

**Quantification of ergodicity in stochastic  
homogenization: optimal bounds via spectral gap on  
Glauber dynamics — long version**

Antoine Gloria, Stefan Neukamm, Felix Otto

► **To cite this version:**

Antoine Gloria, Stefan Neukamm, Felix Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics — long version. [Research Report] 2013, pp.106. <hal-00782075>

**HAL Id: hal-00782075**

**<https://hal.inria.fr/hal-00782075>**

Submitted on 29 Jan 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics — long version

A. Gloria <sup>\*</sup>      S. Neukamm<sup>†</sup>      F. Otto<sup>‡</sup>

January 29, 2013

## Abstract

We study the effective large-scale behavior of discrete elliptic equations on the lattice  $\mathbb{Z}^d$  with random coefficients. The theory of stochastic homogenization relates the random but stationary field of coefficients with a deterministic matrix of effective coefficients. This is done via the corrector problem, which can be viewed as a highly degenerate elliptic equation on the infinite-dimensional space of admissible coefficient fields. In this contribution we develop quantitative methods for the corrector problem assuming that the ensemble of coefficient fields satisfies a spectral gap estimate w. r. t. a Glauber dynamics. As a main result we prove an optimal estimate for the decay in time of the parabolic equation associated to the corrector problem (i. e. for the “random environment as seen from a random walker”). As a corollary we obtain existence and moment bounds for stationary correctors (in dimension  $d > 2$ ) and optimal estimates for regularized versions of the corrector (in dimensions  $d \geq 2$ ). We also give a self-contained proof for a new estimate on the gradient of the parabolic, variable-coefficient Green’s function, which is a crucial analytic ingredient in our method.

As an application, we study the approximation of the homogenized coefficients via a representative volume element. The approximation introduces two types of errors. Based on our quantitative methods, we develop an error analysis that gives optimal bounds in terms of scaling in the size of the representative volume element — even for large ellipticity ratios.

---

<sup>\*</sup>Université Libre de Bruxelles (ULB), Brussels, Belgium & Project-team SIMPAF, Inria Lille - Nord Europe, Villeneuve d’Ascq, France, agloria@ulb.ac.be

<sup>†</sup>Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany, neukamm@mis.mpg.de

<sup>‡</sup>Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany, felix.otto@mis.mpg.de

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>General framework and notation</b>	<b>6</b>
<b>3</b>	<b>Green's function representation</b>	<b>7</b>
<b>4</b>	<b>Quantification of ergodicity via spectral gap</b>	<b>16</b>
4.1	The periodic ensemble . . . . .	22
4.2	Proofs . . . . .	24
<b>5</b>	<b>Decay of the variable-coefficient semigroup</b>	<b>28</b>
5.1	Proofs . . . . .	33
<b>6</b>	<b>Application to stochastic homogenization</b>	<b>41</b>
6.1	Qualitative theory of stochastic homogenization . . . . .	41
6.2	Quantitative estimates for the corrector . . . . .	43
6.3	Approximation of the homogenized coefficients by periodization . . . . .	47
6.3.1	Proofs . . . . .	54
<b>7</b>	<b>Estimates on the parabolic Green's function with uniformly elliptic coefficients</b>	<b>63</b>
7.1	The periodic continuum case . . . . .	63
7.2	The periodic discrete case – Proof of Theorem 3 (b) . . . . .	76
7.3	The discrete, whole space case – Proof of Theorem 3 (a) . . . . .	78
7.4	Discrete Calderón-Zygmund estimate . . . . .	80
7.4.1	The elliptic case . . . . .	82
7.4.2	The parabolic case . . . . .	93

# 1 Introduction

In the 70s Yurinskii [36], Kozlov [21], and Papanicolaou & Varadhan [32] proved first results for the homogenization of elliptic equations with random coefficients. They considered the elliptic operator  $-\nabla \cdot \mathbf{a}(\frac{x}{\varepsilon})\nabla$  (on a domain in  $\mathbb{R}^d$ ) with random, uniformly elliptic coefficients  $\mathbf{a}(x) \in \mathbb{R}^{d \times d}$  and studied its asymptotic behavior in the large scale limit  $\varepsilon \downarrow 0$ . For coefficients that are stationary and ergodic with respect to *horizontal shifts* of the form  $\mathbf{a}(\cdot) \mapsto \mathbf{a}(\cdot + x)$ ,  $x \in \mathbb{R}^d$ , they proved a homogenization result which shows that for  $\varepsilon \downarrow 0$  the rapidly oscillating random coefficients  $\mathbf{a}(\frac{\cdot}{\varepsilon})$  can be replaced by *homogenized coefficients*  $\mathbf{a}_{\text{hom}}$  which are constant in space and deterministic.

Homogenization of PDEs with rapidly varying coefficients has first been studied in the deterministic case, and in particular, in settings with periodic coefficients, see e.g. [34, 27, 3, 7] and the references therein. Many homogenization problems can be understood, analyzed or rephrased in terms of *correctors*—functions that are capable of describing the oscillations induced by the coefficients. Let us briefly comment on the notion of correctors in a simple setting, namely in the case of deterministic, periodic (say with period 1) and symmetric coefficients. In this case the corrector  $\varphi$  associated with a macroscopic direction, say  $e \in \mathbb{R}^d$ , is given by the “periodic” corrector problem

$$(1) \quad -\nabla \cdot \mathbf{a}(y)(e + \nabla\varphi(y)) = 0 \quad \text{on the torus in } \mathbb{R}^d \text{ of size 1.}$$

The effective flux associated with  $e$  is given by the average

$$\mathbf{a}_{\text{hom}}e = \int \mathbf{a}(e + \nabla\varphi)$$

where  $\int$  stands for the spatial average over the torus in  $\mathbb{R}^d$  of size 1. From the PDE point of view,  $\varphi$  is the function that “corrects” the affine function  $x \mapsto e \cdot x$  into an  $\mathbf{a}$ -harmonic function. From the variational point of view,  $\varphi$  describes an optimal profile that adapts the affine function  $x \mapsto e \cdot x$  to the variable-coefficient field  $\mathbf{a}$ . In particular, in the variational case (when  $\mathbf{a}$  is symmetric) the homogenized coefficients are given by the homogenization formula

$$e \cdot \mathbf{a}_{\text{hom}}e = \min_{\varphi} \int (e + \nabla\varphi) \cdot \mathbf{a}(e + \nabla\varphi)$$

where the infimum is taken over all 1-periodic functions  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ .

Stochastic homogenization is, on the one hand, similar to the periodic case discussed above: The analysis is based on the corrector and we can argue, as in the periodic case, by appealing to Tartar’s methods of oscillating test functions. On the other hand, the stochastic setting is different and more challenging: While the corrector in the deterministic periodic case is given by the elliptic PDE (1), which is defined on a finite dimensional torus, in the random case, the corrector problem lives in the probability space of all admissible coefficient fields — which has infinite dimension. As a consequence, quantitative estimates are not abundant in the study of the random corrector, and mostly qualitative tools from ergodicity theory have been employed.

In the present contribution we consider stochastic homogenization in a discrete setting where the continuum domain  $\mathbb{R}^d$  is replaced by the rescaled lattice  $\varepsilon\mathbb{Z}^d$ . We introduce quantitative methods that, in particular, allow for several optimal estimates associated with the corrector problem in the random setting. The methods continue and extend earlier ideas presented in [17, 18] by the first and third author.

To display some of our ideas and results, let us anticipate that in the discrete setting, qualitative homogenization theory yields the homogenization formula

$$(2) \quad e \cdot \mathbf{a}_{\text{hom}}e := \inf_{\phi \in L^2(\Omega)} \langle (e + D\phi) \cdot \mathbf{a}(0)(e + D\phi) \rangle.$$

Here  $\Omega \subset (\mathbb{R}^{d \times d})^{\mathbb{Z}^d}$  denotes the space of admissible coefficient fields  $\mathbf{a} : \mathbb{Z}^d \rightarrow \mathbb{R}^{d \times d}$  that are diagonal, uniformly bounded and elliptic.  $\langle \cdot \rangle$  denotes the expected value of a probability measure defined on  $\Omega$

that we call the ensemble. It is assumed to be stationary w. r. t. horizontal shifts of the coefficients. This means that the shifts  $T_x : \Omega \rightarrow \Omega$ ,  $\mathbf{a} \mapsto \mathbf{a}(\cdot + x)$  are measure preserving for all  $x \in \mathbb{Z}^d$ . In (2)  $D\phi$  denotes what we call the *horizontal derivative*. It is defined for random variables  $\phi$  on  $\Omega$  via

$$D\phi(\mathbf{a}) := (D_1\phi(\mathbf{a}), \dots, D_d\phi(\mathbf{a})), \quad D_i\phi(\mathbf{a}) := \phi(\mathbf{a}(\cdot + e_i)) - \phi(\mathbf{a}).$$

The horizontal derivative is associated with variations obtained by horizontally shifting the coefficient field  $\mathbf{a}$  in the  $d$  independent directions of  $\mathbb{Z}^d$ . It naturally emerges in stochastic homogenization, since it is equivalent to the spatial gradient  $\nabla_i u(\mathbf{a}, x) := u(\mathbf{a}, x + e_i) - u(\mathbf{a}, x)$  for stationary random fields  $u(\mathbf{a}, x)$  (see Section 2 for the details).

The Dirichlet energy  $L^2(\Omega) \ni \phi \mapsto \langle (e + D\phi) \cdot \mathbf{a}(0)(e + D\phi) \rangle$  is convex, quadratic and coercive in the sense that it controls the  $L^2$ -norm of  $D\phi$ . However, since  $\Omega$  is a product space with an infinite number of factors, and the horizontal derivative, morally speaking, only controls  $d$  degrees of freedom, we cannot expect a Poincaré inequality (w. r. t.  $D$ ) to hold on  $L^2(\Omega)$  in general. As a consequence, (2) may not admit a minimizer, and it is not clear at all how to develop a quantitative theory at this stage.

For a quantitative analysis, an additional assumption on the statistics of the coefficient field  $\mathbf{a}$  is needed. Our basic idea is the following: as a replacement for the “missing” Poincaré inequality (w. r. t.  $D$ ) we appeal to a different Poincaré inequality which relies on, what we call, the *vertical derivative*  $\frac{\partial\phi}{\partial y}$  – a discrete version of the usual partial derivative  $\frac{\partial\phi(\mathbf{a})}{\partial\mathbf{a}(y)}$  (see Definition 5 for a precise definition). Since with each site  $y \in \mathbb{Z}^d$  a vertical derivative is associated, the dimensionality of  $\Omega$  and the degrees of freedom that we can control with the vertical derivatives match. Therefore, it is reasonable to assume that we have such a “vertical” Poincaré inequality on  $L^2(\Omega)$ . Indeed, as is well-known independent and identically distributed (i. i. d.) coefficients satisfy that assumption. The “vertical” Poincaré inequality has a natural interpretation in statistical mechanics, where it corresponds to the spectral gap of a “heat bath” Glauber dynamics on the space of coefficient fields  $\Omega$ . For this reason we refer to that Poincaré inequality as the *spectral gap* of Glauber dynamics (SG).

In order to build up a quantitative theory based on (SG), we have to establish a connection between vertical derivatives and the horizontal derivatives. As we shall see in Sections 4 and 5 this is possible in situations when the function under consideration, say  $u$ , is the solution to an elliptic or parabolic equation associated with the elliptic operator  $D^*D$  and  $D^*\mathbf{a}(0)D$ , respectively. In Section 4 we study the constant-coefficient case, i. e. the evolution generated by  $D^*D$ , and find that (SG) allows one to quantify ergodicity. In Section 5 we study the more intricate case of variable coefficients, i. e.  $D^*\mathbf{a}(0)D$ .

A central ingredient in both cases is the observation that for  $u(t)$  given by a parabolic equation, the vertical derivatives  $\frac{\partial u(t)}{\partial y}$  can be connected to  $u(t)$  (and its horizontal derivative) by appealing to the parabolic *Green’s function* on  $\mathbb{Z}^d$ . Seeking for quantitative statements, estimates on the gradient of the parabolic Green’s function are required. Those estimates must reflect the optimal “decay in time” and they should be deterministic or “quenched”, which means we need estimates that are uniform in  $\mathbf{a} \in \Omega$ . We establish the estimates by adapting classical methods from elliptic and parabolic regularity theory, such as Calderón-Zygmund-, Meyers’-, Caccioppoli- and Nash-Aronson-type estimates, to the discrete setting. This is done in Section 7.

In Section 6 we apply our methods to stochastic homogenization. As a result of our methods, we find that for dimensions  $d > 2$  stationary correctors in  $L^2(\Omega)$  exist, and thus, the corrector problem in the probability space  $L^2(\Omega)$

$$(3) \quad D^*\mathbf{a}(0)(e + D\phi) = 0$$

is indeed well-defined. Additionally, we obtain higher moment bounds on the corrector (in dimension  $d > 2$ ) and its gradient (in dimension  $d \geq 2$ ). Since our methods extend to the “non-ergodic” situation when  $\langle \cdot \rangle$  concentrates on  $L$ -periodic coefficients, we obtain optimal error estimates for the approximation of the homogenized coefficients by periodization methods. This is explained in Section 6.3.

Yet, another direct application of our methods is Corollary 8, which quantifies the spectral bottom of  $D^*\mathbf{a}(0)D$ . We observe that (SG) implies that the spectrum is “thin” at the bottom.

In the remainder of this introduction we comment on various connections to previous works. The idea to combine methods coming from stochastic homogenization and from statistical mechanics naturally comes up in the study of scaling limits of Gradient Gibbs Measures. Gradient Gibbs measures can be seen as a model for thermally fluctuating interfaces, see [40] for a review. In fact, our work is inspired by that of Naddaf & Spencer [28] (see also Giacomin, Olla & Spohn [12]) on Gradient Gibbs Measures, which combines all the three concepts of (discrete) spatial, of horizontal, and of vertical derivatives. A main ingredient in [28] are estimates on the gradient of the elliptic Green’s function. Implicitly, they obtain  $\ell^p$ -estimates (for some  $p$  slightly larger than 2) via Meyer’s argument. They use it to establish large-scale Gaussian behavior of the Gradient Gibbs measure via stochastic homogenization for dimension  $d = 2$ . Delmotte & Deuschel [38] obtained optimal annealed estimates on derivatives of the parabolic Green’s function and used them to obtain decay estimates for the space-time covariances for the Ginzburg-Landau Gradient interface model. Recall that *annealed* means that the estimates only hold after taking the expectation (see also [6] for annealed estimates on the elliptic and parabolic Green’s function).

The main novelty of our results is that we use optimal, *deterministic* bounds on the *gradient* of the variable-coefficient Green’s functions. Here deterministic means that the bounds depend on the coefficients only via their ellipticity contrast. The first result in that direction has been obtained by the first and third author in [17], where estimates on the gradient of the *elliptic* Green’s function on  $\mathbb{Z}^d$  are established. In the present contribution, we establish an optimal, deterministic estimate for the gradient of the *parabolic* Green’s function on  $\mathbb{Z}^d$ . The estimate reflects the optimal decay in space-time and is uniform in the coefficients. Since such an estimate can only hold in a spatially averaged sense, we appeal to a weighted  $\ell_x^p$ -space (with  $p$  slightly larger than 2) to capture sufficiently fine information on the decay.

Another source of inspiring ideas are the early works by Papanicolaou & Varadhan [32] and Kipnis & Varadhan [20]. In their qualitative analysis, the authors regularize (3) by introducing an additional lower order term to the corrector problem (3), leading to, what we call the modified corrector problem:

$$(4) \quad \mu\phi_\mu + D^*\mathbf{a}(0)(e + D\phi_\mu), \quad 0 < \mu \ll 1.$$

As a merit of the regularization, the modified corrector problem (4) always admits a unique solution in  $L^2(\Omega)$ . That solution can be used to introduce an approximation  $\mathbf{a}_{\text{hom},\mu}$  for the homogenized coefficients  $\mathbf{a}_{\text{hom}}$ . In [32, 20] the authors appeal to spectral analysis to treat the original corrector problem (3) with the help of its modified version (4). Furthermore, they devise a spectral representation formula for the homogenized coefficients and its approximation  $\mathbf{a}_{\text{hom},\mu}$ .

In the first quantitative paper on stochastic homogenization, Yurinskii [37] considered the dependency of the gradient of the modified corrector  $D\phi_\mu$  on  $\mu$ . For dimension  $d > 2$  and under mixing assumptions on the statistics, he obtained non-optimal decay-in- $\mu$  estimates on the  $L^2(\Omega)$ -distance between the gradient  $D\phi_\mu$  of the modified corrector and that of the original corrector  $\phi$ . His analysis does not rely on spectral analysis. As a central ingredient, Yurinskii appeals to deterministic estimates on the parabolic Green’s function (as apposed to its gradient) that rely on the classical regularity theory.

In [18] the first and third author obtained optimal estimates for the *systematic error*  $|\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{hom},\mu}|$  for dimensions  $d \geq 2$  and i. i. d. coefficients. Recently, Mourrat [26] and the first author and Mourrat [15] obtained further quantitative results for the *systematic error* by appealing to the spectral calculus mentioned above. In particular, in [15], motivated by the spectral representation of  $\mathbf{a}_{\text{hom}}$ , a different “higher order” approximation scheme for  $\mathbf{a}_{\text{hom}}$  is introduced, and optimal error estimates for the associated systematic error are obtained for dimensions  $2 \leq d \leq 6$  in the framework of i. i. d. coefficients. We show in Section 6.2 that the quantitative methods that we develop in the present contribution yield optimal estimates for the systematic error in any dimension and for any ensemble that satisfies the (SG) assumption. In particular, we positively answer a conjecture in [15] on the spectral bottom.

## 2 General framework and notation

**Random coefficient fields.** In this paper we consider linear second order difference equations with uniformly elliptic, bounded random coefficients. We denote the set of admissible coefficient matrices by

$$(5) \quad \Omega_0 := \left\{ \mathbf{a} \in \mathbb{R}^{d \times d} : \mathbf{a} \text{ is diagonal,} \right. \\ \left. \begin{aligned} &\mathbf{a} \text{ is bounded, i. e. } |\mathbf{a}\xi| \leq |\xi| \text{ for all } \xi \in \mathbb{R}^d, \\ &\mathbf{a} \text{ is elliptic, i. e. } \xi \cdot \mathbf{a}\xi \geq \lambda|\xi|^2 \text{ for all } \xi \in \mathbb{R}^d \end{aligned} \right\}.$$

Here and below  $\lambda \in (0, 1)$  denotes the ellipticity constant which is fixed throughout the paper. We equip  $\Omega_0$  with the usual topology of  $\mathbb{R}^{d \times d}$ . A *coefficient field*, denoted by  $\mathbf{a}$ , is a function on  $\mathbb{Z}^d$  taking values in  $\Omega_0$ . It can either be viewed as a mapping  $\mathbf{a} : \mathbb{Z}^d \rightarrow \Omega_0$  or as an element  $\{\mathbb{Z}^d \ni x \mapsto \mathbf{a}(x)\}$  of the infinite Cartesian product  $\Omega := (\Omega_0)^{\mathbb{Z}^d}$ . We equip  $\Omega$  with the product topology, i. e. the coarsest topology for which all the projections  $\Omega \ni \mathbf{a} \mapsto \mathbf{a}(x) \in \Omega_0$ ,  $x \in \mathbb{Z}^d$ , are continuous. For  $L \in \mathbb{N}$  we denote by  $\Omega_L$  the subspace of  $L$ -periodic conductivity fields, (i. e.  $\mathbf{a} \in \Omega$  with  $\mathbf{a}(x + Lz) = \mathbf{a}(x)$  for all  $x, z \in \mathbb{Z}^d$ ).

We describe a *random coefficient field* by equipping  $\Omega$  with a probability measure (hereby, we always assume that the underlying  $\sigma$ -algebra is the Borel- $\sigma$ -algebra, i. e. the coordinate projections  $\Omega \ni \mathbf{a} \mapsto \mathbf{a}(x) \in \mathbb{R}^{d \times d}$  are measurable for all  $x \in \mathbb{Z}^d$ ). Following the convention in statistical mechanics, we call a probability measure on  $\Omega$  also an *ensemble* and denote the associated ensemble average (i. e. the expected value) by  $\langle \cdot \rangle$ . If  $\langle \cdot \rangle$  concentrates on  $\Omega_L$  we call it a *L-periodic ensemble*.

Unless otherwise stated we always assume that  $\langle \cdot \rangle$  is *stationary*, i. e. for all translations  $z \in \mathbb{Z}^d$  the coefficient fields  $\{\mathbb{Z}^d \ni x \mapsto \mathbf{a}(x)\}$  and  $\{\mathbb{Z}^d \ni x \mapsto \mathbf{a}(x + z)\}$  have the same joint distribution. Let  $T_z : \Omega \rightarrow \Omega$ ,  $\mathbf{a}(\cdot) \mapsto \mathbf{a}(\cdot + z)$  denote the shift by  $z$ . Then  $\langle \cdot \rangle$  is stationary, if and only if  $T_z$  is  $\langle \cdot \rangle$ -preserving for all shifts  $z \in \mathbb{Z}^d$ .

**Random variables and stationary random fields.** A random variable is a measurable function on  $\Omega$ . We say a random variable is *local*, if it only depends on the value of the coefficient field at a finite number of sites. We denote by  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ , the usual Banach spaces of random variables with finite  $p$ -th moment. Recall that  $\Omega = \Omega_0^{\mathbb{Z}^d}$  is equipped with the product topology associated with  $\Omega_0 \subset \mathbb{R}^{d \times d}$ . Since  $\Omega_0$  is a metric space, the product topology on  $\Omega$  is metrizable. (Notice that  $(\Omega, |\cdot|_\infty)$ , where  $|\mathbf{a}|_\infty := \sup_{x \in \mathbb{Z}^d} |\mathbf{a}(x)|$ , has a different topology). We denote by  $C^0(\Omega)$  the Banach space of bounded continuous functions on  $\Omega$  equipped with the norm  $\|\zeta\|_\infty := \sup_{\mathbf{a} \in \Omega} |\zeta(\mathbf{a})| < \infty$ . Furthermore,  $C^0(\Omega)$  is dense in every  $L^p(\Omega)$ ,  $1 \leq p < \infty$ .

A *random field*  $\tilde{\zeta}$  is a measurable function on  $\Omega \times \mathbb{Z}^d$ . Here measurable means that  $\tilde{\zeta}(\cdot, x)$  is measurable for all  $x \in \mathbb{Z}^d$ . With any random variable  $\zeta : \Omega \rightarrow \mathbb{R}$  we associate its  $\langle \cdot \rangle$ -stationary extension  $\bar{\zeta} : \Omega \times \mathbb{Z}^d \rightarrow \mathbb{R}$  via  $\bar{\zeta}(\mathbf{a}, x) := \zeta(\mathbf{a}(\cdot + x))$ . Conversely, we say that a random field  $\tilde{\zeta}$  is  $\langle \cdot \rangle$ -stationary if it there exists a random variable  $\zeta$  with  $\tilde{\zeta}(\mathbf{a}, \cdot) = \bar{\zeta}(\mathbf{a}, \cdot)$   $\langle \cdot \rangle$ -almost surely. If  $\langle \cdot \rangle$  is stationary, then the ensemble average of a stationary random field  $\bar{\zeta}$  is independent of  $x \in \mathbb{Z}^d$ ; therefore we simply write  $\langle \bar{\zeta} \rangle$  instead of  $\langle \tilde{\zeta}(x) \rangle$ .

**Spatial and horizontal derivatives.** For fields  $\zeta : \mathbb{Z}^d \rightarrow \mathbb{R}$ , vector fields  $\xi = (\xi_1, \dots, \xi_d) : \mathbb{Z}^d \rightarrow \mathbb{R}^d$  and all  $i = 1, \dots, d$  we define the *spatial derivatives*

$$\begin{aligned} \nabla_i \zeta(x) &:= \zeta(x + e_i) - \zeta(x), & \nabla_i^* \zeta(x) &:= \zeta(x - e_i) - \zeta(x), \\ \nabla \zeta &= (\nabla_1 \zeta, \dots, \nabla_d \zeta), & \nabla^* \xi &= \sum_{i=1}^d \nabla_i^* \xi_i. \end{aligned}$$

Here  $e_1, \dots, e_d$  is the canonical basis of  $\mathbb{R}^d$ ,  $\nabla$  is the discrete gradient for functions on  $\mathbb{Z}^d$  and  $-\nabla^*$  is the discrete divergence for vector fields on  $\mathbb{Z}^d$ . Notice that  $\nabla_i$  and  $\nabla_i^*$  are bounded linear operators

on  $\ell^p(\mathbb{Z}^d)$ ,  $1 \leq p \leq \infty$ . As usual  $\ell^p(\mathbb{Z}^d)$  stands for the Banach space of functions on  $\mathbb{Z}^d$  that are  $p$ -summable if  $1 \leq p < \infty$ , and bounded if  $p = \infty$ . In particular,  $\nabla_i^*$  is precisely the  $\ell^2(\mathbb{Z}^d)$ -adjoint of  $\nabla_i$ , so that we have the *discrete integration by parts formula*

$$(6) \quad \sum_{x \in \mathbb{Z}^d} \nabla \zeta(x) \cdot \xi(x) = \sum_{x \in \mathbb{Z}^d} \zeta(x) \nabla^* \xi(x)$$

for all  $\zeta \in \ell^2(\mathbb{Z}^d)$  and  $\xi \in \ell^2(\mathbb{Z}^d)^d$ . Similarly, for  $L$ -periodic  $\zeta$  and  $\xi$  we have

$$(7) \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla \zeta(x) \cdot \xi(x) = \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \zeta(x) \nabla^* \xi(x).$$

Next, we introduce a similar structure for random variables: For scalar random variables  $\zeta : \Omega \rightarrow \mathbb{R}$ , vector-valued random variables  $\xi = (\xi_1, \dots, \xi_d) : \Omega \rightarrow \mathbb{R}^d$  and  $i = 1, \dots, d$  we define the *horizontal derivatives*

$$D_i \zeta(\mathbf{a}) := \zeta(\mathbf{a}(\cdot + e_i)) - \zeta(\mathbf{a}), \quad D_i^* \zeta(\mathbf{a}) := \zeta(\mathbf{a}(\cdot - e_i)) - \zeta(\mathbf{a}),$$

$$D \zeta := (D_1 \zeta, \dots, D_d \zeta), \quad D^* \xi := \sum_{i=1}^d D_i^* \xi_i.$$

$D_i$  and  $D_i^*$  are bounded linear operators on  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ ; and, since  $\langle \cdot \rangle$  is stationary,  $D_i^*$  is the  $L^2(\Omega)$ -adjoint of  $D_i$ . In particular, for  $\zeta \in L^p(\Omega)$  and  $\xi \in L^q(\Omega)^d$  where  $\frac{1}{p} + \frac{1}{q} = 1$  (with the convention  $\frac{1}{\infty} = 0$ ), we have

$$\langle D \zeta \cdot \xi \rangle = \langle \zeta D^* \xi \rangle.$$

An elementary but important observation is the following. Let  $\overline{(\cdot)}$  denote the mapping that associates a random variable with its stationary extension. Then

$$(8) \quad \nabla_i \bar{\zeta} = \overline{D_i \zeta}, \quad \nabla_i^* \bar{\zeta} = \overline{D_i^* \zeta} \quad \text{and, thus} \quad \nabla^* \nabla \bar{\zeta} = \overline{D^* D \zeta}.$$

### 3 Green's function representation

In the present notes we study the corrector problem in probability, i. e. the following degenerate elliptic equation

$$(9) \quad D^* \mathbf{a}(0)(e + D\phi) = 0 \quad \text{in } L^2(\Omega).$$

Here and below,  $\mathbf{a}(0)$  denotes the coordinate projection  $\Omega \ni \mathbf{a} \mapsto \mathbf{a}(0) \in \Omega_0$  and the ensemble  $\langle \cdot \rangle$  is assumed to be stationary. As we explained in the introduction, equation (9) is not well-posed for general ensembles  $\langle \cdot \rangle$  due to the lack of a Poincaré inequality in  $L^2(\Omega)$  associated with the horizontal derivative  $D$ . Yet, the associated parabolic equation

$$(10) \quad \begin{cases} \frac{d}{dt} u(t) + D^* \mathbf{a}(0) D u(t) = 0 & t > 0, \\ u(t = 0) = \zeta \end{cases}$$

is solvable. Both equations, (9) and (10), are related by means of the formal identity

$$(11) \quad \phi = \int_0^\infty u(t) dt,$$

where for  $u(t)$  we choose the initial condition  $\zeta = -D^* \mathbf{a}(0)e$ . This formal relation becomes rigorous as soon we have suitable decay estimates on  $u(t)$  in  $t$ .



In the course of proving such estimates, it is convenient to work in physical space and with parabolic equations of the form

$$(12) \quad \begin{cases} \partial_t u(t, x) + \nabla^* \mathbf{a}(x) \nabla u(t, x) = g(t, x) & \text{for all } t > 0, x \in \mathbb{Z}^d, \\ u(t = 0, x) = g_0(x) & \text{for all } x \in \mathbb{Z}^d. \end{cases}$$

Since the finite difference operator  $\nabla^* \mathbf{a}(x) \nabla$  shares many properties with the well-studied elliptic operator  $-\nabla \cdot \mathbf{a}(x) \nabla$  in continuum space, many tools from elliptic and parabolic regularity are available. The purpose of this section is to justify that the connection between the parabolic equation in probability space (understood in the sense of a semigroup) and the parabolic equation in physical space (as a PDE) indeed holds.

We start by recalling that the horizontal derivatives  $D$  and  $D^*$  are bounded linear operators on  $X$ , where  $X$  stands for one of the Banach spaces  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ , or  $C^0(\Omega)$ . Hence,

$$D^* \mathbf{a}(0) D : X \rightarrow X$$

defines a bounded linear operator. In the following,  $\mathcal{L}(X)$  denotes the Banach space of bounded linear operators from  $X$  to  $X$  and is equipped with the operator norm. For  $\mathcal{G} \in \mathcal{L}(X)$  the exponential  $\exp(\mathcal{G}) \in \mathcal{L}(X)$  is defined by the series  $\sum_{k=0}^{\infty} \frac{\mathcal{G}^k}{k!}$ . As a consequence

$$\mathbb{R} \ni t \mapsto \exp(-t D^* \mathbf{a}(0) D) := \sum_{k=0}^{\infty} \frac{(-t D^* \mathbf{a}(0) D)^k}{k!} \in \mathcal{L}(X)$$

is a uniformly continuous group of bounded linear operators on  $X$ .

Let  $\zeta \in X$  and consider  $u(t) := \exp(-t D^* \mathbf{a}(0) D) \zeta$ . Due to the properties of the exponential,  $u$  belongs to  $C^\infty(\mathbb{R}, X)$  and solves (10). We argue that the stationary extension  $\bar{u}$  of  $u$  solves the corresponding parabolic equation in physical space. The connection between both points of view is most transparent if (9) holds pointwise for all  $\mathbf{a} \in \Omega$ . Considering  $X = C^0(\Omega)$  is a convenient choice for that purpose.

**Lemma 1.** For  $\zeta \in C^0(\Omega)$  define  $u \in C^\infty(\mathbb{R}, C^0(\Omega))$  by

$$u(t) := \exp(-t D^* \mathbf{a}(0) D) \zeta.$$

Consider the stationary extension  $\bar{u}(t, \mathbf{a}, x) := u(t, \mathbf{a}(\cdot + x))$ . Then for all  $\mathbf{a} \in \Omega$ , the function  $(t, x) \mapsto \bar{u}(t, \mathbf{a}, x)$  belongs to  $C^\infty(\mathbb{R}, \ell^\infty(\mathbb{Z}^d))$  and is a solution to (12) with  $g(t, x) = 0$  and  $g_0(x) := \bar{\zeta}(\mathbf{a}, x)$ .

*Proof.* Since  $u$  is defined via the uniformly continuous group

$$t \mapsto \exp(-t D^* \mathbf{a}(0) D) \in \mathcal{L}(C^0(\Omega)),$$

it belongs to  $C^\infty(\mathbb{R}, C^0(\Omega))$  and satisfies the parabolic equation (10) pointwise in  $\mathbf{a}$ , i. e.

$$(13) \quad \begin{cases} \partial_t u(t, \mathbf{a}) + D^* \mathbf{a}(0) D u(t, \mathbf{a}) = 0 & \text{for all } t \in \mathbb{R}, \mathbf{a} \in \Omega, \\ u(t = 0, \mathbf{a}) = \zeta(\mathbf{a}) & \text{for all } \mathbf{a} \in \Omega. \end{cases}$$

Thus, for all  $\mathbf{a} \in \Omega$  the mapping  $(t, x) \mapsto \bar{u}(t, \mathbf{a}, x)$  belongs to  $C^\infty(\mathbb{R}, \ell^\infty(\mathbb{Z}^d))$ . Since for all  $\mathbf{a} \in \Omega$  the function  $t \mapsto u(t, \mathbf{a})$  is smooth, and because of

$$\overline{D^* \mathbf{a}(0) D u(t)}(\mathbf{a}, x) \stackrel{(8)}{=} \nabla^* \mathbf{a}(x) \nabla \bar{u}(t, \mathbf{a}, x), \quad \bar{u}(t = 0, \mathbf{a}, x) = \bar{\zeta}(\mathbf{a}, x),$$

we deduce that  $(t, x) \mapsto \bar{u}(t, \mathbf{a}, x)$  is a  $C^\infty(\mathbb{R}, \ell^\infty(\mathbb{Z}^d))$  solution to (12) for all  $\mathbf{a} \in \Omega$  as claimed.  $\square$

In the subsequent sections we frequently use the *Green's function representation formula* for solutions to (12). In order to define the parabolic Green's function notice that  $\nabla^* \mathbf{a}(\cdot) \nabla \in \mathcal{L}(\ell^1(\mathbb{Z}^d))$ . Hence the associated exponential  $\mathbb{R} \ni t \mapsto \exp(-t \nabla^* \mathbf{a}(\cdot) \nabla)$  defines a uniformly continuous group of operators on  $\ell^1(\mathbb{Z}^d)$ , and we may define the parabolic Green's function by applying that group to the Dirac function on  $\mathbb{Z}^d$ :

$$\delta(x) := \begin{cases} 1 & \text{for } x = 0, \\ 0 & \text{else.} \end{cases}$$

**Definition 1** (parabolic Green's function). *The parabolic Green's function*

$$G : \mathbb{R}_+ \times \Omega \times \mathbb{Z}^d \times \mathbb{Z}^d \rightarrow \mathbb{R}$$

is defined as follows: For all  $\mathbf{a} \in \Omega$  and  $y \in \mathbb{Z}^d$  we denote by  $(t, x) \mapsto G(t, \mathbf{a}, x, y)$  the function in  $C^\infty(\mathbb{R}, \ell^1(\mathbb{Z}^d))$  given by

$$G(t, \mathbf{a}, \cdot, y) := \exp(-t \nabla^* \mathbf{a}(\cdot) \nabla) \delta(\cdot - y).$$

By the properties of the exponential,  $(t, x) \mapsto G(t, \mathbf{a}, x, y)$  is a  $C^\infty(\mathbb{R}, \ell^1(\mathbb{Z}^d))$ -solution to (12) with right-hand side  $g = 0$  and initial condition  $g_0(x) := \delta(x - y)$ . Hence, with the parabolic Green's function at hand we may represent solutions to (12) via *Duhamel's formula*:

**Lemma 2** (Duhamel's formula). *Let  $T > 0$ ,  $\mathbf{a} \in \Omega$ ,  $g_0 \in \ell^\infty(\mathbb{Z}^d)$ ,  $g \in C([0, T], \ell^\infty(\mathbb{Z}^d))$ . Let  $u \in C^1([0, T], \ell^\infty(\mathbb{Z}^d))$  be a solution to (12). Then*

$$(14) \quad u(t, x) = \sum_{y \in \mathbb{Z}^d} G(t, \mathbf{a}, x, y) g_0(y) + \int_0^t \sum_{y \in \mathbb{Z}^d} G(t - s, \mathbf{a}, x, y) g(s, y) ds.$$

*Proof.* Fix  $t > 0$ . For all  $0 \leq s \leq t$  and all  $x \in \mathbb{Z}^d$ , we consider

$$v(s, x) = \sum_{y \in \mathbb{Z}^d} u(s, y) G(t - s, \mathbf{a}, x, y).$$

Since  $u \in C^1([0, T], \ell^\infty(\mathbb{Z}^d))$  and  $G \in C^\infty(\mathbb{R}_+, \ell^1(\mathbb{Z}^d))$ ,  $v$  is differentiable with respect to  $s \in (0, t)$ , and its derivative is given by

$$\partial_s v(s, x) = \sum_{y \in \mathbb{Z}^d} \partial_s u(s, y) G(t - s, \mathbf{a}, x, y) - \sum_{y \in \mathbb{Z}^d} u(s, y) \partial_t G(t - s, \mathbf{a}, x, y).$$

Using first the equation satisfied by  $G$  and then two discrete integrations by parts that amount to resummations (Abel transform), we are lead to

$$\partial_s v(s, x) = \sum_{y \in \mathbb{Z}^d} (\partial_s u(s, y) + \nabla^* \nabla u(s, y)) G(t - s, \mathbf{a}, x, y).$$

Combined with the defining equation for  $u$ , this turns into

$$\partial_s v(s, x) = \sum_{y \in \mathbb{Z}^d} g(s, y) G(t - s, \mathbf{a}, x, y).$$

We integrate this identity from 0 to  $t$ , and get

$$(15) \quad v(t, x) = v(0, x) + \int_0^t \sum_{y \in \mathbb{Z}^d} g(s, y) G(t - s, \mathbf{a}, x, y) ds.$$

We then turn to the definition of  $v$ . On the one hand,

$$(16) \quad v(0, x) = \sum_{y \in \mathbb{Z}^d} u(0, y) G(t, \mathbf{a}, x, y) = \sum_{y \in \mathbb{Z}^d} g_0(y) G(t, \mathbf{a}, x, y).$$

On the other hand,

$$(17) \quad v(t, x) = \sum_{y \in \mathbb{Z}^d} u(t, y) G(0, \mathbf{a}, x, y) = u(t, x).$$

The Duhamel formula follows from the combination of (15), (16), and (17).  $\square$

An immediate consequence of Duhamel's formula is that  $C^1(\mathbb{R}_+, \ell^\infty(\mathbb{Z}^d))$ -solutions to (12) are unique. Since the  $\ell^p$ -spaces are ordered, in the sense of

$$(18) \quad \|\cdot\|_{\ell^q(\mathbb{Z}^d)} \leq \|\cdot\|_{\ell^p(\mathbb{Z}^d)} \quad \text{for all } 1 \leq p \leq q \leq \infty,$$

we may replace in Duhamel's formula the space  $\ell^\infty(\mathbb{Z}^d)$  by any  $\ell^p(\mathbb{Z}^d)$ ,  $1 \leq p < \infty$ . In particular, the uniqueness property for (12) extends to solutions of class  $C^1(\mathbb{R}_+, \ell^p(\mathbb{Z}^d))$ . In particular, we may characterize  $G$  by (12): For all  $\mathbf{a} \in \Omega$  and  $y \in \mathbb{Z}^d$  the function  $(t, x) \mapsto G(t, \mathbf{a}, x, y)$  is the unique solution in  $C^1(\mathbb{R}_+, \ell^p(\mathbb{Z}^d))$  to (12) with  $g = 0$ ,  $g_0 := \delta(\cdot - y)$  and arbitrary  $1 \leq p \leq \infty$ .

Our results heavily rely on estimates for the Green's function that are based on elliptic and parabolic regularity theory. In particular, we require the Green's function to be non-negative. In the continuum setting this follows from the maximum principle. However, in the discrete case considered here, i. e. for the lattice graph  $\mathbb{Z}^d$  with nearest neighbor edges, the maximum principle is **not** valid for general symmetric, elliptic coefficients. Hence, in that case the Green's function might not be non-negative and several estimates that hold in the continuum case break down. For this reason we assume that the coefficients are diagonal, see (5). The following lemma shows that for diagonal coefficients we have a weak maximum principle:

**Lemma 3** (Weak maximum principle). *Let  $\mathbf{a} \in \Omega$ . Consider  $u \in C^1(\mathbb{R}_+, \ell^p(\mathbb{Z}^d))$ ,  $1 \leq p < \infty$ , with*

$$\partial_t u(t, x) + \nabla_x^* \mathbf{a}(x) \nabla_x u(t, x) \leq 0 \quad \text{for all } t > 0, x \in \mathbb{Z}^d.$$

*Then we have*

$$(19) \quad \sup_{(t,x) \in \mathbb{R}_+ \times \mathbb{Z}^d} u(t, x) = \sup_{x \in \mathbb{Z}^d} u(0, x).$$

*Proof.* First, we note that the  $p$ -summability of  $u(t, \cdot)$  yields

$$(20) \quad \lim_{R \uparrow \infty} \sup_{|x| \geq R} |u(t, x)| = 0 \quad \text{for all } t \geq 0,$$

from which we learn that both sides in (19) are non-negative. Second, since on the l. h. s. of (19) we take the supremum over a larger set than on the r. h. s., the inequality “ $\geq$ ” is trivial. Hence, we only need to show that  $\sup_{(t,x) \in [0, T] \times \mathbb{Z}^d} u(t, x) \leq \sup_{x \in \mathbb{Z}^d} u(0, x)$  for all  $T \gg 1$ . Fix  $T \gg 1$  and consider for  $\varepsilon > 0$  the modified function  $u_\varepsilon(t, x) := u(t, x) - \varepsilon t$  which satisfies the strict inequality

$$(21) \quad \partial_t u_\varepsilon(t, x) + \nabla_x^* \mathbf{a}(x) \nabla_x u_\varepsilon(t, x) < 0.$$

Because  $\lim_{\varepsilon \downarrow 0} \sup_{[0, T] \times \mathbb{Z}^d} u_\varepsilon = \sup_{[0, T] \times \mathbb{Z}^d} u$ , we only need to prove that

$$(22) \quad \sup_{(t,x) \in [0, T] \times \mathbb{Z}^d} u_\varepsilon(t, x) \leq \sup_{\mathbb{Z}^d} u(0, \cdot) \quad \text{for all } 0 < \varepsilon \ll 1.$$

Since the r. h. s. is non-negative, we may assume without loss of generality that

$$(23) \quad \sup_{(t,x) \in [0, T] \times \mathbb{Z}^d} u_\varepsilon(t, x) > 0.$$

As we will prove below the supremum on the l. h. s. of (22) is attained, i. e.

$$(24) \quad \text{there exists } (t_0, x_0) \in [0, T] \times \mathbb{Z}^d \text{ such that } u_\varepsilon(t_0, x_0) = \sup_{(t,x) \in [0, T] \times \mathbb{Z}^d} u_\varepsilon(t, x).$$

From (24) we obtain (22) by showing  $t_0 = 0$  via contradiction: If  $t_0 > 0$ , then, since  $(t_0, x_0)$  is a maximum, we have  $\partial_t u_\varepsilon(t_0, x_0) \geq 0$  and, by diagonality of  $\mathbf{a}$ ,

$$\begin{aligned} \nabla^* \mathbf{a}(x_0) \nabla u_\varepsilon(t_0, x_0) &= \sum_{i=1}^d \mathbf{a}_{ii}(x_0 - e_i) (u_\varepsilon(t_0, x_0) - u_\varepsilon(t_0, x_0 - e_i)) \\ &\quad + \sum_{i=1}^d \mathbf{a}_{ii}(x_0) (u_\varepsilon(t_0, x_0) - u_\varepsilon(t_0, x_0 + e_i)) \\ &\geq 0. \end{aligned}$$

Consequently,  $\partial_t u_\varepsilon(t_0, x_0) + \nabla^* \mathbf{a}(x_0) \nabla u_\varepsilon(t_0, x_0) \geq 0$  in contradiction to (21). To complete the argument it remains to prove (24), which in view of (23) follows from

$$(25) \quad \lim_{R \uparrow \infty} \sup_{t \in [0, T]} \sup_{|x| \geq R} |u(t, x)| = 0.$$

Here comes the argument for (25). Since  $u \in C^1(\mathbb{R}_+, \ell^p(\mathbb{Z}^d))$  for some  $p < \infty$ , we deduce that  $m_R(t) := \sup_{|x| \geq R} |u(t, x)|$  is Lipschitz continuous with a Lipschitz constant that can be chosen independent of  $R$ . Now, (25) follows from (20).  $\square$

**Corollary 1.** *For all  $s, t \geq 0$ ,  $x, y, z \in \mathbb{Z}^d$  the Green's function  $G(t, \mathbf{a}, x, y)$*

- *is non-negative, i. e.*

$$(26a) \quad G(t, \mathbf{a}, x, y) \geq 0,$$

- *is stationary in the sense of*

$$(26b) \quad G(t, \mathbf{a}, x + z, y + z) = G(t, \mathbf{a}(\cdot + z), x, y),$$

- *conserves mass in the sense of*

$$(26c) \quad \sum_{x \in \mathbb{Z}^d} G(t, \mathbf{a}, x, y) = 1,$$

- *satisfies the semigroup property, i. e.*

$$(26d) \quad \sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z) G(s, \mathbf{a}, z, y) = G(t + s, \mathbf{a}, x, y),$$

- *is symmetric, i. e.*

$$(26e) \quad G(t, \mathbf{a}, x, y) = G(t, \mathbf{a}, y, x).$$

*In the constant-coefficient case,  $G(t, x) = G(t, \mathbf{id}, x, 0)$  satisfies in addition the estimate*

$$(26f) \quad G(t, 0) \geq G(t, x).$$

*Proof. Step 1.* Arguments for (26a) – (26d).

Since  $\nabla^* \mathbf{a}(\cdot) \nabla$  is a bounded operator on  $\ell^1(\mathbb{Z}^d)$ , the Green's function  $G$  is characterized as the unique solution to

$$(27) \quad \begin{cases} \partial_t G(t, \mathbf{a}, x, y) + \nabla_x^* \mathbf{a}(x) \nabla_x G(t, \mathbf{a}, x, y) = 0 & \text{for all } t > 0, x, y \in \mathbb{Z}^d, \\ G(t = 0, \mathbf{a}, x, y) = \delta(x - y) & \text{for all } x, y \in \mathbb{Z}^d. \end{cases}$$

By applying Lemma 3 to  $-G$  we get the non-negativity of  $G$ . The stationarity (26b) directly follows from the uniqueness. The semigroup property (26d) follows from Duhamel's formula, cf. Lemma 2. Summation in  $x \in \mathbb{Z}^d$  of the first line in (27) yields  $\partial_t \sum_{x \in \mathbb{Z}^d} G(t, \mathbf{a}, x, y) = 0$ ; hence,  $\sum_{x \in \mathbb{Z}^d} G(t, \mathbf{a}, x, y) = \sum_{x \in \mathbb{Z}^d} G(0, \mathbf{a}, x, y) = 1$  and (26c) follows.

**Step 2.** Argument for the symmetry property (26e).

For  $T < 0$  consider the function

$$v(t, y, y') := \sum_{x \in \mathbb{Z}^d} G(t, x, y) G(T - t, x, y').$$

A direct calculation shows that  $\partial_t v(t, y, y') = 0$ , and thus  $v$  is constant in  $t$ . Hence,

$$\begin{aligned} G(T, y, y') &= \sum_{x \in \mathbb{Z}^d} \delta(x - y) G(T, x, y') \\ &= v(0, y, y') = v(T, y, y') \\ &= \sum_{x \in \mathbb{Z}^d} G(T, x, y) \delta(x - y') = G(T, y', y). \end{aligned}$$

Since  $y, y'$  and  $T$  are arbitrary, (26e) follows.

**Step 3.** Argument for (26f).

Let  $G_1(t, x)$  denote the one-dimensional parabolic constant-coefficient Green's function. We have  $G(t, x) = \prod_{i=1}^d G_1(t, x_i)$  as can be checked directly. Since  $G_1$  is non-negative and symmetric, it suffices to prove that

$$G_1(t, z) \leq G_1(t, 0) \quad \text{for all } z \in \mathbb{Z}, z > 0.$$

In fact we prove the stronger monotonicity property

$$(28) \quad G_1(t, z + 1) - G_1(t, z) \leq 0 \quad \text{for all } z \in \mathbb{Z}, z \geq 0.$$

To that end set  $u_\varepsilon(t, z) := G_1(t, z + 1) - G_1(t, z) - \varepsilon t$ . We need to show that for all  $T > 0$  and  $0 < \varepsilon \ll 1$  we have

$$\sup_{t \in [0, T], z \geq 0} u_\varepsilon(t, z) \leq 0.$$

We give an indirect argument assuming that the l. h. s. is strictly positive. Note that  $u_\varepsilon$  satisfies

$$(29) \quad \begin{cases} \partial_t u_\varepsilon(t, z) + \nabla_z^* \nabla_z u_\varepsilon(t, z) < 0 \\ u_\varepsilon(t = 0, z) = \delta(z + 1) - \delta(z). \end{cases}$$

As in the proof of the weak maximum principle, cf. Lemma 3, we infer from the integrability of  $G_1(t, z + 1) - G_1(t, z)$  and (29) that a strictly positive maximum of  $u_\varepsilon$  in  $[0, T] \times \{z \geq 0\}$  is attained for some  $(t_0, z_0)$ , where either  $t_0 = 0$  or  $z_0 = 0$ . We show that both options lead to a contradiction. Indeed, if the maximum is attained for  $t_0 > 0$  and  $z_0 = 0$ , then  $\partial_t u_\varepsilon(t_0, 0) \geq 0$ , so that

$$\begin{aligned} 0 > \nabla_z^* \nabla_z u_\varepsilon(t_0, 0) &= 2u_\varepsilon(t_0, 0) - u_\varepsilon(t_0, 1) - u_\varepsilon(t_0, -1), \\ &\stackrel{(26e)}{=} 3u_\varepsilon(t_0, 0) - u_\varepsilon(t_0, 1) \geq 0, \end{aligned}$$

which contradicts (29). On the other hand if  $u_\varepsilon$  attains a strictly positive maximum at  $t_0 = 0$ , then we obtain a contradiction with  $u_\varepsilon(t_0 = 0, z) = -\delta(z) \leq 0$ .  $\square$

We also need the  $L$ -periodic parabolic Green's function:

**Definition 2.** For  $L \in \mathbb{N}$  we define  $G_L : \mathbb{R}_+ \times \Omega \times \mathbb{Z}^d \times \mathbb{Z}^d \rightarrow \mathbb{R}$  as follows: For all  $\mathbf{a} \in \Omega$  and  $y \in \mathbb{Z}^d$  we denote by  $(t, x) \mapsto G_L(t, \mathbf{a}, x, y)$  the function in  $C^\infty(\mathbb{R}, \ell^\infty(\mathbb{Z}^d))$  given by

$$G_L(t, \mathbf{a}, \cdot, y) := \exp(-t \nabla^* \mathbf{a}(\cdot) \nabla) \delta_L(\cdot - y)$$

where

$$\delta_L(x) := \sum_{z \in \mathbb{Z}^d} \delta(x + Lz).$$

denotes the Dirac function on the discrete torus of size  $L$ .

**Remark 1.** When  $\mathbf{a}$  is  $L$ -periodic, then  $G_L(t, \mathbf{a}, x, y)$  is  $L$ -periodic in  $x$  and characterized as the unique  $C^1(\mathbb{R}, \ell^\infty(\mathbb{Z}^d))$ -solution to (12) with  $g = 0$  and  $g_0 = \delta_L(\cdot - y)$ . Therefore we call  $G_L$  the  $L$ -periodic, parabolic Green's function.

In the  $L$ -periodic setting, the properties gathered in Corollary 1 take the following form:

**Corollary 2.** Let  $L \in \mathbb{N}$  and  $\mathbf{a} \in \Omega_L$ . In addition to (26a), (26b) and (26e) we have (for all  $s, t \geq 0$ ,  $x, y, z, z' \in \mathbb{Z}^d$ )

- $G_L(t, \mathbf{a}, x, y)$  is  $L$ -periodic in  $x$  and  $y$  in the sense of

$$(30a) \quad G_L(t, \mathbf{a}, x + Lz, y + Lz') = G_L(t, \mathbf{a}, x, y),$$

- $G_L(t, \mathbf{a}, x, y)$  conserves mass in the sense of

$$(30b) \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} G_L(t, \mathbf{a}, x, y) = 1,$$

- $G_L(t, \mathbf{a}, x, y)$  satisfies the semigroup property

$$(30c) \quad \sum_{z \in ([0, L) \cap \mathbb{Z})^d} G_L(t, \mathbf{a}, x, z) G_L(s, \mathbf{a}, z, y) = G_L(t + s, \mathbf{a}, x, y).$$

*Proof.* Since  $\nabla^* \mathbf{a}(\cdot) \nabla$  is a bounded operator on  $\ell^\infty(\mathbb{Z}^d)$ , the function  $(t, x) \mapsto G_L(t, \mathbf{a}, x, y)$  is a solution in  $C^1(\mathbb{R}_+, \ell^\infty(\mathbb{Z}^d))$  to

$$(31) \quad \begin{cases} \partial_t G_L(t, \mathbf{a}, x, y) + \nabla_x^* \mathbf{a}(x) \nabla_x G_L(t, \mathbf{a}, x, y) = 0 & \text{for all } t > 0, x \in \mathbb{Z}^d, \\ G_L(t = 0, \mathbf{a}, x, y) = \delta_L(x - y) & \text{for all } x \in \mathbb{Z}^d, \end{cases}$$

so that we get by Duhamel's formula, cf. Lemma (2), the representation

$$(32) \quad G_L(t, \mathbf{a}, x, y) = \sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z) \delta_L(z - y) = \sum_{z' \in \mathbb{Z}^d} G(t, \mathbf{a}, x, y + Lz').$$

With (32) at hand we shall obtain the claimed properties of  $G_L$  from the corresponding properties of  $G$ . In particular, (26a), (26b) and (26e) are directly inherited from  $G$ . We prove (30a): The  $L$ -periodicity of  $G(t, \mathbf{a}, x, y)$  in  $y$  follows from the  $L$ -periodicity of  $\delta_L$  and (32). Since the coefficients  $\mathbf{a}(\cdot)$  and the initial data  $\delta_L$  are  $L$ -periodic, the shifted function  $(t, x) \mapsto G(t, \mathbf{a}, x + Lz, y)$ ,  $z \in \mathbb{Z}^d$ , solves (31), and thus the  $L$ -periodicity of  $G(t, \mathbf{a}, x, y)$  in  $x$  follows from the uniqueness of the solution to (31).

