centered (ie of vanishing expectation) $(\nabla\sigma)_R$, we obtain (4) also for the vector potential σ.

4.2. **Proof of Lemma 1: decay of the semi-group.** From (57) and (58) in the proof of Theorem 1, we learn that

$$\Big(\int \eta_{\sqrt{2T}}|T\nabla u(T)|^2\Big)^{\frac12} \lesssim \fint_{\frac{T}{2}}^{T} dt \fint_0^{\sqrt{t}} dr\Big(\frac{r}{\sqrt{t}}\Big)^{\frac{d}{2}+1}\int \eta_{2\sqrt{2T}}|(q(t)-\langle q(t)\rangle)_r|,$$

so that (54) yields for all $\alpha < \frac{d}{2}$ and $\nu \in \mathbb{R}$,

$$\Big\langle \exp\Big|\nu\Big(\int \eta_{\sqrt{2T}}|T\nabla u(T)|^2\Big)^{\frac12}\Big|\Big\rangle \lesssim \fint_{\frac{T}{2}}^{T} dt\exp\Big(C\frac{1}{t^\alpha}\nu^2\Big) \leq \exp\Big(C\frac{1}{T^\alpha}\nu^2\Big),$$

from which (12) follows.

Estimate (13) is an immediate consequence of (12) combined with the stationarity of $|\nabla u(T,\cdot)|^2$: for all $\alpha < \frac{d}{4}$, and $T \gg 1$,

$$\langle|\nabla u(T)|^2\rangle = \Big\langle\int \eta_{\sqrt{2T}}|\nabla u(T)|^2\Big\rangle \overset{(12)}{\lesssim} T^{-1-\alpha}.$$

4.3. **Proof of Theorem 2: systematic errors.** We split the proof into two steps.

Step 1. Proof of (6).
The starting point is the representation of $\nabla\phi_T^\kappa$ and $\nabla\phi$ as

$$\nabla\phi = \int_0^\infty \nabla u(\tau)d\tau, \qquad \nabla\phi_T^\kappa = \int_0^\infty \exp_\kappa(\tau,T)\nabla u(\tau)d\tau,$$

where $\exp_\kappa(\cdot,T)$ is defined as the Richardson extrapolation of $\exp_1(\tau,T) := \exp(-\frac{\tau}{T})$ wrt T. The extrapolation has the effect that

$$|1-\exp_\kappa(\tau,T)| \lesssim \min\{(\frac{\tau}{T})^\kappa,1\}, \qquad |\frac{\partial}{\partial_\tau}\exp_\kappa(\tau,T)| \lesssim \frac{1}{T}(\frac{\tau}{T})^{\kappa-1} \quad \text{for all } \tau\geq 0. \tag{61}$$

We then split the integral over $(0,+\infty)$ into three contributions. We start with the contribution on the interval $(0,1)$. By integration by parts,

$$\int_0^1(1-\exp_\kappa(\tau,T))\nabla u(\tau)d\tau = \int_0^1\frac{\partial}{\partial\tau}\exp_\kappa(\tau,T)\int_0^\tau\nabla u(t)dtd\tau + (1-\exp_\kappa(1,T))\int_0^1\nabla u(t)dt.$$

By Cauchy-Schwarz' inequality in the first time integral and in probability, and by stationarity of $\int_0^1\nabla u(t)dt$,

$$\Big\langle\Big(\int_0^1(1-\exp_\kappa(\tau,T))\nabla u(\tau)d\tau\Big)^2\Big\rangle \lesssim \int_0^1|\frac{\partial}{\partial\tau}\exp_\kappa(\tau,T)|^2d\tau\times\int_0^1\Big\langle\int\eta_{\sqrt{\tau}}|\int_0^\tau\nabla u(t)dt|^2\Big\rangle d\tau + |1-\exp_\kappa(1,T)|^2\Big\langle\int\eta|\int_0^1\nabla u(t)dt|^2\Big\rangle.$$

Hence, (49) in Lemma 10 and (61) combine to

$$\Big\langle\Big(\int_0^1(1-\exp_\kappa(\tau,T))\nabla u(\tau)d\tau\Big)^2\Big\rangle \lesssim T^{-2\kappa},$$

which is of higher order than the rhs of (6). We turn now to the contributions on the intervals $(1,T)$ and (T,∞), for which the estimate of the decay of the semi-group (13) in Lemma 1 and (61) for $\kappa>\frac{d}{4}$ yield for all $\alpha<\frac{d}{2}$

$$\begin{aligned}\Big\langle\Big(\int_1^\infty(1-\exp_\kappa(\tau,T))\nabla u(\tau)d\tau\Big)^2\Big\rangle^{\frac12} &\lesssim \int_1^T(\frac{\tau}{T})^\kappa\tau^{-1-\frac{\alpha}{2}}d\tau+\int_T^\infty\tau^{-1-\frac{\alpha}{2}}d\tau\\ &\lesssim T^{-\frac{\alpha}{2}}.\end{aligned}$$

This proves (6).

Step 1. Proof of (7).

This estimate is a direct consequence of (6). By definitions (5) of a^κ_{hT}, and the following definition of $a_{\rm hom}$,

$$e_j\cdot a_{\rm hom}e_i \;=\; e_j\cdot\langle a(\nabla\phi_i+e_i)\rangle \;=\; \langle(\nabla\phi_j'+e_j)\cdot a(\nabla\phi_i+e_i)\rangle,$$

where ϕ_i is solution of (2) for $e=e_i$ (and ϕ_j the solution of (2) for $e=e_j$ and a replaced by its pointwise transpose a') we have

$$e_j\cdot(a^\kappa_{hT}-a_{\rm hom})e_i \;=\; \langle(\nabla\phi'^\kappa_{Tj}-\nabla\phi_j')\cdot a(\nabla\phi^\kappa_{Ti}+e_i)\rangle-\langle(\nabla\phi_j'+e_j)\cdot a(\nabla\phi_i-\nabla\phi^\kappa_{Ti})\rangle.$$

Since

$$\langle(\nabla\phi_j'+e_j)\cdot a(\nabla\phi_i-\nabla\phi^\kappa_{Ti})\rangle=\langle(\nabla\phi_i-\nabla\phi^\kappa_{Ti})\cdot a'(\nabla\phi_j'+e_j)\rangle,$$

the weak form of the corrector equation for ϕ_j' in probability and for ϕ_i then yields (whence the choice of the adjoint corrector in the definition of a^κ_{hT})

$$\langle(\nabla\phi_i-\nabla\phi^\kappa_{Ti})\cdot a'(\nabla\phi_j'+e_j)\rangle \;=\; 0 \;=\langle(\nabla\phi_j'-\nabla\phi'^\kappa_{Tj})\cdot a(\nabla\phi_i+e_i)\rangle,$$

and we conclude

$$e_j\cdot(a^\kappa_{hT}-a_{\rm hom})e_i \;=\; \langle(\nabla\phi'^\kappa_{Tj}-\nabla\phi_j')\cdot a(\nabla\phi^\kappa_{Ti}-\nabla\phi_i)\rangle,$$

so that the claim follows from (6) (used for both a and a').

4.4. **Proof of Corollary 1: thickness of the bottom of the spectrum.** For $\tilde\mu\ge 1$, there is nothing to prove and we assume $T=\frac{1}{\tilde\mu}\ge 1$. The starting point is the spectral theorem which allows one to rewrite the definition of ϕ^κ_{Ti} in the form

$$\phi^\kappa_{Ti} \;=\; g_\kappa(\mathcal{L},T)\mathfrak{d},$$

where $g_1(\mu,T)=(T^{-1}+\mu)^{-1}$, and g_κ is the Richardson extrapolation of g_1 wrt T. Set $g_0(\mu)=\frac{1}{\mu}$. Then, by the spectral theorem,

$$\begin{aligned}\langle\nabla(\phi^\kappa_{Ti}-\phi_i)\cdot a\nabla(\phi^\kappa_{Ti}-\phi_i)\rangle &= \langle(\phi^\kappa_{Ti}-\phi_i)\mathcal{L}(\phi^\kappa_{Ti}-\phi_i)\rangle\\ &= \int_0^\infty \mu(g_\kappa(\mu,T)-g_0(\mu))^2\langle\mathfrak{d}P(d\mu)\mathfrak{d}\rangle.\end{aligned}$$

On the one hand, for $\kappa>\frac{d}{4}$, Theorem 2 yields for all $\alpha<\frac{d}{2}$

$$\langle\nabla(\phi^\kappa_{Ti}-\phi_i)\cdot a\nabla(\phi^\kappa_{Ti}-\phi_i)\rangle \;\le\; \langle|\nabla(\phi^\kappa_{Ti}-\phi_i)|^2\rangle \;\lesssim\; T^{-\alpha}.$$

On the other hand, by induction on κ (see for instance [15, Proof of Lemma 2.5, Step 2]) we have

$$|g_\kappa(\mu,T)-g_0(\mu)| \;\gtrsim\; \frac{T^{-\kappa}}{\mu(T^{-1}+\mu)^\kappa},$$

which we use in the form: for all $\mu\le\frac{1}{T}$,

$$\mu(g_\kappa(\mu,T)-g_0(\mu))^2 \;\gtrsim\; \frac{T^{-2\kappa}}{\mu(T^{-1}+\mu)^{2\kappa}}.$$

The combination of these two estimates directly yields the claim, recalling that $T^{-1} = \tilde{\mu}$,

$$\begin{aligned}\int_0^{\frac{1}{T}} \langle \mathfrak{d} P(d\mu)\mathfrak{d}\rangle &\lesssim T^{-1}\int_0^{\frac{1}{T}} \frac{T^{-2\kappa}}{\mu(T^{-1}+\mu)^{2\kappa}}\langle \mathfrak{d} P(d\mu)\mathfrak{d}\rangle \\ &\lesssim T^{-1}\int_0^{\frac{1}{T}} \mu(g_\kappa(\mu,T)-g_0(\mu))^2\langle \mathfrak{d} P(d\mu)\mathfrak{d}\rangle \\ &\lesssim T^{-1-\alpha}.\end{aligned}$$

5. Proofs of the deterministic auxiliary results

5.1. **Proof of Lemma 2: two-scale expansion.** We start with the uniform ellipticity (25). For $\xi \in \mathbb{R}^d$ we define $\phi_{T\xi} := \xi_i \phi_{Ti}$, so that from (20) we obtain by linearity $\frac{1}{T}\phi_{T\xi} - \nabla \cdot a(\nabla\phi_{T\xi}+\xi) = 0$. Since by uniqueness the solution $\phi_{T\xi}(a,x)$ is shift-invariant or stationary in the sense of $\phi_{T\xi}(a, x+z) = \phi_{T\xi}(a(\cdot + z), x)$, we get from the stationarity of $\langle\cdot\rangle$

$$\frac{1}{T}\langle \phi_{T\xi}^2\rangle + \langle \nabla\phi_{T\xi}\cdot a(\nabla\phi_{T\xi}+\xi)\rangle = 0,$$

which implies in particular

$$\langle(\nabla\phi_{T\xi}+\xi)\cdot a(\nabla\phi_{T\xi}+\xi)\rangle \le \xi\cdot\langle a(\nabla\phi_{T\xi}+\xi)\rangle \overset{(19),(20)}{=} \xi\cdot a_{hT}\xi. \tag{62}$$

On the one hand, (62) yields the lower bound in (25). Indeed, using Jensen's inequality for $\langle\cdot\rangle$ and stationarity of ϕ_ξ in form of $\langle\nabla\phi_\xi\rangle = \nabla\langle\phi_\xi\rangle = 0$, we have

$$\xi\cdot a_{hT}\xi \overset{(62),(1)}{\ge} \lambda\langle|\nabla\phi_{T\xi}+\xi|^2\rangle \ge \lambda|\langle\nabla\phi_{T\xi}\rangle+\xi|^2 = \lambda|\xi|^2.$$

On the other hand, applying the second bound in (1) in form of $(\nabla\phi_{T\xi}+\xi)\cdot a(\nabla\phi_{T\xi}+\xi) \ge |a(\nabla\phi_{T\xi}+\xi)|^2$ and then Jensen's inequality to the effect of $\langle(\nabla\phi_{T\xi}+\xi)\cdot a(\nabla\phi_{T\xi}+\xi)\rangle \ge |\langle a(\nabla\phi_{T\xi}+\xi)\rangle|^2 \overset{(19),(20)}{=} |a_{hT}\xi|^2$, we obtain the second bound in (25).

We continue with the argument for (26) which by (19) takes the form

$$q_{Ti} = \langle q_{Ti}\rangle + \nabla\cdot\sigma_{Ti} + g_{Ti}. \tag{63}$$

We fix and drop the indices Ti. Taking the divergence of (23), ie taking the derivative wrt x_k and summing over $k = 1, \cdots, d$ yields

$$\frac{1}{T}(\nabla\cdot\sigma)_j - \triangle(\nabla\cdot\sigma)_j = \partial_j\nabla\cdot q - \triangle q_j.$$

Inserting (20) yields the identity of vector fields

$$\nabla\cdot\sigma - T\triangle\nabla\cdot\sigma = \nabla\phi - T\triangle q.$$

Adding (24) yields

$$(\mathrm{id} - T\triangle)(\nabla\cdot\sigma + g) = q - \langle q\rangle - T\triangle q,$$

which implies (63) and thus (26) by invertibility of $\mathrm{id} - T\triangle$ on bounded fields; this is the uniqueness statement alluded to after the definition of g, the main ingredient of which is the localized energy estimate (173) in Lemma 10.

Identities (27) and (30) follow immediately from definition (22).

We now give the argument for (28) and eventually (29); for notational simplicity we omit the index T. From (27) we obtain by Leibniz' rule

$$\nabla w = \nabla u - \partial_i v(e_i + \nabla\phi_i) - \phi_i \nabla\partial_i v. \tag{64}$$

We obtain (28) by applying a to this identity and using (20) & (26)

$$\begin{aligned} a\nabla w &= a\nabla u - \partial_i v(a_h e_i + \nabla\cdot\sigma_i + g_i) - \phi_i a\nabla\partial_i v \\ &= a\nabla u - a_h\nabla v - \partial_i v\nabla\cdot\sigma_i - \partial_i v g_i - \phi_i a\nabla\partial_i v \\ &= a\nabla u - a_h\nabla v - \nabla\cdot(\partial_i v\sigma_i) - \partial_i v g_i - (\phi_i a - \sigma_i)\nabla\partial_i v. \end{aligned}$$

Applying $-\nabla\cdot a$ to (64) yields by definition (20) of q_i

$$-\nabla\cdot a\nabla w = -\nabla\cdot a\nabla u + (\nabla\partial_i v)\cdot q_i + \partial_i v\nabla\cdot q_i + \nabla\cdot(\phi_i a\nabla\partial_i v).$$

We now make use of (26) and (20) to obtain

$$-\nabla\cdot a\nabla w = -\nabla\cdot a\nabla u + (\nabla\partial_i v)\cdot(a_h e_i + \nabla\cdot\sigma_i + g_i) + \partial_i v\frac{1}{T}\phi_i + \nabla\cdot(\phi_i a\nabla\partial_i v).$$

Since a_h is constant, $(\nabla\partial_i v)\cdot a_h e_i = \nabla\cdot a_h\nabla v$. In addition, we appeal to the general formula $(\nabla\zeta)\cdot(\nabla\cdot\sigma) = -\nabla\cdot(\sigma\nabla\zeta)$ (which is a consequence of the skew-symmetry of σ in combination with the symmetry of $\nabla^2\zeta$) to obtain

$$-\nabla\cdot a\nabla w = -\nabla\cdot a\nabla u + \nabla\cdot a_h\nabla v - \nabla\cdot(\sigma_i\nabla\partial_i v) + g_i\cdot\nabla\partial_i v + \partial_i v\frac{1}{T}\phi_i + \nabla\cdot(\phi_i a\nabla\partial_i v).$$

Taking the sum of this with $\partial_\tau w \overset{(22)}{=} \partial_t u - \partial_\tau v - \phi_i\partial_i\partial_\tau v$ and appealing to the equations (10) & (18), we obtain (29).

5.2. **Proof of Lemma 3: pointwise estimates of spatial averages.** By definition (18) of v we have $v(t+\tau) = G_h(\tau + r^2) * u(t)$, where $G_h = G_h(\tau, x)$ is the Gaussian heat kernel associated to the constant coefficient operator $\partial_\tau - \nabla\cdot a_{hT}\nabla$, which by (14) turns into $v(t+\tau) = G_h(\tau+r^2) * \nabla\cdot(q(t) - \langle q(t)\rangle)$. From this, we get the representations

$$\begin{aligned} &(\nabla v, \nabla^2 v, \nabla\partial_\tau v)(t+\tau) \\ &\quad= \big(\nabla^2 G_h, \nabla^3 G_h, \nabla^2\nabla\cdot a_{hT}\nabla G_h\big)(\tau+r^2) * (q(t) - \langle q(t)\rangle). \end{aligned}$$

We now use the semi-group property of convolution with Gaussians to distribute the smoothing on the two scales $\sqrt{\tau + r^2}$, R evenly on both factors

$$\begin{aligned} &(\nabla v, \nabla^2 v, \nabla\partial_\tau v)_R(t+\tau) \\ &\quad= \big(\nabla^2 G_h, \nabla^3 G_h, \nabla^2\nabla\cdot a_{hT}\nabla G_h\big)_{\frac{1}{\sqrt{2}}R}\Big(\frac{\tau+r^2}{2}\Big) \\ &\qquad * \Big(G_h\Big(\frac{\tau+r^2}{2}\Big) * (q(t) - \langle q(t)\rangle)\Big)_{\frac{1}{\sqrt{2}}R}. \end{aligned}$$

We estimate the first factor with help of the two independent estimates

$$|\nabla^k G_h(\tau)| \lesssim \frac{1}{\sqrt{\tau}^k} G_h(2\tau) \text{ and } |(\nabla^k G_h(\tau))_{\frac{1}{\sqrt{2}}R}| \lesssim \frac{1}{R^k} G_h(\tau)_R,$$

which we combine to $|(\nabla^k G_h(\frac{\tau+r^2}{2}))_{\frac{1}{\sqrt{2}}R}| \lesssim \frac{1}{\sqrt{\tau+r^2+R^2}^k} G_h(\tau+r^2)_R$, and so obtain

$$\sum_{k=2}^{3} \big|\sqrt{\tau+r^2+R^2}^k (\nabla^{k-1} v(t+\tau))_R\big| + \big|(\tau+r^2+R^2)^2 (\nabla\partial_\tau v(t+\tau))_R\big|$$
$$\lesssim G_h(\tau+r^2)_R * \big|\big(G_h(\frac{\tau+r^2}{2}) * (q(t)-\langle q(t)\rangle)\big)_{\frac{1}{\sqrt{2}}R}\big|. \quad (65)$$

Based on the explicit representation $G_h(\tau,x) = \frac{1}{\sqrt{(4\pi)^d \tau^d \det a_{hT}}} \exp(-\frac{x\cdot a_{hT}^{-1} x}{4\tau})$, and by the bounds (25) on a_{hT}, we see that we may write $G_h(\tau)$ as the convolution of an *isotropic* Gaussian of covariance $\sim \tau$ and a (possibly anisotropic) Gaussian of covariance matrix $\sim \tau$; hence the two convolutions $G_h(\tau)*f$ and $f_{\sqrt{\tau}}$ are comparable in the sense of

$$|G_h(\tau)*f| \lesssim |f_{\frac{1}{C}\sqrt{\tau}}|_{C\sqrt{\tau}} \le |f|_{C\sqrt{\tau}}. \quad (66)$$

With help of the first estimate in (66) (and the semi-group property), we obtain for the rhs of (65)

$$\big|\big(G_h(\frac{\tau+r^2}{2}) * (q(t)-\langle q(t)\rangle)\big)_{\frac{1}{\sqrt{2}}R}\big| \lesssim \big|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+r^2+R^2}}\big|_{C\sqrt{\tau+r^2+R^2}}$$

and then with help of the second estimate

$$G_h(\tau+r^2)_R * \big|G_h(\frac{\tau+r^2}{2}) * (q(t)-\langle q(t)\rangle)_{\frac{1}{\sqrt{2}}R}\big|$$
$$\lesssim \big|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+r^2+R^2}}\big|_{C\sqrt{\tau+r^2+R^2}}.$$

Inserting this into (65) yields the claim of the lemma.

5.3. **Proof of Lemma 4: control of the relative homogenization error.** In the statement of Lemma 4, the average of the lhs is performed with a Gaussian kernel whereas the rhs is localized by an exponential kernel. We shall use in the proof that the exponential kernel dominates the Gaussian kernel. According to (29) in Lemma 2, the error w in the two-scale expansion satisfies an equation of the form

$$\partial_\tau w - \nabla\cdot a\nabla w = f + \nabla\cdot h \quad \text{for } \tau \ge t. \quad (67)$$

We shall now argue that this implies by the localized parabolic energy estimate

$$\Big(\int_t^T \int \eta_{\sqrt{T}} |\nabla w|^2 d\tau\Big)^{\frac{1}{2}}$$
$$\lesssim \Big(\int \eta_{\sqrt{T}} w^2(t)\Big)^{\frac{1}{2}} + \Big(\int_t^T \int \eta_{\sqrt{T}} |(\sqrt{T} f, h)|^2 d\tau\Big)^{\frac{1}{2}}. \quad (68)$$

By scaling and translation invariance, in establishing (68) we may assume $t = 0$ and $T = 1$; we obtain from (67)

$$\frac{d}{dt}\int \eta_1 \frac{1}{2}w^2 = -\int \nabla(\eta_1 w)\cdot(a\nabla w + h) + \int \eta_1 w f.$$

By Leibniz' rule $\nabla(\eta_1 w) = \eta_1\nabla w + w\nabla\eta_1$, the bounds (1) on a, and the exponential form of η_1, we have

$$\begin{aligned}
-\nabla(\eta_1 w)\cdot(a\nabla w + h) &\le -\lambda\eta_1|\nabla w|^2 + \eta_1|\nabla w||h| + |w||\nabla\eta_1|(|\nabla w| + |h|)\\
&\le \eta_1(-\lambda|\nabla w|^2 + |\nabla w||h| + |w|(|\nabla w| + |h|))\\
&\le \eta_1(-\frac{\lambda}{2}|\nabla w|^2 + (|\nabla w| + |w|)|h| + \frac{1}{2\lambda}w^2).
\end{aligned}$$

This yields the differential inequality

$$\frac{d}{dt}\int \eta_1\frac{1}{2}w^2 \le \int \eta_1\big(-\frac{\lambda}{2}|\nabla w|^2 + \frac{1}{2\lambda}w^2 + (|\nabla w| + |w|)(|h| + |f|)\big),$$

which by Young's inequality yields

$$\frac{d}{dt}\int \eta_1 w^2 + \frac{1}{C}\int \eta_1|\nabla w|^2 \le C\int \eta_1(w^2 + f^2 + |h|^2), \tag{69}$$

which we rewrite as

$$\frac{d}{dt}\exp(-Ct)\int \eta_1 w^2 + \frac{1}{C}\exp(-Ct)\int \eta_1|\nabla w|^2 \lesssim C\exp(-Ct)\int \eta_1(f^2 + |h|^2).$$

Integration yields $\int_0^1\int \eta_1|\nabla w|^2 d\tau \lesssim \int \eta_1 w^2(t=0) + \int_0^1\int \eta_1(f^2 + |h|^2)d\tau$, which is the rescaled version of (68).

