

# TINY FLUCTUATIONS OF THE AVERAGING PROCESS AROUND ITS DEGENERATE STEADY STATE

FEDERICO SAU

*University of Trieste*

ABSTRACT. We analyze nonequilibrium fluctuations of the averaging process on  $\mathbb{T}_\varepsilon^d$ , a continuous degenerate Gibbs sampler running over the edges of the discrete  $d$ -dimensional torus. We show that, if we start from a smooth deterministic non-flat interface, recenter, blow-up by a non-standard CLT-scaling factor  $\theta_\varepsilon = \varepsilon^{-(d/2+1)}$ , and rescale diffusively, Gaussian fluctuations emerge in the limit  $\varepsilon \rightarrow 0$ . These fluctuations are purely dynamical, zero at times  $t = 0$  and  $t = \infty$ , and non-trivial for  $t \in (0, \infty)$ . We fully determine the correlation matrix of the limiting noise, non-diagonal as soon as  $d \geq 2$ . The main technical challenge in this stochastic homogenization procedure lies in a LLN for a weighted space-time average of squared discrete gradients. We accomplish this through a Poincaré inequality with respect to the underlying randomness of the edge updates, a tool from Malliavin calculus in Poisson space. This inequality, combined with sharp gradients' second moment estimates, yields quantitative variance bounds without prior knowledge of the limiting mean. Our method avoids higher (e.g., fourth) moment bounds, which seem inaccessible with the present techniques.

## CONTENTS

1. Introduction, model and main results	1
2. Discrete gradients	10
3. Proof of Theorem 2. Variance via Poincaré inequality	18
4. Proof of Theorem 2. Mean	24
5. Proof of Theorem 1	30
Appendix A. Proofs of Lemmas 2.2, 4.1 and 4.2	34
References	37

## 1. INTRODUCTION, MODEL AND MAIN RESULTS

The averaging process on a graph is a continuous-space Markov chain, which is commonly interpreted as an opinion dynamics, a distributed algorithm, or an interface moving through a randomized sequence of deterministic local updates (see, e.g., [BGPS06, AL12, Ald13, MSW22] and references therein). Its dynamics goes as follows. Attach i.i.d. Poisson clocks to edges, and assign real values to vertices; at the arrival times of these clocks, update the values with their average. As time runs, the averaging process converges to a flat

---

*E-mail address:* federico.sau@units.it.

*2020 Mathematics Subject Classification.* Primary 60K35, 35B27; secondary 60J27, 60H15, 60H07.

*Key words and phrases.* Interacting particle systems; scaling limits; stochastic homogenization; hydrodynamic limit; nonequilibrium fluctuations; averaging process; Malliavin calculus.

configuration, and one major problem in the field is that of quantifying the speed of convergence to its degenerate equilibrium in terms of characteristic features of the underlying graph [CDSZ22, QS23, CQS23].

In this paper, we examine scaling limits of the averaging process on the discrete  $d$ -dimensional torus  $\mathbb{T}_\varepsilon^d$ ,  $d \geq 1$ , subjected to a diffusive space-time rescaling. In this setting, the averaging process, say  $u_t^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , evolves by replacing, at rate  $\varepsilon^{-2}$  for each nearest neighbor pair  $x, y \in \mathbb{T}_\varepsilon^d$ , the values  $(u_{t-}^\varepsilon(x), u_{t-}^\varepsilon(y))$  with

$$(u_t^\varepsilon(x), u_t^\varepsilon(y)) := \left( \frac{1}{2} (u_{t-}^\varepsilon(x) + u_{t-}^\varepsilon(y)), \frac{1}{2} (u_{t-}^\varepsilon(x) + u_{t-}^\varepsilon(y)) \right).$$

Alternatively, in the language of stochastic homogenization,  $u_t^\varepsilon$  is the solution to the discrete parabolic problem  $\partial_t u_t^\varepsilon = \nabla_*^\varepsilon \cdot (a_\varepsilon(t, \cdot) \nabla^\varepsilon u_{t-}^\varepsilon)$  on  $\mathbb{T}_\varepsilon^d$ , in which the coefficient field  $a_\varepsilon = (a_\varepsilon(t, x))_{t \geq 0, x \in \mathbb{T}_\varepsilon^d}$  is random, time-dependent, and formally given by

$$a_\varepsilon(t, x) = \frac{\varepsilon^2}{2} \begin{pmatrix} dN_t^{\varepsilon, 1}(x) & & \\ & \ddots & \\ & & dN_t^{\varepsilon, d}(x) \end{pmatrix}, \quad t \geq 0, \quad x \in \mathbb{T}_\varepsilon^d. \quad (1.1)$$

Here,  $dN_t^{\varepsilon, i}(x)$  stands for the increments of the Poisson process of intensity  $\varepsilon^{-2}$  attached to the pair  $x, x + \varepsilon e_i \in \mathbb{T}_\varepsilon^d$ . This homogenization problem is then degenerate, in the sense that the coefficient field  $a_\varepsilon$  is a linear combination of Dirac deltas and, thus, does not satisfy any ellipticity conditions: there is no  $C > 0$  satisfying neither  $a_\varepsilon \geq C$ , nor  $a_\varepsilon \leq C$  (in the matrix sense).

The first ergodic theorems in stochastic homogenization for linear elliptic and parabolic problems date back to the seminal works of Kozlov [Koz79] and Papanicolau and Varadhan [PV81]. Since then, there has been an intense activity in considerably extending these qualitative results. In fact, even quantitative features of the solutions, such as regularity and fluctuations limit theorems, are by now well understood in the uniform ellipticity context (see, e.g., [AKM19, JO22], [GM16, DGO20], and references therein), while the only degenerate examples are limited to supercritical percolation clusters [AD18, Dar21], rigid inclusions [DG22] and log-normal coefficient fields [GQ24, CGQ24].

In this article, we move one step forward by investigating dynamical fluctuations of the random  $u_t^\varepsilon$  in this time-dependent degenerate setting (cf. (1.1)). This result is the content of Theorem 1, which could be regarded as the functional central limit theorem (FCLT) after the (quantitative) law of large numbers (LLNs) established in [Sau23]. As recently worked out in some examples of random walks in time-dependent random environment (see, e.g., [Rho08, BR18], all concerned with LLNs of random walks' probabilities), our FCLT also draws upon the fundamental idea that adding a mixing dynamics potentially tempers spatial degeneracy. In our case, this dynamics consists in having zero conductances which suddenly take the value  $+\infty$  at the occurrence of a Poisson mark. As in most degenerate instances, this mixing mechanisms plays a crucial role in recovering on large scales regularity estimates which would deterministically hold true in the uniform ellipticity setting.

Our approach is probabilistic, based on martingales, as developed by Holley and Stroock [HS78] and successfully exploited within the context of interacting particle systems (see, e.g., the monographs [DMP91, KL99]). Nonetheless, in contrast with most systems studied in that realm, the averaging process  $u_t^\varepsilon$  has no truly random ergodic states (they all consist of deterministic flat configurations). This lack of microscopic fluctuations at equilibrium prevents us to employ methods based on relative entropy, not having a clear notion of local equilibrium to compare the law of  $u_t^\varepsilon$  with (see, e.g., [JM18, JL23] and references therein).

We overcome this obstacle by combining the aforementioned martingale-based approach with two main ingredients, typical of infinite-dimensional stochastic analysis. On the one hand, we control a key quantity — squared discrete gradients of  $u_t^\varepsilon$  — by expressing its expected value as an infinite series of iterated integrals. On the other hand, inspired by a series of recent works in stochastic homogenization, e.g., [DO20, DG20a, DG20b], we leverage the Poisson nature of the updates by operating with tools from Malliavin calculus (in Poisson space). These are the two building blocks in the proof of our second main result, Theorem 2, a LLN for weighted space-time averages of squared discrete gradients. Theorem 2 is also an essential step in the characterization of the limiting Gaussian process from Theorem 1, and reveals the smallness of the fluctuations  $u_t^\varepsilon - \mathbb{E}^\varepsilon[u_t^\varepsilon]$ , as well as the non-diagonal correlation structure of the limiting driving noise (despite the form of the coefficient field  $a_\varepsilon$  in (1.1) with i.i.d. diagonal entries).

In conclusion, as originally proposed in [AL12, Ald13], the averaging process (together with a long list of companions “random averages” models, e.g., [FF98, BRAS06, CF11, Lan12, HH14, BB21, FMZ23], with either site or edge updates) is a natural nonequilibrium system of stochastic moving interfaces, whose quantitative features are still largely unexplored. These include, for instance, a regularity theory and worst-case mixing times on geometric settings other than the torus. The latter problem becomes particularly interesting in light of the works [CDSZ22, CQS23], which show that on the simplest examples of expander graphs the averaging process’ degenerate local update rule dramatically affects the timescales relevant to relaxation, if compared to the homogenized dynamics. We believe the combination of Markov chains’ and stochastic homogenization techniques that we develop here to be fruitful also in this context.

We now present the model and our main results. Before that, let us introduce some general notation. All throughout the article,  $\mathbb{T}^d := (\mathbb{R}/\mathbb{Z})^d$ , with  $d \geq 1$ , is the  $d$ -dimensional torus, while  $\mathbb{T}_\varepsilon^d$  denotes its lattice discretization with mesh size  $\varepsilon \in (0, 1)$ . For simplicity, we always assume  $\varepsilon^{-1} \in \mathbb{N}$ . Moreover, with a slight abuse of notation, we often identify the set  $\mathbb{T}_\varepsilon^d$  with the undirected graph obtained by connecting vertices of the form  $x$  and  $y = x \pm \varepsilon e_i \in \mathbb{T}_\varepsilon^d$  with an edge, where  $e_1, \dots, e_d \in \mathbb{R}^d$  denotes the canonical basis of  $\mathbb{R}^d$ . Hence, letting  $|\cdot|$  denote the usual Euclidean distance on  $\mathbb{R}^d$ , we will refer to sites  $x, y \in \mathbb{T}_\varepsilon^d$  satisfying  $|x - y| = \varepsilon$  as nearest neighbor vertices. Finally, we shall implicitly imply “holds true for all  $\varepsilon \in (0, 1)$  with  $\varepsilon^{-1} \in \mathbb{N}$ ” whenever  $\varepsilon$  appears in a statement without any other indications.

**1.1. Averaging process.** The *averaging process* on  $\mathbb{T}_\varepsilon^d$  is the continuous-time Markov process  $(u_t^\varepsilon)_{t \geq 0}$  which evolves on  $\mathbb{R}^{\mathbb{T}_\varepsilon^d}$  by updating, independently and at rate  $\varepsilon^{-2}$ , the values at nearest neighbor vertices with their average value. A possible rigorous description of the model goes through the introduction of a family of i.i.d. Poisson processes

$$N^\varepsilon = (N_t^{\varepsilon, i}(x))_{t \geq 0, x \in \mathbb{T}_\varepsilon^d, i=1, \dots, d} \quad (1.2)$$

of intensity  $\varepsilon^{-2}$ , with the arrivals of  $N_t^{\varepsilon, i}(x)$  corresponding to an update among the two nearest neighbor vertices  $x$  and  $x + \varepsilon e_i \in \mathbb{T}_\varepsilon^d$ . We write  $\mathbb{P}^\varepsilon$  and  $\mathbb{E}^\varepsilon$  for the associated probability law and expectation, respectively. Hence, for every starting configuration  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ ,  $(u_t^\varepsilon)_{t \geq 0}$  is defined as the unique *càdlàg* solution to the following finite system of SDEs driven by Poisson processes: for every  $x \in \mathbb{T}_\varepsilon^d$  and  $t > 0$ ,

$$du_t^\varepsilon(x) = \sum_{i=1}^d dN_t^{\varepsilon, i}(x) \left[ \frac{1}{2} (u_{t-}^\varepsilon(x) + u_{t-}^\varepsilon(x + \varepsilon e_i)) - u_{t-}^\varepsilon(x) \right]$$

$$\begin{aligned}
& + \sum_{i=1}^d dN_t^{\varepsilon,i}(x - \varepsilon e_i) \left[ \frac{1}{2} (u_{t-}^{\varepsilon}(x) + u_{t-}^{\varepsilon}(x - \varepsilon e_i)) - u_{t-}^{\varepsilon}(x) \right] \\
& = \sum_{i=1}^d dN_t^{\varepsilon,i}(x) \left[ \frac{1}{2} (u_{t-}^{\varepsilon}(x + \varepsilon e_i) - u_{t-}^{\varepsilon}(x)) \right] \\
& + \sum_{i=1}^d dN_t^{\varepsilon,i}(x - \varepsilon e_i) \left[ \frac{1}{2} (u_{t-}^{\varepsilon}(x - \varepsilon e_i) - u_{t-}^{\varepsilon}(x)) \right] , \tag{1.3}
\end{aligned}$$

with  $dN_t^{\varepsilon,i}(x) = N_t^{\varepsilon,i}(x) - N_{t-}^{\varepsilon,i}(x)$  denoting the increments of the Poisson process. Further, note that the spacing  $\varepsilon$  in  $\mathbb{T}_\varepsilon^d$  and the intensity  $\varepsilon^{-2}$  of the Poisson processes correspond to a diffusive space-time rescaling in (1.3).

Since  $\mathbb{T}_\varepsilon^d$  is a finite graph, it readily follows that, as a strong solution to the system above,  $(u_t^\varepsilon)_{t \geq 0}$  exists unique for all initial conditions  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , and is a Markov process. Moreover, due to the connectedness of  $\mathbb{T}_\varepsilon^d$ ,  $u_t^\varepsilon$  approaches, as  $t \rightarrow \infty$  and for all  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , the flat profile with value  $\langle u_0^\varepsilon \rangle_\varepsilon \in \mathbb{R}$ , the spatial average of  $\mathbb{T}_\varepsilon^d$ ,

$$u_t^\varepsilon(x) \xrightarrow{t \rightarrow \infty} \langle u_0^\varepsilon \rangle_\varepsilon := \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} u_0^\varepsilon(y) ,$$

$\mathbb{P}^\varepsilon$ -a.s. and uniformly over  $x \in \mathbb{T}_\varepsilon^d$ . In particular, we observe that, despite  $u_t^\varepsilon$  is, generically, a random element of  $\mathbb{R}^{\mathbb{T}_\varepsilon^d}$  for all times  $t > 0$ , the corresponding unique steady state is deterministic and identically equal to  $\langle u_0^\varepsilon \rangle_\varepsilon$ . Hence, the averaging dynamics is trivial at equilibrium, registering, in particular, no fluctuations.

In this work, we analyze asymptotics as  $\varepsilon \rightarrow 0$  of out-of-equilibrium fluctuations for this model, which, due to the degeneracy of the steady state, lacks a non-trivial notion of *local equilibrium* — a fundamental feature in the theory of hydrodynamic limits for interacting particle systems (see, e.g., [KL99]). In words, we shall find a scaling factor  $\theta_\varepsilon \rightarrow \infty$  as  $\varepsilon \rightarrow 0$  and a non-trivial process  $(\mathcal{Y}_t)_{t \geq 0}$  describing well, in the diffusive regime as  $\varepsilon \rightarrow 0$ , the nonequilibrium fluctuations encoded in the centered fields

$$\mathcal{Y}_t^\varepsilon := \theta_\varepsilon \left\{ \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} (u_t^\varepsilon(x) - \mathbb{E}^\varepsilon[u_t^\varepsilon(x)]) \delta_x \right\} , \tag{1.4}$$

after taking spatial averages against suitable test functions. Note that we consider a deterministic initial configuration  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , so to observe only dynamical fluctuations.

Maybe surprisingly at first, it turns out that, instead of the most standard CLT-scaling  $\varepsilon^{-d/2}$ , the averaging process' nonequilibrium fluctuations require an extra factor  $\varepsilon^{-1}$ , thus,

$$\theta_\varepsilon = \varepsilon^{-(d/2+1)} .$$

In this sense, fluctuations are unusually small in our context. Next to this, the limiting fluctuations  $\mathcal{Y}_t$  become Gaussian, have rather explicit space-time correlations, and satisfy  $\mathcal{Y}_0 = \lim_{t \rightarrow \infty} \mathcal{Y}_t = 0$ .

**1.2. Tiny nonequilibrium fluctuations.** In order to best describe the limit  $\mathcal{Y}_t$ , let us further manipulate the equations (1.3) defining  $u_t^\varepsilon$ . By adopting the standard notation, for all  $g \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ ,  $x \in \mathbb{T}_\varepsilon^d$  and  $i = 1, \dots, d$ ,

$$\begin{aligned}
\nabla^{\varepsilon,i} g(x) & := \varepsilon^{-1} (g(x + \varepsilon e_i) - g(x)) , & \nabla^\varepsilon g(x) & := (\nabla^{\varepsilon,i} g(x))_{i=1,\dots,d} , \\
\nabla_*^{\varepsilon,i} g(x) & := \varepsilon^{-1} (g(x) - g(x - \varepsilon e_i)) , & \nabla_*^\varepsilon g(x) & := (\nabla_*^{\varepsilon,i} g(x))_{i=1,\dots,d} ,
\end{aligned} \tag{1.5}$$

$$\Delta_\varepsilon g(x) := \sum_{i=1}^d \nabla_*^{\varepsilon,i} \nabla^{\varepsilon,i} g(x) = \sum_{i=1}^d \varepsilon^{-2} (g(x + \varepsilon e_i) + g(x - \varepsilon e_i) - 2g(x)) , \quad (1.6)$$

and letting  $\text{diag}(a) \in \mathbb{R}^{d \times d}$  denote the diagonal matrix having  $a = (a^i)_{i=1,\dots,d} \in \mathbb{R}^d$  as diagonal elements, the system (1.3) reads as

$$\begin{aligned} du_t^\varepsilon(x) &= \frac{\varepsilon}{2} \sum_{i=1}^d \{ dN_t^{\varepsilon,i}(x) \nabla^{\varepsilon,i} u_{t-}^\varepsilon(x) - dN_t^{\varepsilon,i}(x - \varepsilon e_i) \nabla^{\varepsilon,i} u_{t-}^\varepsilon(x - \varepsilon e_i) \} \\ &= \frac{\varepsilon^2}{2} \sum_{i=1}^d \nabla_*^{\varepsilon,i} (dN_t^{\varepsilon,i} \nabla^{\varepsilon,i} u_{t-}^\varepsilon)(x) \\ &= \frac{\varepsilon^2}{2} \nabla_*^\varepsilon \cdot (\text{diag}(dN_t^{\varepsilon,\cdot}) \nabla^\varepsilon u_{t-}^\varepsilon)(x) , \end{aligned}$$

or, equivalently, by passing to the compensated Poisson processes

$$\bar{N}_t^{\varepsilon,i}(x) := N_t^{\varepsilon,i}(x) - \varepsilon^{-2} t , \quad t \geq 0 , \quad x \in \mathbb{T}_\varepsilon^d , \quad i = 1, \dots, d , \quad (1.7)$$

as

$$du_t^\varepsilon(x) = \frac{1}{2} \Delta_\varepsilon u_t^\varepsilon(x) dt + \frac{\varepsilon^2}{2} \nabla_*^\varepsilon \cdot (\text{diag}(d\bar{N}_t^{\varepsilon,\cdot}) \nabla u_{t-}^\varepsilon)(x) . \quad (1.8)$$

Hence, next to a smoothening Laplacian term  $\frac{1}{2} \Delta_\varepsilon u_t^\varepsilon$ , the stochastic dynamics is driven by an i.i.d. diagonal noise which multiplies the discrete gradients  $\nabla^\varepsilon u_{t-}^\varepsilon$  of the random averaging process. The gradient  $\nabla_*^\varepsilon$  accounts for the conservative nature of the noise.

Our main task is to establish, after centering and scaling with  $\theta_\varepsilon$ , a homogenization principle for this noise term, in which two main effects equally contribute: on the one side, the random gradients  $\nabla^\varepsilon u_{t-}^\varepsilon$  get replaced by their deterministic counterparts; on the other side, the stochasticity lost in the previous step is superseded by the appearance of a non-diagonal structure of the limiting noise. More precisely, we show that the limiting fluctuations  $\mathcal{Y}_t$  solve (in a weak sense) the following SPDE on  $\mathbb{T}^d$

$$d\mathcal{Y}_t = \frac{1}{2} \Delta \mathcal{Y}_t - \frac{1}{2} \nabla \cdot (\xi \nabla u_t) , \quad (1.9)$$

where:

- $\nabla u_t$  is the gradient of the deterministic solution  $u_t$  to the heat equation on  $\mathbb{T}^d$

$$\partial_t u = \frac{1}{2} \Delta u ; \quad (1.10)$$

- $\xi = (\xi^{i,j})_{i,j=1,\dots,d}$  is a matrix-valued space-time white noise whose covariance is informally given, for some  $\mathbf{a} = \mathbf{a}(d) \in (0, 1]$ , by

$$\mathbb{E}[\xi_t^{i,j}(x) \xi_s^{k,\ell}(y)] = \delta_0(t-s) \delta_0(x-y) \mathbf{1}_{i=k} \mathbf{1}_{j=\ell} \left( (1 + \mathbf{a}) \mathbf{1}_{i=j} + \frac{1 - \mathbf{a}}{d-1} \mathbf{1}_{i \neq j} \right) , \quad (1.11)$$

for all  $x, y \in \mathbb{T}^d$ ,  $t, s \geq 0$ , and  $i, j, k, \ell = 1, \dots, d$ ;

- a weak solution to (1.9) is meant to be a solution to the corresponding martingale problem (cf. (1.16), (5.2)).

The precise statement of this convergence result is the content of the following theorem, before which we need to introduce some notation. For all  $k \in \mathbb{N}_0 \cup \{\infty\}$ , let  $\mathcal{C}^k(\mathbb{T}^d)$  be the space of  $k$ -continuous differentiable functions on  $\mathbb{T}^d$ . We write  $\mathcal{C}(\mathbb{T}^d) = \mathcal{C}^0(\mathbb{T}^d)$ . For all  $\alpha \in \mathbb{R}$ ,  $H^\alpha(\mathbb{T}^d)$  denotes the  $\alpha^{\text{th}}$ -order fractional Sobolev space on  $\mathbb{T}^d$ , defined as the closure of  $\mathcal{C}^\infty(\mathbb{T}^d)$  with respect to the following Hilbertian (semi)norm

$$\|f\|_{H^\alpha(\mathbb{T}^d)} := \left( \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^\alpha |\langle f | \phi_m \rangle_{L^2(\mathbb{T}^d)}|^2 \right)^{\frac{1}{2}} , \quad f \in \mathcal{C}^\infty(\mathbb{T}^d) . \quad (1.12)$$

Here,  $\phi_m(x) := e^{2\pi im \cdot x}$ , for all  $x \in \mathbb{T}^d$  and  $m \in \mathbb{Z}^d$ . It is well-known that  $H^\alpha(\mathbb{T}^d)$  is a Hilbert space, and that, for  $\alpha = k \in \mathbb{N}_0$ ,  $H^\alpha(\mathbb{T}^d)$  coincides with the usual Sobolev space  $W^{2,k}(\mathbb{T}^d)$ . Moreover,  $H^\alpha(\mathbb{T}^d) \hookrightarrow H^\beta(\mathbb{T}^d)$  for all  $\alpha > \beta$ , and  $\mathcal{Y}_t^\varepsilon \in H^{-\alpha}(\mathbb{T}^d)$  for all  $\alpha > d/2$ . For more details and properties, we refer to, e.g., [JM18, Section C]. Finally, we write  $\mathcal{D}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$  (resp.  $\mathcal{C}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$ ) for the space of càdlàg (resp. continuous) trajectories taking values in  $H^{-\alpha}(\mathbb{T}^d)$  (see, e.g., [Bil99, §2–3]).