Next, we prove (30b). By the integration by parts formula for  $L$ -periodic functions we have

$$\frac{d}{dt} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} G(t, \mathbf{a}, x, y) = - \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_x^* \mathbf{a}(x) \nabla_x G(t, \mathbf{a}, x, y) = 0,$$

and (30b) follows from  $\sum_{x \in ([0, L] \cap \mathbb{Z})^d} G(t = 0, \mathbf{a}, x, y) = 1$ . Next we deduce the semigroup property (30c) from (32) and (26d):

$$\begin{aligned}
G_L(t + s, \mathbf{a}, x, y) &\stackrel{(32)}{=} \sum_{z \in \mathbb{Z}^d} G(t + s, \mathbf{a}, x, z) \delta_L(z - y) \\
&\stackrel{(26d)}{=} \sum_{z' \in \mathbb{Z}^d} \sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z') G(s, \mathbf{a}, z', z) \delta_L(z - y) \\
&\stackrel{(32)}{=} \sum_{z' \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z') G_L(s, \mathbf{a}, z', y) \\
&\stackrel{(30a)}{=} \sum_{z'' \in \mathbb{Z}^d} \sum_{z \in ([0, L] \cap \mathbb{Z})^d} G(t, \mathbf{a}, x, z + Lz'') G_L(s, \mathbf{a}, z, y) \\
&\stackrel{(32)}{=} \sum_{z \in ([0, L] \cap \mathbb{Z})^d} G_L(t, \mathbf{a}, x, z) G_L(s, \mathbf{a}, z, y).
\end{aligned}$$

□

The  $L$ -periodic parabolic Green's function is characterized by the spatial parabolic equation on the integer torus of size  $L$ :

**Corollary 3.** *Let  $L \in \mathbb{N}$  and  $\mathbf{a} \in \Omega_L$ . Then for all  $y \in \mathbb{Z}^d$ ,  $(t, x) \mapsto G_L(t, \mathbf{a}, x, y)$  is characterized as the unique solution in  $C^1(\mathbb{R}_+, \ell_x^\infty(\mathbb{Z}^d))$  to*

$$(33) \quad \left\{ \begin{array}{l} \partial_t G_L(t, \mathbf{a}, x, y) + \nabla_x^* \mathbf{a}(x) \nabla_x G_L(t, \mathbf{a}, x, y) = 0 \\ \qquad \qquad \qquad \text{for all } t > 0, x \in ([0, L] \cap \mathbb{Z})^d, \\ G_L(t = 0, \mathbf{a}, x, y) = \delta_L(x - y) \\ \qquad \qquad \qquad \text{for all } x \in ([0, L] \cap \mathbb{Z})^d, \\ G_L(t, \mathbf{a}, x, y) \text{ is } L\text{-periodic in } x. \end{array} \right.$$

*Proof.* We already know from the proof of Corollary 2 that  $G_L$  is uniquely determined by the parabolic equation (31). In particular, since the solution is (31) is  $L$ -periodic, it also solves (33). Hence, it suffices to argue that any solution to (33) solves (31). But this immediately follows from (33) and the periodicity assumption. □

**Remark 2.** *In the constant-coefficient case the Green's function is shift invariant, i. e.  $G(t, \mathbf{a}, x, y) = G(t, \mathbf{a}, x - y, 0)$ . When no confusion occurs we shall write*

$$G(t, x) := G(t, \mathbf{id}, x, 0) \quad \text{and} \quad G_L(t, x) := G_L(t, \mathbf{id}, x, 0).$$

We call  $G(t, x)$  and  $G_L(t, x)$  the whole space and  $L$ -periodic constant-coefficient Green's functions.

We conclude this section with some estimates for the constant-coefficient parabolic Green's function  $G(t, x)$ .

**Lemma 4.** *The constant-coefficient parabolic Green's function  $G(t, x)$  satisfies*

$$(34) \quad (t + 1)^{-\frac{d}{2}} \lesssim \sum_{x \in \mathbb{Z}^d} G^2(t, x) \lesssim (t + 1)^{-\frac{d}{2}},$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depend on the dimension  $d$ .

*Proof of Lemma 4.* The upper estimate can be easily obtained by a Nash-Aronson type argument. However, we do not present this simple argument here, since in Section 7 we shall establish stronger estimates that contain the upper estimate of (34) as a special case. In fact, we have for  $\alpha < \infty$  the spatially weighted estimate

$$(35) \quad \sum_{x \in \mathbb{Z}} \left( \omega^{\frac{\alpha}{2}}(t, x) G(t, x) \right)^2 \lesssim (t+1)^{-\frac{1}{2}}, \quad \omega(t, x) := \frac{|x|^2}{t+1} + 1$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\alpha$  and  $d$ . We establish that estimate along the lines of the proof of Theorem 3 (b) in the more intricate case of the variable-coefficient,  $L$ -periodic parabolic Green's function. In particular, see (173) for the  $L$ -periodic, continuum version of (35), Section 7.2 for the passage from the continuum to the discrete Green's function, and Section 7.3 for the passage from the  $L$ -periodic to the whole space estimate. The claimed upper estimates is obtained from (35) for  $\alpha = 0$ .

Next, we prove the lower estimate. For  $t > 0$  let  $R(t) > 0$  denote a radius with the property

$$(36) \quad \sum_{\substack{x \in \mathbb{Z}^d \\ |x| > R(t)}} G(t, x) \leq \frac{1}{2}.$$

Because of (26c), such radius indeed exists. By Jensen's inequality we have

$$\frac{1}{R^d(t)} \left( \sum_{|x| \leq R(t)} G(t, x) \right)^2 \leq \sum_{|x| \leq R(t)} G^2(t, x),$$

and thus

$$\begin{aligned} \sum_{x \in \mathbb{Z}^d} G^2(t, x) &\geq \sum_{|x| \leq R(t)} G^2(t, x) \geq \frac{1}{R^d(t)} \left( \sum_{|x| \leq R(t)} G(t, x) \right)^2 = \frac{1}{R^d(t)} \left( 1 - \sum_{|x| > R(t)} G(t, x) \right)^2 \\ &\stackrel{(36)}{\geq} \frac{1}{4R^d(t)}. \end{aligned}$$

It remains to argue that we can choose  $R(t)$  such that

$$(37) \quad R(t) \leq C\sqrt{t+1}$$

with  $C \lesssim 1$  to be chosen below. It is easy to see that (35) implies (37): On the one hand we have

$$\begin{aligned} \sum_{|x| > C\sqrt{t+1}} G(t, x) &= \sum_{|x| > C\sqrt{t+1}} G(t, x) \omega^{\frac{\alpha}{2}}(t, x) \omega^{-\frac{\alpha}{2}}(t, x) \\ &\leq \left( \sum_{|x| > C\sqrt{t+1}} \left( G(t, x) \omega^{\frac{\alpha}{2}}(t, x) \right)^2 \right)^{\frac{1}{2}} \left( \sum_{|x| > C\sqrt{t+1}} \omega^{-\alpha}(t, x) \right)^{\frac{1}{2}} \\ &\stackrel{(35)}{\lesssim} (t+1)^{-\frac{d}{4}} \left( \sum_{|x| > C\sqrt{t+1}} \omega^{-\alpha}(t, x) \right)^{\frac{1}{2}}. \end{aligned}$$

On the other hand, for  $\alpha > \frac{d}{2}$  we have

$$\begin{aligned} \sum_{|x| > C\sqrt{t+1}} \omega^{-\alpha}(t, x) &\leq \sum_{|x| > C\sqrt{t+1}} \left( \frac{t+1}{|x|^2} \right)^\alpha \\ &\lesssim \int_{\{|x| > C\sqrt{t+1}\}} \left( \frac{t+1}{|x|^2} \right)^\alpha dx = (t+1)^{\frac{d}{2}} \int_{|z| > C} |z|^{-2\alpha} dz. \end{aligned}$$



Hence, we obtain  $\sum_{|x|>C\sqrt{t+1}} G(t, x) \lesssim \int_{|z|>C} |z|^{-2\alpha} dz$  and (37) follows, since  $\int_{|z|>C} |z|^{-2\alpha} dz \rightarrow 0$  as  $C \uparrow \infty$ .  $\square$

**Corollary 4.** *The  $L$ -periodic constant-coefficient Green's function  $G_L(t, x)$  satisfies*

$$(38) \quad \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |G_L(t, x) - L^{-d}|^2 \lesssim \begin{cases} (t+1)^{-\frac{d}{2}} & t \leq L^2, \\ L^{-d} \exp(-c_0 \frac{t}{L^2}) & t \geq L^2. \end{cases}$$

Above  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $d$ , and  $c_0$  denotes a constant that only depends on  $d$ .

*Proof of Corollary 4.* The claimed estimate follows from the estimates (173) and (174) that we establish in Section 7.  $\square$

## 4 Quantification of ergodicity via spectral gap

In this section we quantify ergodicity by the spectral gap of a Glauber dynamics on coefficient fields. We recall the notion of ergodicity and introduce a characterization that is based on the decay of the semigroup generated by  $D^*D$  in  $L^2(\Omega)$ . In the first part of this section we introduce a spectral gap estimate (SG) which is associated with the ‘‘heat bath’’ Glauber dynamics that appears in statistical mechanics. As the main result of this section we prove that the decay of the semigroup can be quantified by appealing to (SG). As a corollary we find that (infinite) ensembles that satisfy (SG) are ergodic. In the second part of this section we discuss  $L$ -periodic ensembles. We apply the results to independent and identically distributed coefficient fields.

We start with preliminaries and basic facts of ergodic theory. In the following a central role is played by the shift operator  $T_x : \Omega \rightarrow \Omega$ ,  $\mathbf{a} \mapsto \mathbf{a}(\cdot + x)$  defined for all shifts  $x \in \mathbb{Z}^d$ . By definition, an ensemble is stationary if and only if  $T_x$  is measure preserving for all shifts  $x \in \mathbb{Z}^d$ .

**Definition 3** (shift invariance). *Let  $\langle \cdot \rangle$  be stationary. A measurable set  $A \subset \Omega$  is called shift invariant if*

$$A = T_x A := \{T_x \mathbf{a} : \mathbf{a} \in A\} \quad \text{for all } x \in \mathbb{Z}^d.$$

A measurable function  $\zeta : \Omega \rightarrow \mathbb{R}$  is called shift invariant, if

$$\zeta(T_x \mathbf{a}) = \zeta(\mathbf{a}) \quad \text{for all } x \in \mathbb{Z}^d \text{ and } \langle \cdot \rangle\text{-almost every } \mathbf{a} \in \Omega.$$

We denote by  $\text{INV}^2(\Omega)$  the set of shift invariant functions in  $L^2(\Omega)$ , and by  $L_0^2(\Omega)$  the orthogonal complement of  $\text{INV}^2(\Omega)$  in  $L^2(\Omega)$ .

Let us recall the pointwise multiparameter ergodic theorem:

**Theorem 1** (ergodic theorem). *Let  $\langle \cdot \rangle$  be stationary and  $\zeta \in L^2(\Omega)$ . Then we have*

$$\lim_{R \rightarrow \infty} \frac{1}{R^d} \sum_{x \in ([0, R] \cap \mathbb{Z})^d} \zeta \circ T_x = \langle \zeta | \text{INV}^2(\Omega) \rangle \quad \langle \cdot \rangle\text{-almost surely,}$$

where  $\langle \zeta | \text{INV}^2(\Omega) \rangle$  denotes the  $L^2(\Omega)$  projection of  $\zeta$  onto  $\text{INV}^2(\Omega)$ .

The one-parameter version of the theorem goes back to Birkhoff [4]. For the multiparameter version we refer to [1, Theorem 2.4], see also [35]. Originally, the ergodic theorem is stated for  $\zeta \in L^1(\Omega)$ . In that case  $\langle \zeta | \text{INV}^2(\Omega) \rangle$  has to be replaced by the conditional expectation of  $\zeta$  where we condition w. r. t. the smallest  $\sigma$ -algebra that contains all shift invariant measurable subsets of  $\Omega$ .

**Definition 4** (ergodicity). *A stationary ensemble  $\langle \cdot \rangle$  is called ergodic, if*

$$(39) \quad \text{INV}^2(\Omega) = \{\zeta \in L^2(\Omega) : \zeta = \langle \zeta \rangle \text{ almost surely}\}.$$

In the following we relate ergodicity to the properties of the operator  $D^*D$  and of the associated semigroup (see Section 3). Ergodicity can be characterized by the properties of the kernel of  $D^*D$  and by the decay of the semigroup associated with  $D^*D$ . We start with some simple, qualitative observations.

**Lemma 5.** *Let  $\langle \cdot \rangle$  be stationary. We have*

$$(40) \quad \begin{aligned} \text{INV}^2(\Omega) &= \left\{ \zeta \in L^2(\Omega) : D\zeta = 0 \quad \langle \cdot \rangle\text{-almost surely} \right\} \\ &= \left\{ \zeta \in L^2(\Omega) : D^*D\zeta = 0 \quad \langle \cdot \rangle\text{-almost surely} \right\}, \end{aligned}$$

$$(41) \quad L_0^2(\Omega) = \overline{\{D^*\xi : \xi \in L^2(\Omega)^d\}}^{L^2(\Omega)}.$$

*Proof of Lemma 5. Step 1.* Proof of (40)

The relations

$$\begin{aligned} \text{INV}^2(\Omega) &= \left\{ \zeta \in L^2(\Omega) : D\zeta = 0 \quad \langle \cdot \rangle\text{-almost surely} \right\} \\ &\subset \left\{ \zeta \in L^2(\Omega) : D^*D\zeta = 0 \quad \langle \cdot \rangle\text{-almost surely} \right\} \end{aligned}$$

are trivial. It remains to show that the set on the right-hand side is contained in  $\text{INV}^2(\Omega)$ . Let  $\zeta \in L^2(\Omega)$  satisfy  $D^*D\zeta = 0$ . Multiplication with  $\zeta$ , taking the expected value and integration by parts yields

$$0 = \langle \zeta D^*D\zeta \rangle = \langle |D\zeta|^2 \rangle.$$

Hence,  $D\zeta = 0$  almost surely, and thus  $\zeta$  is shift invariant.

**Step 2.** Proof of (41)

Set  $X := \{D^*\xi : \xi \in L^2(\Omega)^d\}$ . Since  $L^2(\Omega) = \text{INV}^2(\Omega) \oplus L_0^2(\Omega)$  is an orthogonal decomposition, it suffices to prove

$$(42) \quad X \subset L_0^2(\Omega) \quad \text{and} \quad X^\perp \subset \text{INV}^2(\Omega).$$

For the first inclusion it suffices to show

$$\langle D^*\xi \varphi \rangle = 0 \quad \text{for all } \xi \in L^2(\Omega)^d, \varphi \in \text{INV}^2(\Omega).$$

Indeed, this follows by integration by parts and the fact that  $D\varphi = 0$  almost surely for  $\varphi \in \text{INV}^2(\Omega)$ . We prove the second inclusion in (42). Let  $\varphi \in X^\perp$ , then  $0 = \langle \varphi D^*\xi \rangle = \langle D\varphi \cdot \xi \rangle$  for all  $\xi \in L^2(\Omega)^d$ . We deduce that  $D\varphi = 0$  almost surely, and thus  $\varphi$  is shift invariant.  $\square$

Since  $D^*D$  is bounded, symmetric and non-negative, the random variable  $u(t) := \exp(-tD^*D)\zeta$  converges as  $t \uparrow \infty$  to the projection of  $\zeta$  onto the kernel of  $D^*D$ , which by Lemma 5 coincides with  $\text{INV}^2(\Omega)$ :

**Lemma 6.** *Let  $\langle \cdot \rangle$  be stationary. For  $\zeta \in L^2(\Omega)$  we have*

$$\lim_{t \uparrow \infty} \langle |\exp(-tD^*D)\zeta - \zeta'|^2 \rangle = 0$$

where  $\zeta = \zeta' + \zeta''$  with  $\zeta' \in \text{INV}^2(\Omega)$  and  $\zeta'' \in L_0^2(\Omega)$ .

*Proof of Lemma 6. Step 1.* Orthogonal decomposition of  $\zeta$ .

By Lemma 5 we have  $D^*D\zeta' = 0$  and thus  $\exp(-tD^*D)\zeta' = \zeta'$ . Moreover,  $\exp(-tD^*D)\zeta'' \in L_0^2(\Omega)$  since  $\zeta'' \in L_0^2(\Omega)$ . Hence,  $\exp(-tD^*D)\zeta = \zeta' + \exp(-tD^*D)\zeta''$  is an orthogonal decomposition for all  $t \in \mathbb{R}$ , and thus  $\langle |\exp(-tD^*D)\zeta - \zeta'|^2 \rangle = \langle |\exp(-tD^*D)\zeta''|^2 \rangle$ . As a consequence, it suffices to prove that

$$(43) \quad \forall \chi \in L_0^2(\Omega) : \quad \lim_{t \uparrow \infty} \langle |\exp(-tD^*D)\chi|^2 \rangle = 0.$$

**Step 2.** A priori estimates.

Let  $\zeta \in L^2(\Omega)$  and set  $u(t) := \exp(-tD^*D)\zeta$ . We claim that

$$(44) \quad \forall t \geq 0 : \langle |u(t)|^2 \rangle \leq \langle \zeta^2 \rangle,$$

$$(45) \quad \lim_{t \uparrow \infty} \langle |Du(t)|^2 \rangle = 0.$$

For the argument recall that  $u(t)$  solves the parabolic equation (10) with initial condition  $\zeta$ . By testing the equation with  $u(t)$  and  $Du(t)$ , we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \langle u(t)^2 \rangle &= \left\langle \frac{d}{dt} u(t) u(t) \right\rangle \stackrel{(10)}{=} -\langle |Du(t)|^2 \rangle \leq 0, \\ \frac{1}{2} \frac{d}{dt} \langle |Du(t)|^2 \rangle &= \left\langle \frac{d}{dt} Du(t) \cdot Du(t) \right\rangle = \left\langle \frac{d}{dt} u(t) D^* Du(t) \right\rangle \\ &\stackrel{(10)}{=} -\langle |D^* Du(t)|^2 \rangle \leq 0. \end{aligned}$$

Integration of the first identity yields (44) and  $\int_0^\infty \langle |Du(t)|^2 \rangle dt \leq \langle \zeta^2 \rangle < \infty$ . By the second estimate  $t \mapsto \langle |Du(t)|^2 \rangle$  is non-increasing so that the latter yields (45).

**Step 3.** Proof of (43).

Let  $\nu > 0$ . By Lemma 5 there exist  $\xi^\nu = (\xi_1^\nu, \dots, \xi_d^\nu) \in L^2(\Omega)^d$  with  $\langle |\chi - D^* \xi^\nu|^2 \rangle \leq \nu$ . We claim that

$$(46) \quad \forall t \in \mathbb{R}_+ \quad \langle |\exp(-tD^*D)(\chi - D^* \xi^\nu)|^2 \rangle \leq \nu$$

$$(47) \quad \lim_{t \uparrow \infty} \langle |\exp(-tD^*D)D^* \xi^\nu|^2 \rangle = 0.$$

Indeed, (46) follows from (44) and the fact that  $D^* \xi^\nu$  approximates  $\chi$ . Estimate (47) can be seen as follows. Since the shift operators  $T_{e_1}, \dots, T_{e_d}$  commute, we get

$$\exp(-tD^*D)D^* \xi^\nu = \sum_{i=1}^d D_i^* \exp(-tD^*D) \xi_i^\nu.$$

Hence,

$$\begin{aligned} \langle |\exp(-tD^*D)D^* \xi^\nu|^2 \rangle &\leq d \sum_{i=1}^d \langle |D_i^* \exp(-tD^*D) \xi_i^\nu|^2 \rangle \\ &\stackrel{\text{stationarity}}{=} d \sum_{i=1}^d \langle |D_i \exp(-tD^*D) \xi_i^\nu|^2 \rangle. \end{aligned}$$

Now, (47) follows from (45) applied with  $\zeta = \xi_i^\nu$ ,  $i = 1, \dots, d$ .

Finally, we argue that (46) and (47) imply (43). Indeed, for all  $\nu > 0$  we have

$$\begin{aligned} \frac{1}{2} \langle |\exp(-tD^*D)\chi|^2 \rangle &\leq \langle |\exp(-tD^*D)(\chi - D^* \xi^\nu)|^2 \rangle \\ &\quad + \langle |\exp(-tD^*D)D^* \xi^\nu|^2 \rangle \\ &\leq \nu + \langle |\exp(-tD^*D)D^* \xi^\nu|^2 \rangle. \end{aligned}$$

Passing first to the limit  $t \uparrow \infty$  and secondly to the limit  $\nu \downarrow 0$  yields (43).  $\square$

By Lemma 5 the kernel of  $D^*D$  is trivial (in the sense that it is one dimensional and only contains the constants) if and only if  $\langle \cdot \rangle$  is ergodic. We obtain the following characterization:

**Corollary 5** (characterization of ergodicity). *Let  $\langle \cdot \rangle$  be stationary. Then the following statements are equivalent:*

(a)  $\langle \cdot \rangle$  is ergodic

(b) for every  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$  we have

$$(48) \quad \lim_{t \uparrow \infty} \langle |\exp(-tD^*D)\zeta|^2 \rangle = 0.$$

The goal of the following discussion is to show that the decay in (48) can be quantified, provided  $\langle \cdot \rangle$  satisfies a spectral gap estimate w. r. t. the Glauber dynamics on coefficients fields. To make this precise we introduce *vertical derivatives* of random variables.

**Definition 5** (vertical derivative). *Let  $\langle \cdot \rangle$  be an arbitrary ensemble. For  $\zeta \in L^2(\Omega)$  and  $y \in \mathbb{Z}^d$  we define*

$$\langle \zeta \rangle_y := \langle \zeta | \{\mathbf{a}(x)\}_{x \in \mathbb{Z}^d \setminus \{y\}} \rangle,$$

i. e. the conditional expectation of  $\zeta$  given  $\mathbf{a}(x)$  for all  $x \in \mathbb{Z}^d \setminus \{y\}$ . We set

$$\frac{\partial \zeta}{\partial y} := \zeta - \langle \zeta \rangle_y.$$

**Remark 3.** 1. The vertical derivative  $\frac{\partial}{\partial y}$  can be seen as a discrete version of the classical partial derivative  $\frac{\partial}{\partial \mathbf{a}(y)}$ . It monitors how sensitively a random variable  $\zeta$  depends on the value of the coefficient field  $\{\mathbb{Z}^d \ni y \mapsto \mathbf{a}(y)\}$  at site  $y$ . Note that due to discreteness the vertical derivative does not satisfy the product rule.

2. The vertical derivative  $\frac{\partial}{\partial y} : L^2(\Omega) \rightarrow L^2(\Omega)$  defines a bounded and symmetric operator on  $L^2(\Omega)$ . In fact,  $\frac{\partial \zeta}{\partial y}$  is precisely the  $L^2(\Omega)$ -orthogonal projection of  $\zeta$  onto the subspace of random variables in  $L^2(\Omega)$  that do not depend on the coefficient  $\mathbf{a}(y)$ .

3. The vertical derivative does not commute with the shift operator. In fact, we have

$$\begin{aligned} \langle \zeta \circ T_x \rangle_y &= \langle \zeta \circ T_x | \{\mathbf{a}(z)\}_{z \neq y} \rangle = \langle \zeta \circ T_x | \{(T_{-x}T_x \mathbf{a})(z)\}_{z \neq y} \rangle \\ &= \langle \zeta \circ T_x | \{(T_x \mathbf{a})(z-x)\}_{z \neq y} \rangle \\ &= \langle \zeta \circ T_x | \{(T_x \mathbf{a})(z')\}_{z' \neq y-x} \rangle \\ &= \langle \zeta \rangle_{y-x} \circ T_x, \end{aligned}$$

which in terms of the stationary extension  $\overline{(\cdot)}$  reads

$$(49) \quad \langle \overline{\zeta}(x) \rangle_y = \overline{\langle \zeta \rangle_{y-x}}(x).$$

Hence,

$$(50) \quad \frac{\partial \overline{\zeta}(x)}{\partial y} = \overline{\left( \frac{\partial \zeta}{\partial (y-x)} \right)}(x).$$

Now we are in position to introduce the key assumption on the ensemble:

**Definition 6** (spectral gap for the Glauber dynamics). *We say  $\langle \cdot \rangle$  satisfies a spectral gap for the Glauber dynamics with constant  $\rho > 0$ , in short  $SG_\infty(\rho)$ , if for all  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$  we have*

$$\langle \zeta^2 \rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle.$$

From the functional analytic point of view, the spectral gap estimate is a Poincaré inequality on  $L^2(\Omega)$  for the vertical derivative  $\frac{\partial}{\partial y}$ . Since each site  $y \in \mathbb{Z}^d$  is endowed with a vertical derivative, the number of degrees of freedom that we control with the right-hand side of the spectral gap estimate matches the dimensionality of the underlying probability space  $\Omega$  – recall that  $\Omega$  is the  $\mathbb{Z}^d$ -fold product of  $\Omega_0$ . Hence,  $SG_\infty(\rho)$  is a reasonable assumption. It has a natural interpretation in statistical mechanics:

**Remark 4** (Glauber dynamics). *In statistical mechanics,  $\mathbf{a} \in \Omega$  is viewed as the configuration of a spin system on the lattice  $\mathbb{Z}^d$  with continuous spin space  $\Omega_0$ . The ensemble  $\langle \cdot \rangle$  is typically described by a Hamiltonian  $H(\mathbf{a})$ , i. e. of the form*

$$\langle \zeta \rangle = \frac{1}{Z} \int_{\Omega_0^{\mathbb{Z}^d}} \zeta(\mathbf{a}) \exp(-H(\mathbf{a})) d\mathbf{a}.$$

*Typically, the dynamics on  $\Omega$  are reversible w. r. t. to the equilibrium measure  $\langle \cdot \rangle$ . A particular simple reversible dynamics is Glauber dynamics, that can be viewed as a “spin flip” dynamics. In the “heat-bath” version, the dynamics associates each site with an exponential clock. If the clock rings, say at site  $y$ , then the spin  $\mathbf{a}(y)$  is updated by randomly drawing a spin from  $\Omega_0$  with a distribution given by  $\langle \cdot \rangle_y(\mathbf{a})$ . Expressed as an evolution on the level of densities w. r. t. the equilibrium measure  $\langle \cdot \rangle$ , the dynamics is generated by the positive, symmetric operator*

$$\sum_{y \in \mathbb{Z}^d} \left( \frac{\partial}{\partial y} \right)^2$$

*and the associated Dirichlet energy is given by*

$$L^2(\Omega) \ni \zeta \mapsto \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle.$$

*This is precisely the right-hand side in the spectral gap estimate.  $SG_\infty(\rho)$  quantifies the gap between the zero eigenvalue and the first non-zero eigenvalue of  $\sum_{y \in \mathbb{Z}^d} \left( \frac{\partial}{\partial y} \right)^2$ , and yields an exponential relaxation rate for the Glauber dynamics to equilibrium.*

A fundamental example for stationary ensembles satisfying  $SG_\infty(\rho)$  are ensembles associated with independent and identically distributed (i. i. d.) coefficients; which means that the coordinate projections  $\Omega \ni \mathbf{a} \mapsto \mathbf{a}(x) \in \Omega_0$ ,  $x \in \mathbb{Z}^d$ , are independent and identically distributed. Obviously, an ensemble  $\langle \cdot \rangle$  is associated with i. i. d. coefficients if and only if it can be written as the product of a probability measure  $\beta$  on  $\Omega_0$ :

**Definition 7.** *Let  $\beta$  be a probability measure on  $\Omega_0$ . The infinite i. i. d. ensemble associated with the base measure  $\beta$  is defined via*

$$(51) \quad \forall \zeta \in C^0(\Omega) : \quad \langle \zeta \rangle = \int_{\Omega} \zeta(\mathbf{a}) \left( \beta^{\otimes \mathbb{Z}^d} \right) (d\mathbf{a}),$$

*where  $\beta^{\otimes \mathbb{Z}^d}$  denotes the  $\mathbb{Z}^d$ -fold product of  $\beta$ .*

**Lemma 7.** *Let  $\langle \cdot \rangle$  be an infinite i. i. d. ensemble. Then  $\langle \cdot \rangle$  satisfies  $SG_\infty(\rho)$  with constant  $\rho = 1$ , and, in particular, is ergodic.*

Unless stated otherwise we postpone the proofs of that and the following results to the end of this section. The program of the following discussion is to show that based on  $SG_\infty(\rho)$  we can quantify ergodicity, in the sense that  $SG_\infty(\rho)$  yields an algebraic rate for the decay in (48), and thus quantifies the thickness of the spectral bottom of  $D^*D$ . We now are ready to state the main results of this section.

**Proposition 1.** *Let  $\langle \cdot \rangle$  be stationary and satisfy  $SG_\infty(\rho)$ . Then for all  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$  we have*

$$\langle (\exp(-tD^*D)\zeta)^2 \rangle^{\frac{1}{2}} \leq \frac{1}{\sqrt{\rho}} \left( \sum_{z \in \mathbb{Z}^d} G^2(t, z) \right)^{\frac{1}{2}} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle^{\frac{1}{2}},$$

*where  $G(t, x)$  denotes the discrete, constant-coefficient parabolic Green’s function, cf. Definition 1.*

**Remark 5.** The combination of Proposition 1 with Lemma 4 yields for all  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$

$$\langle (\exp(-tD^*D)\zeta)^2 \rangle^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{\rho}}(t+1)^{-\frac{d}{4}} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle^{\frac{1}{2}}.$$

**Remark 6** (optimality). Proposition 1 is optimal in the sense that there exists a non-trivial ensemble  $\langle \cdot \rangle$  and non-trivial initial data  $\zeta$  such that the estimate in the proposition holds with equality. Indeed, let  $\langle \cdot \rangle$  be an infinite i. i. d. ensemble associated with some base measure  $\beta$ . Consider  $u(t) := \exp(-tD^*D)\zeta$  with initial data

$$\zeta(\mathbf{a}) := f(\mathbf{a}(0))$$

where  $f : \Omega_0 \rightarrow \mathbb{R}$  is square integrable and mean-free w. r. t.  $\beta$ . Since the ensemble is i. i. d., we have

$$\langle f(\mathbf{a}(0)) \rangle = 0 \quad \text{and} \quad \left( \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle^{\frac{1}{2}} \right)^2 = \text{Var}_\beta(f),$$

where  $\text{Var}_\beta$  denotes the variance w. r. t. to  $\beta$ . By Lemma 7 the ensemble satisfies  $SG_\infty(\rho)$  with constant  $\rho = 1$  and thus Proposition 1 yields

$$(52) \quad \langle u^2(t) \rangle \leq \sum_{z \in \mathbb{Z}^d} G^2(t, z) \text{Var}_\beta(f).$$

On the other hand, we now argue that

$$(53) \quad \langle u^2(t) \rangle = \sum_{z \in \mathbb{Z}^d} G^2(t, z) \text{Var}_\beta(f).$$

The argument starts with the Green's function representation in Lemma 10 below, which leads to

$$\begin{aligned} \langle u^2(t) \rangle &= \left\langle \left( \sum_{z \in \mathbb{Z}^d} G(t, -z) f(\mathbf{a}(z)) \right)^2 \right\rangle \\ &\stackrel{(26e)}{=} \sum_{z \in \mathbb{Z}^d} \sum_{z' \in \mathbb{Z}^d} G(t, z) G(t, z') \langle f(\mathbf{a}(z)) f(\mathbf{a}(z')) \rangle. \end{aligned}$$

Since  $\langle \cdot \rangle$  is i. i. d., the ensemble average on the right-hand side simplifies to

$$\langle f(\mathbf{a}(z)) f(\mathbf{a}(z')) \rangle = \delta(z - z') \text{Var}_\beta(f),$$

and we conclude

$$\langle u^2(t) \rangle = \sum_{z \in \mathbb{Z}^d} \sum_{z' \in \mathbb{Z}^d} G(t, z) G(t, z') \delta(z - z') \text{Var}_\beta(f) = \sum_{z \in \mathbb{Z}^d} G^2(t, z) \text{Var}_\beta(f).$$

In combination with Corollary 5 and the fact that local random variables (i. e. measurable functions on  $\Omega$  that only depend on a finite number of sites) are dense in  $L^2(\Omega)$ , we get as a corollary from Proposition 1:

**Corollary 6.** Let  $\langle \cdot \rangle$  be stationary and satisfy  $SG_\infty(\rho)$ . Then  $\langle \cdot \rangle$  is ergodic.

If the initial data  $\zeta$  are of the form  $\zeta = D^*\xi$ , the decay of the semigroup is better, as the next proposition shows.

**Proposition 2.** Let  $\langle \cdot \rangle$  be stationary and suppose that  $SG_\infty(\rho)$  holds with constant  $\rho > 0$ . For  $\xi \in L^2(\Omega)^d$  consider  $u(t) := \exp(-tD^*D)D^*\xi$ . Then

$$\langle (\exp(-tD^*D)D^*\xi)^2 \rangle^{\frac{1}{2}} \leq \frac{1}{\sqrt{\rho}} \left( \sum_{z \in \mathbb{Z}^d} |\nabla G(t, z)|^2 \right)^{\frac{1}{2}} \sum_{y \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \xi}{\partial y} \right|^2 \right\rangle^{\frac{1}{2}}.$$

**Remark 7.** The gradient of the parabolic Green's function satisfies the estimate

$$(54) \quad \left( \sum_{z \in \mathbb{Z}^d} |\nabla G(t, z)|^2 \right)^{\frac{1}{2}} \lesssim (t+1)^{-\left(\frac{d}{4} + \frac{1}{2}\right)}.$$

This can be easily seen by a Nash-Aronson-type argument that basically relies on Nash's inequality and (26c) (and thus on the weak maximum principle). We do not display that elementary argument here, since in Section 7 we will establish a finer version of (54), namely for a spatially weighted  $\ell^p$ -norm and in the variable-coefficient case, see Theorem 3 (a). Note that estimate (54) is contained in Theorem 3 (a), as can be seen by the argument in Remark 12. In combination with Proposition 2 we get for  $\xi \in L^2(\Omega)^d$  the decay estimate

$$\langle (\exp(-tD^*D)\xi)^2 \rangle^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{\rho}} (t+1)^{-\left(\frac{d}{4} + \frac{1}{2}\right)} \sum_{y \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \xi}{\partial y} \right|^2 \right\rangle^{\frac{1}{2}}.$$

#### 4.1 The periodic ensemble

In the following we consider the periodic case. Periodic ensembles are typically not ergodic (see Lemma 9 below), and thus  $\text{INV}^2(\Omega)$  is in general not trivial. However, we have the following characterization:

**Lemma 8.** Let  $\langle \cdot \rangle$  be stationary and  $L$ -periodic. Then

$$\text{INV}^2(\Omega) = \left\{ \zeta \in L^2(\Omega) : \zeta = \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \zeta \circ T_x \quad \langle \cdot \rangle\text{-almost surely} \right\}.$$

*Proof of Lemma 8.* The inclusion “ $\subset$ ” is trivial. The opposite inclusion follows from the following elementary observation: Since  $\langle \cdot \rangle$  is stationary and  $L$ -periodic, the stationary extension  $\bar{\zeta}(\mathbf{a}, x) := \zeta(T_x \mathbf{a})$  is  $L$ -periodic in  $x$  for almost every  $\mathbf{a} \in \Omega$ . Thus, the empirical average

$$\Omega \ni \mathbf{a} \mapsto \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \bar{\zeta}(\mathbf{a}, x)$$

is shift invariant. □

Recall that  $L^2(\Omega) = \text{INV}^2(\Omega) \oplus L_0^2(\Omega)$  is an  $L^2(\Omega)$ -orthogonal decomposition. By Lemma 5 the set  $\text{INV}^2(\Omega)$  is precisely the kernel of the operator  $D^*D$ , and from Lemma 6 we learn that

$$(55) \quad \forall \zeta \in L_0^2(\Omega) : \quad \lim_{t \uparrow \infty} \langle |\exp(-tD^*D)\zeta|^2 \rangle = 0.$$

In analogy to Proposition 1 we quantify the decay in (55) by appealing to a spectral gap estimate that is adapted to the periodic setting:

**Definition 8** (vertical derivative, periodic case). Let  $\langle \cdot \rangle$  be stationary and  $L$ -periodic. For  $\zeta \in L^2(\Omega)$  and  $y \in \mathbb{Z}^d$  we define

$$\langle \zeta \rangle_{L,y} := \langle \zeta | \{ \mathbf{a}(x) \}_{x \in \mathbb{Z}^d \setminus \{y + L\mathbb{Z}^d\}} \rangle,$$

i. e. the conditional expectation of  $\zeta$  given  $\mathbf{a}(x)$  for all  $x \in \mathbb{Z}^d \setminus \{y + L\mathbb{Z}^d\}$ . We set

$$\frac{\partial \zeta}{\partial_L y} := \zeta - \langle \zeta \rangle_{L,y}.$$

**Remark 8.** The calculus rules for the vertical derivative and horizontal shift (49) and (50) also hold for  $\langle \cdot \rangle_y$  and  $\frac{\partial}{\partial y}$  replaced by  $\langle \cdot \rangle_{L,y}$  and  $\frac{\partial}{\partial_L y}$ .

**Definition 9** (spectral gap estimate, periodic case). We say  $\langle \cdot \rangle$  satisfies a spectral gap estimate with constant  $\rho > 0$  on the torus of size  $L \in \mathbb{N}$ , in short  $SG_L(\rho)$ , if  $\langle \cdot \rangle$  is stationary,  $L$ -periodic and for all  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$  we have

$$\langle \zeta^2 \rangle \leq \frac{1}{\rho} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left( \frac{\partial \zeta}{\partial_L y} \right)^2 \right\rangle.$$

A fundamental example for a  $L$ -periodic ensemble satisfying  $SG_L(\rho)$  is the  $L$ -periodic ensembles associated with i. i. d. coefficients:

**Definition 10.** Let  $\beta$  be a probability measure on  $\Omega_0$ . The  $L$ -periodic i. i. d. ensemble associated with the base measure  $\beta$  is defined via

$$(56) \quad \forall \zeta \in C^0(\Omega) : \quad \langle \zeta \rangle = \int_{\Omega_0^{([0,L) \cap \mathbb{Z})^d}} (\zeta \circ \mathbf{a}_L)(\mathbf{a}) \left( \beta^{\otimes \Lambda^d} \right) (d\mathbf{a}),$$

where  $\mathbf{a}_L : \Omega_0^{([0,L) \cap \mathbb{Z})^d} \rightarrow \Omega_L$  denotes the mapping which associates a coefficient field  $\mathbf{a} \in \Omega_0^{([0,L) \cap \mathbb{Z})^d}$  with its  $L$ -periodic extension  $\mathbf{a}_L(\mathbf{a}) \in \Omega_L$  characterized by  $\mathbf{a}_L(\mathbf{a}) = \mathbf{a}$  on  $([0,L) \cap \mathbb{Z})^d$ .

**Lemma 9.** Let  $\langle \cdot \rangle$  be the  $L$ -periodic i. i. d. ensemble associated with a base measure  $\beta$ . Then  $\langle \cdot \rangle$  is stationary,  $L$ -periodic and satisfies  $SG_L(\rho)$  with constant  $\rho = 1$ . It is ergodic if and only if  $\beta$  is a Dirac measure on  $\Omega_0$ .

**Proposition 3.** Let  $\langle \cdot \rangle$  satisfy  $SG_L(\rho)$  with constant  $\rho > 0$ . Then for all  $\zeta \in L_0^2(\Omega)$  we have

$$\langle (\exp(-tD^*D)\zeta)^2 \rangle^{\frac{1}{2}} \leq \frac{1}{\sqrt{\rho}} \left( \sum_{x \in ([0,L) \cap \mathbb{Z})^d} |G_L(t,x) - \bar{G}_L|^2 \right)^{\frac{1}{2}} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left( \frac{\partial \zeta}{\partial_L y} \right)^2 \right\rangle^{\frac{1}{2}}.$$

where  $\bar{G}_L := L^{-d}$  is the spatial average of  $G_L$  over  $([0,L) \cap \mathbb{Z})^d$ .

In combination with the Green's function estimate Corollary 4 we deduce that for all  $\zeta \in L_0^2(\Omega)$  we have

$$\langle (\exp(-tD^*D)\zeta)^2 \rangle^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{\rho}} (t+1)^{-\frac{d}{4}} \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left( \frac{\partial \zeta}{\partial_L y} \right)^2 \right\rangle^{\frac{1}{2}},$$

Hence, for initial conditions in  $L_0^2(\Omega)$  the proposition yields a decay of the semigroup with the same rate as the one in Proposition 1. However, in contrast to the situation considered in Proposition 1, in the periodic case  $INV^2(\Omega)$  is typically non-trivial (in the sense that it contains more random variables than the constant ones). Thus, Corollary 5 is not applicable.



## 4.2 Proofs

In the proof of the Propositions 1 and 2 we appeal to the Green's function representation formula for solutions to the parabolic equation in the probability space:

**Lemma 10.** For  $\zeta \in L^2(\Omega)$  consider  $u(t) := \exp(-tD^*D)\zeta$ . Then for all  $t > 0$  we have

$$u(t) = \sum_{z \in \mathbb{Z}^d} G(t, z) \bar{\zeta}(z) \quad \langle \cdot \rangle\text{-almost surely.}$$

*Proof of Lemma 10.* By a simple approximation argument that we leave to the reader, it suffices to consider  $\zeta \in C^0(\Omega)$ . Consider the stationary extension  $\bar{u}(t, \mathbf{a}, x) := u(t, \mathbf{a}(\cdot + x))$ . By Lemma 1, for all  $\mathbf{a} \in \Omega$  the mapping  $(t, x) \mapsto \bar{u}(t, \mathbf{a}, x)$  belongs to  $C^\infty(\mathbb{R}_+, \ell^\infty(\mathbb{Z}^d))$  and solves the parabolic equation

$$(57) \quad \begin{cases} \partial_t \bar{u}(t, \mathbf{a}, x) + \nabla^* \nabla \bar{u}(t, \mathbf{a}, x) = 0 & \text{for all } t > 0, x \in \mathbb{Z}^d, \\ \bar{u}(t = 0, \mathbf{a}, x) = \bar{\zeta}(\mathbf{a}, x) & \text{for all } x \in \mathbb{Z}^d. \end{cases}$$

Hence, by appealing to Duhamel's formula, see Lemma 2, we get

$$\bar{u}(t, \mathbf{a}, x) = \sum_{y \in \mathbb{Z}^d} G(t, x - z) \bar{\zeta}(\mathbf{a}, z).$$

Taking  $x = 0$  and using the symmetry property  $G(t, -z) = G(t, z)$  proves the claim.  $\square$

*Proof of Proposition 1.* To ease the notation set

$$u(t) := \exp(-tD^*D)\zeta.$$

**Step 1.** Application of (SG).

By Lemma 10 we have

$$\langle u(t) \rangle = \left\langle \sum_{z \in \mathbb{Z}^d} G(t, z) \bar{\zeta}(z) \right\rangle \stackrel{\text{stationarity}}{=} \langle \zeta \rangle \sum_{z \in \mathbb{Z}^d} G(t, z) = 0$$

for all  $t \geq 0$ , and by (44) we have  $\langle u^2(t) \rangle \leq \langle \zeta^2 \rangle < \infty$ . Hence, we can apply  $\text{SG}_\infty(\rho)$  and obtain

$$(58) \quad \langle u^2(t) \rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial u(t)}{\partial y} \right)^2 \right\rangle.$$

**Step 2.** Formula for  $\frac{\partial u(t)}{\partial y}$ .

We claim that

$$\frac{\partial u(t)}{\partial y} = \sum_{z \in \mathbb{Z}^d} G(t, z) \overline{\left( \frac{\partial \zeta}{\partial(y-z)} \right)}(z).$$

We recall that  $\overline{(\cdot)}$  denotes the mapping that associates the stationary extension with its random variable. By Lemma 10,

$$u(t) = \sum_{z \in \mathbb{Z}^d} G(t, z) \bar{\zeta}(z).$$

Since  $G$  does not depend on the coefficient field  $\mathbf{a}$ , this yields

$$\frac{\partial u(t)}{\partial y} = \sum_{z \in \mathbb{Z}^d} G(t, z) \frac{\partial \bar{\zeta}(z)}{\partial y}.$$

and the claim follows from (50).

**Step 3. Conclusion.**

In view of Steps 1 and 2, we only need to prove the estimate

$$(59) \quad \left( \sum_{y \in \mathbb{Z}^d} \left\langle \left( \sum_{z \in \mathbb{Z}^d} G(t, z) \overline{\left( \frac{\partial \zeta}{\partial(y-z)} \right)}(z) \right)^2 \right\rangle \right)^{\frac{1}{2}} \leq \left( \sum_{z \in \mathbb{Z}^d} G^2(t, z) \right)^{\frac{1}{2}} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle^{\frac{1}{2}}.$$

Here comes the argument: With the new variable  $x = y - z$  the left-hand side of (59) reads

$$\begin{aligned} [\text{L. H. S. of (59)}] &= \left( \sum_{y \in \mathbb{Z}^d} \left\langle \left( \sum_{x \in \mathbb{Z}^d} G(t, y-x) \overline{\left( \frac{\partial \zeta}{\partial x} \right)}(y-x) \right)^2 \right\rangle \right)^{\frac{1}{2}} \\ &\stackrel{\Delta\text{-inequality}}{\leq} \left( \sum_{y \in \mathbb{Z}^d} \left\langle \left( \sum_{x \in \mathbb{Z}^d} \left\langle \left( G(t, y-x) \overline{\left( \frac{\partial \zeta}{\partial x} \right)}(y-x) \right)^2 \right\rangle \right) \right\rangle \right)^{\frac{1}{2}} \\ &\stackrel{G \text{ is deterministic, stationarity}}{=} \sum_{x \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} G^2(t, y-x) \left\langle \left| \frac{\partial \zeta}{\partial x} \right|^2 \right\rangle \right)^{\frac{1}{2}} \\ &= \sum_{x \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \zeta}{\partial x} \right|^2 \right\rangle^{\frac{1}{2}} \left( \sum_{y \in \mathbb{Z}^d} G^2(t, y-x) \right)^{\frac{1}{2}}. \end{aligned}$$

□

*Proof of Proposition 2.* Set  $\zeta := D^* \xi$  and  $u(t) := \exp(-tD^*D)\zeta$ . Since  $\langle \zeta \rangle = 0$  and therefore also  $\langle u(t) \rangle = 0$ , the application of  $\text{SG}_\infty(\rho)$  yields

$$(60) \quad \langle u^2(t) \rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial u(t)}{\partial y} \right)^2 \right\rangle.$$

We claim that

$$(61) \quad \frac{\partial u(t)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, z) \cdot \overline{\left( \frac{\partial \xi}{\partial(y-z)} \right)}(z).$$

Indeed, by Lemma 10,  $u(t) = \sum_{z \in \mathbb{Z}^d} G(t, z) \bar{\zeta}(z)$ . Since  $\bar{\zeta} = \overline{D^* \xi} = \nabla^* \bar{\xi}$ , an integration by parts yields

$$u(t) = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, z) \cdot \bar{\xi}(z).$$

Now (61) can be proved as in Step 2 in the proof of Proposition 1. By repeating the arguments in Step 3 in the proof of Proposition 1 we obtain from (61) the estimate

$$\left( \sum_{y \in \mathbb{Z}^d} \left\langle \left( \sum_{z \in \mathbb{Z}^d} \nabla G(t, z) \cdot \overline{\left( \frac{\partial \xi}{\partial(y-z)} \right)}(z) \right)^2 \right\rangle \right)^{\frac{1}{2}} \leq \sum_{y \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \xi}{\partial y} \right|^2 \right\rangle^{\frac{1}{2}} \left( \sum_{z \in \mathbb{Z}^d} |\nabla G(t, z)|^2 \right)^{\frac{1}{2}}.$$

□

*Proof of Lemma 7.* By virtue of Corollary 6 we only have to show

$$(62) \quad \langle \zeta^2 \rangle \leq \sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle$$

for random variables  $\zeta \in L^2(\Omega)$  with  $\langle \zeta \rangle = 0$ . By an approximation argument it suffices to consider local random variables, i. e.  $\zeta$  only depends on the value of the coefficient field at a finite number of sites  $y$ . For the argument we introduce a couple of notation: Let  $\{y_n\}_{n=1,2,3,\dots}$  be an enumeration of  $\mathbb{Z}^d$ . For  $n = 1, 2, 3, \dots$  let  $\langle \cdot \rangle_{\leq n}$  denote the  $L^2(\Omega)$ -orthogonal projection onto the subspace of  $L^2(\Omega)$ -random variables that do not depend on the coefficients  $\mathbf{a}(y_1), \dots, \mathbf{a}(y_n)$ , and set  $\zeta_n := \langle \zeta \rangle_{\leq n}$  for  $n \geq 1$  and  $\zeta_0 := \langle \zeta \rangle_{\leq 0} := \zeta$ . (Note that  $\langle \cdot \rangle_{\leq n}$  is precisely the conditional expectation where we condition on the value of  $\mathbf{a}$  at all sites except  $y_1, \dots, y_n$ ). We split the proof of (62) into two steps.

**Step 1.** Martingale decomposition.

We claim that

$$(63) \quad \langle \zeta^2 \rangle = \sum_{n=1}^{\infty} \langle (\zeta_{n-1} - \zeta_n)^2 \rangle.$$

Here comes the argument: By construction we have  $\zeta_0 = \zeta$ . Since  $\zeta$  is assumed to be local and mean free we have  $\zeta_N = \langle \zeta \rangle = 0$  for  $N$  large enough, and thus

$$(64) \quad \zeta = \sum_{n=1}^N \zeta_{n-1} - \zeta_n.$$

Taking the square and the expected value yields

$$\langle \zeta^2 \rangle = \sum_{n=1}^N \sum_{m=1}^N \langle (\zeta_{n-1} - \zeta_n)(\zeta_{m-1} - \zeta_m) \rangle.$$

Hence, (63) follows, provided that the random variables  $\{\zeta_{n-1} - \zeta_n\}_{n \in \mathbb{N}}$  are pairwise orthogonal in  $L^2(\Omega)$ . For the argument let  $m > n$ . Since by construction  $\zeta_{m-1} - \zeta_m$  does not depend on  $\mathbf{a}(y_1), \dots, \mathbf{a}(y_{m-1})$  we have

$$(65) \quad \zeta_{m-1} - \zeta_m = \langle \zeta_{m-1} - \zeta_m \rangle_{\leq m-1};$$

and by general properties of the orthogonal projection we get

$$(66) \quad \langle \zeta_{n-1} - \zeta_n \rangle_{\leq m-1} = \langle \langle \zeta \rangle_{\leq n-1} \rangle_{\leq m-1} - \langle \langle \zeta \rangle_{\leq n} \rangle_{\leq m-1} = \langle \zeta \rangle_{\leq m-1} - \langle \zeta \rangle_{\leq m-1} = 0.$$

Thus, since  $\langle \cdot \rangle_{m-1}$  is symmetric, we get

$$\begin{aligned} \langle (\zeta_{m-1} - \zeta_m)(\zeta_{n-1} - \zeta_n) \rangle &\stackrel{(65)}{=} \langle \langle \zeta_{m-1} - \zeta_m \rangle_{\leq m-1} (\zeta_{n-1} - \zeta_n) \rangle \\ &= \langle (\zeta_{m-1} - \zeta_m) \langle \zeta_{n-1} - \zeta_n \rangle_{\leq m-1} \rangle \\ &\stackrel{(66)}{=} 0 \end{aligned}$$

and the claim follows.