We now make use of (68). According to (29) we have $f := \phi_{Ti}(\frac{1}{T} - \partial_\tau)\partial_i v + g_{Ti}\cdot\nabla\partial_i v$ and $h := (\phi_{Ti}a - \sigma_{Ti})\nabla\partial_i v$; according to (30) we have $w(t) = u(t) - v(t) - \phi_{Ti}\partial_i v(t)$. Hence (68) together with the bounds (1) turns into

$$\begin{aligned}
\Big(\int_t^T\int \eta_{\sqrt{T}}|\nabla w|^2 d\tau\Big)^{\frac{1}{2}} &\lesssim \Big(\int \eta_{\sqrt{T}}\big(|u(t) - v(t)|^2 + T|\nabla v(t)|^2|\frac{\phi_T}{\sqrt{T}}|^2\big)\Big)^{\frac{1}{2}}\\
&+ \Big(\int_t^T\int \eta_{\sqrt{T}}|(\nabla v, \sqrt{T}\nabla^2 v, T\nabla\partial_\tau v)|^2|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2 d\tau\Big)^{\frac{1}{2}}.
\end{aligned}$$

Without changing the form of the rhs, we may add the non-divergence terms in the representation (28) of the homogenization error

$$\begin{aligned}
&\Big(\int_t^T\int \eta_{\sqrt{T}}|(\nabla w, a\nabla w + \partial_i v g_{Ti} + (\phi_{Ti}a - \sigma_{Ti})\nabla\partial_i v|^2 d\tau\Big)^{\frac{1}{2}}\\
\lesssim\ &\Big(\int \eta_{\sqrt{T}}\big(|u(t) - v(t)|^2 + T|\nabla v(t)|^2|\frac{\phi_T}{\sqrt{T}}|^2\big)\Big)^{\frac{1}{2}}\\
+\ &\Big(\int_t^T\int \eta_{\sqrt{T}}|(\nabla v, \sqrt{T}\nabla^2 v, T\nabla\partial_\tau v)|^2|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2 d\tau\Big)^{\frac{1}{2}}.
\end{aligned}\tag{70}$$

We now make two observation about the relation between the Gaussian averaging function $G_R(x) = \frac{1}{(\sqrt{2\pi}R)^d}\exp(-\frac{1}{2}|\frac{x}{R}|^2)$ and the exponential localization function $\eta_R(x) = \frac{1}{R^d}\exp(-|\frac{x}{R}|)$: Since the tails of the latter dominate the tails of the former, we have $G_R \lesssim G_{2R} \lesssim \eta_R$; which allows us to also conclude $|\nabla G_R| \lesssim \frac{1}{R}G_{2R} \lesssim \frac{1}{R}\eta_R$. In conjunction with $\eta_R \le (\frac{\sqrt{T}}{R})^d\eta_{\sqrt{T}}$ for $R \le \sqrt{T}$ and Jensen's inequality wrt to the averaging by $\eta_{\sqrt{T}}$, we have for any function $f = f(x)$:

$$|f_R| \lesssim (\frac{\sqrt{T}}{R})^{\frac{d}{2}}\big(\int \eta_{\sqrt{T}}|f|^2\big)^{\frac{1}{2}} \quad \text{and} \quad |(\nabla f)_R| \lesssim \frac{1}{R}(\frac{\sqrt{T}}{R})^{\frac{d}{2}}\big(\int \eta_{\sqrt{T}}|f|^2\big)^{\frac{1}{2}}. \qquad (71)$$

Applying the first estimate, which we use in the weaker form of $(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}|f_R| \lesssim \big(\int \eta_{\sqrt{T}}|f|^2\big)^{\frac{1}{2}}$, to the lhs of (70) and the second estimate, which we rewrite as $(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}|(\nabla f)_R| \lesssim \big(\int \eta_{\sqrt{T}}|\frac{f}{\sqrt{T}}|^2\big)^{\frac{1}{2}}$, to the terms $f = \phi_{Ti}\partial_i v, \partial_i v\sigma_{Ti}$ from (27) & (28), we see that (70) yields

$$\begin{aligned}
&(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\big(\int_t^T |(\nabla u - \nabla v, a\nabla u - a_{hT}\nabla v)_R|^2 d\tau\big)^{\frac{1}{2}}\\
&\lesssim \big(\int \eta_{\sqrt{T}}\big(|u(t)-v(t)|^2 + T|\nabla v(t)|^2|\frac{\phi_T}{\sqrt{T}}|^2\big)\big)^{\frac{1}{2}}\\
&+ \big(\int \eta_{\sqrt{T}}\int_t^T |(\nabla v, \sqrt{T}\nabla^2 v, T\nabla\partial_\tau v)|^2 d\tau|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2\big)^{\frac{1}{2}}.
\end{aligned}$$

Using Cauchy-Schwarz in t on the lhs and $\eta_{\sqrt{T}} \lesssim \exp(-\frac{|x|}{2\sqrt{T}})\eta_{2\sqrt{T}}$ on the rhs we obtain

$$\begin{aligned}
&(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T |(\nabla u - \nabla v, a\nabla u - a_{hT}\nabla v)_R| d\tau \qquad (72)\\
&\lesssim \sqrt{T}\big(\int \eta_{\sqrt{T}}|u(t)-v(t)|^2\big)^{\frac{1}{2}}\\
&+ \Sigma\sup_x\Big\{\exp(-\frac{|x|}{2\sqrt{T}})\Big(T|\nabla v(t)| + \big(T\int_t^T |(\nabla v, \sqrt{T}\nabla^2 v, T\nabla\partial_\tau v)|^2 d\tau\big)^{\frac{1}{2}}\Big)\Big\},
\end{aligned}$$

where we have set for abbreviation $\Sigma := \big(\int \eta_{2\sqrt{T}}|(\frac{\phi_T}{\sqrt{T}}, \frac{\sigma_T}{\sqrt{T}}, g_T)|^2\big)^{\frac{1}{2}}$. In order to estimate the first factor of the last term, we appeal to Lemma 3 with $R = 0$ in form of

$$\begin{aligned}
&T|(\nabla v, \sqrt{T}\nabla^2 v, T\nabla\partial_\tau v)(t+\tau)|\\
&\lesssim (\frac{T}{\tau+r^2} + (\frac{T}{\tau+r^2})^{\frac{3}{2}} + (\frac{T}{\tau+r^2})^2)|(q(t) - \langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+r^2}}|_{C\sqrt{\tau+r^2}},
\end{aligned}$$

with the understanding that the first of the three lhs term is estimated by the first rhs contribution alone. Since $r^2 \le t, \tau \le T$, this yields by Jensen's inequality applied to the rhs

$$\begin{aligned} & T|(\nabla v,\sqrt{T}\nabla^2 v,T\nabla\partial_\tau v)(t+\tau)| \\ \lesssim & \ (\frac{T}{\tau+r^2}+(\frac{T}{\tau+r^2})^2)|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}r}|_{C\sqrt{T}}, \end{aligned}$$

so that we obtain from integrating over $\tau\in(0,T-t)$

$$\begin{aligned} & T|\nabla v(t)| \\ + & \ \big(T\int_t^T|(\nabla v,\sqrt{T}\nabla^2 v,T\nabla\partial_\tau v)|^2 d\tau\big)^{\frac12}\lesssim(\frac{\sqrt{T}}{r})^3|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}r}|_{C\sqrt{T}}. \end{aligned} \tag{73}$$

In order to take the supremum in x, (72), we now need the auxiliary statement

$$\sup_x\Big\{\exp(-|\frac{x}{2\sqrt{T}}|)|f|_{C\sqrt{T}}(x)\Big\}\lesssim\int\eta_{2\sqrt{T}}|f|. \tag{74}$$

Here comes the argument for (74): Wlog we may assume $2\sqrt{T}=1$ so that the estimate turns into

$$\sup_x\Big\{\exp(-|x|)\int\exp(-\frac12|\frac{x-y}{C}|^2)|f(y)|dy\Big\}\lesssim\int\exp(-|y|)|f(y)|dx,$$

which follows from $|x|+\frac12|\frac{x-y}{C}|^2\geq|x|+|x-y|-\frac{C^2}{2}\geq|y|-\frac{C^2}{2}$. Using (74) for $f=|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}r}|$, inserting this into (73) and then into (72) yields

$$\begin{aligned} & (\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T|(\nabla u-\nabla v,a\nabla u-a_{hT}\nabla v)_R|d\tau \\ & \lesssim\sqrt{T}\big(\int\eta_{\sqrt{T}}|u(t)-v(t)|^2\big)^{\frac12}+\Sigma(\frac{\sqrt{T}}{r})^3\int\eta_{2\sqrt{T}}|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}r}|. \end{aligned} \tag{75}$$

We now turn to the first rhs term in (75). Because of (18) we have $v(t)=G_h(r^2)*u(t)$, hence $v(t)$ behaves as $u(t)_r$ so that by the same argument that leads to (86) in the proof of Lemma 5 (see below):

$$\sqrt{T}\Big(\int\eta_{\sqrt{T}}|u(t)-v(t)|^2\Big)^{\frac12}\lesssim\frac{\sqrt{T}}{\sqrt{t}}\frac{r}{\sqrt{t}}\Big(\int\eta_{\sqrt{T}}|t\nabla u(t)|^2\Big)^{\frac12}. \tag{76}$$

We now appeal to Lemma 11 with T replaced by t and the interval $(0,t)$ replaced by $(\frac{t}{2},t)$ and with $p=\frac{d}{2}+2$:

$$\begin{aligned} \Big(\int\eta_{\sqrt{t}}|\sqrt{t}\nabla u(t)|^2\Big)^{\frac12} & \lesssim\fint_{\frac{t}{2}}^t dt'\fint_0^{\sqrt{t/2}}dr'(\frac{r'}{\sqrt{t/2}})^{\frac{d}{2}+2}\int\eta_{2\sqrt{t/2}}|u_{r'}(t')| \\ & \lesssim\fint_{\frac{t}{2}}^t dt'\fint_0^{\sqrt{t'}}dr'(\frac{r'}{\sqrt{t'}})^{\frac{d}{2}+2}\int\eta_{2\sqrt{t}}|u_{r'}(t')|. \end{aligned} \tag{77}$$

Using once more (14) in form of $u(t')=\nabla\cdot(q(t')-\langle q(t')\rangle)$ and properties of the convolution with Gaussians in form of $|(\nabla\cdot(q(t')-\langle q(t')\rangle))_{r'}| = |(\nabla\cdot(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt{2}}r'})_{\frac{1}{\sqrt{2}}r'}| \lesssim \frac{1}{r'}|(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt{2}}r'}|_{r'}$, this yields

$$\Big(\int\eta_{\sqrt{t}}|t\nabla u(t)|^2\Big)^{\frac12}\lesssim \fint_{\frac{t}{2}}^{t}dt'\fint_0^{\sqrt{t'}}dr'(\frac{r'}{\sqrt{t'}})^{\frac{d}{2}+1}\int\eta_{2\sqrt{t}}|(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt{2}}r'}|_{r'}. \quad (78)$$

In order to connect to (76), we need to post-process both sides of (78) by some auxiliary statements, which hold for a function $f=f(x)\geq 0$ and two scales $r\leq\frac12 R$:

$$\text{(holds even for } r\lesssim R) \quad \int\eta_R f_r\sim\int\eta_R f, \quad (79)$$

$$(\frac{R}{r})^{\frac{d}{2}}\int\eta_{2R}(y)\Big(\int\eta_r(x-y)f^2(x)dx\Big)^{\frac12}dy\gtrsim\Big(\int\eta_R f^2\Big)^{\frac12}, \quad (80)$$

$$\int\eta_R(y)\int\eta_r(x-y)f(x)dxdy\sim\int\eta_R f. \quad (81)$$

Let us first give the elementary argument for (79) - (81), for the purpose of which we may assume $R=1$. Clearly, (79) is equivalent to $(\eta_1)_r\sim\eta_1$, that is,

$$\frac{1}{r^d}\int\exp(-\frac12|\frac{y}{r}|^2)\exp(|x|-|x-y|)dy\sim 1,$$

by the triangle inequality we have $\exp(|x|-|x-y|)\in[\exp(-|y|),\exp(|y|)]$, so that the statement follows from $\int\exp(-\frac12|\frac{y}{r}|^2\pm|y|)dy\sim r^d$, which can be seen from applying Young's inequality on $-\frac12|\frac{y}{r}|^2\pm|y|$. The approximate identity (81) is equivalent to $\int\eta_1(y)\eta_r(x-y)dy=\int\eta_r(y)\eta_1(x-y)dy\sim\eta_1(x)$. Since $\frac{\eta_1(x-y)}{\eta_1(x)}=\exp(-|x-y|+|x|)$ $\in[\exp(-|y|),\exp(|y|)]$ this follows from $r\leq\frac12$. For (80) we start with $\eta_1(y)\int\eta_r(x-y)f^2(x)dx\leq(\frac1r)^d\int\eta_1(x)f^2(x)dx$, a consequence of

$$\eta_1(y)\eta_r(x-y)=\frac{1}{r^d}\exp(-|y|-|\frac{x-y}{r}|)\leq\frac{1}{r^d}\exp(-|x|)=\frac{1}{r^d}\eta_1(x),$$

of which we take the square root: $2^d\eta_2(y)\big(\int\eta_r(x-y)f^2(x)dx\big)^{\frac12}\leq(\frac1r)^{\frac{d}{2}}\big(\int\eta_1 f^2\big)^{\frac12}$. We combine this with (81) in form of

$$\int\eta_1 f^2\lesssim\int\eta_1(y)\int\eta_r(x-y)f^2(x)dxdy=\int 2^d\eta_2(y)\eta_2(y)\int\eta_r(x-y)f^2(x)dxdy$$
$$\leq(\frac1r)^{\frac{d}{2}}\Big(\int\eta_1 f^2\Big)^{\frac12}\int\eta_2(y)\Big(\int\eta_r(x-y)f^2(x)dx\Big)^{\frac12}dy.$$

Dividing by $\big(\int\eta_1 f^2\big)^{\frac12}$ yields (80).

Let us use (79) - (81) for the post-processing of (78): We first use (79) for $r=r'$, $R=2\sqrt{t}$ on the rhs of (78) to the effect of $\int\eta_{2\sqrt{t}}|(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt{2}}r'}|_{r'}\lesssim\int\eta_{2\sqrt{t}}|(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt{2}}r'}|$. We then recenter the ensuing estimate in y, multiply with $\eta_{\sqrt{T}}(y)$; an application of (80) on the lhs for $R=\sqrt{T}$ and of (81) for $2R=\sqrt{T}$ on the rhs, both for $r=2\sqrt{t}$, we obtain

$$\begin{aligned}&\Big(\int \eta_{\sqrt{T}}|t\nabla u(t)|^2\Big)^{\frac12}\\ &\lesssim\ (\frac{\sqrt{T}}{\sqrt{t}})^{\frac d2}\fint_{\frac t2}^{t}dt'\fint_0^{\sqrt{t'}}dr'(\frac{r'}{\sqrt{t'}})^{\frac d2+1}\int \eta_{\sqrt{T}}|(q(t')-\langle q(t')\rangle)_{\frac{1}{\sqrt2}r'}|.\end{aligned}$$

By a rescaling of r' we may get rid of the factor $\frac{1}{\sqrt2}$. We insert this into (76)

$$\begin{aligned}&\sqrt{T}\Big(\int \eta_{\sqrt{T}}|u(t)-v(t)|^2\Big)^{\frac12}\\ &\lesssim\ (\frac{\sqrt{T}}{\sqrt{t}})^{\frac d2+1}\frac{r}{\sqrt{t}}\fint_{\frac t2}^{t}dt'\fint_0^{\sqrt{t'}}dr'(\frac{r'}{\sqrt{t'}})^{\frac d2+1}\int \eta_{\sqrt{T}}|(q(t')-\langle q(t')\rangle)_{r'}|\end{aligned}$$

and then into (75), where we redefine $\frac1C r$ by r,

$$\begin{aligned}&(\frac{R}{\sqrt{T}})^{\frac d2+1}\int_t^T|(\nabla u-\nabla v,a\nabla u-a_{hT}\nabla v)_R|d\tau\\ &\lesssim\ (\frac{\sqrt{T}}{\sqrt{t}})^{\frac d2+1}\frac{r}{\sqrt{t}}\fint_{\frac t2}^{t}dt'\fint_0^{\sqrt{t'}}dr'(\frac{r'}{\sqrt{t'}})^{\frac d2+1}\int \eta_{\sqrt{T}}|(q(t')-\langle q(t')\rangle)_{r'}|\\ &+\ \Sigma(\frac{\sqrt{T}}{r})^3\int \eta_{2\sqrt{T}}|(q(t)-\langle q(t)\rangle)_r|,\end{aligned}$$

which is (32).

5.4. **Proof of Corollary 2: optimal decay conditioned on the good event.** By definition (35) of $\mathcal{G}_\delta$, the convexity of $F\mapsto\langle\exp|F|\rangle$ and by stationarity of $|(q(t')-\langle q(t')\rangle)_{r'}|$, which for $f(a,x):=|\nu(q(t')-\langle q(t')\rangle)_{r'}(x)|$ gives $\langle\exp(\int\eta_R(x)f(x)dx)\rangle\le\int dx\eta_R(x)\langle\exp(f(x))\rangle=\langle\exp(f)\rangle$, we obtain from (32) in Lemma 4

$$\begin{aligned}&\Big\langle I(\mathcal{G}_\delta)\exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac d2+1}\int_t^T(\nabla u-\nabla v,a\nabla u-a_{hT}\nabla v)_R d\tau\big|\Big\rangle\\ &\le\ \frac12\fint_{\frac t2}^{t}dt'\fint_0^{\sqrt{t'}}dr'\Big\langle\exp\big|C\nu(\frac{\sqrt{T}}{\sqrt{t}})^{\frac d2+1}\frac{r}{\sqrt{t}}(\frac{r'}{\sqrt{t'}})^{\frac d2+1}(q(t')-\langle q(t')\rangle)_{r'}\big|\Big\rangle\\ &+\ \frac12\Big\langle\exp\big|C\nu\delta(\frac{\sqrt{T}}{r})^3(q(t)-\langle q(t)\rangle)_r\big|\Big\rangle.\end{aligned}$$

Using the elementary inequality $\exp(|F|) \le \exp(F) + \exp(-F)$, we may make use of our assumption (33) (for suitably redefined ν)

$$\begin{aligned}
&\big\langle I(\mathcal{G}_\delta) \exp \big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1} \int_t^T |(\nabla u - \nabla v, a\nabla u - a_{hT}\nabla v)_R| d\tau\big|\big\rangle \\
\le\ & \exp\big(C\Lambda(\frac{\sqrt{T}}{\sqrt{t}})^{d+2}(\frac{r}{\sqrt{t}})^2\nu^2\big) + \exp\big(C\Lambda\delta^2(\frac{\sqrt{T}}{r})^6(\frac{\sqrt{t}}{r})^{d+2}\nu^2\big) \\
\le\ & 2\exp\big(C\Lambda\big((\frac{\sqrt{T}}{\sqrt{t}})^{d+2}(\frac{r}{\sqrt{t}})^2 + \delta^2(\frac{\sqrt{T}}{\sqrt{t}})^6(\frac{\sqrt{t}}{r})^{d+8}\big)\nu^2\big).
\end{aligned}$$

This estimate prompts the optimization of the regularization scale $r \le \sqrt{t}$; in order to keep the exponents simple, we replace the two exponents $d+2$ and 6 of $\frac{\sqrt{T}}{\sqrt{t}}$ by common exponent $d+4$.

5.5. **Proof of Lemma 5: strong smallness of the extended corrector.** In order to avoid indices, we use ($d=3$)-notation when it comes to the vector potential σ_T, that is, we write $\frac{1}{T}\sigma_T - \triangle\sigma_T = \nabla \times q_T$ instead of (23).