**Theorem 1** (Nonequilibrium fluctuations). *Fix  $\alpha > 3 + d/2$  and  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ . Then, there exists  $\mathfrak{a} = \mathfrak{a}(d) \in (0, 1]$  such that, when initializing the averaging process  $(u_t^\varepsilon)_{t \geq 0}$  with  $u_0^\varepsilon = u_0|_{\mathbb{T}_\varepsilon^d}$ , the following convergence in law*

$$(\mathcal{Y}_t^\varepsilon)_{t \geq 0} \xrightarrow{\varepsilon \rightarrow 0} (\mathcal{Y}_t)_{t \geq 0}, \quad \text{in } \mathcal{D}([0, \infty); H^{-\alpha}(\mathbb{T}^d)),$$

holds true for the corresponding fluctuation fields  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$  given in (1.4) with  $\theta_\varepsilon = \varepsilon^{-(d/2+1)}$ , where  $(\mathcal{Y}_t)_{t \geq 0}$  is the unique process in  $\mathcal{C}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$  solving the SPDE (1.9)–(1.11), with:

- initial condition  $\mathcal{Y}_0 = 0$ ;
- $(u_t)_{t \geq 0}$  therein being the solution to (1.10) with initial condition  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ .

**Remark 1.1** (Constant  $\mathfrak{a}$ ). *Since our proofs do not require it, we did not try to extract the precise numerical value of the constant  $\mathfrak{a}$  appearing in Theorem 1. We just mention that*

$$\mathfrak{a} = 1 \quad \text{for } d = 1, \quad \mathfrak{a} = \frac{\pi}{3\pi - 4} \quad \text{for } d = 2,$$

while we have some more abstract expressions for  $d \geq 3$  guaranteeing  $\mathfrak{a} \in (0, 1)$ . For more details, see Remark A.2.

**1.3. A law of large numbers for the squared gradients.** In order to prove Theorem 1, we follow a probabilistic approach, taking advantage of the Markovian nature of the averaging process  $u_t^\varepsilon$ . In formulas, recalling (1.8) and that  $\mathcal{Y}_0^\varepsilon = 0$ , this corresponds to decompose the fields  $\mathcal{Y}_t^\varepsilon$  in (1.4), when tested against a test function  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$ , as

$$\mathcal{Y}_t^\varepsilon(f) = \int_0^t \mathcal{Y}_s^\varepsilon \left( \frac{1}{2} \Delta_\varepsilon f \right) ds + \mathcal{M}_t^\varepsilon(f), \quad t \geq 0, \quad (1.13)$$

where, for  $\theta_\varepsilon = \varepsilon^{-(d/2+1)}$ ,

$$\begin{aligned} \mathcal{M}_t^\varepsilon(f) &:= \int_0^t \theta_\varepsilon \varepsilon^{d+2} \sum_{x \in \mathbb{T}_\varepsilon^d} \left( (\text{diag}(\text{d}\bar{N}_s^{\varepsilon, \cdot}) \nabla u_{s-}^\varepsilon)(x) \right) \cdot \left( \frac{1}{2} \nabla^\varepsilon f(x) \right) \\ &= \int_0^t \varepsilon^{d/2+1} \sum_{i=1}^d \sum_{x \in \mathbb{T}_\varepsilon^d} \left( \text{d}\bar{N}_s^{\varepsilon, i}(x) \nabla^{\varepsilon, i} u_{s-}^\varepsilon(x) \right) \left( \frac{1}{2} \nabla^{\varepsilon, i} f(x) \right), \quad t \geq 0, \end{aligned} \quad (1.14)$$

is a square-integrable martingale (with respect to the filtration generated by the process  $u_t^\varepsilon$ ). The martingale  $\mathcal{M}_t^\varepsilon(f)$  has jumps and may be described through its predictable quadratic variation, referred to as  $\langle \mathcal{M}^\varepsilon(f) \rangle_t$  and explicitly given by

$$\langle \mathcal{M}^\varepsilon(f) \rangle_t = \int_0^t \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \left( \nabla^{\varepsilon, i} u_s^\varepsilon(x) \right)^2 \left( \frac{1}{2} \nabla^{\varepsilon, i} f(x) \right)^2 ds, \quad t \geq 0. \quad (1.15)$$

This derivation uses elementary properties of the compensated Poisson processes in (1.7), such as  $\mathbb{E}[\bar{N}_t^{\varepsilon, i}(x)] = 0$  and  $\langle \bar{N}^{\varepsilon, i}(x) \rangle_t = \varepsilon^{-2}t$ , as well as their independence.

In view of the decomposition in (1.13), the proof of Theorem 1 boils down to proving the following three claims:

- (i) the sequence  $((\mathcal{Y}_t^\varepsilon)_{t \geq 0})_\varepsilon$  is tight in  $\mathcal{D}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$ , for  $\alpha > 3 + d/2$ ;
- (ii) all limiting processes  $(\mathcal{Y}_t)_{t \geq 0}$  are continuous;
- (iii) all limiting processes  $(\mathcal{Y}_t)_{t \geq 0}$  are solutions to the following martingale problem

$$\mathcal{Y}_t(f) = \int_0^t \mathcal{Y}_s(\frac{1}{2}\Delta f) ds + \mathcal{M}_t(f), \quad t \geq 0, \quad f \in \mathcal{C}^\infty(\mathbb{T}^d), \quad (1.16)$$

where  $(\mathcal{M}_t(f))_{t \geq 0}$  is a martingale (with respect to the filtration generated by  $(\mathcal{Y}_t)_{t \geq 0}$ ) whose predictable quadratic variation is deterministic and compatible with the noise structure described in (1.9)–(1.11).

By the form of the limiting martingale and standard uniqueness results due to Holley and Stroock [HS78], (ii) and (iii) together would uniquely characterize the limit as the one described in Theorem 1.

While the proofs of the steps (i) and (ii) above are not particularly problematic, establishing the third step on the identification of the limiting martingale represents the major challenge of the proof of Theorem 1. More specifically, thanks to classical martingale convergence theorems, our main task reduces to prove convergence in probability of the corresponding predictable quadratic variations. Reading out the explicit expression of these quadratic variations from the right-hand side of (1.15), this means proving a weak law of large numbers for weighted space-time averages of squares of discrete gradients of  $u_t^\varepsilon$ . Most of our analysis is devoted to the proof of this result, which, for later reference, we now state in form of a theorem.

**Theorem 2** (LLN for squared discrete gradients). *Fix  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ , and let  $(u_t^\varepsilon)_{t \geq 0}$  and  $(u_t)_{t \geq 0}$  denote, respectively, the averaging process and the solution to the heat equation (1.10), started from  $u_0^\varepsilon = u_0|_{\mathbb{T}_\varepsilon^d}$  and  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ . Then, for all  $g = (g^i) \in (\mathcal{C}([0, \infty) \times \mathbb{T}^d))^d$  and  $t \geq 0$ , we have*

$$\mathbb{E}^\varepsilon \left[ (\Gamma_t^\varepsilon(g) - \Gamma_t(g))^2 \right] \xrightarrow{\varepsilon \rightarrow 0} 0,$$

where

$$\Gamma_t^\varepsilon(g) := \sum_{i=1}^d \int_0^t \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} (\nabla^{\varepsilon, i} u_s^\varepsilon(x))^2 g_s^i(x) ds, \quad (1.17)$$

$$\Gamma_t(g) := \sum_{i,j=1}^d \left( (1 + \mathbf{a}) \mathbf{1}_{i=j} + \frac{1 - \mathbf{a}}{d-1} \mathbf{1}_{i \neq j} \right) \int_0^t \int_{\mathbb{T}^d} (\nabla^j u_s(x))^2 g_s^i(x) dx ds, \quad (1.18)$$

and  $\mathbf{a} = \mathbf{a}(d) \in (0, 1]$  is the constant in Theorem 1 and Remark 1.1.

**1.4. Proof ideas and techniques.** The starting point in the proof of the LLN in Theorem 2 is a variance estimate (Theorem 3). We derive this by tools from Malliavin calculus in Poisson space (see, e.g., [LP11, LP18] and references therein). More specifically, we will resort to one specific tool in that context: the Poincaré inequality. In words, this inequality allows to estimate the variance of a functional  $F = F(N^\varepsilon)$  of the underlying Poisson process in terms of (a norm of) its derivative  $DF$  with respect to an infinitesimal variation of  $N^\varepsilon$ . The functional  $DF$  is referred to as the Malliavin derivative.

As we will show, for the averaging process,  $DF$  comes with a nice probabilistic dynamical interpretation in terms of an evolving discrepancy (Section 3.1). We exploit this point of view to turn this Poincaré inequality into an effective bound. In order to achieve this, we derive some new second-moment estimates for the discrete gradients of the averaging process, and combine them with properties of the averaging process earlier developed in [AL12, QS23, Sau23]. Such improved estimates are key in the proof of Theorem 2, for

controlling the variance of  $\Gamma_t^\varepsilon(g)$ , and are sharp enough to determine the limit of its mean (Theorem 4).

1.4.1. *Poincaré inequality in Poisson space.* In our setting, this Poincaré inequality may be rigorously stated as follows. Recall the Poisson process  $N^\varepsilon$  in (1.2) of the edge updates, which we can describe, more precisely, through its intensity measure  $\gamma_\varepsilon$  on  $\mathbb{Y}_\varepsilon := [0, \infty) \times \mathbb{T}_\varepsilon^d \times \{1, \dots, d\}$  (endowed with the corresponding product Borel  $\sigma$ -field), given by the product measure

$$\gamma_\varepsilon = \varepsilon^{-2} dt \otimes \nu_{\mathbb{T}_\varepsilon^d} \otimes \nu_{\{1, \dots, d\}} .$$

Here,  $dt$  stands for the Lebesgue measure on  $[0, \infty)$ , whereas, for a discrete set  $S$ ,  $\nu_S$  denotes the counting measure on  $S$ . Letting  $\mathbf{N}(\mathbb{Y}_\varepsilon)$  be the space of integer-valued  $\sigma$ -finite measures on  $\mathbb{Y}_\varepsilon$ , we then interpret  $N^\varepsilon$  as a random element in the measurable space  $\mathbf{N}(\mathbb{Y}_\varepsilon)$ , endowed with  $\mathcal{N}^\varepsilon$ , the smallest  $\sigma$ -field making the mappings  $m \in \mathbf{N}(\mathbb{Y}_\varepsilon) \mapsto m(A) \in \mathbb{R}$  measurable, for all measurable sets  $A \subset \mathbb{Y}_\varepsilon$ . Now,  $\mathbb{P}^\varepsilon$  and  $\mathbb{E}^\varepsilon$  stand for the law and expectation of  $N^\varepsilon$ , respectively.

Then, the *Poincaré inequality in Poisson space* reads, for every measurable function  $F : \mathbf{N}(\mathbb{Y}_\varepsilon) \rightarrow \mathbb{R}$ , as

$$\begin{aligned} \text{Var}_\varepsilon(F) &:= \mathbb{E}^\varepsilon[(F(N^\varepsilon) - \mathbb{E}^\varepsilon[F(N^\varepsilon)])^2] \leq \int_{\mathbb{Y}_\varepsilon} \mathbb{E}^\varepsilon[(D_{(t,x,i)}F(N^\varepsilon))^2] d\gamma_\varepsilon(t, x, i) \\ &= \varepsilon^{-(d+2)} \int_0^\infty \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E}^\varepsilon[(D_{(t,x,i)}F(N^\varepsilon))^2] dt . \end{aligned} \quad (1.19)$$

In this formula, for all  $(t, x, i) \in \mathbb{Y}_\varepsilon$ ,  $D_{(t,x,i)}$  is the difference operator defined, for all measurable  $F : \mathbf{N}(\mathbb{Y}_\varepsilon) \rightarrow \mathbb{R}$ , as

$$D_{(t,x,i)}F(m) := F(m + \delta_{(t,x,i)}) - F(m) , \quad m \in \mathbf{N}(\mathbb{Y}_\varepsilon) ,$$

with  $\delta_{(t,x,i)}$  denoting the Dirac delta at  $(t, x, i) \in \mathbb{Y}_\varepsilon$ . Roughly speaking,  $D_{(t,x,i)}$  plays the role of derivative at the “point”  $m \in \mathbf{N}(\mathbb{Y}_\varepsilon)$  in the “direction”  $(t, x, i) \in \mathbb{Y}_\varepsilon$ . In our context, if we keep in mind the interpretation of  $N^\varepsilon$  as identifying the edge updates for the averaging process,  $D_{(t,x,i)}F$  measures the effect on  $F$  of the addition of an update at time  $t$  between the nearest neighbors  $x$  and  $x + \varepsilon e_i \in \mathbb{T}_\varepsilon^d$ .

1.4.2. *Application of Poincaré inequality.* In proving Theorem 2, we will apply the Poincaré inequality (1.19) to functionals  $F$  of the form  $\Gamma_t^\varepsilon(g)$  given as in (1.17), for some  $t > 0$ ,  $g = (g^i)_{i=1, \dots, d} \in (\mathcal{C}([0, \infty) \times \mathbb{T}^d))^d$ , and initial condition  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ . This yields

$$\begin{aligned} \text{Var}_\varepsilon(\Gamma_t^\varepsilon(g)) &\leq \varepsilon^{-(d+2)} \int_0^\infty \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E}^\varepsilon[(D_{(s,x,i)}\Gamma_t^\varepsilon(g))^2] ds \\ &= \varepsilon^{-(d+2)} \int_0^t \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E}^\varepsilon[(D_{(s,x,i)}\Gamma_t^\varepsilon(g))^2] ds , \end{aligned} \quad (1.20)$$

where the identity uses the fact that  $\Gamma_t^\varepsilon(g)$  does not depend on the history after time  $t$ . Moreover, since  $(\Gamma_t^\varepsilon(g))_{t \geq 0}$  is an additive functional of  $(u_t^\varepsilon)_{t \geq 0}$ ,  $D_{(s,x,i)}\Gamma_t^\varepsilon(g)$  may be further simplified, for  $s \leq t$ , as

$$D_{(s,x,i)}\Gamma_t^\varepsilon(g) = \sum_{j=1}^d \int_s^t \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \left\{ (\nabla^{\varepsilon, j} E_{(s,x,i)} u_r^\varepsilon(y))^2 - (\nabla^{\varepsilon, j} u_r^\varepsilon(y))^2 \right\} g_r^j(y) dr ,$$

where, for  $r \geq s$ ,  $E_{(s,x,i)}^\varepsilon u_r^\varepsilon$  is the averaging process following the edge updates  $N^\varepsilon$ , conditioned to be equal, at time  $r = s$ , to

$$y \in \mathbb{T}_\varepsilon^d \longmapsto E_{(s,x,i)}^\varepsilon u_s^\varepsilon(y) := \begin{cases} \frac{1}{2} (u_s^\varepsilon(x) + u_s^\varepsilon(x + \varepsilon e_i)) & \text{if } y = x \text{ or } y = x + \varepsilon e_i \\ u_s^\varepsilon(y) & \text{else .} \end{cases}$$

Analogously to  $D_{(s,x,i)}$  (which acts on functionals  $F$ ), the operator  $E_{(s,x,i)}$  acts on  $(u_r^\varepsilon)_{r \geq 0}$  by imposing an extra averaging-update at vertices  $x$  and  $x + \varepsilon e_i \in \mathbb{T}_\varepsilon^d$  at time  $s \geq 0$ . Furthermore, by linearity of the averaging process,  $E_{(s,x,i)}^\varepsilon u_r^\varepsilon$  may be further decomposed as

$$E_{(s,x,i)}^\varepsilon u_r^\varepsilon = u_r^\varepsilon + w_r^{\varepsilon,(s,x,i)}, \quad r \geq s, \quad (1.21)$$

namely, a sum of the original process  $u_r^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  and a ‘‘discrepancy’’  $w_r^{\varepsilon,(s,x,i)} \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , both evolving as two averaging processes with two different initializations, sharing the same edge updates. In particular, they are not independent. In view of this, we may recast an estimate of the right-hand side of (1.20) into a bound on the evolution of such a discrepancy, and a central part of our analysis deals with establishing precise quantitative estimates on it and its discrete gradients (see Section 3.2 for more details).

**1.4.3. Ultracontractivity and gradient estimates for the averaging process.** Next to the Poincaré inequality in Poisson space, our analysis builds on two functional analytic properties of the averaging flow on the discrete torus: ultracontractivity and second-moment gradient estimates.

*Ultracontractivity.* By ultracontractivity we mean a bound of the following form: for all  $T > 0$  and some  $C = C(d, T) > 0$ ,

$$\mathbb{E}^\varepsilon [ \|u_t^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 ] \leq \langle u_0^\varepsilon \rangle_\varepsilon^2 + C \|u_0^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 \left( (t \vee \varepsilon^2)^{-d/2} \vee 1 \right), \quad t \in (0, T). \quad (1.22)$$

Here,  $\|\cdot\|_{L^p(\mathbb{T}_\varepsilon^d)}$ ,  $p \in [1, \infty]$ , denotes the  $L^p$ -norm on  $\mathbb{T}_\varepsilon^d$  with respect to the uniform measure  $\varepsilon^d$ , i.e., for all  $g \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ ,

$$\|g\|_{L^p(\mathbb{T}_\varepsilon^d)} := \left( \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} |g(x)|^p \right)^{\frac{1}{p}} \text{ for } p \in [1, \infty), \quad \|g\|_{L^\infty(\mathbb{T}_\varepsilon^d)} := \sup_{x \in \mathbb{T}_\varepsilon^d} |g(x)|. \quad (1.23)$$

We write  $\langle \cdot | \cdot \rangle_{L^2(\mathbb{T}_\varepsilon^d)}$  for the inner product in  $L^2(\mathbb{T}_\varepsilon^d)$ .

Estimates of type (1.22) for parabolic equations are classical, dating back to Nash’s work ([Nas58], see also [FS86]), and are well-known also in the graph context (see, e.g., [DSC96]). For the averaging process, they were first derived only recently in [QS23], and further refined on  $\mathbb{T}_\varepsilon^d$ , establishing the stronger inequality [Sau23, Theorem 2.1]

$$\mathbb{E}^\varepsilon [ \|u_t^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 ] \leq C \|u_0^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 (t \vee \varepsilon^2)^{-(d/2+1)}, \quad t \geq 0. \quad (1.24)$$

On the left-hand side above,  $\|\cdot\|_{H^1(\mathbb{T}_\varepsilon^d)}$  reads, for all  $g \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , as

$$\|g\|_{H^1(\mathbb{T}_\varepsilon^d)} := \left( \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d |\nabla^{\varepsilon,i} g(x)|^2 \right)^{\frac{1}{2}}. \quad (1.25)$$

In order to see why (1.24) is stronger than (1.22), it suffices to integrate over time the following identity due to Aldous and Lanoue [AL12, Eq. (2.7)]:

$$\frac{d}{dt} \mathbb{E} [ \|u_t^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 ] = -\frac{1}{2} \mathbb{E} [ \|u_t^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 ], \quad t \geq 0. \quad (1.26)$$

Finally, we observe that the bounds in (1.22) and (1.24) are effective for all  $t = O(1)$ , as well as for very small times  $t \ll 1$ . We emphasize that such estimates for  $t \ll 1$  will play a key role in Lemma 3.3, when proving Theorem 2.

*Gradient estimates.* The second important and new ingredient is a thorough analysis of the dynamics of discrete gradients of the averaging process on the torus on a diffusive space-time scale. This will be carried out in Sections 2 and 4. In a nutshell, this approach goes through:

- (a) first, solving explicitly the SDEs governing the evolution of  $(\nabla^{\varepsilon,i} u_t^\varepsilon)_{t \geq 0} \subset \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  (Lemma 2.1);
- (b) then, reducing asymptotics of the second moments  $\mathbb{E}[(\nabla^{\varepsilon,i} u_t^\varepsilon(x))^2]$  to the analysis of infinitely-many iterated integrals (Propositions 2.1, 4.1 and 4.2).

This latter step is what allows us to determine the non-diagonal form of the limiting correlations (cf. (1.11)). Next to asymptotics, we extract some new pseudo-contractivity and small-time estimates for gradients' second moments of independent interest (Lemmas 2.3–2.5).

It is worth to mention that similar estimates were recently obtained in [Sau23, Proposition 2.2], aiming at capturing the correct order of magnitude of  $\mathbb{E}[(\nabla^{\varepsilon,i} u_t^\varepsilon(x))^2]$  for a wider time window, including the regime  $t \gg 1$  as  $\varepsilon \rightarrow 0$ . However, while the approach in the aforementioned work only provides bounds, in the present article, we are able to determine exact asymptotics by restricting to times  $t = O(1)$ . Finally, our analysis relies solely on second moments, not involving higher (e.g., fourth) moment estimates, whose control seems to be out of reach with the present techniques.

**1.5. Structure of the paper.** The rest of the paper is organized as follows. In Section 2, we collect some technical results concerning discrete gradients of the averaging process which will be used all throughout. In Sections 3 and 4 we prove Theorem 2 on the LLN for  $\Gamma_t^\varepsilon$  defined in (1.17). More specifically, in the former section, we prove that variances vanish as  $\varepsilon \rightarrow 0$ , while in the latter one we establish the convergence of the means. Section 5 is devoted to the proof of Theorem 1. In Appendix A, we prove a few auxiliary results employed in Sections 3 and 4.

## 2. DISCRETE GRADIENTS

In this section, we derive a number of technical results on the random discrete gradients of  $u_t^\varepsilon$  on  $\mathbb{T}_\varepsilon^d$ , of use for the subsequent sections. Before starting, let us introduce some notation.