**Step 2.** Conclusion by exploiting the product structure of  $\langle \cdot \rangle$ .

We claim that for all  $n \in \mathbb{N}$  we have

$$\langle (\zeta_{n-1} - \zeta_n)^2 \rangle \leq \left\langle \left( \frac{\partial \zeta}{\partial y_n} \right)^2 \right\rangle.$$

For the argument recall that the i. i. d. ensemble  $\langle \cdot \rangle$  is a product measure  $\beta^{\otimes \mathbb{Z}^d}$  where the base measure  $\beta$  is a probability measure on  $\Omega_0$ . As a consequence we find that

$$\zeta_n = \langle \langle \zeta \rangle_{y_n} \rangle_{\leq n-1},$$

where  $\langle \cdot \rangle_{y_n}$  is defined in Definition 5. Hence

$$\begin{aligned} \langle (\zeta_{n-1} - \zeta_n)^2 \rangle &= \langle \langle (\zeta - \langle \zeta \rangle_{y_n})^2 \rangle_{\leq n-1} \rangle \\ &\stackrel{\text{Jensen}}{\leq} \langle \langle (\zeta - \langle \zeta \rangle_{y_n})^2 \rangle_{\leq n-1} \rangle = \langle (\zeta - \langle \zeta \rangle_{y_n})^2 \rangle \\ &= \left\langle \left( \frac{\partial \zeta}{\partial y_n} \right)^2 \right\rangle. \end{aligned}$$

□

*Proof of Lemma 9.* The argument for  $\text{SG}_L(\rho)$  is similar to the one in the proof of Lemma 7. Therefore, we only prove that ergodicity of  $\langle \cdot \rangle$  is equivalent to the property that  $\langle \cdot \rangle$  concentrates on a non-random, spatially constant-coefficient field. The latter means that the probability measure  $\beta$  (associated with  $\langle \cdot \rangle$  via (56)) is a Dirac measure on  $\Omega_0$ . If  $\beta$  is a Dirac measure, then  $\langle \cdot \rangle$  is obviously ergodic. Now, assume that  $\beta$  is not a Dirac, which means that there exists a measurable set  $S \subset \Omega_0$  with  $0 < \beta(S) < 1$ . Consider the set  $A := \{\mathbf{a} \in \Omega_L : \mathbf{a}(x) \in S \text{ for all } x \in ([0, L] \cap \mathbb{Z})^d\}$  and let  $I_A$  denote the associated characteristic function. By construction  $A$  is shift invariant, and thus  $I_A \in \text{INV}^2(\Omega)$ . Suppose now that  $\langle \cdot \rangle$  is ergodic. We deduce that  $I_A = \langle I_A \rangle$  almost surely. Since  $I_A$  is  $\{0, 1\}$ -valued, we get  $\langle I_A \rangle \in \{0, 1\}$ . However, by construction  $\langle I_A \rangle = \beta(S)^{L^d} \notin \{0, 1\}$ . Hence,  $\langle \cdot \rangle$  is not ergodic. □

*Proof of Proposition 3.* The proof is similar to the one of Proposition 1. Therefore, we only indicate the modifications. Set  $u(t) := \exp(-tD^*D)\zeta$ . Since  $\zeta \in L_0^2(\Omega)$  we have  $u(t) \in L_0^2(\Omega)$  (cf. proof of Lemma 6) and thus, in particular,  $\langle u(t) \rangle = 0$ . Hence,  $\text{SG}_L(\rho)$  yields

$$(67) \quad \langle u^2(t) \rangle \leq \frac{1}{\rho} \sum_{y \in ([0, L] \cap \mathbb{Z})^d} \left\langle \left( \frac{\partial u(t)}{\partial L y} \right)^2 \right\rangle.$$

By appealing to Lemma 10 and the  $L$ -periodicity of  $\bar{\zeta}$  we get the representation

$$\begin{aligned} u(t) &= \sum_{z \in \mathbb{Z}^d} G(t, z) \bar{\zeta}(z) = \sum_{z' \in L\mathbb{Z}^d} \sum_{z \in ([0, L] \cap \mathbb{Z})^d} G(t, z + z') \bar{\zeta}(z + z') \\ &\stackrel{\text{periodicity}}{=} \sum_{z' \in L\mathbb{Z}^d} \sum_{z \in ([0, L] \cap \mathbb{Z})^d} G(t, z + z') \bar{\zeta}(z) \\ &\stackrel{(32)}{=} \sum_{z \in ([0, L] \cap \mathbb{Z})^d} G_L(t, z) \bar{\zeta}(z). \end{aligned}$$

Since  $\zeta \in L_0^2(\Omega)$ , we have  $\sum_{z \in ([0, L] \cap \mathbb{Z})^d} \bar{\zeta}(z) = 0$ , cf. Lemma 8, and thus the above can be rewritten as

$$u(t) = \sum_{z \in ([0, L] \cap \mathbb{Z})^d} (G_L(t, z) - \bar{G}_L) \bar{\zeta}(z).$$

Now, we proceed as in Step 3 of the proof of Proposition 1. We get

$$\frac{\partial u(t)}{\partial L y} = \sum_{z \in ([0, L] \cap \mathbb{Z})^d} (G_L(t, z) - \bar{G}_L) \overline{\left( \frac{\partial \zeta}{\partial L(y - z)} \right)}(y).$$

The conclusion follows as in Step 3 of the proof of Proposition 1. □

## 5 Decay of the variable-coefficient semigroup

This chapter is the core of the text and the one that contains new results. In the following we consider the variable-coefficient semigroup. We denote by  $\mathbf{a}(0)$  the projection  $\Omega \ni \mathbf{a} \mapsto \mathbf{a}(0) \in \Omega_0$  associated with the site  $x = 0$ . Recall that  $D^*\mathbf{a}(0)D$  is a bounded linear operator on  $C^0(\Omega)$  and thus

$$t \mapsto \exp(-tD^*\mathbf{a}(0)D)$$

defines a uniformly continuous group of bounded operators on  $C^0(\Omega)$  (see Section 3 for the details). Recall that for all initial conditions  $\zeta \in C^0(\Omega)$  the function  $u(t) := \exp(-tD^*\mathbf{a}(0)D)\zeta$  belongs to  $C^\infty(\mathbb{R}_+, C^0(\Omega))$  and solves the parabolic equation

$$(68) \quad \begin{cases} \frac{d}{dt}u(t) + D^*\mathbf{a}(0)Du(t) = 0 & \text{for all } t > 0, \mathbf{a} \in \Omega, \\ u(t=0) = \zeta & \text{for all } \mathbf{a} \in \Omega. \end{cases}$$

This equation defines a stochastic process on the space  $\Omega$  of coefficient fields that can be conveniently interpreted in the context of random walks in random environments: Consider a random walker on  $\mathbb{Z}^d$  with jump rate across a bond  $\{x, x + e_i\}$  given by  $\mathbf{a}_{ii}(x)$ . Then the above process describes the “environment  $\mathbf{a}$  as seen by the walker”.

As a main result we prove an optimal decay estimate for  $\exp(-tD^*\mathbf{a}(0)D)$  with initial conditions in divergence form.

**Theorem 2.** *Assume that either*

$$(69a) \quad \langle \cdot \rangle \text{ is stationary and satisfies } SG_\infty(\rho),$$

or

$$(69b) \quad \langle \cdot \rangle \text{ is stationary, } L\text{-periodic and satisfies } SG_L(\rho).$$

Then there exists an exponent  $1 \leq p_0 < \infty$  that only depends on  $\lambda$  and  $d$  such that for all  $p_0 \leq p < \infty$  and  $t \geq 0$  the function

$$u(t) := \exp(-tD^*\mathbf{a}(0)D)D^*\xi, \quad \xi \in C^0(\Omega)^d,$$

satisfies

$$\langle |u(t)|^{2p} \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-\left(\frac{d}{4} + \frac{1}{2}\right)} \|\xi\|_{\ell_y^1 L_{(\cdot)}^{2p}},$$

where

$$\|\xi\|_{\ell_y^1 L_{(\cdot)}^{2p}} := \begin{cases} \sum_{y \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \xi}{\partial y} \right|^{2p} \right\rangle^{\frac{1}{2p}} & \text{in case of (69a),} \\ \sum_{y \in ([0,L) \cap \mathbb{Z})^d} \left\langle \left| \frac{\partial \xi}{\partial_L y} \right|^{2p} \right\rangle^{\frac{1}{2p}} & \text{in case of (69b),} \end{cases}$$

and  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p, \rho, \lambda$  and  $d$ .

**Remark 9.** *As explained in Remark 7, Proposition 2 combined with gradient estimates on the parabolic (constant coefficient) Green’s function yields the decay estimate*

$$\langle |\exp(-tD^*D)D^*\xi|^2 \rangle^{\frac{1}{2}} \lesssim (t+1)^{-\left(\frac{d}{4} + \frac{1}{2}\right)} \|\xi\|_{\ell_y^1 L_{(\cdot)}^2}.$$

Theorem 2 generalizes this estimate to the variable-coefficient case, i. e. for the group generated by  $D^*\mathbf{a}(0)D$ . In particular the theorem shows that the variable-coefficient semigroup decays with the same algebraic rate as the constant-coefficient semigroup. However, in contrast to Proposition 2, in the variable-coefficient case we have to estimate higher moments. In fact, in the case of high ellipticity ratio of the coefficient field, the exponent  $p$  is required to be large for our argument.

**Remark 10.** By the definition of  $\Omega_0$ , cf. (5), any coefficient field  $\mathbf{a} \in \Omega$  is diagonal in the sense of  $\mathbf{a}_{ij} = 0$  for  $i \neq j$ . As explained in Section 3, the reason for that assumption is that our argument requires estimates based on elliptic and parabolic regularity for discrete difference equations on  $\mathbb{Z}^d$ ; and thus, we need in particular the maximum principle, cf. Lemma 3. In fact, Lemma 3 and Lemma 14 stated below are the only places where we appeal to the diagonality of  $\mathbf{a}$ . We believe that our arguments extend to a suitable subclass of symmetric, uniformly elliptic coefficients that allow for a weak maximum principle.

**Remark 11.** For the application to stochastic homogenization we consider Theorem 2 with initial values of the form

$$\mathfrak{d} := -D^* \mathbf{a}(0) e, \quad e \in \mathbb{R}^d, |e| = 1.$$

Note that  $\mathfrak{d}$  is precisely the right-hand side of (3) – the corrector problem of stochastic homogenization. Note that for  $1 \leq p < \infty$  we have

$$\|\mathfrak{d}\|_{\ell_y^1 L_{\langle \cdot \rangle}^{2p}} = \langle |\mathbf{a}(0) - \langle \mathbf{a}(0) \rangle_0|^{2p} \rangle^{\frac{1}{2p}} \leq 1 - \lambda.$$

Hence, Theorem 2 yields  $\langle |\exp(-tD^* \mathbf{a}(0)D)\mathfrak{d}|^{2p} \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})}$  for  $p_0 \leq p < \infty$ . Combined with Jensen's inequality w. r. t.  $\langle \cdot \rangle$  we obtain the central estimate

$$(70) \quad \forall 1 \leq q < \infty : \quad \langle |\exp(-tD^* \mathbf{a}(0)D)\mathfrak{d}|^q \rangle^{\frac{1}{q}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})}.$$

We split the proof of Theorem 2 into several lemmas. Most of the proofs are postponed to Section 5.1. The starting point is the spectral gap estimate. Since we have to estimate higher moments, we need the following version of (SG):

**Lemma 11** ( $p$ -version of (SG)). *Let  $\langle \cdot \rangle$  either satisfy (69a) or (69b). Then for  $p \in \mathbb{N}$  and any  $\langle \cdot \rangle$ -measurable  $\zeta$  with  $\langle \zeta \rangle = 0$  we have*

$$\langle \zeta^{2p} \rangle^{\frac{1}{2p}} \lesssim \|\zeta\|_{L_{\langle \cdot \rangle}^{2p} \ell_y^2},$$

where

$$\|\zeta\|_{L_{\langle \cdot \rangle}^{2p} \ell_y^2} := \begin{cases} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} & \text{in case of (69a),} \\ \left\langle \left( \sum_{y \in ([0,L] \cap \mathbb{Z})^d} \left( \frac{\partial \zeta}{\partial Ly} \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} & \text{in case of (69b),} \end{cases}$$

and  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p$  and  $\rho$ .

In the course of proving Theorem 2 via Lemma 11, we have to estimate the vertical derivative  $\frac{\partial u}{\partial y}$ . In order to do so, we go back to physical space by considering the stationary extension  $\bar{u}$  of  $u$ . As we observed in Section 3,  $\bar{u}$  is a solution of a variable-coefficient parabolic equation on  $\mathbb{Z}^d$ . Taking the vertical derivative of that equation leads to a characterization of the non-stationary random field  $\frac{\partial \bar{u}(t, \mathbf{a}, x)}{\partial y}$  as the solution to a parabolic equation (with time variable  $t$  and space variable  $x$ -space) for some specific right-hand side  $\xi_1(t, \mathbf{a}, x, y)$  and initial value  $\xi_0(\mathbf{a}, x, y)$ :

$$\begin{cases} \partial_t \frac{\partial \bar{u}(t, \mathbf{a}, x)}{\partial y} + \nabla_x^* \mathbf{a}(x) \nabla_x \frac{\partial \bar{u}(t, \mathbf{a}, x)}{\partial y} = \nabla^* \xi_1(t, \mathbf{a}, x, y) & t > 0, x \in \mathbb{Z}^d, \\ \frac{\partial \bar{u}(t=0, \mathbf{a}, x)}{\partial y} = \nabla^* \xi_0(\mathbf{a}, x, y) & x \in \mathbb{Z}^d. \end{cases}$$

Note that we drop from now on the dependency of  $\bar{u}$ ,  $\xi_1$  and  $\xi_0$  on  $\mathbf{a}$  in the notation. By appealing to Duhamel's formula, cf. Lemma 2, and integration by parts, we get the representation

$$(71) \quad \frac{\partial \bar{u}(t, x)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, x, z) \cdot \xi_0(z, y) + \int_0^t \sum_{z \in \mathbb{Z}^d} \nabla_z G(t-s, x, z) \cdot \xi_1(s, z, y) ds.$$

On a formal level, we easily get explicit formulæ for  $\xi_1$  and  $\xi_0$ : Indeed, if the Leibniz rule for the vertical derivative were valid, we would get

$$\xi_1(t, x, y) \stackrel{\text{formally}}{=} -\frac{\partial \mathbf{a}(x)}{\partial y} \nabla \bar{u}(t, x), \quad \xi_0(t, x, y) = \frac{\partial \bar{\xi}(x)}{\partial y}.$$

Since  $\frac{\partial \mathbf{a}(x)}{\partial y}$  vanishes for  $x \neq y$ , the representation (71) would simplify to

$$\frac{\partial \bar{u}(t, x)}{\partial y} \stackrel{\text{formally}}{=} \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, x, z) \cdot \frac{\partial \bar{\xi}(z)}{\partial y} - \int_0^t \nabla_y G(t-s, x, y) \cdot \frac{\partial \mathbf{a}(y)}{\partial y} \nabla \bar{u}(s, y) ds.$$

However, for a rigorous argument the above identity has to be corrected:

**Lemma 12.** *Let  $\langle \cdot \rangle$  either satisfy (69a) or (69b). Consider*

$$u(t) := \exp(-tD^* \mathbf{a}(0)D) D^* \xi, \quad \xi \in C^0(\Omega)^d.$$

*In the case of (69a) we have*

$$(72) \quad \frac{\partial u(t)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, 0, z) \cdot \overline{\left( \frac{\partial \xi}{\partial(y-z)} \right)}(z) + \int_0^t \nabla_y G(t-s, 0, y) \cdot \bar{g}(s, y) ds$$

*$\langle \cdot \rangle$ -almost surely, where  $g \in C(\mathbb{R}_+, L^\infty(\Omega))$  satisfies*

$$(73) \quad \langle |g(t)|^{2p} \rangle^{\frac{1}{2p}} \leq 2 \langle |Du(t)|^{2p} \rangle^{\frac{1}{2p}}$$

*for all  $p < \infty$ , and  $\overline{(\cdot)}$  stands for stationary extension.*

*In the  $L$ -periodic case, i. e. when  $\langle \cdot \rangle$  satisfies (69b), the statement remains valid for  $\sum_{y \in \mathbb{Z}^d} \frac{\partial}{\partial(x-y)}$  and  $G$  replaced by  $\sum_{z \in ([0, L] \cap \mathbb{Z})^d} \frac{\partial}{\partial_L(y-z)}$  and  $G_L$ , respectively.*

Note the differences between the constant-coefficient formula (61) and the variable-coefficient formula (72):

- the constant-coefficient Green's function is replaced by the variable-coefficient Green's function,
- there is an additional “non-linear” term.

Lemma 12 indicates that decay estimates on  $\nabla G$  in  $t$  in the spirit of (54) are required. For our purpose we need estimates that are uniform in  $\mathbf{a}$ , but nevertheless, are optimal in terms of the exponent in  $t$ . By optimal we mean that the exponent should be identical to the one for the constant-coefficient Green's function. As a consequence, these estimates cannot be pointwise in  $x$ , but rely on integral estimates like for (54). In order to treat the nonlinear term it turns out that we need to capture the decay in  $x$  in a better way than (54) does. We do this by establishing *weighted* integral estimates with weight functions

$$(74) \quad \omega(t, x) := \left( \frac{|x|^2}{t+1} + 1 \right)^{\frac{1}{2}}, \quad \omega_L(t, x) := \left( \frac{\text{dist}^2(x, LZ^d)}{t+1} + 1 \right)^{\frac{1}{2}}$$

where  $\text{dist}(x, LZ^d) := \min_{z \in \mathbb{Z}^d} |x - Lz|$  denotes the distance to 0 on the  $L$ -torus. Finally, in order to treat the nonlinear term, it turns out that we need a slightly stronger estimate than the *square* integral estimate.

**Theorem 3.** *There exists an exponent  $q > 1$  (only depending on  $\lambda$  and  $d$  and that we may choose as close to 1 as we please) such that for any weight exponent  $\alpha < \infty$  we have:*

(a) for all  $\mathbf{a} \in \Omega$

$$\left( \sum_{x \in \mathbb{Z}^d} \left( \omega^\alpha(t, x-y) |\nabla_x G(t, \mathbf{a}, x, y)| \right)^{2q} \right)^{\frac{1}{2q}} \lesssim (t+1)^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2q}},$$

(b) for all  $L \in \mathbb{N}$  and  $\mathbf{a} \in \Omega_L$

$$\left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \left( \omega_L^\alpha(t, x-y) |\nabla_x G_L(t, \mathbf{a}, x, y)| \right)^{2q} \right)^{\frac{1}{2q}} \lesssim (t+1)^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2q}} \exp(-c_0 \frac{t}{L^2}),$$

where  $c_0 > 0$  denotes a constant that only depends on  $\lambda$  and  $d$ .

Above  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\alpha, q, \lambda$  and  $d$ .

The proof of the theorem is presented in Section 7. As we shall see in the proof of Lemma 13, the fact that we obtain optimal decay for  $\nabla G$  in an average that is slightly better than a square average is crucial. In fact,  $p$  in the statement of Theorem 2 will be the dual exponent to  $q$ . This explains why we are forced to estimate high moments of  $u$  even if ultimately we are mostly interested in the second and fourth moment, cf Section 6.

**Remark 12.** 1. Since the constant-coefficient discrete Green's function is on large scales and for large times close to the constant-coefficient continuum Green's function, we see by a scaling argument that

$$\left( \sum_{x \in \mathbb{Z}^d} \left( \omega^\alpha(t, x-y) |\nabla_x G(t, \mathbf{id}, x, y)| \right)^{2q} \right)^{\frac{1}{2q}} \sim (t+1)^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2q}}.$$

Hence the estimate of Theorem 3 is optimal in terms of scaling.

2. We recall that Proposition 2 showed that the rate of decay of the constant-coefficient semigroup for divergence form initial data is determined by the constant-coefficient parabolic Green's function via

$$(75) \quad \left( \sum_{x \in \mathbb{Z}^d} |\nabla G(t, \mathbf{id}, x, y)|^2 \right)^{\frac{1}{2}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})},$$

see Remark 7.

3. The estimate of Theorem 3 is the stronger the higher the exponents  $\alpha$  and  $q$  are chosen. In particular, Theorem 3 contains (75) in the sense that the estimate of Theorem 3 (a) also holds for  $q = 1$  and  $\alpha = 0$  as can be seen as follows:

$$\begin{aligned} \left( \sum_{x \in \mathbb{Z}^d} |\nabla G(t, \mathbf{a}, x, y)|^2 \right)^{\frac{1}{2}} &= \left( \sum_{x \in \mathbb{Z}^d} (\omega^\alpha(t, x-y) |\nabla G(t, \mathbf{a}, x, y)|)^2 \omega^{-2\alpha}(t, x-y) \right)^{\frac{1}{2}} \\ &\stackrel{\text{H\"older}}{\leq} \left( \sum_{x \in \mathbb{Z}^d} (\omega^\alpha(t, x-y) |\nabla G(t, \mathbf{a}, x, y)|)^{\frac{2p}{p-1}} \right)^{\frac{p-1}{2p}} \left( \sum_{x \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, x-y) \right)^{\frac{1}{2p}}. \end{aligned}$$



For  $2p\alpha > d$  we have

$$\begin{aligned}
\sum_{x \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, x-y) &= \sum_{z \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, z) = \sum_{z \in \mathbb{Z}^d} \left( \frac{|z|^2}{t+1} + 1 \right)^{-p\alpha} = \sum_{z \in \mathbb{Z}^d} \left( \frac{t+1}{|z|^2 + t+1} \right)^{p\alpha} \\
(76) \quad &\sim \int_{\mathbb{R}^d} \left( \frac{t+1}{|z|^2 + t+1} \right)^{p\alpha} dz = (t+1)^{\frac{d}{2}} \int_{\mathbb{R}^d} \left( \frac{1}{|y|^2 + 1} \right)^{p\alpha} dy \\
&\stackrel{2p\alpha > d}{\sim} (t+1)^{\frac{d}{2}}.
\end{aligned}$$

In combination with Theorem 3 we obtain

$$\left( \sum_{x \in \mathbb{Z}^d} |\nabla G(t, \mathbf{a}, x, y)|^2 \right)^{\frac{1}{2}} \lesssim (t+1)^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{p-1}{2p}} \times (t+1)^{\frac{d}{2} \frac{1}{2p}} = (t+1)^{-(\frac{d}{4} + \frac{1}{2})}.$$

Combined with the spectral gap estimate in its  $p$ -version and the representation formula of Lemma 12 we get:

**Lemma 13.** *In the situation of Theorem 2 there exists an exponent  $1 \leq p_0 < \infty$  (only depending on  $\lambda$  and  $d$ ), such that for all  $p_0 \leq p < \infty$  we have*

$$\langle u^{2p}(t) \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})} \|\xi\|_{\ell_y^1 L_{(\cdot)}^{2p}} + \int_0^t (t-s+1)^{-(\frac{d}{4} + \frac{1}{2})} \langle |Du(s)|^{2p} \rangle^{\frac{1}{2p}} ds,$$

where  $\|\xi\|_{\ell_y^1 L_{(\cdot)}^{2p}}$  is defined as in Theorem 2 and  $\lesssim$  means  $\leq$  up to a constant that only depends on  $p$ ,  $\rho$ ,  $\lambda$  and  $d$ .

To complete the argument of Theorem 2 we have to gain control over the “non-linear term”  $\int_0^t (t-s+1)^{-(\frac{d}{4} + \frac{1}{2})} \langle |Du(s)|^{2p} \rangle^{\frac{1}{2p}} ds$ . This is done by appealing to a Caccioppoli estimate in probability and an ODE-argument. The following lemma shows that, at least on average,  $\langle |Du(t)|^{2p} \rangle$  has better decay than  $\langle u^{2p}(t) \rangle$  so that there is hope to absorb the “non-linear term” in to the right-hand side.

**Lemma 14** (Caccioppoli). *Let  $\langle \cdot \rangle$  be stationary. For measurable  $\zeta$  consider  $u(t) = \exp(-tD^* \mathbf{a}(0)D)\zeta$ . Then for all  $p \in \mathbb{N}$  we have*

$$\langle |Du(t)|^{2p} \rangle \lesssim -\frac{d}{dt} \langle u^{2p}(t) \rangle$$

where  $\lesssim$  means  $\leq$  up to constant that depends on  $d$  and  $p$ .

The following lemma shows how the absorption of the “non-linear term” is implemented.

**Lemma 15** (ODE-argument). *Let  $1 \leq p, \gamma < \infty$  and  $a(t), b(t) \geq 0$ . Suppose that*

$$(77a) \quad a(t) \leq C_1 \left( (t+1)^{-\gamma} + \int_0^t (t-s+1)^{-\gamma} b(s) ds \right),$$

$$(77b) \quad b^p(t) \leq C_1 \left( -\frac{d}{dt} a^p(t) \right),$$

for some constant  $C_1 < \infty$ . Then

$$a(t) \leq C_2 (t+1)^{-\gamma}$$

where  $C_2 < \infty$  only depends on  $C_1$ ,  $p$  and  $\gamma$ .

*Proof of Theorem 2.* Let  $p_0$  be given by Lemma 13 and fix an exponent  $p_0 \leq p < \infty$ . By homogeneity we may assume without loss of generality that  $\|\xi\|_{\ell_y^1 L_{(\cdot)}^{2p}} = 1$ , so that the asserted estimate reduces to

$$(78) \quad \langle u^{2p}(t) \rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})}.$$

Set

$$a(t) := \langle u^{2p}(t) \rangle^{\frac{1}{2p}}, \quad b(t) := \langle |Du(t)|^{2p} \rangle^{\frac{1}{2p}}, \quad \gamma = \frac{d}{4} + \frac{1}{2}.$$

By Lemma 13 and Lemma 14 we have

$$\begin{aligned} a(t) &\lesssim (t+1)^{-\gamma} + \int_0^t (t-s+1)^{-\gamma} b(s) ds, \\ b^{2p}(t) &\lesssim -\frac{d}{dt} a^{2p}(t). \end{aligned}$$

Hence, Lemma 15 yields (78) in form of  $a(t) \lesssim (t+1)^{-(\frac{d}{4}+\frac{1}{2})}$ .  $\square$

## 5.1 Proofs

*Proof of Lemma 11.* We prove the lemma only in the case (69a), since the argument in the periodic case, i. e. when  $\langle \cdot \rangle$  satisfies (69b), is identical. By an approximation argument left to the reader, it suffices to consider  $\zeta \in C^0(\Omega)$ . The technical difficulty in the argument originates from the failure of the Leibniz rule

$$(79) \quad \frac{\partial \zeta^p}{\partial y} = p \zeta^{p-1} \frac{\partial \zeta}{\partial y}.$$

**Step 1.** Substitute for the Leibniz rule (79):

We claim that for positive integers  $p$

$$(80) \quad \langle \left( \frac{\partial \zeta^p}{\partial y} \right)^2 \rangle \lesssim \langle \zeta^{2(p-1)} \left( \frac{\partial \zeta}{\partial y} \right)^2 \rangle + \langle \left( \frac{\partial \zeta}{\partial y} \right)^{2p} \rangle$$

up to a multiplicative constant that only depends on  $p$  and  $d$ .

Here comes the argument: By definition of  $\frac{\partial}{\partial y}$  we need to show

$$\langle (\zeta^p - \langle \zeta^p \rangle_y)^2 \rangle \lesssim \langle \zeta^{2(p-1)} (\zeta - \langle \zeta \rangle_y)^2 \rangle + \langle (\zeta - \langle \zeta \rangle_y)^{2p} \rangle.$$

Since  $\langle \langle \cdot \rangle_y \rangle = \langle \cdot \rangle$  (by general properties of conditional expectations), it suffices to show

$$\langle (\zeta^p - \langle \zeta^p \rangle_y)^2 \rangle_y \lesssim \left\langle \zeta^{2(p-1)} (\zeta - \langle \zeta \rangle_y)^2 + (\zeta - \langle \zeta \rangle_y)^{2p} \right\rangle_y \quad \langle \cdot \rangle\text{-almost surely.}$$

Since the conditional expectation is an orthogonal projection in  $L^2$ , we have

$$\forall \tilde{\zeta} \in L^2(\Omega) : \quad \langle (\tilde{\zeta} - \langle \tilde{\zeta} \rangle_y)^2 \rangle_y \leq \min_{c \in \mathbb{R}} \langle (\tilde{\zeta} - c)^2 \rangle_y \quad \langle \cdot \rangle\text{-almost surely.}$$

In particular, with  $\tilde{\zeta} = \zeta^p$  and  $c = \langle \zeta \rangle_y^p$  we get  $\langle (\zeta^p - \langle \zeta^p \rangle_y)^2 \rangle_y \leq \langle (\zeta^p - \langle \zeta \rangle_y^p)^2 \rangle_y$   $\langle \cdot \rangle$ -almost surely.

Hence, it suffices to argue that

$$\langle (\zeta^p - \langle \zeta \rangle_y^p)^2 \rangle_y \lesssim \left\langle \zeta^{2(p-1)} (\zeta - \langle \zeta \rangle_y)^2 + (\zeta - \langle \zeta \rangle_y)^{2p} \right\rangle_y \quad \langle \cdot \rangle\text{-almost surely.}$$

The latter follows from the elementary inequality

$$\forall a, b \in \mathbb{R} : \quad (a^p - b^p)^2 \lesssim a^{2(p-1)}(a-b)^2 + (a-b)^{2p},$$

which by scaling can be reduced to  $|b| = 1$  and by symmetry to  $b = 1$ :

$$\forall a \in \mathbb{R} : \quad (a^p - 1)^2 \lesssim a^{2(p-1)}(a-1)^2 + (a-1)^{2p}.$$

The latter is a consequence of the following two elementary inequalities:

$$\begin{aligned} (a^p - 1)^2 &\lesssim a^{2p} + 1 \lesssim (a-1)^{2p} \quad \text{for } |a-1| \geq \frac{1}{2}, \\ (a^p - 1)^2 &\lesssim (a-1)^2 \lesssim a^{2(p-1)}(a-1)^2 \quad \text{for } |a-1| \leq \frac{1}{2}. \end{aligned}$$

**Step 2.** Application of the spectral gap estimate.

For  $p \in \mathbb{N}$  we claim that

$$\langle \zeta^{2p} \rangle \lesssim \langle \zeta^2 \rangle^p + \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right)^p \right\rangle.$$

Since for  $p = 1$  the statement turns into the original version of the spectral gap estimate, we only consider  $p \geq 2$ . The application of  $\text{SG}_\infty(\rho)$  to  $\zeta^p - \langle \zeta^p \rangle$  yields

$$\langle (\zeta^p - \langle \zeta^p \rangle)^2 \rangle \leq \frac{1}{\rho} \sum_{y \in \mathbb{Z}^d} \langle \left( \frac{\partial \zeta}{\partial y} \zeta^p \right)^2 \rangle.$$

With Step 1 and the triangle inequality, this inequality turns into

$$(81) \quad \langle \zeta^{2p} \rangle \lesssim \langle \zeta^p \rangle^2 + \left\langle \zeta^{2(p-1)} \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 + \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^{2p} \right\rangle.$$

We treat each of the three terms on the right-hand side of (81) separately. For the third term we appeal to the discrete  $\ell^{2p}$ - $\ell^2$  estimate

$$\left\langle \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^{2p} \right\rangle \leq \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right)^p \right\rangle.$$

For the second term on the right-hand side in (81) we use Hölder's inequality with exponents  $(\frac{p}{p-1}, p)$ :

$$\langle \zeta^{2(p-1)} \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \rangle \leq \langle \zeta^{2p} \rangle^{\frac{p-1}{p}} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right)^p \right\rangle^{\frac{1}{p}}.$$

We then may use Young's inequality with exponent  $(\frac{p}{p-1}, p)$  to absorb the first factor into the left-hand side of (81). We turn to the first term. For  $p = 2$  there is nothing to do, whereas for  $p > 2$  we apply Hölder's inequality with exponents  $(2\frac{p-1}{p-2}, 2\frac{p-1}{p})$  to  $\langle |\zeta|^p \rangle = \langle |\zeta|^{p\frac{p-2}{p-1}} |\zeta|^{\frac{p}{p-1}} \rangle$ :

$$\langle |\zeta|^p \rangle^2 \leq \langle \zeta^{2p} \rangle^{\frac{p-2}{p-1}} \langle \zeta^2 \rangle^{\frac{p}{p-1}}.$$

Again, we then may use Young's inequality with exponent  $(\frac{p-1}{p-2}, p-1)$  to absorb the first factor into the l. h. s. of (81). The claim now follows from the combination of these four estimates.

**Step 3.** Conclusion.

The application of  $\text{SG}_\infty(\rho)$  to  $\zeta$  combined with the assumption  $\langle \zeta \rangle = 0$  yields

$$\langle \zeta^2 \rangle \leq \frac{1}{\rho} \left\langle \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right\rangle.$$

Hence, by Jensen's inequality in probability we get

$$\langle \zeta^2 \rangle^p \leq \rho^{-p} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial \zeta}{\partial y} \right)^2 \right)^p \right\rangle$$

and the statement of Lemma 11 follows from Step 2.  $\square$

*Proof of Lemma 12.* We only prove the statement in the case of the infinite ensemble, i. e. when  $\langle \cdot \rangle$  satisfies (69a). The argument for periodic ensembles is similar.

By Lemma 1 the stationary extension  $\bar{u}(t, \cdot)$  of  $u$  satisfies (for all  $\mathbf{a} \in \Omega$ ) the spatial parabolic equation

$$(82) \quad \begin{cases} \partial_t \bar{u}(t, x) + \nabla^* \mathbf{a}(x) \nabla \bar{u}(t, x) = 0 & \text{for all } t > 0, x \in \mathbb{Z}^d, \\ \bar{u}(t = 0, x) = \nabla^* \bar{\xi}(x) & \text{for all } x \in \mathbb{Z}^d. \end{cases}$$

We take the vertical derivative  $\frac{\partial}{\partial y}$  of this equation. Using that the vertical derivative  $\frac{\partial}{\partial y}$  and the spatial derivative  $\nabla^*$  commute we arrive at

$$(83) \quad \begin{cases} \partial_t \frac{\partial \bar{u}(t, x)}{\partial y} + \nabla^* \mathbf{a}(x) \nabla \frac{\partial \bar{u}(t, x)}{\partial y} = \nabla^* \xi_1(t, x, y) & \text{for all } t > 0, x \in \mathbb{Z}^d, \\ \frac{\partial \bar{u}(t = 0, x)}{\partial y} = \nabla^* \xi_0(x, y) & \text{for all } x \in \mathbb{Z}^d, \end{cases}$$

where

$$\begin{aligned} \xi_1(t, x, y) &:= \mathbf{a}(x) \nabla \frac{\partial \bar{u}(t, x)}{\partial y} - \frac{\partial}{\partial y} (\mathbf{a}(x) \nabla \bar{u}(t, x)), \\ \xi_0(x, y) &:= \frac{\partial \bar{\xi}(x)}{\partial y}. \end{aligned}$$

Duhamel's formula, cf. Lemma 2, and two integrations by parts yield

$$(84) \quad \frac{\partial \bar{u}(t, 0)}{\partial y} = \sum_{z \in \mathbb{Z}^d} \nabla_z G(t, 0, z) \xi_0(z, y) + \int_0^t \sum_{z \in \mathbb{Z}^d} \nabla_z G(t - s, 0, z) \xi_1(s, z, y) ds.$$

We claim that

$$(85) \quad \xi_0(x, y) = \overline{\left( \frac{\partial \xi}{\partial (y - x)} \right)}(x),$$

$$(86) \quad \xi_1(t, x, y) = \delta(y - x) \bar{g}(t, x),$$

where

$$g(t) := \langle \mathbf{a}(0) Du(t) \rangle_0 - \mathbf{a}(0) \langle Du(t) \rangle_0.$$

Indeed, (85) is an application of (50). We now prove (86): By definition of  $\xi_1$  and the vertical derivative, we have

$$\xi_1(t, x, y) = \langle \mathbf{a}(x) \nabla \bar{u}(t, x) \rangle_y - \mathbf{a}(x) \nabla \langle \bar{u}(t, x) \rangle_y.$$

Identity (86) thus follows from the fact that for all  $y \neq x$  we have

$$\langle \mathbf{a}(x) \nabla \bar{u}(t, x) \rangle_y - \mathbf{a}(x) \langle \nabla \bar{u}(t, x) \rangle_y = 0,$$

and from the properties of the stationary extension:

$$\begin{aligned} & \langle \mathbf{a}(x) \nabla \bar{u}(t, x) \rangle_y - \mathbf{a}(x) \langle \nabla \bar{u}(t, x) \rangle_y \\ & \stackrel{(8)}{=} \overline{\langle \mathbf{a}(0) Du(t)(x) \rangle_y} - \mathbf{a}(x) \overline{\langle Du(t)(x) \rangle_y} \\ & \stackrel{(49)}{=} \delta(y - x) \overline{\langle \mathbf{a}(0) Du(t) \rangle_{y-x} - \mathbf{a}(0) \langle Du(t) \rangle_{y-x}}(x) \\ & = \delta(y - x) \bar{g}(t, x). \end{aligned}$$

Since  $u(t) = \bar{u}(t, 0)$  and thus  $\frac{\partial \bar{u}(t, 0)}{\partial y} = \frac{\partial u(t)}{\partial y}$ , the combination of (84), (85) and (86) yields (72).

It remains to prove (73). By the triangle-inequality in  $L^{2p}(\Omega)$ , the boundedness of  $\mathbf{a}(0)$  by 1 and Jensen's inequality for  $\langle \cdot \rangle_0$ , we get

$$\langle |g(t)|^{2p} \rangle_0^{\frac{1}{2p}} \leq \langle |\langle \mathbf{a}(0) Du(t) \rangle_0|^{2p} \rangle_0^{\frac{1}{2p}} + \langle |\mathbf{a}(0) \langle Du(t) \rangle_0|^{2p} \rangle_0^{\frac{1}{2p}} \leq 2 \langle |Du(t)|^{2p} \rangle_0^{\frac{1}{2p}}.$$

□

*Proof of Lemma 13.* We only prove the statement in the case of the infinite ensemble, i. e. when  $\langle \cdot \rangle$  satisfies (69a). The argument for periodic ensembles is similar.

Since  $\langle D^* \xi \rangle = 0$ , we have  $\langle u(t) \rangle = 0$  for all  $t > 0$ , and therefore Lemma 11 yields

$$\langle u^{2p}(t) \rangle \lesssim \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \frac{\partial u(t)}{\partial y} \right)^2 \right)^p \right\rangle.$$

Using the representation of  $\frac{\partial u(t)}{\partial y}$  derived in Lemma 12, and the triangle inequality w. r. t.  $\langle \left( \sum_{y \in \mathbb{Z}^d} (\cdot)^2 \right)^p \rangle^{\frac{1}{2p}}$ , we obtain

$$(87) \quad \langle u^{2p}(t) \rangle^{\frac{1}{2p}} \leq \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \sum_{z \in \mathbb{Z}^d} |\nabla_z G(t, 0, z)| \left| \overline{\left( \frac{\partial \xi}{\partial(y-z)} \right)}(z) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \\ + \int_0^t \left\langle \left( \sum_{y \in \mathbb{Z}^d} |\nabla_y G(t-s, 0, y)|^2 |\bar{g}(s, y)|^2 \right)^p \right\rangle^{\frac{1}{2p}} ds,$$

where  $\langle |\bar{g}(t, x)|^{2p} \rangle^{\frac{1}{2p}} \leq 2 \langle |Du(t)|^{2p} \rangle^{\frac{1}{2p}}$ . We estimate the first term of the r. h. s. by introducing new variables and appealing to the triangle inequality w. r. t.  $\langle \left( \sum_{y \in \mathbb{Z}^d} (\cdot)^2 \right)^p \rangle^{\frac{1}{2p}}$  as follows:

$$\left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \sum_{z \in \mathbb{Z}^d} |\nabla_z G(t, 0, z)| \left| \overline{\left( \frac{\partial \xi}{\partial(y-z)} \right)}(z) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \\ \stackrel{x:=y-z}{=} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( \sum_{x \in \mathbb{Z}^d} |\nabla_y G(t, 0, y-x)| \left| \overline{\frac{\partial \xi}{\partial x}}(y-x) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \\ \stackrel{\Delta\text{-inequality}}{\leq} \sum_{x \in \mathbb{Z}^d} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( |\nabla_y G(t, 0, y-x)| \left| \overline{\frac{\partial \xi}{\partial x}}(y-x) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \\ \stackrel{x':=y-x}{=} \sum_{x \in \mathbb{Z}^d} \left\langle \left( \sum_{x' \in \mathbb{Z}^d} \left( |\nabla_{x'} G(t, 0, x')| \left| \overline{\frac{\partial \xi}{\partial x}}(x') \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}}.$$

Hence, (87) turns into

$$\langle u^{2p}(t) \rangle^{\frac{1}{2p}} \leq \sum_{x \in \mathbb{Z}^d} \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( |\nabla_y G(t, 0, y)| \left| \overline{\frac{\partial \xi}{\partial x}}(y) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \\ + \int_0^t \left\langle \left( \sum_{y \in \mathbb{Z}^d} |\nabla_y G(t-s, 0, y)|^2 |\bar{g}(s, y)|^2 \right)^p \right\rangle^{\frac{1}{2p}} ds$$

and it thus remains to show that

$$(88) \quad \left\langle \left( \sum_{y \in \mathbb{Z}^d} \left( |\nabla_y G(t, 0, y)| \left| \overline{\frac{\partial \xi}{\partial x}}(y) \right| \right)^2 \right)^p \right\rangle^{\frac{1}{2p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})} \left\langle \left| \frac{\partial \xi}{\partial x} \right|^{2p} \right\rangle^{\frac{1}{2p}},$$

$$(89) \quad \left\langle \left( \sum_{y \in \mathbb{Z}^d} |\nabla_y G(t-s, 0, y)|^2 |\bar{g}(s, y)|^2 \right)^p \right\rangle^{\frac{1}{2p}} \lesssim (t-s+1)^{-(\frac{d}{4} + \frac{1}{2})} \langle |Du(s)|^{2p} \rangle^{\frac{1}{2p}}.$$

We only give the argument for (88), the argument for (89) being similar. Let  $q := \frac{p}{p-1}$  denote the dual exponent to  $p$ , let  $\alpha > 0$  denote some exponent to be fixed later, and let  $\omega(t, x)$  denote the weight defined in (74). By Hölder's inequality with exponents  $(q, p)$  we have

$$\begin{aligned} \sum_{y \in \mathbb{Z}^d} \left( |\nabla_y G(t, 0, y)| \left| \frac{\partial \bar{\xi}}{\partial x}(y) \right| \right)^2 &\leq \left( \sum_{x \in \mathbb{Z}^d} (\omega^\alpha(t, y) |\nabla_y G(t, 0, y)|)^{2q} \right)^{\frac{1}{q}} \left( \sum_{y \in \mathbb{Z}^d} (\omega^{-\alpha}(t, y) \left| \frac{\partial \bar{\xi}}{\partial x}(y) \right|)^{2p} \right)^{\frac{1}{p}} \\ &\stackrel{(26e)}{=} \left( \sum_{y \in \mathbb{Z}^d} (\omega^\alpha(t, y) |\nabla_y G(t, y, 0)|)^{2q} \right)^{\frac{1}{q}} \left( \sum_{y \in \mathbb{Z}^d} (\omega^{-\alpha}(t, y) \left| \frac{\partial \bar{\xi}}{\partial x}(y) \right|)^{2p} \right)^{\frac{1}{p}}. \end{aligned}$$

Hence,

$$\begin{aligned} [\text{L. H. S. of (88)}] &\leq \left( \sup_{\mathbf{a} \in \Omega} \sum_{y \in \mathbb{Z}^d} (\omega^\alpha(t, y) |\nabla_y G(t, y, 0)|)^{2q} \right)^{\frac{1}{2q}} \\ &\quad \times \sum_{x \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, y) \left\langle \left| \frac{\partial \bar{\xi}}{\partial x}(y) \right|^{2p} \right\rangle \right)^{\frac{1}{2p}}. \end{aligned}$$

By stationarity  $\langle \left| \frac{\partial \bar{\xi}}{\partial x}(y) \right|^{2p} \rangle$  is independent of  $y$ , so that the second factor in the product on the right-hand side simplifies to  $\sum_{x \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \bar{\xi}}{\partial x} \right|^{2p} \right\rangle^{\frac{1}{2p}} \left( \sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, y) \right)^{\frac{1}{2p}}$ , and we conclude that

$$\begin{aligned} (90) \quad [\text{L. H. S. of (88)}] &\leq \sup_{\mathbf{a} \in \Omega} \left( \sum_{x' \in \mathbb{Z}^d} (\omega^\alpha(t, y) |\nabla_y G(t, y, 0)|)^{2q} \right)^{\frac{1}{2q}} \left( \sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, y) \right)^{\frac{1}{2p}} \\ &\quad \times \sum_{x \in \mathbb{Z}^d} \left\langle \left| \frac{\partial \bar{\xi}}{\partial x} \right|^{2p} \right\rangle^{\frac{1}{2p}}. \end{aligned}$$

We now address the choice of  $p$  and  $\alpha$ : First  $p$  is chosen so large that its dual exponent  $q$  is within the range of Theorem 3. Second,  $\alpha$  is chosen so large that  $2p\alpha > d$ . Then we have

$$\begin{aligned} \sup_{\mathbf{a} \in \Omega} \left( \sum_{y \in \mathbb{Z}^d} (\omega^\alpha(t, y) |\nabla_y G(t, y, 0)|)^{2q} \right)^{\frac{1}{2q}} &\lesssim (t+1)^{-\left(\frac{d}{2} + \frac{1}{2}\right) + \frac{d}{2} \frac{1}{2q}} \quad \text{by Theorem 3, and} \\ \left( \sum_{y \in \mathbb{Z}^d} \omega^{-2p\alpha}(t, y) \right)^{\frac{1}{2p}} &\lesssim (t+1)^{\frac{d}{2} \frac{1}{2p}} \quad \text{by (76)}. \end{aligned}$$

In combination with (90) the desired estimate (88) follows.  $\square$

*Proof of Lemma 14.* By an approximation argument we may assume that  $\zeta \in C^0(\Omega)$ . In this case  $u(t) = \exp(-tD^* \mathbf{a}(0)D)\zeta$  belongs to  $C^\infty(\mathbb{R}_+, C^0(\Omega))$ .

**Step 1.** Caccioppoli's inequality.

We claim that for all  $t > 0$  and  $p \in \mathbb{N}$

$$\lambda \langle Du^{2p-1} \cdot Du \rangle \leq -\frac{d}{dt} \frac{1}{2p} \langle u^{2p} \rangle.$$

Indeed, since  $u(t) = \exp(-tD^* \mathbf{a}(0)D)\zeta$ , we have

$$(91) \quad \frac{d}{dt} u + D^* \mathbf{a}(0)D u = 0 \quad t > 0.$$

Hence, multiplying by  $u^{2p-1}$ , taking the expectation, and integrating by parts yield

$$\frac{d}{dt} \frac{1}{2p} \langle u^{2p} \rangle = \langle u^{2p-1} \frac{d}{dt} u \rangle \stackrel{(91)}{=} -\langle Du^{2p-1} \cdot \mathbf{a}(0) Du \rangle.$$

Recalling that  $\langle \cdot \rangle$  is diagonal, and that  $\mathbf{a}(0) \geq \lambda \text{id}$ , we have

$$Du^{2p-1} \cdot \mathbf{a}(0) Du = \sum_{i=1}^d \mathbf{a}_{ii}(0) D_i u^{2p-1} D_i u \geq \lambda Du^{2p-1} \cdot Du,$$

due to the elementary inequality  $(a^{2p-1} - b^{2p-1})(a - b) \geq 0$  for all  $a, b \in \mathbb{R}$  and the definition of  $D$ . This proves the claim.

**Step 2.** Conclusion using discreteness.

On the one hand, discreteness makes the argument more complicated because Leibniz' rule

$$(92) \quad Du^{2p-1} \cdot Du = (2p-1)u^{2(p-1)}|Du|^2$$

does not hold. On the other hand, discreteness is crucial because only then we have

$$(93) \quad u^{2(p-1)}|Du|^2 \gtrsim |Du|^{2p}.$$

Instead of finding a proxy for (92), we establish

$$Du^{2p-1} \cdot Du \gtrsim |Du|^{2p},$$

in one step. Here  $\gtrsim$  means  $\geq$  up to a multiplicative constant that only depends on  $p$ . By arguing componentwise, this estimate reduces to

$$(94) \quad (a^{2p-1} - b^{2p-1})(a - b) \gtrsim (a - b)^{2p} \quad \text{for all } a, b \in \mathbb{R}.$$

By scaling and symmetry, this estimate can be reduced to  $b = 1$ :

$$(a^{2p-1} - 1)(a - 1) \gtrsim (a - 1)^{2p}.$$

The latter is a consequence of the following two elementary inequalities:

$$\begin{aligned} (a^{2p-1} - 1)(a - 1) &\sim (a - 1)^2 \gtrsim (a - 1)^{2p} && \text{for } |a - 1| \leq \frac{1}{2}, \\ (a^{2p-1} - 1)(a - 1) &\gtrsim (|a|^{2p-1} + 1)(|a| + 1) \gtrsim a^{2p} + 1 \gtrsim (a - 1)^{2p} && \text{for } |a - 1| \geq \frac{1}{2}. \end{aligned}$$

□

*Proof of Lemma 15.* Our claim can be reformulated as

$$(95) \quad \Lambda(t) := \sup_{0 \leq s \leq t} (1 + s)^\gamma a(s) \lesssim 1.$$

**Step 1.** Two auxiliary estimates.

We claim that

$$(96) \quad \int_{\tau_1}^{\tau_2} b(s) ds \lesssim \begin{cases} (\tau_2 - \tau_1)^{1-\frac{1}{p}} a(\tau_1) & \text{for all } 0 \leq \tau_1 \leq \tau_2, \\ \tau_1^{1-\gamma-\frac{1}{p}} \Lambda(\tau_2) & \text{for all } 1 \leq \tau_1 \leq \tau_2. \end{cases}$$

The first estimate follows from Hölder's inequality, equation (77b), and the non-negativity of  $a$ :

$$(97) \quad \int_{\tau_1}^{\tau_2} b(s) ds \leq (\tau_2 - \tau_1)^{1-\frac{1}{p}} \left( - \int_{\tau_1}^{\tau_2} \frac{d}{dt} a^p(t) dt \right)^{\frac{1}{p}} \leq (\tau_2 - \tau_1)^{1-\frac{1}{p}} a(\tau_1).$$

The second inequality can be deduced from the first one as follows: Let  $N \in \mathbb{N}$  satisfy  $2^{N-1}\tau_1 < \tau_2 \leq 2^N\tau_1$ . We then have

$$\begin{aligned}
\int_{\tau_1}^{\tau_2} b(s) ds &\leq \sum_{n=0}^{N-1} \int_{2^n\tau_1}^{2^{n+1}\tau_1} b(s) ds \stackrel{(97)}{\lesssim} \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}} a(2^n\tau_1) \\
&\leq \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}} (1+2^n\tau_1)^{-\gamma} \Lambda(2^n\tau_1) \\
&\stackrel{\Lambda \text{ monotone}}{\leq} \Lambda(\tau_2) \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\frac{1}{p}} (1+2^n\tau_1)^{-\gamma} \\
&\leq \Lambda(\tau_2) \sum_{n=0}^{N-1} (2^n\tau_1)^{1-\gamma-\frac{1}{p}} \stackrel{\gamma \geq 1}{\lesssim} \Lambda(\tau_2) \tau_1^{1-\gamma-\frac{1}{p}}.
\end{aligned}$$

**Step 2.** A threshold estimate.

Let  $1 \leq \tau \leq \frac{1}{4}t$ . We claim that

$$(t+1)^\gamma a(t) \lesssim 1 + \tau^{1-\frac{1}{p}} + \left( \tau^{1-\gamma-\frac{1}{p}} + \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}} \right) \Lambda(t).$$

First notice that by (77b) the function  $a(\cdot)$  is non-increasing. Hence,

$$\begin{aligned}
(98) \quad a(t) &\leq \frac{1}{2t} \int_{\frac{t}{2}}^t a(t') dt' \\
&\stackrel{(77a)}{\lesssim} \frac{1}{2t} \int_{\frac{t}{2}}^t (t'+1)^{-\gamma} dt' + \frac{1}{2t} \int_{\frac{t}{2}}^t \int_0^{t'} (t'-s+1)^{-\gamma} b(s) ds dt'.
\end{aligned}$$

Clearly, the first term of the right-hand side is estimated by

$$(99) \quad \frac{1}{2t} \int_{\frac{t}{2}}^t (t'+1)^{-\gamma} dt' \lesssim (t+1)^{-\gamma}.$$

For the second term of the right-hand side of (98), we split the inner integral into three contributions that we estimate separately. More precisely, we shall prove that

$$(100) \quad \frac{1}{2t} \int_{\frac{t}{2}}^t \int_0^\tau (t'-s+1)^{-\gamma} b(s) ds dt' \lesssim (t+1)^{-\gamma} \tau^{1-\frac{1}{p}},$$

$$(101) \quad \frac{1}{2t} \int_{\frac{t}{2}}^t \int_\tau^{\frac{t'}{2}} (t'-s+1)^{-\gamma} b(s) ds dt' \lesssim (t+1)^{-\gamma} \tau^{1-\gamma-\frac{1}{p}} \Lambda(t),$$

$$(102) \quad \frac{1}{2t} \int_{\frac{t}{2}}^t \int_{\frac{t'}{2}}^{t'} (t'-s+1)^{-\gamma} b(s) ds dt' \lesssim (t+1)^{-\gamma} \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}} \Lambda(t).$$

Argument for (100): Since  $\tau \leq \frac{t}{4} \leq \frac{t'}{2}$ , we have  $t'-s+1 \geq \frac{t'}{2} + 1$ . Hence,

$$\int_0^\tau (t'-s+1)^{-\gamma} b(s) ds \lesssim (t'+1)^{-\gamma} \int_0^\tau b(s) ds \stackrel{(96)}{\lesssim} (t'+1)^{-\gamma} \tau^{1-\frac{1}{p}} a(0)$$

and (100) follows by (99) and (77a) for  $t=0$  in form of  $a(0) \lesssim 1$ .

