ALEA, Lat. Am. J. Probab. Math. Stat. **21**, 385–430 (2024)
DOI: 10.30757/ALEA.v21-17

Upper bounds on the fluctuations for a class of degenerate convex $\nabla\phi$-interface models

Paul Dario

Laboratoire d'Analyse et de Mathématiques Appliquées (LAMA) and CNRS, Université Paris-Est Créteil, 61 avenue du Général de Gaulle, Créteil, France
E-mail address: paul.dario@u-pec.fr
URL: https://www.math.ens.psl.eu/∼dario/

Abstract. We derive upper bounds on the fluctuations of a class of random surfaces of the $\nabla\phi$-type with convex interaction potentials. The Brascamp-Lieb concentration inequality provides an upper bound on these fluctuations for uniformly convex potentials. We extend these results to twice continuously differentiable convex potentials whose second derivative grows asymptotically like a polynomial and may vanish on an (arbitrarily large) interval. Specifically, we prove that, when the underlying graph is the d-dimensional torus of side length L, the variance of the height is smaller than $C \ln L$ in two dimensions and remains bounded in dimension $d \geq 3$.

The proof makes use of the Helffer-Sjöstrand representation formula (originally introduced by Helffer and Sjöstrand (1994) and used by Naddaf and Spencer (1997) and Giacomin, Olla Spohn (2001) to identify the scaling limit of the model), the anchored Nash inequality (and the corresponding on-diagonal heat kernel upper bound) established by Mourrat and Otto (2016) and Efron's monotonicity theorem for log-concave measures (Efron (1965)).

1. Introduction

The aim of this paper is to obtain fluctuations upper bounds for a class of random surfaces subject to $\nabla\phi$ type interaction arising in statistical physics. These models are used to model phase separation in $\mathbb{R}^{d+1}$, and are defined as follows. For any fixed dimension $d \geq 2$ and integer $L \geq 1$, we let $\mathbb{T}_L := (\mathbb{Z}/(2L+1)\mathbb{Z})^d$ be the d-dimensional torus of side length $2L+1$. We endow the edges of the torus with an orientation, let $E(\mathbb{T}_L)$ be the set of positively oriented edges of $\mathbb{T}_L$, and let V be a potential, i.e., a measurable function $V : \mathbb{R} \to \mathbb{R}$ satisfying suitable properties. The random surface on $\mathbb{T}_L$ with potential V is then the probability measure $\mu_{\mathbb{T}_L}$ on the set of functions $\Omega^\circ_{\mathbb{T}_L} := \left\{ \phi : \mathbb{T}_L \to \mathbb{R} \, : \, \sum_{x \in \mathbb{T}_L} \phi(x) = 0 \right\}$ defined by

$$\mu_{\mathbb{T}_L}(d\phi) := \frac{1}{Z_L} \exp\left(- \sum_{e \in E(\mathbb{T}_L)} V\left(\nabla\phi(e)\right) \right) d\phi, \tag{1.1}$$

Received by the editors February 5th, 2023; accepted January 18th, 2024.
2010 *Mathematics Subject Classification.* 82C41, 82C24, 35K65, 60K37.
Key words and phrases. Random interfaces, Nash inequality, heat kernel.

where the discrete gradient is defined by $\nabla\phi(e) := \phi(y) - \phi(x)$ for the positively oriented edge $e = (x, y)$, $d\phi$ denotes the Lebesgue measure on the space $\Omega^\circ_{\mathbb{T}_L}$ (equipped with the L^2 scalar product) and the normalization constant (or partition function)

$$Z_L := \int_{\Omega^\circ_{\mathbb{T}_L}} \exp\left(- \sum_{e \in E(\mathbb{T}_L)} V\left(\nabla\phi(e)\right) \right) d\phi$$