After recalling that  $\mathcal{N}^\varepsilon$  is the  $\sigma$ -field associated to  $N^\varepsilon$  (cf. Section 1.4.1), we define  $\mathcal{F}^\varepsilon := \mathcal{N}^\varepsilon \otimes \sigma(u_0^\varepsilon)$  as the product  $\sigma$ -field associated to the Poisson process and the averaging process' initial conditions. We write  $(\mathcal{F}_t^\varepsilon)_{t \geq 0}$  for the corresponding filtration, with respect to which  $(u_t^\varepsilon)_{t \geq 0} \subset \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  is adapted. In particular,  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  is  $\mathcal{F}_0^\varepsilon$ -measurable. Furthermore, recalling the definitions of  $\nabla^{\varepsilon,i}$ ,  $\nabla_*^{\varepsilon,i}$ ,  $\nabla^\varepsilon$ ,  $\nabla_*^\varepsilon$  and  $\Delta_\varepsilon$  from (1.5)–(1.6), we introduce the random walk  $(X_t^\varepsilon)_{t \geq 0}$  on  $\mathbb{T}_\varepsilon^d$ , with semigroup  $(P_t^\varepsilon)_{t \geq 0}$  and infinitesimal generator  $\frac{1}{2}\Delta_\varepsilon$ . Moreover, we let  $(p_t^\varepsilon(x, y))_{t \geq 0, x, y \in \mathbb{T}_\varepsilon^d}$  denote the corresponding transition probabilities. In what follows, we will repeatedly use that all discrete gradients commute, namely,

$$\nabla^{\varepsilon,i} \nabla^{\varepsilon,j} = \nabla^{\varepsilon,j} \nabla^{\varepsilon,i}, \quad \nabla^{\varepsilon,i} \nabla_*^{\varepsilon,j} = \nabla_*^{\varepsilon,j} \nabla^{\varepsilon,i}, \quad \nabla_*^{\varepsilon,i} \nabla_*^{\varepsilon,j} = \nabla_*^{\varepsilon,j} \nabla_*^{\varepsilon,i}, \quad (2.1)$$

for all  $i, j = 1, \dots, d$ , from which we readily obtain the following well-known intertwining relations for the semigroup and the gradients on  $\mathbb{T}_\varepsilon^d$ :

$$P_t^\varepsilon \nabla^{\varepsilon,i} = \nabla^{\varepsilon,i} P_t^\varepsilon, \quad P_t^\varepsilon \nabla_*^{\varepsilon,i} = \nabla_*^{\varepsilon,i} P_t^\varepsilon, \quad t \geq 0, \quad i = 1, \dots, d. \quad (2.2)$$

We remark that  $\Delta_\varepsilon$  and  $P_t^\varepsilon$  are symmetric operators on  $L^2(\mathbb{T}_\varepsilon^d)$  (cf. (1.23)) and that, by translation and rotational invariance of the random walk's dynamics on  $\mathbb{T}_\varepsilon^d$ , we have  $p_t^\varepsilon(x, y) = p_t^\varepsilon(y, x) = p_t^\varepsilon(y - x, 0) = p_t^\varepsilon(x - y, 0)$ . Finally, in order to lighten the notation, we write  $\mathbb{P} = \mathbb{P}^\varepsilon$  and  $\mathbb{E} = \mathbb{E}^\varepsilon$  all throughout.

**2.1. Stochastic dynamics of discrete gradients.** As a first step, we write the random discrete gradients of  $u_t^\varepsilon$  in a mild-solution form. For this purpose, recall the definition of compensated Poisson processes  $\bar{N}^\varepsilon = (\bar{N}_t^{\varepsilon, i}(x))_{t \geq 0, x \in \mathbb{T}_\varepsilon^d, i=1, \dots, d}$ , as well as the equations (1.8) solved by  $u_t^\varepsilon$ .

**Lemma 2.1** (Mild solution). *Fix  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ . Then,  $\mathbb{P}$ -a.s., we have*

$$\nabla^{\varepsilon, i} u_t^\varepsilon(x) = P_t^\varepsilon(\nabla^{\varepsilon, i} u_0^\varepsilon)(x) + \sum_{j=1}^d \sum_{y \in \mathbb{T}_\varepsilon^d} \int_0^t b_{t-s}^{\varepsilon, i, j}(x, y) \nabla^{\varepsilon, j} u_{s-}^\varepsilon(y) d\bar{N}_s^{\varepsilon, j}(y), \quad (2.3)$$

for all  $t > 0$ ,  $x \in \mathbb{T}_\varepsilon^d$  and  $i = 1, \dots, d$ , where

$$b_t^\varepsilon(x, y) := \frac{1}{2} (p_t^\varepsilon(x, y + \varepsilon e_i) + p_t^\varepsilon(x, y - \varepsilon e_j) - p_t^\varepsilon(x + \varepsilon e_i - \varepsilon e_j, y) - p_t^\varepsilon(x, y)). \quad (2.4)$$

*Proof.* Passing to the discrete gradients in (1.8), we obtain

$$\begin{aligned} d(\nabla^{\varepsilon, i} u_t^\varepsilon)(x) &= \frac{1}{2} \nabla^{\varepsilon, i} \Delta_\varepsilon u_t^\varepsilon(x) dt + \frac{\varepsilon^2}{2} \nabla^{\varepsilon, i} (\nabla_*^\varepsilon \cdot (\text{diag}(d\bar{N}_t^{\varepsilon, \cdot}) \nabla u_{t-}^\varepsilon))(x) \\ &= \frac{1}{2} \Delta_\varepsilon (\nabla^{\varepsilon, i} u_t^\varepsilon)(x) + \frac{\varepsilon^2}{2} \sum_{j=1}^d \nabla_*^{\varepsilon, j} \nabla^{\varepsilon, i} (d\bar{N}_t^{\varepsilon, j} \nabla^{\varepsilon, j} u_{t-}^\varepsilon)(x), \end{aligned}$$

where for the second step we used the commutation relations (2.1). Note that the above finite system of SDEs (with compensated-Poisson noise) has the form  $df_t = \frac{1}{2} \Delta_\varepsilon f_t dt + F_t$ , hence, recalling that  $P_t^\varepsilon = \exp(\frac{t}{2} \Delta_\varepsilon)$ ,  $\nabla^{\varepsilon, i} u_t^\varepsilon$  admits the following explicit form,  $\mathbb{P}$ -a.s.,

$$\begin{aligned} \nabla^{\varepsilon, i} u_t^\varepsilon &= P_t^\varepsilon(\nabla^{\varepsilon, i} u_0^\varepsilon)(x) + \int_0^t P_{t-s}^\varepsilon \left( \frac{\varepsilon^2}{2} \sum_{j=1}^d \nabla_*^{\varepsilon, j} \nabla^{\varepsilon, i} (d\bar{N}_s^{\varepsilon, j} \nabla^{\varepsilon, j} u_{s-}^\varepsilon) \right) (x) \\ &= P_t^\varepsilon(\nabla^{\varepsilon, i} u_0^\varepsilon)(x) + \sum_{j=1}^d \frac{\varepsilon^2}{2} \int_0^t \nabla_*^{\varepsilon, j} \nabla^{\varepsilon, i} P_{t-s}^\varepsilon (d\bar{N}_s^{\varepsilon, j} \nabla^{\varepsilon, j} u_{s-}^\varepsilon)(x), \end{aligned}$$

where the last step employs the intertwining in (2.2). The desired claim now follows by recalling (2.4) and observing that, for all  $g \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , we have

$$\begin{aligned} \varepsilon^2 \nabla_*^{\varepsilon, j} \nabla^{\varepsilon, i} P_t^\varepsilon g(x) &= \varepsilon (\nabla^{\varepsilon, i} P_t^\varepsilon g(x) - \nabla^{\varepsilon, i} P_t^\varepsilon g(x - \varepsilon e_j)) \\ &= P_t^\varepsilon g(x + \varepsilon e_i) - P_t^\varepsilon g(x) - P_t^\varepsilon g(x - \varepsilon e_j + \varepsilon e_i) + P_t^\varepsilon g(x - \varepsilon e_j) \\ &= \sum_{y \in \mathbb{T}_\varepsilon^d} 2b_t^\varepsilon(x, y) g(y). \end{aligned}$$

This concludes the proof of the lemma.  $\square$

**2.2. Second moments of discrete gradients.** Let us iterate the expression from Lemma 2.1 to obtain formulas for the expectation of  $\nabla^{\varepsilon, i} u_t^\varepsilon(x) \nabla^{\varepsilon, i} v_t^\varepsilon(x)$ , where  $(u_t^\varepsilon)_{t \geq 0}$  and  $(v_t^\varepsilon)_{t \geq 0}$  are two copies of the averaging process starting from  $u_0^\varepsilon$  and  $v_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ , respectively, driven by the same Poisson noise. For this and subsequent results, the following quantity defined in terms of  $b_t^{\varepsilon, i, j}(x, y)$  (see (2.4))

$$q_t^{\varepsilon, i, j}(x, y) := (\varepsilon^{-1} b_t^{\varepsilon, i, j}(x, y))^2, \quad t \geq 0, \quad x, y \in \mathbb{T}_\varepsilon^d, \quad i, j = 1, \dots, d, \quad (2.5)$$

will play a key role, whose main properties are collected in the next lemma and proved in Appendix A.

**Lemma 2.2.** *For all  $t \geq 0$ ,  $x, y \in \mathbb{T}_\varepsilon^d$  and  $i, j = 1, \dots, d$ , we have*

$$q_t^{\varepsilon, i, j}(x, y) = q_t^{\varepsilon, i, j}(y, x). \quad (2.6)$$

Moreover, the following quantity

$$\mathcal{Q}_\varepsilon(t) := \sum_{j=1}^d \sum_{y \in \mathbb{T}_\varepsilon^d} q_t^{\varepsilon, i, j}(x, y) \quad (2.7)$$

depends only on  $d \geq 1$  and  $t \geq 0$  (thus, not on  $x \in \mathbb{T}_\varepsilon^d$  and  $i = 1, \dots, d$ ), and satisfies

$$\int_0^\infty \mathcal{Q}_\varepsilon(t) dt = \frac{1}{2}. \quad (2.8)$$

We state a notational remark.

**Remark 2.1** (Notation). *Here and all throughout, we adopt the following standard notation: for all  $t > 0$  and  $k \in \mathbb{N}$ ,*

$$[0, t]_{>}^k := \left\{ (s_1, \dots, s_k) \in [0, t]^k : s_1 > \dots > s_k \right\},$$

while, for all  $h \in \mathcal{C}([0, t]_{>}^k)$ ,

$$\int_{[0, t]_{>}^k} ds_1 \cdots ds_k h(\cdot) := \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} h(\cdot) ds_k \cdots ds_1.$$

Moreover, within this context, we implicitly identify  $s_0 := t$  and, whenever  $x \in \mathbb{T}_\varepsilon^d$  is a fixed vertex and  $i = 1, \dots, d$  a fixed direction,  $y_0 := x$  and  $j_0 = i$ . We employ analogous notations when replacing 0 (resp.  $t$ ) by  $-\infty$  (resp.  $\infty$ ), in  $[0, t]_{>}^k$ .

**Proposition 2.1.** *For all  $u_0^\varepsilon, v_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$ ,  $t > 0$ ,  $x \in \mathbb{T}_\varepsilon^d$  and  $i = 1, \dots, d$ , we have*

$$\mathbb{E}[\nabla^{\varepsilon, i} u_t^\varepsilon(x) \nabla^{\varepsilon, i} v_t^\varepsilon(x)] = \sum_{k=0}^\infty \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon, v_0^\varepsilon)(x), \quad (2.9)$$

the above series being absolutely convergent. Here, each summand is defined, for  $k = 0$ , as

$$\Pi_t^{\varepsilon, i, 0}(u_0^\varepsilon, v_0^\varepsilon)(x) := P_t^\varepsilon(\nabla^{\varepsilon, i} u_0^\varepsilon)(x) P_t^\varepsilon(\nabla^{\varepsilon, i} v_0^\varepsilon)(x), \quad (2.10)$$

while, for  $k \geq 1$ , as (identifying  $s_0 = t$ ,  $y_0 = x$  and  $j_0 = i$ , cf. Remark 2.1)

$$\Pi_t^{\varepsilon, i, k}(u_0^\varepsilon, v_0^\varepsilon)(x) \quad (2.11)$$

$$:= \int_{[0, t]_{>}^k} ds_1 \cdots ds_k \sum_{y_1, \dots, y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_1, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon, v_0^\varepsilon)(y_k),$$

where  $q^\varepsilon$  is given in (2.5). In particular, for  $u_0^\varepsilon = v_0^\varepsilon$ , we have

$$\mathbb{E}[(\nabla^{\varepsilon, i} u_t^\varepsilon(x))^2] = \sum_{k=0}^\infty \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)(x), \quad \text{with } \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon) := \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon, u_0^\varepsilon). \quad (2.12)$$

*Proof.* The two proofs of (2.9) and (2.12) are essentially the same; for notational convenience, let us discuss in detail the one of (2.12), and only quickly comment on that of (2.9). Recall (2.3) and note that  $P_t^\varepsilon(\nabla^{\varepsilon, i} u_0^\varepsilon)$  is  $\mathcal{F}_0^\varepsilon$ -measurable, while

$$\mathbb{E} \left[ \int_0^t b_{t-s}^{\varepsilon, i, j}(x, y) (d\bar{N}_s^{\varepsilon, j} \nabla^{\varepsilon, j} u_{s-}^\varepsilon)(y) \Big| \mathcal{F}_0^\varepsilon \right] = 0,$$

because  $s \mapsto b_{t-s}^{\varepsilon,i,j}(x,y) \nabla^{\varepsilon,j} u_{s-}^{\varepsilon}(y)$  is predictable and  $s \mapsto \bar{N}_s^{\varepsilon,j}(y)$  is a martingale with respect to  $(\mathcal{F}_s^{\varepsilon})_{s \geq 0}$ . Hence, the left-hand side of (2.12) reads as

$$\begin{aligned} & \mathbb{E}[(\nabla^{\varepsilon,i} u_t^{\varepsilon}(x))^2] \\ &= (P_t^{\varepsilon}(\nabla^{\varepsilon,i} u_0^{\varepsilon})(x))^2 + \mathbb{E}\left[\left(\sum_{j=1}^d \sum_{y \in \mathbb{T}_{\varepsilon}^d} \int_0^t b_{t-s}^{\varepsilon,i,j}(x,y) (d\bar{N}_s^{\varepsilon,j}(y) \nabla^{\varepsilon,j} u_{s-}^{\varepsilon}(y))\right)^2\right] \\ &= (P_t^{\varepsilon}(\nabla^{\varepsilon,i} u_0^{\varepsilon})(x))^2 + \sum_{j=1}^d \sum_{y \in \mathbb{T}_{\varepsilon}^d} \int_0^t (b_{t-s}^{\varepsilon,i,j}(x,y))^2 \varepsilon^{-2} \mathbb{E}[(\nabla^{\varepsilon,j} u_s^{\varepsilon}(y))^2] ds \\ &= (P_t^{\varepsilon}(\nabla^{\varepsilon,i} u_0^{\varepsilon})(x))^2 + \sum_{j=1}^d \sum_{y \in \mathbb{T}_{\varepsilon}^d} \int_0^t q_{t-s}^{\varepsilon,i,j}(x,y) \mathbb{E}[(\nabla^{\varepsilon,j} u_s^{\varepsilon}(y))^2] ds, \end{aligned}$$

where for the second step we used the fact that  $\bar{N}^{\varepsilon}$  is a family of i.i.d. martingales, each with predictable quadratic variation given by  $s \mapsto \varepsilon^{-2} s$ , whereas the third step employs the definition in (2.5). Hence, letting  $U_t^{\varepsilon,i}(x) := \mathbb{E}[(\nabla^{\varepsilon,i} u_t^{\varepsilon}(x))^2]$  and recalling (2.10), we may rewrite the above identity as

$$U_t^{\varepsilon,i}(x) = \Pi_t^{\varepsilon,i,k}(u_0^{\varepsilon})(x) + \sum_{j=1}^d \sum_{y \in \mathbb{T}_{\varepsilon}^d} \int_0^t q_{t-s}^{\varepsilon,i,j}(x,y) U_s^{\varepsilon,j}(y) ds,$$

which, by non-negativity of all three functions involved, may be iterated infinitely often, yielding (2.12), which *a priori* may be infinite. However, recalling (2.10) and (2.8), we first obtain

$$\sup_{t \geq 0} \|\Pi_t^{\varepsilon,j,0}(u_0^{\varepsilon})\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)} \leq \|\nabla^{\varepsilon,j} u_0^{\varepsilon}\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)}^2 < \infty, \quad j = 1, \dots, d,$$

and, then,

$$\mathbb{E}[(\nabla^{\varepsilon,i} u_0^{\varepsilon}(x))^2] \leq \max_{j=1, \dots, d} \|\nabla^{\varepsilon,j} u_0^{\varepsilon}\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)}^2 \sum_{k=0}^{\infty} 2^{-k} = 2 \max_{j=1, \dots, d} \|\nabla^{\varepsilon,j} u_0^{\varepsilon}\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)}^2 < \infty.$$

This proves the desired claim when  $u_0^{\varepsilon} = v_0^{\varepsilon}$ , i.e., (2.12). When  $u_0^{\varepsilon} \neq v_0^{\varepsilon}$ , we similarly have

$$\begin{aligned} & \mathbb{E}[\nabla^{\varepsilon,i} u_t^{\varepsilon}(x) \nabla^{\varepsilon,i} v_t^{\varepsilon}(x)] \\ &= \Pi_t^{\varepsilon,i,0}(u_0^{\varepsilon}, v_0^{\varepsilon})(x) + \sum_{j=1}^d \sum_{y \in \mathbb{T}_{\varepsilon}^d} \int_0^t q_{t-s}^{\varepsilon,i,j}(x,y) \mathbb{E}[\nabla^{\varepsilon,i} u_s^{\varepsilon}(x) \nabla^{\varepsilon,i} v_s^{\varepsilon}(x)] ds, \end{aligned}$$

which, thanks to

$$\sup_{t \geq 0} \|\Pi_t^{\varepsilon,j,0}(u_0^{\varepsilon}, v_0^{\varepsilon})\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)} \leq \|\nabla^{\varepsilon,j} u_0^{\varepsilon}\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)} \|\nabla^{\varepsilon,j} v_0^{\varepsilon}\|_{L^{\infty}(\mathbb{T}_{\varepsilon}^d)} < \infty, \quad j = 1, \dots, d,$$

and the identity in (2.8), may be iterated, yielding (2.9).  $\square$

**2.3. Technical estimates.** All lemmas in this section are required for the proof of Lemma 3.2 below. We start with the following upper bound on integrals of second moments of discrete gradients, which may be thought of as an annealed pseudo-contractivity bound in  $H^1(\mathbb{T}_{\varepsilon}^d)$ .

**Lemma 2.3.** *Recall the definition of  $\|\cdot\|_{H^1(\mathbb{T}_{\varepsilon}^d)}$  from (1.25). Then, for all  $u_0^{\varepsilon} \in \mathbb{R}^{\mathbb{T}_{\varepsilon}^d}$  and  $t > 0$ , we have*

$$\mathbb{E}[\|u_t^{\varepsilon}\|_{H^1(\mathbb{T}_{\varepsilon}^d)}^2] \leq 2 \|u_0^{\varepsilon}\|_{H^1(\mathbb{T}_{\varepsilon}^d)}^2. \quad (2.13)$$

*Proof.* By definition of  $\|\cdot\|_{H^1(\mathbb{T}_\varepsilon^d)}$  and Proposition 2.1, we have

$$\mathbb{E}[\|u_t^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2] = \sum_{k=0}^{\infty} \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) .$$

Hence, the desired claim in (2.13) follows from

$$\varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) \leq 2^{-k} \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 , \quad k \in \mathbb{N}_0 . \quad (2.14)$$

We start with the case  $k = 0$ . Recalling (2.10), we have

$$\begin{aligned} \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \Pi_t^{\varepsilon,i,0}(u_0^\varepsilon)(x) &= \sum_{i=1}^d \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} (P_t^\varepsilon(\nabla^{\varepsilon,i} u_0^\varepsilon)(x))^2 \\ &\leq \sum_{i=1}^d \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} P_t^\varepsilon(\nabla^{\varepsilon,i} u_0^\varepsilon)^2(x) \\ &= \sum_{i=1}^d \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} (\nabla^{\varepsilon,i} u_0^\varepsilon(x))^2 = \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 , \end{aligned} \quad (2.15)$$

where the second step follows by Jensen inequality, while the third step used the symmetry of  $P_t^\varepsilon$  in  $L^2(\mathbb{T}_\varepsilon^d)$ . This proves (2.14) for  $k = 0$ . For the general case  $k \geq 1$ , first recall (2.11). Then,

$$\begin{aligned} &\varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) \\ &= \int_{[0,t]_>^k} ds_1 \cdots ds_k \sum_{y_0, \dots, y_{k-1} \in \mathbb{T}_\varepsilon^d} \sum_{j_0, \dots, j_{k-1}=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \\ &\quad \times \left( \varepsilon^d \sum_{y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_k=1}^d \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon)(y_k) \right) \\ &= \int_{[0,t]_>^k} ds_1 \cdots ds_k \left( \prod_{\ell=1}^k \mathcal{Q}_\varepsilon(s_\ell - s_{\ell-1}) \right) \left( \varepsilon^d \sum_{y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_k=1}^d \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon)(y_k) \right) \\ &\leq \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \int_{[0,t]_>^k} ds_1 \cdots ds_k \left( \prod_{\ell=1}^k \mathcal{Q}_\varepsilon(s_\ell - s_{\ell-1}) \right) \\ &\leq \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \left( \int_0^\infty \mathcal{Q}_\varepsilon(s) ds \right)^k \\ &= 2^{-k} \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 , \end{aligned}$$

where the second step follows from (2.6)–(2.7), the third step used the inequality in (2.15), the fourth step used the non-negativity of  $\mathcal{Q}_\varepsilon$ , whereas the fifth step follows from (2.8). This proves (2.14) for  $k \geq 1$ , thus, concludes the proof of the lemma.  $\square$

The following lemma presents a second bound on the expectation of discrete gradients.