Argument for (101): As above we have

$$\int_\tau^{\frac{t'}{2}} (t'-s+1)^{-\gamma} b(s) ds \lesssim (t'+1)^{-\gamma} \int_\tau^{\frac{t'}{2}} b(s) ds \stackrel{(96)}{\lesssim} (t'+1)^{-\gamma} \tau^{1-\gamma-\frac{1}{p}} \Lambda\left(\frac{t'}{2}\right)$$



and (101) follows by (99) and the monotonicity of  $\Lambda$ .

Argument for (102): Since

$$\begin{aligned} \frac{t}{2} \leq t' \leq t \quad \text{and} \quad \frac{t'}{2} \leq s \leq t' \\ \Leftrightarrow \\ \frac{t}{4} \leq s \leq t \quad \text{and} \quad \max\{s, \frac{t}{2}\} \leq t' \leq \min\{2s, t\}, \end{aligned}$$

we obtain by switching the order of the integrals:

$$\begin{aligned} [\text{L. H. S. of (102)}] &\leq \frac{1}{t} \int_{\frac{t}{4}}^t \int_s^t (t' - s + 1)^{-\gamma} dt' b(s) ds = \frac{1}{t} \int_{\frac{t}{4}}^t \int_0^{t-s} (t'' + 1)^{-\gamma} dt'' b(s) ds \\ &\stackrel{\gamma \geq 1}{\leq} \int_{\frac{t}{4}}^t b(s) ds \times \frac{1}{t} \int_0^t (t'' + 1)^{-1} dt'' \stackrel{(96)}{\lesssim} (t+1)^{1-\gamma-\frac{1}{p}} \Lambda(t) \frac{\ln(t+1)}{t} \\ &\stackrel{t \geq 1}{\lesssim} (t+1)^{-\gamma} \Lambda(t) \frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}}. \end{aligned}$$

The claim of Step 2 follows from the combination of (98)—(102).

**Step 3.** Proof of (95).

For  $\tau \gg 1$ ,  $\frac{\ln(t+1)}{(t+1)^{\frac{1}{p}}}$  is monotone decreasing for  $t \geq \tau$ . Hence for  $\tau \gg 1$ , Step 2 can be upgraded to

$$(t+1)^\gamma a(t) \lesssim 1 + \tau^{1-\frac{1}{p}} + \left( \tau^{1-\gamma-\frac{1}{p}} + \frac{\ln(\tau+1)}{(\tau+1)^{\frac{1}{p}}} \right) \Lambda(t).$$

Because of  $\gamma \geq 1$  and  $p < \infty$ , the expressions  $\tau^{1-\gamma-\frac{1}{p}}$  and  $\frac{\ln(\tau+1)}{(\tau+1)^{\frac{1}{p}}}$  tend to zero as  $\tau \rightarrow \infty$  and  $t \rightarrow \infty$ . Hence, by Step 2 we can find a threshold  $\tau_0 > 0$  only depending on  $p, \gamma$  and  $C_1$  such that for all  $t \geq 4\tau_0$

$$(t+1)^\gamma a(t) \leq C_2 \left( 1 + \tau_0^{1-\frac{1}{p}} \right) + \frac{1}{2} \Lambda(t),$$

where  $C_2$  is a constant only depending on  $p, \gamma$  and  $C_1$ . For all  $t > 0$  we then have

$$\begin{aligned} \Lambda(t) &\leq \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + \sup_{4\tau_0 \leq s \leq t} (1+s)^\gamma a(s) \\ &\leq \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + C_2 \left( 1 + \tau_0^{1-\frac{1}{p}} \right) + \frac{1}{2} \Lambda(t), \end{aligned}$$

that is

$$\Lambda(t) \leq 2 \sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) + 2C_2(1 + \tau_0^{1-\frac{1}{p}}).$$

Since

$$\sup_{0 \leq s \leq 4\tau_0} (1+s)^\gamma a(s) \stackrel{(77b)}{\lesssim} (1+4\tau_0)^{\gamma+1-\frac{1}{p}} a(0),$$

and because  $a(0) \lesssim 1$  by (77a) applied with  $t = 0$ , we deduce that  $\Lambda(t)$  is bounded for all  $t > 0$  by a constant that only depends on  $p, \gamma, C_1$  (and the choice of  $\tau_0$ ).

□

## 6 Application to stochastic homogenization

In this section we apply our quantitative methods to stochastic homogenization. In Section 6.1 we recall some fundamentals of the qualitative stochastic homogenization theory. In Section 6.2 we study the corrector problem of stochastic homogenization and establish several new estimates. Finally, in Section 6.3 we address the problem of approximating the homogenized coefficients by the periodization method, and establish new and optimal error estimates.

### 6.1 Qualitative theory of stochastic homogenization

Let  $B \subset \mathbb{R}^d$  be a bounded domain and  $f$  a bounded, continuous function on  $B$ . For  $\mathbf{a} \in \Omega$  let  $u_\varepsilon(\mathbf{a}, \cdot) : \varepsilon\mathbb{Z}^d \rightarrow \mathbb{R}$  denote the solution to the discrete Dirichlet problem

$$\begin{aligned} \nabla_\varepsilon^* \mathbf{a}\left(\frac{x}{\varepsilon}\right) \nabla_\varepsilon u_\varepsilon(\mathbf{a}, x) &= f(x) & \text{for } x \in \varepsilon\mathbb{Z}^d \cap B, \\ u_\varepsilon(\mathbf{a}, x) &= 0 & \text{for } x \in \varepsilon\mathbb{Z}^d \setminus B. \end{aligned}$$

Above,  $\nabla_\varepsilon$  and  $-\nabla_\varepsilon^*$  are finite difference approximations of the usual gradient and divergence operator, i. e. they are just rescaled versions of the discrete derivatives introduced in Section 2 and defined as follows: For  $g : \varepsilon\mathbb{Z}^d \rightarrow \mathbb{R}$  and  $G = (G_1, \dots, G_d) : \varepsilon\mathbb{Z}^d \rightarrow \mathbb{R}^d$  we set

$$\begin{aligned} \nabla_{\varepsilon,i} g(x) &:= \frac{1}{\varepsilon}(g(x + \varepsilon e_i) - g(x)), & \nabla_{\varepsilon,i}^* g(x) &:= \frac{1}{\varepsilon}(g(x - \varepsilon e_i) - g(x)), \\ \nabla_\varepsilon g(x) &:= (\nabla_{\varepsilon,1} g, \dots, \nabla_{\varepsilon,d} g), & \nabla_\varepsilon^* G(x) &:= \sum_{i=1}^d \nabla_{\varepsilon,i}^* G_i(x). \end{aligned}$$

Classical results show that for appropriate classes of coefficient fields  $\mathbf{a} \in \Omega$  the following homogenization property (H) holds: There exists a homogenized matrix  $\mathbf{a}_{\text{hom}} \in \mathbb{R}_{\text{sym}}^{d \times d}$  such that for all  $f$  as described above, the solutions  $u_\varepsilon(\mathbf{a}, \cdot)$  converge to some  $u_0$  in the sense that

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^d \sum_{x \in \varepsilon\mathbb{Z}^d \cap B} (u_\varepsilon(\mathbf{a}, x) - u_0(x))^2 = 0,$$

with  $u_0$  defined as the unique weak solution to

$$(103) \quad \begin{aligned} -\nabla \cdot \mathbf{a}_{\text{hom}} \nabla u_0 &= f & \text{in } B, \\ u_0 &= 0 & \text{on } \partial B. \end{aligned}$$

Trivially, (H) holds for coefficient fields that are constant in space. If  $\mathbf{a} \in \Omega$  is variable in space,  $\nabla_\varepsilon u_\varepsilon$  rapidly oscillates and the passage to the macroscopic limit becomes non-trivial. By  $G$ -compactness of general linear, symmetric elliptic operators (see [34, Proposition 3], property (H) holds in general only along subsequences, and different limits might emerge. For periodic coefficients this is not the case. Indeed, for  $\mathbf{a} \in \Omega_L$  property (H) holds with  $\mathbf{a}_{\text{hom}}$  given by the periodic homogenization formula

$$\forall e \in \mathbb{R}^d : \quad e \cdot \mathbf{a}_{\text{hom}} e = \min_{\varphi(x) \text{ } L\text{-periodic}} \frac{1}{L^d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} (e + \nabla \varphi(x)) \cdot \mathbf{a}(x) (e + \nabla \varphi(x)).$$

The first results for random coefficients have independently been obtained by Kozlov [21] and Papanicolaou & Varadhan [32]. For the rest of this section we assume that  $\langle \cdot \rangle$  is stationary and ergodic. Then  $\langle \cdot \rangle$ -almost surely the random field  $u_\varepsilon : \Omega \times \varepsilon\mathbb{Z}^d \rightarrow \mathbb{R}$  converges to the deterministic limit  $u_0$  determined by (103), where the homogenized coefficients are defined as follows.

**Definition 11.** For  $\langle \cdot \rangle$  stationary we define  $\mathbf{a}_{\text{hom}} \in \mathbb{R}_{\text{sym}}^{d \times d}$  by the formula

$$(104) \quad \forall e \in \mathbb{R}^d : \quad e \cdot \mathbf{a}_{\text{hom}} e := \inf_{\phi \in L^2(\Omega)} \langle (e + D\phi) \cdot \mathbf{a}(0) (e + D\phi) \rangle.$$

Fix some arbitrary direction  $e \in \mathbb{R}^d$  with  $|e| = 1$ . A solution  $\phi \in L^2(\Omega)$  to the minimization problem (104) is called a *stationary corrector* associated with  $\langle \cdot \rangle$  and the direction  $e$ . If it exists, then it is unique up to the addition of a shift-invariant function (i. e. a constant in the ergodic case) and is characterized by the *corrector problem*

$$(105) \quad D^* \mathbf{a}(0) D \phi = \mathfrak{d} \quad \text{in } L^2(\Omega),$$

with right-hand side

$$(106) \quad \mathfrak{d} := -D^* \mathbf{a}(0) e.$$

Note that (105) is just the Euler-Lagrange equation associated with the minimization problem defining  $\mathbf{a}_{\text{hom}}$ , see (104). Sometimes it is more convenient to consider the *corrector problem in physical space*, which characterizes  $\phi$  on the level of its stationary extension: If  $\phi$  is a solution to (105), then its stationary extension  $\bar{\phi}$  is the unique stationary random field with finite second moment (and mean given by  $\langle \phi \rangle$ ) solving

$$(107) \quad \nabla^* \mathbf{a}(x) \nabla \bar{\phi}(\mathbf{a}, x) = -\nabla^* \mathbf{a}(x) e \quad \text{for all } x \in \mathbb{Z}^d \text{ and } \langle \cdot \rangle\text{-almost every } \mathbf{a} \in \Omega.$$

Vice versa, any stationary solution to (107) with finite second moment yields a stationary corrector.

However, due to the lack of a Poincaré inequality in  $L^2(\Omega)$  for the horizontal derivative  $D$ , stationary correctors may not exist for general ergodic ensembles. Yet, we may still look for the gradient of a corrector in the form of a “curl free” random vector field. To make this precise we introduce the Hilbert space

$$V(\Omega) := \overline{\{ D \phi : \phi \in L^2(\Omega) \}}^{L^2(\Omega)^d}.$$

**Lemma 16.** *Let  $\langle \cdot \rangle$  be stationary. For all  $e \in \mathbb{R}^d$  the functional*

$$V(\Omega) \ni \chi \mapsto \langle (e + \chi) \cdot \mathbf{a}(0) (e + \chi) \rangle$$

*admits a unique minimizer. The minimizer is characterized by the Euler-Lagrange equation*

$$(108) \quad D^* \mathbf{a}(0) (e + \chi) = 0 \quad \text{for almost every } \mathbf{a} \in \Omega.$$

*It satisfies the a priori estimate  $\lambda \langle |\chi|^2 \rangle^{\frac{1}{2}} \leq |e|$ , and one has*

$$(109) \quad e \cdot \mathbf{a}_{\text{hom}} e = \langle (e + \chi) \cdot \mathbf{a}(0) (e + \chi) \rangle$$

*and*

$$(110) \quad \lambda |e|^2 \leq e \cdot \mathbf{a}_{\text{hom}} e \leq |e|^2.$$

*Proof.* The existence and uniqueness of a solution to (108) follows from Riesz’ representation theorem based on the comparability of the standard scalar product on  $L^2(\Omega)^d$  with  $\langle \chi \cdot \mathbf{a}(0) \chi \rangle$ :

$$\lambda \langle |\chi|^2 \rangle \leq \langle \chi \cdot \mathbf{a}(0) \chi \rangle \leq \langle |\chi|^2 \rangle.$$

This also yields the a priori estimate. The characterizing property of the Euler-Lagrange equation and the formula for  $\mathbf{a}_{\text{hom}}$  follow by density of  $\{ D \phi : \phi \in L^2(\Omega) \}$  in  $V(\Omega)$ . Finally, the bounds in (110) follow on the one hand from  $e \cdot \mathbf{a}_{\text{hom}} e \leq \langle e \cdot \mathbf{a}(0) e \rangle \leq |e|^2$  and, on the other hand from

$$\forall \chi \in V(\Omega) : \quad \langle (e + \chi) \cdot \mathbf{a}(0) (e + \chi) \rangle \geq \lambda \langle |e + \chi|^2 \rangle \geq \lambda |e|^2.$$

The latter holds since  $\langle e \cdot D \phi \rangle = 0$  for all  $\phi \in L^2(\Omega)$  and thus  $\langle e \cdot \chi \rangle = 0$  for all  $\chi \in V(\Omega)$  by density.  $\square$

Note that  $V(\Omega) \subset \{\chi \in L^2(\Omega)^d : D_i \chi_j = D_j \chi_i \text{ and } \langle \chi_i \rangle = 0 \text{ for all } i, j = 1, \dots, d\}$ . Hence the stationary extension  $\bar{\chi}$  is a curl-free vector field on  $\mathbb{Z}^d$ . By the Poincaré lemma in  $\mathbb{Z}^d$  we can represent  $\bar{\chi}$  by the gradient of a *non-stationary* random field: There exists a function  $\varphi_0 : \Omega \times \mathbb{Z}^d \rightarrow \mathbb{R}$  such that  $\langle \cdot \rangle$ -almost surely  $\bar{\chi} = \nabla \varphi_0$  and  $\varphi_0(x=0) = 0$ . Hence, (109) can be written in the form of

$$(111) \quad e \cdot \mathbf{a}_{\text{hom}} e = \langle (e + \nabla \varphi_0) \cdot \mathbf{a}(e + \nabla \varphi_0) \rangle.$$

By construction,  $\varphi_0$  satisfies a corrector problem in physical space: for  $\langle \cdot \rangle$ -almost every  $\mathbf{a} \in \Omega$  we have

$$(112a) \quad \nabla^* \mathbf{a}(x) \nabla \varphi_0(\mathbf{a}, x) = -\nabla^* \mathbf{a}(x) e \quad \text{in } \mathbb{Z}^d,$$

and

$$(112b) \quad \varphi_0(\mathbf{a}, x=0) = 0,$$

$$(112c) \quad \nabla \varphi_0 \text{ is stationary,} \quad \langle \nabla \varphi_0 \rangle = 0.$$

In fact, for stationary and ergodic ensembles, the random field  $\varphi_0$  (and thus  $\chi$  from Lemma 16) is characterized by (112a) – (112c): As shown by [21, 32] in the continuum case and [24, 23] in the discrete case, the corrector problem (112a) – (112c) admits a unique solution, which we call the (*non-stationary*) *corrector anchored at  $x=0$* .

With  $\varphi_0$  at hand, the proof of (H) can be deduced by Tartar’s method of oscillating test functions. In this respect it is a crucial observation that  $\varphi_0$  has almost surely sublinear growth on spatial average, i. e.

$$(113) \quad \limsup_{R \rightarrow \infty} \frac{1}{R^d} \sum_{x \in ([0, R] \cap \mathbb{Z})^d} \left| \frac{\varphi_0(\mathbf{a}, x)}{R} \right|^2 = 0 \quad \text{for } \langle \cdot \rangle\text{-almost every } \mathbf{a} \in \Omega.$$

Indeed, this can be deduced from (112a) – (112c) with the help of a maximal function estimate from ergodic theory, see [23, Theorem 10] for a proof in the continuum case, [33, p.226 ff] for the discrete case, see also [32] and [24] for a proof of sublinearity in the form of  $\langle \left( \frac{\varphi_0(x)}{|x|} \right)^2 \rangle \rightarrow 0$  as  $x \rightarrow \infty$ . Let us remark that the maximal function estimate does not yield any rate for the convergence in (113).

## 6.2 Quantitative estimates for the corrector

In this section we apply our quantitative methods to the *corrector equation* of stochastic homogenization (105). As explained in the previous section, solutions to (105) may not exist for general stationary and ergodic ensembles. However, as a consequence of Theorem 2, we find that in dimensions  $d > 2$  and for ensembles  $\langle \cdot \rangle$  satisfying (69a) or (69b), stationary correctors do exist:

**Proposition 4.** *Let  $d > 2$ , and  $\langle \cdot \rangle$  either satisfy (69a) or (69b). Consider*

$$\phi = \int_0^\infty \exp(-tD^* \mathbf{a}(0)D) \mathfrak{d} dt,$$

where  $\mathfrak{d}$  is given by (106). Then the integral converges absolutely in  $L^2(\Omega)$  and  $\phi$  is the unique solution with zero mean to (105). Moreover, we have

$$\langle |\phi|^p \rangle^{\frac{1}{p}} \lesssim 1 \quad \text{for all } 1 \leq p < \infty.$$

Here,  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p, \rho, \lambda$  and  $d$ .

*Proof.* To shorten the notation we set

$$u(t) = \exp(-tD^* \mathbf{a}(0)D) \mathfrak{d}.$$

Since for any  $1 \leq p < \infty$  we have

$$\sum_{y \in \mathbb{Z}^d} \left\langle \left( \frac{\partial \mathfrak{d}}{\partial y} \right)^{2p} \right\rangle^{\frac{1}{2p}} \leq 2,$$

Theorem 2 combined with Jensen's inequality in probability yields

$$\langle |u(t)|^p \rangle^{\frac{1}{p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})} \quad \text{for all } 1 \leq p < \infty.$$

Since  $\frac{d}{4} + \frac{1}{2} > 1$  for  $d > 2$ , the right-hand side is integrable and thus  $\phi$  is a well-defined function in  $L^p(\Omega)$  with  $\langle |\phi|^p \rangle^{\frac{1}{p}} \lesssim 1$ . By construction we have

$$\begin{aligned} \mathfrak{d} &= u(0) = - \int_0^\infty \frac{d}{dt} u(t) dt = \int_0^\infty D^* \mathbf{a}(0) D u(t) dt \\ &= D^* \mathbf{a}(0) D \int_0^\infty u(t) dt = D^* \mathbf{a}(0) D \phi, \end{aligned}$$

which is (105). Since  $\langle \mathfrak{d} \rangle = 0$ , we conclude that  $\langle \phi \rangle = 0$ .  $\square$

**Remark 13.** *With the stationary corrector  $\phi$  at hand, we can easily recover the non-stationary corrector  $\varphi_0$  (defined via (112a) – (112c)), by setting  $\varphi_0(\mathbf{a}, x) := \bar{\phi}(\mathbf{a}, x) - \bar{\phi}(\mathbf{a}, 0)$ . Moreover, with help of  $\phi$  we deduce that  $\varphi_0$  has not only sublinear growth on average, but is bounded on average: In fact, we have*

$$\limsup_{R \rightarrow \infty} \frac{1}{R^d} \sum_{x \in ([0, R] \cap \mathbb{Z})^d} |\bar{\phi}(\mathbf{a}, x)|^2 = \langle \bar{\phi}^2 \rangle \quad \text{for } \langle \cdot \rangle\text{-almost every } \mathbf{a} \in \Omega$$

by ergodicity.

In dimension  $d = 2$  we do not expect (105) to have a solution. Therefore, we consider for all  $\mu > 0$  the modified corrector equation, as standard in stochastic homogenization theory (e. g. see [32]):

$$(114) \quad \mu \phi_\mu + D^* \mathbf{a}(0) D \phi_\mu = \mathfrak{d},$$

where  $\mathfrak{d}$  is given by (106). The modified corrector  $\phi_\mu$  will also play a crucial role in the discussion of an approximation scheme for  $\mathbf{a}_{\text{hom}}$ , cf. Section 6.3. Equation (114) has a unique solution in  $L^2(\Omega)$  as can be deduced from the Riesz representation theorem. In fact, since  $\mathfrak{d} \in C^0(\Omega)$ , (114) even makes sense in  $C^0(\Omega)$  — thus from a deterministic point of view:

**Lemma 17.** *For every  $\mu > 0$  the modified corrector problem (114) admits a unique solution in  $C^0(\Omega)$ . It is given by the formula*

$$(115) \quad \phi_\mu = \int_0^\infty \exp(-\mu t) \exp(-t D^* \mathbf{a}(0) D) \mathfrak{d} dt.$$

*Proof.* The modified corrector problem (114) admits a unique solution in  $L^2(\Omega)$  as can be seen by appealing to Riesz' representation theorem. Since  $C^0(\Omega) \subset L^2(\Omega)$ , it suffices to argue that  $\phi_\mu$  defined via (115) is a solution to (114) in  $C^0(\Omega)$ . To that end set  $u(t) := \exp(-t D^* \mathbf{a}(0) D) \mathfrak{d}$ . Since  $\mathfrak{d} \in C^0(\Omega)$  and  $\exp(-t D^* \mathbf{a}(0) D)$  is a uniformly continuous group of bounded operators on  $C^0(\Omega)$ , we deduce that  $u \in C^\infty(\mathbb{R}, C^0(\Omega))$  satisfies (13) with initial value  $\zeta = \mathfrak{d}$ . Since, as we will see below,

$$(116) \quad |u(t, \mathbf{a})| \leq 2 \quad \text{for all } t > 0 \text{ and } \mathbf{a} \in \Omega,$$

$\phi_\mu$  is well-defined and belongs to  $C^0(\Omega)$ . Furthermore, we have

$$\begin{aligned} D^* \mathbf{a}(0) D \phi_\mu(\mathbf{a}) &= \int_0^\infty \exp(-\mu t) D^* \mathbf{a}(0) D u(t, \mathbf{a}) dt = - \int_0^\infty \exp(-\mu t) \frac{d}{dt} u(t, \mathbf{a}) dt \\ &= -\mu \int_0^\infty \exp(-\mu t) u(t, \mathbf{a}) dt + u(0, \mathbf{a}) = -\mu \phi_\mu(\mathbf{a}) + \mathfrak{d}(\mathbf{a}), \end{aligned}$$

which is the modified corrector problem (114). Hence, it remains to prove (116). By Lemma 1, the stationary extension  $\bar{u}$  satisfies the parabolic equation (13), and thus admits by Duhamel's formula, cf. Lemma 2, the representation  $\bar{u}(t, \mathbf{a}, x) = -\sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z) \nabla^* \mathbf{a}(z) e$ . Since  $u(t, \mathbf{a}) = \bar{u}(t, \mathbf{a}, 0)$ , an integration by parts yields  $u(t, \mathbf{a}) = -\sum_{z \in \mathbb{Z}^d} \nabla_z G(t, \mathbf{a}, 0, z) \cdot \mathbf{a}(z) e$ . From (26a) and (26c) we learn that

$$|u(t, \mathbf{a})| \leq \sum_{z \in \mathbb{Z}^d} |\nabla_z G(t, \mathbf{a}, 0, z)| \leq 2 \sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, 0, z) = 2.$$

□

If the ensemble satisfies  $\text{SG}_\infty(\rho)$ , we obtain the following quantitative estimate, which is uniform in  $\mu$  for  $d > 2$ :

**Proposition 5.** *Let  $d \geq 2$  and  $\mu > 0$ . Assume that  $\langle \cdot \rangle$  either satisfies (69a) or (69b). Let  $\phi_\mu$  be defined via (115). Then for all  $1 \leq p < \infty$  we have*

$$\langle |\phi_\mu|^p \rangle^{\frac{1}{p}} \lesssim \begin{cases} \ln^{\frac{1}{2}}(\frac{1}{\mu} + 1) & \text{for } d = 2 \text{ and } 1 \leq p \leq 2, \\ \ln(\frac{1}{\mu} + 1) & \text{for } d = 2 \text{ and } p > 2, \\ 1 & \text{for } d > 2, \end{cases}$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p, \rho, \lambda$  and  $d$ .

**Remark 14.** *The estimate in the case  $d = 2$  and  $p \leq 2$  is optimal (in terms of the exponent  $\frac{1}{2}$  of the logarithmic divergence), while it is not for  $d = 2$  and  $p > 2$ .*

*Proof of Proposition 5.* To shorten the notation we set

$$u_\mu(t) := \exp(-\mu t) \exp(-t D^* \mathbf{a}(0) D) \mathfrak{d}.$$

We only consider the difficult case of  $d = 2$ . The estimate for  $p \in [1, 2)$  follows by Cauchy-Schwarz from the estimate for  $p = 2$ . The estimate for  $p = 2$  can be seen as follows: Since  $D^* \mathbf{a}(0) D$  is symmetric, we have for all  $t, s \geq 0$ ,

$$\begin{aligned} & \langle u_\mu(s) u_\mu(t) \rangle \\ &= \left\langle \left( \exp(-s\mu) \exp(-s D^* \mathbf{a}(0) D) \mathfrak{d} \right) \left( \exp(-\mu t) \exp(-t D^* \mathbf{a}(0) D) \mathfrak{d} \right) \right\rangle \\ &= \left\langle \left( \exp(-\frac{s+t}{2}\mu) \exp(-\frac{s+t}{2} D^* \mathbf{a}(0) D) \mathfrak{d} \right)^2 \right\rangle \\ &= \langle u_\mu^2(\frac{s+t}{2}) \rangle. \end{aligned}$$

Hence,

$$\begin{aligned} (117) \quad \left\langle \left( \int_0^\infty u_\mu(t) dt \right)^2 \right\rangle &= \int_0^\infty \int_0^\infty \langle u_\mu(s) u_\mu(t) \rangle dt ds \\ &= \int_0^\infty \int_0^\infty \langle u_\mu^2(\frac{s+t}{2}) \rangle dt ds \\ &= 2 \int_0^\infty \int_{\frac{s}{2}}^\infty \langle u_\mu^2(\tau) \rangle d\tau ds \\ &\stackrel{\text{Theorem 2}}{\lesssim} \int_0^\infty \int_{\frac{s}{2}}^\infty \exp(-2\mu\tau) (\tau + 1)^{-2} d\tau ds \\ &\leq \int_0^\infty \exp(-\mu s) \int_{\frac{s}{2}}^\infty (\tau + 1)^{-2} d\tau ds \\ &\lesssim \int_0^\infty \exp(-\mu s) (s + 1)^{-1} ds \\ (118) \quad &\lesssim \ln(\frac{1}{\mu} + 1). \end{aligned}$$

The last inequality follows from the two elementary estimates

$$\int_0^{\frac{1}{\mu}} \exp(-\mu t)(t+1)^{-1} dt \leq \int_0^{\frac{1}{\mu}} (t+1)^{-1} dt \leq \ln\left(\frac{1}{\mu} + 1\right),$$

and

$$\int_{\frac{1}{\mu}}^{\infty} \exp(-\mu t)(t+1)^{-1} dt \leq \mu \int_{\frac{1}{\mu}}^{\infty} \exp(-\mu t) dt \leq 1.$$

This completes the argument for  $p = 2$ . Next, we consider the case  $p > 2$  which we obtain by a simpler argument that relies on the triangle inequality in  $L^2(\Omega)$  and does not appeal to the semigroup structure:

$$\begin{aligned} \langle \phi_\mu^2 \rangle^{\frac{1}{p}} &\leq \int_0^\infty \exp(-\mu t) \langle |\exp(-tD^* \mathbf{a}(0)D) \mathfrak{d}|^p \rangle^{\frac{1}{p}} dt \stackrel{\text{Theorem 2}}{\lesssim} \int_0^\infty \exp(-\mu t)(t+1)^{-1} dt \\ &\stackrel{(118)}{\lesssim} \ln\left(\frac{1}{\mu} + 1\right). \end{aligned}$$

□

Our quantitative methods also yield bounds for higher moments of the gradient field  $\chi$  introduced in Lemma 16:

**Proposition 6.** *Let  $d \geq 2$  and let  $\langle \cdot \rangle$  either satisfy (69a) or (69b). We set*

$$\chi = \int_0^\infty D \exp(-tD^* \mathbf{a}(0)D) \mathfrak{d} dt,$$

where  $\mathfrak{d}$  is given by (106). Then  $\chi \in V(\Omega)$  satisfies (108). Furthermore, we have

$$\langle |\chi|^p \rangle^{\frac{1}{p}} \lesssim 1 \quad \text{for all } 1 \leq p < \infty.$$

Here,  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p, \rho, \lambda$  and  $d$ .

*Proof.* For dimension  $d > 2$  the proposition directly follows from Proposition 4, and the fact that  $D$  is a bounded operator on  $L^p(\Omega)$ . We present the argument for dimension  $d = 2$ . For convenience we set

$$u(t) = \exp(-tD^* \mathbf{a}(0)D) \mathfrak{d}.$$

By Theorem 2 we have

$$(119) \quad \langle |u(t)|^p \rangle^{\frac{1}{p}} \lesssim (t+1)^{-(\frac{d}{4} + \frac{1}{2})} = (t+1)^{-1}$$

for all  $1 \leq p < \infty$ . We fix  $\gamma \in (p-1, p)$  to the effect of

$$(120) \quad \int_0^\infty (t+1)^{-\frac{\gamma}{p-1}} < \infty \quad \text{and} \quad \int_0^\infty (t+1)^{\gamma-1-p} < \infty.$$

Then

$$\begin{aligned} \int_0^\infty \langle |Du(t)|^p \rangle^{\frac{1}{p}} dt &\stackrel{\Delta\text{-inequality}}{\leq} \int_0^\infty \langle |Du(t)|^p \rangle^{\frac{1}{p}} (t+1)^{\frac{\gamma}{p}} (t+1)^{-\frac{\gamma}{p}} dt \\ &\stackrel{\text{Hölder}}{\leq} \left( \int_0^\infty \langle |Du(t)|^p \rangle (t+1)^\gamma dt \right)^{\frac{1}{p}} \left( \int_0^\infty (t+1)^{-\frac{\gamma}{p-1}} dt \right)^{\frac{p-1}{p}} \\ &\stackrel{(120)}{\lesssim} \int_0^\infty -\frac{d}{dt} \langle |u(t)|^p \rangle (t+1)^\gamma dt \\ &\stackrel{\langle |Du(0)|^p \rangle \lesssim 1}{\lesssim} \left( \int_0^\infty \langle |u(t)|^p \rangle (t+1)^{\gamma-1} dt \right)^{\frac{1}{p}} + 1 \\ &\stackrel{(119)}{\lesssim} \left( \int_0^\infty (t+1)^{\gamma-1-p} dt \right)^{\frac{1}{p}} + 1 \\ &\stackrel{(120)}{\lesssim} 1. \end{aligned}$$

Hence, the integral  $\chi = \int_0^\infty Du(t) dt$  converges in  $L^p(\Omega)$  and  $\langle |\chi|^p \rangle^{\frac{1}{p}} \lesssim 1$ . In particular, it converges in  $L^2(\Omega)$  and thus

$$\chi = L^2(\Omega)\text{-}\lim_{T \uparrow \infty} \int_0^T Du(t) dt = L^2(\Omega)\text{-}\lim_{T \uparrow \infty} D \int_0^T u(t) dt \in V(\Omega).$$

Since

$$D^* \mathbf{a}(0) \chi = \int_0^\infty D^* \mathbf{a}(0) Du(t) dt = - \int_0^\infty \frac{d}{dt} u(t) dt \stackrel{(119)}{=} u(0) = -D^* \mathbf{a}(0) e,$$

$\chi$  indeed solves (108). □

### 6.3 Approximation of the homogenized coefficients by periodization

For a discussion of the literature on error estimates, in particular the pertinent work by Yurinskii [37] and Naddaf & Spencer [29], we refer to [17, Section 1.2].

Let  $\langle \cdot \rangle$  be an ergodic ensemble, and fix an arbitrary direction  $e \in \mathbb{R}^d$ . In terms of numerical approximations, the representations of  $\mathbf{a}_{\text{hom}}$  given by (109) and (111) are of no immediate practical use: In order to obtain  $\chi$  in the form of  $\bar{\chi} = \nabla \varphi_0$  (where  $\varphi_0$  denotes the corrector anchored at  $x = 0$ ), the corrector problem (112a) – (112c) has to be solved

- for *every realization* of the coefficients  $\mathbf{a} \in \Omega$  and
- in the *whole space*  $\mathbb{Z}^d$ .

In order to treat the infinite dimensionality of  $\Omega$ , it is natural to appeal to ergodicity (in the sense that spatial averages are equal to ensemble averages, see Theorem 1). This suggests to replace (109) by

$$(121) \quad \frac{1}{L^d} \sum_{x \in \mathbb{Z}^d} \bar{E}(x) \eta\left(\frac{x}{L}\right), \quad \text{where } \bar{E}(x) = \bar{E}(\mathbf{a}, x) := (e + \nabla \varphi_0(\mathbf{a}, x)) \cdot \mathbf{a}(x) (e + \nabla \varphi_0(\mathbf{a}, x)),$$

and  $\eta$  is a suitable averaging function, that is,  $\eta : \mathbb{R}^d \rightarrow [0, 1]$  is compactly supported and  $\int_{\mathbb{R}^d} \eta dx = 1$ . Indeed, since  $\bar{E}$  is a stationary random field (due to (112c)) and  $\langle \cdot \rangle$  is ergodic, Birkhoff's ergodic theorem implies that the right-hand side in (121) converges to  $e \cdot \mathbf{a}_{\text{hom}} e$  for  $\langle \cdot \rangle$ -almost every  $\mathbf{a} \in \Omega$ . On top of that, one expects  $\bar{E}(\mathbf{a}, x)$  to display a decay (w. r. t.  $x$ ) of correlations over large distances, so that (121) seems to yield a good approximation for  $L \gg 1$ .

Hence, for  $e \cdot \mathbf{a}_{\text{hom}} e$  we only need to know the gradient of the corrector, i. e.  $\bar{\chi}(\mathbf{a}, \cdot)$  resp.  $\nabla \varphi_0(\mathbf{a}, \cdot)$ , for a single realization of the coefficients. However, for that we still have to solve (112a) in the whole space  $\mathbb{Z}^d$  – at least for a single realization of the coefficients. (In fact, in order to enforce (112c) it might still be necessary to consider the whole ensemble  $\langle \cdot \rangle$ ). In order to overcome these difficulties, we approximate the solution to the corrector problem by the solution to an equation posed on a bounded domain, completed by suitable boundary conditions. There are several choices possible (homogeneous Dirichlet or Neumann boundary conditions for instance). In this contribution we consider periodic boundary conditions which nicely interplay with stationarity, and naturally emerge in the following “mathematical version” of the *periodic representative volume element* scheme: For some large  $L$  we “approximate” the stationary ergodic ensemble  $\langle \cdot \rangle$  by a stationary  $L$ -periodic ensemble  $\langle \cdot \rangle_L$  that is characterized by the same (or at least similar) specifications (think of  $\langle \cdot \rangle_L \approx \langle \cdot \rangle_{\{\mathbf{a}(x)\}_{x \in ([0, L) \cap \mathbb{Z}]^d}}$ ); see also Remark 15 below), and we replace (121) by

**Definition 12.** *Let  $\langle \cdot \rangle$  be stationary and  $L$ -periodic for some  $L \in \mathbb{N}$ . We define the random matrix  $\mathbf{a}_{\text{av}, L} : \Omega_L \rightarrow \mathbb{R}_{\text{sym}}^{d \times d}$  via*

$$(122) \quad e \cdot \mathbf{a}_{\text{av}, L} e = \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z}]^d} \overline{(e + \chi) \cdot \mathbf{a}(0) (e + \chi)}(x)$$

where  $e \in \mathbb{R}^d$ , and where  $\chi$  is associated with  $\langle \cdot \rangle$  and  $e$  through (108).



The average (122) can also be viewed as the right-hand side of (121) with the averaging function  $\eta$  being the characteristic function of the unit cube  $[0, 1)^d$ . In contrast to that, for non-periodic ensembles better approximation properties appear to require a smooth averaging functions, as discussed in [17, 18].

For  $L$ -periodic ensembles, a solution to the corrector problem (105) can unambiguously be obtained by appealing to the corresponding spatial,  $L$ -periodic corrector problem. In that way we also obtain estimates for the second moment of  $\phi$  directly, without assuming  $\text{SG}_L(\rho)$ . However, that approach does not yield estimates on the corrector independent of the period  $L$ .

**Lemma 18.** *Let  $d \geq 2$ ,  $L \in \mathbb{N}$  and let  $\langle \cdot \rangle$  denote a stationary,  $L$ -periodic ensemble. For all  $e \in \mathbb{R}^d$  with  $|e| = 1$ , and  $\mathbf{a} \in \Omega_L$  let  $\bar{\phi}(\mathbf{a}, \cdot)$  denote the unique solution to*

$$(123) \quad \begin{cases} \nabla^* \mathbf{a}(x)(\nabla \bar{\phi}(\mathbf{a}, x) + e) = 0 & \text{in } \mathbb{Z}^d \\ \text{subject to } \bar{\phi}(\mathbf{a}, \cdot) \text{ } L\text{-periodic and } \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \bar{\phi}(\mathbf{a}, x) = 0. \end{cases}$$

Then  $\bar{\phi}$  is  $\langle \cdot \rangle$ -stationary, and  $\phi(\mathbf{a}) := \bar{\phi}(\mathbf{a}, 0)$  satisfies the corrector problem (105). In addition we have  $\langle \phi^2 \rangle \lesssim L^2$ , where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\lambda$  and  $d$ .

*Proof.* Since there is a spatial Poincaré inequality for periodic functions of  $([0, L) \cap \mathbb{Z})^d$  with zero mean, (123) admits a unique solution by the Riesz representation theorem. The energy estimate, uniform ellipticity, and the Poincaré inequality yield

$$(124) \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \bar{\phi}(\mathbf{a}, x)^2 \lesssim L^2 \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\nabla \bar{\phi}(\mathbf{a}, x)|^2 \lesssim L^{d+2}.$$

From the uniqueness of the solution to (123) we deduce that  $\bar{\phi}$  is  $\langle \cdot \rangle$ -stationary and thus  $\bar{\phi}$  is identical to the stationary extension to  $\phi$ . By stationarity we have

$$\langle \phi^2 \rangle = \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \langle |\bar{\phi}(x)|^2 \rangle \stackrel{(124)}{\lesssim} L^2.$$

In particular,  $\phi \in L^2(\Omega)$ , and (123) combined with (8) yields (105).  $\square$

In contrast to  $\mathbf{a}_{\text{hom}}$  (associated with the “infinite” ensemble  $\langle \cdot \rangle$ ), the approximation  $\mathbf{a}_{\text{av}, L}$  (associated with an  $L$ -periodic ensemble  $\langle \cdot \rangle_L$ ) can be computed for each realization  $\mathbf{a}$  separately by solving a minimization problem on the discrete torus  $([0, L) \cap \mathbb{Z})^d$ .

**Lemma 19.** *Let  $\langle \cdot \rangle$  be stationary and  $L$ -periodic. Then for all  $e \in \mathbb{R}^d$  we have*

$$e \cdot \mathbf{a}_{\text{av}, L}(\mathbf{a})e = \min_{\varphi(x) \text{ } L\text{-periodic}} \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} (e + \nabla \varphi(x)) \cdot \mathbf{a}(x)(e + \nabla \varphi(x))$$

$\langle \cdot \rangle$ -almost surely.

*Proof.* From Lemma 18 we learn that  $\chi = D\phi$ . Moreover, the stationary extension  $\bar{\phi}$  satisfies (123), which is precisely the Euler-Lagrange equation associated with the energy

$$\varphi \mapsto \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} (e + \nabla \varphi(x)) \cdot \mathbf{a}(x)(e + \nabla \varphi(x))$$

defined for  $L$ -periodic functions on  $\mathbb{Z}^d$ . Hence,

$$\begin{aligned} e \cdot \mathbf{a}_{\text{av}, L}e &= \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} (e + \nabla \bar{\phi}(x)) \cdot \mathbf{a}(x)(e + \nabla \bar{\phi}(x)) \\ &= \min_{\varphi(x) \text{ } L\text{-periodic}} \frac{1}{L^d} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} (e + \nabla \varphi(x)) \cdot \mathbf{a}(x)(e + \nabla \varphi(x)). \end{aligned}$$

$\square$

Let  $\mathbf{a}_{\text{hom},L}$  denote the homogenized coefficient associated with the  $L$ -periodic ensemble  $\langle \cdot \rangle_L$  via Definition 11. As a consequence of (104), (124), and stationarity we have

$$\mathbf{a}_{\text{hom},L} = \langle \mathbf{a}_{\text{av},L} \rangle_L.$$

One might wonder whether a better approximation to  $\mathbf{a}_{\text{hom},L}$  could be obtained by taking the average in (122) over larger domains. However, since an  $L$ -periodic ensemble is typically not ergodic, this is not the case. In fact, by the Riemann-Lebesgue lemma we have for  $\langle \cdot \rangle_L$ -almost every  $\mathbf{a} \in \Omega$

$$\lim_{R \uparrow \infty} \frac{1}{R^d} \sum_{x \in ([0,R) \cap \mathbb{Z})^d} (e + \bar{\chi}(\mathbf{a}, x)) \cdot \mathbf{a}(x) (e + \bar{\chi}(\mathbf{a}, x)) = e \cdot \mathbf{a}_{\text{av},L}(\mathbf{a})e,$$

where the right-hand side still fluctuates around its average  $e \cdot \mathbf{a}_{\text{hom},L}e$ .

In the following we discuss the mean-square error introduced by the approximation based on periodization. From now on,  $\langle \cdot \rangle$  denotes a stationary “infinite” ensemble and  $\langle \cdot \rangle_L$  an associated stationary,  $L$ -periodic ensemble. We require  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$  to satisfy  $\text{SG}_\infty(\rho)$  and  $\text{SG}_L(\rho)$ , respectively. We denote by  $\mathbf{a}_{\text{hom}}$  and  $\mathbf{a}_{\text{hom},L}$  the homogenized coefficients associated with  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$ , via Definition 11, and write  $\mathbf{a}_{\text{av},L}$  for the approximation associated with the  $L$ -periodic ensemble introduced in Definition 12.

**Remark 15.** *At this stage we have not yet explained how to obtain  $\langle \cdot \rangle_L$  from  $\langle \cdot \rangle$ . In the present contribution, we address this question only in the case when  $\langle \cdot \rangle$  is the infinite i. i. d. ensemble associated with a base measure  $\beta$  (see Definition 7). In that case, the right choice for  $\langle \cdot \rangle_L$  is simply the  $L$ -periodic i. i. d. ensemble associated with the same base measure  $\beta$ , cf. Definition 10. However, most parts of the error analysis hold in general — without knowing the precise coupling between  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$  —, so that we only restrict our self to i. i. d. ensembles when it is needed. The construction of  $\langle \cdot \rangle_L$  from  $\langle \cdot \rangle$  for more general ensembles will be treated in a forthcoming work (see [16]).*

We are interested in the mean square error  $\langle |\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{av},L}|^2 \rangle_L$  which splits naturally into two parts:

$$\begin{aligned} \langle |\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{av},L}|^2 \rangle_L &= \langle |\mathbf{a}_{\text{av},L} - \langle \mathbf{a}_{\text{av},L} \rangle_L|^2 \rangle + \langle |\langle \mathbf{a}_{\text{av},L} \rangle_L - \mathbf{a}_{\text{hom}}|^2 \rangle \\ &= \langle |\mathbf{a}_{\text{av},L} - \langle \mathbf{a}_{\text{av},L} \rangle_L|^2 \rangle + |\mathbf{a}_{\text{hom},L} - \mathbf{a}_{\text{hom}}|^2 \\ (125) \qquad \qquad \qquad &= (\text{random error})^2 + (\text{systematic error})^2. \end{aligned}$$

Following the terminology of [17], we call the (square roots of the) first and second term on the right-hand side the *random error* and *systematic error*, respectively. The *random error* is due to the lack of ergodicity of the  $L$ -periodic ensemble and measures the fluctuation of  $\mathbf{a}_{\text{av},L}$  around its average. The *systematic error* is of different nature: By replacing the infinite ensemble  $\langle \cdot \rangle$  with the  $L$ -periodic ensemble  $\langle \cdot \rangle_L$ , we typically introduce artificial long-range correlations (in order to enforce the periodicity of  $\langle \cdot \rangle_L$ ), and thus induce the *systematic error* — although  $\mathbf{a}_{\text{hom}}$  and  $\mathbf{a}_{\text{hom},L}$  are obtained from  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$  by the same mechanism, namely by Definition 11. In the following we discuss both errors separately. Most proofs are postponed to the end of this section.

We start with the random error, which can be rewritten as the variance of  $\mathbf{a}_{\text{av},L}$ . Using the bound on the quartic moment of the gradient of the corrector derived in the previous section (see Proposition 6), we obtain the following optimal variance estimate:

**Proposition 7** (Optimal variance estimate). *Let  $d \geq 2$ ,  $\langle \cdot \rangle$  be stationary,  $L$ -periodic and satisfy  $\text{SG}_L(\rho)$ , and let  $\mathbf{a}_{\text{av},L}$  be as in Definition 12. Then for all  $e \in \mathbb{R}^d$ ,  $|e| = 1$ , we have*

$$\text{Var}[e \cdot \mathbf{a}_{\text{av},L}e] \lesssim L^{-d}$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\rho, \lambda$  and  $d$ .

Hence, the random error decays at the rate  $L^{-\frac{d}{2}}$  of the central limit theorem. Since the random error is due to fluctuations, its effect can be reduced by empirical averaging:

**Corollary 7.** Let  $d \geq 2$ ,  $\langle \cdot \rangle$  be stationary,  $L$ -periodic and satisfy  $SG_L(\rho)$ . For  $N \in \mathbb{N}$  consider the random matrix

$$\mathbf{a}_{\text{av},L,N} = \frac{1}{N} \sum_{i=1}^N \mathbf{a}_{\text{av},L}(\mathbf{a}^i),$$

where  $\mathbf{a}^1, \dots, \mathbf{a}^N$  denote  $N$  independent realizations of the coefficient field according to  $\langle \cdot \rangle$ , and  $\mathbf{a}_{\text{av},L}$  is as in Definition 12. Then for all  $e \in \mathbb{R}^d$ ,  $|e| = 1$ , we have

$$\text{Var}[e \cdot \mathbf{a}_{\text{av},L,N} e] \lesssim \frac{1}{N} L^{-d},$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\rho, \lambda$  and  $d$ .

Next, we discuss the systematic error  $|\mathbf{a}_{\text{hom},L} - \mathbf{a}_{\text{hom}}|$ . Let us fix some arbitrary  $e \in \mathbb{R}^d$  with  $|e| = 1$ . For clarity of the presentation, we denote by  $\chi$  the gradient of the corrector associated with  $e$  and  $\langle \cdot \rangle$ , and by  $\chi_L$  the gradient of the corrector associated with  $e$  and  $\langle \cdot \rangle_L$  through (108). We then recall that

$$\begin{aligned} e \cdot \mathbf{a}_{\text{hom}} e &= \langle (e + \chi) \cdot \mathbf{a}(0)(e + \chi) \rangle, \\ e \cdot \mathbf{a}_{\text{hom},L} e &= \langle (e + \chi_L) \cdot \mathbf{a}(0)(e + \chi_L) \rangle_L. \end{aligned}$$

In order to compare  $\mathbf{a}_{\text{hom}}$  and  $\mathbf{a}_{\text{hom},L}$ , the ensembles  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$  have to be *coupled*. In this contribution we do not discuss that issue in general. Instead, we give some simple ideas that are enough to treat the case of i. i. d. coefficients, for which a natural coupling is given by “naïve” periodization. For  $L \in \mathbb{N}$  we consider the mapping

$$(126) \quad \mathbf{a}_L : \Omega \rightarrow \Omega_L, \quad \mathbf{a} \mapsto \mathbf{a}_L(\mathbf{a}) := L\text{-periodic extension of } \mathbf{a}(\cdot)|_{([\frac{-L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d}.$$

**Lemma 20.** Let  $\langle \cdot \rangle$  and  $\langle \cdot \rangle_L$  denote the infinite and  $L$ -periodic i. i. d. ensemble associated with a base measure  $\beta$ . Then for any  $\langle \cdot \rangle_L$ -stationary function  $\zeta$  we have  $\langle \zeta \rangle_L = \langle \zeta \circ \mathbf{a}_L \rangle$ .

In particular, we deduce from Lemma 20 that

$$(127) \quad e \cdot \mathbf{a}_{\text{hom},L} e - e \cdot \mathbf{a}_{\text{hom}} e = \langle (e + \chi) \cdot \mathbf{a}(0)(e + \chi) - (e + \chi_L \circ \mathbf{a}_L) \cdot \mathbf{a}(0)(e + \chi_L \circ \mathbf{a}_L) \rangle,$$

where we use that  $\mathbf{a}(0) \circ \mathbf{a}_L = \mathbf{a}(0)$  on  $\Omega$ . Hence, in order to compare  $\mathbf{a}_{\text{hom}}$  to  $\mathbf{a}_{\text{hom},L}$ , we need to compare  $\chi$  to  $\chi_L \circ \mathbf{a}_L$ . We do this at the level of their stationary extensions, which means that we have to compare the gradients of the correctors on the whole space  $\mathbb{Z}^d$ .

This is a priori a rather subtle task due to the periodic boundary conditions, which generate long range correlations. In order to reduce these correlations, we artificially introduce a zero-order term in the corrector problem, following the original idea of Papanicolaou & Varadhan [32]. For all  $\mu > 0$  let  $\phi_\mu$  be given by (115) and recall that  $\phi_\mu$  is a solution to (114). We define the *modified homogenized coefficients*  $\mathbf{a}_{\text{hom},\mu}$  and  $\mathbf{a}_{\text{hom},L,\mu}$  via

$$\begin{aligned} e \cdot \mathbf{a}_{\text{hom},\mu} e &= \langle (e + D\phi_\mu) \cdot \mathbf{a}(0)(e + D\phi_\mu) \rangle, \\ e \cdot \mathbf{a}_{\text{hom},L,\mu} e &= \langle (e + D\phi_\mu) \cdot \mathbf{a}(0)(e + D\phi_\mu) \rangle_L, \end{aligned}$$

and split the systematic error into the three contributions

$$(128) \quad \begin{aligned} \text{systematic error} &= |e \cdot \mathbf{a}_{\text{hom},L} e - e \cdot \mathbf{a}_{\text{hom}} e| \\ &\leq |e \cdot \mathbf{a}_{\text{hom},L} e - e \cdot \mathbf{a}_{\text{hom},L,\mu} e| \\ &\quad + |e \cdot \mathbf{a}_{\text{hom},L,\mu} e - e \cdot \mathbf{a}_{\text{hom},\mu} e| \\ &\quad + |e \cdot \mathbf{a}_{\text{hom}} e - e \cdot \mathbf{a}_{\text{hom},\mu} e|. \end{aligned}$$

The first and last terms are what we call *systematic sub-errors*, in the spirit of [18] and [15]. These errors will essentially depend on  $\mu$  (and not on  $L$ ). The second term is what we call the *coupling*

error. This is the only term which relates  $\langle \cdot \rangle_L$  to  $\langle \cdot \rangle$  and where the relative scaling of  $\mu$  and  $L$  matters. It is on the level of  $\mathbf{a}_{\text{hom},\mu}$  that we use the coupling (127) by periodization. In order to estimate the systematic error, we estimate each of the three terms, and then *optimize in  $\mu$* . In the present contribution we treat the systematic sub-errors in detail, and treat the *coupling error* only for i. i. d. ensembles — a more thorough analysis will appear in [16] for a larger class of ensembles.

We turn now to the two *systematic sub-errors*.

**Proposition 8.** *Let  $\langle \cdot \rangle$  denote either a stationary ensemble that satisfies  $SG_\infty(\rho)$ , or a stationary,  $L$ -periodic ensemble that satisfies  $SG_L(\rho)$ . For all  $\mu \in (0, 1]$  we have*

$$|e \cdot \mathbf{a}_{\text{hom}} e - e \cdot \mathbf{a}_{\text{hom},\mu} e| \lesssim \begin{cases} \mu^{\frac{d}{2}} & \text{for } 2 \leq d < 4 \\ \mu^2 \ln\left(\frac{1}{\mu} + 1\right) & \text{for } d = 4 \\ \mu^2 & \text{for } d > 4. \end{cases}$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\rho$ ,  $\lambda$  and  $d$ .