Step 1. We claim that for all $t \le T$ and $0 < \delta \le 1$

$$\Big(\int \eta_{\sqrt{T}r_0}|\frac{\phi_t}{\sqrt{t}}|^2\Big)^{\frac{1}{2}} \lesssim \frac{1}{\delta}\Big(\int \eta_{\sqrt{T}r_0}|(q_t - \langle q_t\rangle)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}} + \delta, \tag{82}$$

$$\Big(\int \eta_{\sqrt{T}r_0}|\frac{\sigma_t}{\sqrt{t}}|^2\Big)^{\frac{1}{2}} \lesssim \Big(\int \eta_{\sqrt{T}r_0}|(q_t - \langle q_t\rangle)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}} + \delta, \tag{83}$$

$$\Big(\int \eta_{\sqrt{T}r_0}|g_t|^2\Big)^{\frac{1}{2}} \lesssim \Big(\int \eta_{\sqrt{T}r_0}|(q_t - \langle q_t\rangle - \nabla\phi_t)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}} + \delta. \tag{84}$$

Here comes the argument: It is enough to prove these estimates for $\sqrt{T} = \sqrt{t}$: We replace the origin by y and then apply $\int \eta_{\sqrt{T}r_0}(y)\cdot dy$, using $\eta_{\sqrt{T}r_0} * \eta_{\sqrt{t}r_0} \lesssim \eta_{\sqrt{T}r_0}$ (for $\sqrt{t} \ll \sqrt{T}$, cf. (81)) for our exponential localization function. By scale invariance, it is enough to establish them for t such that $\sqrt{t}r_0 = 1$, we shall thus write (ϕ, σ, g) and η instead of (ϕ_t, σ_t, g_t) and $\eta_{\sqrt{t}}$, respectively. For all functions (ϕ, σ, g) we start from the splitting

$$\int \eta|(\phi,\sigma,g)|^2 \lesssim \int \eta|(\phi,\sigma,g)_\delta|^2 + \int \eta|(\phi,\sigma,g) - (\phi,\sigma,g)_\delta|^2. \tag{85}$$

Since the exponential localization function η dominates the Gaussian convolution kernel, we have for the second term

$$\int \eta|(\phi,\sigma,g) - (\phi,\sigma,g)_\delta|^2 \lesssim \delta^2 \int \eta|\nabla(\phi,\sigma,g)|^2. \tag{86}$$

Indeed, if G_r denotes the Gaussian kernel of scale r, we have

$$(\phi_\delta - \phi)(x) = \int_0^1 \int \nabla\phi(x+sz)\cdot zG_\delta(z)dzds;$$

since $|zG_\delta(z)|\lesssim \delta G_{2\delta}(z)$, this yields $|\phi_\delta-\phi|\lesssim\delta\int_0^1|\nabla\phi|_{2\delta s}ds$, and thus by Jensen's inequality $|\phi_\delta-\phi|^2\lesssim\delta^2\int_0^1|\nabla\phi|_{2\delta s}^2ds$ and therefore by the symmetry of the convolution operator

$$\int\eta|\phi-\phi_\delta|^2\lesssim\delta^2\int_0^1\int\eta|\nabla\phi|_{2\delta s}^2ds\overset{(79)}{\lesssim}\delta^2\int\eta|\nabla\phi|^2,$$

which is (86). Thanks to the massive term in the equations (20), (23), (24) for ϕ, σ, and g, we may localize the elliptic energy estimates (see (174) in the proof of Lemma 10 where the radius $r_0\le C$ is introduced, which does not appear here because we have set $\sqrt{t}r_0=1$)

$$\int\eta|(\nabla\phi,q)|^2\lesssim\int\eta|ae|^2\lesssim 1,\tag{87}$$

and

$$\int\eta|\nabla\sigma|^2\lesssim\int\eta|q|^2,\quad\int\eta|\nabla g|^2\lesssim\int\eta|(\nabla\phi,q-\langle q\rangle)|^2,$$

which together with $|\langle q\rangle|\le\langle|q|^2\rangle\lesssim 1$, which follows from (87) by stationarity, yield

$$\int\eta|(\nabla\phi,\nabla\sigma,\nabla g,q-\langle q\rangle)|^2\lesssim 1.$$

It remains to estimate the first term in (85). This is easy for (σ,g): Since these satisfy constant-coefficient equations we have $\frac{1}{t}\sigma_\delta-\triangle\sigma_\delta=\nabla\times(q-\langle q\rangle)_\delta$ and $\frac{1}{t}g_\delta-\triangle g_\delta=(q-\langle q\rangle-\nabla\phi)_\delta$, so that by the same localized elliptic energy estimate

$$\int\eta|\sigma_\delta|^2\lesssim\int\eta|(q-\langle q\rangle)_\delta|^2,\quad\int\eta|g_\delta|^2\lesssim\int\eta|(q-\langle q\rangle-\nabla\phi)_\delta|^2.$$

For ϕ we note that $\phi=\nabla\cdot(q-\langle q\rangle)$, cf. (20), and thus also $\phi_{\sqrt{2}\delta}=(\nabla\cdot(q-\langle q\rangle)_\delta)_\delta$ by the semi-group property of convolution with the Gaussian. Since the gradient of a Gaussian is estimated by the Gaussian with twice the variance, cf. $|\nabla G_\delta(z)|=|\frac{z}{\delta}G_\delta(z)|\lesssim G_{\sqrt{2}\delta}(z)$, this implies $|\phi_{\sqrt{2}\delta}|\lesssim\frac{1}{\delta}|(q-\langle q\rangle)_\delta|_{\sqrt{2}\delta}$, and thus by Jensen's inequality $|\phi_{\sqrt{2}\delta}|^2\lesssim\frac{1}{\delta^2}|(q-\langle q\rangle)_\delta|_{\sqrt{2}\delta}^2$. Hence we have

$$\int\eta|\phi_{\sqrt{2}\delta}|^2\lesssim\frac{1}{\delta^2}\int\eta_{\sqrt{2}\delta}|(q-\langle q\rangle)_\delta|^2\overset{(79)}{\lesssim}\frac{1}{\delta^2}\int\eta|(q-\langle q\rangle)_\delta|^2.$$

It remains to replace δ by $\sqrt{2}\delta$ in (85) when it comes to ϕ.

Step 2. We claim that for $t\le T$

$$\begin{aligned}&\Big(\int\eta_{\sqrt{T}r_0}|(\frac{\phi_T-\phi_t}{\sqrt{T}},\frac{\sigma_T-\sigma_t}{\sqrt{T}},\nabla\phi_T-\nabla\phi_t,q_T-q_t)|^2\Big)^{\frac12}\\&\lesssim\ \Big(\int\eta_{\sqrt{T}r_0}|(\frac{\phi_t}{\sqrt{t}},\frac{\sigma_t}{\sqrt{t}},(q_t-\langle q_t\rangle)_{\sqrt{t}})|^2\Big)^{\frac12}.\end{aligned}\tag{88}$$

Here comes the argument. We note that by equation (20), $u:=\phi_T-\phi_t$ satisfies $\frac{1}{T}u-\nabla\cdot a\nabla u=(\frac{1}{t}-\frac{1}{T})\phi_t$. Using (20) in form of $\frac{1}{t}\phi_t=\nabla\cdot(q_t-\langle q_t\rangle)$, we split the rhs according to $\phi_t = (\phi_t - \phi_{t,\sqrt{t}}) + t\nabla\cdot(q_t - \langle q_t\rangle)_{\sqrt{t}}$. We now perform a localized elliptic energy estimate which means testing with $\eta_{\sqrt{T}r_0}u$; thanks to the massive term we obtain as in the argument for (174)

$$\int \eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|^2 \lesssim (1-\frac{t}{T})\int \eta_{\sqrt{T}r_0}u\frac{1}{t}(\phi_t-\phi_{t,\sqrt{t}})+\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\sqrt{t}}|^2. \quad (89)$$

For the contribution of the first rhs term we use symmetry of the convolution operator

$$(1-\frac{t}{T})\int \eta_{\sqrt{T}r_0}u\frac{1}{t}(\phi_t-\phi_{t,\sqrt{t}}) = (1-\frac{t}{T})\int \big(\eta_{\sqrt{T}r_0}u - (\eta_{\sqrt{T}r_0}u)_{\sqrt{t}}\big)\frac{1}{t}\phi_t.$$

Denoting by $G_{\sqrt{t}}$ the Gaussian convolution kernel on scale $\sqrt{t}$, we have

$$\begin{aligned}&\big(\eta_{\sqrt{T}r_0}u - (\eta_{\sqrt{T}r_0}u)_{\sqrt{t}}\big)(x)\\ &= \int_0^1\int \big(\eta_{\sqrt{T}r_0}\nabla u + u\nabla\eta_{\sqrt{T}r_0}\big)(x+sz)\cdot zG_{\sqrt{t}}(z)dzds\end{aligned}$$

and thus by $|zG_{\sqrt{t}}(z)| \lesssim \sqrt{t}G_{2\sqrt{t}}(z)$, $|\nabla\eta_{\sqrt{T}r_0}| \lesssim \frac{1}{\sqrt{T}}\eta_{\sqrt{T}r_0}$, and Jensen's inequality

$$\begin{aligned}&\big|\eta_{\sqrt{T}r_0}u - (\eta_{\sqrt{T}r_0}u)_{2\sqrt{t}}\big|(x)\\ &\lesssim \sqrt{t}\int_0^1\int \big(\eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|\big)(x+sz)G_{2\sqrt{t}}(z)dzds\\ &\lesssim \sqrt{t}\int_0^1 \big(\eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|\big)_{2\sqrt{ts}}(x)ds\end{aligned}$$

and thus by Cauchy-Schwarz' inequality, the symmetry of the convolution operator $\cdot_{2\sqrt{ts}}$, and Jensen's inequality

$$\begin{aligned}&\big|(1-\frac{t}{T})\int \eta_{\sqrt{T}r_0}u\frac{1}{t}(\phi_t-\phi_{t,\sqrt{t}})\big|\\ &\lesssim \int_0^1\int \big(\eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|\big)_{2\sqrt{ts}}|\frac{\phi_t}{\sqrt{t}}|ds\\ &\lesssim \Big(\int \eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|^2\Big)^{\frac12}\Big(\int_0^1\int (\eta_{\sqrt{T}r_0})_{2\sqrt{ts}}(\frac{\phi_t}{\sqrt{t}})^2ds\Big)^{\frac12}\\ &\overset{(79)}{\lesssim} \Big(\int \eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|^2\Big)^{\frac12}\Big(\int \eta_{\sqrt{T}r_0}(\frac{\phi_t}{\sqrt{t}})^2\Big)^{\frac12}.\end{aligned}$$

Equipped with this estimate we return to (89) and obtain for $\phi_T - \phi_t$ and $|q_T - q_t| \le |\nabla(\phi_T - \phi_t)|$ as desired

$$\begin{aligned}&\Big(\int \eta_{\sqrt{T}r_0}|(\frac{\phi_T-\phi_t}{\sqrt{T}},\nabla\phi_T-\nabla\phi_t,q_T-q_t)|^2\Big)^{\frac12}\\ &\lesssim \Big(\int \eta_{\sqrt{T}r_0}(\frac{\phi_t}{\sqrt{t}})^2\Big)^{\frac12} + \Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\sqrt{t}}|^2\Big)^{\frac12}. \quad (90)\end{aligned}$$

We now turn to the contribution of σ. As for ϕ, we note that by (23), $u:=\sigma_T-\sigma_t$ satisfies $\frac{1}{T}u-\triangle u=(1-\frac{t}{T})\frac{1}{t}\sigma_t+\nabla\times(q_T-q_t)$ and split the first rhs term into $\frac{1}{t}\sigma_t=\frac{1}{t}(\sigma_t-\sigma_{t,\sqrt{t}})+\frac{1}{t}\sigma_{t,\sqrt{t}}$. Applying the convolution $\cdot_{\sqrt{t}}$ to $\frac{1}{t}\sigma_t-\triangle\sigma_t=\nabla\times q_t$ leads to

$$\frac{1}{t}\sigma_{t,\sqrt{t}}-\triangle\sigma_{t,\sqrt{t}}=\nabla\times(q_t-\langle q_t\rangle)_{\sqrt{t}}. \tag{91}$$

Hence we write

$$\begin{aligned}&\frac{1}{T}u-\triangle u \qquad (92)\\ &= (1-\frac{t}{T})\big(\frac{1}{t}(\sigma_t-\sigma_{t,\sqrt{t}})+\triangle\sigma_{t,\sqrt{t}}+\nabla\times(q_t-\langle q_t\rangle)_{\sqrt{t}}\big)+\nabla\times(q_T-q_t).\end{aligned}$$

By the argument (173) we may exponentially localize the energy estimate for (91) on scale $\sqrt{t}r_0$, and a fortiori on scale $\sqrt{T}r_0$

$$\Big(\int\eta_{\sqrt{T}r_0}|(\frac{\sigma_{t,\sqrt{t}}}{\sqrt{t}},\nabla\sigma_{t,\sqrt{t}})|^2\Big)^{\frac{1}{2}}\lesssim\Big(\int\eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\sqrt{t}}|^2\Big)^{\frac{1}{2}}. \tag{93}$$

By the argument for the localized elliptic energy estimate (173) applied to (92) we have

$$\begin{aligned}&\int\eta_{\sqrt{T}r_0}|(\frac{u}{\sqrt{T}},\nabla u)|^2\\ &\lesssim (1-\frac{t}{T})\int\eta_{\sqrt{T}r_0}u\frac{1}{t}(\sigma_t-\sigma_{t,\sqrt{t}})+\int\eta_{\sqrt{T}r_0}|(\nabla\sigma_{t,\sqrt{t}},(q_t-\langle q_t\rangle)_{\sqrt{t}},q_T-q_t)|^2.\end{aligned}$$

Treating the first rhs term as in case of ϕ, we obtain

$$\begin{aligned}&\Big(\int\eta_{\sqrt{T}r_0}|\frac{\sigma_T-\sigma_t}{\sqrt{T}}|^2\Big)^{\frac{1}{2}}\\ &\lesssim \Big(\int\eta_{\sqrt{T}r_0}|(\frac{\sigma_t}{\sqrt{t}},\nabla\sigma_{t,\sqrt{t}},(q_t-\langle q_t\rangle)_{\sqrt{t}},q_T-q_t)|^2\Big)^{\frac{1}{2}}. \qquad (94)\end{aligned}$$

The combination of (90), and (93) inserted into (94) yields (88).

Step 3. We claim that provided $t\le T$ and $0<\delta<1$

$$\begin{aligned}&\Big(\int\eta_{\sqrt{T}r_0}|(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},g_T)|^2\Big)^{\frac{1}{2}} \qquad (95)\\ &\lesssim \frac{1}{\delta}\Big(\Big(\int\eta_{\sqrt{T}r_0}|(\nabla\phi_t,q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}}+\langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\rangle^{\frac{1}{2}}\Big)+\delta.\end{aligned}$$

Here comes the argument: For the estimate on (ϕ_T,σ_T), we use (88) in conjunction with the triangle inequality

$$\begin{aligned}&\Big(\int\eta_{\sqrt{T}r_0}|(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},\nabla\phi_T-\nabla\phi_t,q_T-q_t)|^2\Big)^{\frac{1}{2}}\\ &\lesssim \Big(\int\eta_{\sqrt{T}r_0}|(\frac{\phi_t}{\sqrt{t}},\frac{\sigma_t}{\sqrt{t}},(q_t-\langle q_t\rangle)_{\sqrt{t}}|^2\Big)^{\frac{1}{2}},\end{aligned}$$

and insert (82) & (83):

$$\Big(\int \eta_{\sqrt{T}r_0}|(\frac{\phi_T}{\sqrt{T}},\frac{\sigma_T}{\sqrt{T}},\nabla\phi_T-\nabla\phi_t,q_T-q_t)|^2\Big)^{\frac{1}{2}}$$
$$\lesssim \frac{1}{\delta}\Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}}+\delta, \tag{96}$$

where we made use of the following monotonicity in the convolution scale, based on the semi-group property $q_{\sqrt{t'}}=(q_{\sqrt{t}})_{\sqrt{t'-t}}$ for $t'\geq t$ of convolution with Gaussian yielding by Jensen's inequality $|q_{\sqrt{t'}}|^2\leq |q_{\sqrt{t}}|^2_{\sqrt{t'-t}}$ and thus by (79)

$$\int \eta_{\sqrt{T}r_0}|q_{\sqrt{t'}}|^2\lesssim\int \eta_{\sqrt{T}r_0}|q_{\sqrt{t}}|^2. \tag{97}$$

For later use we note that because of stationarity of the fields q_T-q_t and $(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}$, taking the square expectation of (96) implies in particular

$$|\langle q_T\rangle-\langle q_t\rangle|\leq\langle|q_T-q_t|^2\rangle^{\frac{1}{2}}\ \lesssim\ \frac{1}{\delta}\langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\rangle^{\frac{1}{2}}+\delta. \tag{98}$$

We now turn to the estimate of g_T. We use (84) for $t=T$ and obtain by the triangle inequality

$$\Big(\int \eta_{\sqrt{T}r_0}|g_T|^2\Big)^{\frac{1}{2}}$$
$$\lesssim \Big(\int \eta_{\sqrt{T}r_0}|(q_T-q_t,\nabla\phi_T-\nabla\phi_t,q_t-\langle q_t\rangle,\nabla\phi_t)_{\delta\sqrt{T}}|^2\Big)^{\frac{1}{2}}$$
$$+|\langle q_T\rangle-\langle q_t\rangle|+\delta.$$

By the monotonicity (97) in the convolution scale this yields

$$\Big(\int \eta_{\sqrt{T}r_0}|g_T|^2\Big)^{\frac{1}{2}}$$
$$\lesssim \Big(\int \eta_{\sqrt{T}r_0}|(q_T-q_t,\nabla\phi_T-\nabla\phi_t)|^2\Big)^{\frac{1}{2}}$$
$$+\Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle,\nabla\phi_t)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}}+|\langle q_T\rangle-\langle q_t\rangle|+\delta.$$

Inserting (96) and (98) into this yields (95).

Step 4. Post-processing.
We first note that we may get rid of the $\nabla\phi_t$-term on the rhs of (95): Indeed, by (20) we have $\nabla\phi_t=t\nabla\nabla\cdot(q_t-\langle q_t\rangle)$ and thus $(\nabla\phi_t)_{\sqrt{2}\delta\sqrt{t}}=t(\nabla\nabla\cdot(q_t-\langle q_t\rangle)_{\delta\sqrt{t}})_{\delta\sqrt{t}}$, yielding $|(\nabla\phi_t)_{\sqrt{2}\delta\sqrt{t}}|\lesssim\frac{1}{\delta^2}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}})|_{2\delta\sqrt{t}}$ and therefore $|(\nabla\phi_t)_{\sqrt{2}\delta\sqrt{t}}|^2\lesssim\frac{1}{\delta^4}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}})|^2_{2\delta t}$. Since $(\eta_{\sqrt{T}r_0})_{2\delta\sqrt{t}}\overset{(79)}{\lesssim}\eta_{\sqrt{T}r_0}$ this implies as desired

$$\Big(\int \eta_{\sqrt{T}r_0}|(\nabla\phi_t)_{\sqrt{2}\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}}\lesssim\frac{1}{\delta^2}\Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\Big)^{\frac{1}{2}}. \tag{99}$$

Next we argue that

$$\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2 \lesssim \frac{1}{\delta^{\frac{d}{2}}}\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|, \tag{100}$$

which amounts to showing

$$\sup|(q_t-\langle q_t\rangle)_r| \lesssim (\frac{\sqrt{t}}{r})^{\frac{d}{2}} \tag{101}$$

for $r\le\sqrt{t}$, and which entails by stationarity

$$\langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\rangle \lesssim \frac{1}{\delta^{\frac{d}{2}}}\langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|\rangle. \tag{102}$$

In order to establish (101), we note that by Hölder's inequality and since the exponential localization function dominates the Gaussian kernel of the same scale

$$|(q_t)_r| \lesssim \Big(\int \eta_r|q_t|^2\Big)^{\frac12} \lesssim (\frac{\sqrt{t}}{r})^{\frac{d}{2}}\Big(\int \eta_{\sqrt{t}}|q_t|^2\Big)^{\frac12} \lesssim (\frac{\sqrt{t}}{r})^{\frac{d}{2}}, \tag{103}$$

where the last estimate follows from the uniform bound (51) $(\int \eta_{\sqrt{t}}|q_t|^2)^{\frac12}\lesssim 1$ from Lemma 10. Taking the expectation of (103), we obtain by stationarity $|\langle q_t\rangle| = |\langle (q_t)_r\rangle| \le \langle|(q_t)_r|\rangle \lesssim (\frac{\sqrt{t}}{r})^{\frac{d}{2}}$ so that also $|(q_t-\langle q_t\rangle)_r|\lesssim(\frac{\sqrt{t}}{r})^{\frac{d}{2}}$. Since this holds with the origin replaced by any point, we obtain (101).

The estimates (99), (100) & (102) combine to the following estimate on the rhs of (95)

$$\begin{aligned}
&\frac{1}{\delta}\Big(\int \eta_{\sqrt{T}r_0}|(\nabla\phi_t, q_t-\langle q_t\rangle)_{\sqrt{2}\delta t}|^2 + \langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|^2\rangle\Big)^{\frac12}\\
&\lesssim \frac{1}{\delta^{\frac{d}{4}+3}}\Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta t}| + \langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|\rangle\Big)^{\frac12}\\
&\lesssim \frac{1}{\delta^{\frac{d}{2}+7}}\Big(\int \eta_{\sqrt{T}r_0}|(q_t-\langle q_t\rangle)_{\delta t}| + \langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|\rangle\Big) + \delta,
\end{aligned}$$

where we used Young's inequality in the last step.

5.6. **Proof of Corollary 3: control of the bad event.** By the deterministic estimate (36) of the modified corrector in Lemma 5 (where we may assume $r_0\ge 2$), (37) amounts to showing that the centered random variable $F := |(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}| - \langle|(q_t-\langle q_t\rangle)_{\delta\sqrt{t}}|\rangle$, seen as a stationary field, satisfies the estimate

$$\log\Big\langle I\Big((\int \eta_{\sqrt{T}r_0}F \ge \delta^{\frac{d}{2}+8}\Big)\Big\rangle \lesssim -(\frac{\sqrt{T}}{\sqrt{t}}\frac{1}{\log(e+\frac{\sqrt{T}}{\sqrt{t}})})^d\delta^{d+16}.$$

Replacing the scale $\delta\sqrt{t}$ by a general $r \le \sqrt{t}$ and then redefining δ, this amounts to showing

$$\log\Big\langle I\Big(\big(\int \eta_{\sqrt{T}r_0}\big(|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle\big) \ge \delta\Big)\Big\rangle$$
$$\lesssim -\big(\frac{\sqrt{T}}{\sqrt{t}}\frac{1}{\log(e+\frac{\sqrt{T}}{\sqrt{t}})}\big)^d\delta^2.$$

Because of Chebychef's inequality in form of $\log\langle I(F\ge\delta)\rangle \le -\nu\delta + \log\langle\exp(\nu F)\rangle$ applied to the centered random variable $F := \int \eta_{\sqrt{T}r_0}\big(|(q_t-\langle q_t\rangle)_r - \langle|(q_t-\langle q_t\rangle)_r|\rangle\big)$, this is a consequence of the Gaussian estimate

$$\log\Big\langle \exp\Big(\nu\int \eta_{\sqrt{T}r_0}\big(|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle\big)\Big)\Big\rangle \lesssim \big(\frac{\sqrt{t}}{\sqrt{T}}\log(e+\frac{\sqrt{T}}{\sqrt{t}})\big)^d\nu^2$$

and subsequent optimization in ν. We note that because of $\eta_{\sqrt{T}r_0} \sim (\eta_{\sqrt{T}r_0})_{\sqrt{T}}$, cf. (79), the convexity of $F \mapsto \log\langle\exp(F)\rangle$, and the stationarity of $|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle$, this follows from

$$\log\Big\langle \exp\Big(\nu\big(|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle\big)_{\sqrt{T}}\Big)\Big\rangle \lesssim \big(\frac{\sqrt{t}}{\sqrt{T}}\log(e+\frac{\sqrt{T}}{\sqrt{t}})\big)^d\nu^2. \quad (104)$$

We now give the argument for estimate (104). We will use two facts established in Lemma 10

- For $r \le \sqrt{t}$, the centered random variable $|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle$ is approximately local on scale $\sqrt{t}$, cf. (50).
- It is also uniformly bounded in the sense of $\big(|(q_t-\langle q_t\rangle)_r| - \langle|(q_t-\langle q_t\rangle)_r|\rangle\big)_{\sqrt{t}} \lesssim 1$, cf. (51), which in particular implies a Gaussian bound, cf. (46) with $\Lambda = 1$.

Hence by stochastic cancellations, cf. (47) with $R = \sqrt{T}$ in Lemma 9, we indeed obtain (104).

5.7. **Proof of Lemma 6: nonlinear estimate.** Recall the decomposition

$$\big(\int_0^T \nabla u d\tau, q(T)\big)_R = F_0 + F_1,$$

where

$$F_0 := \big(\int_0^t \nabla u d\tau, q(t)\big)_R + (\mathrm{id}, a_{hT})\big(\nabla\int_t^T v d\tau\big)_R,$$
$$F_1 := \big(\int_t^T \nabla u - \nabla v d\tau, \int_t^T a\nabla u - a_{hT}\nabla v d\tau\big)_R,$$

cf. (15), (16), and (17). We divide the proof into four steps. We start by the main stochastic ingredient in the first step. It is then used to control both F_0 and F_1 in the second and third steps. We conclude in the last step.