**Lemma 2.4.** *For all  $u_0^\varepsilon, v_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  and  $t > 0$ , we have*

$$\sum_{i=1}^d \mathbb{E}[\langle \nabla^{\varepsilon,i} u_t^\varepsilon \mid \nabla^{\varepsilon,i} v_t^\varepsilon \rangle_{L^2(\mathbb{T}_\varepsilon^d)}] \leq 2 \left( \max_{i=1,\dots,d} \|\nabla^{\varepsilon,i} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \right) \left( \sum_{i=1}^d \|\nabla^{\varepsilon,i} v_0^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right).$$

*Proof.* Recall (2.10). Since

$$\left| \Pi_t^{\varepsilon,i,0}(u_0^\varepsilon, v_0^\varepsilon)(x) \right| \leq \|\nabla^{\varepsilon,i} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} P_t^\varepsilon |\nabla^{\varepsilon,i} v_0^\varepsilon|(x),$$

we obtain

$$\begin{aligned} \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \left| \Pi_t^{\varepsilon,i,0}(u_0^\varepsilon, v_0^\varepsilon)(x) \right| &\leq \max_{i=1,\dots,d} \|\nabla^{\varepsilon,i} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \left( \varepsilon^d \sum_{i=1}^d \sum_{x \in \mathbb{T}_\varepsilon^d} P_t^\varepsilon |\nabla^{\varepsilon,i} v_0^\varepsilon|(x) \right) \\ &= \max_{i=1,\dots,d} \|\nabla^{\varepsilon,i} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \left( \sum_{i=1}^d \|\nabla^{\varepsilon,i} v_0^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right), \end{aligned} \quad (2.16)$$

where for the last step we used the invariance of  $P_t^\varepsilon$ , i.e.,  $\langle P_t^\varepsilon g \rangle_\varepsilon = \langle g \rangle_\varepsilon$ , for all  $g \in \mathbb{T}_\varepsilon^d$ . Finally, recalling (2.9), we get

$$\begin{aligned} \sum_{i=1}^d \mathbb{E}[\langle \nabla^{\varepsilon,i} u_t^\varepsilon \mid \nabla^{\varepsilon,i} v_t^\varepsilon \rangle_{L^2(\mathbb{T}_\varepsilon^d)}] &= \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d |\mathbb{E}[\nabla^{\varepsilon,i} u_t^\varepsilon(x) \nabla^{\varepsilon,i} v_t^\varepsilon(x)]| \\ &\leq \sum_{k=0}^{\infty} \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \left| \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon, v_0^\varepsilon)(x) \right| \\ &\leq \max_{i=1,\dots,k} \|\nabla^{\varepsilon,i} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \left( \sum_{i=1}^d \|\nabla^{\varepsilon,i} v_0^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right) \sum_{k=0}^{\infty} 2^{-k}, \end{aligned}$$

where for the last step we first used (2.16) and then (2.8). This concludes the proof of the lemma.  $\square$

The following estimate turns out to be useful for small times and concentrated initial conditions. In what follows, we let  $B_\sigma(x) \subset \mathbb{T}^d$  denote the open ball of radius  $\sigma \in (0, \frac{1}{2})$  around  $x \in \mathbb{T}^d$ , and exploit the following classical exit-time estimate for the random walk  $(X_t^\varepsilon)_{t \geq 0}$  (see, e.g., [DSPS21, Eq. (4.4)] and references therein): for all  $t \in (0, 1)$ ,  $x \in \mathbb{T}_\varepsilon^d$  and  $\rho \in (0, \frac{1}{2})$ ,

$$\sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} p_t^\varepsilon(x, y) \leq \mathbf{P}_x^\varepsilon \left( \sup_{s \in [0, t]} |X_s^\varepsilon - x| > \rho \right) \leq c_1 \exp \left( -\frac{c_2 \rho}{t^{1/2} \vee \varepsilon} \right), \quad (2.17)$$

for some constants  $c_1, c_2 > 0$  depending only on  $d \geq 1$ .

**Lemma 2.5.** *Let the initial condition  $u_0^\varepsilon \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  be vanishing outside  $B_\sigma(x)$ , for some  $\sigma \in (0, \frac{1}{2})$  and  $x \in \mathbb{T}_\varepsilon^d$ . Then, for all  $t \in (0, 1)$ ,  $\rho \in (\sigma + \varepsilon, \frac{1}{2})$  and  $n \in \mathbb{N}$ , we have*

$$\sum_{i=1}^d \mathbb{E}[\|\mathbf{1}_{B_\rho^c(x)} \nabla^{\varepsilon,i} u_t^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2] \leq C \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \varepsilon^{-2} \left\{ \varepsilon^{-2} \exp \left( -\frac{C'(\rho - \sigma)}{n(t^{1/2} \vee \varepsilon)} \right) + 2^{-n} \right\}, \quad (2.18)$$

for some constants  $C, C' > 0$  depending only on  $d \geq 1$ .

*Proof.* Recall (2.12). Then, for every integer  $n \geq 1$ , the left-hand side of (2.18) reads as

$$\begin{aligned}
\sum_{i=1}^d \mathbb{E}[\|\mathbb{1}_{B_\rho^c(x)} \nabla^{\varepsilon,i} u_t^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2] &= \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \mathbb{E}[(\nabla^{\varepsilon,i} u_t^\varepsilon(y))^2] \\
&= \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \sum_{k=0}^n \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \sum_{k=n+1}^{\infty} \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) \\
&\leq \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \sum_{k=0}^n \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) + \sum_{k=n+1}^{\infty} \left( \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) \right) \\
&\leq \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \sum_{k=0}^n \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) + 2^{-n} \|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2, \tag{2.19}
\end{aligned}$$

where the last step used (2.14) to bound the expression in parenthesis, and  $2^{-n} = \sum_{k=n+1}^{\infty} 2^{-k}$ . Recall (2.10). By Jensen inequality,  $\text{supp}(u_0^\varepsilon) \subset B_\sigma(x)$ , and the crude bound

$$\max_{j=1,\dots,d} \|\nabla^{\varepsilon,j} u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq 2\varepsilon^{-1} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)},$$

we get, for all  $s \in (0, 1)$ ,  $j = 1, \dots, d$ , and  $y \in \mathbb{T}_\varepsilon^d$ ,

$$\Pi_s^{\varepsilon,j,0}(u_0^\varepsilon)(y) \leq 2\varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \sum_{\substack{z \in \mathbb{T}_\varepsilon^d \\ |x-z| < \sigma + \varepsilon}} p_s^\varepsilon(y, z). \tag{2.20}$$

Thanks to the exit-time estimate (2.17), we have, for  $y \in \mathbb{T}_\varepsilon^d$  satisfying  $|x - y| > \sigma + \varepsilon$ ,

$$\Pi_s^{\varepsilon,j,0}(u_0^\varepsilon)(y) \leq 2\varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \left\{ c_1 \exp\left(-\frac{c_2(|x-y| - \sigma - \varepsilon)}{s^{1/2} \sqrt{\varepsilon}}\right) \right\}. \tag{2.21}$$

As a consequence, we obtain, for  $k = 0$  and for some  $C_0 = C_0(d) > 0$ ,

$$\varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \Pi_t^{\varepsilon,i,0}(u_0^\varepsilon)(y) \leq C_0 \varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \exp\left(-\frac{c_2(\rho - \sigma - \varepsilon)}{t^{1/2} \sqrt{\varepsilon}}\right). \tag{2.22}$$

For  $k = 1, \dots, n$ , recalling (2.11), we have

$$\begin{aligned}
&\varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(y) \\
&= \varepsilon^d \int_{[0,t]_{>}^k} ds_1 \cdots ds_k \sum_{\substack{y_0, \dots, y_k \in \mathbb{T}_\varepsilon^d \\ |y_0 - x| > \rho}} \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon)(y_k).
\end{aligned}$$

By splitting the summation over  $y_k \in \mathbb{T}_\varepsilon^d$  into  $\{y_k \in \mathbb{T}_\varepsilon^d : |x - y_k| < \sigma + \frac{\rho - \sigma - \varepsilon}{2n}\}$  and its complement, the estimates in (2.20) and (2.21) yield

$$\begin{aligned} \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)(y) &\leq 2\varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \\ &\times \left( I_t^{\varepsilon, k}(\sigma + \frac{\rho - \sigma - \varepsilon}{2n}, \rho) + c_1 \exp\left(-\frac{c_2(\frac{\rho - \sigma - \varepsilon}{2n} - \varepsilon)}{t^{1/2} \vee \varepsilon}\right) J_t^{\varepsilon, k}(\sigma + \frac{\rho - \sigma - \varepsilon}{2n}, \rho) \right), \end{aligned} \quad (2.23)$$

where, for all  $0 \leq \rho_1 \leq \rho_2 \leq \frac{1}{2}$ ,

$$I_t^{\varepsilon, k}(\rho_1, \rho_2) := \int_{[0, t]^k} ds_1 \cdots ds_k \sum_{\substack{y_0, \dots, y_k \in \mathbb{T}_\varepsilon^d \\ |y_0| \geq \rho_2, |y_k| < \rho_1}} \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right), \quad (2.24)$$

$$J_t^{\varepsilon, k}(\rho_1, \rho_2) := \int_{[0, t]^k} ds_1 \cdots ds_k \sum_{\substack{y_0, \dots, y_k \in \mathbb{T}_\varepsilon^d \\ |y_0| \geq \rho_2, |y_k| \geq \rho_1}} \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right). \quad (2.25)$$

Note that we used the translation invariance of the transition probabilities of  $(X_t^\varepsilon)_{t \geq 0}$  to get rid of the dependence on  $x \in \mathbb{T}_\varepsilon^d$ . Let us extend the above definitions to the case  $k = 0$  as follows:

$$I_t^{\varepsilon, 0}(\rho_1, \rho_2) := 0, \quad J_t^{\varepsilon, 0}(\rho_1, \rho_2) := 1. \quad (2.26)$$

We first estimate  $J_t^{\varepsilon, k}$  in (2.25). By (2.8), we readily obtain, for all  $0 \leq \rho_1 \leq \rho_2 \leq \frac{1}{2}$ ,

$$J_t^{\varepsilon, k}(\rho_1, \rho_2) \leq J_t^{\varepsilon, k}(0, 0) \leq d \varepsilon^{-d} 2^{-k}. \quad (2.27)$$

We now deal with  $I_t^{\varepsilon, k}$  in (2.24), for all  $k \leq n \in \mathbb{N}$ . Define

$$\lambda = \lambda(\rho_1, \rho_2, n) := \frac{\rho_2 - \rho_1}{2n}.$$

Recalling the definitions (2.4) and (2.5) and the exit-time estimate (2.17), we have, for all  $t \in (0, 1)$  and  $0 \leq \rho_1 \leq \rho_2 \leq \frac{1}{2}$ ,

$$\max_{i=1, \dots, d} \sup_{\substack{z \in \mathbb{T}_\varepsilon^d \\ |x-z| \geq \rho_2}} \int_0^t \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| < \rho_1}} \sum_{j=1}^d q_s^{\varepsilon, i, j}(z, y) ds \leq C_1 \varepsilon^{-2} t \exp\left(-\frac{c_2(\rho_2 - \rho_1)}{t^{1/2} \vee \varepsilon}\right), \quad (2.28)$$

for some  $C_1 = C_1(d) > 0$ . Analogously, by splitting the summation over  $y_{k-1} \in \mathbb{T}_\varepsilon^d$  into  $\{y_{k-1} \in \mathbb{T}_\varepsilon^d : |y_{k-1}| < \rho_1 + \lambda\}$  and its complement, we obtain

$$\begin{aligned} I_t^{\varepsilon, k}(\rho_1, \rho_2) &= \int_{[0, t]^{k-1}} ds_1 \cdots ds_{k-1} \sum_{\substack{y_0, \dots, y_{k-1} \in \mathbb{T}_\varepsilon^d \\ |y_0| > \rho_2}} \sum_{j_0, \dots, j_{k-1}=1}^d \left( \prod_{\ell=1}^{k-1} q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \\ &\quad \times \int_0^{s_{k-1}} \sum_{\substack{y_k \in \mathbb{T}_\varepsilon^d \\ |y_k| < \rho_1}} \sum_{j_k=1}^d q_{s_{k-1}-s_k}^{\varepsilon, j_{k-1}, j_k}(y_{k-1}, y_k) ds_k \\ &\leq 2^{-1} I_t^{\varepsilon, k-1}(\rho_1 + \lambda, \rho_2) + C_1 \varepsilon^{-2} t \exp\left(-\frac{c_2 \lambda}{t^{1/2} \vee \varepsilon}\right) J_t^{\varepsilon, k-1}(\rho_1 + \lambda, \rho_2) \end{aligned}$$

$$\leq 2^{-1} I_t^{\varepsilon, k-1}(\rho_1 + \lambda, \rho_2) + C_2 2^{-k} \varepsilon^{-(d+2)} t \exp\left(-\frac{c_2 \lambda}{t^{1/2} \sqrt{\varepsilon}}\right),$$

where for the first inequality we used (2.8) and (2.28), while for the second one we used (2.27). Here,  $C_2 = C_2(d) > 0$ . Finally, noting that  $\rho_2 - (\rho_1 + n\lambda) > \lambda$  and recalling (2.26), we may iterate the above inequality, so to obtain

$$I_t^{\varepsilon, k}(\rho_1, \rho_2) \leq C_2 k 2^{-k} \varepsilon^{-(d+2)} t \exp\left(-\frac{c_2(\rho_2 - \rho_1)}{2n(t^{1/2} \sqrt{\varepsilon})}\right), \quad \text{for all } k \leq n. \quad (2.29)$$

By combining the estimates in (2.22), (2.23), (2.27) and (2.29), we get, for some  $C_3, C_4, C_5 > 0$  depending only on  $d \geq 1$ ,

$$\begin{aligned} & \sum_{k=0}^n \left( \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{i=1}^d \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)(y) \right) \\ & \leq C_4 \varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \exp\left(-\frac{C_3(\rho - \sigma - \varepsilon)}{2n(t^{1/2} \sqrt{\varepsilon})}\right) \left( 1 + \varepsilon^{-2} t \sum_{k=1}^n k 2^{-k} + \sum_{k=1}^n 2^{-k} \right) \\ & \leq C_5 \varepsilon^{-4} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \exp\left(-\frac{C_3(\rho - \sigma - \varepsilon)}{2n(t^{1/2} \sqrt{\varepsilon})}\right), \end{aligned}$$

where for the last step we used  $t < 1$ . Inserting this bound into (2.19), and estimating  $\|u_0^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \leq C_6 \varepsilon^{-2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2$ , for some  $C_6 = C_6(d) > 0$ , we get the desired result.  $\square$

### 3. PROOF OF THEOREM 2. VARIANCE VIA POINCARÉ INEQUALITY

The main goal of this section is to provide a quantitative control of the right-hand side of (1.20) — thus, of the variance of  $\Gamma_t^\varepsilon(g)$  — and prove that it vanishes as  $\varepsilon \rightarrow 0$ . We fix  $t > 0$ ,  $g = (g^i)_{i=1, \dots, d} \in (\mathcal{C}([0, \infty) \times \mathbb{T}^d))^d$ , and the initial conditions  $u_0 \in \mathcal{C}(\mathbb{T}^d)$  and  $u_0^\varepsilon = u_0|_{\mathbb{T}_\varepsilon^d} \in \mathbb{R}^{\mathbb{T}_\varepsilon^d}$  all throughout the section. Finally,  $\|\cdot\|_\infty$  indicates uniform norms, e.g.,

$$\|g\|_\infty := \max_{i=1, \dots, d} \sup_{t \geq 0} \sup_{x \in \mathbb{T}^d} |g^i(t, x)| \quad \text{and} \quad \|u_0\|_\infty := \sup_{x \in \mathbb{T}^d} |u_0(x)|. \quad (3.1)$$

**Theorem 3** (Variance). *Let  $\Gamma_t^\varepsilon(g)$  be given as in (1.17). Then, for all  $h \in (0, 1)$ ,*

$$\varepsilon^{-(d+2)} \int_0^\infty \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E}[(D_{(s, x, i)} \Gamma_t^\varepsilon(g))^2] ds \leq C \|g\|_\infty^2 \|u_0\|_\infty^4 (\varepsilon^{d/2})^{(1-h)\frac{d}{d+2}}, \quad (3.2)$$

for some  $C = C(d, h) > 0$ .

We break the proof of Theorem 3 into steps. We start by estimating a conditional version of the expectation on the left-hand side of (3.2).

**Proposition 3.1.** *We have,  $\mathbb{P}$ -a.s., for all  $s \in (0, t)$ ,  $x \in \mathbb{T}_\varepsilon^d$ ,  $i = 1, \dots, d$ , and  $h \in (0, 1)$ ,*

$$\mathbb{E}[(D_{(s, x, i)} \Gamma_t^\varepsilon(g))^2 | \mathcal{F}_s^\varepsilon] \leq C \varepsilon^{d+2} \|g\|_\infty^2 \|u_0\|_\infty^2 (\nabla^{\varepsilon, i} u_s^\varepsilon(x))^2 (\varepsilon^{d/2})^{(1-h)\frac{d}{d+2}},$$

where  $C = C(d, h) > 0$ .

Before presenting the proof of Proposition 3.1, we use it to prove Theorem 3.

*Proof of Theorem 3.* The tower property and Proposition 3.1 yield the following bound on the left-hand side of (3.2):

$$\begin{aligned} & \varepsilon^{-(d+2)} \int_0^\infty \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E} \left[ \mathbb{E} \left[ (D_{(s,x,i)} \Gamma_t^\varepsilon(g))^2 \mid \mathcal{F}_s^\varepsilon \right] \right] ds \\ & \leq C \|g\|_\infty^2 \|u_0\|_\infty^2 (\varepsilon^{d/2})^{(1-h)\frac{d}{d+2}} \int_0^\infty \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \mathbb{E} \left[ (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2 \right] ds . \end{aligned}$$

The desired claim now follows by Aldous-Lanoue identity (1.26), which allows to write the above time integral as

$$2 \|u_0^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \leq 2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 .$$

This concludes the proof of the theorem.  $\square$

The rest of this section is devoted to the proof of Proposition 3.1.

**3.1. Dynamics of the discrepancy.** Let us adopt the following shorthand notation: for all  $s \in (0, t)$ ,  $x \in \mathbb{T}_\varepsilon^d$ , and  $i = 1, \dots, d$ ,

$$\begin{aligned} \Lambda_{(s,x,i)}^\varepsilon & := (D_{(s,x,i)} \Gamma_t^\varepsilon(g))^2 \\ & = \left( \int_s^t \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d \left\{ (\nabla^{\varepsilon,j} u_r^\varepsilon(y) + \nabla^{\varepsilon,j} w_r^{\varepsilon,(s,x,i)}(y))^2 - (\nabla^{\varepsilon,j} u_r^\varepsilon(y))^2 \right\} g_r^j(y) dr \right)^2 , \end{aligned} \quad (3.3)$$

where we recall from (1.21) that  $(w_r^{\varepsilon,(s,x,i)})_{r \geq 0}$  denotes the discrepancy created by the extra update at time  $s$  at the vertices  $x$  and  $x + \varepsilon e_i$ . From now on, we shall drop  $(s, x, i)$  from the notation, and simply write, e.g.,

$$\Lambda^\varepsilon = \Lambda_{(s,x,i)}^\varepsilon \quad \text{and} \quad w_r^\varepsilon = w_r^{\varepsilon,(s,x,i)} .$$

Next, we split the expression in (3.3) into three terms: for all  $\delta \in (\varepsilon^2, 1)$ ,

$$\begin{aligned} \Lambda^\varepsilon & = \left( \int_s^t \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d \left\{ (\nabla^{\varepsilon,j} w_r^{\varepsilon,(s,x,i)}(y))^2 + 2 \nabla^{\varepsilon,j} w_r^{\varepsilon,(s,x,i)}(y) \nabla^{\varepsilon,j} u_r^\varepsilon(y) \right\} g_r^j(y) dr \right)^2 \\ & \leq 2 \left( \int_s^t \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d (\nabla^{\varepsilon,j} w_r^\varepsilon(y))^2 g_r^j(y) dr \right)^2 \\ & \quad + 16 \left( \int_s^{(s+\delta) \wedge t} \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d \nabla^{\varepsilon,j} w_r^\varepsilon(y) \nabla^{\varepsilon,j} u_r^\varepsilon(y) g_r^j(y) dr \right)^2 \\ & \quad + 16 \left( \int_{(s+\delta) \wedge t}^t \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d \nabla^{\varepsilon,j} w_r^\varepsilon(y) \nabla^{\varepsilon,j} u_r^\varepsilon(y) g_r^j(y) dr \right)^2 \\ & =: 2 \Lambda_1^\varepsilon + 16 \Lambda_2^{\varepsilon,\delta} + 16 \Lambda_3^{\varepsilon,\delta} , \end{aligned} \quad (3.4)$$

where we used twice the elementary inequality  $(a + b)^2 \leq 2a^2 + 2b^2$ ,  $a, b \in \mathbb{R}$ . Our task is to analyze the dynamics of the discrepancy  $(w_r^\varepsilon)_{r \geq s}$ , bounding the conditional expectation of each of these three terms. For this purpose, let us observe that, while  $w_r^\varepsilon \equiv 0$  for  $r < s$ , for  $r = s$  we have

$$w_s^\varepsilon = \frac{\varepsilon}{2} \nabla^{\varepsilon,i} u_s^\varepsilon(x) (\mathbb{1}_x - \mathbb{1}_{x+\varepsilon e_i}) , \quad (3.5)$$

from which we obtain

$$\|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 = \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} (w_s^\varepsilon(y))^2 = \frac{1}{2} \varepsilon^{d+2} (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2. \quad (3.6)$$

**3.2. Proof strategy.** We estimate the three terms in (3.4) in three separate lemmas in the subsequent section. As an overview of the proof strategy, we observe that all three lemmas fundamentally exploit the time integrals. In this way, we avoid to estimate fourth moments of discrete gradients, for which we are not able to recover convergent recursive inequalities, as done for the second moments.

In this spirit, the first term  $A_1^\varepsilon$  is handled rather easily by squaring the time integral, applying Cauchy-Schwarz, and gain  $L^2$ -norms of  $w_s^\varepsilon$  from the time integration and Aldous-Lanoue identity (1.26). The resulting upper bound depends on the discrepancy only via  $\|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^4 \approx \varepsilon^{2(d+2)} \ll \varepsilon^{d+2}$  (see (3.6)), which therefore suffices. The third term  $A_3^{\varepsilon,\delta}$ , although it contains both  $w_r^\varepsilon$  and  $u_r^\varepsilon$ , is dealt with analogously. Nevertheless, if we were to take  $\delta = 0$  (as in  $A_1^\varepsilon$ ), this strategy would yield  $\|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)} \|u_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \approx \varepsilon^{d+2}$ , which is too poor for our purposes. This explains the necessity of introducing a small burn-in time of size  $\varepsilon^2 \ll \delta \ll \varepsilon$ , producing a regularization effect of the Dirac-like discrepancy  $w_s^\varepsilon$ . This smoothening is quantified in terms of the ultracontractivity of the averaging process (1.22).