Since the analysis for the infinite ensemble and for the periodic ensemble are identical, we display the proof for the systematic sub-error associated with the infinite ensemble  $\langle \cdot \rangle$  only. To this aim we take advantage of the formulation of the systematic error in terms of spectral representation (see [15]). Since the elliptic operator  $D^* \mathbf{a}(0) D$  is bounded, symmetric, and non-negative on  $L^2(\Omega)$ , the spectral theorem yields the existence of a spectral measure  $P(d\nu)$  such that for all  $\zeta \in L^2(\Omega)$ , and suitable continuous functions  $f$ , we have

$$f(D^* \mathbf{a}(0) D) \zeta = \int_0^\infty f(\nu) P(d\nu) \zeta.$$

Hence, with  $\mathfrak{d}$  given by (106), one may write the modified corrector in terms of the spectral resolution

$$\phi_\mu = (\mu + D^* \mathbf{a}(0) D)^{-1} \mathfrak{d} = \int_0^\infty \frac{1}{\mu + \nu} P(d\nu) \mathfrak{d}.$$

This permits to prove an alternative formula for  $\mathbf{a}_{\text{hom},\mu}$ . Indeed, we have

$$\langle (e + D\phi_\mu) \cdot \mathbf{a}(0) (e + D\phi_\mu) \rangle = \langle e \cdot \mathbf{a}(0) e \rangle + 2\langle D\phi_\mu \cdot \mathbf{a}(0) (e + D\phi_\mu) \rangle + \langle D\phi_\mu \cdot \mathbf{a}(0) D\phi_\mu \rangle,$$

which in view of (114) takes the form

$$(129) \quad \langle (e + D\phi_\mu) \cdot \mathbf{a}(0) (e + D\phi_\mu) \rangle = \langle e \cdot \mathbf{a}(0) e \rangle - 2\mu \langle \phi_\mu^2 \rangle - \langle D\phi_\mu \cdot \mathbf{a}(0) D\phi_\mu \rangle.$$

From the spectral representation of  $\mathcal{L}$  we get

$$\begin{aligned} \langle \phi_\mu^2 \rangle &= \int_0^\infty \frac{1}{(\mu + \nu)^2} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle, \\ \langle D\phi_\mu \cdot \mathbf{a}(0) D\phi_\mu \rangle &= \int_0^\infty \frac{\nu}{(\mu + \nu)^2} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle. \end{aligned}$$

Hence,

$$e \cdot \mathbf{a}_{\text{hom},\mu} e = \langle e \cdot \mathbf{a}(0) e \rangle - \int_0^\infty \frac{2\mu + \nu}{(\mu + \nu)^2} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.$$

Note that  $\nu \mapsto \frac{1}{\nu}$  is  $\langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle$ -integrable on  $(0, +\infty)$ . Indeed, by Fatou's Lemma we have

$$\int_0^\infty \frac{1}{\nu} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \leq \liminf_{\mu \downarrow 0} \int_0^\infty \frac{\mu}{(\mu + \nu)^2} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle = \lim_{\mu \downarrow 0} \langle D\phi_\mu \cdot \mathbf{a}(0) D\phi_\mu \rangle < \infty,$$

where in the last step we used the a priori estimate  $\langle D\phi_\mu \mathbf{a}(0) D\phi_\mu \rangle \leq 1$ . As a consequence we get by the Lebesgue dominated convergence theorem,

$$e \cdot \mathbf{a}_{\text{hom}} e = \lim_{\mu \rightarrow 0} e \cdot \mathbf{a}_{\text{hom},\mu} e = \langle e \cdot \mathbf{a}(0) e \rangle - \int_0^\infty \frac{1}{\nu} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.$$

Hence, the difference between  $e \cdot \mathbf{a}_{\text{hom},\mu}e$  and  $e \cdot \mathbf{a}_{\text{hom}}e$  has the representation

$$(130) \quad e \cdot (\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{hom},\mu})e = \mu^2 \int_0^\infty \frac{1}{\nu(\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

The dominant contribution to this integral comes from the bottom of the spectrum: Indeed, using the fundamental theorem of calculus and Fubini's theorem we shall show in the proof of Proposition 8 that

$$\int_0^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \leq (k+3) \int_{\tilde{\nu}=0}^1 \frac{1}{\tilde{\nu}^2(\mu + \tilde{\nu})^{k+2}} \int_{\nu=0}^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle d\tilde{\nu} + \langle \mathfrak{d}^2 \rangle.$$

Hence, for proving a bound on the systematic error, it is enough to have a good estimate on the quantity

$$\int_0^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle = \langle \mathfrak{d}P([0, \tilde{\nu}])\mathfrak{d} \rangle$$

for  $\tilde{\nu} \leq 1$ .

The desired estimate directly follows from the estimate on the semigroup, cf. Theorem 2:

**Corollary 8.** *Let  $d \geq 2$ . Let  $\langle \cdot \rangle$  denote either a stationary ensemble that satisfies  $SG_\infty(\rho)$ , or a stationary,  $L$ -periodic ensemble that satisfies  $SG_L(\rho)$ . We denote by  $P_L(d\nu)$  the spectral measure of the operator  $D^*\mathbf{a}(0)D$ , and by  $\mathfrak{d} \in L^2(\Omega)$  the right-hand side of the corrector problem  $-D^*\mathbf{a}(0)e$  in some direction  $e \in \mathbb{R}^d$  with  $|e| = 1$ . Then, for all  $0 \leq \tilde{\nu} \leq 1$ , we have*

$$\int_0^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \lesssim \tilde{\nu}^{\frac{d}{2}+1},$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\rho$ ,  $\lambda$  and  $d$ .

This corollary of Theorem 2 answers positively a conjecture of [15].

Before we turn to the complete estimate of the error  $\langle |\mathbf{a}_{\text{av},L} - \mathbf{a}_{\text{hom}}|^2 \rangle$ , we present the estimate on the coupling error in the i. i. d. case (the result for more general ensembles will be discussed in a forthcoming work [16]).

**Proposition 9.** *Let  $d \geq 2$ ,  $L \in \mathbb{N}$ . Let  $\langle \cdot \rangle$  be the infinite i. i. d. ensemble associated with some base measure  $\beta$ , and let  $\langle \cdot \rangle_L$  denote the  $L$ -periodic i. i. d. ensemble associated with  $\beta$ . Then, for all  $\mu \in (0, 1]$  and any direction  $e \in \mathbb{R}^d$ ,  $|e| = 1$ , we have*

$$(131) \quad |e \cdot \mathbf{a}_{\text{hom},L,\mu}e - e \cdot \mathbf{a}_{\text{hom},\mu}e| \lesssim \sqrt{\mu}^{-(d+1)} \exp(-c_0\sqrt{\mu}L),$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\lambda$  and  $d$ , and  $c_0 > 0$  is a constant depending only on  $\lambda$  and  $d$ .

**Remark 16.** *The estimate in Proposition 9 relies on a uniform estimate on the gradient of the elliptic Green's function  $G_\mu$  associated with the elliptic operator  $\mu + \nabla_x \mathbf{a}(x) \nabla_x$  of the modified corrector problem in physical space. The key feature of that estimate is an exponential decay in  $\sqrt{\mu}|x|$  away from the diagonal. In the proof of Proposition 9 we content ourselves with a simple-minded estimate that we obtain by combining the discrete estimate  $|\nabla G_\mu| \lesssim G_\mu$  with a (non-optimal) pointwise estimate for the elliptic Green's function itself. The advantage of that reasoning is that it is easy to prove. However, in that way only get the exponent  $(d+1)$  of the prefactor in (131) – which is not optimal. Yet, as we shall see in Lemma 21 the prefactor only induces a logarithmic correction and does not play a role for the complete error estimate.*

We continue with our discussion of the complete error estimate. We are in position to give an estimate for the systematic error, which is optimal (up to a logarithm) in dimensions  $d = 2, 3$ . The combination of (128), Proposition 8, and Proposition 9 yields for  $d = 2, 3$ :

$$(132) \quad \langle |\mathbf{a}_{\text{hom},L} - \mathbf{a}_{\text{hom}}|^2 \rangle^{\frac{1}{2}} \lesssim \sqrt{\mu}^d + \frac{1}{\sqrt{\mu}^{d+1}} \exp(-c_0\sqrt{\mu}L).$$

We optimize that expression in  $\mu$ :

**Lemma 21.** For  $\alpha > 0$ ,  $c_0 > 0$  and  $L \gg 1$  we have

$$\left(\frac{\ln L}{L}\right)^d \lesssim \min_{\mu > 0} \left( \sqrt{\mu}^d + \frac{1}{\sqrt{\mu}^\alpha} \exp(-c_0 \sqrt{\mu} L) \right) \lesssim \left(\frac{\ln L}{L}\right)^d$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $\alpha$ ,  $c_0$  and  $d$ .

Hence, in dimension  $d = 2, 3$  we established that the systematic error decays with the rate  $-d$  (up to a logarithmic correction). In combination with (125) and Corollary 7 we get for  $d = 2, 3$  an optimal estimate of the complete error:

$$(133) \quad \langle |\mathbf{a}_{\text{av},L,N} - \mathbf{a}_{\text{hom}}|^2 \rangle^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{N}} L^{-\frac{d}{2}} + L^{-d} \ln^d L.$$

However, Proposition 8, that itself is optimal for all  $d$ , does not yield (133) for  $d > 3$ , since the scaling exponent of the systematic sub-error  $\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{hom},\mu}$  saturates at  $d = 4$ . In order to reduce the systematic sub-error, one needs to use higher order approximations of the homogenized coefficients than  $\mathbf{a}_{\text{hom},L,\mu}$ , as introduced in [15]. Note that they – as  $\mathbf{a}_{\text{hom},L,\mu}$  itself – are just an ingredient to the *proof* of the complete error estimate, and are (at least here) not seen as a refined approximation scheme. We present here a variant of this approach based on a Richardson extrapolation scheme. As a motivation for this, consider for a moment a function  $f_0(\mu)$  with expansion at 0 of order  $M \in \mathbb{N}$ :

$$f_0(\mu) = f_0(0) + c_2 \mu^2 + \dots + c_N \mu^M + O(\mu^{M+1}).$$

For all  $k = 1, 2, 3, \dots$  we iteratively define the Richardson extrapolations

$$f_k(\mu) = \frac{1}{2^{k+1} - 1} \left( 2^{k+1} f_{k-1}(\mu) - f_{k-1}(2\mu) \right).$$

Then it is easy to check that

$$f_k(\mu) = f_0(0) + \begin{cases} o(\mu^{k+1}) & \text{for } k < M, \\ O(\mu^M) & \text{for } k \geq M. \end{cases}$$

We note that the  $k$ th extrapolation can be written as a superposition of the terms  $f_0(\mu), \dots, f_0(2^k \mu)$ . In particular, there exist universal coefficients  $a_{k,0}, \dots, a_{k,k}$  with

$$f_k(\mu) = \sum_{\ell=0}^k a_{k,\ell} f_0(2^\ell \mu).$$

We make use of the Richardson extrapolation to approximate homogenized coefficients as follows:

**Proposition 10.** Let  $d \geq 2$ . Let  $\langle \cdot \rangle$  denote either a stationary ensemble that satisfies  $SG_\infty(\rho)$ , or a stationary,  $L$ -periodic ensemble that satisfies  $SG_L(\rho)$ . For all  $\mu > 0$  we consider the superposition

$$\mathbf{a}_{\text{hom},\mu}^k := \sum_{\ell=0}^k a_{k,\ell} \mathbf{a}_{\text{hom},2^\ell \mu}.$$

Then for all  $k > \frac{d}{2} - 2$  we have

$$|\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{hom},\mu}^k| \lesssim \mu^{\frac{d}{2}}$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $k, \rho$  and  $d$ .

We may finally state the complete error estimate:

**Theorem 4.** Let  $d \geq 2$ ,  $L, N \in \mathbb{N}$ . Let  $\langle \cdot \rangle$  be the infinite i. i. d. ensemble associated with some base measure  $\beta$ , and let  $\langle \cdot \rangle_L$  denote the  $L$ -periodic i. i. d. ensemble associated with  $\beta$ . Then we have

$$\langle |\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{av},L,N}|^2 \rangle_L^{\frac{1}{2}} \lesssim \frac{1}{\sqrt{N}} L^{-\frac{d}{2}} + L^{-d} \ln^d L,$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\lambda$  and  $d$ .

The proof relies on the following triangle inequality on the systematic error:

$$\begin{aligned} \text{systematic error} &= |e \cdot \mathbf{a}_{\text{hom},L} e - e \cdot \mathbf{a}_{\text{hom}} e| \\ &\leq |e \cdot \mathbf{a}_{\text{hom},L} e - e \cdot \mathbf{a}_{\text{hom},L,\mu}^k e| \\ &\quad + |e \cdot \mathbf{a}_{\text{hom},L,\mu}^k e - e \cdot \mathbf{a}_{\text{hom},\mu}^k e| \\ &\quad + |e \cdot \mathbf{a}_{\text{hom}} e - e \cdot \mathbf{a}_{\text{hom},\mu}^k e| \end{aligned}$$

with  $k$  chosen so that  $k > \frac{d}{2} - 2$ . The first and last terms of the r. h. s. are dealt with in Proposition 10, whereas the coupling error is again controlled by Proposition 9 since  $e \cdot \mathbf{a}_{\text{hom},L,\mu}^k e - e \cdot \mathbf{a}_{\text{hom},\mu}^k e$  is just a linear combination of terms  $e \cdot \mathbf{a}_{\text{hom},L,2^j\mu} e - e \cdot \mathbf{a}_{\text{hom},2^j\mu} e$  (for  $j \in \{1, \dots, k\}$ ), to which Proposition 9 applies.

### 6.3.1 Proofs

*Proof of Proposition 7.* By Definition 12 and Lemma 18 we have

$$(134) \quad \mathcal{E} = e \cdot \mathbf{a}_{\text{av},L} e = \frac{1}{L^d} \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \overline{(e + D\phi) \cdot \mathbf{a}(0)(e + D\phi)}(x).$$

In the following we drop the subindex  $L$  in the notation for  $\frac{\partial}{\partial L y}$  and  $\langle \cdot \rangle_{L,y}$ .

**Step 1.** Estimate for  $\nabla \frac{\partial \bar{\phi}}{\partial y}$ .

We claim that

$$(135) \quad \sum_{x \in ([0,L) \cap \mathbb{Z})^d} |\nabla \frac{\partial \bar{\phi}}{\partial y}(x)|^2 \lesssim \langle |e + \nabla \bar{\phi}(y)|^2 \rangle_y.$$

Since  $\langle \cdot \rangle$  is  $L$ -periodic, both  $\bar{\phi}$  and its derivative  $\frac{\partial \bar{\phi}}{\partial y}$  are  $L$ -periodic. By Lemma 18,  $\bar{\phi}$  is characterized by the  $L$ -periodic corrector problem (123), which in its weak form reads

$$(136) \quad \sum_{x \in ([0,L) \cap \mathbb{Z})^d} \nabla \varphi(x) \cdot \mathbf{a}(x)(e + \nabla \bar{\phi}(\mathbf{a}, x)) = 0$$

for all  $L$ -periodic test functions  $\varphi : \mathbb{Z}^d \rightarrow \mathbb{R}$ . By taking the vertical derivative of (136) w. r. t.  $y$ , and using the modified Leibniz rule

$$\frac{\partial(\zeta_1 \zeta_2)}{\partial y} = \zeta_1 \frac{\partial \zeta_2}{\partial y} + \zeta_1 \langle \zeta_2 \rangle_y - \langle \zeta_1 \zeta_2 \rangle_y,$$

we obtain

$$\sum_{x \in ([0,L) \cap \mathbb{Z})^d} \nabla \varphi(x) \cdot \mathbf{a}(x) \nabla \frac{\partial \bar{\phi}}{\partial y}(x) = -\nabla \varphi(y) \cdot \left( \mathbf{a}(y) \langle e + \nabla \bar{\phi}(y) \rangle_y - \langle \mathbf{a}(y)(e + \nabla \bar{\phi}(y)) \rangle_y \right)$$

for all  $y \in ([0,L) \cap \mathbb{Z})^d$ . For  $\varphi = \frac{\partial \bar{\phi}}{\partial y}$  this yields

$$(137) \quad \begin{aligned} &\lambda \sum_{x \in ([0,L) \cap \mathbb{Z})^d} |\nabla \frac{\partial \bar{\phi}}{\partial y}(x)|^2 \\ &\leq -\nabla \frac{\partial \bar{\phi}}{\partial y}(y) \cdot \left( \mathbf{a}(y) \langle e + \nabla \bar{\phi}(y) \rangle_y - \langle \mathbf{a}(y)(e + \nabla \bar{\phi}(y)) \rangle_y \right) \\ &\lesssim |\nabla \frac{\partial \bar{\phi}}{\partial y}(y)| \langle |e + \nabla \bar{\phi}(y)| \rangle_y. \end{aligned}$$

In particular, (137) implies

$$|\nabla \frac{\partial \bar{\phi}}{\partial y}(y)|^2 \lesssim \langle |e + \nabla \bar{\phi}(y)|^2 \rangle_y.$$

Inserting this into (137) yields the desired estimate.

**Step 2.** Estimate for  $\frac{\partial \mathcal{E}}{\partial y}$  (cf. (134)).

We claim that

$$\left| \frac{\partial \mathcal{E}}{\partial y} \right| \lesssim L^{-d} \left( |e + \nabla \bar{\phi}(y)|^2 + \langle |e + \nabla \bar{\phi}(y)|^2 \rangle_y \right).$$

Indeed, by the modified Leibniz rule

$$\frac{\partial(\zeta_1 \zeta_2)}{\partial y} = \zeta_1 \frac{\partial \zeta_2}{\partial y} + \frac{\partial \zeta_1}{\partial y} \zeta_2 - \frac{\partial \zeta_1}{\partial y} \frac{\partial \zeta_2}{\partial y} - \left\langle \frac{\partial \zeta_1}{\partial y} \frac{\partial \zeta_2}{\partial y} \right\rangle_y,$$

we have

$$\begin{aligned} \frac{\partial \mathcal{E}}{\partial y} &= \frac{1}{L^d} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \mathbf{a}(x) : \frac{\partial}{\partial y} \left( (e + \nabla \bar{\phi}(x)) \otimes (e + \nabla \bar{\phi}(x)) \right) \\ &\quad + \frac{1}{L^d} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \frac{\partial \mathbf{a}(x)}{\partial y} : (e + \nabla \bar{\phi}(x)) \otimes (e + \nabla \bar{\phi}(x)) \\ &\quad - \frac{1}{L^d} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \frac{\partial \mathbf{a}(x)}{\partial y} : \frac{\partial}{\partial y} \left( (e + \nabla \bar{\phi}(x)) \otimes (e + \nabla \bar{\phi}(x)) \right) \\ &\quad - \frac{1}{L^d} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \left\langle \frac{\partial \mathbf{a}(x)}{\partial y} : \frac{\partial}{\partial y} \left( (e + \nabla \bar{\phi}(x)) \otimes (e + \nabla \bar{\phi}(x)) \right) \right\rangle_y. \end{aligned}$$

For convenience, we denote by  $I_1$  the first term of the r. h. s., and by  $I_2$  the sum of the other three terms. Since

$$\left| \frac{\partial \mathbf{a}(x)}{\partial y} \right| \leq \begin{cases} 2 & \text{if } x - y \in LZ^d, \\ 0 & \text{else,} \end{cases}$$

and by using  $\left| \frac{\partial \zeta}{\partial y} \right| \leq |\zeta| + \langle |\zeta| \rangle_y$ , we get

$$|I_2| \lesssim L^{-d} \left( |\nabla \bar{\phi}(y) + e|^2 + \langle |\nabla \bar{\phi}(y) + e|^2 \rangle_y \right).$$

For the estimate of  $I_1$  we again appeal to the modified Leibniz rule:

$$(138) \quad \begin{aligned} I_1 &= \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \mathbf{a}(x) : \frac{\partial}{\partial y} \left( (e + \nabla \bar{\phi}(x)) \otimes (e + \nabla \bar{\phi}(x)) \right) \\ &= 2 \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \mathbf{a}(x) : (e + \nabla \bar{\phi}(x)) \otimes \nabla \frac{\partial \bar{\phi}}{\partial y}(x) \end{aligned}$$

$$(139) \quad \left. \begin{aligned} &- \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \mathbf{a}(x) : \nabla \frac{\partial \bar{\phi}}{\partial y}(x) \otimes \nabla \frac{\partial \bar{\phi}}{\partial y}(x) \\ &- \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \mathbf{a}(x) : \left\langle \nabla \frac{\partial \bar{\phi}}{\partial y}(x) \otimes \nabla \frac{\partial \bar{\phi}}{\partial y}(x) \right\rangle_y. \end{aligned} \right\}$$

Notice that (138) is precisely the weak form of the Euler-Lagrange equation associated with the minimization problem that defines  $\mathbf{a}_{\text{av}, L}$ , cf. (136). Since  $x \mapsto \frac{\partial \bar{\phi}}{\partial y}(x)$  is an admissible test function, (138) vanishes. For the estimate of (139) we appeal to Step 1:

$$|(139)| \lesssim \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\nabla \frac{\partial \bar{\phi}}{\partial y}(x)|^2 + \langle |\nabla \frac{\partial \bar{\phi}}{\partial y}(x)|^2 \rangle_y \lesssim \langle |e + \nabla \bar{\phi}(y)|^2 \rangle_y.$$



**Step 3.** Conclusion via spectral gap estimate.

We apply the spectral gap estimate to  $e \cdot \mathbf{a}_{L,\text{hom}}e$ , use Step 2, and Jensen's inequality to bound the variance of  $e \cdot \mathbf{a}_{L,\text{hom}}e$  by the quartic moment of  $D\phi$ :

$$\begin{aligned} \text{Var}[e \cdot \mathbf{a}_{L,\text{hom}}e] &= \langle (\mathcal{E} - \langle \mathcal{E} \rangle)^2 \rangle \leq \frac{1}{\rho} \sum_{y \in ([0,L] \cap \mathbb{Z})^d} \left\langle \left( \frac{\partial \mathcal{E}}{\partial y} \right)^2 \right\rangle \\ &\stackrel{\text{Step 2}}{\lesssim} L^{-2d} \sum_{y \in ([0,L] \cap \mathbb{Z})^d} \langle |e + \nabla \bar{\phi}(y)|^4 \rangle + \langle \langle |e + \nabla \bar{\phi}(y)|^2 \rangle_y^2 \rangle \\ &\stackrel{\text{stationarity, Jensen}}{=} L^{-2d} \sum_{y \in ([0,L] \cap \mathbb{Z})^d} \langle |e + D\phi|^4 \rangle = L^{-d} \langle |e + D\phi|^4 \rangle. \end{aligned}$$

Since  $\langle |D\phi|^4 \rangle = \langle |\chi|^4 \rangle \lesssim 1$  by Proposition 6, the statement follows.  $\square$

*Proof of Corollary 8.* We simply apply the semigroup  $t \mapsto \exp(-tD^* \mathbf{a}(0)D)$  of Theorem 2 to  $\mathfrak{d}$ , and define for all  $t \geq 0$

$$u(t) := \exp(-tD^* \mathbf{a}(0)D) \mathfrak{d}.$$

By the spectral theorem,  $u(t) = \int_0^\infty \exp(-t\nu) P(d\nu) \mathfrak{d}$ , so that for all  $t > 0$ ,

$$\langle u^2(t) \rangle = \int_0^\infty \exp(-2t\nu) \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \geq \exp(-2) \int_0^{\frac{1}{t}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.$$

Corollary 8 thus follows from the estimate  $\langle u^2(t) \rangle \lesssim (t+1)^{-(\frac{d}{2}+1)}$  of Theorem 2.  $\square$

*Proof of Proposition 8.* We essentially follow the argument of [15]. The starting point is formula (130):

$$e \cdot (\mathbf{a}_{\text{hom}} - \mathbf{a}_{\text{hom},\mu})e = \mu^2 \int_0^\infty \frac{1}{\nu(\mu + \nu)^2} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.$$

Let us directly consider a more general version of this integral (which will be needed in the proof of Proposition 10): For all  $k = 0, 1, 2, 3, \dots$  we consider

$$\int_0^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle.$$

We split the integral into two parts to single out the bottom of the spectrum:

$$\begin{aligned} &\int_0^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \\ &= \int_0^1 \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle + \int_1^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \\ (140) \quad &\leq \int_0^1 \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle + \int_1^\infty \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle. \end{aligned}$$

To make the bottom of the spectrum explicitly appear, we write the first integrand as an integral: By the fundamental theorem of calculus we have for all  $f \in C^1((0, 1])$

$$\begin{aligned} \int_0^1 f(\nu) \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle &= - \int_{\nu=0}^1 \int_{\tilde{\nu}=\nu}^1 f'(\tilde{\nu}) d\tilde{\nu} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle + f(1) \int_{\nu=0}^1 \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle \\ &= - \int_{\tilde{\nu}=0}^1 f'(\tilde{\nu}) \int_{\nu=0}^{\tilde{\nu}} \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle d\tilde{\nu} + f(1) \int_{\nu=0}^1 \langle \mathfrak{d} P(d\nu) \mathfrak{d} \rangle. \end{aligned}$$

Applied with  $f(\nu) = \frac{1}{\nu(\mu+\nu)^{k+2}}$  we get

$$\begin{aligned}
& \int_0^1 \frac{1}{\nu(\mu+\nu)^{k+2}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle \\
&= \int_{\tilde{\nu}=0}^1 \frac{\mu + (k+3)\tilde{\nu}}{\tilde{\nu}^2(\mu+\tilde{\nu})^{k+3}} \int_{\nu=0}^{\tilde{\nu}} d\langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle d\tilde{\nu} + \frac{1}{(\mu+1)^{k+2}} \int_{\nu=0}^1 \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle \\
(141) \quad &\leq (k+3) \int_{\tilde{\nu}=0}^1 \frac{1}{\tilde{\nu}^2(\mu+\tilde{\nu})^{k+2}} \int_{\nu=0}^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle d\tilde{\nu} + \int_0^1 \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle.
\end{aligned}$$

From (140), (141), and the identity  $\int_0^\infty \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle = \langle \mathfrak{d}^2 \rangle$ , we deduce that

$$(142) \quad \int_0^\infty \frac{1}{\nu(\mu+\nu)^{k+2}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle \leq (k+3) \int_{\tilde{\nu}=0}^1 \frac{1}{\tilde{\nu}^2(\mu+\tilde{\nu})^{k+2}} \int_{\nu=0}^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle d\tilde{\nu} + \langle \mathfrak{d}^2 \rangle.$$

Let us assume that the bottom of the spectrum satisfies an estimate of the form

$$(143) \quad \int_0^{\tilde{\nu}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle \lesssim \tilde{\nu}^\gamma \quad \text{for some } \gamma \geq 0.$$

Then, by the change of variables  $\mu u = \tilde{\nu}$ , (142) turns into

$$\begin{aligned}
\int_0^\infty \frac{1}{\nu(\mu+\nu)^{k+2}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle &\lesssim \int_0^1 \frac{\tilde{\nu}^{\gamma-2}}{(\mu+\tilde{\nu})^{k+2}} d\tilde{\nu} + 1 \\
&= \mu^{\gamma-3-k} \int_0^{\frac{1}{\mu}} \frac{u^{\gamma-2}}{(1+u)^{k+2}} du + 1.
\end{aligned}$$

For  $\gamma \geq 2$  this estimate can be simplified to

$$(144) \quad \int_0^\infty \frac{1}{\nu(\mu+\nu)^{k+2}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle \lesssim \mu^{\gamma-3-k} \int_0^{\frac{1}{\mu}} (1+u)^{\gamma-k-4} + 1.$$

The claim of the proposition follows via (130) from (144) for  $k=0$ , using that (143) holds with  $\gamma = \frac{d}{2} + 1$  by Corollary 8.  $\square$

*Proof of Lemma 21.* The upper estimate follows with  $\sqrt{\mu} = \frac{\alpha+d}{c_0} L^{-1} \ln(\frac{L}{\ln L})$ . Indeed, we have

$$\begin{aligned}
(145) \quad \sqrt{\mu}^d + \frac{1}{\sqrt{\mu}^\alpha} \exp(-c_0 \sqrt{\mu} L) &\lesssim \left(\frac{\ln L}{L}\right)^d + \left(\frac{L}{\ln L}\right)^\alpha \exp((\alpha+d) \ln \frac{\ln L}{L}) \\
&\lesssim \left(\frac{\ln L}{L}\right)^d + \left(\frac{L}{\ln L}\right)^\alpha \left(\frac{\ln L}{L}\right)^{\alpha+d} \lesssim \left(\frac{\ln L}{L}\right)^d.
\end{aligned}$$

For the lower estimate let  $\mu_*$  denote the minimizer of  $\sqrt{\mu}^d + \sqrt{\mu}^{-\alpha} \exp(-c_0 \sqrt{\mu} L)$ . By (145) we have

$$\sqrt{\mu_*}^d + \frac{1}{\sqrt{\mu_*}^\alpha} \exp(-c_0 \sqrt{\mu_*} L) \lesssim \left(\frac{\ln L}{L}\right)^d,$$

and thus

$$(146) \quad \sqrt{\mu_*} \lesssim \frac{\ln L}{L},$$

$$(147) \quad \frac{1}{\sqrt{\mu_*}^\alpha} \exp(-c_0 \sqrt{\mu_*} L) \leq 1 \quad \text{for } L \gg 1.$$

By taking the logarithm in (147) we get  $-c_0 \sqrt{\mu_*} L \leq \alpha \ln \sqrt{\mu_*}$ , and thus

$$(148) \quad \sqrt{\mu_*} \geq \frac{\alpha \ln(\frac{1}{\sqrt{\mu_*}})}{c_0 L}.$$

By taking the logarithm of the reciprocal inequality of (146), we get  $\ln(\frac{1}{\sqrt{\mu_*}}) \gtrsim \ln L - \ln(\ln L) \geq \frac{1}{2} \ln L$  for  $L \gg 1$ , which combines with (148) to  $\sqrt{\mu_*} \gtrsim \frac{\ln L}{L}$ , and thus

$$\sqrt{\mu_*}^d + \sqrt{\mu_*}^{-\alpha} \exp(-c_0 \sqrt{\mu_*} L) \geq \sqrt{\mu_*}^d \gtrsim \left(\frac{\ln L}{L}\right)^d.$$

□

For the proofs of Proposition 9 and 10, which deal with the coupling error, we have to estimate the difference  $\phi_\mu \circ \mathbf{a}_L - \phi_\mu$ . We do that on the level of the stationary extension. As a preparation we establish a Green's representation formula for bounded solutions to elliptic equations of the form

$$(149) \quad \mu\varphi + \nabla^* \mathbf{a} \nabla \varphi = f \quad \text{in } \mathbb{Z}^d.$$

**Lemma 22.** *Let  $\mathbf{a} \in \Omega$  and  $\varphi(x), f(x)$  be bounded functions on  $\mathbb{Z}^d$  related by (149). Then*

$$\varphi(x) = \sum_{z \in \mathbb{Z}^d} G_\mu(\mathbf{a}, x, z) f(z)$$

where

$$G_\mu(\mathbf{a}, x, z) := \int_0^\infty \exp(-\mu t) G(t, \mathbf{a}, x, z) dt$$

denotes the modified elliptic Green's function.

*Proof.* Since  $\sum_{z \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z) = 1$  by (26c), we deduce that  $G_\mu(\mathbf{a}, x, z)$  is in  $\ell_x^1(\mathbb{Z}^d)$ :

$$\mu \sum_{x \in \mathbb{Z}^d} G_\mu(\mathbf{a}, x, z) = \mu \int_0^\infty \exp(-\mu t) \sum_{x \in \mathbb{Z}^d} G(t, \mathbf{a}, x, z) dt = 1.$$

Moreover, it satisfies the equation

$$\mu G_\mu(\mathbf{a}, x, z) + \nabla_x^* \mathbf{a}(x) \nabla G_\mu(\mathbf{a}, x, z) = \delta(x - z).$$

Indeed, we have

$$\begin{aligned} \nabla_x^* \mathbf{a}(x) \nabla_x G(\mathbf{a}, x, z) &= \int_0^\infty \exp(-\mu t) \nabla_x^* \mathbf{a}(x) \nabla_x G(t, \mathbf{a}, x, z) dt \\ &= \delta(x - z) - \mu G_\mu(\mathbf{a}, x, z). \end{aligned}$$

After these preparations we multiply (149) with  $G_\mu(\mathbf{a}, x, z)$ , sum in  $x$  and integrate by parts:

$$\sum_{x \in \mathbb{Z}^d} f(x) G_\mu(\mathbf{a}, x, z) = \sum_{x \in \mathbb{Z}^d} \varphi(x) (\mu G_\mu(\mathbf{a}, x, z) + \nabla_x^* \mathbf{a}(x) \nabla G_\mu(\mathbf{a}, x, z)) = \varphi(z).$$

□

We need the following estimate on  $G_\mu(x, z)$ :

**Lemma 23.** *For  $\mu \in (0, 1]$  and  $x, z \in \mathbb{Z}^d$  we have*

$$G_\mu(\mathbf{a}, x, z) \lesssim \frac{1}{\sqrt{\mu}} \exp(-c_0 \sqrt{\mu} |x - z|)$$

where  $c_0$  denotes a positive constant that only depends on  $\lambda$  and  $d$ .

*Proof.* It suffices to consider the case  $z = 0$ . The elliptic Green's function  $G_\mu(x) := G_\mu(\mathbf{a}, x, 0)$  satisfies

$$(150) \quad (\mu + \nabla^* \mathbf{a}(x) \nabla) G_\mu(x) = \delta(x).$$

For the cut-off parameter  $M$  we introduce the exponentially growing test-function  $\zeta_M(x) := \exp(c_0 \sqrt{\mu} \min\{|x|, M\})$  where  $c_0$  is a small positive constant to be chosen later. For convenience we assume that  $c_0 \sqrt{\mu} \leq 1$ . As can be easily checked we have

$$(151) \quad |\nabla_i \zeta_M(x)| \lesssim c_0 \sqrt{\mu} \zeta_M(x) \quad \text{and} \quad \zeta_M(x - e_i) \lesssim \zeta_M(x).$$

Since  $G_\mu$  is summable, testing (150) with the bounded function  $G_\mu \zeta_M$  and an integration by parts yield

$$\begin{aligned} 1 &= \mu \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M + \sum_{x \in \mathbb{Z}^d} \nabla G_\mu \cdot \mathbf{a} \nabla (G_\mu \zeta_M) \\ &\stackrel{\text{diagonality, ellipticity}}{\geq} \mu \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M + \lambda \sum_{x \in \mathbb{Z}^d} |\nabla G_\mu|^2 \zeta_M \\ &\quad - \sum_{x \in \mathbb{Z}^d} \sum_{i=1}^d |\nabla_i G_\mu| |G_\mu(x + e_i)| |\nabla_i \zeta_M|. \end{aligned}$$

Rearranging terms, application of (151) and Cauchy-Schwarz yield

$$\mu \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M + \lambda \sum_{x \in \mathbb{Z}^d} |\nabla G_\mu|^2 \zeta_M \leq C c_0 \sqrt{\mu} \left( \sum_{x \in \mathbb{Z}^d} |\nabla G_\mu|^2 \zeta_M \right)^{\frac{1}{2}} \left( \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M \right)^{\frac{1}{2}} + 1$$

for some positive constant  $C > 0$  that only depends on the dimension  $d$ . By appealing to Young's inequality we may absorb the second factor on the right-hand side into the first term of the l. h. s.:

$$\frac{\mu}{2} \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M \leq \left( \frac{C^2 c_0^2}{2} - \lambda \right) \sum_{x \in \mathbb{Z}^d} |\nabla G_\mu|^2 \zeta_M + 1.$$

Since  $C$  is independent of  $c_0$ , we can make the prefactor on the right-hand side negative by choosing  $c_0$  sufficiently small. We get

$$\sum_{x \in \mathbb{Z}^d} G_\mu^2 \exp(c_0 \sqrt{\mu} |x|) = \lim_{M \uparrow \infty} \sum_{x \in \mathbb{Z}^d} G_\mu^2 \zeta_M \lesssim \frac{1}{\mu},$$

which combined with the discrete  $\ell^\infty$ - $\ell^2$ -estimate leads to

$$\begin{aligned} G_\mu(x) &= \exp(-\frac{c_0}{2} \sqrt{\mu} |x|) G_\mu(x) \exp(\frac{c_0}{2} \sqrt{\mu} |x|) \\ &\leq \exp(-\frac{c_0}{2} \sqrt{\mu} |x|) \left( \sum_{y \in \mathbb{Z}^d} G_\mu^2 \exp(c_0 \sqrt{\mu} |y|) \right)^{\frac{1}{2}} \\ &\lesssim \exp(-\frac{c_0}{2} \sqrt{\mu} |x|) \frac{1}{\sqrt{\mu}}. \end{aligned}$$

□

*Proof of Proposition 9.* By appealing to the coupling  $\mathbf{a}_L : \Omega \rightarrow \Omega_L$  we have

$$\begin{aligned} e \cdot \mathbf{a}_{\text{hom}, L, \mu} e &= \langle (e + D\phi_\mu) \cdot \mathbf{a}(0) (e + D\phi_\mu) \rangle_L \\ &\stackrel{\text{Lemma 20}}{=} \left\langle \left\{ (e + D\phi_\mu) \cdot \mathbf{a}(0) (e + D\phi_\mu) \right\} \circ \mathbf{a}_L \right\rangle \\ &\stackrel{\mathbf{a}(0) \circ \mathbf{a}_L = \mathbf{a}(0)}{=} \langle (e + D\phi_\mu \circ \mathbf{a}_L) \cdot \mathbf{a}(0) (e + D\phi_\mu \circ \mathbf{a}_L) \rangle, \end{aligned}$$

so that

$$\begin{aligned}
& e \cdot (\mathbf{a}_{\text{hom},L,\mu} - \mathbf{a}_{\text{hom},\mu})e \\
&= \langle (e + D\phi_\mu \circ \mathbf{a}_L) \cdot \mathbf{a}(0)(e + D\phi_\mu \circ \mathbf{a}_L) - (e + D\phi_\mu) \cdot \mathbf{a}(0)(e + D\phi_\mu) \rangle \\
&= \langle (D\phi_\mu \circ \mathbf{a}_L - D\phi_\mu) \cdot \mathbf{a}(0)(2e + D\phi_\mu + D\phi_\mu \circ \mathbf{a}_L) \rangle \\
&\stackrel{\text{Lemma 20}}{\lesssim} \langle |D\phi_\mu \circ \mathbf{a}_L - D\phi_\mu|^2 \rangle^{1/2} (1 + \langle |D\phi_\mu|^2 \rangle^{1/2} + \langle |D\phi_\mu|^2 \rangle_L^{1/2}) \\
&\lesssim \langle |D\phi_\mu \circ \mathbf{a}_L - D\phi_\mu|^2 \rangle^{1/2}.
\end{aligned}$$

For the last line we used the boundedness of  $\langle |D\phi_\mu|^2 \rangle$  and  $\langle |D\phi_\mu|^2 \rangle_L$ , which we obtain by the standard  $L^2$ -a priori estimate from (114). Hence it remains to estimate  $\langle |D\phi_\mu \circ \mathbf{a}_L - D\phi_\mu|^2 \rangle^{1/2}$ . In fact, by discreteness, it suffices to estimate  $\langle |\phi_\mu \circ \mathbf{a}_L - \phi_\mu|^2 \rangle^{1/2}$ . We shall do this on the level of the stationary extension. To that end we introduce the non-stationary random-field

$$\psi_{L,\mu}(\mathbf{a}, x) := \overline{\phi_\mu}(\mathbf{a}_L(\mathbf{a}), x) - \overline{\phi_\mu}(\mathbf{a}, x),$$

with help of which the desired estimate can be written as

$$(152) \quad \langle |\psi_{L,\mu}(\mathbf{a}, 0)|^2 \rangle^{1/2} \lesssim \mu^{-\frac{d+1}{2}} \exp(-\frac{c_0}{4} \sqrt{\mu} L).$$

For the proof of (152) let  $G_\mu(\mathbf{a}, 0, z) = \int_0^\infty \exp(-\mu t) G(t, \mathbf{a}, 0, z) dt$  denote the elliptic Green's function. We claim that

$$(153) \quad |\psi_{L,\mu}(\mathbf{a}, 0)| \leq \sum'_z |\nabla_z G_\mu(\mathbf{a}, 0, z)| |\nabla \overline{\phi_\mu}(\mathbf{a}_L(\mathbf{a}), z) + e|,$$

where  $\sum'_z$  stands for  $\sum_{z \in \mathbb{Z}^d \setminus ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d}$ . Indeed, since  $\phi_\mu$  is a  $C^0(\Omega)$  solution to (114), cf. Lemma 17, we deduce that  $x \mapsto \overline{\phi_\mu}(\mathbf{a}, x)$  is a bounded solution to (149) with right-hand side  $f(x) = -\nabla^* \mathbf{a}(x)e$ . Hence,  $x \mapsto \psi_{L,\mu}(\mathbf{a}, x)$  is a bounded solution to (149) with right-hand side  $f(x) = \nabla^*(\mathbf{a} - \mathbf{a}_L(\mathbf{a}))(x)(e + \nabla \overline{\phi_\mu}(\mathbf{a}_L(\mathbf{a}), x))$  and Lemma 22 yields

$$\psi_{L,\mu}(\mathbf{a}, 0) = \sum_{z \in \mathbb{Z}^d} \nabla_z G_\mu(\mathbf{a}, 0, z) \cdot (\mathbf{a} - \mathbf{a}_L(\mathbf{a}))(z)(e + \nabla \overline{\phi_\mu}(\mathbf{a}_L(\mathbf{a}), z)).$$

Now, (153) follows from

$$|(\mathbf{a} - \mathbf{a}_L(\mathbf{a}))(x)| \leq \begin{cases} 0 & \text{for } x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d, \\ 1 & \text{else.} \end{cases}$$

With (153) at hand we can draw the conclusion by appealing to Lemma 23. Indeed, by discreteness Lemma 23 yields the estimate

$$|\nabla G_\mu(\mathbf{a}, 0, z)| \lesssim \frac{1}{\sqrt{\mu}} \exp(-c_0 \sqrt{\mu} |z|),$$

which is uniform in  $\mathbf{a}$ . Combined with (153) we get

$$\begin{aligned}
\langle \psi_{L,\mu}^2(\cdot, 0) \rangle &\lesssim \left\langle \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) |\nabla \overline{\phi}_\mu(\mathbf{a}_L(\cdot), z) + e| \right)^2 \right\rangle \\
&\stackrel{\text{Cauchy-Schwarz}}{\leq} \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \right) \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \langle |\nabla \overline{\phi}_\mu(\mathbf{a}_L(\cdot), z) + e|^2 \rangle \\
&\stackrel{\Delta\text{-inequality, } |e|=1}{\lesssim} \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \right) \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) (\langle |\nabla \overline{\phi}_\mu(\mathbf{a}_L(\cdot), z)|^2 \rangle + 1) \\
&\stackrel{\text{Lemma 20}}{=} \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \right) \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) (\langle |\nabla \overline{\phi}_\mu(\cdot, z)|^2 \rangle_L + 1) \\
&\stackrel{(8)}{=} \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \right) \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) (\langle |D\phi_\mu|^2 \rangle_L + 1) \\
&\stackrel{\langle |D\phi_\mu|^2 \rangle_L \lesssim 1}{\lesssim} \left( \sum'_z \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) \right)^2 \lesssim \left( \int_{|z| \geq \frac{L}{2}} \frac{1}{\sqrt{\mu}} \exp(-c_0\sqrt{\mu}|z|) dz \right)^2 \\
&\lesssim \mu^{-(d+1)} \exp(-\frac{c_0}{2}\sqrt{\mu}L),
\end{aligned}$$

and thus (152) follows.  $\square$

*Proof of Proposition 10.* For convenience we write

$$f_0(\mu) := e \cdot \mathbf{a}_{\text{hom},\mu} e,$$

and set for all  $k \in \mathbb{N}$ ,

$$f_{k+1}(\mu) = \frac{1}{2^{k+2} - 1} \left( 2^{k+2} f_k(\mu) - f_k(2\mu) \right).$$

**Step 1.** Spectral representation of  $f_k(\mu)$ .

We claim that for all  $k \in \mathbb{N}_0$  we have

$$f_k(\mu) - e \cdot \mathbf{a}_{\text{hom}} e = \mu^{k+2} \int_0^\infty \frac{p_k(\nu, \mu)}{\nu(2^0\mu + \nu)^2 \dots (2^k\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle,$$

where  $p_k$  denotes a linear combination of monomials  $\mu^i \nu^j$  with  $i + j = k$  (we follow the convention that  $p_0 \in \mathbb{R}$ ).

We proceed by induction. For  $k = 0$ , the claim is identical to (130). Assume that the claim holds for some  $k \in \mathbb{N}_0$ , that is

$$f_k(\mu) = e \cdot \mathbf{a}_{\text{hom}} e + \mu^{k+2} \int_0^\infty \frac{p_k(\nu, \mu)}{\nu(\mu + \nu)^2 \dots (2^k\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.$$

By definition of  $f_{k+1}(\mu)$ , this yields

$$\begin{aligned}
&f_{k+1}(\mu) - e \cdot \mathbf{a}_{\text{hom}} e \\
&= \frac{1}{2^{k+2} - 1} \int_0^\infty \frac{2^{k+2} \mu^{k+2} p_k(\mu, \nu)}{\nu(\mu + \nu)^2 \dots (2^k\mu + \nu)^2} - \frac{(2\mu)^{k+2} p_k(2\mu, \nu)}{\nu(2\mu + \nu)^2 \dots (2^k(2\mu) + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle \\
&= \frac{\mu^{k+2}}{2^{k+2} - 1} \int_0^\infty \frac{2^{k+2} p_k(\mu, \nu) (2^{k+1}\mu + \nu)^2 - 2^{k+2} p_k(2\mu, \nu) (\mu + \nu)^2}{\nu(\mu + \nu)^2 \dots (2^{k+1}\mu + \nu)^2} \langle \mathfrak{d}P(d\nu)\mathfrak{d} \rangle.
\end{aligned}$$

Both terms in the numerator of the fraction are linear combination of monomials  $\mu^i \nu^j$  of total degree  $i + j = k + 2$ . Denoting by  $c_k$  the coefficient of  $\nu^k$  in  $p_k(\mu, \nu)$ , it follows that the coefficients of  $\nu^{k+2}$  in the two terms of the numerator are  $c_k 2^{k+2}$  and  $-c_k 2^{k+2}$ . Hence the polynomial of the numerator (which has degree  $k + 2$ ) is divisible by  $\mu$ . This proves the claim for  $k + 1$ , and concludes the induction argument.

**Step 2.** Conclusion.

By Step 1 we have

$$f_k(\mu) - e \cdot \mathbf{a}_{\text{hom}} e = \mu^{k+2} \int_0^\infty \frac{p_k(\nu, \mu)}{\nu(\mu + \nu)^2 \dots (2^k \mu + \nu)^2} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle.$$

Since  $p_k(\nu, \mu)$  is a sum of monomials of total degree  $k$ , we have

$$|p_k(\nu, \mu)| \lesssim (\mu + \nu)^k$$

for all  $\nu, \mu \geq 0$ . Hence,

$$|f_k(\mu) - e \cdot \mathbf{a}_{\text{hom}} e| \lesssim \mu^{k+2} \int_0^\infty \frac{1}{\nu(\mu + \nu)^{k+2}} \langle \mathfrak{d}P(d\nu) \mathfrak{d} \rangle.$$

Using (144) with  $\gamma = \frac{d}{2} + 1$  given by Corollary 8, this turns into

$$|f_k(\mu) - e \cdot \mathbf{a}_{\text{hom}} e| \lesssim \mu^{\frac{d}{2}} \int_0^{\frac{1}{\mu}} (1 + u)^{\frac{d}{2} - k - 3} + \mu^{k+2}$$

The conclusion follows from the integrability of  $u \mapsto (1 + u)^{\frac{d}{2} - k - 3}$  on  $(0, \infty)$  for  $k > \frac{d}{2} - 2$ .  $\square$

## 7 Estimates on the parabolic Green's function with uniformly elliptic coefficients

In this section we prove Theorem 3 that collects the required estimate on the variable-coefficients parabolic Green's function. For the sake of clarity we first discuss the  $L$ -periodic, spatially continuum case, where the lattice  $\mathbb{Z}^d$  is replaced by  $\mathbb{R}^d$ . In contrast to the discrete case, in the continuum case the classical Leibniz' and chain rules are available. Moreover, the elliptic and parabolic Calderón-Zygmund estimates that we need for the Meyers' estimate are well-known. In Sections 7.2 and 7.3 we indicate the modifications needed in the discrete  $L$ -periodic and the discrete whole space case. In Section 7.4 we establish the required discrete Calderón-Zygmund estimates.

### 7.1 The periodic continuum case

In this section,  $G_L(t, \mathbf{a}, x, y)$  refers to the  $L$ -periodic Green's function on  $\mathbb{R}^d$ . When no confusion occurs, we suppress the argument  $\mathbf{a}$  in the notation. Note that  $\bar{G}_L := L^{-d}$  is the spatial average of  $G_L(t, \cdot, y)$  for all  $t > 0$  and  $y \in \mathbb{R}^d$ .

Throughout this section we denote by  $\omega_0$  the  $L$ -periodic weight  $\omega_0(x) := \text{dist}(x, LZ^d)$  that we use exclusively use in the continuum case. The upcoming Lemma 24 states the necessary pointwise bounds on the Green's function itself.

**Lemma 24** (continuum case). *For any weight exponent  $\alpha < \infty$  we have*

$$(154) \quad G_L(t, x, y) \lesssim t^{-\frac{d}{2}} \left( \frac{\omega_0^2(x-y)}{t} + 1 \right)^{-\frac{\alpha}{2}} \quad \text{for } t \lesssim L^2,$$

$$(155) \quad |G_L(t, x, y) - \bar{G}_L| \lesssim L^{-d} \exp(-c_0 \frac{t}{L^2}) \quad \text{for } t \gtrsim L^2.$$

Here,  $\lesssim$  means  $\leq$  up to a constant that depends on  $\alpha$  (next to  $\lambda, d$ ), and  $c_0 > 0$  denotes a generic constant that only depends on  $\lambda$  and  $d$ .

**Remark 17.** *In the continuum whole space case, we have the well-known stronger result of*

$$G(t, x, y) \lesssim t^{-\frac{d}{2}} \exp(-c_0 \frac{|x-y|^2}{t}),$$

first established by Nash [30], see Fabes & Stroock [11] for a stream-lined approach. In Lemma 24, we don't establish this Gaussian but only a super-algebraic decay. The main reason for this is that we display a line of arguments that extends to the spatially discrete case: In the discrete case, we can only expect exponential instead of Gaussian tails. Luckily, the proof of Lemma 25, whose result we are primarily interested in, only requires this super-algebraic decay.

Let us give the short argument why in the time-continuous but spatially discrete case, we can only expect exponential tails (incidentally, in the fully discrete case, the Gaussian tails are replaced by a compact support). To this purpose, we consider the case of  $d = 1$  and  $a = 1$ . We will now argue that in this case,

$$(156) \quad G(t, x) := G(t, x, 0) \geq \exp(-x \ln \frac{2x}{t}) \quad \text{for } x \gg t.$$

Indeed, we have by the semigroup property and non-negativity of  $G$  for all  $\epsilon > 0$  and  $n \in \mathbb{N}$ :

$$G(\epsilon(n+1), n+1) = \sum_{x \in \mathbb{Z}} G(\epsilon n, (n+1) - x) G(\epsilon, x) \geq G(\epsilon n, n) G(\epsilon, 1),$$

and thus by induction over  $n$

$$(157) \quad G(\epsilon n, n) \geq G(\epsilon, 1)^n.$$



On the other hand, from  $\frac{d}{dt}|_{t=0}G(t,1) = -2G(0,1) + G(0,0) + G(0,2) = 1$  we learn that  $G(\epsilon,1) \geq \frac{1}{2}\epsilon$  for  $\epsilon \ll 1$ . Inserting this into (157) we obtain

$$G(\epsilon n, n) \geq \left(\frac{1}{2}\epsilon\right)^n = \exp\left(-n \ln \frac{2}{\epsilon}\right) \quad \text{for } \epsilon \ll 1,$$

which translates into (156).

The precise asymptotics of  $G(t,x)$  (in the case  $d = 1, a = 1$ ) have been computed in Pang [31]. In the general case, a lower bound behaving like (156) has been derived by Delmotte [8]. For the upper bound and a partial Aronson lower bound see also [12, Propositions B.3 and B.4].

The reasons why we include a proof of Lemma 24 are the following:

- The proof is elementary, and some of its elements will be referred to in the proof of Lemma 25 (whose result is new, we believe).
- We could not find the discrete, *periodic* case in the literature. Delmotte's more involved proof (it is more involved since it captures the exact tail behavior, see Remark 17) in the discrete case is restricted to the case of the whole space (i. e.  $L = \infty$ ). It seemed more convenient to carry out the extension of the simple proof of Lemma 24 to the discrete case than to carry out the extension of Delmotte's involved proof to the periodic case.