Step 1. Claim that stochastic cancellations lead to

$$
\begin{aligned}
&\big\langle \exp \big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_R\big|\big\rangle \\
\lesssim \quad & \exp\big(C(\frac{\sqrt{t}}{\sqrt{T}})^d \Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)\nu^2\big) \quad \text{for all } R\le\sqrt{T}. \qquad (105)
\end{aligned}
$$

Here comes the argument for (105). By Definition 1 of $\Lambda(t)$, we have

$$
\begin{aligned}
&\big\langle \exp\big(\nu(\frac{r}{\sqrt{t}})^{\frac{d}{2}+1}(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_r\big)\big\rangle \\
&\qquad\qquad \le \exp\big(\Lambda(t)\nu^2\big) \quad \text{for } r\le\sqrt{t}. \qquad (106)
\end{aligned}
$$

By the elementary estimate $\exp|F| \le \exp(-F)+\exp(F)$, and by redefining ν, it follows

$$
\begin{aligned}
&\big\langle \exp \big|\nu(\frac{r}{\sqrt{T}})^{\frac{d}{2}+1}(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_r\big|\big\rangle \\
&\qquad\qquad \lesssim \exp\big((\frac{\sqrt{t}}{\sqrt{T}})^{d+2}\Lambda(t)\nu^2\big) \quad \text{for } r\le\sqrt{t}.
\end{aligned}
$$

By the trivial estimate $(\frac{\sqrt{t}}{\sqrt{T}})^{d+2} \le (\frac{\sqrt{t}}{\sqrt{T}})^d\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)$ this yields (105) in the range of $R\le\sqrt{t}$. If we specify (106) to $r=\sqrt{t}$ we obtain for the centered random variable $(\int_0^t\nabla u d\tau, q(t)-\langle q(t)\rangle)_{\sqrt{t}}$ the Gaussian bound $\big\langle \exp\big(\nu(\int_0^t\nabla u d\tau, q(t)-\langle q(t)\rangle)_{\sqrt{t}}\big)\big\rangle \le \exp\big(\Lambda(t)\nu^2\big)$. Since in addition, it is approximately local on scale $\sqrt{t}$, see (48) in Lemma 10, we obtain the following stochastic cancellations, see Lemma 9:

$$
\begin{aligned}
&\big\langle \exp\big(\nu(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_R\big)\big\rangle \\
&\qquad \le \exp\big(C\Lambda(t)(\frac{\sqrt{t}}{R})^d\log^d(e+\frac{1}{\Lambda(t)}(\frac{R}{\sqrt{t}})^d)\nu^2\big) \quad \text{for } R\ge\sqrt{t}. \qquad (107)
\end{aligned}
$$

By a redefinition of ν and using $\exp|F|\le\exp(F)+\exp(-F)$ we rewrite this as

$$
\begin{aligned}
&\big\langle \exp \big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}}(\int_0^t \nabla u d\tau, q(t) - \langle q(t)\rangle)_R\big|\big\rangle \\
&\qquad \lesssim \exp\big(C(\frac{\sqrt{t}}{\sqrt{T}})^d\Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{R}{\sqrt{t}})^d)\nu^2\big) \quad \text{for } R\ge\sqrt{t},
\end{aligned}
$$

which entails (105) in the remaining range $\sqrt{t}\le R\le\sqrt{T}$ by the trivial estimates and $(\frac{R}{\sqrt{T}})^{\frac{d}{2}} \ge (\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}$ and $\log(e+\frac{1}{\Lambda(t)}(\frac{R}{\sqrt{t}})^d)\le\log(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)$.

Step 2. Control of the dominant term F_0, cf. (16).

We start with an estimate of the second contribution to F_0 as follows

$$\begin{aligned}\Big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T(\mathrm{id},a_{hT})(\nabla v)_R d\tau\big|\Big\rangle\\ \lesssim \exp\big(C(\frac{\sqrt{t}}{\sqrt{T}})^d\Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)\nu^2\big).\end{aligned} \tag{108}$$

To this purpose we appeal to the deterministic estimate (31) from Lemma 3, which we just use for $k=2$ and which we simplify by giving up the scale r^2 in the prefactor and in $(q(t)-\langle q(t)\rangle)_{\sqrt{\tau+r^2+R^2}}$ (using Jensen's inequality and the semi-group property of the convolution with a Gaussian for the latter):

$$|(\nabla v(t+\tau))_R| \lesssim \frac{1}{\tau+R^2}|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+R^2}}|_{C\sqrt{\tau+r^2+R^2}}.$$

We rewrite this as

$$\begin{aligned}&(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}|(\nabla v(t+\tau))_R|\\ &\quad\lesssim \omega_R(\tau)\,(\frac{\sqrt{\tau+R^2}}{\sqrt{T}})^{\frac{d}{2}+1}|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+R^2}}|_{C\sqrt{\tau+r^2+R^2}},\end{aligned}$$

where $\omega_R(\tau):=\frac{1}{\tau+R^2}(\frac{R}{\sqrt{\tau+R^2}})^{\frac{d}{2}+1}$ satisfies $\int_0^\infty\omega_R d\tau\sim 1$, and thus, after correction by a multiplicative constant, defines a convex combination in τ. By the convexity of $F\mapsto\langle\exp|F|\rangle$ and stationarity of $|q(t)-\langle q(t)\rangle|_{\frac{1}{C}\sqrt{\tau+R^2}}$ this yields

$$\begin{aligned}&\Big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T(\nabla v)_R d\tau\big|\Big\rangle\\ &\quad\lesssim \int_0^{T-t}d\tau\,\omega_R\Big\langle \exp\big(C\nu(\frac{\frac{1}{C}\sqrt{\tau+R^2}}{\sqrt{T}})^{\frac{d}{2}+1}|(q(t)-\langle q(t)\rangle)_{\frac{1}{C}\sqrt{\tau+R^2}}|\big)\Big\rangle.\end{aligned}$$

Since for $\tau\le T$ we have $\frac{1}{C}\sqrt{\tau+R^2}\le\sqrt{T}$, we may apply (105) with R replaced by $\frac{1}{C}\sqrt{\tau+R^2}$, and using once more $\int_0^\infty\omega_R d\tau\sim 1$ we obtain (108) using the upper bound (25) on a_{hT}.

We now complete the estimate of the dominant term F_0, cf. (16):

$$\Big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(F_0-\langle F_0\rangle)\big|\Big\rangle \lesssim \exp\big(C(\frac{\sqrt{t}}{\sqrt{T}})^d\Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)\nu^2\big). \tag{109}$$

Because of $F_0=(\int_0^t\nabla u d\tau, q(t))_R+\int_t^T(\mathrm{id},a_{hT})(\nabla v)_R d\tau$, this is an immediate consequence of (105) and (108) and the convexity of $\exp|F|$ in form of $\exp|F+F'|\le\frac{1}{2}\exp|2F|+\frac{1}{2}\exp|2F'|$.

Step 3. Control of the homogenization error F_1, cf. (17).

Roughly speaking, this is a combination of Corollary 2 (with Λ replaced by the quantity $\sup_{t'\in[\frac{t}{2},t]}\Lambda(t')$), which with our abbreviation takes the form of

$$\big\langle I(\mathcal{G}_\delta)\exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}F_1\big|\big\rangle \lesssim \exp\big(C\sup_{t'\in[\frac{t}{2},t]}\Lambda(t')(\frac{\sqrt{T}}{\sqrt{t}})^{d+4}\delta^{\frac{4}{d+10}}\nu^2\big), \tag{110}$$

and estimate (37) in Corollary 3, which we use with t replaced by t_0:

$$\big\langle I(\mathcal{G}_\delta^c)\big\rangle \lesssim \exp\big(-\frac{1}{C}(\frac{\sqrt{T}}{\sqrt{t_0}})^d\frac{1}{\log^d(e+\frac{\sqrt{T}}{\sqrt{t_0}})}\delta^{d+16}\big).$$

In order to conclude we need an estimate of F_1 conditioned on the bad event $\mathcal{G}_\delta^c$. To this purpose, we split $F_1=\int_t^T((\mathrm{id},a)\nabla u)_R d\tau-\int_t^T(\mathrm{id},a_{hT})(\nabla v)_R d\tau$; for the second summand, we appeal to (108). For the first summand, we obtain by the uniform estimate (49) of Lemma 10

$$\begin{aligned}
(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\Big|\int_t^T((\mathrm{id},a)\nabla u)_R d\tau\Big| &\le (\frac{R}{\sqrt{T}})^{\frac{d}{2}}\Big|((\mathrm{id},a)\int_t^T\nabla u d\tau)_R\Big| \\
&\overset{(1),(71)}{\lesssim} \Big(\int\eta_{\sqrt{T}}\Big|\int_t^T\nabla u d\tau\Big|^2\Big)^{\frac{1}{2}} \\
&\le \Big(\int\eta_{\sqrt{T}}\Big|\int_0^T\nabla u d\tau\Big|^2\Big)^{\frac{1}{2}}+\Big(\int\eta_{\sqrt{T}}\Big|\int_0^t\nabla u d\tau\Big|^2\Big)^{\frac{1}{2}}\lesssim 1.
\end{aligned}$$

(Note that we have also $\int\eta_{\sqrt{T}}|\int_0^t\nabla u d\tau|^2\lesssim 1$, since this can be obtained from (49) in form of $\int\eta_{\sqrt{t}}|\int_0^t\nabla u d\tau|^2\lesssim 1$ by replacing $\eta_{\sqrt{t}}$ by $\eta_{\sqrt{t}}(\cdot-y)$, averaging over y the wrt $\eta_{\sqrt{T}}(y)$, and appealing to (79).) The combination of these observations yields

$$\begin{aligned}
&\big\langle I(\mathcal{G}_\delta^c)\exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}F_1\big|\big\rangle \\
&\le \frac{1}{2}\big\langle I(\mathcal{G}_\delta^c)\exp\big|2\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T((\mathrm{id},a)\nabla u)_R d\tau\big|\big\rangle \\
&+\frac{1}{2}\big\langle\exp\big|2\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\int_t^T(\mathrm{id},a_{hT})(\nabla v)_R d\tau\big|\big\rangle \\
&\lesssim \exp\big(-\frac{1}{C}(\frac{\sqrt{T}}{\sqrt{t_0}})^d\frac{1}{\log^d(e+\frac{\sqrt{T}}{\sqrt{t_0}})}\delta^{d+16}\big)\ \exp\big(C|\nu|\big) \\
&+\exp\big(C(\frac{\sqrt{t}}{\sqrt{T}})^d\Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d)\nu^2\big).
\end{aligned}$$

Adding this to (110), and applying Young's inequality on the argument of the rhs exponential we obtain for the homogenization error

$$\big\langle\exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}F_1\big|\big\rangle\lesssim\exp(C\Lambda\nu^2), \tag{111}$$

where we have set for abbreviation

$$\Lambda := \sup_{t'\in[\frac{t}{2},t]} \Lambda(t')(\frac{\sqrt{T}}{\sqrt{t}})^{d+4}\delta^{\frac{4}{d+10}} + (\frac{\sqrt{t_0}}{\sqrt{T}})^d \log^d(e+\frac{\sqrt{T}}{\sqrt{t_0}})\frac{1}{\delta^{d+16}}$$
$$+ C(\frac{\sqrt{t}}{\sqrt{T}})^d \Lambda(t)\log^d(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^d), \tag{112}$$

and which already has the form of the propagation rule (40).

Step 4. Conclusion.
We first claim that (111) implies

$$\big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(F_1-\langle F_1\rangle)\big|\big\rangle \lesssim \exp(C\Lambda\nu^2). \tag{113}$$

Indeed, by Jensen's inequality in form of $\exp|\nu\langle F\rangle| \le \langle \exp|\nu F|\rangle$ we obtain from (111) that $|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\langle F_1\rangle| \le \Lambda\nu^2+C$, which by optimization in ν yields $|(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\langle F_1\rangle| \lesssim \sqrt{\Lambda}$ and thus $\langle \exp|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}\langle F_1\rangle|\rangle \lesssim \exp(C\sqrt{\Lambda}|\nu|) \lesssim \exp(C\Lambda\nu^2)$, so that (113) follows from (111). Adding (113) to (109), we reach

$$\big\langle \exp\big|\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(F_0+F_1-\langle F_0+F_1\rangle)\big|\big\rangle \lesssim \exp(C\Lambda\nu^2).$$

Since however $F_0+F_1-\langle F_0+F_1\rangle$ is centered, this improves to

$$\big\langle \exp\big(\nu(\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(F_0+F_1-\langle F_0+F_1\rangle)\big)\big\rangle \le \exp(C\Lambda\nu^2)$$

by (150) in the proof of Lemma 9. In view of $(\int_0^T \nabla u d\tau, q(T)-\langle q(T)\rangle)_R = F_0 + F_1 - \langle F_0+F_1\rangle$, cf. (15), and the fact that $\int_0^T \nabla u d\tau$ is centered by shift-invariance of u and stationarity of $\langle\cdot\rangle$, this yields $\Lambda(T) \le C\Lambda$ by Definition 1 of $\Lambda(T)$. In view of the definition (112) of Λ, this implies the propagation estimate (40).

5.8. **Proof of Lemma 7: buckling lemma.** We split the proof into five steps.

Step 1. Reformulation of the proviso (41).
We introduce the following average of $\{\Lambda(t)\}_t$:

$$\Lambda_T := \int_0^\infty \exp(-\frac{t}{T})\max\{(\frac{t}{T})^{\frac{d}{2}+1},(\frac{t}{T})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{T}, \tag{114}$$

where the notation $\{\Lambda_T\}_T$ is to evoke the relation between $\{q(t)\}_t$ and $\{q_T\}_T$. We claim that (41) in Lemma 6 is met when

$$\Lambda_{t_0} \ll \frac{\delta^{2d+18}}{\log^d(e+\frac{1}{\delta})}. \tag{115}$$

Here comes the argument: We recall that q_T is a convex combination of $\{q(t)\}_t$ in the sense of $q_T = \int_0^\infty \exp(-\frac{t}{T})q(t)\frac{dt}{T}$, cf. (21), which we use in form of $(q_T - \langle q_T\rangle)_R$

$= \int_0^\infty \exp(-\frac{t}{T})(q(t)-\langle q(t)\rangle)_R\frac{dt}{T}$, which in turn implies by Jensen's inequality $\langle|(q_T-\langle q_T\rangle)_R|^2\rangle \le \int_0^\infty \exp(-\frac{t}{T})\langle|(q(t)-\langle q(t)\rangle)_R|^2\rangle\frac{dt}{T}$. We split the integral into two parts

$$\begin{aligned}\langle|(q_T-\langle q_T\rangle)_R|^2\rangle \le &\int_0^{R^2} \exp(-\frac{t}{T})\langle|(q(t)-\langle q(t)\rangle)_R|^2\rangle\frac{dt}{T}\\ &+\int_{R^2}^\infty \exp(-\frac{t}{T})\langle|(q(t)-\langle q(t)\rangle)_R|^2\rangle\frac{dt}{T}.\end{aligned} \tag{116}$$

Because of $\langle|F|^2\rangle \le \lim_{\nu\downarrow 0}\frac{1}{\nu^2}\log\langle\exp(\nu F)\rangle$ for a centered random variable F, the second integral is directly estimated by $\int_{R^2}^\infty \exp(-\frac{t}{T})(\frac{t}{R^2})^{\frac{d}{2}+1}\Lambda(t)\frac{dt}{T}$ according to Definition 1 of $\Lambda(t)$. Turning to the first rhs term of (116) we appeal to the stochastic cancellations already used in (107) in the proof of Lemma 6:

$$\Big\langle\big|(q(t)-\langle q(t)\rangle)_R\big|^2\Big\rangle \lesssim \Lambda(t)(\frac{\sqrt{t}}{R})^d\log^d(e+\frac{1}{\Lambda(t)}(\frac{R}{\sqrt{t}})^d)\quad\text{for } R\ge\sqrt{t}. \tag{117}$$

In order to overcome the nonlinearity in $\Lambda(t)$ coming from the logarithm in (117), we use the elementary estimate $\Lambda\log^d(e+\frac{1}{\Lambda}) \lesssim \Lambda\log^d(e+\frac{1}{\delta})+\delta$ for any $\Lambda,\delta\lesssim 1$ (which we may apply to $\Lambda=\Lambda(t)(\frac{\sqrt{t}}{R})^d \le \Lambda(t)\ll 1$ because of our anchoring, and which by taking the d-th root and a redefinition of Λ,δ follows from $\Lambda\log(e+\frac{1}{\Lambda})$ $\lesssim \Lambda\log(e+\frac{1}{\delta})+\delta$, which in terms of $a=\frac{1}{\Lambda}, a_0=\frac{1}{\delta}\gtrsim 1$ can be rewritten as $\log(e+a)$ $\lesssim \log(e+a_0)+\frac{a}{a_0}$, which follows from concavity of $a\mapsto\log a$ for $a\ge a_0$ and from monotonicity for $a\le a_0$). Equipped with this elementary estimate, we rewrite (117) as

$$\Big\langle\big|(q(t)-\langle q(t)\rangle)_R\big|^2\Big\rangle \lesssim \Lambda(t)(\frac{t}{R^2})^{\frac{d}{2}}\log^d(\frac{1}{\delta})+\delta\quad\text{for } R\ge\sqrt{t},\quad \delta\ll 1.$$

Inserting also this into (116) we obtain

$$\langle|(q_T-\langle q_T\rangle)_R|^2\rangle \le \log^d(\frac{1}{\delta})\int_0^\infty \exp(-\frac{t}{T})\max\{(\frac{t}{R^2})^{\frac{d}{2}+1},(\frac{t}{R^2})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{T}+\delta.$$

Multiplying this with $(\frac{R^2}{T})^{\frac{d}{2}+1}\le(\frac{R^2}{T})^{\frac{d}{2}}$ for $R\le\sqrt{T}$, we obtain by definition (114) of Λ_T

$$(\frac{R^2}{T})^{\frac{d}{2}+1}\langle|(q_T-\langle q_T\rangle)_R|^2\rangle \le \log^d(\frac{1}{\delta})\Lambda_T+(\frac{R^2}{T})^{\frac{d}{2}+1}\delta\quad\text{for }\ R\le\sqrt{T},\ \ \delta\ll 1. \tag{118}$$

We now use (118) with δ replaced by δ^{d+16}, with T replaced by t_0, and with R replaced by $\delta\sqrt{t_0}$:

$$\delta^{d+2}\langle|(q_{t_0}-\langle q_{t_0}\rangle)_{\delta\sqrt{t_0}}|^2\rangle \le \log^d(\frac{1}{\delta})\Lambda_{t_0}+\delta^{d+2}\delta^{d+16}\quad\text{for }\ \delta\ll 1. \tag{119}$$

From this we learn that (115) implies (41).

Step 2. Reformulation of the propagation rule.
It is convenient to have $\Lambda(t')\sim\Lambda(t)$ for $t'\sim t$. We shall thus replace $\Lambda(t)$ by $\sup_{T\ge t}(\frac{\sqrt{t}}{\sqrt{T}})^{d+4}\Lambda(T)$, which stays finite by Lemma 8. Since the rhs of the propagation

estimate (40) does not increase faster than $\sqrt{T}^{d+4}$, it also holds for the $\Lambda(t)$ replaced by $\sup_{T\ge t}(\frac{\sqrt{t}}{\sqrt{T}})^{d+4}\Lambda(T)$. Hence we will from now on assume that $\Lambda(t)$ satisfies

$$\frac{1}{\sqrt{T}^{d+4}}\Lambda(T)\le\frac{1}{\sqrt{t}^{d+4}}\Lambda(t)\quad\text{for } T\ge t, \tag{120}$$

which allows to simplify the propagation estimate (40) in as far as we may replace $\sup_{t'\in[\frac{t}{2},t]}\Lambda(t')$ (and $\Lambda(t)$) by $\Lambda(\frac{t}{2})$. Relabelling $\frac{t}{2}$ by t (and noting that the propagation rule is trivial unless $T\gg t$), it simplified to

$$\begin{aligned}\Lambda(T)&\lesssim(\frac{\sqrt{t}}{\sqrt{T}})^{d}\Lambda(t)\log^{d}(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^{d})\\&+\delta^{\frac{4}{d+10}}(\frac{\sqrt{T}}{\sqrt{t}})^{d+4}\Lambda(t)+\frac{1}{\delta^{d+16}}(\frac{\sqrt{t_0}}{\sqrt{T}}\log(e+\frac{\sqrt{T}}{\sqrt{t_0}}))^{d}.\end{aligned}\tag{121}$$

In terms of the multiplicative increment $M\gg 1$ in $T=Mt$, and using the elementary inequality $\log(e+\frac{1}{\Lambda(t)}(\frac{\sqrt{T}}{\sqrt{t}})^{d})\le\log(e+\frac{1}{\Lambda(t)})+\log(e+(\frac{\sqrt{T}}{\sqrt{t}})^{d})$, this reads as

$$\begin{aligned}\Lambda(Mt)\ &\lesssim\ \frac{1}{M^{\frac{d}{2}}}\Lambda(t)\log^{d}(e+\frac{1}{\Lambda(t)})\\&+\ \big(\frac{\log^{d}(e+M)}{M^{\frac{d}{2}}}+\delta^{\frac{4}{d+10}}M^{\frac{d}{2}+2}\big)\Lambda(t)\\&+\ \frac{1}{\delta^{d+16}}(\frac{\sqrt{t_0}}{\sqrt{M}}\log(e+\frac{\sqrt{M}}{\sqrt{t_0}}))^{d},\end{aligned}$$

where in the last summand, which is monotone increasing in $\frac{\sqrt{t_0}}{\sqrt{t}}$, we replaced $\frac{\sqrt{t_0}}{\sqrt{Mt}}$ by the latter. This motivates the choice of δ such that $\delta^{\frac{4}{d+10}}M^{\frac{d}{2}+2}=\frac{1}{M^{\frac{d}{2}}}$:

$$\begin{aligned}\Lambda(Mt)\ &\lesssim\ \frac{1}{M^{\frac{d}{2}}}\Lambda(t)\log^{d}(e+\frac{1}{\Lambda(t)})+\frac{\log^{d}(e+M)}{M^{\frac{d}{2}}}\Lambda(t)\\&+\ M^{p}(\frac{\sqrt{t_0}}{\sqrt{t}}\log(e+\frac{\sqrt{t}}{\sqrt{t_0}}))^{d},\end{aligned}\tag{122}$$

which in view of (115) holds provided

$$\Lambda_{t_0}\ll\frac{1}{M^{p}},\tag{123}$$

where $p=p(d)\gg 1$ denotes a generic exponent.