The second term  $A_2^{\varepsilon,\delta}$  concerns the remaining part of the time integral left over by  $A_3^{\varepsilon,\delta}$ , namely, from time  $r = s$  to  $r = s + \delta$ . This term is the most delicate. Instead of relying on Cauchy-Schwarz inequality, this time we must leverage the fact that, for small times and when starting from highly concentrated data,  $L^1$ -norms display a better decay than  $L^2$ -norms. In order to turn this observation into effective bounds, we crucially employ: (i) a localization argument (Lemma 2.5, with  $\delta^{1/2} \ll \rho \ll 1$ ); (ii) pseudo-contractivity estimates for discrete gradients (Lemmas 2.3 and 2.4).

**3.3. Proofs.** We split the proof of Proposition 3.1 into three lemmas, one for each term in (3.4). We start with the first term in (3.4). Here,  $C_1 = C_1(d) > 0$ .

**Lemma 3.1** (Estimate of  $A_1^\varepsilon$ ). *We have*

$$\mathbb{E}[A_1^\varepsilon \mid \mathcal{F}_s^\varepsilon] \leq C_1 \varepsilon^{2d+2} \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2.$$

*Proof.* Since

$$A_1^\varepsilon \leq 2 \|g\|_\infty^2 \int_s^t \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \int_r^t \|w_{r'}^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 dr' dr,$$

the tower property and Aldous-Lanoue identity in (1.26) yield

$$\begin{aligned} \mathbb{E}[A_1^\varepsilon \mid \mathcal{F}_s^\varepsilon] &\leq 2 \|g\|_\infty^2 \int_s^t \mathbb{E} \left[ \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \int_r^t -2 \frac{d}{dr'} \mathbb{E} \left[ \|w_{r'}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_r^\varepsilon \right] dr' \mid \mathcal{F}_s^\varepsilon \right] dr \\ &\leq 4 \|g\|_\infty^2 \int_s^t \mathbb{E} \left[ \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \|w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon \right] dr \\ &\leq 4 \|g\|_\infty^2 \|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \int_s^t \mathbb{E} \left[ \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon \right] dr \\ &\leq 8 \|g\|_\infty^2 \|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^4, \end{aligned}$$

where the third step used that  $r \mapsto \|w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2$  is deterministically non-increasing for  $r \geq s$ . Finally, we obtain the desired inequality by (3.6), which yields

$$\|w_s^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^4 = 2^{-2} \varepsilon^{2(d+2)} (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^4$$

$$\begin{aligned} &\leq \varepsilon^{2d+2} \|u_s^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2 \\ &\leq \varepsilon^{2d+2} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2, \end{aligned}$$

where the last step follows by the  $\mathbb{P}$ -a.s. monotonicity of  $L^\infty(\mathbb{T}_\varepsilon^d)$ -norms for the averaging process.  $\square$

The estimate of the second term in (3.4) is the most involved one. Here,  $C_2, C'_2 > 0$  depend only on  $d \geq 1$ .

**Lemma 3.2** (Estimate of  $A_2^\varepsilon$ ). *We have, for all  $\delta \in (\varepsilon^2, 1)$ ,  $\rho \in (4\varepsilon, \frac{1}{2})$  and  $n \in \mathbb{N}$ ,*

$$\mathbb{E}[A_2^{\varepsilon,\delta} \mid \mathcal{F}_s^\varepsilon] \leq C_2 T^{\varepsilon,\delta,\rho,n} \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2,$$

where

$$T^{\varepsilon,\delta,\rho,n} := \varepsilon^{-2} \delta^2 \left\{ (\rho\varepsilon)^d + \varepsilon^{-2} \exp\left(-C'_2 \frac{\rho}{n \delta^{1/2}}\right) + 2^{-n} \right\}.$$

*Proof.* By the tower property and

$$\max_{j=1,\dots,d} \|\nabla^{\varepsilon,j} u_r^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq 2\varepsilon^{-1} \|u_r^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq 2\varepsilon^{-1} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}, \quad r \geq 0,$$

we get

$$\mathbb{E}[A_2^{\varepsilon,\delta} \mid \mathcal{F}_s^\varepsilon] \leq 4\varepsilon^{-1} \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \int_s^{s+\delta} \sum_{j=1}^d \mathbb{E}[\|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} Z_r^{\varepsilon,\delta} \mid \mathcal{F}_s^\varepsilon] dr, \quad (3.7)$$

where

$$Z_r^{\varepsilon,\delta} := \int_r^{s+\delta} \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d |\mathbb{E}[\nabla^{\varepsilon,j} w_{r'}^\varepsilon(y) \nabla^{\varepsilon,j} u_{r'}^\varepsilon(y) \mid \mathcal{F}_r^\varepsilon]| dr', \quad r \in (s, s+\delta). \quad (3.8)$$

We apply Lemma 2.4 to the integrand in (3.8), so to obtain

$$\begin{aligned} Z_r^{\varepsilon,\delta} &\leq 2\delta \left( \max_{j=1,\dots,d} \|\nabla^{\varepsilon,j} u_r^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \right) \left( \sum_{j=1}^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right) \\ &\leq 4\delta (\varepsilon^{-1} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}) \left( \sum_{j=1}^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right). \end{aligned}$$

By combining this estimate with (3.7), we get

$$\mathbb{E}[A_2^{\varepsilon,\delta} \mid \mathcal{F}_s^\varepsilon] \leq 16\varepsilon^{-2} \delta \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \int_s^{s+\delta} \mathbb{E} \left[ \left( \sum_{j=1}^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right)^2 \mid \mathcal{F}_s^\varepsilon \right] dr. \quad (3.9)$$

We now split the  $L^1$ -norm above as follows: for any  $\rho \in (4\varepsilon, \frac{1}{2})$  and  $j = 1, \dots, d$ ,

$$\begin{aligned} \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} &= \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| \leq \rho}} |\nabla^{\varepsilon,j} w_r^\varepsilon(y)| + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} |\nabla^{\varepsilon,j} w_r^\varepsilon(y)| \\ &= \|\mathbf{1}_{\bar{B}_\rho(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} + \|\mathbf{1}_{B_\rho^c(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}, \end{aligned}$$

from which we obtain, by Cauchy-Schwarz inequality,

$$\left( \sum_{j=1}^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right)^2 \leq C'_1 \sum_{j=1}^d \left\{ \|\mathbf{1}_{\bar{B}_\rho(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 + \|\mathbf{1}_{B_\rho^c(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 \right\}$$

$$\begin{aligned}
&\leq C'_1 \sum_{j=1}^d \left\{ \|\mathbb{1}_{\bar{B}_\rho(x)}\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 + \|\mathbb{1}_{B_\rho^\varepsilon(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \right\} \\
&\leq C'_2 \sum_{j=1}^d \left\{ \rho^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 + \|\mathbb{1}_{B_\rho^\varepsilon(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \right\} \\
&= C'_2 \left\{ \rho^d \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{j=1}^d (\nabla^{\varepsilon,j} w_r^\varepsilon(y))^2 \right\},
\end{aligned}$$

for some constants  $C'_1, C'_2 > 0$  depending only on  $d \geq 1$ . Taking expectations, the above estimate yields

$$\begin{aligned}
&\mathbb{E} \left[ \left( \sum_{j=1}^d \|\nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)} \right)^2 \middle| \mathcal{F}_s^\varepsilon \right] \\
&\leq C'_2 \left\{ \rho^d \mathbb{E}[\|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon] + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{j=1}^d \mathbb{E}[(\nabla^{\varepsilon,j} w_r^\varepsilon(y))^2 \mid \mathcal{F}_s^\varepsilon] \right\} \\
&\leq C'_2 \left\{ 2\rho^d \|w_s^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{j=1}^d \mathbb{E}[(\nabla^{\varepsilon,j} w_r^\varepsilon(y))^2 \mid \mathcal{F}_s^\varepsilon] \right\} \\
&= C'_2 \left\{ 2\rho^d c \varepsilon^d (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2 + \varepsilon^d \sum_{\substack{y \in \mathbb{T}_\varepsilon^d \\ |x-y| > \rho}} \sum_{j=1}^d \mathbb{E}[(\nabla^{\varepsilon,j} w_r^\varepsilon(y))^2 \mid \mathcal{F}_s^\varepsilon] \right\}, \quad (3.10)
\end{aligned}$$

where the second step follows from Lemma 2.3, while the third step used (3.5), which ensures that, for some  $c = c(d) > 0$ ,

$$\|w_s^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 = \varepsilon^d \sum_{y \in \mathbb{T}_\varepsilon^d} \sum_{j=1}^d (\nabla^{\varepsilon,j} w_s^\varepsilon(y))^2 = c \varepsilon^d (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2.$$

Observe that  $w_s^\varepsilon$  is non-zero only in a  $2\varepsilon$ -neighborhood of  $x \in \mathbb{T}_\varepsilon^d$ . Moreover, by (3.5),

$$\|w_s^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 = \frac{\varepsilon^2}{2} (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2.$$

Henceforth, Lemma 2.5 applied to  $(w_r^\varepsilon)_{r \geq s}$  yields, for all  $n \in \mathbb{N}$  and  $r \in (s, s + \delta)$ ,

$$\sum_{j=1}^d \mathbb{E}[\|\mathbb{1}_{B_\rho^\varepsilon(x)} \nabla^{\varepsilon,j} w_r^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon] \leq C (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2 \left\{ \varepsilon^{-2} \exp\left(-\frac{C' \rho}{n \delta^{1/2}}\right) + 2^{-n} \right\}. \quad (3.11)$$

By combining (3.9), (3.10) and (3.11), we get the desired estimate.  $\square$

We now bound the third term in (3.4). Here,  $C_3 = C_3(d) > 0$ .

**Lemma 3.3** (Estimate of  $\Lambda_3^{\varepsilon,\delta}$ ). *We have, for all  $\delta \in (\varepsilon^2, 1)$ ,*

$$\mathbb{E}[\Lambda_3^{\varepsilon,\delta} \mid \mathcal{F}_s^\varepsilon] \leq C_3 \varepsilon^{2d+2} \delta^{-d/2} \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon,i} u_s^\varepsilon(x))^2.$$

*Proof.* Since  $A_3^{\varepsilon, \delta} = 0$  whenever  $s + \delta \geq t$ , let us assume  $s + \delta < t$ . Then, by repeatedly applying Cauchy-Schwarz inequality, we get

$$\begin{aligned} & \mathbb{E}[A_3^{\varepsilon, \delta} \mid \mathcal{F}_{s+\delta}^\varepsilon] \\ & \leq \|g\|_\infty^2 \mathbb{E} \left[ \left( \int_{s+\delta}^t \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)} \|u_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)} \, dr \right)^2 \middle| \mathcal{F}_{s+\delta}^\varepsilon \right] \\ & \leq \|g\|_\infty^2 \mathbb{E} \left[ \left( \int_{s+\delta}^t \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \, dr \right) \left( \int_{s+\delta}^t \|u_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \, dr \right) \middle| \mathcal{F}_{s+\delta}^\varepsilon \right] \\ & \leq \|g\|_\infty^2 \mathbb{E} \left[ \left( \int_{s+\delta}^t \|w_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \right)^2 \middle| \mathcal{F}_{s+\delta}^\varepsilon \right]^{\frac{1}{2}} \mathbb{E} \left[ \left( \int_{s+\delta}^t \|u_r^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 \right)^2 \middle| \mathcal{F}_{s+\delta}^\varepsilon \right]. \end{aligned}$$

By estimating both expectations as already done in the proof of Lemma 3.1, we get

$$\begin{aligned} \mathbb{E}[A_3^{\varepsilon, \delta} \mid \mathcal{F}_{s+\delta}^\varepsilon] & \leq 8 \|g\|_\infty^2 \|w_{s+\delta}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \|u_{s+\delta}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \\ & \leq 8 \|g\|_\infty^2 \|w_{s+\delta}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2. \end{aligned}$$

By the tower property, we obtain

$$\mathbb{E}[A_3^{\varepsilon, \delta} \mid \mathcal{F}_s^\varepsilon] = \mathbb{E}[\mathbb{E}[A_3^{\varepsilon, \delta} \mid \mathcal{F}_{s+\delta}^\varepsilon] \mid \mathcal{F}_s^\varepsilon] \leq 8 \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \mathbb{E}[\|w_{s+\delta}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon].$$

Since  $\langle w_s^\varepsilon \rangle_\varepsilon = 0$  (cf. (3.5)), the last expectation may be further estimated thanks to (1.22):

$$\mathbb{E}[\|w_{s+\delta}^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2 \mid \mathcal{F}_s^\varepsilon] \leq C \|w_s^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 \delta^{-d/2}.$$

Recalling (3.5), we have

$$\|w_s^\varepsilon\|_{L^1(\mathbb{T}_\varepsilon^d)}^2 = \varepsilon^{2d+2} (\nabla^{\varepsilon, i} u_s^\varepsilon(x))^2,$$

and, thus, the desired result.  $\square$

We conclude this section with the proof of Proposition 3.1.

*Proof of Proposition 3.1.* We exhibit a choice of  $\delta = \delta(\varepsilon) \in (\varepsilon^2, 1)$  yielding

$$\mathbb{E}[A_\ell^{\varepsilon, \delta} \mid \mathcal{F}_s^\varepsilon] \leq C \varepsilon^{d+2} \|g\|_\infty^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 (\nabla^{\varepsilon, i} u_s^\varepsilon(x))^2 V_\ell^\varepsilon, \quad \ell \in \{1, 2, 3\}, \quad (3.12)$$

for some  $V_\ell^\varepsilon \leq C' (\varepsilon^{d/2})^{(1-h)\frac{d}{d+2}}$ , for all  $h \in (0, 1)$  and some  $C' = C'(d, h) > 0$ . The inequality in (3.4) would conclude the proof of the proposition.

For what concerns  $\ell = 1$ , we have  $A_1^{\varepsilon, \delta} = A_1^\varepsilon$  and, by Lemma 3.1, we obtain (3.12) with  $V_1^\varepsilon = \varepsilon^{-(d+2)} \varepsilon^{2d+2} = \varepsilon^d$ . Consider the case  $\ell = 3$ . Letting  $\delta = \varepsilon^{2(1-a)}$ ,  $a \in (0, 1)$ , Lemma 3.3 yields (3.12) with

$$V_3^\varepsilon = \varepsilon^{-(d+2)} (\varepsilon^{2d+2} \delta^{-d/2}) = \varepsilon^{-d-2+2d+2-d+ad} = \varepsilon^{ad}.$$

By setting  $a = (1-h)\frac{1}{2}\frac{d}{d+2}$  for some small  $h \in (0, 1)$ , we obtain the desired claim for  $\ell = 3$ . Consider  $\ell = 2$ , and let  $\rho = \varepsilon^{1-b}$ , for  $b = \frac{1}{2}\frac{d}{d+2} \in (a, 1)$ , as well as  $n = \lceil \log \varepsilon \rceil^2$ . Then, by inserting these choices into the claim of Lemma 3.2, we obtain

$$\begin{aligned} V_2^\varepsilon & = \varepsilon^{-(d+2)} \varepsilon^{-2+4-4a} \left\{ \varepsilon^{2d-bd} + \varepsilon^{-2} \exp \left( -C_2' \frac{\varepsilon^{-(b-a)}}{\log^2 |\varepsilon|} \right) + 2^{-\log^2 |\varepsilon|} \right\} \\ & = \varepsilon^{-4a+d-bd} + \varepsilon^{-d-2-4a} \exp \left( -C_2' \frac{\varepsilon^{-(b-a)}}{\lceil \log \varepsilon \rceil^2} \right) + 2^{-\lceil \log \varepsilon \rceil^2}. \end{aligned} \quad (3.13)$$

Since  $b = \frac{1}{2} \frac{d}{d+2}$  and  $a = (1-h)b$ , the first term in (3.13) is smaller than  $(\varepsilon^{d/2})^{\frac{d}{d+2}} \leq V_3^\varepsilon$ . Observe that the choices  $b > a$  and  $n = \lceil \log \varepsilon \rceil^2$  ensure that the second and third terms in (3.13) are bounded above, uniformly over  $\varepsilon \in (0, 1)$ , by, e.g.,  $C''V_1^\varepsilon = C''\varepsilon^d$ , for some  $C'' = C''(d, h) > 0$ . This proves (3.12) for  $\ell = 2$  and, thus, concludes the proof.  $\square$

#### 4. PROOF OF THEOREM 2. MEAN

In view of Theorem 3 on the variance of  $\Gamma_t^\varepsilon(g)$ , the proof of Theorem 2 is complete as soon as we show convergence of the corresponding means. This fact is summarized in the following theorem. Also in this section, we fix a test function  $g = (g^i)_{i=1, \dots, d} \in (\mathcal{C}([0, \infty) \times \mathbb{T}^d))^d$ , as well as the initial condition  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ , and consider  $u_0^\varepsilon := u_0|_{\mathbb{T}_\varepsilon^d}$ .

**Theorem 4 (Mean).** *Let  $\Gamma_t^\varepsilon(g)$  and  $\Gamma_t(g)$  be given as in (1.17) and (1.18), respectively. Then, for all  $t > 0$ , we have*

$$\mathbb{E}[\Gamma_t^\varepsilon(g)] \xrightarrow{\varepsilon \rightarrow 0} \Gamma_t(g) .$$

We break the proof of Theorem 4 into two main steps: Propositions 4.1 and 4.2. We remark that, while the proof of Theorem 3 was crucially exploiting the time integration in the definition of  $\Gamma_t^\varepsilon(g)$ , this time we establish a pointwise convergence in both time and space variables. Here, the series representation in terms of  $\Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)$  for expectations of squared gradients (Proposition 2.1) plays a prominent role. For such expressions, we establish two claims. First, we introduce an approximation  $\tilde{\Pi}_t^{\varepsilon, i, k}(u_0^\varepsilon)$  of  $\Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)$  (Proposition 4.1). Then, we prove a limit theorem for these approximations (Proposition 4.2).

**Proposition 4.1.** *Recall (2.11). Then, for all  $t > 0$  and  $i = 1, \dots, d$ , we have*

$$\sum_{k=0}^{\infty} \left| \Pi_t^{\varepsilon, i, k}(u_0^\varepsilon)(x) - \tilde{\Pi}_t^{\varepsilon, i, k}(u_0^\varepsilon)(x) \right| \xrightarrow{\varepsilon \rightarrow 0} 0 , \quad (4.1)$$

uniformly over  $x \in \mathbb{T}_\varepsilon^d$ . Here,  $\tilde{\Pi}_t^{\varepsilon, i, k} := \Pi_t^{\varepsilon, i, k}$  for  $k = 0$ , while, for  $k \geq 1$ ,

$$\begin{aligned} & \tilde{\Pi}_t^{\varepsilon, i, k}(u_0^\varepsilon)(x) \\ & := \sum_{j=1}^d \Pi_t^{\varepsilon, j, 0}(u_0^\varepsilon)(x) \int_{[0, t]_{>}^k} ds_1 \cdots ds_k \sum_{y_1, \dots, y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_1, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \mathbb{1}_{j_k=j} . \end{aligned} \quad (4.2)$$

**Proposition 4.2.** *Recall (4.2). Then, there exists  $\mathfrak{a} = \mathfrak{a}(d) \in (0, 1]$  such that, for all  $t > 0$  and  $i = 1, \dots, d$ , we have*

$$\sum_{k=0}^{\infty} \tilde{\Pi}_t^{\varepsilon, i, k}(u_0^\varepsilon)(x) \xrightarrow{\varepsilon \rightarrow 0} \sum_{j=1}^d \left( (1 + \mathfrak{a}) \mathbb{1}_{i=j} + \frac{1 - \mathfrak{a}}{d - 1} \mathbb{1}_{i \neq j} \right) (\nabla^j u_t(x))^2 , \quad (4.3)$$

uniformly over  $x \in \mathbb{T}_\varepsilon^d$ .

We present the proofs of these two propositions in the subsequent section. Before that, we conclude this part by proving Theorem 4.

*Proof of Theorem 4.* From (1.17) and Proposition 2.1, we have

$$\begin{aligned} \mathbb{E}[\Gamma_t^\varepsilon(g)] &= \sum_{i=1}^d \int_0^t \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \mathbb{E}[(\nabla^{\varepsilon, i} u_s^\varepsilon(x))^2] g_s^i(x) ds \\ &= \sum_{i=1}^d \int_0^t \varepsilon^d \sum_{x \in \mathbb{T}_\varepsilon^d} \left\{ \sum_{k=0}^{\infty} \Pi_s^{\varepsilon, i, k}(u_0^\varepsilon)(x) \right\} g_s^i(x) ds . \end{aligned}$$

By the uniform boundedness of  $g$  and the estimate in Lemma 2.3, we may apply the dominated convergence theorem, so that the convergence in Theorem 4 reduces to that of the term in curly brackets above, for all  $s > 0$  and uniformly over  $x \in \mathbb{T}_\varepsilon^d$ . Recalling the expression of  $I_t(g)$  from (1.18), this is precisely the content of Propositions 4.1 and 4.2.  $\square$

**4.1. Proofs.** We start with two lemmas on some additional properties of the functions  $q^\varepsilon$  defined in (2.5). The first of this lemma is a refinement of Lemma 2.2 in the multidimensional case  $d \geq 2$ . We defer the proof of the following lemma to Appendix A.