The next lemma contains the optimal estimate on  $\nabla_x G$ .

**Lemma 25.** *There exists an exponent  $p > 1$  (only depending on  $\lambda$  and  $d$  and that we may choose as close to 1 as we please) such that for any weight exponent  $\alpha < \infty$  we have*

$$\left( \int_{[0,L]^d} \left( \left( \frac{\omega_0^2(x-y)}{t} + 1 \right)^{\frac{\alpha}{2}} |\nabla_x G_L(t,x,y)| \right)^{2p} dx \right)^{\frac{1}{2p}} \lesssim t^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp\left(-c_0 \frac{t}{L^2}\right),$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\alpha$  and  $p$  (next to  $\lambda$  and  $d$ ).

*Proof of Lemma 24.* For convenience we shall write  $G$  and  $\bar{G}$  instead of  $G_L$  and  $\bar{G}_L$ , respectively.

**Step 1.** Unweighted  $L^2$ -bounds.

We claim that

$$(158) \quad \int G(t,x,y)^2 dx \lesssim t^{-\frac{d}{2}} \quad \text{for } t \lesssim L^2,$$

$$(159) \quad \int (G(t,x,y) - \bar{G})^2 dx \lesssim L^{-d} \exp\left(-c_0 \frac{t}{L^2}\right) \quad \text{for } t \gtrsim L^2.$$

All spatial integrals, if not stated otherwise, extend over the torus of side length  $L$ . Because all estimates depend on the coefficients  $\mathbf{a}$  only through their ellipticity bound  $\lambda$ , we may appeal to translational invariance in order to restrict ourselves to  $y = 0$ .

We start by claiming that it is sufficient to show

$$(160) \quad \int (G(t,x,0) - \bar{G})^2 dx \lesssim t^{-\frac{d}{2}} \quad \text{for all } t,$$

$$(161) \quad \int (G(t,x,0) - \bar{G})^2 dx \lesssim \exp\left(-c_0 \frac{t-T}{L^2}\right) \int (G(T,x,0) - \bar{G})^2 dx$$

for  $t \geq T$ .

Indeed, (158) follows from (160) since by the triangle inequality and  $\bar{G} = L^{-d}$  we have

$$\int G(t,x,0)^2 dx \leq 2 \int (G(t,x,0) - \bar{G})^2 dx + 2L^{-d}.$$

In particular, (160) implies  $\int (G(T, x, 0) - \bar{G})^2 dx \lesssim L^{-d}$  for  $T = L^2$ . Hence (161) yields (159) — recall that  $c_0$  denotes a *generic* constant.

We now turn to the proof of (160) & (161). Here and in the sequel,  $G$  stands for  $G(t, x, 0)$ . We note that

$$\begin{aligned}
\frac{d}{dt} \int \frac{1}{2} (G - \bar{G})^2 dx &= \int (G - \bar{G}) \partial_t G dx \\
&= - \int \nabla G \cdot a \nabla G dx \\
(162) \qquad \qquad \qquad &\leq -\lambda \int |\nabla G|^2 dx.
\end{aligned}$$

Here and in the sequel  $\nabla G$  stands for  $\nabla_x G(t, x, 0)$ . We combine (162) with the Poincaré inequality for functions with mean value zero in form of

$$\int (G - \bar{G})^2 dx \lesssim L^2 \int |\nabla G(t, x, 0)|^2 dx.$$

This yields the differential inequality  $\frac{d}{dt} \int (G - \bar{G})^2 dx \lesssim -\frac{1}{L^2} \int (G - \bar{G})^2 dx$ . The latter can be rewritten as  $\frac{d}{dt} (\exp(c_0 \frac{t}{L^2}) \int (G - \bar{G})^2 dx) \leq 0$ , yielding (161) after integration.

In order to derive (160), we combine (162) with the interpolation estimate (to be established below)

$$(163) \qquad \left( \int (u - \bar{u})^2 dx \right)^{\frac{1}{2}} \lesssim \left( \left( \int |\nabla u|^2 dx \right)^{\frac{1}{2}} \right)^{\frac{d}{d+2}} \left( \int |u - \bar{u}| dx \right)^{\frac{2}{d+2}},$$

where  $\bar{u}$  denotes the spatial average of the function  $u(x)$ . We apply (163) to  $u(x) = G(t, x, 0)$  and use  $\int |G(t, x, 0) - \bar{G}| dx \leq 2 \int |G(t, x, 0)| dx = 2 \int G(t, x, 0) dx = 2$ . Thus (162) and (163) combine to the differential inequality

$$\frac{d}{dt} \int (G - \bar{G})^2 dx \lesssim - \left( \int (G - \bar{G})^2 dx \right)^{1 + \frac{2}{d}}.$$

The latter can be rewritten as  $\frac{d}{dt} (\int (G - \bar{G})^2 dx)^{-\frac{2}{d}} \gtrsim 1$ , yielding (160) after integration.

We finally turn to the argument for the interpolation estimate (163). In case of  $d > 2$ , it is a consequence of the Poincaré-Sobolev inequality with mean-value zero, and of Hölder's inequality:

$$\begin{aligned}
\left( \int |u - \bar{u}|^{\frac{2d}{d-2}} dx \right)^{\frac{d-2}{2d}} &\lesssim \left( \int |\nabla u|^2 dx \right)^{\frac{1}{2}}, \\
\left( \int |u - \bar{u}|^2 dx \right)^{\frac{1}{2}} &\lesssim \left( \left( \int |u - \bar{u}|^{\frac{2d}{d-2}} dx \right)^{\frac{d-2}{2d}} \right)^{\frac{d}{d+2}} \left( \int |u - \bar{u}| dx \right)^{\frac{2}{d+2}}.
\end{aligned}$$

The case of  $d = 2$  is slightly more subtle, in particular in the periodic case. We first argue that we may replace the spatial average  $\bar{u}$  by the median  $M$  in (163), i. e.

$$(164) \qquad \left( \int (u - M)^2 dx \right)^{\frac{1}{2}} \lesssim \left( \left( \int |\nabla u|^2 dx \right)^{\frac{1}{2}} \right)^{\frac{1}{2}} \left( \int |u - M| dx \right)^{\frac{1}{2}}.$$

Recall that the median  $M$  is characterized by

$$(165) \qquad |\{u > M\}| \leq \frac{1}{2} L^2 \quad \text{and} \quad |\{u \geq M\}| \geq \frac{1}{2} L^2,$$

where  $|A|$  denotes the 2-dimensional volume of a subset  $A$  of the  $L$ -torus. Indeed, on the one hand, since  $\bar{u}$  is the  $L^2$ -orthogonal projection of  $u$  onto constants we have  $\int (u - \bar{u})^2 dx \leq \int (u - M)^2 dx$ . On

the other hand, we have  $\int |u - M| dx \leq 3 \int |u - \bar{u}| dx$ ; this follows via the triangle inequality in  $L^1$  from  $L^2|\bar{u} - M| \leq 2 \int |u - \bar{u}| dx$ ; to see the latter, assume w. l. o. g. that  $M > \bar{u}$ , we note that by definition (165) we have  $\frac{1}{2}L^2 \leq |\{u \geq M\}| \leq \frac{1}{M - \bar{u}} \int |u - \bar{u}| dx$ .

We now give the argument for (164). W. l. o. g. we may assume  $M = 0$  and split (164) into positive and negative parts; this amounts to showing

$$(166) \quad \int_{u>0} u^2 dx \lesssim \left( \int_{u>0} |\nabla u|^2 dx \right)^{\frac{1}{2}} \int_{u>0} u dx$$

under the assumption  $|\{u > 0\}| \leq \frac{1}{2}L^d$ . Because of this assumption, we have by the isoperimetric inequality applied to  $\max\{u, 0\}^2$  and Cauchy-Schwarz' inequality

$$(167) \quad \int_{u>0} u^4 dx \lesssim \left( \int_{u>0} |\nabla u|^2 dx \right)^2 \int_{u>0} u^2 dx.$$

On the other hand, we have by Hölder's inequality

$$(168) \quad \int_{u>0} u^2 dx \leq \left( \int_{u>0} u^4 dx \right)^{\frac{1}{3}} \left( \int_{u>0} u dx \right)^{\frac{2}{3}}.$$

Estimates (167) and (168) combine into (166).

**Step 2.** Weighted  $L^2$ -bounds.

We shall prove that for any weight exponent  $\alpha < \infty$

$$(169) \quad \int \left( \left( \frac{\omega_0^2(x-y)}{t} + 1 \right)^{\frac{\alpha}{2}} G(t, x, y) \right)^2 dx \lesssim t^{-\frac{d}{2}} \quad \text{for } t \lesssim L^2.$$

We start with a couple of reductions. As in Step 1, we restrict ourselves to  $y = 0$ . Since  $\left( \frac{\omega_0^2}{t} + 1 \right)^{\frac{\alpha}{2}}$  is monotone in  $\alpha$ , we may assume that  $\alpha$  is large, say  $\alpha > \frac{d}{2}$ . By combining the pointwise estimate

$$\left( \frac{\omega_0^2}{t} + 1 \right)^{\alpha} G^2 \lesssim t^{-\alpha} \omega_0^{2\alpha} G^2 + G^2,$$

with (158), we deduce that it is enough to show that for all  $\alpha > \frac{d}{2}$

$$(170) \quad \int \omega_0^{2\alpha} G^2 dx \lesssim t^{\alpha - \frac{d}{2}} \quad \text{for } t \lesssim L^2.$$

We now turn to the proof of (170) and note that using the ellipticity of  $\mathbf{a}$ , we have

$$(171) \quad \begin{aligned} & \frac{d}{dt} \int \omega_0^{2\alpha} \frac{1}{2} G^2 dx \\ &= \int \omega_0^{2\alpha} G \partial_t G dx \\ &= - \int \nabla(\omega_0^{2\alpha} G) \cdot \mathbf{a} \nabla G dx \\ &= - \int \omega_0^{2\alpha} \nabla G \cdot \mathbf{a} \nabla G dx - 2\alpha \int \omega_0^{2\alpha-1} G \nabla \omega_0 \cdot \mathbf{a} \nabla G dx \\ &\stackrel{|\nabla \omega_0| \leq 1}{\leq} -\lambda \int \omega_0^{2\alpha} |\nabla G|^2 dx + 2\alpha \int \omega_0^{2\alpha-1} G |\nabla G| dx \\ &\stackrel{\text{Young}}{\leq} \frac{\alpha^2}{\lambda} \int \omega_0^{2\alpha-2} G^2 dx \\ &\stackrel{\text{Hölder}}{\leq} \frac{\alpha^2}{\lambda} \left( \int \omega_0^{2\alpha} G^2 dx \right)^{1-\frac{1}{\alpha}} \left( \int G^2 dx \right)^{\frac{1}{\alpha}}. \end{aligned}$$

Hence we obtain the differential inequality

$$\frac{d}{dt} \left( \int \omega_0^{2\alpha} G^2 dx \right)^{\frac{1}{\alpha}} \lesssim \left( \int G^2 dx \right)^{\frac{1}{\alpha}},$$

which by Step 1 turns into

$$(172) \quad \frac{d}{dt} \left( \int \omega_0^{2\alpha} G dx \right)^{\frac{1}{\alpha}} \lesssim t^{-\frac{d}{2\alpha}} \quad \text{for } t \lesssim L^2.$$

Because  $\frac{d}{2\alpha} < 1$  and since  $\int \omega_0^{2\alpha} G dx$  vanishes for  $t = 0$  (where again we use  $2\alpha > d$ ), the integration of this inequality yields (170).

**Step 3.** Weighted  $L^\infty$ -bounds.

In this step, we establish the statement of Lemma 24, that is

$$(173) \quad G(t, x, y) \lesssim t^{-\frac{d}{2}} \left( \frac{\omega_0^2(x-y)}{t} + 1 \right)^{-\frac{\alpha}{2}} \quad \text{for } t \lesssim L^2,$$

$$(174) \quad |G(t, x, y) - \bar{G}| \lesssim L^{-d} \exp(-c_0 \frac{t}{L^2}) \quad \text{for } t \gtrsim L^2.$$

This upgrade of Step 1 & 2 is based on the semigroup property, i. e.

$$G(t, x, y) = \int G(t/2, x, z) G(t/2, z, y) dz,$$

which using the symmetry of  $G(t, \cdot, \cdot)$  we rewrite as

$$(175) \quad G(t, x, y) = \int G(t/2, z, x) G(t/2, z, y) dz.$$

We start by deriving (174) from (159) and thus assume that  $t \gtrsim L^2$ . For that purpose, we note that because of  $\int G(t/2, z, x) dz = \int G(t/2, z, y) dz = \int \bar{G} dz = 1$ , (175) can be rewritten as

$$G(t, x, y) - \bar{G} = \int (G(t/2, z, x) - \bar{G})(G(t/2, z, y) - \bar{G}) dz.$$

An application of Cauchy Schwarz' inequality yields

$$|G(t, x, y) - \bar{G}| \leq \left( \int (G(t/2, x, z) - \bar{G})^2 dz \int (G(t/2, z, y) - \bar{G})^2 dz \right)^{\frac{1}{2}}.$$

Inserting (159) (with  $t$  replaced by  $t/2 \gtrsim L^2$ ) gives (174) (recall that  $c_0$  denotes a generic constant).

We turn to (173). By the triangle inequality on the  $L$ -torus  $\omega_0(x-y) \leq \omega_0(x-z) + \omega_0(z-y)$ , that we use in the form  $\omega_0^2(x-y) \leq 2\omega_0^2(x-z) + 2\omega_0^2(z-y)$ , we obtain the inequality

$$(176) \quad \omega(t, x-y) \leq \omega(t/2, z-x) \omega(t/2, z-y)$$

for the weight

$$(177) \quad \omega(t, x) := \left( \frac{\omega_0^2(x)}{t} + 1 \right)^{\frac{1}{2}}.$$

Combining (175) and (176) yields the inequality

$$\begin{aligned} & \omega^\alpha(t, x-y) G(t, x, y) \\ & \leq \int \omega^\alpha(t/2, z-x) G(t/2, z, x) \omega^\alpha(t/2, z-y) G(t/2, z, y) dz \\ & \leq \left( \int (\omega^\alpha(t/2, z-x) G(t/2, z, x))^2 dz \int (\omega^\alpha(t/2, z-y) G(t/2, z, y))^2 dz \right)^{\frac{1}{2}}. \end{aligned}$$

Inserting (169) (with  $t$  replaced by  $t/2 \lesssim L^2$ ) yields (173). □

*Proof of Lemma 25.* We recall that  $\omega$  is defined in (177), and we note for further reference that

$$(178) \quad |\nabla\omega(t, x)| \leq t^{-\frac{1}{2}},$$

$$(179) \quad \omega(t, x) \lesssim 1 \quad \text{for } t \gtrsim L^2.$$

For convenience we write  $G$  instead of  $G_L$  and  $\int$  instead of  $\int_{[0, L]^d}$ .

**Step 1.**  $L_{tx}^2$  estimate of  $\nabla G$ .

We shall prove that

$$\frac{1}{T} \int_T^{2T} \int (\omega^\alpha(t, x-y) |\nabla G(t, x, y)|)^2 dx dt \lesssim T^{-\frac{d}{2}-1} \exp(-c_0 \frac{T}{L^2}).$$

As in the proof of Lemma 24, we may assume  $y = 0$ . In view of (179), this statement splits into

$$(180) \quad \frac{1}{T} \int_T^{2T} \int \omega^2 |\nabla G|^2 dx dt \lesssim T^{-\frac{d}{2}-1} \quad \text{for } T \lesssim L^2,$$

$$(181) \quad \frac{1}{T} \int_T^{2T} \int |\nabla G|^2 dx dt \lesssim L^{-d} T^{-1} \exp(-c_0 \frac{T}{L^2}) \quad \text{for } T \gtrsim L^2.$$

(Again, we need to redefine  $c_0$  when replacing  $L^{-d} T^{-1} \exp(-c_0 \frac{T}{L^2})$  by  $T^{-\frac{d}{2}-1} \exp(-c_0 \frac{T}{L^2})$  for  $T \gtrsim L^2$ .)

We first note that integrating (162), cf. the proof of Lemma 24, from  $T$  to  $2T$  yields

$$(182) \quad \int_T^{2T} \int |\nabla G|^2 dx dt \lesssim \int (G(T, x, 0) - \bar{G})^2 dx.$$

Combined with (159), cf. Step 1 of the proof of Lemma 24, it yields (181).

We now turn to (180). By (182) and (158), it is enough to show

$$(183) \quad \int_T^{2T} \int \omega_0^{2\alpha} |\nabla G|^2 dx dt \lesssim T^{\alpha-\frac{d}{2}} \quad \text{for } T \lesssim L^2$$

for  $\alpha = 0$  and  $\alpha$  sufficiently large. For  $\alpha = 0$ , this is again a consequence of (182) and (158) so that we focus on the case  $\alpha \gg 1$ . The starting point is (171), cf. the proof of Lemma 24. This time we use Young's inequality differently:

$$(184) \quad \begin{aligned} & \frac{d}{dt} \int \omega_0^{2\alpha} \frac{1}{2} G^2 dx \\ & \leq -\lambda \int \omega_0^{2\alpha} |\nabla G|^2 dx + 2\alpha \int \omega_0^{2\alpha-1} G |\nabla G| dx \\ & \leq -\frac{\lambda}{2} \int \omega_0^{2\alpha} |\nabla G|^2 dx + \frac{4\alpha^2}{\lambda} \int \omega_0^{2\alpha-2} G^2 dx, \end{aligned}$$

which we integrate between  $T$  and  $2T$ :

$$\begin{aligned} & \int_T^{2T} \int \omega_0^{2\alpha} |\nabla G(t, x, 0)|^2 dx dt \\ & \lesssim \int \omega_0^{2\alpha} G(T, x, 0)^2 dx + \int_T^{2T} \int \omega_0^{2\alpha-2} G(t, x, 0)^2 dx dt. \end{aligned}$$

Combined with (170), cf. Step 2 in the proof of Lemma 24, it yields the desired estimate (183) for  $\alpha > \frac{d}{2} + 1$ .

**Step 2.**  $L_t^\infty L_x^2$  estimate of  $\nabla G$ .

We claim that

$$(185) \quad \int (\omega^\alpha(t, x-y) |\nabla G(t, x, y)|)^2 dx \lesssim t^{-\frac{d}{2}-1} \exp(-c_0 \frac{t}{L^2}).$$

(The reader may wonder why we pass *twice* from a time-averaged to a pointwise-in-times estimate: here in Step 2 and later in Step 6. The reason is that we need Step 2 in Step 5.) As for the passage from Step 2 to Step 3 in the proof of Lemma 24, this upgrade of Step 1 follows from the semigroup property (175), that we use in the form of

$$G(T, x, y) = \frac{3}{T} \int_{T/3}^{2T/3} \int G(t, x, z)G(T-t, z, y)dzdt.$$

We differentiate this identity w. r. t.  $x$ :

$$\nabla G(T, x, y) = \frac{3}{T} \int_{T/3}^{2T/3} \int G(T-t, z, y)\nabla G(t, x, z)dzdt.$$

Using  $\frac{3}{T} \int_{T/3}^{2T/3} \int G(T-t, z, y)dzdt = 1$ , we obtain by Jensen's inequality

$$(186) \quad |\nabla G(T, x, y)|^2 \leq \frac{3}{T} \int_{T/3}^{2T/3} \int G(T-t, z, y)|\nabla G(t, x, z)|^2dzdt.$$

Like in Step 3 of Lemma 24, cf. (176), we appeal to the triangle inequality  $\omega_0(x-y) \leq \omega_0(x-z) + \omega_0(z-y)$ . By Young's inequality we get  $\omega_0^2(x-y) \leq \frac{1}{\theta}\omega_0^2(x-z) + \frac{1}{1-\theta}\omega_0^2(z-y)$  with  $\theta = \frac{t}{T} \in [0, 1]$  and therefore obtain

$$(187) \quad \omega(T, x-y) \leq \omega(t, x-z)\omega(T-t, y-z).$$

Integrating (186) on the torus and using (187) yield

$$(188) \quad \begin{aligned} & \int \omega^{2\alpha}(T, x-y)|\nabla G(T, x, y)|^2dx \\ & \leq \frac{3}{T} \int_{T/3}^{2T/3} \int \omega^{2\alpha}(T-t, y-z)G(T-t, z, y) \\ & \quad \int \omega^{2\alpha}(t, x-z)|\nabla G(t, x, z)|^2dx dzdt. \end{aligned}$$

We then need to estimate the term  $\omega^{2\alpha}(T-t, z-y)G(T-t, z, y) = \omega^{2\alpha}(t', z-y)G(t', z, y)$  in (188). To this aim, we use Lemma 24:

- Case  $t' \lesssim L^2$ : Estimate (154) shows that  $G(t', z, y) \lesssim t'^{-\frac{d}{2}}\omega^{-\beta}(t', z-y)$  for any weight exponent  $\beta < \infty$ . Taking in particular for  $\beta = 2\alpha + r$  for some  $r > d$  yields

$$\omega^{2\alpha}(t', z-y)G(t', z, y) \lesssim t'^{-\frac{d}{2}}\omega^{-r}(t', z-y).$$

- Case  $t' \gtrsim L^2$ : Estimate (179) shows that  $G(t', x, y) \lesssim L^{-d}$ . Combined with  $\omega^{2\alpha}(t', z-y) \lesssim 1$ , this yields

$$\omega^{2\alpha}(t', z-y)G(t', z, y) \lesssim L^{-d}.$$

Since  $t' = T-t \in [T/3, 2T/3]$  we thus obtain

$$(189) \quad \omega^{2\alpha}(T-t, z-y)G(T-t, z, y) \lesssim \left\{ \begin{array}{ll} T^{-\frac{d}{2}}\omega^{-r}(T, z-y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\}.$$

We note for further reference that the r. h. s. of (189) has unit integral:

$$(190) \quad \left\{ \begin{array}{l} \int T^{-\frac{d}{2}}\omega^{-r}(T, z-y)dz \\ \int L^{-d}dz \end{array} \right\} \lesssim 1.$$

This is obvious for the second expression; for the first expression we note that the choice  $r > d$  implies that

$$(191) \quad \int \omega^{-r}(t, x) dx \lesssim t^{\frac{d}{2}}.$$

This follows from

$$(192) \quad \begin{aligned} \int_{(-\frac{L}{2}, \frac{L}{2})^d} (1 + \frac{\omega_0^2(x)}{t})^{-\frac{r}{2}} dx &\leq \int_{\mathbb{R}^d} (1 + \frac{|x|^2}{t})^{-\frac{r}{2}} dx \\ &= t^{\frac{d}{2}} \int_{\mathbb{R}^d} (1 + |\hat{x}|^2)^{-\frac{r}{2}} d\hat{x} \stackrel{r>d}{\lesssim} t^{\frac{d}{2}}. \end{aligned}$$

Inserting (189) into (188) (and exchanging the  $z$  and  $t$  integrations) yields

$$\begin{aligned} &\int \omega^{2\alpha}(T, x - y) |\nabla G(T, x, y)|^2 dx \\ &\lesssim \int \left\{ \begin{array}{ll} T^{-\frac{d}{2}} \omega^{-r}(T, z - y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\} \\ &\quad \times \frac{1}{T} \int_{T/3}^{2T/3} \int \omega^{2\alpha}(t, x - z) |\nabla G(t, x, z)|^2 dx dt dz. \end{aligned}$$

Inserting the statement of Step 1 (with  $T$  replaced by  $T/3$  and  $y$  replaced by  $z$ ) into this estimate yields

$$\begin{aligned} &\int \omega^{2\alpha}(T, x - y) |\nabla G(T, x, y)|^2 dx \\ &\lesssim \int \left\{ \begin{array}{ll} T^{-\frac{d}{2}} \omega^{-r}(T, z - y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\} T^{-\frac{d}{2}-1} \exp(-c_0 \frac{T}{L^2}) dz. \end{aligned}$$

We now appeal to (190) to obtain (185), the statement of Step 2.

**Step 3.** Meyers' estimate.

We claim that there exists a  $p > 1$  (with  $p - 1$  as small as we please), only depending on  $\lambda$  and  $d$ , such that for any  $u(t, x)$ ,  $g(t, x)$  that are  $L$ -periodic in  $x$ , compactly supported in  $t$ , and related by the parabolic equation

$$(193) \quad \partial_t u - \nabla \cdot \mathbf{a} \nabla u = \nabla \cdot g,$$

we have

$$(194) \quad \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \lesssim \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}}.$$

We reproduce this classical estimate (see [25] for the continuum elliptic case), since we need to reproduce it in the discrete periodic situation, where we could not find a reference. The main ingredient is the parabolic Calderón-Zygmund estimate (for periodic boundary data): Let  $u(t, x)$ ,  $g(t, x)$  be  $L$ -periodic in  $x$  and compactly supported in  $t$  related by the *constant-coefficient* equation

$$\partial_t u - \Delta u = \nabla \cdot g.$$

Provided  $1 < 2p < \infty$ , we then have

$$\left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \leq C(2p, d) \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}}.$$

A scaling argument ( $\hat{t} = \frac{1+\lambda}{2}t$ ,  $\hat{u} = \frac{1+\lambda}{2}u$ ) shows that the above can also be formulated as follows: For

$$(195) \quad \partial_t u - \frac{1+\lambda}{2} \Delta u = \nabla g$$

we have

$$(196) \quad \frac{1+\lambda}{2} \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \leq C(2p, d) \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}}.$$

We start by noting that we may assume

$$(197) \quad C(2, d) = 1.$$

Indeed, (195) implies

$$\frac{d}{dt} \int \frac{1}{2} u^2 dx = -\frac{1+\lambda}{2} \int |\nabla u|^2 dx + \int \nabla u \cdot g dx.$$

Because of the compact support in time this yields

$$\frac{1+\lambda}{2} \int_{-\infty}^{\infty} \int |\nabla u|^2 dx dt = \int_{-\infty}^{\infty} \int \nabla u \cdot g dx dt.$$

The Cauchy-Schwarz inequality now shows that  $C(2, d) = 1$  in (196).

On the other hand, complex interpolation shows that we may assume that

$$C(2p, d) \leq C(2, d)^{1-\theta} C(2p_0, d)^\theta \quad \text{for } \frac{1}{2p} = (1-\theta)\frac{1}{2} + \theta\frac{1}{2p_0} \text{ and } \theta \in [0, 1].$$

The combination of this with (197) yields  $\lim_{p \downarrow 1} C(2p, d) = 1$ . In particular, there exists a  $p > 1$  with  $p-1$  arbitrarily small and only depending on  $\lambda > 0$  and  $d$  such that

$$(198) \quad C(2p, d) < \frac{1+\lambda}{1-\lambda}.$$

We now claim that this  $p$  has the desired property, that is, (194) for solutions of (193). For that purpose, we rewrite the variable-coefficient equation (193) as a perturbation of the constant-coefficient equation (195):

$$\partial_t u - \frac{1+\lambda}{2} \Delta u = \nabla \cdot \left( g + \left( a - \frac{1+\lambda}{2} \text{id} \right) \nabla u \right).$$

The application of (196) and the triangle inequality in  $L_{tx}^{2p}$  then yield

$$(199) \quad \begin{aligned} & \frac{1+\lambda}{2} \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \\ & \leq C(2p, d) \left( \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}} \right. \\ & \quad \left. + \left( \int_{-\infty}^{\infty} \int \left| \left( a - \frac{1+\lambda}{2} \text{id} \right) \nabla u \right|^{2p} dx dt \right)^{\frac{1}{2p}} \right). \end{aligned}$$

We note that because of  $\lambda \text{id} \leq a \leq \text{id}$  we have  $-\frac{1-\lambda}{2} \text{id} \leq a - \frac{1+\lambda}{2} \text{id} \leq \frac{1-\lambda}{2} \text{id}$  and thus  $\left| \left( a - \frac{1+\lambda}{2} \text{id} \right) \nabla u \right| \leq \frac{1-\lambda}{2} |\nabla u|$ . Hence (199) turns into

$$\begin{aligned} & \frac{1+\lambda}{2} \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \\ & \leq C(2p, d) \left( \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}} + \frac{1-\lambda}{2} \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \right). \end{aligned}$$



We now see that (198) is exactly the condition needed to absorb the second term of the r. h. s. in the l. h. s..

**Step 4.** Generalized Meyers' estimate.

We claim that for any  $u(t, x)$ ,  $g(t, x)$ , and  $f(t, x)$  that are  $L$ -periodic in  $x$  and compactly supported in  $t$  with

$$(200) \quad \partial_t u - \nabla \cdot \mathbf{a} \nabla u = \nabla \cdot g + f,$$

we have

$$\begin{aligned} & \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \\ & \lesssim \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}} + \left( \int_{-\infty}^{\infty} \left( \int |f|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}}. \end{aligned}$$

Let  $\bar{u}(t)$ ,  $\bar{f}(t)$  denote the spatial averages of  $u$  and  $f$ , respectively, over the period cell. Because of periodicity of  $u$ ,  $g$ , and  $f$ , the integration of (200) yields  $\frac{d\bar{u}}{dt} = \bar{f}$ , so that we may rewrite (200) as

$$(201) \quad \partial_t(u - \bar{u}) - \nabla \cdot \mathbf{a} \nabla(u - \bar{u}) = \nabla \cdot g + (f - \bar{f}).$$

In order to apply Step 3 to  $u - \bar{u}$ , we have to write  $f - \bar{f}$  as a divergence field  $\nabla \cdot h$ . To this purpose, we let  $v(t, x)$  denote the solution of the periodic Poisson problem

$$(202) \quad -\Delta v = f - \bar{f} \quad \text{and} \quad \bar{v} = 0.$$

Because of  $\frac{2pd}{d-2p+d} = 2p$  we have by the Sobolev embedding

$$(203) \quad \left( \int |\nabla v|^{2p} dx \right)^{\frac{1}{2p}} \lesssim \left( \int |\nabla^2 v|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{2pd}}.$$

Moreover, because of  $\frac{2pd}{2p+d} > 1$  (which follows from  $d \geq 2$  and  $p > 1$ ) we may apply the *elliptic Calderón-Zygmund* estimate to the Poisson problem (202) (cf. Lemma 26 below for the discrete version):

$$(204) \quad \left( \int |\nabla^2 v|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{2pd}} \lesssim \left( \int |f|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{2pd}}.$$

We then learn from (202), (203), and (204) that for  $h := -\nabla v$  we have

$$(205) \quad \begin{aligned} & \nabla \cdot h = f - \bar{f} \quad \text{and} \\ & \left( \int_{-\infty}^{\infty} \int |h|^{2p} dx dt \right)^{\frac{1}{2p}} \lesssim \left( \int_{-\infty}^{\infty} \left( \int |f|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}}. \end{aligned}$$

Hence in view of (201) we indeed have  $\partial_t(u - \bar{u}) - \nabla \cdot \mathbf{a} \nabla(u - \bar{u}) = \nabla \cdot (g + h)$  so that by Step 3 and the triangle inequality in  $L_{tx}^{2p}$  we obtain

$$\begin{aligned} & \left( \int_{-\infty}^{\infty} \int |\nabla u|^{2p} dx dt \right)^{\frac{1}{2p}} \\ & \lesssim \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}} + \left( \int_{-\infty}^{\infty} \int |h|^{2p} dx dt \right)^{\frac{1}{2p}}. \end{aligned}$$

Inserting (205) into this estimate yields the claim of Step 4.

**Step 5.**  $L_{tx}^{2p}$ -estimate on  $\nabla G$ .

We shall prove in this step that

$$\left( \frac{1}{T} \int_T^{2T} \int (\omega^\alpha(t, x-y) |\nabla G(t, x, y)|)^{2p} dx dt \right)^{\frac{1}{2p}} \lesssim T^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2}).$$

In order to treat the cases of  $T \leq L^2$  and  $T \geq L^2$  at once, we introduce the notation

$$G'(t, x, y) := \begin{cases} G(t, x, y) & \text{for } T \leq L^2 \text{ and } t < 4L^2 \\ G(t, x, y) - \bar{G} & \text{for } T \geq L^2 \text{ and } t > \frac{1}{2}L^2 \end{cases}.$$

W. l. o. g. we assume  $y = 0$ . Select a temporal cut-off function  $\eta(t)$  for the interval  $[T, 2T]$  that is supported in  $[\frac{T}{2}, 4T]$ , say the piecewise affine function

$$(206) \quad \eta(t) = \begin{cases} 0 & \text{for } t \leq \frac{T}{2} \\ \frac{2t}{T} - 1 & \text{for } \frac{T}{2} \leq t \leq T \\ 1 & \text{for } T \leq t \leq 2T \\ 2 - \frac{t}{2T} & \text{for } 2T \leq t \leq 4T \\ 0 & \text{for } 4T \leq t \end{cases}.$$

We apply Step 4 to  $u(t, x) = \eta(t)\omega^\alpha(T, x)G'(t, x, 0)$ . We note that by the Leibniz rule and the defining equation for  $G$ ,

$$(207) \quad \begin{aligned} \partial_t u - \nabla \cdot \mathbf{a} \nabla u &= \eta \omega^\alpha (\partial_t G - \nabla \cdot \mathbf{a} \nabla G) + \frac{d\eta}{dt} \omega^\alpha G' - \eta \nabla \omega^\alpha \cdot \mathbf{a} \nabla G - \eta \nabla \cdot (G' \mathbf{a} \nabla \omega^\alpha) \\ &= \frac{d\eta}{dt} \omega^\alpha G' - \eta \nabla \omega^\alpha \cdot \mathbf{a} \nabla G - \eta \nabla \cdot (G' \mathbf{a} \nabla \omega^\alpha). \end{aligned}$$

This equation takes the form of (200) with  $g = \mathbf{a}g_0$ ,  $f = f_1 + f_2$ , where

$$g_0 := -(\nabla \omega^\alpha) \eta G', \quad f_1 := \frac{d\eta}{dt} \omega^\alpha G', \quad f_2 := -\eta \nabla \omega^\alpha \cdot \mathbf{a} \nabla G.$$

We also note that

$$\nabla u = \eta \omega^\alpha \nabla G - g_0.$$

Hence in order to conclude the proof of this step using the Meyers' estimate of Step 4, we need to control the following three terms

$$\begin{aligned} \left( \int_{-\infty}^{\infty} \int |g|^{2p} dx dt \right)^{\frac{1}{2p}} &\leq \left( \int_{-\infty}^{\infty} \int |g_0|^{2p} dx dt \right)^{\frac{1}{2p}} \\ &\leq \left( \int_{\frac{T}{2}}^{4T} \int |G' \nabla \omega^\alpha|^{2p} dx dt \right)^{\frac{1}{2p}}, \\ \left( \int_{-\infty}^{\infty} \left( \int |f_1|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}} &\leq \frac{2}{T} \left( \int_{\frac{T}{2}}^{4T} \left( \int |\omega^\alpha G'|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}}, \\ \left( \int_{-\infty}^{\infty} \left( \int |f_2|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}} &\leq \left( \int_{\frac{T}{2}}^{4T} \left( \int (|\nabla \omega^\alpha| |\nabla G|)^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}} \end{aligned}$$

by the quantity  $T^{\frac{1}{2p}} T^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2})$ . In view of (178), these three estimates reduce to

$$\begin{aligned} \left( \int_{\frac{T}{2}}^{4T} \int |\omega^{\alpha-1} G'|^{2p} dx dt \right)^{\frac{1}{2p}} &\lesssim T^{\frac{1}{2p} - \frac{d}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2}), \\ \left( \int_{\frac{T}{2}}^{4T} \left( \int |\omega^\alpha G'|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}} &\lesssim T^{\frac{1}{2p} - \frac{d}{2} + \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2}), \\ \left( \int_{\frac{T}{2}}^{4T} \left( \int (\omega^{\alpha-1} |\nabla G|)^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} dt \right)^{\frac{1}{2p}} &\lesssim T^{\frac{1}{2p} - \frac{d}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2}). \end{aligned}$$

We shall in fact establish the stronger pointwise-in-time estimates

$$(208) \quad \left( \int |\omega^{\alpha-1} G'|^{2p} dx \right)^{\frac{1}{2p}} \lesssim t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{t}{L^2}),$$

$$(209) \quad \left( \int |\omega^\alpha G'|^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} \lesssim t^{-\frac{d}{2} + \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{t}{L^2}),$$

$$(210) \quad \left( \int (\omega^{\alpha-1} |\nabla G|)^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{d}} \lesssim t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{t}{L^2}).$$

We note that (208) and (209) can be reduced to the simpler statement: For every exponent  $1 \leq q < \infty$  and every  $0 < \alpha < \infty$  we have

$$(211) \quad \left( \int |\omega^\alpha G'|^q dx \right)^{\frac{1}{q}} \lesssim t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{q}} \exp(-c_0 \frac{t}{L^2}).$$

For this purpose, we distinguish the cases  $T \leq L^2$  and  $T \geq L^2$ . In the latter (for which  $t \geq \frac{1}{2} L^2$ ), Lemma 24 shows that  $|G(t, x, 0) - \bar{G}| \lesssim L^{-d} \exp(-c_0 \frac{t}{L^2})$ , cf. (155). Together with (179), this implies

$$\left( \int |\omega^\alpha (G - \bar{G})|^q dx \right)^{\frac{1}{q}} \lesssim L^{-d + \frac{d}{q}} \exp(-c_0 \frac{t}{L^2}) \stackrel{t \geq L^2}{\lesssim} t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{q}} \exp(-c_0 \frac{t}{L^2})$$

— recall that  $c_0$  denotes a generic constant whose value may change from line to line. This settles (211) in the case of  $T \geq L^2$ .

In the case of  $T \leq L^2$  (for which  $t \leq 4L^2$ ), Lemma 24 (with  $\alpha$  replaced by  $\alpha + r/q$  for some  $r > d$ ) yields  $\omega^\alpha(t, x) G(t, x, 0) \lesssim t^{-\frac{d}{2}} \omega(t, x)^{-r/q}$ , cf. (154). Hence we obtain using in addition (191) that

$$\left( \int (\omega^\alpha G)^q dx \right)^{\frac{1}{q}} \lesssim t^{-\frac{d}{2}} \left( \int \omega^{-r} dx \right)^{\frac{1}{q}} \stackrel{r > d, (191)}{\lesssim} t^{-\frac{d}{2}} t^{\frac{d}{2} \frac{1}{q}} \stackrel{t \leq L^2}{\lesssim} t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{q}} \exp(-c_0 \frac{t}{L^2}).$$

This establishes (211) in the case of  $t \leq L^2$ .

We now turn to (210). We note that for  $p < \frac{d}{d-2}$  in case of  $d > 2$  (and all  $p < \infty$  in case of  $d = 2$ ) we

have  $\frac{pd}{2p+d} < 1$ , and we assume that  $p$  was chosen close enough to 1. Hence we have

$$\begin{aligned}
& \left( \int (\omega^{\alpha-1} |\nabla G|)^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{2pd}} \\
&= \left( \int \omega^{-\frac{2pd}{2p+d}} (\omega^\alpha |\nabla G|)^{\frac{2pd}{2p+d}} dx \right)^{\frac{2p+d}{2pd}} \\
&\stackrel{\frac{2pd}{2p+d} < 2, \text{H\"older}}{\leq} \left( \int \omega^{-\frac{2pd}{2p+d-pd}} dx \right)^{\frac{2p+d-pd}{2pd}} \left( \int (\omega^\alpha |\nabla G|)^2 dx \right)^{\frac{1}{2}} \\
&\stackrel{\frac{2pd}{2p+d-pd} > d, (191)}{\lesssim} t^{\frac{d}{2} \frac{2p+d-pd}{2pd}} \left( \int (\omega^\alpha |\nabla G|)^2 dx \right)^{\frac{1}{2}} \\
&\stackrel{(185)}{\lesssim} t^{\frac{d}{2} \frac{2p+d-pd}{2pd}} \left( t^{-\frac{d}{2}-1} \exp(-c_0 \frac{t}{L^2}) \right)^{\frac{1}{2}} \\
&\lesssim t^{-\frac{d}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{t}{L^2}).
\end{aligned}$$

This establishes (210).

**Step 6.**  $L_t^\infty L_x^{2p}$ -estimate for  $\nabla G$ .

We finally prove the desired estimate

$$\left( \int (\omega^\alpha(T, x-y) |\nabla G(T, x, y)|)^{2p} dx \right)^{\frac{1}{2p}} \lesssim T^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{T}{L^2}).$$

We essentially repeat the argument of Step 2. The only differences are:

- In (186), we use Jensen's inequality applied to  $\mathbb{R}^d \ni g \mapsto |g|^{2p}$ :

$$|\nabla G(T, x, y)|^{2p} \leq \frac{3}{T} \int_{T/3}^{2T/3} \int G(T-t, z, y) |\nabla G(t, x, z)|^{2p} dz dt.$$

- In (188), we multiply with  $\omega^{2p\alpha}$ :

$$\begin{aligned}
& \int \omega^{2p\alpha}(T, x-y) |\nabla G(T, x, y)|^{2p} dx \\
&\leq \frac{3}{T} \int_{T/3}^{2T/3} \int \omega^{2p\alpha}(T-t, y-z) G(T-t, z, y) \\
&\quad \int \omega^{2p\alpha}(t, x-z) |\nabla G(t, x, z)|^{2p} dx dz dt.
\end{aligned} \tag{212}$$

- In (189), we replace  $2\alpha$  by  $2p\alpha$ :

$$(213) \quad \omega^{2p\alpha}(T-t, z-y) G(T-t, z, y) \lesssim \left\{ \begin{array}{ll} T^{-\frac{d}{2}} \omega^{-r}(T, z-y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\}.$$

Inserting (213) into (212) yields

$$\begin{aligned}
& \int \omega^{2p\alpha}(T, x-y) |\nabla G(T, x, y)|^{2p} dx \\
&\lesssim \int \left\{ \begin{array}{ll} T^{-\frac{d}{2}} \omega^{-r}(T, z-y) & \text{for } T \leq L^2 \\ L^{-d} & \text{for } T \geq L^2 \end{array} \right\} \\
&\quad \times \frac{1}{T} \int_{T/3}^{2T/3} \int \omega^{2p\alpha}(t, x-z) |\nabla G(t, x, z)|^{2p} dx dt dz.
\end{aligned}$$

The conclusion then follows from Step 5 and estimate (190).

□

## 7.2 The periodic discrete case – Proof of Theorem 3 (b)

The proof of Theorem 3 is similar to the proof of Lemma 25 up to some changes due to discreteness. In the following we indicate the required modifications and start with some general remarks.

- The most intricate modification are the Calderón-Zygmund estimates that we use both for the elliptic and parabolic constant-coefficient equations. The passage from continuum to discrete here is carried out in Section 7.4.
- Hölder's inequality, Poincaré, Sobolev, and the isoperimetric estimates hold verbatim.
- In the discrete case the weight  $\omega_0$  is replaced by

$$\omega_0(x) := \text{dist}(x, LZ^d) + 1.$$

The purpose of this modification is to have

$$(214) \quad \omega_0(x \pm e_i) = \text{dist}(x \pm e_i, LZ^d) + 1 \lesssim \text{dist}(x, LZ^d) + 1 = \omega_0(x),$$

while we still have

$$(215) \quad |\nabla \omega_0(x)| \lesssim 1,$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $d$ .

- The classical Leibniz rule turns into the identity

$$(216) \quad \nabla(uv) = (\nabla u)v + [u]\nabla v,$$

where  $[u] := \sum_{i=1}^d u(\cdot + e_i)(e_i \otimes e_i)$ . Furthermore, the chain rule is replaced by the inequality (for  $\beta \geq 1$ )

$$(217) \quad \forall a, b \geq 0 : |a^\beta - b^\beta| \lesssim (a^{\beta-1} + b^{\beta-1})|a - b|,$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\beta$ .

- Regarding discrete derivatives of the weight  $\omega_0$  we have (for  $2\alpha \geq 1$ )

$$(218) \quad \begin{aligned} |\nabla_i \omega_0^{2\alpha}(x)| &\stackrel{(217)}{\lesssim} (\omega_0^{2\alpha-1}(x + e_i) + \omega_0^{2\alpha-1}(x)) |\nabla_i \omega_0(x)| \\ &\stackrel{(215)}{\lesssim} (\omega_0^{2\alpha-1}(x + e_i) + \omega_0^{2\alpha-1}(x)) \\ &\stackrel{(214)}{\lesssim} \omega_0^{2\alpha-1}(x). \end{aligned}$$

where  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\alpha$  and  $d$ .

The discrete version of Lemma 24 reads as follows: For any weight exponent  $\alpha < \infty$  we have

$$(219) \quad \begin{aligned} G_L(t, x, y) &\lesssim (t+1)^{-\frac{d}{2}} \left( \frac{\omega_0^2(x-y)}{t+1} + 1 \right)^{-\frac{\alpha}{2}} \quad \text{for } t \lesssim L^2, \\ |G_L(t, x, y) - \bar{G}_L| &\lesssim L^{-d} \exp(-c_0 \frac{t}{L^2}) \quad \text{for } t \gtrsim L^2. \end{aligned}$$

Notice that in contrast to the continuum case, the discrete parabolic Green's function has no singularity for  $t = 0$ . In particular, we have the elementary estimate for  $p \geq 1$

$$(220) \quad \left( \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |G_L(t, x, y)|^p \right)^{\frac{1}{p}} \leq \sum_{x \in ([0, L) \cap \mathbb{Z})^d} G_L(t, x, y) = \sum_{x \in ([0, L) \cap \mathbb{Z})^d} G_L(t, x, y) = 1$$

which yields the improvement in (219) with  $t \rightsquigarrow (t+1)$ . The proof of (219) is similar to Lemma 24. The following modifications are needed:

- In all the statements,  $t$  should be replaced by  $t + 1$ . This can be done for (158) and (160) using (220). These improved estimates are then propagated to (169), where  $t$  may be replaced by  $t + 1$ . We note that the initial data associated with the l. h. s. in (172), i. e. the integral  $\int \omega_0^{2\alpha} G_L dx$ , does not vanish for  $t = 0$ , but evaluates to 1. As a consequence, in (170),  $t$  has to be replaced by  $t + 1$ ; and thus, in (173) and (177) as well.
- The continuum Caccioppoli estimate (171) is replaced by the following calculation. Below  $\sum$  stands for summation in  $x \in ([0, L] \cap \mathbb{Z})^d$ , and  $\lesssim$  means  $\leq$  up to a constant that only depends on  $\alpha$  and  $d$ . We start with the identity

$$\begin{aligned}
& \frac{d}{dt} \sum \omega_0^{2\alpha} \frac{1}{2} G_L^2 \\
&= - \sum \nabla(\omega_0^{2\alpha} G_L) \cdot \mathbf{a} \nabla G_L \\
&\stackrel{(216)}{=} - \sum \omega_0^{2\alpha} \nabla G_L \cdot \mathbf{a} \nabla G_L - \sum_{i,j=1}^d G_L(\cdot + e_i) \nabla_i \omega_0^{2\alpha} \mathbf{a}_{ij} \nabla_j G_L,
\end{aligned}$$

and the estimate

$$\begin{aligned}
& \sum_{i,j=1}^d \sum G_L(x + e_i) |\nabla_i \omega_0^{2\alpha}(x)| \mathbf{a}_{ij}(x) |\nabla_j G_L(x)| \\
&\stackrel{(218)}{\lesssim} \sum_{i,j=1}^d \sum G_L(x + e_i) (\omega_0(x + e_i) + \omega_0(x))^{2\alpha-1} \mathbf{a}_{ij}(x) |\nabla_j G_L(x)| \\
&\stackrel{(214)}{\leq} \sum_{i,j=1}^d \sum G_L(x + e_i) \omega_0^{\alpha-1}(x + e_i) \omega_0^\alpha(x) |\nabla_i G_L(x)| \\
&\leq \left( \sum \omega_0^{2\alpha-2}(x) G_L^2(x) \right)^{\frac{1}{2}} \left( \sum \omega_0^{2\alpha}(x) |\nabla G_L(x)|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

By appealing to Young's inequality, the combination of both yields the discrete version of (171):

$$\begin{aligned}
& \frac{d}{dt} \sum \omega_0^{2\alpha}(x) \frac{1}{2} G_L^2(x) \lesssim \sum \omega_0^{2\alpha-2}(x) G_L^2(x) \\
&\leq \frac{1}{\lambda} \left( \sum \omega_0^{2\alpha}(x) G_L^2(x) \right)^{1-\frac{1}{\alpha}} \left( \sum G_L^2(x) \right)^{\frac{1}{\alpha}}.
\end{aligned}$$

- Inequality (176) (and more generally (187)) holds in the slightly weaker form of

$$(221) \quad \omega(t, x - y) \leq \sqrt{2} \omega(\theta t, z - x) \omega((1 - \theta)t, z - y) \quad \text{for all } \theta \in [0, 1].$$

Indeed, from the triangle-inequality on the  $L$ -torus, we obtain

$$\omega_0^2(x - y) \leq 2(\omega_0^2(z - x) + \omega_0^2(z - y)),$$

and thus

$$\frac{\omega_0^2(x - y)}{t + 1} \leq 2 \left( \frac{\omega_0^2(z - x)}{\theta t + 1} + \frac{\omega_0^2(z - y)}{(1 - \theta)t + 1} \right),$$

which yields (221).

The discrete version of Lemma 25 reads as follows: For any weight exponent  $0 < \alpha < \infty$  we have

$$(222) \quad \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \left( \left( \frac{\omega_0(x - y)^2}{t + 1} + 1 \right)^{\frac{\alpha}{2}} |\nabla_x G_L(t, x, y)| \right)^{2p} \right)^{\frac{1}{2p}} \lesssim (t + 1)^{-\frac{d}{2} - \frac{1}{2} + \frac{d}{2} \frac{1}{2p}} \exp(-c_0 \frac{t}{L^2}).$$

Note that  $\omega(t, x) \sim \frac{\omega_0(x-y)^2}{t+1} + 1$ , so that (222) is equivalent to Theorem 3 (b). The adaptation to the discrete setting is made in two steps: The statement for  $t \gg 1$  is proved along the lines of the proof of Lemma 25, whereas the statement for  $t \lesssim 1$  follows from soft arguments. We first indicate the adaptations of the proof of Lemma 25 to the discrete setting for  $t \gg 1$ , and then turn to the soft argument.

- Inequality (184) can be adapted to the discrete case similarly to (171).
- Step 2 holds verbatim; in particular, (192) remains valid, since by periodicity we have (for  $r > d$ )

$$\sum_{x \in ([0, L] \cap \mathbb{Z})^d} \left( \frac{\omega_0^2(x)}{t} + 1 \right)^{-\frac{r}{2}} = \sum_{x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d} \left( \frac{(|x| + 1)^2}{t} + 1 \right)^{-\frac{r}{2}} \leq \sum_{x \in \mathbb{Z}^d} \left( \frac{(|x| + 1)^2}{t} + 1 \right)^{-\frac{r}{2}} \lesssim t^{-\frac{d}{2}}.$$

- In Step 3, the parabolic equation (193) in continuum space is replaced by its discrete version

$$\partial_t u + \nabla^* \mathbf{a} \nabla u = \nabla^* g.$$

With the discrete Calderón-Zygmund estimate at hand (see Lemma 27 below) estimate (194) follows as in the continuum case.

- In Step 4, equations (200), (201) and (202) have to be replaced by their discrete versions (with  $-\nabla \cdot \rightsquigarrow \nabla^*$ ).
- In Step 5 we have to replace (207) by

$$\begin{aligned} \partial_t u + \nabla^* \mathbf{a} \nabla u \\ = \eta \omega^\alpha (\partial_t G_L + \nabla^* \mathbf{a} \nabla G_L) + \nabla^* \mathbf{a} g_0 + f_1 + f_2, \end{aligned}$$

where (using the notation introduced in (216))

$$\begin{aligned} g_0 &:= \eta [G'_L] \nabla \omega^\alpha, & f_1 &:= \frac{d\eta}{dt} \omega^\alpha G'_L, \\ f_2 &:= \eta \sum_{i,j=1}^d \nabla_i^* \omega^\alpha \mathbf{a}_{ij}(\cdot - e_j) \nabla_j G_L(\cdot - e_j). \end{aligned}$$

Appealing to (214) and (215), the expressions  $g_0, f_1$  and  $f_2$  can be estimated as in the continuum case.