Step 3. Campanato iteration.
The form of the estimate (122) suggests a Campanato iteration; however, the first rhs term which because of the logarithm is (slightly) nonlinear in $\Lambda(t)$ requires an additional twist. Introducing the abbreviation $\Phi(\Lambda):=\Lambda\log^{d}(e+\frac{1}{\Lambda})$ we note that Φ is concave (for $\Lambda\ll 1$ which is satisfied because of our assumption in (42)) so

that $\Phi(\Lambda) \le \Phi'(\Lambda_0)(\Lambda-\Lambda_0)+\Phi(\Lambda_0)$. Provided the increment M is, in relation to the threshold $\Lambda_0 \ll 1$, chosen such that $\frac{1}{M^{\frac{d}{2}}}\Phi(\Lambda_0) \ll \Lambda_0$, which turns into

$$\log\frac{1}{\Lambda_0} \ll M^{\frac12}, \tag{124}$$

the iteration (122), using $\Phi'(\Lambda_0) = \log^d(e+\frac{1}{\Lambda_0}) - \frac{d}{1+e\Lambda_0}\log^{d-1}(e+\frac{1}{\Lambda_0}) \lesssim \log^d\frac{1}{\Lambda_0}$, turns into

$$\begin{aligned}(\Lambda(Mt)-\Lambda_0)_+ \ &\le\ C_0\frac{\log^d\frac{1}{\Lambda_0}+\log^d M}{M^{\frac{d}{2}}}(\Lambda(t)-\Lambda_0)_+ \\ &+\ CM^p(\frac{\sqrt{t_0}}{\sqrt{t}}\log(e+\frac{\sqrt{t}}{\sqrt{t_0}}))^d. \end{aligned}\tag{125}$$

We now are in a position to use Campanato's iteration: For given exponent $\alpha < \frac{d}{2}$ we make sure that the multiplicative increment M is so large that $C_0\frac{\log^d M}{M^{\frac{d}{2}}} \le \frac12\frac{1}{M^\alpha}$, which just requires $M \gg 1$, and then Λ_0 not too small such that $C_0\frac{\log^d\frac{1}{\Lambda_0}}{M^{\frac{d}{2}}} \le \frac12\frac{1}{M^\alpha}$, which amounts to a strengthening of (124) to

$$\log\frac{1}{\Lambda_0} \ll M^{\frac12-\frac{\alpha}{d}}. \tag{126}$$

With these provisions, (125) turns into

$$(\Lambda(Mt)-\Lambda_0)_+ \le \frac{1}{M^\alpha}(\Lambda(t)-\Lambda_0)_+ + CM^p(\frac{\sqrt{t_0}}{\sqrt{t}}\log(e+\frac{\sqrt{t}}{\sqrt{t_0}}))^d.$$

We use this for $t = M^n t_0$ to obtain the iteration

$$(\Lambda(M^{n+1}t_0)-\Lambda_0)_+ \le \frac{1}{M^\alpha}(\Lambda(M^n t_0)-\Lambda_0)_+ + CM^p(\frac{1+n\log M}{\sqrt{M}^n})^d,$$

which recursively implies

$$\begin{aligned}&(\Lambda(M^n t_0)-\Lambda_0)_+ \\ &\lesssim\ (\frac{1}{M^\alpha})^n(\Lambda(t_0)-\Lambda_0)_+ + M^p\sum_{m=0}^{n-1}(\frac{1}{M^\alpha})^{n-1-m}(\frac{1+m\log M}{\sqrt{M}^m})^d \\ &\lesssim\ (\frac{1}{M^\alpha})^n\big((\Lambda(t_0)-\Lambda_0)_+ + M^p\log^d M\big),\end{aligned}$$

where we have used $\alpha < \frac{d}{2}$ in the last step. Using (120) we obtain the continuum version of the above

$$(\Lambda(t)-\Lambda_0)_+ \lesssim M^{\frac{d}{2}+2}(\frac{t_0}{t})^\alpha\big((\Lambda(t_0)-\Lambda_0)_+ + M^p\log^d M\big),$$

which we rewrite as

$$\Lambda(t) \lesssim \Lambda_0 + M^{\frac{d}{2}+2}(\frac{t_0}{t})^\alpha\big(\Lambda(t_0)+M^p\log^d M\big).$$

Since the threshold Λ_0 was arbitrary besides the constraint (126), this yields

$$\Lambda(t) \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^{\frac{d}{2}+2}(\frac{t_0}{t})^\alpha\big(\Lambda(t_0) + M^p \log^d M\big).$$

After a redefinition of the generic exponent $p = p(d)$ we may rewrite the above as

$$\Lambda(t) \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{t})^\alpha(\Lambda(t_0) + 1), \tag{127}$$

which holds provided (123).

Step 4. Buckling on level of $\{\Lambda_t\}_t$.
In order to buckle, we re-express (127) in terms of Λ_t rather than $\Lambda(t)$. We note that it follows from the definition (114) of Λ_T that $\Lambda_T \gtrsim \frac{1}{T}\int_{\frac{T}{2}}^{T}\Lambda(t)dt$, so that by (120) we always have $\Lambda(t) \lesssim \Lambda_t$. Hence we may replace $\Lambda(t_0)$ by Λ_{t_0} on the rhs of (127):

$$\Lambda(t) \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{t})^\alpha(\Lambda_{t_0} + 1). \tag{128}$$

This allows us to insert (123) so that (128) simplifies to

$$\Lambda(t) \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{t})^\alpha. \tag{129}$$

For $T \geq t_0$, we multiply this estimate by $\frac{1}{T}\exp(-\frac{t}{T})\max\{(\frac{t}{T})^{\frac{d}{2}+1}, (\frac{t}{T})^{\frac{d}{2}}\}$ and integrate over $t \in (t_0, \infty)$ to obtain thanks to $\alpha < \frac{d}{2} + 1$

$$\int_{t_0}^{\infty}\exp(-\frac{t}{T})\max\{(\frac{t}{T})^{\frac{d}{2}+1}, (\frac{t}{T})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{T} \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{T})^\alpha.$$

Together with the trivial estimate

$$\begin{aligned}&\int_{0}^{t_0}\exp(-\frac{t}{T})\max\{(\frac{t}{T})^{\frac{d}{2}+1}, (\frac{t}{T})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{T}\\ &\lesssim (\frac{t_0}{T})^{\frac{d}{2}+1}\int_{0}^{t_0}\exp(-\frac{t}{t_0})\max\{(\frac{t}{t_0})^{\frac{d}{2}+1}, (\frac{t}{t_0})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{t_0} \overset{(114)}{\leq} (\frac{t_0}{T})^{\frac{d}{2}+1}\Lambda_{t_0},\end{aligned}$$

combined with (123) to $\int_0^{t_0}\exp(-\frac{t}{T})\max\{(\frac{t}{T})^{\frac{d}{2}+1}, (\frac{t}{T})^{\frac{d}{2}}\}\Lambda(t)\frac{dt}{T} \ll \frac{1}{M^p}(\frac{t_0}{T})^{\frac{d}{2}+1} \ll M^p(\frac{t_0}{T})^\alpha$ (where we used again $\alpha < \frac{d}{2} + 1$), we obtain by definition (114) of Λ_T

$$\Lambda_T \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{T})^\alpha.$$

Hence we have for all $t \geq t_0 \gg 1$ and $M \gg 1$

$$\Lambda_t \lesssim \exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) + M^p(\frac{t_0}{t})^\alpha \quad \text{provided} \quad \Lambda_{t_0} \ll \frac{1}{M^p}.$$

Given $t \geq t_0$ we now choose M so large that $\exp(-\frac{1}{C}M^{\frac{1}{2}-\frac{\alpha}{d}}) \lesssim M^p(\frac{t_0}{t})^\alpha$, which certainly is the case for $M^{\frac{1}{2}-\frac{\alpha}{d}} \gg \log(e + \frac{t}{t_0})$, so that the above turns into

$$\Lambda_t \lesssim (\frac{t_0}{t})^\alpha \log^p\frac{t}{t_0} \quad \text{provided} \quad \Lambda_{t_0} \ll \frac{1}{\log^p(e + \frac{t}{t_0})},$$

where for notational simplicity, we redefined $p = p(d,\alpha)$ to be $\frac{p}{\frac12-\frac{\alpha}{d}}$. Slightly redefining $\alpha < \frac{d}{2}$ to absorb the logarithm, we see that this implies

$$\Lambda_t \lesssim (\frac{t_0}{t})^\alpha \quad \text{provided} \quad \Lambda_{t_0} \ll \frac{1}{\log^p(e+\frac{t}{t_0})}.$$

Since for $t \gg t_0$ we have $(\frac{t_0}{t})^\alpha \ll \frac{1}{\log^p \frac{t}{t_0}}$, this implication is self-sustaining provided we have $\Lambda_{t_0} \ll 1$. This is best seen as follows. Define $t_*(t) := t_0 \log t$. Then

$$\Lambda_t \lesssim (\frac{t_*(t)}{t})^\alpha \quad \text{provided} \quad \Lambda_{t_*(t)} \ll \frac{1}{\log^p(e+\log t)},$$

the validity of which propagates due to the chain of inequalities

$$\frac{1}{\log^p(e+\log t)} \gg (\frac{\log\log t}{\log t})^\alpha \sim (\frac{t_*(t_*(t))}{t_*(t)})^\alpha \gtrsim \Lambda_{t_*(t)}$$

and the assumption $\Lambda_{t_0} \ll 1$. Slightly redefining $\alpha < \frac{d}{2}$ to absorb the logarithm, we have as desired

$$\Lambda_t \lesssim (\frac{t_0}{t})^\alpha \quad \text{for all } t \geq t_0 \quad \text{provided} \quad \Lambda_{t_0} \ll 1. \tag{130}$$

Step 5. Back to $\{\Lambda(t)\}_t$.
We now go back to (129) and (123), in which we relabel (t, t_0) by (T, t):

$$\Lambda(T) \lesssim \exp(-\frac{1}{C}M^{\frac12-\frac{\alpha}{d}}) + M^p(\frac{t}{T})^\alpha \quad \text{provided} \quad \Lambda_t \ll \frac{1}{M^p}.$$

By (130), provided we chose t just so large that $(\frac{t_0}{t})^\alpha \ll \frac{1}{M^p}$, this implies

$$\Lambda(T) \lesssim \exp(-\frac{1}{C}M^{\frac12-\frac{\alpha}{d}}) + M^p(\frac{t_0}{T})^\alpha \quad \text{provided} \quad \Lambda_{t_0} \ll 1,$$

where we recall that p is a generic exponent whose value may change from line to line. We finally choose M just so large that $\exp(-\frac{1}{C}M^{\frac12-\frac{\alpha}{d}}) \leq (\frac{t_0}{T})^\alpha$ so that the above turns into $\Lambda(T) \lesssim (\frac{t_0}{T})^\alpha \log^p \frac{T}{t_0}$ with a redefined $p = p(d,\alpha)$. As above, by a redefinition of α, this may be rewritten as

$$\Lambda(T) \lesssim (\frac{t_0}{T})^\alpha \quad \text{for all } T \geq t_0 \text{ provided} \quad \Lambda_{t_0} \ll 1.$$

It remains to argue that $\Lambda(t_0) \ll 1$ entails $\Lambda_{t_0} \ll 1$. To this purpose, for some $\delta \ll 1$, we split definition (114) into $\frac{1}{t_0}\int_0^{\delta t_0} \exp(-\frac{t}{t_0})(\frac{t}{t_0})^{\frac{d}{2}}\Lambda(t)dt$, which by the uniform bound (43) on $\Lambda(t)$ is $\lesssim \delta$, and into $\frac{1}{t_0}\int_{\delta t_0}^\infty \exp(-\frac{t}{t_0})\max\{(\frac{t}{t_0})^{\frac{d}{2}}, (\frac{t}{t_0})^{\frac{d}{2}+1}\}\Lambda(t)dt$, which by (120) is $\lesssim \int_{\delta t_0}^\infty \exp(-\frac{t}{t_0})\max\{(\frac{t}{t_0})^{\frac{d}{2}}, (\frac{t}{t_0})^{\frac{d}{2}+1}\}(\frac{t}{\delta t_0})^{\frac{d}{2}+2}\frac{dt}{t_0}\Lambda(\delta t_0) \lesssim \frac{1}{\delta^{\frac{d}{2}+2}}\Lambda(\delta t_0) \ll 1$ for t_0 large enough in function of $\delta > 0$.

5.9. **Proof of Lemma 8: anchoring lemma.** This lemma relies on the following two ingredients. The first ingredient is stochastic and based on H-convergence: For all $\delta > 0$,

$$\lim_{T\uparrow\infty}\langle|(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_{\delta\sqrt{T}}|\rangle = 0, \tag{131}$$

where the limit (131) is uniform wrt the ensemble $\langle\cdot\rangle$, up to a dependence on the dimension d and the ellipticity ratio $\lambda > 0$ (the argument is displayed in the last four steps of the proof). The second ingredient is deterministic and relies on the uniform estimate (49) from Lemma 10; together with (71) we obtain

$$(\frac{R}{\sqrt{T}})^{\frac{d}{2}}|(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_R| \lesssim 1. \tag{132}$$

Step 1. Proof of the claim.
We argue that we may restrict ourselves to the range $1 \le |\nu| \lesssim 1$ and $R \sim \sqrt{T}$ in the definition (39) of $\Lambda(T)$ (with (r,t) replaced by (R,T)) when proving (43). Indeed, on the one hand, since for any centered random variable F like our $F_R(T)$ $:= (\frac{R}{\sqrt{T}})^{\frac{d}{2}+1}(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_R$ we have the inequality $\sup_\nu \frac{1}{\nu^2}\log\langle\exp(\nu F)\rangle$ $\lesssim \sup_{|\nu|\ge 1}\frac{1}{\nu^2}\log\langle\exp(\nu F)\rangle$ (see the argument for (150) in the proof of Lemma 9), we may restrict to $|\nu| \ge 1$. On the other hand, we obtain from (132) that in particular $|F_R(T)| \lesssim \frac{R}{\sqrt{T}}$ which for $|\nu| \ge 1$ and since $R \le \sqrt{T}$ entails $\frac{1}{\nu^2}\log\langle\exp(\nu F_R(T))\rangle$ $\lesssim \frac{1}{|\nu|}\frac{R}{\sqrt{T}} \lesssim \min\{\frac{1}{|\nu|}, \frac{R}{\sqrt{T}}\}$. While this already implies the uniform bound in (43) it shows that for the convergence, it is enough to establish for any $\delta > 0$

$$\lim_{T\uparrow\infty}\sup_{\delta\sqrt{T}\le R\le\sqrt{T}, 1\le|\nu|\le\frac{1}{\delta}}\frac{1}{\nu^2}\log\big\langle\exp\big(\nu F_R(T)\big)\big\rangle = 0.$$

We note that by the convexity of $F \mapsto \log\langle\exp(F)\rangle$ in conjunction with the semi-group property of convolution with a Gaussian $f_R = (f_{\delta\sqrt{T}})_{\sqrt{R^2-\delta^2T}}$ applied to our stationary random field $F(T) := (\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)$ we have the inequality $\log\langle\exp(\nu F(T)_R)\rangle \le \log\langle\exp(\nu F(T)_{\delta\sqrt{T}})\rangle$. Hence it is enough to show

$$\lim_{T\uparrow\infty}\log\big\langle\exp\big|\delta^{-1}F(T)_{\delta\sqrt{T}}\big|\big\rangle = 0. \tag{133}$$

By (132) in form of $\delta^{\frac{d}{2}}|F(T)_{\delta\sqrt{T}}| \lesssim 1$ which yields $\exp\big|\delta^{-1}F(T)_{\delta\sqrt{T}}\big| \le \exp(C\delta^{-1-\frac{d}{2}})$, we obtain (133) from (131) by Lebesgue's dominated convergence.

The last four steps are devoted to proving (131), and in particular its uniformity wrt $\langle\cdot\rangle$.

Step 2. Rescaling: We have

$$\langle|(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)_{\delta\sqrt{T}}|^2\rangle = \langle|(\int_0^1 \nabla u d\tau, q(1) - \langle q(1)\rangle_{\sqrt{T}})_\delta|^2\rangle_{\sqrt{T}}, \tag{134}$$

where the measure $\langle\cdot\rangle_{\sqrt{T}}$ denotes the push-forward of $\langle\cdot\rangle$ under the map $a \mapsto \hat{a}$, where $\hat{a}(\hat{x}) = a(\sqrt{T}\hat{x})$ is the coefficient field rescaled by $\sqrt{T}$. Here comes the argument:

This transformation of the coefficient field is based on the spatial change of variables $x = \sqrt{T}\hat{x}$, which we complement with the temporal change of variables $t = T\hat{t}$, so that we have for the elliptic operator $T(\partial_t - \nabla \cdot a\nabla) = (\partial_{\hat{t}} - \hat{\nabla} \cdot \hat{a}\hat{\nabla})$. In view of the transformation of the initial condition $\sqrt{T}\nabla \cdot (ae) = \hat{\nabla}(\hat{a}e)$, the solution of the parabolic initial value problem (10) transforms according to $\sqrt{T}u = \hat{u}$, which leads to $T\nabla u = \hat{\nabla}\hat{u}$ and thus $\int_0^T \nabla u dt = \int_0^1 \hat{\nabla}\hat{u} d\hat{t}$ as well as $q(T) = \hat{q}(1)$. This establishes (134).

Step 3. Continuity wrt to H-convergence.
For any $\delta > 0$, the expression

$$\big(\int_0^1 \nabla u d\tau, q(1) = a(\int_0^1 \nabla u d\tau + e)\big)_\delta,$$

seen as a function of the λ-uniform coefficient field a, is continuous wrt to H-convergence. We recall that the topology defining H-convergence is the coarsest topology on the space Ω of λ-uniformly elliptic coefficient fields $a = a(x)$ for which the following functionals F are continuous: For any ball $B \subset \mathbb{R}^d$, any two vector fields $h, \tilde{h} \in L^2(B)$ we consider $F = \int \tilde{h} \cdot (\nabla u, a\nabla u)$, where u is the Lax-Milgram solution of $-\nabla \cdot a\nabla u = \nabla \cdot h$ in B that vanishes outside of B.

Here comes the argument: We will argue step by step that certain classes of function(al)s are continuous wrt to H-convergence. We start by replacing the elliptic equation $-\nabla \cdot a\nabla u = \nabla \cdot h$ by the resolvent equation

$$-zu - \nabla \cdot a\nabla u = \nabla \cdot h \tag{135}$$

for any $z \in \mathbb{C}$ in the concave sector $\{(\mathcal{R}z)_+ < \lambda|z|\}$, where $\mathcal{R}z$ denotes the real part of z, and $(\mathcal{R}z)_+$ the positive part of $\mathcal{R}z$. We then replace the ball B by the whole space $\mathbb{R}^d$, localizing the topology with help of our exponential weight function η, which requires restricting to a version of the sector shifted to the left. We then use the standard complex-variable argument to pass from the resolvent in this shifted sector to the semi-group, which in turn yields the result.

Passing to the resolvent equation primarily requires the following estimate for (135)

$$\big(\int |\nabla u|^2\big)^{\frac{1}{2}} \le \frac{|z| + (\mathcal{R}z)_+}{\lambda|z| - (\mathcal{R}z)_+}\big(\int |h|^2\big)^{\frac{1}{2}}, \tag{136}$$

an elementary estimate at the origin of the notion of sectorial operators. Here comes the argument for (136): Testing (135) with the complex conjugate $\bar{u}$ yields the identity

$$-z\int |u|^2 + \int \nabla\bar{u} \cdot a\nabla u = \int \nabla\bar{u} \cdot h. \tag{137}$$

Of this equation we take the real part; noting that because a is real the first inequality in (1) implies $\nabla\bar{u} \cdot a\nabla u \ge \lambda|\nabla u|^2$, we obtain with help of Cauchy-Schwarz' inequality on the rhs of (137) the estimate

$$-(\mathcal{R}z)\|u\|^2 + \lambda\|\nabla u\|^2 \le \|\nabla u\|\|h\|, \tag{138}$$

where we temporarily introduced the notation $\|\cdot\|$ for the L^2-norm of functions and vector fields (note that it does not matter whether we take $L^2(B)$ or $L^2(\mathbb{R}^d)$ since all functions and fields can be assumed to vanish outside B). We also may use (137) to estimate the first lhs term as follows, using the upper bound $|a\xi|\le|\xi|$ provided by (1):

$$|z|\|u\|^2\le\|\nabla u\|^2+\|\nabla u\|\|h\|. \tag{139}$$

Rewriting (138) as $\lambda\|\nabla u\|^2\le\|\nabla u\|\|h\|+(\mathcal{R}z)_+\|u\|^2$, inserting (139), and dividing by $\|\nabla u\|$ yields $\lambda\|\nabla u\|\le\|h\|+\frac{(\mathcal{R}z)_+}{|z|}(\|\nabla u\|+\|h\|)$, which entails (136).

Let us now argue in favor of H-continuity of $a\mapsto\int\tilde h\cdot(\nabla u,a\nabla u)$, where $h,\tilde h\in L^2(\mathbb{R}^d)$ and u is the Lax-Milgram solution of (135) in some ball B (always with the understanding of vanishing boundary conditions). To probe continuity we give ourselves a sequence $\{a_n\}_{n\uparrow\infty}$ of λ-uniformly elliptic coefficient fields that H-converge to a on $\mathbb{R}^d$ and thus a fortiori on B and have to show that $(\nabla u_n,a_n\nabla u_n)$ weakly converges to $(\nabla u,a\nabla u)$ in L^2. According to (136), the corresponding $\{\nabla u_n\}_n$ are bounded in L^2 and thus, after passing to a subsequence which we don't indicate in our notation, weakly convergent to some ∇u. Since the domain B is bounded, we obtain by Rellich's compactness theorem that u_n strongly converges to u in L^2. Consider the Lax-Milgram solution $\tilde u_n$ defined through $-\nabla\cdot a_n\nabla\tilde u_n=\nabla\cdot h+zu$ in B. On the one hand, rewriting the equation for u_n as $-\nabla\cdot a_n\nabla u_n=\nabla\cdot h+zu_n$, we learn from the energy estimate in conjunction with Poincaré's estimate that $\nabla\tilde u_n-\nabla u_n$ and thus also $a_n\nabla\tilde u_n-a_n\nabla u_n$ converges strongly (and a fortiori weakly) to zero in L^2; in particular the weak limit of $\nabla\tilde u_n$ has to agree with the weak limit ∇u of ∇u_n. On the other hand, by definition of H-convergence, $(\nabla\tilde u_n,a_n\nabla\tilde u_n)$ weakly converges to $(\nabla u,a\nabla u)$, so that $(\nabla\tilde u_n,a_n\nabla\tilde u_n)$ has to weakly converge to $(\nabla u,a\nabla u)$ too, where u must be the Lax-Milgram solution of $-\nabla\cdot a\nabla u=\nabla\cdot h+zu$ in B.