**Lemma 4.1.** *Fix  $d \geq 2$ . Then, for all  $x \in \mathbb{T}_\varepsilon^d$ ,  $t \geq 0$ , and  $i, j = 1, \dots, d$ , the quantity*

$$\mathcal{Q}_\varepsilon^{i,j}(t) := \sum_{y \in \mathbb{T}_\varepsilon^d} q_t^{\varepsilon,i,j}(x, y) \quad (4.4)$$

*depends only on  $d \geq 2$ ,  $t \geq 0$  and  $\mathbb{1}_{i=j}$ . Moreover, letting*

$$\int_0^\infty \mathcal{Q}_\varepsilon^{i,j}(t) dt =: \begin{cases} \mathbf{b}_\varepsilon & \text{if } i = j \\ \mathbf{c}_\varepsilon & \text{if } i \neq j \end{cases}, \quad (4.5)$$

*we have  $\mathbf{b}_\varepsilon, \mathbf{c}_\varepsilon \in (0, \frac{1}{2}]$  and the following limits*

$$\mathbf{b} := \lim_{\varepsilon \rightarrow 0} \mathbf{b}_\varepsilon, \quad \mathbf{c} := \lim_{\varepsilon \rightarrow 0} \mathbf{c}_\varepsilon \quad (4.6)$$

*exist in  $(0, \frac{1}{2}]$ .*

**Remark 4.1.** *Comparing  $\mathcal{Q}_\varepsilon^{i,j}$  and  $\mathcal{Q}_\varepsilon$  given, respectively, in (4.4) and (2.7), we have*

$$\sum_{j=1}^d \mathcal{Q}_\varepsilon^{i,j}(t) = \mathcal{Q}_\varepsilon(t), \quad t \geq 0, \quad i = 1, \dots, d, \quad (4.7)$$

*which, together with Lemma 4.1 and (2.8), yields*

$$\mathbf{b}_\varepsilon + (d-1)\mathbf{c}_\varepsilon = \frac{1}{2}, \quad \mathbf{b} + (d-1)\mathbf{c} = \frac{1}{2}. \quad (4.8)$$

The following result refines property (2.8) when the extremes of integration are not  $t = 0$  and  $t = \infty$ , but rather  $t = T > 0$  and  $t = \infty$ . As already done for Lemmas 2.2 and 4.1, we postpone its proof to Appendix A.

**Lemma 4.2.** *There exists a constant  $c_3 = c_3(d) > 0$  satisfying, for all  $T > 0$ ,*

$$\int_T^\infty \mathcal{Q}_\varepsilon(t) dt \leq c_3 \varepsilon^3 (T \vee \varepsilon^2)^{-3/2}. \quad (4.9)$$

We are now ready to prove one of the two propositions of the section. All throughout, we employ the notation of Lemma 4.1 introduced for the case  $d \geq 2$ ; if  $d = 1$ , by writing, e.g.,  $\mathcal{Q}_\varepsilon^{i,j}$ ,  $\mathbf{b}_\varepsilon$ ,  $\mathbf{c}_\varepsilon$ , we actually mean  $\mathcal{Q}_\varepsilon$ ,  $\frac{1}{2}$ , 0, respectively.

*Proof of Proposition 4.2.* In view of Lemma 4.1, we may adopt the following shorthand notation for the time integral in (4.2): for all integers  $k \geq 1$ , and for all  $t \geq 0$  and  $i, j = 1, \dots, d$ ,

$$R_t^{\varepsilon,i,j,k} := \int_{[0,t]_{>}^k} ds_1 \cdots ds_k \sum_{j_0, j_1, \dots, j_k=1}^d \left( \prod_{\ell=1}^k \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s_{\ell-1} - s_\ell) \right) \mathbb{1}_{j_0=i} \mathbb{1}_{j_k=j}, \quad (4.10)$$

with the usual convention that  $s_0 := t$ . Similarly, we write

$$R_\infty^{\varepsilon,i,j,k} := \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k \int_0^\infty \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s) ds \right) \mathbb{1}_{j_0=i} \mathbb{1}_{j_k=j}. \quad (4.11)$$

Further, set  $R_t^{\varepsilon,i,j,k} = R_\infty^{\varepsilon,i,j,k} = \mathbb{1}_{i=j}$  for  $k = 0$ . With these definitions, we readily obtain

$$R_t^{\varepsilon,i,j,k} \leq R_\infty^{\varepsilon,i,j,k}, \quad k \geq 0. \quad (4.12)$$

We now claim that, for all  $t > 0$  and some  $C = C(d) > 0$ ,

$$(1 - Ct^{-3/2}\varepsilon^{3/2})^k R_\infty^{\varepsilon,i,j,k} \leq R_t^{\varepsilon,i,j,k}, \quad k \leq \lfloor \varepsilon^{-1} \rfloor. \quad (4.13)$$

Indeed, for all  $k \leq \lfloor \varepsilon^{-1} \rfloor$ ,  $R_t^{\varepsilon,i,j,k}$  is bounded below by

$$\sum_{j_0, \dots, j_k=1}^d \mathbb{1}_{j_0=i} \mathbb{1}_{j_k=j} \int_{t-\varepsilon t}^t ds_1 \mathcal{Q}_\varepsilon^{j_0, j_1}(ds_1) \int_{s_1-\varepsilon t}^{s_1} ds_2 \cdots \int_{s_{k-1}-\varepsilon t}^{s_{k-1}} ds_k \mathcal{Q}_\varepsilon^{j_{k-1}, j}(s_{k-1} - s_k), \quad (4.14)$$

where we used  $\mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell} \geq 0$  and  $k \leq \lfloor \varepsilon^{-1} \rfloor$  (guaranteeing that  $s_{k-1} - \varepsilon t \geq 0$ ). Thanks to Lemma 4.2, we estimate each nested integral as follows: for all  $\ell = 1, \dots, k$ ,

$$\begin{aligned} \int_{s_{\ell-1}-\varepsilon t}^{s_{\ell-1}} \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s_{\ell-1} - s_\ell) ds_\ell &= \int_0^\infty \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s) ds - \int_{t\varepsilon}^\infty \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s) ds \\ &\geq \int_0^\infty \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s) ds - c_3 t^{-3/2} \varepsilon^{3/2} \\ &\geq (1 - Ct^{-3/2} \varepsilon^{3/2}) \int_0^\infty \mathcal{Q}_\varepsilon^{j_{\ell-1}, j_\ell}(s) ds, \end{aligned}$$

where the first inequality used Lemma 4.2, whereas for the last inequality we employed the non-degeneracy of the limits (4.6) (thus, ensuring  $\liminf_{\varepsilon \rightarrow 0} \int_0^\infty \mathcal{Q}_\varepsilon^{i,j}(s) ds > 0$ ) and chose a sufficiently large constant  $C = C(d) > 0$ . Inserting these bounds into (4.14), we obtain (4.13). In conclusion, by combining (4.12) and (4.13), we get

$$\begin{aligned} \sum_{k=0}^\infty \left| R_t^{\varepsilon,i,j,k} - R_\infty^{\varepsilon,i,j,k} \right| &\leq \sum_{k=1}^{\lfloor \varepsilon^{-1} \rfloor} \left( 1 - (1 - C't^{-3/2} \varepsilon^{3/2})^k \right) R_\infty^{\varepsilon,i,j,k} + \sum_{k=\lfloor \varepsilon^{-1} \rfloor + 1}^\infty R_\infty^{\varepsilon,i,j,k} \\ &\leq 1 - (1 - C't^{-3/2} \varepsilon^{3/2})^{\lfloor \varepsilon^{-1} \rfloor} + 2^{-\lfloor \varepsilon^{-1} \rfloor}, \end{aligned}$$

where the last inequality used (4.7) and property (2.8). In conclusion, we obtain

$$\sum_{k=0}^\infty \left| R_t^{\varepsilon,i,j,k} - R_\infty^{\varepsilon,i,j,k} \right| \xrightarrow{\varepsilon \rightarrow 0} 0, \quad (4.15)$$

for all  $t > 0$  and  $i, j = 1, \dots, d$ .

We now prove that, for all  $i, j = 1, \dots, d$ , we have

$$\sum_{k=1}^\infty R_\infty^{\varepsilon,i,j,k} \xrightarrow{\varepsilon \rightarrow 0} \sum_{k=1}^\infty R_\infty^{i,j,k} := \frac{1}{1 - \mathbf{b} + \mathbf{c}} \times \begin{cases} \mathbf{b} + \mathbf{c} & \text{if } i = j \\ 2\mathbf{c} & \text{if } i \neq j, \end{cases} \quad (4.16)$$

where  $\mathbf{b}, \mathbf{c} \in (0, \frac{1}{2}]$  are the limits in (4.6).

Recall the definition of  $R_\infty^{\varepsilon,i,j,k}$ ,  $k \in \mathbb{N}$ , from (4.11). Thanks to Lemma 4.1,  $R_\infty^{\varepsilon,i,j,k}$  may be precisely determined by counting how many times adjacent indices in the sequence  $(j_0, j_1, \dots, j_{k-1}, j_k) \in \{1, \dots, d\}^{k+1}$  coincide. For this purpose, let  $N(j_0, \dots, j_k) \in$

$\{0, 1, \dots, k\}$  denote the number of adjacent indices with the same value, i.e.,

$$N(j_0, \dots, j_k) := \sum_{\ell=1}^k \mathbb{1}_{j_{\ell-1}=j_\ell}.$$

Then, by (4.5), we have

$$R_\infty^{\varepsilon, i, j, k} = \sum_{j_0, \dots, j_k=1}^d \left( \mathbf{b}_\varepsilon^{N(j_0, \dots, j_k)} \mathbf{c}_\varepsilon^{k-N(j_0, \dots, j_k)} \right) \mathbb{1}_{j_0=i} \mathbb{1}_{j_k=j}.$$

Clearly, the above quantity depends on  $i, j = 1, \dots, d$  only through  $\mathbb{1}_{i=j}$ ; therefore, we may introduce the following shorthand notation: for all  $k \in \mathbb{N}$ ,

$$F_k^\varepsilon := R_\infty^{\varepsilon, i, j, k} \quad \text{if } i = j, \quad G_k^\varepsilon := R_\infty^{\varepsilon, i, j, k} \quad \text{if } i \neq j.$$

When  $d \geq 2$ , such quantities satisfy

$$F_1^\varepsilon = \mathbf{b}_\varepsilon, \quad G_1^\varepsilon = \mathbf{c}_\varepsilon,$$

as well as the following one-step recursive formula: for  $k \geq 2$ ,

$$\begin{aligned} F_k^\varepsilon &= \mathbf{b}_\varepsilon F_{k-1}^\varepsilon + (\mathbf{c}_\varepsilon + (d-2)\mathbf{c}_\varepsilon) G_{k-1}^\varepsilon = \mathbf{b}_\varepsilon F_{k-1}^\varepsilon + \left(\frac{1}{2} - \mathbf{b}_\varepsilon\right) G_{k-1}^\varepsilon \\ G_k^\varepsilon &= \mathbf{c}_\varepsilon F_{k-1}^\varepsilon + (\mathbf{b}_\varepsilon + (d-2)\mathbf{c}_\varepsilon) G_{k-1}^\varepsilon = \mathbf{c}_\varepsilon F_{k-1}^\varepsilon + \left(\frac{1}{2} - \mathbf{c}_\varepsilon\right) G_{k-1}^\varepsilon, \end{aligned}$$

where for the second set of identities we used relation (4.8). In other words, letting

$M_\varepsilon := \begin{pmatrix} \mathbf{b}_\varepsilon & \frac{1}{2} - \mathbf{b}_\varepsilon \\ \mathbf{c}_\varepsilon & \frac{1}{2} - \mathbf{c}_\varepsilon \end{pmatrix}$ , we just obtained

$$\begin{pmatrix} F_k^\varepsilon \\ G_k^\varepsilon \end{pmatrix} = M_\varepsilon \begin{pmatrix} F_{k-1}^\varepsilon \\ G_{k-1}^\varepsilon \end{pmatrix} = \dots = M_\varepsilon^{k-1} \begin{pmatrix} \mathbf{b}_\varepsilon \\ \mathbf{c}_\varepsilon \end{pmatrix},$$

from which we get

$$\sum_{k=1}^{\infty} M_\varepsilon^{k-1} \begin{pmatrix} \mathbf{b}_\varepsilon \\ \mathbf{c}_\varepsilon \end{pmatrix} = (\mathbb{I} - M_\varepsilon)^{-1} \begin{pmatrix} \mathbf{b}_\varepsilon \\ \mathbf{c}_\varepsilon \end{pmatrix}, \quad \text{with } \mathbb{I} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In conclusion, since

$$(\mathbb{I} - M_\varepsilon)^{-1} = \frac{1}{1 - \mathbf{b}_\varepsilon + \mathbf{c}_\varepsilon} \begin{pmatrix} 1 + 2\mathbf{c}_\varepsilon & 1 - 2\mathbf{b}_\varepsilon \\ 2\mathbf{c}_\varepsilon & 2 - 2\mathbf{b}_\varepsilon \end{pmatrix},$$

the left-hand side of (4.16) equals, depending on whether  $i = j$  or  $i \neq j$ ,

$$\frac{1}{1 - \mathbf{b}_\varepsilon + \mathbf{c}_\varepsilon} \times \begin{cases} \mathbf{b}_\varepsilon + \mathbf{c}_\varepsilon & \text{if } i = j \\ 2\mathbf{c}_\varepsilon & \text{if } i \neq j. \end{cases} \quad (4.19)$$

Taking  $\varepsilon \rightarrow 0$ , (4.6) yields (4.16) for  $d \geq 2$ . When  $d = 1$ ,  $i \neq j$  is not possible; hence, we have  $F_k^\varepsilon = \mathbf{b}_\varepsilon^k$  for all  $k \in \mathbb{N}$ . Since  $\mathbf{b}_\varepsilon = \frac{1}{2}$  (see Remark 4.1), the left-hand side of (4.16) equals  $\sum_{k=1}^{\infty} F_k^\varepsilon = \sum_{k=1}^{\infty} 2^{-k} = 1$ . Since  $\mathbf{b} = \frac{1}{2}$  and  $\mathbf{c} = 0$  when  $d = 1$ , the right-hand side of (4.16) equals 1, as desired.

Now, recall that  $\nabla^j u_t = \nabla^j P_t u_0 = P_t \nabla^j u_0$ , for all  $j = 1, \dots, d$ , and similarly for the  $\varepsilon$ -semigroup and corresponding gradients. By the functional CLT for  $(X_t^\varepsilon)_{t \geq 0}$ , the assumptions  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$  with  $u_0^\varepsilon = u_0|_{\mathbb{T}_\varepsilon^d}$  ensure

$$\sup_{t \geq 0} \sup_{x \in \mathbb{T}_\varepsilon^d} \left| (\nabla^{\varepsilon, j} P_t^\varepsilon u_0^\varepsilon(x))^2 - (\nabla^j u_t(x))^2 \right| \xrightarrow{\varepsilon \rightarrow 0} 0, \quad j = 1, \dots, d. \quad (4.20)$$

In view of (4.15), (4.16) and (4.20), the claim in (4.3) follows by a triangle inequality. Indeed, recalling the definitions (4.2), (4.10), (4.11), (2.10), and the uniform boundedness of  $\Pi_t^{\varepsilon,j,0}(u_0^\varepsilon)$ , we have, as  $\varepsilon \rightarrow 0$ ,

$$\sum_{k=0}^{\infty} \tilde{\Pi}_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) = \sum_{j=1}^d \Pi_t^{\varepsilon,j,0}(u_0^\varepsilon)(x) \sum_{k=0}^{\infty} R_t^{\varepsilon,i,j,k}$$

$$(4.20) \implies \approx \sum_{j=1}^d (\nabla^j u_t(x))^2 \sum_{k=0}^{\infty} R_t^{\varepsilon,i,j,k}$$

$$(4.15) \implies \approx \sum_{j=1}^d (\nabla^j u_t(x))^2 \sum_{k=0}^{\infty} R_\infty^{\varepsilon,i,j,k}$$

$$(4.16) \implies \approx \sum_{j=1}^d (\nabla^j u_t(x))^2 \sum_{k=0}^{\infty} R_\infty^{i,j,k},$$

where we defined  $R_\infty^{i,j,k} := \mathbf{1}_{i=j}$  for  $k = 0$ . Observe that the above convergences hold for all  $t > 0$ ,  $i = 1, \dots, d$ , and uniformly over  $x \in \mathbb{T}_\varepsilon^d$ . This proves (4.3) with  $\mathbf{a} = \sum_{k=1}^{\infty} R_\infty^{i,i,k}$ , thus, concluding the proof of the proposition.  $\square$

**Remark 4.2.** *The proof of Proposition 4.2 — in particular, relations (4.16) and (4.19) — allows us to express the value  $\mathbf{a} = \mathbf{a}(d) \in (0, \frac{1}{2}]$  in terms of the limits  $\mathbf{b}, \mathbf{c} \in (0, \frac{1}{2}]$  in (4.6), combined with the relation (4.8):*

$$\mathbf{a} := \frac{\mathbf{b} + \mathbf{c}}{1 - \mathbf{b} + \mathbf{c}} = \frac{1 - 2(d-2)\mathbf{c}}{1 + 2d\mathbf{c}}. \quad (4.21)$$

*Proof of Proposition 4.1.* Recalling (2.10), (2.11) and (4.2), we have, for all  $k \in \mathbb{N}$ ,

$$\begin{aligned} & \Pi_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) - \tilde{\Pi}_t^{\varepsilon,i,k}(u_0^\varepsilon)(x) \\ &= \int_{[0,t]_{>}^k} ds_1 \cdots ds_{k-1} \sum_{y_1, \dots, y_{k-1} \in \mathbb{T}_\varepsilon^d} \sum_{j_1, \dots, j_{k-1}=1}^d \left( \prod_{\ell=1}^{k-1} q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \\ & \quad \times \int_0^{s_{k-1}} ds_k \sum_{y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_k=1}^d q_{s_{k-1}-s_k}^{\varepsilon, j_{k-1}, j_k}(y_{k-1}, y_k) \left\{ \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon)(y_k) - \Pi_t^{\varepsilon, j_k, 0}(u_0^\varepsilon)(x) \right\}. \end{aligned}$$

Further, by  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ ,  $u_0^\varepsilon = u_0|_{\mathbb{T}_\varepsilon^d}$  and the functional CLT in (4.20), we get

$$\left| \Pi_{s_k}^{\varepsilon, j_k, 0}(u_0^\varepsilon)(y_k) - \Pi_t^{\varepsilon, j_k, 0}(u_0^\varepsilon)(x) \right| \leq C_1 |P_{s_k} \nabla^{j_k} u_0(y_k) - P_t \nabla^{j_k} u_0(x)| \leq C_2,$$

for some  $C_1, C_2 > 0$  depending only on  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ . Hence, the left-hand side of (4.1) is bounded above by

$$C_1 \sum_{k=1}^{\infty} K_t^{\varepsilon,k}(u_0)(x) \leq \left( C_1 \sum_{k=1}^{\lceil \log \varepsilon \rceil^2} K_t^{\varepsilon,k}(u_0)(x) \right) + C_2 2^{-\lceil \log \varepsilon \rceil^2}, \quad (4.22)$$

where

$$K_t^{\varepsilon,k}(u_0)(x) := \int_{[-\infty, t]_{>}^{k-1}} ds_1 \cdots ds_{k-1} \sum_{y_1, \dots, y_{k-1} \in \mathbb{T}_\varepsilon^d} \sum_{j_0, \dots, j_{k-1}=1}^d \left( \prod_{\ell=1}^{k-1} q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right)$$

$$\times \int_{-\infty}^{s_{k-1}} ds_k \sum_{y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_k=1}^d q_{s_{k-1}-s_k}^{\varepsilon, j_{k-1}, j_k}(y_{k-1}, y_k) |P_{s_k \vee 0} \nabla^{j_k} u_0(y_k) - P_t \nabla^{j_k} u_0(x)|. \quad (4.23)$$

From now on, the proof ingredients are similar to those of Lemma 2.5 and Proposition 4.2. More in detail, we show, as already done in Proposition 4.2, that the integrals appearing in the definition (4.23) of  $K_t^{\varepsilon, k}(u_0)(x)$  may be restricted around their upper extremes, e.g., integrate  $s_k$  over  $[s_{k-1} - \varepsilon t, s_k]$  rather than on the whole  $[0, s_{k-1}]$ . Then, we conclude by exploiting the space-time continuity of  $P_s \nabla^j u_0(x)$  and the localization arguments employed in the proof of Lemma 2.5. We sketch this part of the proof below.