We conclude with the soft argument in the regime  $t \lesssim 1$ . Since  $|\nabla_i G_L(x)| \leq |G_L(x + e_i)| + |G_L(x)|$ , (222) simply follows from the combination of (219) and (220).  $\square$

### 7.3 The discrete, whole space case – Proof of Theorem 3 (a)

*Proof of Theorem 3 (a).* The estimate follows from the periodic case, i. e. Theorem 3 (b), by an approximation argument. To make this precise, fix  $\mathbf{a} \in \Omega$ , and exponents  $1 \leq p < \infty$ ,  $\alpha > 0$ . For  $L \in \mathbb{N}$ , let  $\mathbf{a}_L \in \Omega_L$  denote the unique  $L$ -periodic coefficient field that satisfies

$$\mathbf{a}(x) = \mathbf{a}_L(x) \quad \text{for all } x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d.$$

We introduce the short hand notation

$$G(t, x) := G(t, \mathbf{a}, x, y) \quad \text{and} \quad G_L(t, x) := G_L(t, \mathbf{a}_L, x, y).$$

Moreover, let  $\zeta_L$  denote the cut-off function

$$\zeta_L(x) := \begin{cases} 1 & \text{for } x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d, \\ 0 & \text{else.} \end{cases}$$

By Theorem 3 (b) we have

$$(223) \quad (t+1)^{-\frac{d}{2}-\frac{1}{2}+\frac{1}{2p}\frac{d}{2}} \stackrel{\text{Theorem 3 (b)}}{\gtrsim} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\omega_L^\alpha(t, x) |\nabla G_L(t, x)|)^{2p} \right)^{\frac{1}{2p}} \\ \stackrel{\text{periodicity}}{=} \left( \sum_{x \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d} (\omega^\alpha(t, x) |\nabla G_L(t, x)|)^{2p} \right)^{\frac{1}{2p}} \\ = \left( \sum_{x \in \mathbb{Z}^d} (\zeta_L(x) \omega^\alpha(t, x) |\nabla G_L(t, x)|)^{2p} \right)^{\frac{1}{2p}}.$$

We claim that it suffices to prove that

$$(224) \quad \lim_{L \uparrow \infty} G_L(t, x) = G(t, x) \quad \text{for all } x \in \mathbb{Z}^d \text{ and } t > 0.$$

Indeed, (224) clearly implies that

$$\lim_{L \uparrow \infty} \zeta_L^p(x) \omega^{\alpha p}(t, x) |\nabla G_L(t, x)|^p = \omega^{\alpha p}(t, x) |\nabla G(t, x)|^p \quad \text{for all } x \in \mathbb{Z}^d \text{ and } t > 0.$$

Combined with (223) we obtain the desired estimate by the lemma of Fatou. It remains to prove (224).

We first claim that for all  $L \gg 1$  we have

$$(225) \quad \forall x \in \mathbb{Z}^d : \quad |G_L(t, x)| + |\partial_t G_L(t, x)| + |\partial_t^2 G_L(t, x)| + |\partial_t^3 G_L(t, x)| \lesssim 1$$

where  $\lesssim$  means  $\leq$  up to a constant only depending on  $d$ . Indeed, by periodicity and conservation of mass we have

$$(226) \quad |G_L(t, x)| \leq \sum_{x' \in ([-\frac{L}{2}, \frac{L}{2}] \cap \mathbb{Z})^d} |G_L(t, x')| = 1.$$

By combining  $\partial_t G_L(t, x) = -\nabla^* \mathbf{a}_L(x) \nabla G_L(t, x)$  with the discrete estimate  $|\nabla^* \mathbf{a}_L(x) \nabla G_L(t, x)| \lesssim \max_{x' \in \{x, x \pm e_i\}} |G_L(t, x')|$  and (226) we get  $|\partial_t G_L(t, x)| \lesssim 1$ . Since  $\partial_t^k G_L(t, x) = -\nabla^* \mathbf{a}_L(x) \nabla \partial_t^{k-1} G_L(t, x)$  we can bootstrap the argument and finally get (225).

Now fix some time horizon  $T > 0$ . By virtue of (225), for each  $x \in \mathbb{Z}^d$  the coordinate functions  $[0, T] \ni t \mapsto G_L(t, x)$  are uniformly bounded in  $C^3([0, T])$ . Hence, by appealing to Arzelà-Ascoli and a diagonal sequence argument (where we use that  $\mathbb{Z}^d$  is countable), for each  $x \in \mathbb{Z}^d$  there exists a function  $\tilde{G}(\cdot, x) \in C^2([0, T])$  such that (as  $L \uparrow \infty$  and up to a subsequence)

$$(227) \quad \left\{ \begin{array}{l} G_L(t, x) \rightarrow \tilde{G}(t, x) \\ \partial_t G_L(t, x) \rightarrow \partial_t \tilde{G}(t, x) \end{array} \right\} \quad \text{for all } t \in [0, T].$$

Combined with  $\lim_{L \uparrow \infty} \mathbf{a}_L(x) = \mathbf{a}(x)$  we get

$$\partial_t \tilde{G}(t, x) + \nabla^* \mathbf{a}(x) \nabla \tilde{G}(t, x) = \lim_{L \uparrow \infty} \left( \partial_t G_L(t, x) + \nabla^* \mathbf{a}_L(x) \nabla G_L(t, x) \right) = 0, \\ \tilde{G}(t, x) = \lim_{L \uparrow \infty} G_L(t, x) = \delta(x - y).$$



In addition we deduce from (225) that  $|\partial_t^2 \tilde{G}(t, x)| \lesssim 1$ , and thus  $\tilde{G}$  is a  $C^1([0, T], \ell^\infty(\mathbb{Z}^d))$ -solution to the parabolic equation (12) with r. h. s.  $g = 0$  and initial condition  $g_0 = \delta(\cdot - y)$ . Now, Duhamel's formula, see Lemma 2, yields  $\tilde{G}(t, x) = G(t, \mathbf{a}, x, y)$  for all  $(t, x) \in [0, T] \times \mathbb{Z}^d$ . In particular, the limit is unique and thus the convergence holds for the entire sequence. Because the time horizon  $T$  was arbitrary (224) follows.  $\square$

## 7.4 Discrete Calderón-Zygmund estimate

In this subsection, we establish the discrete, periodic Calderon-Zygmund estimates, both in the elliptic and parabolic case (note that also the elliptic version is needed in Step 4 of the proof of the discrete version of Lemma 25).

**Lemma 26** (discrete,  $L$ -periodic, elliptic Calderón-Zygmund). *Let  $u : \mathbb{Z}^d \rightarrow \mathbb{R}$ ,  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$  and  $g : \mathbb{Z}^d \rightarrow \mathbb{R}^d$  be  $L$ -periodic, and assume that  $\sum_{x \in ([0, L] \cap \mathbb{Z})^d} f(x) = 0$ .*

(a) *Suppose that*

$$\nabla^* \nabla u = f \quad \text{on } \mathbb{Z}^d.$$

*Then for all  $1 < p < \infty$*

$$\sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\nabla^2 u(x)|^p \lesssim \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |f(x)|^p.$$

(b) *Suppose that*

$$(228) \quad \nabla^* \nabla u = \nabla^* g \quad \text{on } \mathbb{Z}^d.$$

*Then for all  $1 < p < \infty$*

$$(229) \quad \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\nabla u(x)|^p \lesssim \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |g(x)|^p.$$

*Above  $\lesssim$  means  $\leq$  up to a constant that only depends on  $p$  and  $d$ .*

**Lemma 27** (discrete,  $L$ -periodic, parabolic Calderón-Zygmund). *Let  $u : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}^d$  be compactly supported in  $t$  and  $L$ -periodic in  $x$ . Suppose that  $u(t, x)$  is smooth in  $t$  and related to  $g$  via the equation*

$$(230) \quad \partial_t u + \nabla^* \nabla u = \nabla^* g.$$

*Then we have for all  $1 < p < \infty$*

$$\left( \int_{\mathbb{R}} \sum_{x \in \mathbb{Z}^d} |\nabla u(t, x)|^p dt \right)^{\frac{1}{p}} \lesssim \left( \int_{\mathbb{R}} \sum_{x \in \mathbb{Z}^d} |g(t, x)|^p dt \right)^{\frac{1}{p}},$$

*where  $\lesssim$  means up to a constant that only depends on  $p$  and  $d$ .*

The proofs of the previous two statements are given in Sections 7.4.1 and 7.4.2.

Of course, it is not a surprise that these estimates, which are classical in the continuum case — hold true. We give a complete proof here only for the sake of staying self-contained. For the elliptic case, there are (at least) two strategies to establish the discrete result: 1) Mimic the continuum proof based on multiplier estimates; 2) reduce the discrete case to the continuum case. The strategy 1) was applied in [5, 13] and also sketched in [17] in the context of stochastic homogenization. Here, we will give the full argument based on strategy 2).

The reason for displaying this argument is that it can be easily extended — as we will show — to the semi-discrete parabolic case, which has been less treated in the literature. For the discrete or semi-discrete parabolic case, there are at least three strategies: 1) Mimic the continuum proof based on multiplier estimates, as carried out in the fully discrete case in [5]; 2) reduce to the continuum case (as done here); 3) reduce to the discrete elliptic case via semigroup arguments based on  $\mathcal{R}$ -boundedness, as carried out in [14] (see also [10] for a standard reference to  $\mathcal{R}$ -boundedness).

So both in the elliptic and parabolic case, our strategy is to deduce the discrete CZ estimates from the continuum ones. In fact, it is more convenient to pass from the discrete to the continuum CZ estimate on the level of the whole space estimates (instead of the periodic estimates), cf. Lemma 28 and Lemma 30. It is then a fairly soft argument to pass from whole space to periodic setting, especially on the discrete level, cf. Lemma 29 and Lemma 27. In passing from the discrete to the continuum case in the whole space setting, we use Fourier methods. In order to be able to argue pointwise in the dual space, it is necessary to work with functions that are integrable in discrete space (and time). However, a typical solution of a whole-space Poisson problem is not integrable, even for a compactly supported right hand side. It is thus convenient to introduce a massive term, that ensures this integrability (also in time in the parabolic case). Provided that the estimates do not depend on the constant in front of the massive term — which is the case — it requires only a soft argument to pass from the massive to the mass-less case, cf. Lemma 29 and Lemma 27. Hence it is on the level of the massive, whole-space case where we pass from the continuum to the discrete version.

For the readers convenience we recall the classical continuum Calderón-Zygmund estimates, a modern proof of which can be found for instance in the textbook *N. V. Krylov. Lectures on Elliptic and Parabolic Equations in Sobolev Spaces*.

**Theorem 5.** 1. Let  $\mu > 0$ ,  $1 < p < \infty$ ,  $f \in L^p(\mathbb{R}^d)$  and  $u \in W^{2,p}(\mathbb{R}^d)$  be the unique solution to

$$\mu u - \Delta u = f \quad \text{in } \mathbb{R}^d.$$

Then we have

$$\int_{\mathbb{R}^d} |\nabla^2 u(x)|^p dx \lesssim \int_{\mathbb{R}^d} |f(x)|^p dx,$$

where  $\lesssim$  stands for  $\leq$  up to a constant that only depends on  $d$  and  $p$ .

2. Let  $\mu > 0$ ,  $1 < p < \infty$  and  $f \in L_t^p(\mathbb{R}, L_x^p(\mathbb{R}^d))$  be compactly supported in  $t$ . Let  $u \in L_t^p(\mathbb{R}, W_x^{2,p}(\mathbb{R}^d))$  with  $\partial_t u \in L_t^p(\mathbb{R}, L^p(\mathbb{R}^d))$  be the unique solution to

$$\mu u + \partial_t u - \Delta u = f \quad \text{in } \mathbb{R} \times \mathbb{R}^d.$$

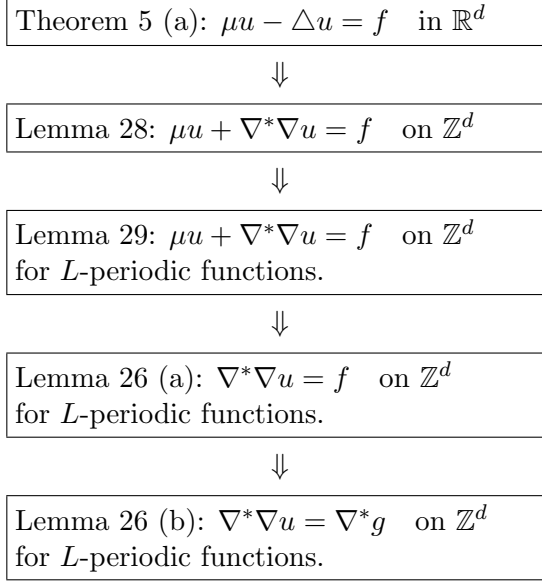
Then we have

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\nabla^2 u(t, x)|^p dx dt \lesssim \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |f(t, x)|^p dx dt,$$

where  $\nabla^2$  denotes the Hessian and  $\lesssim$  stands for  $\leq$  up to a constant that only depends on  $d$  and  $p$ .

### 7.4.1 The elliptic case

In the following diagram we summarize the general procedure that we follow to obtain the discrete Calderón-Zygmund estimate for (228) from the continuum, whole space version in Theorem 5:



**Lemma 28.** Let  $\mu > 0$ ,  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$  be compactly supported and let  $u \in \ell^2(\mathbb{Z}^d)$  denote the unique solution to

$$(231) \quad \mu u + \nabla^* \nabla u = f \quad \text{on } \mathbb{Z}^d.$$

Then we have for all  $1 < p < \infty$

$$(232) \quad \sum_{x \in \mathbb{Z}^d} |\nabla^2 u(x)|^p \lesssim \sum_{x \in \mathbb{Z}^d} |f(x)|^p$$

where  $(\nabla^2 u)_{j\ell} := \nabla_j^* \nabla_\ell u$  and  $\lesssim$  means up to a constant that only depends on  $p$  and  $d$ .

We postpone the proof of Lemma 28 to the end of this section.

**Lemma 29.** Let  $u : \mathbb{Z}^d \rightarrow \mathbb{R}$  and  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$  be  $L$ -periodic, and assume that  $\sum_{x \in ([0, L) \cap \mathbb{Z})^d} f(x) = 0$ . Let  $\mu > 0$  and suppose that

$$\mu u + \nabla^* \nabla u = f \quad \text{on } \mathbb{Z}^d.$$

Then for all  $1 < p < \infty$

$$\sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\nabla^2 u(x)|^p \lesssim \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |f(x)|^p,$$

where  $\lesssim$  means up to a constant that only depends on  $p$  and  $d$ .

*Proof of Lemma 29.* The argument relies on a simple truncation argument. For  $N \in \mathbb{N}$  let  $\chi_N$  denote the characteristic function of the  $N$ -fold periodicity cell  $([-\frac{NL}{2} - 1, \frac{NL}{2} + 1) \cap \mathbb{Z})^d$ . Note that

$$(233) \quad (\mu + \nabla^* \nabla)(u \chi_N) = f \chi_N + f_N$$

where  $f_N := \nabla^* \nabla(u \chi_N) - (\nabla^* \nabla u) \chi_N$  is the commutator of the Laplacian and the truncation. By the properties of  $\chi_N$  and discreteness we have the elementary estimate

$$|f_N(x)| \lesssim \begin{cases} \|u\|_{\ell^\infty} & \text{for } x \in ([-\frac{NL}{2} - 2, \frac{NL}{2} + 2) \cap \mathbb{Z})^d \setminus ([-\frac{NL}{2}, \frac{NL}{2}) \cap \mathbb{Z})^d, \\ 0 & \text{else,} \end{cases}$$

and thus

$$(234) \quad \sum_{x \in \mathbb{Z}^d} |f_N|^p \lesssim (NL)^{d-1} \|u\|_{\ell^\infty}^p.$$

Since the support of the r. h. s.  $f\chi_N + f_N$  is compact (it is contained in  $([-\frac{(N+1)L}{2}, \frac{(N+1)L}{2}) \cap \mathbb{Z})^d$ ), we can apply Lemma 28 to (233). Combined with the triangle inequality, the periodicity of  $f$  and (234) we obtain

$$\begin{aligned} \sum_{x \in \mathbb{Z}^d} |\nabla^* \nabla(u\chi_N)|^p &\lesssim \sum_{x \in ([-\frac{(N+1)L}{2}, \frac{(N+1)L}{2}) \cap \mathbb{Z})^d} |f\chi_N + f_N|^p \\ &\lesssim \sum_{x \in ([-\frac{(N+1)L}{2}, \frac{(N+1)L}{2}) \cap \mathbb{Z})^d} |f|^p + \sum_{x \in \mathbb{Z}^d} |f_N|^p \\ &\leq (N+1)^d \sum_{x \in \Lambda'_L} |f|^p + (NL)^{d-1} \|u\|_{\ell^\infty}^p. \end{aligned}$$

Since  $\nabla^* \nabla(u\chi_N) = \nabla^* \nabla u$  on  $([-\frac{NL}{2}, \frac{NL}{2}) \cap \mathbb{Z})^d$ , we get by periodicity of  $u$

$$\begin{aligned} \sum_{x \in ([-\frac{L}{2}, \frac{L}{2}) \cap \mathbb{Z})^d} |\nabla^* \nabla u|^p &= \frac{1}{N^d} \sum_{x \in ([-\frac{NL}{2}, \frac{NL}{2}) \cap \mathbb{Z})^d} |\nabla^* \nabla(u\chi_N)|^p \\ &\lesssim \frac{(N+1)^d}{N^d} \sum_{x \in ([-\frac{L}{2}, \frac{L}{2}) \cap \mathbb{Z})^d} |f|^p + \frac{(NL)^{d-1}}{N^d} \|u\|_{\ell^\infty}^p, \end{aligned}$$

so that the desired estimate follows by  $N \uparrow \infty$ . □

*Proof of Lemma 26. Step 1.* Argument for Lemma 29  $\Rightarrow$  (a).

This follows from the fact that the unique  $L$ -periodic mean-zero solution to  $\nabla^* \nabla u = f$  is the pointwise limit under  $\mu \downarrow 0$  of the unique  $L$ -periodic mean-zero solution of  $\mu u + \nabla^* \nabla u = f$ . The latter is a consequence of the convergence of the matrix  $\mu + \nabla^* \nabla$  to the invertible matrix  $\nabla^* \nabla$ .

**Step 2.** Argument for (a)  $\Rightarrow$  (b).

Let  $q = \frac{p}{p-1}$  denote the dual exponent of  $p$ . By duality we have

$$\left( \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\nabla_j u|^p \right)^{\frac{1}{p}} = \max_{\tilde{f}} \left| \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_j u \tilde{f} \right|,$$

where the maximum is taken among all  $L$ -periodic functions  $\tilde{f}(x)$  with  $\sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\tilde{f}|^q = 1$ . Hence, it suffices to prove that

$$(235) \quad \left| \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_j u \tilde{f} \right| \lesssim \left( \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |g|^p \right)^{\frac{1}{p}} \left( \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\tilde{f}|^q \right)^{\frac{1}{q}}.$$

Since  $u$  is  $L$ -periodic, an integration by parts yields

$$\forall \bar{f} \in \mathbb{R} : \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_j u \bar{f} = 0,$$

and thus, it suffices to prove (235) for  $L$ -periodic functions with mean value zero. Let  $\tilde{f}$  denote such a function and let  $\tilde{u}$  be the unique  $L$ -periodic solution to

$$(236) \quad \nabla^* \nabla \tilde{u} = \tilde{f} \quad \text{and} \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \tilde{u} = 0.$$

By (a) we have

$$(237) \quad \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\nabla^2 \tilde{u}|^q \lesssim \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\tilde{f}|^q.$$

By testing (236) with  $\nabla_j^* \tilde{f} = \nabla_j^* \nabla^* \nabla \tilde{u}$ , integrations by parts yield

$$\begin{aligned} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u \tilde{f} &= \sum_{x \in ([0, L] \cap \mathbb{Z})^d} u \nabla_j^* \tilde{f} = \sum_{x \in ([0, L] \cap \mathbb{Z})^d} u \nabla_j^* \nabla^* \nabla \tilde{u} = \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\nabla^* \nabla u) \nabla_j^* \tilde{u} \\ &= \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla^* g \nabla_j^* \tilde{u} = \sum_{x \in ([0, L] \cap \mathbb{Z})^d} g \cdot \nabla_j^* \nabla \tilde{u}. \end{aligned}$$

Hence, (235) follows from Hölder's inequality and (237):

$$\begin{aligned} \left| \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u \tilde{f} \right| &\lesssim \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |g|^p \right)^{\frac{1}{p}} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\nabla_j^* \nabla \tilde{u}|^q \right)^{\frac{1}{q}} \\ &\stackrel{(237)}{\lesssim} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |g|^p \right)^{\frac{1}{p}} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\tilde{f}|^q \right)^{\frac{1}{q}}. \end{aligned}$$

□

Next, we present the proof of Lemma 28 which is based on a reduction to the classical Calderón-Zygmund estimate. In the following we denote by  $\partial_n$ ,  $n = 1, \dots, d$  and  $\Delta$  the continuum partial derivatives and continuum Laplace operator. In the proof of Lemma 28 we appeal to what we call the “discrete Fourier transform”  $(\mathcal{F}_{dis}g)(k)$  of a discrete function  $g(x)$ ,  $x \in \mathbb{Z}^d$ :

$$(238) \quad (\mathcal{F}_{dis}g)(k) := (2\pi)^{-\frac{d}{2}} \sum_{x \in \mathbb{Z}^d} g(x) \exp(-ik \cdot x), \quad k \in \mathbb{R}^d.$$

The sum on the r. h. s. is required to converge absolutely which is the case when  $g \in \ell^1(\mathbb{Z}^d)$ . The “discrete Fourier transform”  $(\mathcal{F}_{dis}g)(k)$  is accompanied by the continuum inverse Fourier transform  $(\mathcal{F}^{-1}F)(x)$  of a function  $F(k)$ ,  $k \in \mathbb{R}^d$ :

$$(239) \quad (\mathcal{F}^{-1}F)(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} F(k) \exp(ik \cdot x) dk, \quad x \in \mathbb{R}^d.$$

If  $\chi$  denotes the characteristic function of the “Brillouin zone”  $(-\pi, \pi)^d$ , we note that  $\mathbb{R}^d \ni x \mapsto (\mathcal{F}^{-1}(\chi \mathcal{F}_{dis}g))(x)$  provides an extension of  $\mathbb{Z}^d \ni x \mapsto g(x)$  to  $\mathbb{R}^d$  since for all  $x \in \mathbb{Z}^d$ ,

$$(240) \quad g(x) = (\mathcal{F}^{-1}(\chi \mathcal{F}_{dis}g))(x).$$

Indeed, formula (240) can be seen as follows. Identity (238) defines a  $2\pi$ -periodic function  $(\mathcal{F}_{dis}g)(k)$  and we learn from (239) that  $\{\mathcal{F}^{-1}(\chi \mathcal{F}_{dis}g)(x)\}_{x \in \mathbb{Z}^d}$  are the Fourier coefficients of this function. On the other hand, we learn from (238) that  $\{g(x)\}_{x \in \mathbb{Z}^d}$  are the Fourier coefficients of the same function. This yields (240).

In the proof of Lemma 28 the role of  $g$  is played by  $u$  – the unique solution to (231) in  $\ell^2(\mathbb{Z}^d)$  with compactly supported r. h. s.  $f$ . We will argue below that

$$(241) \quad \mu \|u\|_{\ell^1(\mathbb{Z}^d)} \leq \|f\|_{\ell^1(\mathbb{Z}^d)},$$

so that thanks to  $\mu > 0$  we have  $\|u\|_{\ell^1(\mathbb{Z}^d)} < \infty$  and thus (238) is pointwise well-defined. Estimate (241) can be seen by appealing to the discrete elliptic Green's function associated with the operator

$\mu + \nabla^* \nabla$ . For that and future purposes, we gather some properties of the elliptic and parabolic Green's function associated with the discrete and continuum Laplacian. We denote by  $G^{cont}(t, x)$  and  $G(t, x)$  the continuum and discrete parabolic Green's function associated with  $\partial_t - \Delta$  and  $\partial_t + \nabla^* \nabla$ , respectively. For  $\mu > 0$  we denote by  $G_\mu^{cont}(t, x)$  and  $G_\mu(t, x)$  the continuum and discrete Green's function associated with the massive Laplacian  $\mu + \partial_t - \Delta$  and  $\mu + \partial_t + \nabla^* \nabla$ . Note that

$$(242) \quad G_\mu^{cont}(t, x) = \exp(-\mu t) G^{cont}(t, x) \quad \text{and} \quad G_\mu(t, x) = \exp(-\mu t) G(t, x).$$

The continuum and discrete elliptic Green's functions associated with  $\mu - \Delta$  and  $\mu + \nabla^* \nabla$  are denoted by  $G_\mu^{cont}(x)$  and  $G_\mu(x)$ , respectively. As easily follows from the defining equations, they can be obtained from their parabolic counterparts via integration in time:

$$G_\mu^{cont}(x) = \int_0^\infty G_\mu(t, x) dt \quad (\text{for } x \neq 0), \quad G_\mu(x) = \int_0^\infty G_\mu(t, x) dt.$$

Since  $G^{cont}(t, x)$  and  $G(t, x)$  are non-negative, and since

$$\int_{\mathbb{R}^d} G(t, x) dx = \lim_{t \downarrow 0} \int_{\mathbb{R}^d} G(t, x) dx = 1, \quad \sum_{x \in \mathbb{Z}^d} G(t, x) = \sum_{x \in \mathbb{Z}^d} G(0, x) = 1,$$

we deduce that

$$(243) \quad \begin{aligned} \mu \sum_{x \in \mathbb{Z}^d} G_\mu(x) &= \mu \int_0^\infty \sum_{x \in \mathbb{Z}^d} G_\mu(t, x) dt = \mu \int_0^\infty \sum_{x \in \mathbb{Z}^d} \exp(-\mu t) G(t, x) dt \\ &= \mu \int_0^\infty \exp(-\mu t) dt = 1, \end{aligned}$$

and similarly

$$(244) \quad \mu \int_{\mathbb{R}^d} G_\mu^{cont}(x) dx = 1.$$

Now (241) follows from the combination of (243) with the Green's function representation formula  $u(x) = \sum_{y \in \mathbb{Z}^d} G_\mu(x - y) f(y)$ .

*Proof of Lemma 28.* Obviously, it suffices to prove

$$(245) \quad \|\nabla_j^* \nabla_\ell v\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|f\|_{\ell^q(\mathbb{Z}^d)}$$

where  $j, \ell \in \{1, \dots, d\}$  are arbitrary but from now on fixed indices.

We split the proof into eight steps. In the first step, we reformulate the result in terms of an  $\ell^p$  estimate on the level of Fourier multipliers. The strategy is then to deduce the result in this form from its well-known continuum version by treating the low and high wave numbers separately. To this aim, we derive in Step 2 a representation formula for the difference between the discrete and continuum Fourier multipliers. In the third step we introduce a band restriction operator, and shall prove in the sixth step that for long wavelengths, this difference is controlled in  $L^p$ . To this aim we need  $\ell^1$  and  $L^1$  bounds on multipliers, which we prove in the fifth step. The uniformity of the estimates w. r. t.  $\mu$  are obtained in the fourth step and basically relies on (243) & (244). In the seventh step we show that the  $L^p$  and  $\ell^p$  norms are equivalent for long wavelengths, so that the desired  $\ell^p$  bound on the discrete Fourier multipliers follows. We then conclude the proof in Step 8 by treating the case of short wavelengths.

**Step 1.** Formulation of the result in terms of a Fourier multiplier  $\mathbf{m}_\mu \in \mathbb{R}^{d \times d}$ .

We claim that (245) is equivalent to

$$(246) \quad \left( \sum_{x \in \mathbb{Z}^d} |(\mathcal{F}^{-1}(\chi_{\mathbf{m}_\mu} \mathcal{F}_{dis} f))(x)|^p \right)^{\frac{1}{p}} \lesssim \left( \sum_{x \in \mathbb{Z}^d} |f(x)|^p \right)^{\frac{1}{p}},$$

where the meromorphic multiplier  $\mathbf{m}_\mu(k)$  is given by

$$(247) \quad \mathbf{m}_\mu(k) := \frac{(\exp(-ik_j) - 1)(\exp(ik_\ell) - 1)}{\mu + \sum_{n=1}^d |\exp(ik_n) - 1|^2}, \quad k \in \mathbb{R}^d.$$

Indeed, because of

$$\begin{aligned} (\mathcal{F}_{dis} \nabla_{n'} u)(k) &= (\exp(ik_{n'}) - 1)(\mathcal{F}_{dis} u)(k), \\ (\mathcal{F}_{dis} \nabla^* \nabla u)(k) &= \sum_{n=1}^d |\exp(ik_n) - 1|^2 (\mathcal{F}_{dis} u)(k) \end{aligned}$$

for all  $k \in \mathbb{R}^d$ , the relation  $(\mu + \nabla^* \nabla)u(x) = f(x)$  for all  $x \in \mathbb{Z}^d$ , translates into

$$(\mathcal{F}_{dis} u)(k) = \frac{1}{\mu + \sum_{n=1}^d |\exp(ik_n) - 1|^2} (\mathcal{F}_{dis} f)(k),$$

which yields  $(\mathcal{F}_{dis} \nabla_j^* \nabla_\ell u)(k) = \mathbf{m}_\mu(k) (\mathcal{F}_{dis} f)(k)$  with  $\mathbf{m}_\mu(k)$  given by (247). In view of (240), this can be reformulated as  $\nabla_j^* \nabla_\ell u(x) = (\mathcal{F}^{-1}(\chi \mathbf{m}_\mu \mathcal{F}_{dis} f))(x)$ ,  $x \in \mathbb{Z}^d$ . This establishes the equivalence of (245) and (246).

**Step 2.** Representation of the difference between the discrete Fourier multiplier  $\mathbf{m}_\mu$  and the continuum Fourier multiplier  $\mathbf{m}_\mu^{cont} := \frac{k_j k_\ell}{\mu + |k|^2}$ .

We introduce the auxiliary function

$$(248) \quad h(z) := \begin{cases} \frac{\exp(iz) - 1}{iz} & \text{for } z \neq 0 \\ 1 & \text{for } z = 0 \end{cases}, \quad \text{for } z \in \mathbb{C},$$

and note that  $h(z)$  and  $\frac{1}{h(z)}$  are analytic in the disk  $|z| < 2\pi$ . In particular, we have

$$(249) \quad \frac{1}{h(z)} = 1 - \frac{iz}{2} - z^2 \hat{h}(z) \quad \text{for some smooth function } \hat{h}(z) \in \mathbb{C} \text{ on the disc } |z| < 2\pi.$$

We claim that for  $k \in (-2\pi, 2\pi)$  we have

$$(250) \quad \mathbf{m}_\mu - \mathbf{m}_\mu^{cont} = \mathbf{m}_\mu (\mathbf{m}^* + (1 - \mathfrak{R}_\mu^{cont}) \mathbf{m}^{**}),$$

$$(251) \quad \mathbf{m}_\mu = (1 - \mathfrak{R}_\mu) \mathbf{m}_0,$$

where  $\mathbf{m}_0$  is the discrete multiplier for  $\mu = 0$ , i. e.

$$\mathbf{m}_0(k) := \frac{(\exp(-ik_j) - 1)(\exp(ik_\ell) - 1)}{\sum_{n=1}^d |\exp(ik_n) - 1|^2},$$

where  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$  account for the deviation of the discrete from the continuum multipliers, i. e.

$$\begin{aligned} \mathbf{m}^*(k) &:= 1 - \frac{1}{h(-k_j)h(k_\ell)}, \\ \mathbf{m}^{**}(k) &:= \frac{1}{|k|^2} \sum_{n=1}^d k_n^2 \left( \frac{1 - |h(k_n)|^2}{h(-k_j)h(k_\ell)} \right), \end{aligned}$$

and where  $\mathfrak{R}_\mu$  and  $\mathfrak{R}_\mu^{cont}$  account for the difference between  $\mu > 0$  and  $\mu = 0$  in the discrete and the continuum case, respectively, i. e.

$$\begin{aligned} \mathfrak{R}_\mu(k) &:= \frac{\mu}{\mu + \sum_{n=1}^d |\exp(ik_n) - 1|^2}, \\ \mathfrak{R}_\mu^{cont}(k) &:= \frac{\mu}{\mu + |k|^2}. \end{aligned}$$

Indeed, with the help of  $h(z)$ , (247) can be rewritten as

$$\mathbf{m}_\mu(k) = \frac{k_j k_\ell h(-k_j) h(k_\ell)}{\mu + \sum_{n=1}^d k_n^2 |h(k_n)|^2}.$$

Hence, we have

$$\begin{aligned} \mathbf{m}_\mu - \mathbf{m}_\mu^{\text{cont}} &= \mathbf{m}_\mu \left( 1 - \frac{\mathbf{m}_\mu^{\text{cont}}}{\mathbf{m}_\mu} \right) = \mathbf{m}_\mu \left( 1 - \frac{\mu + \sum_{n=1}^d k_n^2 |h(k_n)|^2}{(\mu + |k|^2) h(-k_j) h(k_\ell)} \right) \\ &= \mathbf{m}_\mu \left( 1 - \frac{\mu}{\mu + |k|^2} \frac{1}{h(-k_j) h(k_\ell)} - \frac{\sum_{n=1}^d k_n^2 |h(k_n)|^2}{(\mu + |k|^2) h(-k_j) h(k_\ell)} \right). \end{aligned}$$

We rearrange the r. h. s. using the following two identities

$$\begin{aligned} 1 - \frac{\mu}{\mu + |k|^2} \frac{1}{h(-k_j) h(k_\ell)} &= \left( 1 - \frac{1}{h(-k_j) h(k_\ell)} \right) + \left( 1 - \frac{\mu}{\mu + |k|^2} \right) \frac{1}{h(-k_j) h(k_\ell)} \\ &= \mathbf{m}^* + (1 - \mathfrak{R}_\mu^{\text{cont}}) \frac{1}{|k|^2} \sum_{n=1}^d \frac{k_n^2}{h(-k_j) h(k_\ell)} \end{aligned}$$

and

$$\begin{aligned} \frac{\sum_{n=1}^d k_n^2 |h(k_n)|^2}{(\mu + |k|^2) h(-k_j) h(k_\ell)} &= \left( 1 - \frac{\mu}{\mu + |k|^2} \right) \frac{1}{|k|^2} \sum_{n=1}^d \frac{k_n^2 |h(k_n)|^2}{h(-k_j) h(k_\ell)} \\ &= (1 - \mathfrak{R}_\mu^{\text{cont}}) \frac{1}{|k|^2} \sum_{n=1}^d \frac{k_n^2 |h(k_n)|^2}{h(-k_j) h(k_\ell)}, \end{aligned}$$

so that we get

$$\mathbf{m}_\mu - \mathbf{m}_\mu^{\text{cont}} = \mathbf{m}_\mu \left( \mathbf{m}^* + (1 - \mathfrak{R}_\mu^{\text{cont}}) \frac{1}{|k|^2} \sum_{n=1}^d k_n^2 \left( \frac{1 - |h(k_n)|^2}{h(-k_j) h(k_\ell)} \right) \right).$$

This proves (250). The identity (251) can be seen as follows:

$$\mathbf{m}_\mu = \frac{\sum_{n=1}^d |\exp(ik_n) - 1|^2}{\mu + \sum_{n=1}^d |\exp(ik_n) - 1|^2} \mathbf{m}_0 = \left( 1 - \frac{\mu}{\mu + \sum_{n=1}^d |\exp(ik_n) - 1|^2} \right) \mathbf{m}_0.$$

By virtue of (249) we deduce that  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$  are linear combinations of terms of the form

$$(252) \quad ik_{n'} \varphi(k),$$

and

$$(253) \quad ik_{n'} \frac{k_n^2}{|k|^2} \varphi(k),$$

respectively. Here,  $n, n' \in \{1, \dots, d\}$  denote generic indices and  $\varphi(k)$  denotes a generic complex-valued function that is smooth on  $k \in (-2\pi, 2\pi)^d$ . The indices  $n, n'$  and the function  $\varphi$  may vary from term to term, but they do not depend on  $\mu$ .

**Step 3.**  $L^1$ -estimates for  $\mathfrak{R}_\mu^{\text{cont}}$  and  $\mathfrak{R}_\mu$  uniform in  $\mu$ .

We claim that

$$(254) \quad \begin{aligned} \int |\mathcal{F}^{-1}(\mathfrak{R}_\mu^{\text{cont}})| dx &\lesssim 1, \\ \sum_{x \in \mathbb{Z}^d} |\mathcal{F}^{-1}(\mathfrak{R}_\mu \chi)| &\lesssim 1. \end{aligned}$$



Since  $\mathfrak{R}_\mu^{cont}$  and  $\mathfrak{R}_\mu$  are the continuum and discrete Fourier multipliers associated with  $\mu(\mu - \Delta)^{-1}$  and  $\mu(\mu + \nabla^* \nabla)^{-1}$ , we have

$$\begin{aligned}\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont})(x) &= \mu G_\mu^{cont}(x) && \text{for a.e. } x \in \mathbb{R}^d, \\ \mathcal{F}^{-1}(\mathfrak{R}_\mu \chi)(x) &= \mu G_\mu(x) && \text{for all } x \in \mathbb{Z}^d.\end{aligned}$$

Hence, the statement directly follows from the uniform estimates (243) and (244).

**Step 4.** Truncation of high wave-numbers.

In this step we introduce a cut-off function that truncates (high) wave-numbers with  $L^{-1} \ll 1$ . Let  $k \mapsto \eta_1(k)$  denote a fixed smooth function such that

$$(255) \quad \eta_1(k) = 1 \text{ for } k \in [-\pi, \pi]^d \quad \text{and} \quad \eta_1 \text{ is compactly supported in } (-2\pi, 2\pi)^d.$$

For all  $L \geq 1$  we set

$$\eta_L(k) := \eta_1(Lk).$$

Let  $\varphi$  be some smooth function on  $(-2\pi, 2\pi)^d$ , and consider the Fourier transforms

$$K_L := \mathcal{F}^{-1}(\varphi \eta_L), \quad \hat{K}_L := \mathcal{F}^{-1}(\varphi(\frac{\cdot}{L}) \eta_1).$$

By the well-known relationship between (inverse) Fourier transform and dilation, we obtain

$$(256) \quad K_L(x) = \frac{1}{L^d} \hat{K}_L\left(\frac{x}{L}\right).$$

Since  $\varphi$  is smooth on  $(-2\pi, 2\pi)^d$ , and since  $\eta_1$  is compactly supported in  $(-2\pi, 2\pi)^d$ , the family  $\{\mathbb{R}^d \ni k \mapsto \varphi(\frac{k}{L}) \eta_1(k)\}_{L \geq 1}$  of functions is bounded in the Schwartz class (that is, the semi-norms  $\sup_{\mathbb{R}^d} |k^\alpha D^\alpha F|$  of  $F(k) = \varphi(\frac{k}{L}) \eta_1(k)$ , controlling weights  $k^\alpha = k_1^{\alpha_1} \cdots k_d^{\alpha_d}$  and derivatives  $D^\alpha = \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$ , are bounded for  $L \geq 1$  and for any multi index  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ ). Thus the family  $\{\mathbb{R}^d \ni x \mapsto \hat{K}_L(x)\}_{L \geq 1}$  is bounded in the Schwartz class as well, so that

$$(257) \quad \int |D^\alpha \hat{K}_L| d\hat{x} \lesssim 1$$

up to a multiplicative constant which only depends on  $\alpha$  and  $\varphi$  next to  $\eta_1$ ; here and in the sequel,  $\int$  stands for  $\int_{\mathbb{R}^d}$ .

**Step 5.**  $L^1$ -estimate of the truncated multipliers  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$ .

This is the key step in the proof. Let  $\eta_L$  be the cut-off function of Step 4. We consider the inverse Fourier transforms of the truncated multipliers  $\mathbb{R}^d \ni k \mapsto \mathbf{m}^*(k) \eta_L(k)$  and  $\mathbb{R}^d \ni k \mapsto \mathbf{m}^{**}(k) \eta_L(k)$ , and claim that for  $L \geq 1$

$$(258a) \quad \int |\mathcal{F}^{-1}(\mathbf{m}^* \eta_L)| dx \lesssim L^{-1},$$

$$(258b) \quad \int |\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L)| dx \lesssim L^{-1}.$$

We start with (258a). Since  $\mathbf{m}^*$  is a linear combination of terms of the form (252) we only need to show that for any smooth function  $\varphi$  on  $(-2\pi, 2\pi)^d$  we have

$$(259) \quad \int |\mathcal{F}^{-1}(ik_n \varphi \eta_L)| dx \lesssim L^{-1}.$$

By standard properties of the (inverse) Fourier transform we get

$$\mathcal{F}^{-1}(ik_n \varphi \eta_L) = \partial_n \mathcal{F}^{-1}(\varphi \eta_L) = \partial_n K_L,$$

where  $K_L$  (and  $\hat{K}_L$  below) are as in Step 4. Hence, Step 4 and the change of variables  $x \rightsquigarrow L\hat{x}$  yield

$$\int \mathcal{F}^{-1}(ik_n \varphi \eta_L) dx = \int |\partial_n K_L| dx = L^{-1} \int |\partial_n \hat{K}_L| d\hat{x} \stackrel{(257)}{\lesssim} L^{-1}.$$

Next we treat  $\mathbf{m}^{**}$ , see (258b). By virtue of (253), it suffices to prove that for any smooth function  $\varphi$  on  $(-2\pi, 2\pi)^d$ ,

$$(260) \quad \int |\mathcal{F}^{-1}(k_n^2 ik_{n'} \frac{1}{|k|^2} \varphi \eta_L)| dx \lesssim L^{-1}.$$

We start by recalling some standard properties of the (inverse) Fourier transform (239), differentiation, convolution, and the fundamental solution of the Laplace equation:

$$\begin{aligned} \mathcal{F}^{-1}(k_n^2 k_{n'} \frac{1}{|k|^2} \varphi \eta_L) &= -\partial_n^2 \mathcal{F}^{-1}(k_{n'} \frac{1}{|k|^2} \varphi \eta_L), \\ \mathcal{F}^{-1}(k_{n'} \frac{1}{|k|^2} \varphi \eta_L) &= (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1}(k_{n'} \frac{1}{|k|^2}) * \mathcal{F}^{-1}(\varphi \eta_L), \\ \mathcal{F}^{-1}(ik_{n'} \frac{1}{|k|^2}) &= -\frac{1}{|S^{d-1}|} \frac{x_{n'}}{|x|^d}, \end{aligned}$$

that combine to

$$(261) \quad \mathcal{F}^{-1}(k_n^2 ik_{n'} \frac{1}{|k|^2} \varphi \eta_L) = \frac{(2\pi)^{\frac{d}{2}}}{|S^{d-1}|} \partial_n^2 \left( \frac{x_{n'}}{|x|^d} * K_L \right),$$

where  $K_L$  is associated with  $\varphi$  as in Step 4, and  $|S^{d-1}|$  is the volume of the  $(d-1)$ -dimensional hypersphere. We now prove that (260) follows from the representation formulas (261) & (256) combined with the estimate (257). For that purpose, we introduce a spatial cut-off function  $\zeta_L$  at scale  $L$  by rescaling a smooth function  $\zeta_1$  chosen so that

$$\zeta_1(x) = 1 \text{ for } |x| \leq 1 \quad \text{and} \quad \zeta_1(x) = 0 \text{ for } |x| \geq 2;$$

that is, we set

$$\zeta_L(x) := \zeta_1\left(\frac{x}{L}\right).$$

By the triangle inequality in  $L^1$ , estimate (260) is a consequence of the two estimates

$$(262) \quad \int \left| \partial_n^2 \left( (\zeta_L \frac{x_{n'}}{|x|^d}) * K_L \right) \right| dx \lesssim L^{-1},$$

$$(263) \quad \int \left| \partial_n^2 \left( ((1 - \zeta_L) \frac{x_{n'}}{|x|^d}) * K_L \right) \right| dx \lesssim L^{-1},$$

which we prove separately. We start with (262): Using the relationship between convolution on the one side and differentiation and integration on the other side, and using the change of variables  $x \rightsquigarrow L\hat{x}$ , we obtain

$$\begin{aligned} \int \left| \partial_n^2 \left( (\zeta_L \frac{x_{n'}}{|x|^d}) * K_L \right) \right| dx &= \int \left| (\zeta_L \frac{x_{n'}}{|x|^d}) * (\partial_n^2 K_L) \right| dx \\ &\leq \int \left| \zeta_L \frac{x_{n'}}{|x|^d} \right| dx \int |\partial_n^2 K_L| dx \\ &= L \int \left| \zeta_1 \frac{\hat{x}_{n'}}{|\hat{x}|^d} \right| d\hat{x} L^{-2} \int |\partial_n^2 \hat{K}_L| d\hat{x}. \end{aligned}$$

Since  $\int \left| \zeta_1 \frac{\hat{x}_{n'}}{|\hat{x}|^d} \right| d\hat{x} < \infty$  (due to the integrability of  $x \mapsto \frac{\hat{x}_{n'}}{|\hat{x}|^d}$  for  $|\hat{x}| \lesssim 1$ ), and since (257) implies that  $\int \left| \partial_n^2 \hat{K}_L \right| d\hat{x}$  is uniformly bounded for  $L \geq 1$ , this establishes (262). The argument for (263) is similar:

$$\begin{aligned} & \int \left| \partial_n^2 \left( ((1 - \zeta_L) \frac{x_{n'}}{|x|^d}) * K_L \right) \right| dx \\ &= \int \left| (\partial_n^2 ((1 - \zeta_L) \frac{x_{n'}}{|x|^d})) * K_L \right| dx \\ &\leq \int \left| \partial_n^2 ((1 - \zeta_L) \frac{x_{n'}}{|x|^d}) \right| dx \int |K_L| dx \\ &= L^{-1} \int \left| \partial_n^2 ((1 - \zeta_1) \frac{\hat{x}_{n'}}{|\hat{x}|^d}) \right| d\hat{x} \int |\hat{K}_L| d\hat{x}. \end{aligned}$$

This implies (263) since this time  $\int \left| \partial_n^2 ((1 - \zeta_1) \frac{\hat{x}_{n'}}{|\hat{x}|^d}) \right| d\hat{x} < \infty$  (due to the integrability of  $\partial_n^2 \frac{\hat{x}_{n'}}{|\hat{x}|^d}$  for  $|\hat{x}| \gtrsim 1$ ) and using again that (257) implies that  $\int |\hat{K}_L| d\hat{x}$  is uniformly bounded for  $L \geq 1$ .

**Step 6.**  $L^p$ -boundedness of the multiplier  $\mathfrak{m}_\mu$  on functions with band restriction.

We claim that there exists an  $L \gg 1$  (that can be chosen uniformly in  $\mu > 0$ ) such that for any field  $F(k) \in \mathbb{C}$  that is band-restricted in the sense of

$$(264) \quad F(k) \text{ is supported in } k \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right]^d,$$

(so that  $F(k)$  is a Fourier transform of a discrete or continuum function that has no length scales smaller than  $L$ ), we have

$$(265) \quad \|\mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p} \lesssim \|\mathcal{F}^{-1}F\|_{L^p},$$

where  $\|g\|_{L^p} := \left(\int_{\mathbb{R}^d} |g(x)|^p dx\right)^{\frac{1}{p}}$ .

We will use Step 2 and Step 5 to reduce this statement to its well-established counterpart with continuum Fourier multiplier  $\mathfrak{m}_\mu^{\text{cont}}(k) = \frac{k_j k_\ell}{\mu + |k|^2}$ , that is,

$$(266) \quad \|\mathcal{F}^{-1}(\mathfrak{m}_\mu^{\text{cont}} F)\|_{L^p} \lesssim \|\mathcal{F}^{-1}F\|_{L^p},$$

which indeed holds even without the band restriction (264), see Theorem 5 (a). By the triangle inequality in  $L^p$  we have

$$\|\mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p} \leq \|\mathcal{F}^{-1}(\mathfrak{m}_\mu^{\text{cont}} F)\|_{L^p} + \|\mathcal{F}^{-1}((\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})F)\|_{L^p}.$$

Hence, in view of (266), it is enough to show that the band restriction (264) implies

$$(267) \quad \|\mathcal{F}^{-1}((\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})F)\|_{L^p} \lesssim L^{-1} \|\mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p},$$

which allows us to absorb this error term into the l. h. s. of (265).

The remainder of the step is devoted to the proof of (267). By the band restriction (264) and the definition (255) of the cut-off  $\eta_L$ , we have

$$(\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})F = (\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})\eta_L F.$$

By Step 2, cf. (250), this takes the form

$$(\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})F = \left( \mathfrak{m}^* + (1 - \mathfrak{R}_\mu^{\text{cont}})\mathfrak{m}^{**} \right) \eta_L \mathfrak{m}_\mu F.$$

By the triangle inequality we then get

$$(268) \quad \begin{aligned} \|\mathcal{F}^{-1}((\mathfrak{m}_\mu - \mathfrak{m}_\mu^{\text{cont}})F)\|_{L^p} &\leq \|\mathcal{F}^{-1}(\mathfrak{m}^* \eta_L \mathfrak{m}_\mu F)\|_{L^p} + \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p} \\ &\quad + \|\mathcal{F}^{-1}(\mathfrak{R}_\mu^{\text{cont}} \mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p}. \end{aligned}$$

By Young's convolution estimate and (254) we have for the last r. h. s. term

$$\begin{aligned} \|\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont} \mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p} &= (2\pi)^{\frac{d}{2}} \|\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont}) * \mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p} \\ &\leq (2\pi)^{\frac{d}{2}} \|\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont})\|_{L^1} \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p} \\ &\lesssim \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p}. \end{aligned}$$

In combination with (268) we get

$$\begin{aligned} \|\mathcal{F}^{-1}((\mathfrak{m}_\mu - \mathfrak{m}_\mu^{cont})F)\|_{L^p} &\lesssim \|\mathcal{F}^{-1}(\mathfrak{m}^* \eta_L \mathfrak{m}_\mu F)\|_{L^p} + \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L \mathfrak{m}_\mu F)\|_{L^p} \\ &= (2\pi)^{\frac{d}{2}} \|\mathcal{F}^{-1}(\mathfrak{m}^* \eta_L) * \mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p} + \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L) * \mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p} \\ &= (2\pi)^{\frac{d}{2}} \left( \|\mathcal{F}^{-1}(\mathfrak{m}^* \eta_L)\|_{L^1} + \|\mathcal{F}^{-1}(\mathfrak{m}^{**} \eta_L)\|_{L^1} \right) \|\mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{L^p} \end{aligned}$$

and (267) follows from (258a), and (258b).

**Step 7.**  $\ell^p$ -boundedness of the multiplier  $\mathfrak{m}_\mu$  on functions with band restriction.

We claim that there exists an  $L \gg 1$  (that can be chosen uniformly in  $\mu > 0$ ) such that for any field  $F(k) \in \mathbb{C}$  that is band-restricted in the sense of (264), we have

$$(269) \quad \|\mathcal{F}^{-1}(\mathfrak{m}_\mu F)\|_{\ell^p} \lesssim \|\mathcal{F}^{-1}F\|_{\ell^p},$$

where  $\|g\|_{\ell^p} := \left( \sum_{x \in \mathbb{Z}^d} |g(x)|^p \right)^{\frac{1}{p}}$ .

In view of Step 6, it is enough to show that the band restriction (264) implies the equivalence of the discrete and continuum norms, that is,

$$(270) \quad \|g\|_{\ell^p} \sim \|g\|_{L^p}, \quad \text{where } g := \mathcal{F}^{-1}F,$$

which we split into the two estimates

$$(271) \quad \|g\|_{\ell^p} \lesssim \|g\|_{L^p},$$

$$(272) \quad \|g\|_{L^p} \lesssim \|g\|_{\ell^p}.$$

We start with the estimate (271): By the Sobolev embedding into Hölder spaces, there exists an  $n \in \mathbb{N}$  only depending on  $d$  such that

$$(273) \quad |g(0)| \leq \sup_{(0,1)^d} |g| \lesssim \left( \int_{(0,1)^d} |g|^p dx \right)^{\frac{1}{p}} + \left( \int_{(0,1)^d} |\nabla^n g|^p dx \right)^{\frac{1}{p}}.$$

Appealing to translation invariance and the triangle inequality in  $\ell^p$  yields

$$\|g\|_{\ell^p} \lesssim \|g\|_{L^p} + \|\nabla^n g\|_{L^p}.$$

In order to obtain (271), it remains to argue that the band restriction (264) implies  $\|\nabla^n g\|_{L^p} \lesssim \|g\|_{L^p}$ . In fact, we have the stronger statement

$$(274) \quad \|\nabla^n g\|_{L^p} \lesssim L^{-n} \|g\|_{L^p}.$$

Estimate (274) can be seen as follows: In view of the band restriction (264) of  $F$  and the definition (255) of the cut-off  $\eta_L$  we have  $g = \mathcal{F}^{-1}F = \mathcal{F}^{-1}(\eta_L F) = (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1} \eta_L * \mathcal{F}^{-1}F = \phi_L * g$ , where the family of kernels  $\{\phi_L := \mathcal{F}^{-1} \eta_L\}_{L \geq 1}$  satisfies  $\phi_L(x) = \frac{1}{L^d} \phi_1\left(\frac{x}{L}\right)$  and  $\phi_1$  is in the Schwartz class (being the inverse Fourier transform of the Schwartz class function  $k \mapsto (2\pi)^{\frac{d}{2}} \eta_1(k)$ ). We thus obtain the representation  $\nabla^n g = \nabla^n(\phi_L * g) = (\nabla^n \phi_L) * g$  with  $\nabla^n \phi_L(x) = L^{-n} \frac{1}{L^d} \nabla^n \phi_1\left(\frac{x}{L}\right)$ , which yields the inequality

$$\|\nabla^n g\|_{L^p} \leq \|\nabla^n \phi_L\|_{L^1} \|g\|_{L^p} = L^{-n} \|\nabla^n \phi_1\|_{L^1} \|g\|_{L^p},$$

and thus the estimate (274), since  $\phi_1$  is in the Schwartz class.