For $r\ge\frac{1}{\lambda}$, we now argue in favor of H-continuity of $a\mapsto\int\eta_r\tilde h\cdot(\nabla u,a\nabla u)$ where the vector field h and the function f satisfy $\int\eta_r|(h,f)|^2<\infty$ and the vector field $\tilde h$ is in the dual space $\int\frac{1}{\eta_r}|\tilde h|^2<\infty$, and where u is related to (h,f) by the resolvent equation

$$-zu-\nabla\cdot a\nabla u=\nabla\cdot h+f \tag{140}$$

on the whole space $\mathbb{R}^d$ and $z\in\mathbb{C}$ lies in the shifted concave sector

$$\mathcal{R}(z+1)<\frac{\lambda}{4}|z+1|, \tag{141}$$

which is slightly less concave then for (136). As above, this relies essentially on the following a priori estimate for (140)

$$\begin{aligned}&\Big(\int\eta_r(|u|^2+|\nabla u|^2)\Big)^{\frac12}\\ &\le\ \frac{2|z+1|+\mathcal{R}(z+1)}{\frac{\lambda}{2}|z+1|-2\mathcal{R}(z+1)}\Big(\Big(\int\eta_r|f|^2\Big)^{\frac12}+\Big(\int\eta_r|h|^2\Big)^{\frac12}\Big),\end{aligned} \tag{142}$$

which holds for both a ball B and the whole space $\mathbb{R}^d$ and in particular provides well-posedness in the latter case. We first turn to the argument for (142); to this purpose, we shift the equation rather than the sector z is in, and thus consider

$$-zu+(\mathrm{id}-\nabla\cdot a\nabla)u=\nabla\cdot h+f.$$

Testing this equation with $\eta_r\bar u$ and using Leibniz' rule yields the identity

$$\begin{aligned}-z\int\eta_r|u|^2+\int\big(\eta_r(|u|^2+\nabla\bar u\cdot a\nabla u)+\bar u\nabla\eta_r\cdot a\nabla u\big)&\\ =\int\big(\eta_r(\bar uf-\nabla\bar u\cdot h)-\bar u\nabla\eta_r\cdot h\big).&\end{aligned}$$

Using next to the bounds (1) on a that for our exponential localization function $|\nabla\eta_r|\le\frac{1}{r}\eta_r$, we have the following inequalities for $r\ge\frac{1}{\lambda}$:

$$\begin{aligned}\int\big(\eta_r(|u|^2+\nabla\bar u\cdot a\nabla u)+\bar u\nabla\eta_r\cdot a\nabla u\big)&\ \ge\ \frac{\lambda}{2}\int\eta_r(|u|^2+|\nabla u|^2),\\ \Big|\int\big(\eta_r(|u|^2+\nabla\bar u\cdot a\nabla u)+\bar u\nabla\eta_r\cdot a\nabla u\big)\Big|&\ \le\ 2\int\eta_r(|u|^2+|\nabla u|^2),\\ \Big|\int\big(\eta_r(\bar uf-\nabla\bar u\cdot h)-\bar u\nabla\eta_r\cdot h\big)\Big|&\ \le\ \Big(\int\eta_r(|u|^2+|\nabla u|^2)\Big)^{\frac12}\\ &\ \times\ \Big(\Big(\int\eta_r|f|^2\Big)^{\frac12}+\Big(\int\eta_r|h|^2\Big)^{\frac12}\Big).\end{aligned}$$

We now may argue as for (136) with $\big(\int|\nabla u|^2\big)^{\frac12}$ replaced by $\big(\int\eta_r(|u|^2+|\nabla u|^2)\big)^{\frac12}$, the lower bound λ in (138) replaced by $\frac{\lambda}{2}$, the upper bound 1 in (139) replaced by 2, and the size of the rhs $\big(\int|\nabla u|^2\big)^{\frac12}$ replaced by $\big(\int\eta_r|f|^2\big)^{\frac12}+\big(\int\eta_r|h|^2\big)^{\frac12}$. Hence this yields

$$\Big(\int\eta_r(|u|^2+|\nabla u|^2)\Big)^{\frac12}\le\frac{2|z|+\mathcal{R}z}{\frac{\lambda}{2}|z|-2\mathcal{R}z}\Big(\Big(\int\eta_r|f|^2\Big)^{\frac12}+\Big(\int\eta_r|h|^2\Big)^{\frac12}\Big),$$

which turns into (142) after shifting back.

Equipped with the estimate (142), we now may turn to the H-continuity of the new class of functionals F. We establish continuity by showing that such an F is the limit of the corresponding functionals $\{F_R\}_{R\uparrow\infty}$, for which u is replaced by the Lax-Milgram solution u_R for (140) with the whole space replaced by the centered ball B_R of radius R, and which we thus know to be continuous by the previous argument. Since we establish this limit in the uniform topology, the continuity of F_R transfers to F. More precisely, we shall show that

$$\Big(\int\eta_r|\nabla u_R-\nabla u|^2\Big)^{\frac12}\lesssim\Big(\frac{1}{R}+\exp\big(-\frac{R}{2r}\big)\Big)\Big(\int\eta_r(|f|^2+|h|^2)\Big)^{\frac12}. \tag{143}$$

To this purpose we consider $w=u_R-u$ which satisfies the homogeneous equation $-zw-\nabla\cdot a\nabla w=0$, but only in B_R (without vanishing bc). We thus consider $\tilde w=\tilde\eta w$, where $\tilde\eta$ is a smooth cut-off for $B_{\frac{R}{2}}$ in B_R, and note that it satisfies

$-z\tilde{w}-\nabla\cdot a\nabla\tilde{w}=\nabla\cdot\tilde{h}+\tilde{f}$, now on all of $\mathbb{R}^d$, but with the rhs $\tilde{h}:=-wa\nabla\tilde{\eta}$ and $\tilde{f}:=-\nabla\tilde{\eta}\cdot a\nabla w$. Hence an application of (142) yields

$$\int\eta_r|\nabla\tilde{w}|^2\lesssim\int\eta_r(|\tilde{f}|^2+|\tilde{h}|^2),$$

which in view of the choice of $\tilde{\eta}$ turns into

$$\int_{B_{\frac{R}{2}}}\eta_r|\nabla w|^2\lesssim\frac{1}{R^2}\int\eta_r(|w|^2+|\nabla w|^2).$$

Using $\eta_r(x)\lesssim\exp(-\frac{|x|}{2r})\eta_{2r}(x)$ and $\sup_{x\in B^c_{\frac{R}{2}}}\exp(-\frac{|x|}{2r})\le\exp(-\frac{R}{4r})$ we have

$$\int_{B^c_{\frac{R}{2}}}\eta_r|\nabla w|^2\lesssim\exp(-\frac{R}{4r})\int\eta_r|\nabla w|^2.$$

The combination of the two last estimates yields by definition of $w=u_R-u$

$$\begin{aligned}&\Big(\int\eta_r|\nabla u_R-\nabla u|^2\Big)^{\frac12}\\ \lesssim\ &\big(\frac{1}{R}+\exp(-\frac{R}{2r})\big)\Big(\int\eta_r(|u_R|^2+|u|^2+|\nabla u_R|^2+|\nabla u|^2)\Big)^{\frac12}.\end{aligned}$$

Applying once more (142) to u and u_R separately, we obtain (143).

We now pass from the resolvent to the semi-group. We fix a function f with $\int\eta_r|f|^2<\infty$, and for z within the sector (141) consider the solution u_z of $-zu_z-\nabla\cdot a\nabla u_z=f$ in $\mathbb{R}^d$. Let Γ be the set $\mathcal{R}(z+2)=\frac{\lambda}{4}|z+2|$ oriented in such a way that it positively circles around the positive axis, which is well within the sector (141). Hence in view of (142), for $t>0$ the integral $\int_\Gamma e^{-tz}u_z dz$ converges absolutely with values in the weighted L^2-space $(\int\eta_r(|\cdot|^2+|\nabla\cdot|^2))^{\frac12}$. It is then an easy consequence of $\frac{1}{2\pi i}\int_\Gamma e^{-tz}dz=0$ that

$$u(t)=\frac{1}{2\pi i}\int_\Gamma e^{-tz}u_z dz$$

solves the equation $\partial_t u-\nabla\cdot a\nabla u=0$, and of Cauchy's integral theorem in form of $\lim_{t\downarrow0}\frac{1}{2\pi i}\int_\Gamma\frac{e^{-tz}}{\mu-z}dz=1$ for $\mu>-2$ that $u(t=0)=f$. In particular, we have

$$\int\tilde{h}\cdot(\nabla u(t),a\nabla u(t))=\frac{1}{2\pi i}\int_\Gamma e^{-tz}\int\tilde{h}\cdot(\nabla u_z,a\nabla u_z)\,dz,$$

converging absolutely for any $\tilde{h}$ with $\int\frac{1}{\eta_r}|\tilde{h}|^2<\infty$. Hence by Lebesgue's theory, the H-continuity of $a\mapsto\int\tilde{h}\cdot(\nabla u_z,a\nabla u_z)$ transfers to $a\mapsto\int\tilde{h}\cdot(\nabla u(t),a\nabla u(t))$.

Finally, the H-continuity of $(\nabla\int_0^1 ud\tau,q(1))_\delta$ is a consequence of the previous general statement. Indeed, we consider $U(a;t,x)=\int_0^t ud\tau+x\cdot e$ and note that it solves the homogeneous equation $\partial_t U-\nabla\cdot a\nabla U=0$ with initial data $U(t=0)=f$ where $f(x)=x\cdot e$. Since clearly $\int\eta_r f^2<\infty$ and $\int\frac{1}{\eta_r}|G_\delta|^2<\infty$, where G_δ

denotes the Gaussian of variance $\delta > 0$, we have that $(\int_0^1 \nabla u d\tau + e, q(1))_\delta = \big(\int G_\delta \nabla U(1), \int G_\delta a \nabla U(1)\big)$ is an H-continuous function.

Step 4. We now argue that a stationary ensemble $\langle\cdot\rangle$ that is of range r for all $r > 0$ is deterministic in the sense that

$$a = \langle a\rangle \quad \text{ae in } \mathbb{R}^d \quad \text{almost surely.} \tag{144}$$

Here comes the argument: For a scale ρ, we consider the convolution with a Dirac sequence $\tilde\eta_\rho(x) = \frac{1}{\rho^d}\tilde\eta_1(\frac{x}{\rho})$, where $\tilde\eta_1$ is supported in B_1, $\tilde\eta_1 \geq 0$, and $\int \tilde\eta_1 = 1$:

$$a_\rho(y) = \int \tilde\eta_\rho(x-y)a(x)dx = \int \tilde\eta_\rho(x)a(x+y)dx. \tag{145}$$

We first note that for any $y \in \mathbb{R}^d$ and $\rho > 0$, $a \mapsto a_\rho(y)$ is continuous wrt H-convergence: Indeed, $e_i \cdot a_\rho(y)e_j$ is of the form $\int \tilde h \cdot a\nabla u$, provided we make the following choices.

- For the ball: $B = B_{2\rho}(y)$,
- for the square integrable vector field $\tilde h$ testing the weak convergence: $\tilde h(x) = \tilde\eta_\rho(x-y)e_i$,
- and for the square integrable vector field h defining the equation for u: $h = -a\nabla(\zeta x_j)$ where ζ is a cut-off for $B_\rho(y)$ in $B_{2\rho}(y)$ so that $u = \zeta x_j$ and thus $\nabla u = e_j$ in $B_\rho(y)$, the support of $x \mapsto \tilde\eta_\rho(x-y)$.

By definition of stationarity of $\langle\cdot\rangle$, we learn from the second representation in (145)

$$\langle a_\rho\rangle := \langle a_\rho(y)\rangle \quad \text{does not depend on } y \in \mathbb{R}^d.$$

By definition of range r, we learn from the fact that in view of (145), $a_\rho(y)$ depends on a only through $a_{|B_\rho(y)}$:

$$\langle a_\rho(y) \otimes a_\rho(y')\rangle = \langle a_\rho(y)\rangle \otimes \langle a_\rho(y')\rangle \quad \text{for } |y-y'| \geq r + 2\rho.$$

Both last statements combine to

$$\langle (a_\rho(y) - \langle a_\rho\rangle) : (a_\rho(y') - \langle a_\rho\rangle)\rangle = 0 \quad \text{for } |y-y'| \geq r+2\rho. \tag{146}$$

Denoting by the subscript R another convolution, we compute

$$\langle |(a_\rho)_R - \langle a_\rho\rangle|^2\rangle = \int\int \tilde\eta_R(y)\tilde\eta_R(y')\langle (a_\rho(y) - \langle a_\rho\rangle) : (a_\rho(y') - \langle a_\rho\rangle)\rangle.$$

By the upper bound in (1), which implies $|a(x)| \leq d$ for the Frobenius norm, this yields in conjunction with (146)

$$\langle |(a_\rho)_R - \langle a_\rho\rangle|^2\rangle \leq d \sup|\tilde\eta_1||B_1|(\frac{r+2\rho}{R})^d,$$

which we integrate over an arbitrary ball B to obtain

$$\langle \int_B |(a_\rho)_R - \langle a_\rho\rangle|^2\rangle \lesssim |B|(\frac{r+2\rho}{R})^d. \tag{147}$$

We first let ρ tend to zero; since a_ρ converges to the Lebesgue-measurable and bounded a strongly in L^2_{loc}, we learn that also $\langle a_\rho\rangle$ converges to a limit we call $\langle a\rangle$, so that (147) turns into $\langle \int_B |a_R - \langle a\rangle|^2\rangle \lesssim |B|(\frac{r}{R})^d$. Since $r > 0$ was arbitrary, this yields $\langle \int_B |a_R - \langle a\rangle|^2\rangle = 0$. We finally let R go to zero to obtain $\langle \int_B |a - \langle a\rangle|^2\rangle = 0$. Since this holds for any ball B, we first learn that $\langle a\rangle$ does not depend on B and then deduce (144).

Step 5. Indirect argument for (131) by compactness.
We assume that (131) fails, which means that there exists a sequence $\langle\cdot\rangle_n$ of admissible ensembles and a sequence of times $T_n \uparrow \infty$ such that

$$\liminf_{n\uparrow\infty} \big\langle\big|(\int_0^{T_n} \nabla u d\tau, q(T_n) - \langle q(T_n)\rangle_n)_{\delta\sqrt{T_n}}\big|^2\big\rangle_n > 0. \tag{148}$$

We write for brevity $\langle\cdot\rangle_{\hat n}$ for the rescaled ensemble $\langle\cdot\rangle_{n,\sqrt{T_n}}$ introduced in Step 1. We note that the space Ω of λ-uniformly elliptic coefficient fields $a = a(x)$ on $\mathbb{R}^d$ is compact under H-convergence: For $\mathbb{R}^d$ replaced by a fixed bounded open set this is explicit in the work of Murat & Tartar, see [25, Theorem 6.5]. In our case like in theirs, compactness relies on the fact that the topology is characterized by the continuity of the maps F made explicit in Step 3, which are indexed by vector fields $h, \tilde h \in L^2(\mathbb{R}^d)$ and, in our case, by balls $B \subset \mathbb{R}^d$; and that these maps can be approximated in the uniform topology by a countable set of maps, which are generated by a countable dense set in $L^2(\mathbb{R}^d)$, and in our case, by the balls with rational center and radius. The H-limits coming from different balls are compatible thanks to the locality of that notion, see [25, Lemma 10.5]. Since Ω is compact, the space of probability measures on the topological space Ω (by which we of course mean those that respect the topology) is also compact. Hence after extraction of a subsequence, which we don't indicate in our notation, we may assume that our sequence of ensembles $\langle\cdot\rangle_{\hat n}$ weak-* converges to a probability measure $\langle\cdot\rangle$, which means that for all continuous functions F on Ω, we have $\lim_{n\uparrow\infty}\langle F\rangle_{\hat n} = \langle F\rangle$. By definition of the latter, this weak-* convergence preserves the stationarity and the property of having a given range. Since $\langle\cdot\rangle_{\hat n}$ has range $\frac{1}{\sqrt{T_n}} \downarrow 0$, we have that the limiting measure is both stationary and of range r for any $r > 0$. Hence by Step 4, it is deterministic in the sense of (144). The latter implies $u = 0$ and thus

$$(\nabla u(1), q(1)) = (0, \langle a\rangle e) \quad \text{ae in } \mathbb{R}^d \quad \text{almost surely,}$$

and thus in particular

$$\langle|(\int_0^1 \nabla u d\tau, q(1) - \langle q(1)\rangle)_\delta|^2\rangle = 0.$$

Since by Step 3, $(\int_0^1 \nabla u d\tau, q(1))_\delta$ is a continuous function on Ω, this implies

$$\lim_{n\uparrow\infty}\langle|(\int_0^1 \nabla u d\tau, q(1) - \langle q(1)\rangle_{\hat n})_\delta|^2\rangle_{\hat n} = 0,$$

which by Step 2 is in contradiction with (148).

6. Proofs of the stochastic auxiliary results

6.1. **Proof of Lemma 9: decay of averages of approximately local fields.**

We split the proof into four steps.

Step 1. CLT scaling in case of exact locality.
We start with the auxiliary statement

$$\begin{aligned} &\log\langle\exp(\nu g)\rangle \leq \Lambda\nu^2 + |\nu| \quad \text{for all } \nu \\ &\Longrightarrow \quad \log\langle\exp(\nu g_R)\rangle \lesssim \Lambda\frac{\nu^2}{R^d} + |\nu| \quad \text{for all } \nu \end{aligned} \tag{149}$$

for a (possibly non-centered) random function g that is *exactly* one-local in the sense of

$$g(a) = g(\tilde{a}) \quad \text{provided} \quad a = \tilde{a} \text{ on } B_1.$$

In order to establish (149), we write $g_R = g_R(0) = \int_{\mathbb{R}^d}(G_R g)(x)dx$, where G_R denotes the Gaussian with variance R, as convex combination of a *sum* of independent random variables:

$$g_R = \fint_{[0,2)^d} \sum_{z\in\mathbb{Z}^d} 2^d (G_R g)(2z+x)dx.$$

Note that for fixed x, the points in the lattice $\{2z+x\}_{z\in\mathbb{Z}^d}$ have distance at least 2 and thus the numbers $\{g(2z+x)\}_{z\in\mathbb{Z}^d}$ and therefore $\{2^d(G_R g)(2z+x)\}_{z\in\mathbb{Z}^d}$ are independent since g is one-local and a has range unity. Together with the convexity of $F\mapsto \log\langle\exp(F)\rangle$ this implies

$$\begin{aligned} \log\langle\exp(\nu g_R)\rangle &\leq \fint_{[0,2)^d} \log\Big\langle \exp\big(\nu\sum_{z\in\mathbb{Z}^d} 2^d(G_R g)(2z+x)\big)\Big\rangle dx \\ &= \fint_{[0,2)^d} \sum_{z\in\mathbb{Z}^d} \log\langle\exp(\nu 2^d (G_R g)(2z+x))\rangle dx. \end{aligned}$$

By the assumption in (149) with ν replaced by $2^d G_R(2z+x)\nu$ in conjunction with stationarity of g, this yields the conclusion in (149) in form of

$$\begin{aligned} \log\langle\exp(\nu g_R)\rangle &\leq \fint_{[0,2)^d} \sum_{z\in\mathbb{Z}^d} \Lambda(4^d G_R^2(2z+x)\nu^2 + 2^d G_R(2z+x)|\nu|)dx \\ &= \Lambda 2^d \int_{\mathbb{R}^d} G_R^2 dx \nu^2 + \int_{\mathbb{R}^d} G_R dx|\nu| \sim \Lambda\frac{\nu^2}{R^d} + |\nu|. \end{aligned}$$

Step 2. A property of centered random variables.
We continue with the auxiliary statement for a centered random variable g:

$$\log\langle\exp(\nu g)\rangle \leq \nu^2 \text{ for all } |\nu|\geq 1 \quad \Longrightarrow \quad \log\langle\exp(\nu g)\rangle \lesssim \nu^2 \text{ for all } \nu. \tag{150}$$

To this purpose we note that

$$\begin{aligned}\frac{d}{d\nu}\log\langle\exp(\nu g)\rangle &= \frac{\langle g\exp(\nu g)\rangle}{\langle\exp(\nu g)\rangle},\\ \frac{d^2}{d\nu^2}\log\langle\exp(\nu g)\rangle &= \frac{\langle g^2\exp(\nu g)\rangle}{\langle\exp(\nu g)\rangle}-\big(\frac{\langle g\exp(\nu g)\rangle}{\langle\exp(\nu g)\rangle}\big)^2.\end{aligned}$$

Since g is centered, we infer from Jensen's inequality that $\langle\exp(\nu g)\rangle\ge\exp(\nu\langle g\rangle)=1$ and from the first formula that

$$\log\langle\exp(\nu g)\rangle_{|\nu=0}=0,\quad \frac{d}{d\nu}_{|\nu=0}\log\langle\exp(\nu g)\rangle=0.$$

From the second formula, the lower bound $\langle\exp(\nu g)\rangle\ge 1$, and Jensen's inequality, we deduce that

$$0\le\frac{d^2}{d\nu^2}\log\langle\exp(\nu g)\rangle\le\langle g^2(\exp(g)+\exp(-g))\rangle\quad\text{for } |\nu|\le 1.$$

By the real variable inequality $g^2(\exp(g)+\exp(-g))\le(\exp(g)+\exp(-g))^2=\exp(2g)+\exp(-2g)+2$, this implies $\langle g^2(\exp(g)+\exp(-g))\rangle\le\langle\exp(2g)\rangle+\langle\exp(-2g)\rangle+2\le 2(\exp(2^2)+1)$ by the assumption in (150). Hence the second derivative of $\log\langle\exp(\nu g)\rangle$ in $\nu\in[-1,1]$ is bounded, while the function vanishes to first order in the origin. This establishes the conclusion in (150).