By splitting the domain of integration of the first integral in (4.23) into  $\{s_1 < t - \varepsilon t\}$  and its complement, we get, for some  $C_3 = C_3(u_0) > 0$ ,

$$K_t^{\varepsilon, k}(u_0)(x) \leq C_3 L_t^{\varepsilon, k} + K_t^{\varepsilon, k, 1}(u_0)(x), \quad (4.24)$$

where, for all  $m = 1, \dots, k$ ,

$$\begin{aligned} K_t^{\varepsilon, k, m}(u_0)(x) &:= \int_{[-\infty, t]_{>}^k} ds_1 \cdots ds_k \sum_{y_1, \dots, y_k \in \mathbb{T}_\varepsilon^d} \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \\ &\times |P_{s_k} \nabla^{j_k} u_0(y_k) - P_t \nabla^{j_k} u_0(x)| \left( \prod_{\ell=1}^m \mathbb{1}_{[s_{\ell-1}-\varepsilon t, s_{\ell-1}]}(s_\ell) \right), \end{aligned} \quad (4.25)$$

and

$$L_t^{\varepsilon, k} := d \left( \int_{t\varepsilon}^{\infty} \mathcal{Q}_\varepsilon(s) ds \right) \left( \int_0^{\infty} \mathcal{Q}_\varepsilon(s) ds \right)^{k-1} \leq dc_3 t^{-3/2} \varepsilon^{3/2} 2^{-(k-1)}. \quad (4.26)$$

Note that in (4.26) we used Lemmas 2.2 and 4.2 to obtain the inequality. Moreover, with the notation in (4.25), we may define  $K_t^{\varepsilon, k, 0} := K_t^{\varepsilon, k}$ . Applying this same strategy to  $K_t^{\varepsilon, k, 1}(u_0)(x)$ , by splitting the integral with respect to  $s_2 \in (-\infty, s_1)$  into  $\{s_2 < s_1 - \varepsilon t\}$  and its complement, we get

$$K_t^{\varepsilon, k, 1}(u_0)(x) \leq C_3 L_t^{\varepsilon, k} + K_t^{\varepsilon, k, 2}(u_0)(x).$$

Thus, by iterating  $k$  times the above inequality, (4.24) yields

$$K_t^{\varepsilon, k}(u_0)(x) \leq C_3 k L_t^{\varepsilon, k} + K_t^{\varepsilon, k, k}(u_0)(x). \quad (4.27)$$

Since  $k \leq \lceil |\log \varepsilon|^2 \rceil$ , we further get, for all  $\rho \in (0, \frac{1}{2})$ ,

$$\begin{aligned} K_t^{\varepsilon, k, k}(u_0)(x) &\leq d2^{-k} \left\{ \max_{j=1, \dots, d} \sup_{\substack{r, s \geq 0 \\ |r-s| < \varepsilon |\log \varepsilon|^2}} \|P_s \nabla^j u_0 - P_r \nabla^j u_0\|_{L^\infty(\mathbb{T}^d)} \right\} \\ &+ d2^{-k} \left\{ \max_{j=1, \dots, d} \sup_{s \geq 0} \sup_{\substack{x, y \in \mathbb{T}^d \\ |x-y| \leq \rho}} |P_s \nabla^j u_0(y) - P_s \nabla^j u_0(x)| \right\} \\ &+ C_4 \int_{[0, \varepsilon |\log \varepsilon|^2 t]_{>}^k} ds_1 \cdots ds_k \sum_{\substack{y_1, \dots, y_k \in \mathbb{T}_\varepsilon^d \\ |y_k - x| > \rho}} \sum_{j_0, \dots, j_k=1}^d \left( \prod_{\ell=1}^k q_{s_{\ell-1}-s_\ell}^{\varepsilon, j_{\ell-1}, j_\ell}(y_{\ell-1}, y_\ell) \right) \\ &=: d2^{-k} \left\{ \omega(\varepsilon |\log \varepsilon|^2) + \vartheta(\rho) \right\} + C_4 I_{\varepsilon |\log \varepsilon|^2 t}^{\varepsilon, k}(\varepsilon, \rho). \end{aligned} \quad (4.28)$$

Note that the time integral above coincides with  $I_t^{\varepsilon,k}(\rho_1, \rho_2)$  introduced in (2.24) — and estimated in (2.29) — with the following choices:

$$\rho_1 = \varepsilon, \quad \rho_2 = \rho, \quad t = \varepsilon |\log \varepsilon|^2 t, \quad n = \lceil |\log \varepsilon| \rceil^2.$$

As a consequence, summing over  $k = 1, \dots, \lfloor \varepsilon^{-1/2} \rfloor$ , (4.27), (4.26), (4.28) and (2.29) yield

$$\begin{aligned} & \sum_{k=1}^{\lceil |\log \varepsilon| \rceil^2} K_t^{\varepsilon,k}(u_0)(x) \\ & \leq C_5 \left( t^{-3/2} \varepsilon^{3/2} + \omega(\varepsilon |\log \varepsilon|^2) + \vartheta(\rho) + \varepsilon^{-(d+2)} \exp\left(-\frac{C' \rho}{|\log \varepsilon|^3 \varepsilon^{1/2}}\right) \right), \end{aligned}$$

for some  $C_5 > 0$  depending only on  $d \geq 1$  and  $u_0 \in \mathcal{C}^2(\mathbb{T}^d)$ . Taking, e.g.,  $\rho = \varepsilon^{1/3}$ , the above bound and the space-time continuity of  $P_s \nabla^j u_0$  ensure that the left-hand side above vanishes as  $\varepsilon \rightarrow 0$ , for all  $t > 0$ . The desired claim now follows from (4.22).  $\square$

## 5. PROOF OF THEOREM 1

In this section, we prove the functional CLT for the distribution-valued *càdlàg* process  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$  defined in (1.4). Let us recall from (1.13)–(1.14) the decomposition of  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$  into a drift and a martingale term:

$$\mathcal{Y}_t^\varepsilon(f) = \int_0^t \mathcal{Y}_s^\varepsilon \left( \frac{1}{2} \Delta_\varepsilon f \right) ds + \mathcal{M}_t^\varepsilon(f), \quad t \geq 0, \quad f \in \mathcal{C}^\infty(\mathbb{T}^d). \quad (5.1)$$

Observe that, by (1.15) and the notation introduced in (1.17),  $(\mathcal{U}_t^\varepsilon)_{t \geq 0}$  defined as

$$\mathcal{U}_t^\varepsilon(f) := \mathcal{M}_t^\varepsilon(f)^2 - \Gamma_t^\varepsilon \left( \frac{1}{2} \nabla^\varepsilon f \right), \quad t \geq 0, \quad f \in \mathcal{C}^\infty(\mathbb{T}^d), \quad (5.2)$$

is also a distribution-valued martingale (with respect to the natural filtration of  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$ ). Recall that all these fields depend on the initial conditions

$$u_0 \in \mathcal{C}^2(\mathbb{T}^d) \quad \text{and} \quad u_0^\varepsilon \equiv u_0|_{\mathbb{T}_\varepsilon^d} \in \mathbb{R}^{\mathbb{T}_\varepsilon^d},$$

which we fix all throughout this section.

As already mentioned in Section 1.3, the proof of Theorem 1 may be divided into steps (i)–(iii): tightness of the sequence, continuity of the limit points, and characterization of the limit, respectively. As we will see, these three steps build on Theorem 2 and two main estimates, which we present and prove in the next subsection.

**5.1. Main estimates.** Define, for all  $g \in \mathbb{R}^{\mathbb{T}_\varepsilon}$ ,  $\|g\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)} := \max_{i=1, \dots, d} \|\nabla^{\varepsilon,i} g\|_{L^\infty(\mathbb{T}_\varepsilon^d)}$ . This definition naturally extends to functions  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$ .

**Lemma 5.1** (Second moments of fields). *For all  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$  and  $t \geq 0$ , we have*

$$\mathbb{E}[\mathcal{Y}_t^\varepsilon(f)^2] \leq \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2 \quad (5.3)$$

$$\mathbb{E}[\mathcal{M}_t^\varepsilon(f)^2] \leq \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}^2 \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}^2. \quad (5.4)$$

*Proof.* We start with the proof of (5.3). By expanding the square and taking expectation, we obtain

$$\mathbb{E}[\mathcal{Y}_t^\varepsilon(f)^2] = \theta_\varepsilon^2 \varepsilon^{2d} \sum_{x, y \in \mathbb{T}_\varepsilon^d} \{ \mathbb{E}[u_t^\varepsilon(x) u_t^\varepsilon(y)] - \mathbb{E}[u_t^\varepsilon(x)] \mathbb{E}[u_t^\varepsilon(y)] \} f(x) f(y). \quad (5.5)$$

Writing  $u_t^\varepsilon(x)$  as the mild solution of (1.8), we get

$$u_t^\varepsilon(x) = P_t^\varepsilon u_0^\varepsilon(x) + \sum_{i=1}^d \frac{\varepsilon^2}{2} \int_0^t P_{t-s}^\varepsilon \nabla_*^{\varepsilon,i} (d\bar{N}_s^{\varepsilon,i} \nabla^{\varepsilon,i} u_{s-}^\varepsilon)(x)$$

$$= P_t^\varepsilon u_0^\varepsilon(x) + \sum_{i=1}^d \frac{\varepsilon^2}{2} \int_0^t \sum_{z \in \mathbb{T}_\varepsilon^d} \nabla_*^{\varepsilon,i} p_{t-s}^\varepsilon(\cdot, z)(x) \nabla^{\varepsilon,i} u_{s-}^\varepsilon(z) d\bar{N}_s^{\varepsilon,i}(z).$$

Since  $\bar{N}^\varepsilon = (\bar{N}_t^{\varepsilon,i}(x))_{t \geq 0, x \in \mathbb{T}_\varepsilon^d, i=1, \dots, d}$  is a family of i.i.d. martingales satisfying  $\mathbb{E}[(\bar{N}_t^{\varepsilon,i}(x))^2] = \varepsilon^{-2} t$ , we obtain

$$\begin{aligned} \mathbb{E}[u_t^\varepsilon(x) u_t^\varepsilon(y)] &= \mathbb{E}[u_t^\varepsilon(x)] \mathbb{E}[u_t^\varepsilon(y)] \\ &+ \sum_{i=1}^d \frac{\varepsilon^2}{4} \int_0^t \sum_{z \in \mathbb{T}_\varepsilon^d} \mathbb{E}[(\nabla^{\varepsilon,i} u_s^\varepsilon(z))^2] (\nabla_*^{\varepsilon,i} p_{t-s}^\varepsilon(\cdot, z)(x) \nabla_*^{\varepsilon,i} p_{t-s}^\varepsilon(\cdot, z)(y)) ds. \end{aligned}$$

Inserting this identity into (5.5) and recalling that  $\theta_\varepsilon = \varepsilon^{-(d/2+1)}$ , we get

$$\begin{aligned} \mathbb{E}[\mathcal{Y}_t^\varepsilon(f)^2] &= \frac{\varepsilon^d}{4} \sum_{z \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \int_0^t \mathbb{E}[(\nabla^{\varepsilon,i} u_s^\varepsilon(z))^2] \left( \sum_{x \in \mathbb{T}_\varepsilon^d} (\nabla_*^{\varepsilon,i} p_{t-s}^\varepsilon(\cdot, z)(x)) f(x) \right)^2 ds \\ &= \varepsilon^d \sum_{z \in \mathbb{T}_\varepsilon^d} \sum_{i=1}^d \int_0^t \mathbb{E}[(\nabla^{\varepsilon,i} u_s^\varepsilon(z))^2] \left( \frac{1}{2} P_{t-s}^\varepsilon \nabla^{\varepsilon,i} f(z) \right)^2 ds. \end{aligned} \quad (5.6)$$

By the estimate  $\|P_{t-s}^\varepsilon \nabla^{\varepsilon,i} f\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq \|\nabla^{\varepsilon,i} f\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}$ ,

$$\mathbb{E}[\mathcal{Y}_t^\varepsilon(f)^2] \leq \frac{1}{2} \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}^2 \int_0^t \frac{1}{2} \mathbb{E}[\|u_s^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2] ds \leq \frac{1}{2} \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}^2 \|u_0^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon^d)}^2, \quad (5.7)$$

where the second inequality used Aldous-Lanoue identity (1.26). This yields (5.3).

For what concerns (5.4), since the process in (5.2) is a mean-zero martingale, we get

$$\mathbb{E}[\mathcal{M}_t^\varepsilon(f)^2] = \mathbb{E}[\Gamma_t^\varepsilon(\frac{1}{2} \nabla^\varepsilon f)].$$

Noting  $\Gamma_t^\varepsilon(\frac{1}{2} \nabla^\varepsilon f)$  equals (5.6) with  $P_{t-s}^\varepsilon \nabla^{\varepsilon,i} f$  replaced by  $\nabla^{\varepsilon,i} f$  (cf. (1.17)), the argument leading from (5.6) to (5.7) proves (5.4).  $\square$

**Lemma 5.2** (Size of jumps). *For all  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$ , we have*

$$\mathbb{E}\left[\sup_{t \geq 0} |\mathcal{Y}_t^\varepsilon(f) - \mathcal{Y}_{t-}^\varepsilon(f)|\right] \leq \varepsilon^{d/2} \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}.$$

*Proof.* By the Poisson nature of the averaging dynamics, there is  $\mathbb{P}$ -a.s. at most one update at the time. Suppose to observe an update over the nearest neighbor vertices  $x$  and  $y = x + \varepsilon e_i \in \mathbb{T}_\varepsilon^d$  at time  $t \geq 0$ ; then,

$$\begin{aligned} &|\mathcal{Y}_t^\varepsilon(f) - \mathcal{Y}_{t-}^\varepsilon(f)| \\ &= \theta_\varepsilon \varepsilon^d \left| \frac{1}{2} (u_{t-}^\varepsilon(x) + u_{t-}^\varepsilon(y)) (f(x) + f(y)) - u_{t-}^\varepsilon(x) f(x) - u_{t-}^\varepsilon(y) f(y) \right| \\ &= \theta_\varepsilon \varepsilon^d \left| \frac{1}{2} (u_{t-}^\varepsilon(x) - u_{t-}^\varepsilon(y)) (f(x) - f(y)) \right| \\ &\leq \theta_\varepsilon \varepsilon^{d+1} \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)} \|u_{t-}^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \\ &\leq \theta_\varepsilon \varepsilon^{d+1} \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)} \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}, \end{aligned}$$

where the first inequality used  $|f(x) - f(y)| \leq \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}$ , while the second one used the monotonicity  $\|u_s^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \leq \|u_0^\varepsilon\|_{L^\infty(\mathbb{T}_\varepsilon^d)}$ . Recalling  $\theta_\varepsilon = \varepsilon^{-(d/2+1)}$  concludes the proof.  $\square$

**5.2. Martingale convergence theorem.** In view of the martingale decomposition in (5.1)–(5.2) and the conclusions of Theorem 2 and Lemma 5.2, the classical martingale convergence theorem (see, e.g, [EK86, Theorem 7.1.9, p. 339]) applies to our case. For the reader's convenience, we collect its consequences in the following proposition.

**Proposition 5.1.** *Recall (1.18). Then, for all  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$ , there exists a unique (in distribution) real-valued square-integrable continuous martingale  $M^f = (M_t^f)_{t \geq 0}$  (with respect to its natural filtration) with  $M_0^f = 0$ , predictable quadratic variation equal to  $\Gamma_t(\frac{1}{2}\nabla f)$ , and Gaussian independent increments. Moreover, recalling (1.14), (1.15), and (1.17), we have, for all  $f \in \mathcal{C}^\infty(\mathbb{T}^d)$ ,*

$$(\mathcal{M}_t^\varepsilon(f))_{t \geq 0} \xrightarrow{\varepsilon \rightarrow 0} M_t^f, \quad \text{in } \mathcal{D}([0, \infty); \mathbb{R}). \quad (5.8)$$

*Proof.* Since  $(\Gamma_t(\frac{1}{2}\nabla f))_{t \geq 0}$  is a continuous deterministic non-negative function, [EK86, Theorem 7.1.1, p. 338] proves the first assertion on the existence and uniqueness of the martingale  $(M_t^f)_{t \geq 0}$ .

As for the claim in (5.8), it suffices to observe that the hypotheses of [EK86, Theorem 7.1.4, p. 339], as well as conditions (1.16)–(1.19) therein, hold true in our case. Indeed, the predictable quadratic variations  $(\Gamma_t^\varepsilon(\frac{1}{2}\nabla^\varepsilon f))_{t \geq 0}$  of  $(\mathcal{M}_t^\varepsilon(f))_{t \geq 0}$  are continuous non-negative processes. Hence, the hypotheses and conditions (1.16) and (1.18) are fulfilled. Condition (1.17) therein follows from

$$|\mathcal{Y}_t^\varepsilon(f) - \mathcal{Y}_{t-}^\varepsilon(f)| = |\mathcal{M}_t^\varepsilon(f) - \mathcal{M}_{t-}^\varepsilon(f)|, \quad (5.9)$$

which holds because the drift term in (5.1) is continuous, and Lemma 5.2. Finally, condition (1.19) is a consequence of the triangle inequality

$$\begin{aligned} & \left| \Gamma_t^\varepsilon(\tfrac{1}{2}\nabla^\varepsilon f) - \Gamma_t(\tfrac{1}{2}\nabla f) \right| \leq \left| \Gamma_t^\varepsilon(\tfrac{1}{2}\nabla^\varepsilon f) - \Gamma_t^\varepsilon(\tfrac{1}{2}\nabla f) \right| + \left| \Gamma_t^\varepsilon(\tfrac{1}{2}\nabla f) - \Gamma_t(\tfrac{1}{2}\nabla f) \right| \\ & \leq \|f\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)} \max_{i=1, \dots, d} \|\nabla^{\varepsilon, i} f - \nabla^i f\|_{L^\infty(\mathbb{T}_\varepsilon^d)} \int_0^t \|u_s^\varepsilon\|_{H^1(\mathbb{T}_\varepsilon^d)}^2 ds + \left| \Gamma_t^\varepsilon(\tfrac{1}{2}\nabla f) - \Gamma_t(\tfrac{1}{2}\nabla f) \right|, \end{aligned}$$

taking expectation, Aldous-Lanoue identity (1.26), and Theorem 2. This concludes the proof of the proposition.  $\square$

**5.3. Tightness and continuity of limits.** In view of the decomposition (5.1), tightness and continuity of the limit points for  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$  in  $\mathcal{D}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$  is equivalent to the same property for the corresponding drift and martingale terms. Since we already established convergence for the martingale when tested against smooth test functions (Proposition 5.1), the tightness proof for this term may be simplified. For this purpose, let us start by recalling from [Bil99] a useful characterization of tightness with continuous limit points for general *càdlàg* processes.

**Proposition 5.2** (Tightness & continuous limits). *Let  $(\Xi, \|\cdot\|)$  be a Banach space. A sequence of  $\Xi$ -valued *càdlàg* processes  $((Y_t^\varepsilon)_{t \geq 0})_\varepsilon$  is tight in  $\mathcal{D}([0, \infty); \Xi)$  and such that all limits are continuous if and only if the following three conditions hold true: for all  $T > 0$  and  $\gamma > 0$ ,*

$$\lim_{\zeta \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{t \in [0, T]} \|Y_t^\varepsilon\| > \zeta \right) = 0, \quad (5.10)$$

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{\substack{s, t \in [0, T] \\ |t-s| < \delta}} \|Y_t^\varepsilon - Y_s^\varepsilon\| > \gamma \right) = 0, \quad (5.11)$$

and

$$\lim_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{t \in [0, T]} \|Y_t^\varepsilon - Y_{t^-}^\varepsilon\| > \gamma \right) = 0. \quad (5.12)$$

*Proof.* Just combine Theorems 13.2 and 13.4 in [Bil99] with the inequalities in (12.7)–(12.9) therein.  $\square$

**Remark 5.1.** Thanks to the definition of the norm  $\|\cdot\|_{H^{-\alpha}(\mathbb{T}^d)}$  in (1.12), Proposition 5.2 simplifies if  $(\Xi, \|\cdot\|) = (H^{-\alpha}(\mathbb{T}^d), \|\cdot\|_{H^{-\alpha}(\mathbb{T}^d)})$ , for some  $\alpha > 0$ . Indeed, while the first condition in Proposition 5.2, namely (5.10), remains unchanged, the second condition (5.11) may be replaced by

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{\substack{s, t \in [0, T] \\ |t-s| < \delta}} |Y_t^\varepsilon(\phi_m) - Y_s^\varepsilon(\phi_m)| > \gamma \right) = 0, \quad m \in \mathbb{Z}^d.$$

An analogous simplification holds for the third condition (5.12).

We now have all we need to prove the desired claim for the sequence  $((\mathcal{Y}_t^\varepsilon)_{t \geq 0})_\varepsilon$ .

**Proposition 5.3.** The sequence  $((\mathcal{Y}_t^\varepsilon)_{t \geq 0})_\varepsilon$  is tight in  $\mathcal{D}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$ , for all  $\alpha > 3 + d/2$ , and all limit points are continuous.

*Proof.* By (5.1), it suffices to verify the conditions in Proposition 5.2 for the drift and martingale terms separately. Recall (3.1), and fix  $T > 0$ ,  $\gamma > 0$  all throughout the proof.

We start with the drift term, for which we adopt the following shorthand notation:

$$\mathcal{A}_t^\varepsilon(f) := \int_0^t \mathcal{Y}_s^\varepsilon \left( \frac{1}{2} \Delta_\varepsilon f \right) ds, \quad t \geq 0, \quad f \in \mathcal{C}^\infty(\mathbb{T}^d).$$

As for the first condition, namely (5.10), Cauchy-Schwarz inequality yields

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{t \in [0, T]} \|\mathcal{A}_t^\varepsilon\|_{H^{-\alpha}(\mathbb{T}^d)}^2 \right] &\leq T \int_0^T \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-\alpha} \limsup_{\varepsilon \rightarrow 0} \mathbb{E} \left[ |\mathcal{Y}_s^\varepsilon \left( \frac{1}{2} \Delta_\varepsilon \phi_m \right)|^2 \right] ds \\ &\leq (2\pi)^4 T^2 \|u_0\|_\infty^2 \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-(\alpha-3)}, \end{aligned}$$

where the second inequality used (5.3) in Lemma 5.1 and  $\limsup_{\varepsilon \rightarrow 0} \|\Delta \phi_m\|_{\text{Lip}(\mathbb{T}_\varepsilon^d)}^2 \leq (2\pi)^4 (1 + |m|^2)^3$ . Since  $\alpha > 3 + d/2$ , the right-hand side above is finite. Therefore, Markov inequality yields (5.10) for the drift. For the second condition (5.11), a similar argument yields, for all  $m \in \mathbb{Z}^d$ ,

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{\substack{s, t \in [0, T] \\ |t-s| < \delta}} |\mathcal{A}_t^\varepsilon(\phi_m) - \mathcal{A}_s^\varepsilon(\phi_m)| \right] \leq \delta T (2\pi)^4 (1 + |m|^2)^3 \|u_0\|_\infty^2,$$

which vanishes as  $\delta \rightarrow 0$ . By Markov inequality and Remark 5.1, this suffices to prove (5.11). The third condition (5.12) is trivially satisfied since the drift is continuous.

Turning to the martingale term  $(\mathcal{M}_t^\varepsilon)_{t \geq 0}$ , we have

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, T]} \|\mathcal{M}_t^\varepsilon\|_{H^{-\alpha}(\mathbb{T}^d)}^2 \right] &\leq \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-\alpha} \mathbb{E} \left[ \sup_{t \in [0, T]} |\mathcal{M}_t^\varepsilon(\phi_m)|^2 \right] \\ &\leq \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-\alpha} \mathbb{E} [ |\mathcal{M}_T^\varepsilon(\phi_m)|^2 ] \\ &\leq (2\pi)^2 \|u_0\|_\infty^2 \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-(\alpha-1)}, \end{aligned}$$

where the second step used Doob inequality, while the third one used (5.4) in Lemma 5.1. The right-hand side above is finite because  $\alpha > 1 + d/2$ . This proves the first condition (5.10). Now observe that Proposition 5.1, (5.9), and Lemma 5.2 ensure that, for all  $m \in \mathbb{Z}^d$ , the sequence  $((\mathcal{M}_t^\varepsilon(\phi_m))_{t \geq 0})_\varepsilon$  satisfies all three conditions in Proposition 5.2 with  $(\Xi, \|\cdot\|) = (\mathbb{C}, |\cdot|)$ . Remark 5.1 concludes the proof.  $\square$

**5.4. Final step.** We now have all we need to prove Theorem 1. Indeed, the decomposition of  $(\mathcal{Y}_t^\varepsilon)_{t \geq 0}$  in (5.1), and Propositions 5.1 and 5.3 ensure that any limit process, say  $(\mathcal{Y}_t)_{t \geq 0}$ , have paths in  $\mathcal{C}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$ , for all  $\alpha > 3 + d/2$ , and solve the martingale problem (1.16) in  $\mathcal{C}([0, \infty); H^{-\alpha}(\mathbb{T}^d))$  with  $\mathcal{Y}_0 = 0$ . Moreover, the martingale  $(\mathcal{M}_t)_{t \geq 0}$  in (1.16) satisfies

$$(\mathcal{M}_t(f))_{t \geq 0} = (M_t^f)_{t \geq 0}, \quad f \in \mathcal{C}^\infty(\mathbb{T}^d),$$

where  $(M_t^f)_{t \geq 0}$  is defined in Proposition 5.1, and the above identity is meant in distribution.