We now turn to (272). By the above Sobolev embedding into Hölder spaces, cf. (274), we also have

$$\begin{aligned} \left( \int_{(0,1)^d} |g - g(0)|^p dx \right)^{\frac{1}{p}} &\lesssim \sup_{(0,1)^d} |\nabla g| \\ &\lesssim \left( \int_{(0,1)^d} |\nabla g|^p dx \right)^{\frac{1}{p}} + \left( \int_{(0,1)^d} |\nabla^{n+1} g|^p dx \right)^{\frac{1}{p}}, \end{aligned}$$

which yields

$$\left( \int_{(0,1)^d} |g|^p dx \right)^{\frac{1}{p}} \lesssim |g(0)| + \left( \int_{(0,1)^d} |\nabla g|^p dx \right)^{\frac{1}{p}} + \left( \int_{(0,1)^d} |\nabla^{n+1} g|^p dx \right)^{\frac{1}{p}}$$

and then

$$\|g\|_{L^p} \lesssim \|g\|_{\ell^p} + \|\nabla g\|_{L^p} + \|\nabla^{n+1} g\|_{L^p}.$$

In view of (274), we have  $\|\nabla g\|_{L^p} + \|\nabla^{n+1} g\|_{L^p} \lesssim (L^{-1} + L^{-n-1})\|g\|_{L^p} \ll \|g\|_{L^p}$ , so that we may absorb this error term into the l. h. s. .

**Step 8. Conclusion.**

We prove (246), that is

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \lesssim \|f\|_{\ell^p}.$$

It remains to treat the wave lengths that are of order unity, that is, of the grid spacing. We fix an  $L \gg 1$  so large that the conclusion of Step 7 holds with  $L$  replaced by  $L/2$ . We apply Step 7 to  $F = \eta_L \chi_{\mathcal{F}_{dis}} f$ , note that by definition (255) of  $\eta_L$ ,  $F$  is supported in  $(-\frac{2\pi}{L}, \frac{2\pi}{L})^d$  so that (269) holds with  $L$  replaced by  $L/2$ . Using the triangle inequality in  $\ell^p$  and  $\mathcal{F}^{-1}(\chi_{\mathcal{F}_{dis}} f) = f$  on  $\mathbb{Z}^d$  (cf. (240)), we obtain

$$\begin{aligned} &\|\mathcal{F}^{-1}(\mathbf{m}_\mu \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \\ &\leq \|\mathcal{F}^{-1}(\mathbf{m}_\mu \eta_L \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} + \|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \\ &\stackrel{(269)}{\lesssim} \|\mathcal{F}^{-1}(\eta_L \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} + \|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \\ &\leq \|f\|_{\ell^p} + \|\mathcal{F}^{-1}((1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} + \|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p}. \end{aligned}$$

It thus remains to show that

$$(275) \quad \|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} + \|\mathcal{F}^{-1}((1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \lesssim \|f\|_{\ell^p}.$$

We only give the argument for the first l. h. s. term, the second one is treated in a similar way. By the relation between the discrete Fourier transform  $[-\pi, \pi]^d \ni k \mapsto (\mathcal{F}_{dis} f)(k) =: F(k)$ , its inverse  $\mathbb{Z}^d \ni x \mapsto (\mathcal{F}^{-1}(\chi F))(x)$ , cf. (240), and the discrete convolution  $*_{dis}$  we have the identity

$$\begin{aligned} \mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f) &= \mathcal{F}^{-1}(\chi \mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f) \\ &= (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1}(\chi \mathbf{m}_\mu (1 - \eta_L)) *_{dis} \mathcal{F}^{-1}(\chi_{\mathcal{F}_{dis}} f) \\ &= (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1}(\chi \mathbf{m}_\mu (1 - \eta_L)) *_{dis} f \quad \text{on } \mathbb{Z}^d, \end{aligned}$$

and thus the estimate

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi_{\mathcal{F}_{dis}} f)\|_{\ell^p} \leq (2\pi)^{\frac{d}{2}} \|\mathcal{F}^{-1}(\chi \mathbf{m}_\mu (1 - \eta_L))\|_{\ell^1} \|f\|_{\ell^p}.$$

Recall that  $\mathbf{m}_\mu = \mathbf{m}_0 - \mathfrak{R}_\mu \mathbf{m}_0$ . Since  $\eta_L$  is smooth, equal to 1 in a neighborhood of the origin and has support in  $(-\pi, \pi)^d$ , and because  $\mathbf{m}_0$  is smooth in  $\mathbb{R}^d \setminus (2\pi\mathbb{Z})^d$  and  $2\pi$ -periodic, the function  $\chi \mathbf{m}_0 (1 - \eta_L) = \mathbf{m}_0 (1 - \eta_L) 1_{(-\pi, \pi)^d}$  can be extended by periodicity to a smooth function on  $\mathbb{R}^d$ . Therefore,

its Fourier coefficients  $\mathcal{F}^{-1}(\chi \mathbf{m}_0(1 - \eta_L)) = (2\pi)^{-\frac{d}{2}} \int_{(-\pi, \pi)^d} \mathbf{m}_0(1 - \eta_L) \exp(ik \cdot x) dk$ ,  $x \in \mathbb{Z}^d$ , decay faster than any negative power of  $|x|$ , and in particular  $\|\mathcal{F}^{-1}(\chi \mathbf{m}_0(1 - \eta_L))\|_{\ell^1} < \infty$ . Combined with Step 3, we deduce that

$$\begin{aligned} \|\mathcal{F}^{-1}(\chi \mathbf{m}_\mu(1 - \eta_L))\|_{\ell^1} &= \|\mathcal{F}^{-1}((1 - \mathfrak{R}_\mu) \chi \mathbf{m}_0(1 - \eta_L))\|_{\ell^1} \\ &\stackrel{(254)}{\lesssim} \|\mathcal{F}^{-1}(\chi \mathbf{m}_0(1 - \eta_L))\|_{\ell^1} \lesssim 1 \end{aligned}$$

as desired.  $\square$

#### 7.4.2 The parabolic case

The general strategy to obtain the discrete, periodic Calderón-Zygmund estimate for (230) from Theorem 5 (b) is similar to the elliptic case. However, in contrast to the elliptic case, we directly pass (in the  $L$ -periodic case) from the Calderón-Zygmund estimate for  $(\mu + \partial_t + \nabla^* \nabla)^{-1} f$  to the estimate for  $(\partial_t + \nabla^* \nabla)^{-1}(\nabla^* g)$ , i. e. we skip the statement that corresponds to Lemma 26 (a) in the elliptic setting. The parabolic versions of Lemma 28 and Lemma 29 read:

**Lemma 30.** *Let  $\mu > 0$ ,  $f : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  be compactly supported and let  $u : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  be smooth in  $t$  and related to  $f$  via the equation*

$$(276) \quad \mu u + \partial_t u + \nabla^* \nabla u = f.$$

Then we have for all  $1 < p < \infty$

$$\left( \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} |\nabla^2 u(t, x)|^p dt \right)^{\frac{1}{p}} \lesssim \left( \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} |f(t, x)|^p dt \right)^{\frac{1}{p}},$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p$  and  $d$ .

We postpone the proof of Lemma 30 to the end of this section and directly turn to the periodic case.

**Lemma 31.** *Let  $\mu > 0$ ,  $u, f : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  be  $L$ -periodic in  $x$ , smooth and compactly supported in  $t$ . Suppose that  $u$  is related to  $f$  via the equation*

$$(277) \quad \mu u + \partial_t u + \nabla^* \nabla u = f.$$

Then we have for all  $1 < p < \infty$

$$\left( \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\nabla^2 u(t, x)|^p dt \right)^{\frac{1}{p}} \lesssim \left( \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |f(t, x)|^p dt \right)^{\frac{1}{p}},$$

where  $\lesssim$  means  $\leq$  up to a multiplicative constant that only depends on  $p$  and  $d$ .

*Proof of Lemma 31.* The proof is similar to the proof of Lemma 29 where the elliptic case is treated. It relies on a truncation argument. Let  $\chi_N$  be defined as in the proof of Lemma 29. We have

$$(\mu + \partial_t + \nabla^* \nabla)(u \chi_N) = f \chi_N + f_N,$$

where  $f_N := \nabla^* \nabla(u \chi_N) - \chi_N \nabla^* \nabla u$ . By appealing to the standard a priori estimate

$$\int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |u(t, x)|^2 dt \leq \frac{1}{\mu^2} \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |f(t, x)|^2 dt$$

we deduce, as in the proof of Lemma 29, that

$$\int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} |f_N(t, x)|^p dt \leq C(NL)^{d-1}$$

where here and below  $C$  denotes a generic constant that is independent of  $N$ . Since  $u_{\chi_N}$ ,  $f_{\chi_N}$  and  $f_N$  are compactly supported in space and time, we can apply Lemma 28, and, as in the proof of Lemma 29 we arrive at

$$\int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |\nabla^2 u(t, x)|^p dt \lesssim \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |f(t, x)|^p dt + C \frac{(NL)^{d-1}}{N^d}.$$

Now the desired estimate follows by  $N \rightarrow \infty$ .  $\square$

*Proof of Lemma 27.* As in the elliptic case, cf. Lemma 26, the proof relies on a duality argument. Set  $q := \frac{p}{p-1}$ . For fixed  $j = 1, \dots, d$  we only need to prove that

$$(278) \quad \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_j u f dt \right| \lesssim \|g\|_{L_t^p \ell_x^p} \|f\|_{L_t^q \ell_x^q}$$

for all  $f(t, x)$   $L$ -periodic in  $x$  and compactly supported in  $t$ . Here and below  $\|\cdot\|_{L_t^p \ell_x^p}$  denotes the norm

$$\|g\|_{L_t^p \ell_x^p} := \left( \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} |g(t, x)|^p dt \right)^{\frac{1}{p}}.$$

In the first step we establish (278) (modulo an additional lower order term on the r. h. s.) for test functions of the form  $f = (\mu - \partial_t + \nabla^* \nabla)v$  with  $v$  smooth and compactly supported in time. In the second step we argue by a density argument that it is indeed sufficient to consider such test functions.

**Step 1.** The duality argument.

Let  $v(t, x)$  be smooth and compactly supported in  $t$ ,  $L$ -periodic in  $x$ , and assume that

$$(279) \quad \sum_{x \in ([0, L) \cap \mathbb{Z})^d} v(t, x) = 0 \quad \text{for all } t \in \mathbb{R}.$$

For  $\mu > 0$  define  $f(t, x)$  via

$$(280) \quad f := (\mu - \partial_t + \nabla^* \nabla)v.$$

We claim that for any  $\mu > 0$  we have

$$(281) \quad \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L) \cap \mathbb{Z})^d} \nabla_j u f dt \right| \lesssim \|g\|_{L_t^p \ell_x^p} \|f\|_{L_t^q \ell_x^q} + \mu L^2 \|\nabla u\|_{L_t^p \ell_x^p} \|f\|_{L_t^q \ell_x^q}.$$

For the argument note that the function  $(t, x) \mapsto v(-t, x)$  satisfies (277) with r. h. s.  $(t, x) \mapsto f(-t, x)$ . Hence, Lemma 31 yields

$$(282) \quad \|\nabla^2 v\|_{L_t^q \ell_x^q} \lesssim \|f\|_{L_t^q \ell_x^q}.$$

Now we turn to (278). Integration by parts yields

$$\begin{aligned}
(283) \quad & \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u f \, dt \stackrel{(280)}{=} \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u (\mu - \partial_t + \nabla^* \nabla) v \, dt \\
&= \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} ((\mu + \partial_t + \nabla^* \nabla) \nabla_j u) v \, dt \\
&= \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\partial_t + \nabla^* \nabla) u \nabla_j^* v \, dt + \mu \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} u \nabla_j^* v \, dt \\
&\stackrel{(230)}{=} \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} g \cdot \nabla \nabla_j^* v \, dt + \mu \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\nabla_j u) v \, dt.
\end{aligned}$$

We estimate the first term on the r. h. s. by appealing to Hölder's inequality and (282):

$$(284) \quad \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} g \cdot \nabla \nabla_j^* v \, dt \right| \lesssim \|g\|_{L_t^p \ell_x^p} \|f\|_{L_t^q \ell_x^q}.$$

Next we estimate the second term on the r. h. s. in (283). By applying Hölder's inequality and twice Poincaré's inequality to  $v$  (using (279)), we have

$$\begin{aligned}
& \mu \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\nabla_j u) v \, dt \right| \\
& \leq \mu \|\nabla u\|_{L_t^p \ell_x^p} \|v\|_{L_t^q \ell_x^q} \\
& \lesssim \mu L^2 \|\nabla u\|_{L_t^p \ell_x^p} \|\nabla^2 v\|_{L_t^q \ell_x^q} \\
& \stackrel{(282)}{\lesssim} \mu L^2 \|\nabla u\|_{L_t^p \ell_x^p} \|f\|_{L_t^q \ell_x^q}.
\end{aligned}$$

**Step 2.** The density argument.

Let  $f(t, x)$  be smooth and compactly supported in  $t$ ,  $L$ -periodic in  $x$ , and suppose that  $f$  has vanishing spatial average  $\sum_{x \in ([0, L] \cap \mathbb{Z})^d} f(t, x) = 0$ . We claim that for every  $\delta > 0$  and  $\mu > 0$  there exists a function  $v(t, x)$  that is smooth and compactly supported in  $t$ , satisfying the mean-free condition (279) and

$$\|f - (\mu - \partial_t + \nabla^* \nabla) v\|_{L_t^q \ell_x^q} \leq \delta.$$

For the argument recall that  $G_\mu(t, x) := \exp(-\mu t) G(t, x)$  denotes the parabolic Green's function (on  $\mathbb{Z}^d$ ) associated with the operator  $\mu + \partial_t + \nabla^* \nabla$  (see (242)). Now, consider the function  $\hat{v}(t, x)$  given by

$$\hat{v}(t, x) := \int_{-\infty}^t \sum_{y \in \mathbb{Z}^d} G_\mu(t - s, x - y) \hat{f}(s, y) \, ds,$$

where  $\hat{f}(t, x) := f(-t, x)$ . By construction  $\hat{v}$  is smooth in  $t$ , vanishes for sufficiently negative  $t$  and satisfies the mean-free condition (279). Since  $\hat{f}(t, x)$  is  $L$ -periodic in  $x$ ,  $\hat{v}(t, x)$  is  $L$ -periodic in  $x$  as well, and we have

$$(\mu + \partial_t + \nabla^* \nabla) \hat{v}(t, x) = \hat{f}(t, x).$$

In order to obtain a compact support in time, we introduce a cut-off for  $t \geq 2T \gg 1$ . Let  $\eta(t)$  be a smooth function with

$$\eta(t) = \begin{cases} 1 & \text{for } t \leq 1, \\ 0 & \text{for } t \geq 2, \end{cases} \quad \text{and} \quad |\partial_t \eta| \lesssim 1.$$



For  $T \gg 1$  set  $\eta_T(t) := \eta(\frac{t}{T})$ . Since  $\hat{f}$  is compactly supported in time, we have for  $T \gg 1$

$$(\mu + \partial_t + \nabla^* \nabla)(\eta_T \hat{v}) = \hat{f} + \hat{v} \partial_t \eta_T.$$

Hence, the statement follows with  $v(t, x) := \hat{v}(-t, x) \eta_T(-t)$  for some  $T \gg 1$ , provided we can show that

$$(285) \quad \lim_{T \uparrow \infty} \|\hat{v} \partial_t \eta_T\|_{L_t^q \ell_x^q} = 0.$$

For the proof of (285) choose  $T_0$  sufficiently large such that  $\hat{f}(t, \cdot) = 0$  for all  $|t| \geq T_0$ . Note that

$$(\mu + \partial_t + \nabla^* \nabla) \hat{v}(t, x) = 0 \quad \text{for } t \geq T_0 \text{ and } x \in \mathbb{Z}^d.$$

Multiplication with  $\hat{v}$  and summation in  $x$  shows that  $\frac{d}{dt} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(t, x))^2 \leq 0$  for all  $t \geq T_0$ , and thus

$$(286) \quad \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(t, x))^2 \leq \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(T_0, x))^2 \quad \text{for all } t \geq T_0.$$

Now we estimate (285) for  $T \geq T_0$ . We first consider the case  $q \geq 2$ . Since  $|\partial_t \eta_T| \lesssim \frac{1}{T}$  and  $\partial_t \eta_T$  is supported in  $[T, 2T]$  we get with the discrete  $\ell^q$ - $\ell^2$ -estimate

$$\begin{aligned} \|\hat{v} \partial_t \eta_T\|_{L_t^q \ell_x^q}^q &\lesssim \frac{1}{T^q} \int_T^{2T} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\hat{v}(t, x)|^q dt \\ &\leq \frac{1}{T^q} \int_T^{2T} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(t, x))^2 \right)^{\frac{q}{2}} dt \\ &\stackrel{(286)}{\leq} \frac{1}{T^q} \int_T^{2T} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(T_0, x))^2 \right)^{\frac{q}{2}} dt \\ &= T^{q-1} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(T_0, x))^2 \right)^{\frac{q}{2}}. \end{aligned}$$

The r. h. s. converges to 0 as  $T \uparrow \infty$ , which completes the argument for  $q \geq 2$ . The argument for  $1 < q < 2$  is similar – the only difference is that we replace the discrete  $\ell^q$ - $\ell^2$ -estimate by Hölder's inequality with exponents  $(\frac{2}{q}, \frac{2}{2-q})$ :

$$\frac{1}{T^q} \int_T^{2T} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} |\hat{v}(t, x)|^q dt \leq \frac{L^{\frac{d(2-q)}{2}}}{T^q} \int_T^{2T} \left( \sum_{x \in ([0, L] \cap \mathbb{Z})^d} (\hat{v}(t, x))^2 \right)^{\frac{q}{2}} dt.$$

### Step 3. Conclusion.

By a standard density argument, it suffices to prove (278) for functions  $f(t, x)$  that are smooth and compactly supported in  $t$  and  $L$ -periodic in  $x$ . Since  $\nabla_j u$  is the gradient of a  $L$ -periodic function, we have  $\sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u = 0$ , so that it suffices to consider test functions  $f(t, x)$  that have vanishing spatial average in addition. By Step 2 we can find for  $\delta, \mu > 0$  test functions  $f_{\delta, \mu}$  such that

$$(287) \quad \|f - f_{\delta, \mu}\|_{L_t^q \ell_x^q} \leq \delta$$

and  $f_{\delta,\mu}$  is of the form  $(\mu - \partial_t + \nabla^* \nabla)v_{\delta,\mu}$  where  $v_{\delta,\mu}(t, x)$  satisfies the assumptions of Step 1. Hence, from (287) and (281) we deduce that

$$\begin{aligned} & \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u f dt \right| \\ & \stackrel{(287)}{\leq} \left| \int_{-\infty}^{\infty} \sum_{x \in ([0, L] \cap \mathbb{Z})^d} \nabla_j u f_{\delta,\mu} dt \right| + \delta \|\nabla u\|_{L_t^p \ell_x^p} \\ & \stackrel{(281)}{\lesssim} \|g\|_{L_t^p \ell_x^p} \|f_{\delta,\mu}\|_{L_t^q \ell_x^q} + \mu L^2 \|f_{\delta,\mu}\|_{L_t^q \ell_x^q} \|\nabla u\|_{L_t^p \ell_x^p} + \delta \|\nabla u\|_{L_t^p \ell_x^p}. \end{aligned}$$

Since  $\|f_{\delta,\mu}\|_{L_t^q \ell_x^q} \leq \|f\|_{L_t^q \ell_x^q} + \delta$  by (287), we may pass to the limits  $\mu \downarrow 0$  and  $\delta \downarrow 0$  and obtain (278).  $\square$

Now we turn to the proof of Lemma 30 which is the core of the matter.

*Proof of Lemma 30.* We follow the same strategy as for the proof of Lemma 28. Let us anticipate that w. r. t. the proof of Lemma 28 there is one additional step (cf. Step 8 below). We introduce a Fourier transform  $\mathbb{R} \times \mathbb{R}^d \ni (\tau, k) \mapsto (\mathcal{F}_{dis} g)(\tau, k)$  in time and space of a (semi-)discrete function  $\mathbb{R} \times \mathbb{Z}^d \ni (t, x) \mapsto g(t, x)$  via

$$(\mathcal{F}_{dis} g)(\tau, k) := (2\pi)^{-\frac{d+1}{2}} \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} g(t, x) \exp(-i(\tau t + k \cdot x)) dt.$$

We also introduce the inverse Fourier transform  $\mathbb{R} \times \mathbb{R}^d \ni (t, x) \mapsto (\mathcal{F}^{-1} F)(t, x)$  of a function  $\mathbb{R} \times \mathbb{R}^d \ni (\tau, k) \mapsto F(\tau, k)$  via

$$(\mathcal{F}^{-1} F)(x) = (2\pi)^{-\frac{d+1}{2}} \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} F(\tau, k) \exp(i(\tau t + k \cdot x)) dk d\tau.$$

We note that  $\mathbb{R} \times \mathbb{R}^d \ni (t, x) \mapsto \mathcal{F}^{-1}(\chi \mathcal{F}_{dis} g)(t, x)$  (where  $\chi(k)$  is defined as in Lemma 28) provides an extension of  $\mathbb{R} \times \mathbb{Z}^d \ni (t, x) \mapsto g(t, x)$  on  $\mathbb{R} \times \mathbb{R}^d$  since

$$(288) \quad g(t, x) = (\mathcal{F}^{-1}(\chi \mathcal{F}_{dis} g))(t, x) \quad \text{for } (t, x) \in \mathbb{R} \times \mathbb{Z}^d.$$

Note that the Fourier transform of  $u$  is well-defined: Indeed, as a solution to (276) (with compactly supported r. h. s.  $f$  and positive  $\mu$ ),  $u$  belongs to  $L^1(\mathbb{R}, \ell^1(\mathbb{Z}^d))$  as can be seen by appealing to the Green's function representation, cf. Lemma 2, and the fact that the discrete parabolic Green's function satisfies  $G_\mu \in L^1(\mathbb{R}, \ell^1(\mathbb{Z}^d))$ , see Step 3 of the proof of Lemma 28.

**Step 1.** Formulation of the result in terms of a Fourier multiplier  $\mathbf{m}_\mu$ .

We claim that

$$\left( \int \sum_{x \in \mathbb{Z}^d} |\nabla_j^* \nabla_\ell u(t, x)|^p dt \right)^{\frac{1}{p}} \lesssim \left( \int \sum_{x \in \mathbb{Z}^d} |f(t, x)|^p dt \right)^{\frac{1}{p}}$$

is equivalent to

$$(289) \quad \left( \int \sum_{x \in \mathbb{Z}^d} |(\mathcal{F}^{-1}(\chi \mathbf{m}_\mu \mathcal{F}_{dis} f))(t, x)|^p dt \right)^{\frac{1}{p}} \lesssim \left( \int \sum_{x \in \mathbb{Z}^d} |f(x)|^p dt \right)^{\frac{1}{p}},$$

where the meromorphic multiplier  $\mathbf{m}(\tau, k)$  is given by

$$(290) \quad \mathbf{m}_\mu(\tau, k) := \frac{(\exp(ik_j) - 1)(\exp(-ik_\ell) - 1)}{\mu + i\tau + \sum_{n=1}^d |\exp(ik_n) - 1|^2}, \quad (\tau, k) \in \mathbb{R} \times \mathbb{R}^d.$$

Next to the ingredients of Step 1 of Lemma 28, this relies on the identity  $\mathcal{F}_{dis}\partial_t u = i\tau\mathcal{F}_{dis}u$ .

**Step 2.** Representation of the difference between the discrete Fourier multiplier  $\mathbf{m}_\mu$  and the continuum Fourier multiplier  $\mathbf{m}_\mu^{cont} := \frac{k_j k_\ell}{\mu + i\tau + |k|^2}$ .

We claim that

$$\begin{aligned}\mathbf{m}_\mu - \mathbf{m}_\mu^{cont} &= \mathbf{m}_\mu (\mathbf{m}^* + (1 - \mathfrak{R}_\mu^{cont})\mathbf{m}^{**}) \\ &= (1 - \mathfrak{R}_\mu)\mathbf{m}_0 (\mathbf{m}^* + (1 - \mathfrak{R}_\mu^{cont})\mathbf{m}^{**})\end{aligned}$$

where  $\mathbf{m}_0$  denotes the (semi-)discrete multiplier associated with  $\mu = 0$ , i. e.

$$\mathbf{m}_0(\tau, k) := \frac{(\exp(ik_j) - 1)(\exp(-ik_\ell) - 1)}{i\tau + \sum_{n=1}^d |\exp(ik_n) - 1|^2},$$

where  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$  account for the deviation of the difference and continuum multipliers, i. e.

$$\begin{aligned}\mathbf{m}^*(k) &:= 1 - \frac{1}{h(k_j)h(-k_\ell)}, \\ \mathbf{m}^{**}(\tau, k) &:= \frac{1}{i\tau + |k|^2} \sum_{n=1}^d k_n^2 \left( \frac{|h(k_n)|^2 - 1}{h(k_j)h(-k_\ell)} \right),\end{aligned}$$

and where  $\mathfrak{R}_\mu$  and  $\mathfrak{R}_\mu^{cont}$  account for the difference between  $\mu > 0$  and  $\mu = 0$  in the discrete and continuum case, respectively, i. e.

$$\begin{aligned}\mathfrak{R}_\mu(\tau, k) &:= \frac{\mu}{\mu + i\tau + \sum_{n=1}^d |\exp(ik_n) - 1|^2}, \\ \mathfrak{R}_\mu^{cont}(\tau, k) &:= \frac{\mu}{\mu + i\tau + |k|^2},\end{aligned}$$

and  $h$  is defined in (248). The identity follows by a calculation similar to the one in Step 2 in the proof of Lemma 28. Note that  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$  are linear combination of terms of the two types

$$(291) \quad ik_{n'}\varphi(k) \quad \text{and} \quad k_n^2 ik_{n'} \frac{1}{i\tau + |k|^2} \varphi(k),$$

respectively, where  $\varphi(k)$  is smooth for  $k \in (-2\pi, 2\pi)^d$  and  $n, n' \in \{1, \dots, d\}$ .

**Step 3.**  $L^1$ -estimate of  $\mathfrak{R}_\mu^{cont}$  and  $\mathfrak{R}_\mu$  uniform in  $\mu$ .

Since  $\mathfrak{R}_\mu^{cont}$  and  $\mathfrak{R}_\mu$  are the continuum and discrete Fourier multipliers associated with  $\mu(\mu + \partial_t - \Delta)^{-1}$  and  $\mu(\mu + \partial_t + \nabla^* \nabla)^{-1}$ , we have

$$\begin{aligned}\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont})(t, x) &= \mu G_\mu^{cont}(t, x) \quad \text{for a.e. } x \in \mathbb{R}^d, t > 0, \\ \mathcal{F}^{-1}(\mathfrak{R}_\mu \chi)(t, x) &= \mu G_\mu(t, x) \quad \text{for all } x \in \mathbb{Z}^d, t > 0.\end{aligned}$$

and  $\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont})(t, x) = \mathcal{F}^{-1}(\mathfrak{R}_\mu)(t, x) = 0$  for  $t \leq 0$ . Hence, similar to Step 3 in the proof of Lemma 28:

$$(292) \quad \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont})| dx dt = \mu \int_0^{\infty} \int_{\mathbb{R}^d} G_\mu^{cont}(t, x) dx dt \stackrel{(244)}{=} 1,$$

$$(293) \quad \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} |\mathcal{F}^{-1}(\mathfrak{R}_\mu)| dt = \mu \int_0^{\infty} \sum_{x \in \mathbb{Z}^d} G_\mu(t, x) dt \stackrel{(243)}{=} 1.$$

**Step 4.** Treatment of high wave numbers.

This is the main new element w. r. t. the elliptic case in Lemma 28. We claim that for any two fields  $v : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \times \mathbb{Z}^d \rightarrow \mathbb{R}$  that are related by

$$(294) \quad \partial_t v + \nabla^* \nabla v = g,$$

decaying sufficiently fast and banded in the sense of

$$(295) \quad (\mathcal{F}_{dis}v)(\tau, k) = (\mathcal{F}_{dis}g)(\tau, k) = 0 \quad \text{for } k \in \left(-\frac{2\pi}{L}, \frac{2\pi}{L}\right)^d$$

for some fixed  $L$  (chosen below in Step 8), we have

$$\|\nabla_j^* \nabla \ell v\|_{L_t^p \ell_x^p} \lesssim \|g\|_{L_t^p \ell_x^p}.$$

It will be more convenient to establish that result in the form of  $\|\nabla \tilde{v}\|_{L_t^p \ell_x^p} \lesssim \|g\|_{L_t^p \ell_x^p}$  where  $\tilde{v} := \nabla_j^* v$  for some fixed  $j \in \{1, \dots, d\}$ . Note that  $\tilde{v}$  decays sufficiently fast, is banded and related to  $g$  by

$$(296) \quad \partial_t \tilde{v} + \nabla^* \nabla \tilde{v} = \nabla_j^* g.$$

We multiply equation (296) with  $\tilde{v}|\tilde{v}|^{p-2}$ , sum over  $x \in \mathbb{Z}^d$  and integrate the r. h. s. by parts:

$$\frac{d}{dt} \sum_{x \in \mathbb{Z}^d} \frac{1}{p} |\tilde{v}|^p + \sum_{x \in \mathbb{Z}^d} \nabla \tilde{v} \cdot \nabla (\tilde{v}|\tilde{v}|^{p-2}) = \sum_{x \in \mathbb{Z}^d} g \nabla_j (\tilde{v}|\tilde{v}|^{p-2}).$$

Integrating in  $t$  yields

$$\int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} \nabla \tilde{v} \cdot \nabla (\tilde{v}|\tilde{v}|^{p-2}) dt = \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} g \nabla_j (\tilde{v}|\tilde{v}|^{p-2}) dt.$$

On the one hand, we have (by a calculation using the discrete proxy (94) for the continuum chain rule)

$$\begin{aligned} & \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} \nabla \tilde{v} \cdot \nabla (\tilde{v}|\tilde{v}|^{p-2}) dt \\ & \gtrsim \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} \sum_{n=1}^d (|\tilde{v}|^{p-2}(x + e_j) + |\tilde{v}|^{p-2}(x)) (\tilde{v}(x + e_n) - \tilde{v}(x))^2 dt \\ & \gtrsim \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} \sum_{n=1}^d |\tilde{v}(x + e_n) - \tilde{v}(x)|^p dt \\ & \gtrsim \|\nabla \tilde{v}\|_{L_t^p \ell_x^p}^p. \end{aligned}$$

On the other hand, we have by Cauchy-Schwarz' inequality and discreteness

$$\begin{aligned} \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} g \nabla_j (\tilde{v}|\tilde{v}|^{p-2}) dt & \leq \|g\|_{L_t^p \ell_x^p} \|\nabla_j (\tilde{v}|\tilde{v}|^{p-2})\|_{L_t^{\frac{p}{p-1}} \ell_x^{\frac{p}{p-1}}} \\ & \leq \|g\|_{L_t^p \ell_x^p} 2 \|\tilde{v}|\tilde{v}|^{p-2}\|_{L_t^{\frac{p}{p-1}} \ell_x^{\frac{p}{p-1}}} \\ & = \|g\|_{L_t^p \ell_x^p} 2 \|\tilde{v}\|_{L_t^p \ell_x^p}^{p-1}. \end{aligned}$$

It thus remains to argue that the bandedness (295) implies

$$\|\tilde{v}\|_{L_t^p \ell_x^p} \lesssim \|\nabla \tilde{v}\|_{L_t^p \ell_x^p}.$$

For that purpose, we may neglect the  $t$ -dependence: Below,  $\mathcal{F}_{dis,x}$  and  $\mathcal{F}_x^{-1}$  denote the *spatial* Fourier transform and the inverse *spatial* Fourier transform as defined in (239), and we write  $*_x$  and  $*_{dis,x}$  for the continuum and discrete spatial convolution in  $x$ . We show that for a discrete function  $\tilde{v} : \mathbb{Z}^d \rightarrow \mathbb{R}$  with

$$(\mathcal{F}_{dis,x} \tilde{v})(k) = 0 \quad \text{for } k \in \left(-\frac{2\pi}{L}, \frac{2\pi}{L}\right)^d$$

we have

$$\|\tilde{v}\|_{\ell^p} \lesssim \|\nabla \tilde{v}\|_{\ell^p};$$

or, equivalently, in Fourier space: For any  $[-\pi, \pi]^d \ni k \mapsto F(k) \in \mathbb{C}$  with

$$(297) \quad F(k) = 0 \quad \text{for } k \in \left(-\frac{2\pi}{L}, \frac{2\pi}{L}\right)^d,$$

we have

$$\|\mathcal{F}_x^{-1}(\chi F)\|_{\ell^p} \lesssim \sum_{n=1}^d \|\mathcal{F}_x^{-1}((\exp(ik_n) - 1)\chi F)\|_{\ell^p}.$$

In fact, we even have the stronger statement

$$\|\mathcal{F}_x^{-1}(\chi F)\|_{\ell^p} \lesssim \|\mathcal{F}_x^{-1}((\exp(ik_1) - 1)\chi F)\|_{\ell^p}.$$

This can be seen as follows: Because of (297) and the support properties of  $\eta_L(k)$ , cf. (255), we may write  $\chi F = \tilde{m}(\exp(ik_1) - 1)\chi F$  with a multiplier  $\tilde{m}(k) := \frac{1 - \eta_{2L}(k)}{\exp(ik_1) - 1}$ . Hence  $\mathcal{F}_x^{-1}(\chi F) = (2\pi)^{\frac{d}{2}} \mathcal{F}_x^{-1}(\chi \tilde{m}) *_{dis, x} \mathcal{F}_x^{-1}((\exp(ik_1) - 1)\chi F)$  on  $\mathbb{Z}^d$  which yields

$$\|\mathcal{F}_x^{-1}(\chi F)\|_{\ell^p} \leq (2\pi)^{\frac{d}{2}} \|\mathcal{F}_x^{-1}(\chi \tilde{m})\|_{\ell^1} \|\mathcal{F}_x^{-1}((\exp(ik_1) - 1)\chi F)\|_{\ell^p}.$$

Since  $\chi \tilde{m}$  is smooth and  $2\pi$ -periodic in  $k \in [-\pi, \pi]^d$ , we have  $\|\mathcal{F}_x^{-1}(\chi \tilde{m})\|_{\ell^1} < \infty$ , cf. Step 8 in the proof of Lemma 28.

**Step 5.**  $L^1$ -estimate of the truncated multipliers  $\mathbf{m}^*$  and  $\mathbf{m}^{**}$ .

Let  $\eta_L$ ,  $K_L$  and  $\hat{K}_L$  be as in Step 4 in the proof of Lemma 28. We claim that for  $L \geq 1$

$$\begin{aligned} \int |\mathcal{F}_x^{-1}(\mathbf{m}^* \eta_L)| dx &\lesssim L^{-1}, \\ \int \int |\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L)| dx dt &\lesssim L^{-1}, \end{aligned}$$

where here and in the sequel,  $\int$  stands for  $\int_{\mathbb{R}^d}$  and  $\int \int$  stands for  $\int_{-\infty}^{\infty} \int_{\mathbb{R}^d}$ . The first estimate is identical to (258a). The second estimate reduces to

$$(298) \quad \int \int |\mathcal{F}^{-1}(k_n^2 i k_{n'} \frac{1}{i\tau + |k|^2} \varphi \eta_L)| dx dt \lesssim L^{-1}.$$

As a new element w. r. t. Step 5 in the proof of Lemma 28, we need the Fourier transform of the heat kernel, that is  $(\mathcal{F}^{-1} \frac{1}{i\tau + |k|^2})(t, x) = G^{cont}(t, x) = (\pi t)^{-\frac{d}{2}} \exp(-\frac{|x|^2}{4t})$  (for  $t > 0$  and  $= 0$  for  $t \leq 0$ ). With it, we obtain the representation

$$(299) \quad \mathcal{F}^{-1}(k_n^2 i k_{n'} \frac{1}{i\tau + |k|^2} \varphi \eta_L) = (2\pi)^{\frac{d}{2}} (\pi t)^{-\frac{d}{2}} \partial_n^2 \partial_{n'} \left( \exp(-\frac{|x|^2}{4t}) *_x \mathcal{F}_x^{-1}(\varphi \eta_L) \right).$$

We now deduce (298) from that representation and the estimate (257). We split (299) into

$$\begin{aligned} \int_0^{L^2} t^{-\frac{d}{2}} \int \left| \exp(-\frac{|x|^2}{4t}) *_x (\partial_n^2 \partial_{n'} K_L) \right| dx dt &\lesssim L^{-1}, \\ \int_{L^2}^{\infty} t^{-\frac{d}{2}} \int \left| (\partial_n^2 \partial_{n'} \exp(-\frac{|x|^2}{4t})) *_x K_L \right| dx dt &\lesssim L^{-1}. \end{aligned}$$

These two estimates are easily seen to be true:

$$\begin{aligned}
& \int_0^{L^2} t^{-\frac{d}{2}} \int \left| \exp\left(-\frac{|x|^2}{4t}\right) *_x (\partial_n^2 \partial_{n'} K_L) \right| dx dt \\
& \leq \int_0^{L^2} t^{-\frac{d}{2}} \int |\exp\left(-\frac{|x|^2}{4t}\right)| dx dt \int |\partial_n^2 \partial_{n'} K_L| dx \\
& \stackrel{x=\sqrt{4ty}, x=L\hat{x}}{=} 2^d \int_0^{L^2} dt \int \exp(-|y|^2) dy L^{-3} \int |\partial_n^2 \partial_{n'} \hat{K}_L| d\hat{x} \\
& \stackrel{(257)}{\lesssim} L^{-1}
\end{aligned}$$

and

$$\begin{aligned}
& \int_{L^2}^\infty t^{-\frac{d}{2}} \int \left| (\partial_n^2 \partial_{n'} \exp\left(-\frac{|x|^2}{4t}\right)) *_x K_L \right| dx dt \\
& \leq \int_{L^2}^\infty t^{-\frac{d}{2}} \int |\partial_n^2 \partial_{n'} \exp\left(-\frac{|x|^2}{4t}\right)| dx dt \int |K_L| dx \\
& \stackrel{x=\sqrt{4ty}, x=L\hat{x}}{\lesssim} \int_{L^2}^\infty t^{-\frac{3}{2}} dt \int |\partial_{y_n}^2 \partial_{y_{n'}} \exp(-|y|^2)| dy \int |\hat{K}_L| d\hat{x} \\
& \stackrel{(257)}{\lesssim} L^{-1}.
\end{aligned}$$

**Step 6.**  $L_{tx}^p$ -boundedness of the multiplier  $\mathbf{m}_\mu$  on functions with band restriction.

We claim that there exists an  $L \gg 1$  such that for any field  $F(\tau, k) \in \mathbb{C}$  that is band-restricted in the sense that for all  $\tau \in \mathbb{R}$ ,

$$(300) \quad F(\tau, k) \text{ is supported in } k \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right]^d,$$

we have

$$(301) \quad \|\mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p} \lesssim \|\mathcal{F}^{-1}F\|_{L_{tx}^p},$$

where  $\|g\|_{L_{tx}^p} := \left(\int_{-\infty}^\infty \int_{\mathbb{R}^d} |g(t, x)|^p dx dt\right)^{\frac{1}{p}}$ .

From the continuum counterpart of (301) with continuum Fourier multiplier  $\mathbf{m}_\mu^{cont}(\tau, k) = \frac{k_j k_\ell}{\mu + i\tau + |k|^2}$

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu^{cont} F)\|_{L_{tx}^p} \lesssim \|\mathcal{F}^{-1}F\|_{L_{tx}^p},$$

see Theorem 5 (b), and the triangle inequality, we see that it is enough to show that the band restriction (300) implies

$$(302) \quad \|\mathcal{F}^{-1}((\mathbf{m}_\mu - \mathbf{m}_\mu^{cont})F)\|_{L_{tx}^p} \lesssim L^{-1} \|\mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p}.$$

We now address (302). Thanks to (300), the definition (255) of the cut-off  $\eta_L$  and the decomposition from Step 2, we have

$$(\mathbf{m}_\mu - \mathbf{m}_\mu^{cont})F = (\mathbf{m}^* + (1 - \mathfrak{R}_\mu^{cont})\mathbf{m}^{**})(\eta_L \mathbf{m}_\mu F).$$

By the triangle inequality, the convolution estimate and (292) we then get

$$\begin{aligned}
(303) \quad & \|\mathcal{F}^{-1}((\mathbf{m}_\mu - \mathbf{m}_\mu^{cont})F)\|_{L_{tx}^p} \\
& \leq \|\mathcal{F}^{-1}(\mathbf{m}^* \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} + \|\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} \\
& \quad + \|\mathcal{F}^{-1}(\mathfrak{R}_\mu^{cont} \mathbf{m}^{**} \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} \\
& \stackrel{(292)}{\lesssim} \|\mathcal{F}^{-1}(\mathbf{m}^* \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} + \|\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p}.
\end{aligned}$$

For the first term on the r. h. s. we appeal to the spatial convolution estimate

$$\begin{aligned}\|\mathcal{F}^{-1}(\mathbf{m}^* \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} &= (2\pi)^{\frac{d}{2}} \|\mathcal{F}_x^{-1}(\mathbf{m}^* \eta_L) *_x \mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p} \\ &\leq (2\pi)^{\frac{d}{2}} \|\mathcal{F}_x^{-1}(\mathbf{m}^* \eta_L)\|_{L_x^1} \|\mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p},\end{aligned}$$

while for the second term on the r. h. s. of (303) we appeal to the space-time convolution estimate

$$\begin{aligned}\|\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L \mathbf{m}_\mu F)\|_{L_{tx}^p} &= (2\pi)^{\frac{d+1}{2}} \|\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L) * \mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p} \\ &\leq (2\pi)^{\frac{d+1}{2}} \|\mathcal{F}^{-1}(\mathbf{m}^{**} \eta_L)\|_{L_{tx}^1} \|\mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_{tx}^p},\end{aligned}$$

so that (302) follows from the estimates derived in Step 5.

**Step 7.**  $L_t^p \ell_x^p$ -boundedness of the multiplier  $\mathbf{m}_\mu$  on functions with band restriction.

We claim that there exists an  $L \gg 1$  such that for any field  $F(k) \in \mathbb{C}$  that is band-restricted in the sense of (300), we have

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu F)\|_{L_t^p \ell_x^p} \lesssim \|\mathcal{F}^{-1} F\|_{L_t^p \ell_x^p},$$

where  $\|g\|_{L_t^p \ell_x^p} := \left( \int_{-\infty}^{\infty} \sum_{x \in \mathbb{Z}^d} |g(x)|^p dt \right)^{\frac{1}{p}}$ .

Step 7 follows from Step 6 via the equivalence of discrete and continuum norms, that is,

$$\|g\|_{L_t^p \ell_x^p} \sim \|g\|_{L_t^p \ell_x^p}, \quad \text{where } g := \mathcal{F}^{-1} F.$$

In fact, we even have

$$\|g(t, \cdot)\|_{\ell_x^p} \sim \|g(t, \cdot)\|_{\ell_x^p} \quad \text{for all } t.$$

This follows immediately from (270), cf. Step 7 in the proof of Lemma 28, since (300) implies that  $(\mathcal{F}_x g(t, \cdot))(k)$  is supported in  $[-\frac{\pi}{L}, \frac{\pi}{L}]^d$  for all  $t$ .

**Step 8.** Conclusion.

In view of Step 1, it is enough to prove that

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} \lesssim \|f\|_{L_t^p \ell_x^p}.$$

As in Step 8 of the proof of Lemma 28, in order to pass from Step 7 to Step 8, we shall show that

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} + \|\mathcal{F}^{-1}((1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} \lesssim \|f\|_{L_t^p \ell_x^p}.$$

We split this statement into

$$(304) \quad \|\mathcal{F}^{-1}(\mathbf{m}_\mu (1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} \lesssim \|\mathcal{F}^{-1}((1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p},$$

$$(305) \quad \|\mathcal{F}^{-1}((1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} \lesssim \|f\|_{L_t^p \ell_x^p}.$$

Since  $1 - \eta_L$  is just a function of  $k$  and not  $\tau$ , the second estimate (305) can be established as (275), that is, based on the identity

$$\begin{aligned}\mathcal{F}^{-1}((1 - \eta_L) \chi \mathcal{F}_{dis} f) &= \mathcal{F}^{-1}(\chi(1 - \eta_L) \chi \mathcal{F}_{dis} f) \\ &= (2\pi)^{\frac{d}{2}} \mathcal{F}_x^{-1}(\chi(1 - \eta_L)) *_x \mathcal{F}^{-1}(\chi \mathcal{F}_{dis} f) \\ &= (2\pi)^{\frac{d}{2}} \mathcal{F}_x^{-1}(\chi(1 - \eta_L)) *_x \mathcal{F}^{-1} f \quad \text{on } \mathbb{R} \times \mathbb{Z}^d\end{aligned}$$

and thus the estimate

$$\|\mathcal{F}^{-1}((1 - \eta_L) \chi \mathcal{F}_{dis} f)\|_{L_t^p \ell_x^p} \leq (2\pi)^{\frac{d}{2}} \|\mathcal{F}_x^{-1}(\chi(1 - \eta_L))\|_{\ell_x^1} \|f\|_{L_t^p \ell_x^p}.$$

However, we need a new argument for (304). We first note that in view of the property (255) of the cut-off function  $\eta_L$ , (304) can be rephrased as follows: For any field  $F(\tau, k) \in \mathbb{C}$  (playing the role of  $(1 - \eta_L) \mathcal{F}_{dis} g$ ) that is banded in the sense of

$$F(\tau, k) = 0 \quad \text{for } k \in \left(-\frac{2\pi}{L}, \frac{2\pi}{L}\right)^d,$$

(that is, it is the discrete Fourier transform of a function  $g(t, x)$  that does not contain wave length  $\geq 2L$  — that is  $\gtrsim 1$ , since  $L$  is fixed) we have

$$\|\mathcal{F}^{-1}(\mathbf{m}_\mu \chi F)\|_{L_t^p \ell_x^p} \lesssim \|\mathcal{F}^{-1}(\chi F)\|_{L_t^p \ell_x^p}.$$

Since  $\mathbf{m}_\mu = (1 - \mathfrak{R}_\mu)\mathbf{m}_0$ , and  $\|\mathfrak{R}_\mu\|_{L_t^1 \ell_x^1} \leq 1$  by Step 3, it even suffices to show that

$$\|\mathcal{F}^{-1}(\mathbf{m}_0 \chi F)\|_{L_t^p \ell_x^p} \lesssim \|\mathcal{F}^{-1}(\chi F)\|_{L_t^p \ell_x^p}.$$

Since by Step 1  $v := \mathcal{F}^{-1}(\mathbf{m}_0 \chi F)$  is related with  $g := \mathcal{F}^{-1}(\chi F)$  via (294), the desired estimate follows from Step 4.  $\square$

## Acknowledgement

Felix Otto wants to acknowledge the hospitality of the University of Paris-Sud (Orsay) – most of the presented material was covered by a Hadamard lecture Felix Otto gave in 2012 at that institution, which in turn was an extended version of a minitutorial at the SIAM PDE conference of 2011.



## References

- [1] M. Akcoglu and U. Krengel. Ergodic theorems for superadditive processes. *Journal für die reine und angewandte Mathematik* 323:53–67, 1981.
- [2] D. Bakry, T. Coulhon, M. Ledoux and L. Saloff-Coste. Sobolev inequalities in disguise. *Indiana Univ. Math. J.* 44(4):1032–1074, 1995.
- [3] A. Bensoussan, J.-L. Lions and G. Papanicolaou. Asymptotic analysis for periodic structures. *Studies in Mathematics and its Applications, Vol. 5, North-Holland Publishing Co.*, 1978.
- [4] G. D. Birkhoff. Proof of the ergodic theorem. *Proc Natl Acad Sci USA* 17(12):656–660, 1931.
- [5] M. Bondesson. Interior a priori estimates in discrete  $L_p$  norms for solutions of parabolic and elliptic difference equations. *Ann. Mat. Pura Appl., IV. Ser.* 95:1-43, 1973.
- [6] J. G. Conlon and A. Naddaf. Greens Functions for Elliptic and Parabolic Equations with Random Coefficients- *New York J. Math.*, 6:153225, 2000.
- [7] G. Dal Maso. An introduction to  $\Gamma$ -convergence. *Progress in Nonlinear Differential Equations and their Applications, 8*, 1993.
- [8] T. Delmotte. Estimations pour les chaînes de Markov réversibles. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1053–1058, 1997.
- [9] T. Delmotte. Inégalité de Harnack elliptique sur les graphes. *Colloq. Math.*, 72(1):19–37, 1997.
- [10] R. Denk, M. Hieber and J. Prüss, Jan.  $\mathcal{R}$ -boundedness, Fourier multipliers and problems of elliptic and parabolic type. *Mem. Am. Math. Soc.*, Vol. 788, 2003
- [11] E. B. Fabes and D. W. Stroock. A new proof of Moser’s parabolic harnack inequality using the old ideas of Nash. *Archive for Rational Mechanics and Analysis*, 96(4):327–338, 1986.
- [12] G. Giacomin, S. Olla and H. Spohn. Equilibrium Fluctuations for  $\nabla\varphi$  Interface Model. *Ann. Probab.* 29(3):1138–1172, 2001.
- [13] M. Geissert. Discrete Maximal  $L_p$  Regularity for Finite Element Operators. *SIAM Journal on Numerical Analysis*, 44(2):677–698, 2005
- [14] M. Geissert. Maximal  $L_p$  regularity for parabolic difference equations. *Math. Nachr.*, 279(16):1522-2616, 2006
- [15] A. Gloria and J.-C. Mourrat. Spectral measure and approximation of homogenized coefficients. *Probab. Theory Relat. Fields*, 154(1), 287–326, 2012.
- [16] A. Gloria, S. Neukamm and F. Otto. Approximation of effective coefficients by periodization in stochastic homogenization. in preparation.
- [17] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3), 779–856, 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, 2012.
- [19] M. Grüter and K.-O. Widman. The Green function for uniformly elliptic equations. *Manuscripta Math.*, 37:303–342, 1982.
- [20] C. Kipnis and S.R.S. Varadhan. Central limit theorem for additive functional of reversible Markov processes and applications to simple exclusion. *Commun. Math. Phys.*, 104:1–19, 1986.

- [21] S.M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.
- [22] S.M. Kozlov. Averaging of difference schemes. *Math. USSR Sbornik*, 57(2):351–369, 1987.
- [23] S.M. Kozlov. The method of averaging and walks in inhomogeneous environments. *Russian Math. Surveys* 40 (2):73–145, 1985.
- [24] R. Künnemann. The diffusion limit for reversible jump processes on  $\mathbb{Z}^d$  with ergodic random bond conductivities. *Commun. Math. Phys.*, 90:27–68, 1983.
- [25] N. Meyers. An  $L^p$ -estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 17(3):189–206, 1963.
- [26] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(11), 294–327, 2011.
- [27] F. Murat and L. Tartar. H-Convergence, Cours Peccot, 1977. Reprinted in *Topics in the Mathematical Modeling of Composite Materials* (A. Cherkaev and R. Kohn eds.), Birkhuser, Boston, 1997, pp. 21-43.
- [28] Ali Naddaf, Thomas Spencer, On homogenization and scaling limits of some gradient perturbations of a massless free field, *Comm. Math Phys.* 183:55–84, 1997.
- [29] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. Preprint, 1998.
- [30] J. Nash Continuity of Solutions of Parabolic and Elliptic Equations *American Journal of Mathematics*, 80(4):931–954, 1958.
- [31] M. M. H. Pang, Heat kernels of graphs. *J. London Math. Soc. (2)*, 47(1):50–64, 1993
- [32] 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 *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.
- [33] V. Sidoravicius and A. S. Sznitman. Quenched invariance principles for walks on clusters of percolations or among random conductances. *Probability Theory and Related Fields*, 129(2):219–244, 2004.
- [34] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Annali della Scuola Normale Superiore di Pisa (3)*, 22(4):571–597, 1968.
- [35] A. A. Tempel'man. Ergodic theorems for general dynamical systems. *Trudy Moskov. Mat. Obšč.*, 26:95–132, 1972.
- [36] V.V. Yurinskii. *Vilnius Conference Abstracts*, Sept. 78, 54.
- [37] V.V. Yurinskii. Averaging of symmetric diffusion in random medium. *Sibirskii Matematicheskii Zhurnal*, 27(4):167–180, 1986.
- [38] T. Delmotte and J.-D. Deuschel, On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to  $\nabla\phi$  interface model, *Probab. Theory Relat. Fields* **133**, 358390 (2005).
- [39] J.-D. Deuschel, G. Giacomin and D. Ioffe, Large deviations and concentration properties for random interface models, *Probab. Theory Relat. Fields* **117**, 49111 (2000).

- [40] T. Funaki, Stochastic interface models, *Lectures on Probability Theory and Statistics*, Lecture Notes in Mathematics Volume 1869, 2005, pp 103-274.
- [41] B. Helffer and J. Sjöstrand, On the correlation for Kac-like models in the convex case, *J. Stat. Mech* **74**, 349-409 (1994).