Step 3. CLT scaling for approximately local random variables.
Suppose the centered random variable g is approximately one-local in the sense that for $r\ge 1$

$$|g(a)-g(\tilde a)|\le\exp(-r)\quad\text{provided } a=\tilde a \text{ in } B_r \tag{151}$$

and has Gaussian moments in the sense of

$$\log\langle\exp(\nu g)\rangle\le\Lambda\nu^2 \tag{152}$$

for some constant $\Lambda\le 1$. Provided $R\gg\log(e+\frac{1}{\Lambda})$, we claim that its spatial average g_R has Gaussian bounds:

$$\log\langle\exp(\nu g_R)\rangle\lesssim\frac{\Lambda}{R^d}(\log\frac{R^d}{\Lambda})^d\nu^2. \tag{153}$$

Here comes the argument: We first note that we may assume $\Lambda\ll 1$, since otherwise we scale down g by a large constant, which only improves (151), and the effect of which can be absorbed in the assumption $R\gg\log(e+\frac{1}{\Lambda})\sim\log\frac{1}{\Lambda}$ and which is controlled in (153) (by rescaling ν). For arbitrary $1\le r\le R$ to be eventually fixed, we introduce the auxiliary random variable $\tilde g(a)=g(\tilde a)$ where $\tilde a$ denotes the restriction of a on B_r, extended by id. According to assumption (151), we have

$$|g(a)-\tilde g(a)|\le\exp(-r)\quad\text{for all } a, \tag{154}$$

and thus also for $(\tilde g_R)(a)=\int\tilde g(a(z+\cdot))G_R(z)dz$, where G_R is the Gaussian of scale R,

$$|g_R(a)-\tilde g_R(a)|\le\int|g(a(z+\cdot))-\tilde g(a(z+\cdot))|G_R(z)dz\overset{(154)}{\le}\exp(-r). \tag{155}$$

Together with (152), (154) yields

$$\log\langle\exp(\nu\tilde g)\rangle \le \Lambda\nu^2 + \exp(-r)|\nu|.$$

By construction, $\tilde g$ is exactly local on scale r, so that (149) (in the rescaled version of $x = r\hat x$, $R = r\hat R$) implies

$$\log\langle\exp(\nu\tilde g_R)\rangle \lesssim \Lambda(\frac{r}{R})^d\nu^2 + \exp(-r)|\nu|.$$

Appealing to (155), this yields

$$\log\langle\exp(\nu g_R)\rangle \ \ \lesssim \ \ \Lambda(\frac{r}{R})^d\nu^2 + \exp(-r)|\nu|.$$

In order to apply Step 2 we rewrite this

$$\log\langle\exp(\nu g_R)\rangle \ \ \le \ \ \big(\Lambda^{\frac12}(\frac{r}{R})^{\frac d2}\nu\big)^2 + \frac{2}{\Lambda^{\frac12}}(\frac{R}{r})^{\frac d2}\exp(-r)|\Lambda^{\frac12}(\frac{r}{R})^{\frac d2}\nu|.$$

Since $\langle g_R\rangle = \langle g\rangle = 0$, we may apply (150) (in its rescaled version of $g = \frac{1}{\Lambda^{\frac12}}(\frac{R}{r})^{\frac d2}\hat g$) as long as

$$\frac{1}{\Lambda^{\frac12}}(\frac{R}{r})^{\frac d2}\exp(-r) \le 1$$

which is satisfied by $r = \log(\frac{R^{\frac d2}}{\Lambda^{\frac12}}) \ge 1$ (which in turn is admissible in the sense of $r \le R$, since the latter means $\log\frac{R^d}{\Lambda} \le 2R$ and follows from our assumption $R \gg \log\frac{1}{\Lambda}$) to the effect of

$$\log\langle\exp(\nu g_R)\rangle \lesssim \Lambda(\frac{r}{R})^d\nu^2 \sim \Lambda(\frac{\log\frac{R^d}{\Lambda}}{R})^d\nu^2.$$

Step 4. Conclusion.
We appeal to Step 3 with $g = g_{\sqrt t}$ in its rescaled version with $x = \sqrt t\hat x$, $R = \sqrt t\hat R$ and obtain from (153)

$$\log\langle\exp(\nu(g_{\sqrt t})_R)\rangle \ \ \lesssim \ \ \frac{\Lambda\sqrt t^d}{R^d}(\log\frac{R^d}{\Lambda\sqrt t^d})^d\nu^2.$$

The application of (153) is legitimate provided $\hat R = \frac{R}{\sqrt t} \gg \log(e+\frac{1}{\Lambda})$. Because of the latter regime and the semi-group property of the convolution with Gaussians $(g_{\sqrt t})_R = g_{\sqrt{R^2+t}}$, the above turns into (47) for $R \gg \sqrt t\log(e+\frac1\Lambda)$.

For (47) in the remaining borderline range $\sqrt t \le R \lesssim \sqrt t\log(e+\frac1\Lambda)$, we note that by the semi-group property, the convexity of $F\mapsto\log\langle\exp(F)\rangle$ and Jensen's inequality, $\log\langle\exp(\nu g_R)\rangle$ is monotone decreasing in R. Hence (46) implies in particular $\log\langle\exp(\nu g_R)\rangle \le \Lambda\nu^2$ for $R \ge \sqrt t$. We conclude by noting that $\Lambda\nu^2 \lesssim \frac{\Lambda\sqrt t^d}{R^d}\log^d(e+\frac{R^d}{\Lambda\sqrt t^d})\nu^2$. Indeed, this is equivalent to $\frac{\Lambda^{\frac1d}\sqrt t}{R}\log(e+\frac{R}{\Lambda^{\frac1d}\sqrt t}) \gtrsim \Lambda^{\frac1d}$ and thus $\frac{\Lambda^{\frac1d}\sqrt t}{R} \gtrsim \Lambda^{\frac1d}\frac{1}{\log(e+\frac1\Lambda)}$, which in turn is a consequence of our borderline regime $R \lesssim \sqrt t\log(e+\frac1\Lambda)$.

7. Proofs of inner-regularity results and uniform bounds

7.1. **Proof of Lemma 11: control of strong norms by averages.** By scaling we may wlog assume $T=1$ and write $\eta=\eta_1$ for brevity.

Step 1. Localized parabolic higher-order energy estimates: For all $\tau\in(0,1)$ and $\beta\geq 0$,

$$\int\eta|\nabla u|^2(\tau) \;\lesssim\; \frac{1}{\tau}\int\eta u^2(\frac{\tau}{2}), \tag{156}$$

$$\int\eta|\nabla u|^2(1) \;\lesssim\; \int_0^{\frac{1}{2}}(2t)^\beta\int\eta u^2(t)dt, \tag{157}$$

where we recall that $\eta(x)=\exp(-|x|)$ denotes our localization function. The argument would be straightforward for symmetric a in which case one computes $\frac{d}{dt}\frac{1}{2}\int\eta\nabla u\cdot a\nabla u$ to find that this energy is non-increasing up to lower-order terms, ie terms that involve $\nabla\eta$. In the general case, one has to pass via $\frac{d}{dt}\frac{1}{2}\int\eta u^2$ and $\frac{d}{dt}\frac{1}{2}\int\eta(\nabla\cdot a\nabla u)^2$, and their interpolation, instead. Indeed, we have from the equation $\partial_t u-\nabla\cdot a\nabla u=0$, integration by parts, and Leibniz' rule

$$\frac{d}{dt}\frac{1}{2}\int\eta u^2=-\int(\eta\nabla u\cdot a\nabla u+u\nabla\eta\cdot\nabla u),$$

and thus by the bounds (1) on a, the property $|\nabla\eta|\leq\eta$ for our exponential localization function, and Young's inequality

$$\frac{d}{dt}\int\eta u^2\leq-\int\eta(\lambda|\nabla u|^2-\frac{1}{\lambda}u^2). \tag{158}$$

Since a does not depend on time, ∂_t commutes with the parabolic operator $\partial_t-\nabla\cdot a\nabla$ so that $v:=\nabla\cdot a\nabla u=\partial_t u$ is also a solution of the homogeneous equation to which we may apply (158). For a parameter $\delta>0$ still to be chosen, this yields in combination

$$\begin{aligned}&\frac{d}{dt}\big(\delta t^2\int\eta v^2+\int\eta u^2\big) \\ &\leq-\lambda\big(\delta t^2\int\eta|\nabla v|^2+\int\eta|\nabla u|^2\big)+2\delta t\int\eta v^2+\frac{1}{\lambda}\big(\delta t^2\int\eta v^2+\int\eta u^2\big).\end{aligned} \tag{159}$$

In order to absorb $\int\eta v^2$ into $\int\eta|\nabla v|^2$ and $\int\eta|\nabla u|^2$, we appeal to the following interpolation estimate

$$\int\eta v^2\leq 2\big(\int\eta|\nabla v|^2\int\eta|\nabla u|^2\big)^{\frac{1}{2}}+\int\eta|\nabla u|^2, \tag{160}$$

which follows from integration by parts and Leibniz' rule in form of $\int\eta v^2=-\int(\eta\nabla v\cdot a\nabla u+v\nabla\eta\cdot a\nabla u)$, the upper bound (1) on a and the choice of η, which yield $\int\eta v^2\leq\int\eta(|\nabla v||\nabla u|+|v||\nabla u|)$, and the inequalities of Cauchy-Schwarz and Young.

By choosing $\delta \ll 1$, we learn from applying once more Young's inequality to the rhs of (160) that (159) turns into the differential inequality

$$\frac{d}{dt}\big(\delta t^2 \int \eta v^2 + \int \eta u^2\big) \le \frac{1}{\lambda}\big(\delta t^2 \int \eta v^2 + \int \eta u^2\big). \tag{161}$$

We now integrate (161) rewritten in form of $\frac{d}{dt}\big(\exp(-\frac{t}{\lambda})\big(\delta t^2 \int \eta v^2 + \int \eta u^2\big)\big) \le 0$ over $t \in (0,\tau)$ to obtain

$$\big(\tau^2 \int \eta v^2(\tau) + \int \eta u^2(\tau)\big) \lesssim \int \eta u^2(0). \tag{162}$$

We use once more an interpolation inequality, this time

$$\lambda\tau \int \eta|\nabla u|^2(\tau) \le 2\big(\tau^2 \int \eta v^2(\tau) \int \eta u^2(\tau)\big)^{\frac{1}{2}} + \frac{\tau}{\lambda}\int \eta u^2(\tau), \tag{163}$$

which follows from the lower bound on a in (1) in form of $\lambda \int \eta|\nabla u|^2 \le \int \eta \nabla u \cdot a\nabla u$, integration by parts and Leibniz' rule in form of $\int \eta\nabla u \cdot a\nabla u = -\int \eta u v - \int u\nabla\eta \cdot a\nabla u$, the upper bound on a in (1), $|\nabla\eta| \le \eta$, and the inequalities of Cauchy-Schwarz and Young. Applying once more Young's inequality to (163) and inserting the result into (162) yields for all $\tau \in (0,1)$,

$$\tau \int \eta|\nabla u|^2(\tau) \lesssim \int \eta u^2(0).$$

By a rescaling in space and time we may use this with the time interval $(0,\tau)$ replaced by $(\frac{\tau}{2},\tau)$, which yields (156). In combination with (158), we obtain (157). Indeed, (158) in form of $\frac{d}{dt}\big(\exp(-\frac{t}{\lambda})\int \eta u^2\big) \le 0$ implies that, for all $\beta \ge 0$, $\int \eta u^2(\frac{1}{2}) \lesssim \int_0^{\frac{1}{2}} (2t)^\beta \int \eta u^2 dt$, from which the claim follows.

Step 2. Interpolation estimate. For an arbitrary function u we have

$$\begin{aligned}
&\Big(\int \eta u^2\Big)^{\frac{1}{2}} \\
&\lesssim \Big[\Big(\int \eta|\nabla u|^2\Big)^{\frac{1}{2}}\Big]^{1-\frac{1}{\frac{d}{2}+p+2}} \Big[\int_0^1 dr r^p \int \eta_2|u_r|\Big]^{\frac{1}{\frac{d}{2}+p+2}} \\
&\quad + \int_0^1 dr r^p \int \eta_2|u_r|.
\end{aligned} \tag{164}$$

Here comes the argument: For $r \ll 1$ we split and estimate like for (82)-(84) in the proof of Lemma 5:

$$\int \eta u^2 \lesssim \int \eta(u-u_{\sqrt{2}r})^2 + \int \eta(u_{\sqrt{2}r})^2 \overset{(86)}{\lesssim} r^2 \int \eta|\nabla u|^2 + \int \eta(u_{\sqrt{2}r})^2. \tag{165}$$

We turn to the second rhs term and seek to replace the L^2-norm by an L^1-norm (at the expense of decreasing the convolution scale and increasing the averaging scale) in the sense of

$$\Big(\int \eta(u_{\sqrt{2}r})^2\Big)^{\frac{1}{2}} \lesssim \frac{1}{r^{\frac{d}{2}}}\int \eta_2|u_r|. \tag{166}$$

Starting from the semi-group property and Jensen's inequality in form of $|u_{\sqrt{2}r}| = |(u_r)_r| \le |u_r|_r$ we obtain since the Gaussian convolution kernel is dominated by the exponential localization function of the same scale

$$|u_{\sqrt{2}r}| \lesssim \int \eta_r |u_r| \lesssim \frac{1}{r^d}\int \eta_2 |u_r|.$$

Evoking translation invariance to replace the origin by a general point y, this implies by the specific properties of the exponential localization function

$$\eta_2(y)|(u_{\sqrt{2}r})(y)| \lesssim \frac{1}{r^d}\int \eta_2(y)\eta_2(x-y)|u_r(x)|dx \lesssim \frac{1}{r^d}\int \eta_2(x)|u_r(x)|dx,$$

so that using $\eta \lesssim \eta_2^2$ we obtain

$$\Big(\int \eta (u_{\sqrt{2}r})^2\Big)^{\frac{1}{2}} \lesssim \big(\int \eta_2 |u_{\sqrt{2}r}|\big)^{\frac{1}{2}}\big(\frac{1}{r^d}\int \eta_2|u_r|\big)^{\frac{1}{2}},$$

which turns into (166) by appealing once more to $|u_{\sqrt{2}r}| \le |u_r|_r$ and (79).

Inserting (166) into (165) yields for all $r \ll 1$

$$\Big(\int \eta u^2\Big)^{\frac{1}{2}} \lesssim r\Big(\int \eta|\nabla u|^2\Big)^{\frac{1}{2}} + \frac{1}{r^{\frac{d}{2}}}\int \eta_2|u_r|. \tag{167}$$

By the monotonicity property (97) we have $\int \eta_2|u_r| \lesssim \int \eta_2|u_{r'}|$ for any $r' \le r \lesssim 1$ and thus

$$\int \eta_2|u_r| \lesssim \frac{1}{r^{p+1}}\int_0^r dr' r'^p \int \eta_2|u_{r'}|.$$

Inserting this into (167) yields the interpolation inequality in its additive form

$$\Big(\int \eta u^2\Big)^{\frac{1}{2}} \lesssim r\Big(\int \eta|\nabla u|^2\Big)^{\frac{1}{2}} + \frac{1}{r^{\frac{d}{2}+p+1}}\int_0^1 dr' r'^p \int \eta_2|u_{r'}|,$$

from which we obtain the desired form (164) by optimization in $r \ll 1$.

Step 3. Conclusion.

Let us momentarily introduce the abbreviations

$$E_1(t) := \int \eta|\nabla u|^2, \quad E_0(t) := \int \eta u^2, \quad E_{-1}(t) := (\int_0^1 r^p \int \eta|u_r|dr)^2,$$

so that the results of the two previous steps turn into

$$E_1(t) \lesssim \frac{1}{t}E_0(\frac{t}{2}) \tag{168}$$

and the "algebraic" relation

$$E_0 \le C(E_1^{1-\theta}E_{-1}^{\theta} + E_{-1}), \tag{169}$$

where we have set $\theta := \frac{1}{\frac{d}{2}+p+2} \in (0,1)$. From (168), we infer by integration for any exponent $\alpha > 0$

$$\int_0^1 t^{1+\alpha}E_1 dt \lesssim \int_0^1 t^\alpha E_0(\frac{t}{2})dt. \tag{170}$$

From (169) we infer by Hölder's inequality for $\alpha=\frac{1-\theta}{\theta}$

$$\int_0^1 t^\alpha E_0 dt \lesssim \Big(\int_0^1 t^{1+\alpha}E_1 dt\Big)^{1-\theta}\Big(\int_0^1 E_{-1}dt\Big)^{\theta}+\int_0^1 E_{-1}dt. \tag{171}$$

Inserting (170) into (171), we obtain in a first stage obtain thanks to $\theta>0$ by Young's inequality $\int_0^1 t^\alpha E_0 dt\lesssim \int_0^1 E_{-1}dt$ which then may be upgraded to the desired result by using (157) for $\beta=\alpha$.

7.2. **Proof of Lemma 10: uniform bounds and approximate locality.** We start with (50) & (51) in the first step, then prove (48) & (49) in a second step under the validity of an auxiliary estimate which we prove in the last step.

Step 1. Proof of the statements (50) & (51) on approximate locality and boundedness of $\big(|(q_T-\langle q_T\rangle)_r|-\langle|(q_T-\langle q_T\rangle)_r|\rangle\big)_{\sqrt{T}}$.
Since $\langle|(q_T-\langle q_T\rangle)_r|\rangle$ is deterministic (and by stationarity controlled by $|(q_T-\langle q_T\rangle)_r|_{\sqrt{T}}$), it is enough to consider $|(q_T-\langle q_T\rangle)_r|_{\sqrt{T}}=|q_{T,r}-\langle q_T\rangle|_{\sqrt{T}}$ both for the uniform bound and the approximate locality. By scaling, we may wlog assume $T=1$ and will thus drop the index $T=1$ when writing $q=q_{T=1}$, $\eta=\eta_{\sqrt{T}=1}$, so that we reduced the locality statement to

$$||q_r-\langle q\rangle|_1(a)-|q_r-\langle q\rangle|_1(\tilde a)|\lesssim \exp(-\frac{R}{C})\quad\text{provided } a=\tilde a \text{ on } B_R. \tag{172}$$

From the defining equation $\phi-\nabla\cdot a(\nabla\phi+e)=0$ we obtain by the localized elliptic energy estimate

$$\int \eta_{2r_0}|\nabla\phi|^2\lesssim 1 \tag{173}$$

with the r_0 fixed in (174) below. Also for later use, let us give the short argument for this localized elliptic energy estimate for the slightly more general equation $v-\nabla\cdot a\nabla v=f+\nabla\cdot g$. We test this equation with $\eta_r v$ with scale r to be specified below and obtain

$$\int(\eta_r v^2+\nabla(\eta_r v)\cdot a\nabla v)=\int(\eta_r vf-\nabla(\eta_r v)\cdot g).$$

Using Leibniz' rule and the bounds (1) on a, this yields

$$\int \eta_r(v^2+|\nabla v|^2)\lesssim\int |\nabla\eta_r||v|(|\nabla v|+|g|)+\eta_r(|v||f|+|\nabla v||g|).$$

Because of $|\nabla\eta_r|\le\frac{1}{r}\eta_r$ for our exponential averaging function, we may absorb the first rhs term into the lhs for $r\gg1$ to the effect of

$$\int \eta_r(v^2+|\nabla v|^2)\lesssim\int \eta_r\big(|v|(|f|+|g|)+|\nabla v||g|\big),$$

which with help of Young's inequality turns into

$$\int \eta_r(v^2+|\nabla v|^2)\lesssim\int \eta_r(f^2+|g|^2)\quad\text{for } r\ge r_0 \tag{174}$$

for some $r_0 \leq C$. We now obtain the estimate (173) by applying (174) to $v = \phi$, $f = 0$, and $g = ae$ (and the upper bound (1) on a).

From (173) we also obtain

$$\int \eta |q|^2 \lesssim \int \eta_{2r_0} |q|^2 \lesssim 1, \tag{175}$$

which settles the first uniform bound in (51). We now compare $\phi = \phi(a)$ with $\tilde\phi := \phi(\tilde a)$; from the difference of the equations

$$(\phi - \tilde\phi) - \nabla \cdot \tilde a \nabla(\phi - \tilde\phi) = \nabla \cdot (a - \tilde a)(\nabla\phi + e)$$

we get from the localized elliptic energy estimate (174) that $\int \eta_{r_0} |\nabla(\phi - \tilde\phi)|^2 \lesssim \int \eta_{r_0} |(a-\tilde a)(\nabla\phi + e)|^2$. Since $q - \tilde q = \tilde a \nabla(\phi - \tilde\phi) + (a - \tilde a)(\nabla\phi + e)$ this yields in particular

$$\int \eta |q - \tilde q|^2 \lesssim \int \eta_{r_0} |q - \tilde q|^2 \lesssim \int \eta_{r_0} |(a - \tilde a)(\nabla\phi + e)|^2 \lesssim \int_{B_R^c} \eta_{r_0} |q|^2, \tag{176}$$

where we used the assumption $a = \tilde a$ on B_R.

We now post-process (176) in two ways: On the one hand, we use that for our exponential localization function we have $\eta_{r_0} \sim \eta_{2r_0}^2$ so that by (175)

$$\int_{B_R^c} \eta_{r_0} |q|^2 \lesssim (\sup_{B_R^c} \eta_{2r_0}) \int \eta_{2r_0} |q|^2 \lesssim \exp(-\frac{R}{2r_0}). \tag{177}$$

On the other hand, we have $|f_1|^2 \overset{(71)}{\lesssim} \int \eta |f|^2$ and $\int \eta |f_r|^2 \leq \int \eta (|f|^2)_r \overset{(79)}{\lesssim} \int \eta |f|^2$ for all $r \leq 1$. Hence we have by the triangle inequality

$$\begin{aligned} ||q_r - \langle q\rangle|_1 - |\tilde q_r - \langle q\rangle|_1|^2 &\lesssim \int \eta ||q_r - \langle q\rangle| - |\tilde q_r - \langle q\rangle||^2 \\ &\leq \int \eta |(q - \tilde q)_r|^2 \lesssim \int \eta |q - \tilde q|^2. \end{aligned} \tag{178}$$

Inserting (177) and (178) into (176) yields (172). The same post-processing settles also the second uniform bound in (51) based on the first uniform bound.

Step 2. Proof of the statements (48) & (49) on the approximate locality and uniform boundedness of $(\int_0^T \nabla u d\tau, q(T) - \langle q(T)\rangle)$.
By scaling, it is again enough to consider $\sqrt{T} = 1$ and as above we write $\eta = \eta_{\sqrt{T}=1}$. For the use in several lemmas, we note that a solution of

$$\partial_t v - \nabla \cdot a \nabla v = \nabla \cdot g$$

satisfies (69) with $f = 0$, cf. Lemma 4, which we rewrite as

$$\frac{d}{dt} \exp(-Ct) \int \eta v^2 + \frac{1}{C} \exp(-Ct) \int \eta |\nabla v|^2 \leq C \exp(-Ct) \int \eta |g|^2. \tag{179}$$

In case of $v(t=0)=0$, this yields the a priori estimate

$$\int_0^1\int\eta|\nabla v|^2dt\lesssim\int_0^1\int\eta|g|^2. \tag{180}$$

We now apply this to $U(t):=\int_0^t ud\tau$, which is characterized by the initial value problem

$$\partial_t U-\nabla\cdot a\nabla U=\nabla\cdot ae,\quad U(t=0)=0.$$

From (180) applied to $v=U$ and $g=ae$, which also holds with η replaced by η_2, since we only used $|\nabla\eta_2|\lesssim\eta_2$, that

$$\int_0^1\int\eta_2|(\nabla U,q=a(\nabla U+e))|^2dt\lesssim\int_0^1\int\eta_2|ae|^2dt\sim 1. \tag{181}$$

As in the elliptic case above, we compare $U=U(a,t,x)$ to $\tilde U:=U(\tilde a,t,x)$ and note that the difference satisfies

$$\partial_t(U-\tilde U)-\nabla\cdot\tilde a\nabla(U-\tilde U)=\nabla\cdot\big((a-\tilde a)(\nabla U+e)\big),\quad (U-\tilde U)(t=0)=0,$$

so that once more by (180)

$$\int_0^1\int\eta|(\nabla(U-\tilde U),q-\tilde q)|^2dt\lesssim\int_0^1\int\eta|a-\tilde a|^2|q|^2dt\lesssim\sup_{B_R^c}\eta_2\int_0^1\int\eta_2|q|^2dt,$$

where in the second estimate, we used $\eta\sim\eta_2^2$ and our assumption in form of $\operatorname{supp}(a-\tilde a)\subset B_R^c$. We now insert (181) and obtain as intermediate result

$$\int_0^1\int\eta|(\nabla(U-\tilde U),q-\tilde q)|^2dt\lesssim\exp(-\frac{R}{2}). \tag{182}$$

In order to upgrade the time-averaged estimates (182) & (181) to the pointwise-in-time estimates (48) & (49), we need the following a priori estimate

$$\int\eta|\partial_t\nabla U|^2=\int\eta|\nabla u|^2\lesssim 1\quad\text{for } t\in[\frac{1}{2},1], \tag{183}$$

which we prove in the last step.