is chosen so that $\mu_{\mathbb{T}_L}$ is a probability distribution. The model (1.1) is known as the $\nabla\phi$ interface model or discrete Ginzburg-Landau model and has received considerable attention since its introduction in the seminal work of Brascamp et al. (1975) (see Funaki (2005); Velenik (2006) and Section 1.2). A natural property to investigate on the model is the question of its *localization* or *delocalization*, that is, to establish whether the variance $\mathrm{Var}_{\mu_{\mathbb{T}_L}}[\phi(0)]$ remains bounded or diverges to infinity as L tends to infinity. Explicit computations available in the case $V(x) = x^2$, i.e., in the case of the discrete Gaussian free field, show that this variance diverges as the size L of the torus tends to infinity in two dimensions (the random surface is said to be *delocalized*, and the divergence is in fact logarithmic in L), and remains bounded uniformly in L in higher dimensions (the random surface is then said to be *localized*). Brascamp et al. (1975) conjectured that this result should remain valid for any potential V satisfying $\int_{\mathbb{R}} \exp(-pV(x))\, dx < \infty$ for all $p > 0$ and obtained a sharp (up to multiplicative constant) upper bound on the fluctuations of the random surface, using the celebrated Brascamp-Lieb concentration inequality, for twice-continuously differentiable potentials satisfying $\inf V'' > 0$ and for a class of convex potential with quadratic growth. Since the results of Brascamp et al. (1975), the localization and delocalization upper bounds have been extended to various settings including:

- Non-convex potentials arising as a perturbation of uniformly convex potentials by Cotar et al. (2009); Cotar and Deuschel (2012);
- Non-convex potentials which are a perturbation of uniformly convex potentials and are amenable to renormalization group analysis by Adams et al. (2016, 2019), Hilger (2016, 2020b,a), Adams and Koller (2023) and Bauerschmidt et al. (2024+a,+) (for the discrete Gaussian model in the latter case);
- Potentials which can be written as a mixture of Gaussians by Biskup and Kotecký (2007), Biskup and Spohn (2011), Brydges and Spencer (2012) and Ye (2019);
- Convex potentials satisfying that the set $\{x \in \mathbb{R} : V''(x) = 0\}$ has Lebesgue measure 0 by Magazinov and Peled (2022);
- The potential $V(x) = |x|$ using the infra-red bound of Bricmont et al. (1982) (this case is also covered in Brydges and Spencer (2012)).

Lower bounds on the fluctuations of the random surface have been established in a much more general setting, and Mermin-Wagner type arguments have been used successfully to prove logarithmic lower bounds for the variance of the height in two dimensions for a large class of potentials including all the twice-continuously differentiable V by Brascamp et al. (1975); Dobrushin and Shlosman (1980); Fröhlich and Pfister (1981); Ioffe et al. (2002), as well as for models with hard-core constraint Miłoś and Peled (2015). Section 1.2 discusses additional results beyond the questions of localization and delocalization (such as hydrodynamic limit, scaling limit, strict convexity of the surface tension, decay of covariances, large deviations) which have been proved for this model.

In this article, we are interested the class of convex potentials whose second derivative grows like a polynomial, formally defined in Assumption 1.1 below.

Assumption 1.1. We assume that $V : \mathbb{R} \to \mathbb{R}$ is a potential satisfying the assumptions:

(i) *Regularity and convexity:* we assume that V is twice-continuously differentiable and convex;
(ii) *Growth of the second derivative:* we assume that the second derivative of V satisfies a power-law growth condition: there exist an exponent $r > 2$ and two constants $c_+, c_- \in (0, \infty)$ such

that

$$0 < c_- \leq \liminf_{|x|\to\infty} \frac{V''(x)}{|x|^{r-2}} \leq \limsup_{|x|\to\infty} \frac{V''(x)}{|x|^{r-2}} \leq c_+ < \infty.$$

The main theorem of this paper establishes that the variance of the random surface grows at most logarithmically fast in two dimensions and remains bounded in dimensions 3 and higher for the class of potentials satisfying Assumption 1.1.