The proof of Theorem 1 ends as soon as we show that such a limit martingale problem has a unique solution. For this purpose, we introduce  $\mathcal{S}(\mathbb{T}^d) := \bigcap_{\alpha \in \mathbb{R}} H^\alpha(\mathbb{T}^d)$  and  $\mathcal{S}(\mathbb{T}^d)' := \bigcup_{\alpha \in \mathbb{R}} H^\alpha(\mathbb{T}^d)$ . Since the embedding  $H^\alpha(\mathbb{T}^d) \hookrightarrow H^\beta(\mathbb{T}^d)$  is Hilbert-Schmidt for all  $\alpha > \beta + d/2$ ,  $(\mathcal{S}(\mathbb{T}^d), L^2(\mathbb{T}^d), \mathcal{S}(\mathbb{T}^d)')$  defines a countably Hilbert nuclear triple. Hence, it suffices to establish uniqueness of solutions in the larger space  $\mathcal{C}([0, \infty); \mathcal{S}(\mathbb{T}^d)')$ . By the Gaussian nature of the problem, the latter is precisely covered by Holley-Stroock theory [HS78] (see also [KL99, §11.4] or [JM18, Section C.5]).

#### APPENDIX A. PROOFS OF LEMMAS 2.2, 4.1 AND 4.2

In this appendix, we prove all results contained in Lemmas 2.2, 4.1 and 4.2 concerned with the quantity first introduced in (2.5), which we recall here for the reader's convenience: for all  $t \geq 0$ ,  $x, y \in \mathbb{T}_\varepsilon^d$  and  $i, j = 1, \dots, d$ ,

$$q_t^{\varepsilon, i, j}(x, y) := \frac{\varepsilon^{-2}}{4} (p_t^\varepsilon(x - \varepsilon e_j, y) + p_t^\varepsilon(x + \varepsilon e_i, y) - p_t^\varepsilon(x + \varepsilon e_i - \varepsilon e_j, y) - p_t^\varepsilon(x, y))^2. \quad (\text{A.1})$$

All throughout this section, we will exploit the invariance and product structure of the random walk  $(X_t^\varepsilon)_{t \geq 0}$  on  $\mathbb{T}_\varepsilon^d$ : for all  $t \geq 0$  and  $x, y \in \mathbb{T}_\varepsilon^d$  with  $x = (x^i)_{i=1, \dots, d}$ ,

$$p_t^\varepsilon(x, y) = p_t^\varepsilon(y, x) = p_t^\varepsilon(0, x - y), \quad p_t^\varepsilon(0, x) = \prod_{i=1}^d \pi_t^\varepsilon(x^i), \quad (\text{A.2})$$

where  $\pi_t^\varepsilon$  denotes the distribution of the continuous-time random walk on  $\mathbb{T}_\varepsilon$  (i.e., one-dimensional), started from the origin and with nearest-neighbor jump rates equal to  $\frac{1}{2}\varepsilon^{-2}$ .

As already mentioned at the beginning of Section 4.1, Lemma 4.1 essentially generalizes Lemma 2.2. For this reason, we find more convenient to prove these two lemmas together.

*Proofs of Lemmas 2.2 and 4.1.* We start by expressing  $q^\varepsilon$  in (A.1) in terms of  $\pi^\varepsilon$ , introduced in (A.2): for all  $t \geq 0$ ,  $x = (x^i)_{i=1, \dots, d} \in \mathbb{T}_\varepsilon^d$ , and  $i, j = 1, \dots, d$  with  $i \neq j$ ,

$$\begin{aligned} q_t^{\varepsilon, i, j}(0, x) &= \left( \prod_{\substack{\ell=1 \\ \ell \neq i, j}}^d \pi_t^\varepsilon(x^\ell) \right)^2 \frac{\varepsilon^{-2}}{4} (\pi_t^\varepsilon(x^i + \varepsilon) - \pi_t^\varepsilon(x^i))^2 (\pi_t^\varepsilon(x^j - \varepsilon) - \pi_t^\varepsilon(x^j))^2 \\ &= \left( \prod_{\substack{\ell=1 \\ \ell \neq i, j}}^d \pi_t^\varepsilon(x^\ell) \right)^2 \varepsilon^2 \left( \frac{1}{2} \nabla^\varepsilon \pi_t^\varepsilon(x^i) \right)^2 \left( \frac{1}{2} \nabla_*^\varepsilon \pi_t^\varepsilon(x^j) \right)^2, \end{aligned}$$

while, for  $i = j = 1, \dots, d$ ,

$$\begin{aligned} q_t^{\varepsilon, i, i}(0, x) &= \left( \prod_{\substack{\ell=1 \\ \ell \neq i}}^d \pi_t^\varepsilon(x^\ell) \right)^2 \frac{\varepsilon^{-2}}{4} (\pi_t^\varepsilon(x^i + \varepsilon) + \pi_t^\varepsilon(x^i - \varepsilon) - 2\pi_t^\varepsilon(x^i))^2 \\ &= \left( \prod_{\substack{\ell=1 \\ \ell \neq i}}^d \pi_t^\varepsilon(x^\ell) \right)^2 \varepsilon^2 \left( \frac{1}{2} \Delta_\varepsilon \pi_t^\varepsilon(x^i) \right)^2, \end{aligned}$$

where all  $\varepsilon$ -gradients and corresponding laplacians here are one-dimensional. Define, for all  $t \geq 0$ ,

$$\begin{aligned} \mathcal{R}_\varepsilon(t) &:= \varepsilon^{-1} \|\pi_t^\varepsilon\|_{L^2(\mathbb{T}_\varepsilon)}^2 = \pi_{2t}^\varepsilon(0), \\ \mathcal{S}_\varepsilon(t) &:= 2 \left\| \frac{1}{2} \nabla^\varepsilon \pi_t^\varepsilon \right\|_{L^2(\mathbb{T}_\varepsilon)}^2 = 2 \left\| \frac{1}{2} \nabla_*^\varepsilon \pi_t^\varepsilon \right\|_{L^2(\mathbb{T}_\varepsilon)}^2, \\ \mathcal{T}_\varepsilon(t) &:= \varepsilon \left\| \frac{1}{2} \Delta_\varepsilon \pi_t^\varepsilon \right\|_{L^2(\mathbb{T}_\varepsilon)}^2. \end{aligned} \tag{A.3}$$

Since  $\Delta_\varepsilon = \nabla^\varepsilon \nabla_*^\varepsilon = \nabla_*^\varepsilon \nabla^\varepsilon$  and  $\frac{d}{dt} \pi_t^\varepsilon = \frac{1}{2} \Delta_\varepsilon \pi_t^\varepsilon$ , we have

$$\mathcal{S}_\varepsilon(t) = -\frac{\varepsilon}{2} \mathcal{R}'_\varepsilon(t), \quad \mathcal{T}_\varepsilon(t) = -\frac{\varepsilon}{2} \mathcal{S}'_\varepsilon(t). \tag{A.4}$$

Now, summing over  $x \in \mathbb{T}_\varepsilon^d$ , we obtain

$$\mathcal{Q}_\varepsilon^{i, j}(t) := \sum_{x \in \mathbb{T}_\varepsilon^d} q_t^{\varepsilon, i, j}(0, x) = \begin{cases} \mathcal{R}_\varepsilon(t)^{d-2} \mathcal{S}_\varepsilon(t)^2 & \text{if } i \neq j \ (d \geq 2) \\ \mathcal{R}_\varepsilon(t)^{d-1} \mathcal{T}_\varepsilon(t) & \text{if } i = j \ (d \geq 1). \end{cases} \tag{A.5}$$

The form of the right-hand side above proves the first claim in Lemma 4.1.

Further summing over  $j = 1, \dots, d$ , we get

$$\begin{aligned} \mathcal{Q}_\varepsilon(t) &:= \sum_{j=1}^d \mathcal{Q}_\varepsilon^{i, j}(t) = \mathcal{R}_\varepsilon(t)^{d-1} \mathcal{T}_\varepsilon(t) + (d-1) \mathcal{R}_\varepsilon(t)^{d-2} \mathcal{S}_\varepsilon(t)^2 \\ &= -\frac{\varepsilon}{2} (\mathcal{R}_\varepsilon(t)^{d-1} \mathcal{S}_\varepsilon(t))', \end{aligned} \tag{A.6}$$

where the second step follows from (A.4). Hence, since  $\mathcal{R}_\varepsilon(0) = 1$ ,  $\mathcal{S}_\varepsilon(0) = \varepsilon^{-1}$ , and  $\mathcal{S}_\varepsilon(t) \rightarrow 0$  as  $t \rightarrow \infty$ , we get

$$\int_0^\infty \mathcal{Q}_\varepsilon(t) dt = -\frac{\varepsilon}{2} \int_0^\infty (\mathcal{R}_\varepsilon(t)^{d-1} \mathcal{S}_\varepsilon(t))' dt = \frac{\varepsilon}{2} (\mathcal{R}_\varepsilon(0)^{d-1} \mathcal{S}_\varepsilon(0)) = \frac{1}{2}. \tag{A.7}$$

This is precisely the main claim in Lemma 2.2 (see also its reformulation in Remark 4.1). This settles the analysis when  $d = 1$ .

In order to prove the final claim in Lemma 4.1 on

$$\mathfrak{b}_\varepsilon = \int_0^\infty \mathcal{R}_\varepsilon(t)^{d-1} \mathcal{T}_\varepsilon(t) dt, \quad \mathfrak{c}_\varepsilon = \int_0^\infty \mathcal{R}_\varepsilon(t)^{d-2} \mathcal{S}_\varepsilon(t)^2 dt, \tag{A.8}$$

and the existence of their limits (see (4.5)–(4.6)), fix  $d \geq 2$  and first observe that, since  $\mathcal{R}_\varepsilon(t)$ ,  $\mathcal{S}_\varepsilon(t)$  and  $\mathcal{Q}_\varepsilon(t)$  are all strictly positive for all  $t \geq 0$ , (A.5) and (A.7) ensure that  $\mathfrak{b}_\varepsilon, \mathfrak{c}_\varepsilon \in (0, \frac{1}{2})$ . Furthermore, by the first identity in (4.8), it suffices to show that either  $\mathfrak{b}_\varepsilon$  or  $\mathfrak{c}_\varepsilon$  converges to a value in  $(0, \frac{1}{2})$ . Let us focus on  $\mathfrak{c}_\varepsilon$ .

Recall (A.3). From Laplace inversion formulas (see, e.g., [Cox89, Eq. (2.3)]), we know

$$\mathcal{R}_\varepsilon(t) = \varepsilon \sum_{j=0}^{\varepsilon^{-1}-1} \exp \{ -2 \psi(\varepsilon j) \varepsilon^{-2} t \}, \quad \text{with } z \in \mathbb{T} \mapsto \psi(z) := 1 - \cos(2\pi z).$$

This and the first identity in (A.4) yield

$$\begin{aligned} \mathbf{c}_\varepsilon &= \int_0^\infty \varepsilon^{d-2} \sum_{j_1, \dots, j_d=0}^{\varepsilon^{-1}-1} \psi(\varepsilon j_1) \psi(\varepsilon j_2) \exp \{ -2 (\psi(\varepsilon j_1) + \dots + \psi(\varepsilon j_d)) \varepsilon^{-2} t \} dt \\ &= \frac{\varepsilon^d}{2} \sum_{j_1, \dots, j_d}^{\varepsilon^{-1}-1} \frac{\psi(\varepsilon j_1) \psi(\varepsilon j_2)}{\psi(\varepsilon j_1) + \dots + \psi(\varepsilon j_d)}, \end{aligned}$$

which converges, as  $\varepsilon \rightarrow 0$ , to

$$\mathbf{c} := \frac{1}{2} \int_{\mathbb{T}^d} \frac{\psi(x^1) \psi(x^2)}{\psi(x^1) + \dots + \psi(x^d)} dx, \quad \text{with } x = (x^i)_{i=1, \dots, d} \in \mathbb{T}^d. \quad (\text{A.9})$$

Since the above integral is strictly positive, this completes the proof of Lemma 4.1.  $\square$

**Remark A.1** (Properties of  $\mathbf{b}$  and  $\mathbf{c}$ ). *The above proof also reveals the behaviors of the coefficients  $\mathbf{b}$  and  $\mathbf{c}$  as functions of the dimension  $d \geq 1$ .*

- (1) Monotonicity. *Since  $\mathcal{R}_\varepsilon(t) \in (0, 1)$  and  $\mathcal{S}_\varepsilon(t), \mathcal{T}_\varepsilon(t) > 0$  for all  $t > 0$ , (A.5) and (A.8) state that both  $\mathbf{b}_\varepsilon$  and  $\mathbf{c}_\varepsilon$  — and, thus, their limits  $\mathbf{b}$  and  $\mathbf{c}$ , too — are strictly decreasing with  $d \geq 1$ .*
- (2) Bounds. *By (4.8) and  $\mathbf{b} \geq 0$ , we have the upper bound  $\mathbf{c} \leq \frac{1}{2(d-1)}$ . Since  $\psi(z) \in [0, 2]$  and  $\int_{\mathbb{T}} \psi(z) dz = 1$ , the integral in (A.9) yields the lower bound  $\mathbf{c} \geq \frac{1}{4d}$ .*
- (3) 2D. *When  $d = 2$ , the value of the integral in (A.9) is explicit:*

$$\mathbf{c} = \frac{1}{2} \int_{\mathbb{T}} \int_{\mathbb{T}} \frac{\psi(x^1) \psi(x^2)}{\psi(x^1) + \psi(x^2)} dx^1 dx^2 = \frac{1}{2} - \frac{1}{\pi}, \quad \mathbf{b} = \frac{1}{\pi},$$

where for the second identity we used (4.8).

**Remark A.2** (Formula for  $\mathbf{a}$ ). *By combining (4.21) and (A.9), we get, for all  $d \geq 2$ ,*

$$\mathbf{a} = \mathbf{a}(d) = \frac{1 - (d-2) \int_{\mathbb{T}^d} \frac{\psi(x^1) \psi(x^2)}{\psi(x^1) + \dots + \psi(x^d)} dx}{1 + d \int_{\mathbb{T}^d} \frac{\psi(x^1) \psi(x^2)}{\psi(x^1) + \dots + \psi(x^d)} dx}, \quad \text{with } x = (x^i)_{i=1, \dots, d} \in \mathbb{T}^d.$$

We conclude the appendix with the short proof of Lemma 4.2.

*Proof of Lemma 4.2.* By (A.6) and arguing as for (A.7), we have

$$\int_T^\infty \mathcal{Q}_\varepsilon(t) dt = \frac{\varepsilon}{2} (\mathcal{R}_\varepsilon(T)^{d-1} \mathcal{S}_\varepsilon(T)) \leq \frac{\varepsilon}{2} \mathcal{S}_\varepsilon(T), \quad T > 0,$$

where for the inequality we used  $\mathcal{R}_\varepsilon \leq 1$ . The desired estimate in (4.9) is a consequence of the fact that  $\mathcal{S}_\varepsilon(T)$  is, up to normalization, the Dirichlet form of the diffusively-rescaled continuous-time random walk on  $\mathbb{T}_\varepsilon^n$  with  $n = 1$ , which is well-known to have the power law behavior  $T^{-(n+1/2)} = T^{-3/2}$  for  $T \in (\varepsilon^2, 1)$ . For a simple proof of this fact, see, e.g., [BB21, Eq. (5.9)].  $\square$

**Acknowledgements.** *The author thanks Assaf Shapira and Matteo Quattropani for fruitful discussions. The author acknowledges financial support by “Microgrants 2022”, funded by Regione FVG, legge LR 2/2011.*

## REFERENCES

- [AD18] Scott Armstrong and Paul Dario. Elliptic regularity and quantitative homogenization on percolation clusters. *Comm. Pure Appl. Math.*, 71(9):1717–1849, 2018.
- [AKM19] Scott Armstrong, Tuomo Kuusi, and Jean-Christophe Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der mathematischen Wissenschaften*. Springer, Cham, 2019.
- [AL12] David Aldous and Daniel Lanoue. A lecture on the averaging process. *Probab. Surv.*, 9:90–102, 2012.
- [Ald13] David Aldous. Interacting particle systems as stochastic social dynamics. *Bernoulli*, 19(4):1122–1149, 2013.
- [BB21] Sayan Banerjee and Krzysztof Burdzy. Rates of convergence to equilibrium for potlatch and smoothing processes. *Ann. Probab.*, 49(3):1129–1163, 2021.
- [BGPS06] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Trans. Inform. Theory*, 52(6):2508–2530, 2006.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [BR18] Marek Biskup and Pierre-François Rodriguez. Limit theory for random walks in degenerate time-dependent random environments. *J. Funct. Anal.*, 274(4):985–1046, 2018.
- [BRAS06] Márton Balázs, Firas Rassoul-Agha, and Timo Seppäläinen. The random average process and random walk in a space-time random environment in one dimension. *Comm. Math. Phys.*, 266(2):499–545, 2006.
- [CDSZ22] Sourav Chatterjee, Persi Diaconis, Allan Sly, and Lingfu Zhang. A phase transition for repeated averages. *Ann. Probab.*, 50(1):1–17, 2022.
- [CF11] Giacomo Como and Fabio Fagnani. Scaling limits for continuous opinion dynamics systems. *Ann. Appl. Probab.*, 21(4):1537–1567, 2011.
- [CGQ24] Nicolas Clozeau, Antoine Gloria, and Siguang Qi. Quantitative homogenization for log-normal coefficients. *arXiv:2403.00168*, 2024.
- [Cox89] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in  $\mathbf{Z}^d$ . *Ann. Probab.*, 17(4):1333–1366, 1989.
- [CQS23] Pietro Caputo, Matteo Quattropiani, and Federico Sau. Cutoff for the averaging process on the hypercube and complete bipartite graphs. *Electron. J. Probab.*, 28:Paper No. 100, 31, 2023.
- [Dar21] Paul Dario. Optimal corrector estimates on percolation cluster. *Ann. Appl. Probab.*, 31(1):377–431, 2021.
- [DG20a] Mitia Duerinckx and Antoine Gloria. Multiscale functional inequalities in probability: concentration properties. *ALEA Lat. Am. J. Probab. Math. Stat.*, 17(1):133–157, 2020.
- [DG20b] Mitia Duerinckx and Antoine Gloria. Multiscale functional inequalities in probability: constructive approach. *Ann. H. Lebesgue*, 3:825–872, 2020.
- [DG22] Mitia Duerinckx and Antoine Gloria. Quantitative homogenization theory for random suspensions in steady Stokes flow. *J. Éc. polytech. Math.*, 9:1183–1244, 2022.
- [DGO20] Mitia Duerinckx, Antoine Gloria, and Felix Otto. The structure of fluctuations in stochastic homogenization. *Comm. Math. Phys.*, 377(1):259–306, 2020.
- [DMP91] Anna De Masi and Errico Presutti. *Mathematical methods for hydrodynamic limits*, volume 1501 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991.
- [DO20] Mitia Duerinckx and Felix Otto. Higher-order pathwise theory of fluctuations in stochastic homogenization. *Stoch. Partial Differ. Equ. Anal. Comput.*, 8(3):625–692, 2020.
- [DSC96] P. Diaconis and L. Saloff-Coste. Nash inequalities for finite Markov chains. *J. Theoret. Probab.*, 9(2):459–510, 1996.
- [DSPS21] Lorenzo Dello Schiavo, Lorenzo Portinale, and Federico Sau. Scaling limits of random walks, harmonic profiles, and stationary non-equilibrium states in Lipschitz domains. *Ann. Appl. Probab. (to appear)*, *arXiv:2112.14196*, 2021.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.
- [FF98] P. A. Ferrari and L. R. G. Fontes. Fluctuations of a surface submitted to a random average process. *Electron. J. Probab.*, 3:no. 6, 34, 1998.
- [FMZ23] Luiz Renato Fontes, Mariela Pentón Machado, and Leonel Zuaznábar. Scaling limit of an equilibrium surface under the random average process. *arXiv:2310.04252*, 2023.

- [FS86] E. B. Fabes and D. W. Stroock. A new proof of Moser’s parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338, 1986.
- [GM16] Yu Gu and Jean-Christophe Mourrat. Scaling limit of fluctuations in stochastic homogenization. *Multiscale Model. Simul.*, 14(1):452–481, 2016.
- [GQ24] Antoine Gloria and Siguang Qi. Quantitative homogenization for log-normal coefficients via malliavin calculus: the one-dimensional case. *arXiv:2402.19182*, 2024.
- [HH14] Olle Häggström and Timo Hirscher. Further results on consensus formation in the Deffuant model. *Electron. J. Probab.*, 19:no. 19, 26, 2014.
- [HS78] Richard A. Holley and Daniel W. Stroock. Generalized Ornstein-Uhlenbeck processes and infinite particle branching Brownian motions. *Publ. Res. Inst. Math. Sci.*, 14(3):741–788, 1978.
- [JL23] Milton Jara and Claudio Landim. The stochastic heat equation as the limit of a stirring dynamics perturbed by a voter model. *Ann. Appl. Probab.*, 33(6A):4163–4209, 2023.
- [JM18] Milton Jara and Otávio Menezes. Non-equilibrium fluctuations of interacting particle systems. *arXiv:1810.09526*, 2018.
- [JO22] Marc Josien and Felix Otto. The annealed Calderón-Zygmund estimate as convenient tool in quantitative stochastic homogenization. *J. Funct. Anal.*, 283(7):Paper No. 109594, 74, 2022.
- [KL99] Claude Kipnis and Claudio Landim. *Scaling limits of interacting particle systems*, volume 320 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1999.
- [Koz79] S. M. Kozlov. The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327, 1979.
- [Lan12] N. Lanchier. The critical value of the Deffuant model equals one half. *ALEA Lat. Am. J. Probab. Math. Stat.*, 9(2):383–402, 2012.
- [LP11] Günter Last and Mathew D. Penrose. Poisson process Fock space representation, chaos expansion and covariance inequalities. *Probab. Theory Related Fields*, 150(3-4):663–690, 2011.
- [LP18] Günter Last and Mathew Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.
- [MSW22] Ramis Movassagh, Mario Szegedy, and Guanyang Wang. Repeated averages on graphs. *Ann. Appl. Probab. (to appear)*, *arXiv:2205.04535*, 2022.
- [Nas58] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.
- [PV81] 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-New York, 1981.
- [QS23] Matteo Quattropiani and Federico Sau. Mixing of the averaging process and its discrete dual on finite-dimensional geometries. *Ann. Appl. Probab.*, 33(2):936–971, 2023.
- [Rho08] Rémi Rhodes. On homogenization of space-time dependent and degenerate random flows. II. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(4):673–692, 2008.
- [Sau23] Federico Sau. Concentration and local smoothness of the averaging process. *arXiv:2311.14176*, 2023.