We upgrade (182): From the latter in form of

$$\int_{\frac12}^1\int\eta|\nabla(U-\tilde U)|^2dt\lesssim\exp(-\frac{R}{2})$$

and (183) in form of

$$\sup_{t\in(\frac12,1)}\int\eta|\partial_t\nabla(U-\tilde U)|^2\lesssim 1$$

we obtain via the elementary interpolation inequality

$$\int\eta|g|^2_{|t=1}\lesssim\frac{1}{\delta}\int_{1-\delta}^1\int\eta|g|^2dt+\delta^2\sup_{t\in(1-\delta,1)}\int\eta|\partial_t g|^2\quad\text{for }\delta\ll 1$$

which we use for $g = \nabla(U - \tilde{U})$, that

$$\int \eta |\nabla(U - \tilde{U})|^2_{|t=1} \lesssim \frac{1}{\delta} \exp(-\frac{R}{2}) + \delta^2.$$

The choice of $\delta = \exp(-\frac{R}{6})$ yields

$$\int \eta |(\nabla U - \nabla\tilde{U}, q - \tilde{q})|^2_{|t=1} \sim \int \eta |\nabla(U - \tilde{U})|^2_{|t=1} \lesssim \exp(-\frac{R}{3}).$$

Because of $|(\nabla U)_1 - (\nabla\tilde{U})_1|, |(q - \langle q\rangle)_1 - (\tilde{q} - \langle q\rangle)_1| \overset{(71)}{\lesssim} \big(\int \eta |(\nabla U - \nabla\tilde{U}, q - \tilde{q})|^2\big)^{\frac{1}{2}}$, this yields (48) in its rescaled version. The upgrade of (181) is similar.

Step 3. Proof of (183).
We split the statement into two:

$$\sup_{t\in[\frac{1}{2},1]} \int \eta |\nabla u|^2 \lesssim \int_0^1 t \int \eta u^2 dt \quad \text{and} \tag{184}$$

$$\int_0^1 t \int \eta u^2 dt \lesssim 1. \tag{185}$$

We start with (184), which can be obtained by time shifting in conjunction with parabolic rescaling of time-space from

$$\int \eta_R |\nabla u|^2_{|t=1} \lesssim \int_0^1 \int \eta_R u^2 dt \tag{186}$$

for any solution of the homogeneous equation $\partial_t u - \nabla \cdot a\nabla u = 0$ and any $R \geq 1$. Note that (186) follows as (157) in the proof of Lemma 11, which only required $|\nabla \eta_R| \leq \eta_R$, which is satisfied due to $R \geq 1$. For the same reason (186) also holds with η_R replaced by $\frac{1}{\eta_R}$ for any $R \geq 1$:

$$\int \frac{1}{\eta_R} |\nabla u|^2_{|t=1} \lesssim \int_0^1 \int \frac{1}{\eta_R} u^2 dt.$$

If in addition, the initial data are given by $u(t = 0) = u_0$, we obtain from (158) with η_R replaced by $\frac{1}{\eta_R}$ that $\int_0^1 \int \frac{1}{\eta_R} u^2 dt \lesssim \int \frac{1}{\eta_R} u_0^2$ so that the above turns into

$$\int \frac{1}{\eta_R} |\nabla u|^2_{|t=1} \lesssim \int \frac{1}{\eta_R} u_0^2.$$

By parabolic rescaling, this yields $t \int \frac{1}{\eta_{\sqrt{t}R}} |\nabla u(t)|^2 \lesssim \int \frac{1}{\eta_{\sqrt{t}R}} u_0^2$¿ For $t \leq 1$ we may choose $R = \frac{1}{\sqrt{t}} \geq 1$ to the effect of

$$t \int \frac{1}{\eta} |\nabla u(t)|^2 \lesssim \int \frac{1}{\eta} u_0^2. \tag{187}$$

Clearly, (187) also holds with a replaced by its transpose a^*; note that transposition exactly preserves (1). If now we have a solution of the inhomogeneous $\partial_t v - \nabla \cdot a^*\nabla v = f$ with homogeneous initial data $v_{|t=0} = 0$, we may write v in terms of the semi-group $S^*(t)$ for $-\nabla\cdot a^*\nabla$ as $v_{|t=1}=\int_0^1 S^*(1-t)f(t)dt$, so that (187), which amounts to the semi-group estimate $t\int\frac{1}{\eta}|\nabla S^*(t)u_0|^2\lesssim\int\frac{1}{\eta}u_0^2$, entails

$$\int\frac{1}{\eta}|\nabla v|^2_{|t=1}\lesssim\int\frac{1}{1-t}\int\frac{1}{\eta}|f|^2dt.$$

The dualization (by $t\rightsquigarrow 1-t$) of this estimate yields for our solution u of the homogeneous equation $\partial_t u-\nabla\cdot a\nabla u=0$ with inhomogeneous initial data $u(t=0)=\nabla\cdot ae$ that

$$\int t\int\eta|u|^2dt\lesssim\int\eta|ae|^2,$$

which turns into (187).

7.3. **Proof of Lemma 12: quantified equi-integrability.** We proceed into steps.

Step 1. Reduction.
By scale and translation invariance, it is enough to show for a solution of

$$\partial_t u-\nabla\cdot a\nabla u=0\quad\text{for}\quad t<0 \tag{188}$$

that

$$\Big(\int_{-\frac14}^0\int\eta_R|\nabla u|^2dxdt\Big)^{\frac12}\lesssim\Big(\frac{1}{R}\Big)^{\frac d2-\varepsilon}\Big(\int_{-1}^0\int\eta_1|\nabla u|^2dxdt\Big)^{\frac12}\quad\text{for } R\le 1. \tag{189}$$

Step 2. Meyers estimate.
Estimate (189) relies on the parabolic Meyer's estimate, which states that for the inhomogeneous equation

$$\partial_t v-\nabla\cdot a\nabla v=\nabla\cdot g$$

on the entire time-space $\mathbb{R}\times\mathbb{R}^d$ with compactly supported v and g, there exists an exponent $p=p(d,\lambda)>2$ (possible very close to 2) such that we have the Calderon-Zygmund type estimate

$$\Big(\int\int|\nabla v|^pdxdt\Big)^{\frac1p}\lesssim\Big(\int\int|g|^pdxdt\Big)^{\frac1p}. \tag{190}$$

We need to post-process this estimate in the following way: For

$$\partial_t v-\nabla\cdot a\nabla v=f+\nabla\cdot g \tag{191}$$

with compactly supported v, f and g, we claim

$$\Big(\int\int|\nabla v|^pdxdt\Big)^{\frac1p}\lesssim\Big(\int\Big(\int|f|^qdx\Big)^{\frac pq}dt\Big)^{\frac1q}+\Big(\int\int|g|^pdxdt\Big)^{\frac1p}, \tag{192}$$

where the exponent q is determined by scaling $q:=\frac{dp}{d+p}$. In fact, we will use (192) in the non-scale invariant form of

$$\Big(\int\int|\nabla v|^pdxdt\Big)^{\frac1p}\quad\lesssim\quad\sup_t\Big(\Big(\int f^2dx\Big)^{\frac12}+\Big(\int|g|^pdx\Big)^{\frac1p}\Big),$$
$$\text{provided }\operatorname{supp}(v,f,g)\subset[-1,0]\times\{|x|\le 1\}, \tag{193}$$

where we used that wlog we may assume $p \leq \frac{2d}{d-2}$ so that $q \leq 2$. Here comes the argument for (192): By the scale and translation invariance of the estimate, we may wlog assume that all functions involved in (191) have compact support in $[-1,0] \times \{|x| \leq 1\}$. Hence f must have vanishing space-time average. In other words, introducing the spatial average $\bar{f}(t) := \int f dx$ we have $\int \bar{f} dt = 0$. On the one hand, therefore there exists $\bar{F} = \bar{F}(t)$ supported in $t \in [-1,0]$ and such that $\frac{d\bar{F}}{dt} = \bar{f}$. By supp$f \subset [-1,0] \times \{|x| \leq 1\}$, we clearly have

$$(\int |\bar{F}|^p dt)^{\frac{1}{p}} \leq \sup_t |\bar{F}| \lesssim \int\int |f| dx dt \lesssim (\int (\int |f|^q dx)^{\frac{p}{q}} dt)^{\frac{1}{p}}. \tag{194}$$

On the one hand, $\bar{F}$ allows us to rewrite (191) as $\partial_t(v - \eta\bar{F}) - \nabla \cdot a\nabla(v - \eta\bar{F}) = f - \eta\bar{f} + \nabla \cdot (g + \bar{F}a\nabla\eta)$, where $\eta = \eta(x)$ is a cut-off function for $\{|x| \leq 1\}$ in $\{|x| \leq 2\}$ with $\int \eta dx = 1$. Fixing the time t, this allows us to solve the Neumann problem

$$-\triangle w = f - \bar{f}\eta \text{ for } |x| < 2, \quad -x \cdot \nabla w = 0 \text{ for } |x| = 2$$

on the ball $\{|x| \leq 2\}$. By maximal regularity for this Neumann problem and Sobolev's inequality (using $p = \frac{dq}{d-q}$) we have

$$(\int_{|x|\leq 2} |\nabla w|^p)^{\frac{1}{p}} \lesssim (\int_{|x|\leq 2} |f - \bar{f}\eta|^q dx)^{\frac{1}{q}} \lesssim (\int |f|^q dx)^{\frac{1}{q}}. \tag{195}$$

Note that by the Neumann boundary condition, when extending $-\nabla w$ trivially outside of $\{|x| \leq 2\}$ to a vector field δg it still solves $\nabla \cdot \delta g = f - \bar{f}\eta$, now in all space. Also the estimate (195) turns into

$$(\int\int |\delta g|^p dx dt)^{\frac{1}{p}} \lesssim (\int (\int |f|^q dx)^{\frac{p}{q}} dt)^{\frac{1}{p}}. \tag{196}$$

Hence with the abbreviations $\tilde{v} := v - \bar{F}\eta$ and $\tilde{g} := g + \bar{F}a\nabla\eta + \delta g$, we obtain $\partial_t\tilde{v} - \nabla \cdot a\nabla\tilde{v} = \nabla \cdot \tilde{g}$, so that an application of (190) yields by the triangle inequality, by the smoothness and compact support of η, and by the bound (1) on a,

$$(\int\int |\nabla v|^p dx dt)^{\frac{1}{p}} \lesssim (\int |\bar{F}|^p dt)^{\frac{1}{p}} + (\int\int |g|^p dx dt)^{\frac{1}{p}} + (\int\int |\delta g|^p dx dt)^{\frac{1}{p}}.$$

Inserting (194) and (196) into this yields (192).

Step 3. Energy estimates. We now turn to a solution of (188) and claim that there exists a constant c such that

$$\sup_{t\in(-\frac{1}{2},0)} \Big((\int_{|x|\leq\frac{1}{2}} |\nabla u|^2 dx)^{\frac{1}{2}} + (\int_{|x|\leq\frac{1}{2}} |u - c|^p dx)^{\frac{1}{p}} \Big) \lesssim (\int_{-1}^0 \int_{|x|\leq 1} |\nabla u|^2)^{\frac{1}{2}}. \tag{197}$$

We start with the first estimate in (197) and to this purpose select a cut-off function $\eta = \eta(x)$ for $\{|x| \leq \frac{1}{2}\}$ in $\{|x| \leq 1\}$. From the equation for u we obtain

$$\frac{d}{dt}\frac{1}{2}\int \eta^2 \nabla u \cdot a\nabla u dx = -\int \nabla \cdot (\eta^2 a\nabla u)(\nabla \cdot a\nabla u) dx.$$

Using $-\nabla\cdot(\eta^2 a\nabla u)(\nabla\cdot a\nabla u) = -\eta^2(\nabla\cdot a\nabla u)^2 - 2\eta(\nabla\eta\cdot a\nabla u)(\nabla\cdot a\nabla u) \le (\nabla\eta\cdot a\nabla u)^2 \overset{(1)}{\le} |\nabla\eta|^2|\nabla u|^2$, we get from integrating in time over (t_0,t) and then averaging over $t_0 \in (-1,-\frac{1}{2})$

$$\sup_{t\in(-\frac{1}{2},0)} \int \eta^2\nabla u\cdot a\nabla u dx \le 2\int_{-1}^{-\frac{1}{2}}\int \eta^2\nabla u\cdot a\nabla u dxdt + \int_{-1}^{0}\int |\nabla\eta|^2|\nabla u|^2 dxdt.$$

Using now the lower bound on a, cf. (1), and the definition of η, we obtain the first estimate in (197).

We now turn to the second estimate in (197). We select a smooth function $\eta = \eta(x)$ supported in $\{|x| \le \frac{1}{2}\}$ with $\int \eta dx = 1$. Denoting by $\bar u = \bar u(t)$ the average of u as defined by $\bar u = \int \eta u dx$ we have by the Poincaré-Sobolev inequality

$$\Big(\int_{|x|\le\frac{1}{2}} |u-\bar u|^p dx\Big)^{\frac{1}{p}} \lesssim \Big(\int_{|x|\le\frac{1}{2}} |\nabla u|^2 dx\Big)^{\frac{1}{2}},$$

where we use again that $p \le \frac{2d}{d-2}$. Hence in view of the first estimate in (197), it remains to show

$$\inf_c \sup_{t\in(-\frac{1}{2},0)} |\bar u - c| \lesssim \Big(\int_{-\frac{1}{2}}^{0} \Big|\frac{d\bar u}{dt}\Big|^2 dt\Big)^{\frac{1}{2}} \lesssim \Big(\int_{-\frac{1}{2}}^{0}\int_{|x|\le\frac{1}{2}} |\nabla u|^2 dxdt\Big)^{\frac{1}{2}}.$$

The latter follows from the equation which gives $\frac{d\bar u}{dt} = \int \nabla\eta\cdot a\nabla u dx$ via the Cauchy-Schwarz inequality.

Step 4. Conclusion. We now may conclude and to this purpose consider $v = \eta(u-c)$ where $\eta = \eta(t,x)$ is a cut-off function for $[-\frac{1}{4},0]\times\{|x|\le\frac{1}{4}\}$ in $[-\frac{1}{2},0]\times\{|x|\le\frac{1}{2}\}$ and c is the constant in (197). We note that by (188), v satisfies (191) with $f = (u-c)\partial_t\eta - \nabla\eta\cdot a\nabla u$ and $g = -(u-c)a\nabla\eta$, and that v, f, and g have support in $(-\frac{1}{2},0)\times\{|x|\le\frac{1}{2}\}$. Hence by an application of (193) we obtain

$$\Big(\int_{-\frac{1}{4}}^{0}\int_{|x|\le\frac{1}{4}} |\nabla u|^p dxdt\Big)^{\frac{1}{p}} \lesssim \sup_{t\in(-\frac{1}{2},0)} \Big(\Big(\int_{|x|\le 1} |\nabla u|^2 dx\Big)^{\frac{1}{2}} + \Big(\int |u-c|^p dx\Big)^{\frac{1}{p}}\Big).$$

Inserting (197) now yields the "reverse Hölder" inequality

$$\Big(\int_{-\frac{1}{4}}^{0}\int_{|x|\le\frac{1}{4}} |\nabla u|^p dxdt\Big)^{\frac{1}{p}} \lesssim \Big(\int_{-1}^{0}\int_{|x|\le 1} |\nabla u|^2 dxdt\Big)^{\frac{1}{2}}.$$

In combination with Hölder's inequality in form of

$$\Big(\int_{-\frac{1}{4}}^{0}\int_{|x|\le r} |\nabla u|^2 dxdt\Big)^{\frac{1}{2}} \lesssim r^{d(\frac{1}{2}-\frac{1}{p})}\Big(\int_{-\frac{1}{4}}^{0}\int_{|x|\le r} |\nabla u|^p dxdt\Big)^{\frac{1}{p}},$$

we obtain

$$\Big(\int_{-\frac{1}{4}}^{0}\int_{|x|\le r} |\nabla u|^2 dxdt\Big)^{\frac{1}{2}} \lesssim r^{\varepsilon}\Big(\int_{-1}^{0}\int_{|x|\le 1} |\nabla u|^2 dxdt\Big)^{\frac{1}{2}} \tag{198}$$

with $\varepsilon = \varepsilon(d,\lambda) := d(\frac{1}{2} - \frac{1}{p}) > 0$ for $r \le \frac{1}{4}$ and also for $r \le 1$ (since the estimate is trivial for $r \in [\frac{1}{4}, 1]$).

It remains to introduce the exponential averaging functions η_R and η_1. For given $R \le \frac{1}{2}$ we appeal to the co-area formula in form of $\int \eta_R |\nabla u|^2 dx = \frac{1}{R^d}\int \exp(-\frac{|x|}{R})|\nabla u|^2 dx = \frac{1}{R^d}\int_0^\infty \exp(-\frac{r}{R})\int_{|x|=r}|\nabla u|^2 dx dr = \frac{1}{R^{d+1}}\int_0^\infty \exp(-\frac{r}{R})\int_{|x|\le r}|\nabla u|^2 dx dr$. From splitting the last integral into $r \le 1$ and $r \ge 1$ and using (198) on the first part we obtain

$$\int_{-\frac{1}{4}}^0 \int \eta_R |\nabla u|^2 dx dt \lesssim \frac{1}{R^{d+1}}\int_0^1 \exp(-\frac{r}{R}) r^{2\varepsilon} dr \int_{-1}^0 \int_{|x|\le 1} |\nabla u|^2 dx dt + \frac{1}{R^{d+1}}\int_1^\infty \exp(-\frac{r}{R})\int_{-1}^0\int_{|x|\le r}|\nabla u|^2 dx dt dr. \tag{199}$$

For the first rhs term we use $\frac{1}{R^{d+1}}\int_0^1 \exp(-\frac{r}{R})r^{2\varepsilon}dr \le \frac{1}{R^{d+1}}\int_0^\infty \exp(-\frac{r}{R})r^{2\varepsilon}dr \sim \frac{1}{R^{d-2\varepsilon}}$ and $\int_{|x|\le 1}|\nabla u|^2 dx \lesssim \int \eta_1 |\nabla u|^2 dx$ so that

$$\frac{1}{R^{d+1}}\int_0^1 \exp(-\frac{r}{R}) r^{2\varepsilon} dr \int_{-1}^0 \int_{|x|\le 1}|\nabla u|^2 dx dt \lesssim \frac{1}{R^{d-2\varepsilon}}\int_{-1}^0\int \eta_1 |\nabla u|^2 dx dt. \tag{200}$$

For the second rhs term in (199) we note that for $R \le \frac{1}{2}$ and $r \ge 1$ we have $\exp(-\frac{r}{R}) \lesssim R^{1+2\varepsilon}\exp(-r)$ (since in this range $\exp(-\frac{r}{R})\exp(r) = \exp(-(\frac{1}{R}-1)r) \le \exp(-\frac{1}{2R})$) so that we have for the second term

$$\frac{1}{R^{d+1}}\int_1^\infty \exp(-\frac{r}{R})\int_{-1}^0\int_{|x|\le r}|\nabla u|^2 dx dt dr \lesssim \frac{1}{R^{d-2\varepsilon}}\int_0^\infty \exp(-r)\int_{-1}^0\int_{|x|\le r}|\nabla u|^2 dx dt dr,$$

which by the co-area formula yields

$$\frac{1}{R^{d+1}}\int_1^\infty \exp(-\frac{r}{R})\int_{-1}^0\int_{|x|\le r}|\nabla u|^2 dx dt dr \lesssim \frac{1}{R^{d-2\varepsilon}}\int_{-1}^0\int \eta_1|\nabla u|^2 dx dt.$$

Inserting this and (200) into (199) yields (189).

Acknowledgements

We thank Peter Bella and Julian Fischer for suggestions on the manuscript. AG acknowledges financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2014-2019 Grant Agreement QUANTHOM 335410).

References

[1] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Preprint*, 2015. arXiv:1507.06935.

[2] S. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *to appear in Arch. Ration. Mech. Anal.*, 2015. arXiv:1411.3668.

[3] S. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *to appear in Ann. Sci. Éc. Norm. Supér*, 2015. arXiv:1406.0996.

[4] M. Avellaneda and F.H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40:803–847, 1987.

[5] P. Bella, A. Giunti, and F. Otto. Quantitative stochastic homogenization : local control of homogenization error through correctors. *Preprint*, 2015. arXiv:1504.02487.

[6] P. Bella and F. Otto. Corrector estimates for elliptic systems with random periodic coefficients. *Preprint*, 2014. arXiv:1409.5271.

[7] J. G. Conlon and A. Naddaf. Greens functions for elliptic and parabolic equations with random coefficients. *New York J. Math.*, 6:153–225, 2000.

[8] J. G. Conlon and T. Spencer. Strong convergence to the homogenized limit of elliptic equations with random coefficients. *Trans. Amer. Math. Soc.*, 366:1257–1288, 2014.

[9] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.

[10] M. Duerinckx and A. Gloria. Quantitative stochastic homogenization via weighted functional inequalities. *in preparation.*

[11] J. Fischer and F. Otto. Sublinear growth of the corrector in stochastic homogenization: optimal stochastic estimates for slowly decaying correlations. *Preprint*, 2015. arXiv:1508.00025.

[12] A. Gloria and J.-C. Mourrat. Spectral measure and approximation of homogenized coefficients. *Probab. Theory Related Fields*, 154:287–326, 2012.

[13] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators and homogenization. *Preprint*, 2014. arXiv:1409.2678.

[14] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Inventiones mathematicae*, 199(2):455–515, 2015.

[15] A. Gloria and J. Nolen. A quantitative central limit theorem for the effective conductance on the discrete torus. *to appear in Comm. Pure Appl. Math.*, 2015. arXiv:1410.5734.

[16] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *to appear in Journ. Eur. Math. Soc. (JEMS).* arXiv:1409.0801.

[17] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 05 2011.

[18] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 02 2012.

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

[20] D. Marahrens and F. Otto. Annealed estimates on the green's function. *Probability theory and related fields.* arXiv:1304.4408.

[21] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47:294–327, 2011.

[22] F. Murat and L. Tartar. H-convergence. In *Topics in the mathematical modelling of composite materials*, volume 31 of *Progr. Nonlinear Differential Equations Appl.*, pages 21–43. Birkhäuser Boston, Boston, MA, 1997.

[23] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. *Unpublished preprint*, 1998.

[24] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random Fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloquia Mathematica Societatis Jnos Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[25] L. Tartar. *The general theory of homogenization*, volume 7 of *Lecture Notes of the Unione Matematica Italiana.* Springer-Verlag, Berlin; UMI, Bologna, 2009. A personalized introduction.

[26] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27:167–180, 215, 1986.

(Antoine Gloria) Université Libre de Bruxelles (ULB), Brussels, Belgium, and Inria, Villeneuve d'Ascq, France
E-mail address: agloria@ulb.ac.be

(Felix Otto) Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, Germany
E-mail address: otto@mis.mpg.de