Theorem 1.2 (Localization and Delocalization)**.** *Under Assumption 1.1, there exists a constant $C := C(d,V) < \infty$ such that, for any $L \geq 2$,*

$$\mathrm{Var}_{\mathbb{T}_L}[\phi(0)] \leq \begin{cases} C \ln L & \text{if } d = 2, \\ C & \text{if } d \geq 3. \end{cases}$$

Remark 1.3. The convexity of the potential V implies that the measure (1.1) is log-concave. Since log-concavity is a property which is closed under marginalization (by the Prékopa-Leindler inequality Prékopa (1971, 1973); Leindler (1972)), this implies that the distribution of the height $\phi(0)$ is also log-concave. Since the tail of a log-concave distribution decays at least exponentially fast on the scale of its standard deviation, the result of Theorem 1.2 can be extended from a bound on the variance to a bound on exponential moments.

Remark 1.4. It is plausible that the techniques developed in this article can be further extended to obtain more precise properties on the behavior of the model (such as its hydrodynamic and scaling limits). These questions are further discussed in Section 1.3 below.

1.1. *Outline of the proof.* In order to highlight the main ideas and techniques used to prove Theorem 1.2, we present below a sketch of the argument for potentials satisfying the following assumptions: we assume that $V : \mathbb{R} \to \mathbb{R}$ is twice-continuously differentiable, convex and that there exists $c_1 \in (0,1)$ such that

$$0 < c_1 \leq \liminf_{|x|\to\infty} V''(x) \quad \text{and} \quad \sup_{x\in\mathbb{R}} V''(x) \leq 1. \tag{1.2}$$

Note that this is more restrictive than Assumption 1.1; the full argument will require some notational and technical adjustments.

1.1.1. *The Helffer-Sjöstrand representation formula.* One of the main tools used to prove fluctuation upper bounds is the Helffer-Sjöstrand representation formula, initially introduced by Helffer and Sjöstrand (1994) and used by Naddaf and Spencer (1997) and Giacomin et al. (2001) in order to identify the scaling limit of the model, and by Deuschel et al. (2000) to establish a large deviation principle for the model (among other results, see Section 1.2). In the setting of this paper, the formula reads as follows. Let ϕ_L be the stationary Langevin dynamic associated with the Gibbs measure $\mu_{\mathbb{T}_L}$, i.e., the solution of the system of stochastic differential equations

$$\begin{cases} d\phi_L(t,x) = \nabla \cdot V'\left(\nabla\phi_L\right)(t,x) + \sqrt{2}dB_t(x) & \text{for } (t,x) \in (0,\infty) \times \mathbb{T}_L, \\ \phi_L(0,x) = \phi(x) & \text{for } x \in \mathbb{T}_L, \end{cases} \tag{1.3}$$

where $\{B_t(x) : t \geq 0, x \in \mathbb{T}_L\}$ is collection of independent Brownian motions, and the initial condition ϕ is sampled according to $\mu_{\mathbb{T}_L}$ independently of the Brownian motions. Then, one has the identity

$$\mathrm{Var}_{\mu_{\mathbb{T}_L}}[\phi(0)] = \mathbb{E}\left[\int_0^\infty P_{\mathbf{a}}(t,0)\,dt\right], \tag{1.4}$$

where $P_{\mathbf{a}}$ is the heat kernel associated with the discrete parabolic equation (using the notation of Section 2)

$$\begin{cases} \partial_t P_{\mathbf{a}}(t,x) - \nabla \cdot \mathbf{a} \nabla P_{\mathbf{a}}(t,x) = 0 & \text{for } (t,x) \in (0,\infty) \times \mathbb{T}_L, \\ P_{\mathbf{a}}(0,x) = \delta_0(x) - \frac{1}{|\mathbb{T}_L|} & \text{for } x \in \mathbb{T}_L, \end{cases}$$

with the environment $\mathbf{a}(t,e) := V''(\nabla \phi_L(t,e))$.

As has been observed in Helffer and Sjöstrand (1994); Naddaf and Spencer (1997); Deuschel et al. (2000); Giacomin et al. (2001), the Helffer-Sjöstrand representation formula can be combined with tools of elliptic regularity, in the form of on-diagonal heat kernel estimates, to prove upper bounds on the fluctuations on the random surface. For instance, if the potential V is assumed to be uniformly convex, i.e., if $0 < c_- \leq V'' \leq 1$, then one has the bound $c_- \leq \mathbf{a}(t,e) \leq 1$. In this setting the parabolic equation arising from the Helffer-Sjöstrand representation formula is uniformly elliptic, and this property is sufficient to prove the following on-diagonal upper bound on the heat kernel

$$P_{\mathbf{a}}(t,0) \leq \frac{C}{(1+t)^{\frac{d}{2}}} \exp\left(-\frac{t}{CL^2}\right). \tag{1.5}$$

Integrating the bound (1.5) over the times $t \in [0,\infty)$ and using the identity (1.4) yields the variance estimate stated in Theorem 1.2.

The proof of Theorem 1.2 follows the strategy described in the previous paragraph, but some additional arguments are required to take into account that the second derivative of a potential V satisfying (1.2) can vanish.

1.1.2. *The on-diagonal heat kernel upper bound in degenerate environment of Mourrat and Otto.* Under Assumption (1.2), the upper bound on the fluctuations of the random surface can be obtained by first extending the on-diagonal upper bound for the heat kernel (1.5) to *degenerate environments*, i.e., environments $\mathbf{a} : (0,\infty) \times E(\mathbb{T}_L) \to [0,1]$ which may vanish (or take values arbitrarily close to 0). This question has received significant attention from the mathematical community (see Section 1.2), and, in this article, we rely on the approach of Mourrat and Otto (2016) and of Biskup and Rodriguez (2018) who respectively proved an on-diagonal upper bound for the heat kernel and a quenched invariance principle for a large class of *dynamic degenerate environments*. The exact result of the former (stated in infinite volume) can be found in Mourrat and Otto (2016, Theorem 4.2). Their proof could be adapted to the setting considered here, and would show the following result. Given an environment $\mathbf{a} : (0,\infty) \times E(\mathbb{T}_L) \to [0,1]$, if we define the *moderated environment* by

$$w(t,e) := \int_t^\infty \frac{\mathbf{a}(s,e)}{(1+s-t)^4}\, ds, \tag{1.6}$$

then there exists a function $t \mapsto \mathscr{M}_t \in [1,\infty]$ depending only on the dimension d and the moderated environment w such that, for any $t \geq 0$,

$$P_{\mathbf{a}}(t,0) \leq \frac{\mathscr{M}_t}{(1+t)^{\frac{d}{2}}} \exp\left(-\frac{t}{\mathscr{M}_t L^2}\right). \tag{1.7}$$

The dependency of the function $\mathscr{M}$ on the parameter w is explicit and it satisfies the following property: if we assume that the environment $\mathbf{a}$ is random, that its law is stationary with respect to both space and time translations and reversible, and if, for any $k \in \mathbb{Z}_+$ and any $(t,e) \in (0,\infty) \times E(\mathbb{T}_L)$,

$$\mathbb{E}\left[w(t,e)^{-k}\right] < \infty, \tag{1.8}$$

then $(\mathscr{M}_t)_{t\geq 0}$ is a stationary process and, for any $k \in \mathbb{N}$,

$$\mathbb{E}\left[\mathscr{M}_t^k\right] < \infty. \tag{1.9}$$

The result of Mourrat and Otto can thus be applied to establish upper bounds on the fluctuations of random surfaces as follows: by the Helffer-Sjöstrand representation formula (noting that the law of the environment $\mathbf{a}(t,e) := V''(\nabla\phi_L(t,e))$ is stationary with respect to both the space and time variables), we see that, if the moment assumption (1.8) can be verified, then the inequality (1.9) implies that

$$\begin{aligned}\mathrm{Var}_{\mathbb{T}_L}\left[\phi(0)\right] = \mathbb{E}\left[\int_0^\infty P_{\mathbf{a}}(t,0)\,dt\right] &\leq \mathbb{E}\left[\int_0^\infty \frac{\mathscr{M}_t}{(1+t)^{\frac{d}{2}}}\exp\left(-\frac{t}{\mathscr{M}_t L^2}\right)dt\right]\\ &\leq \begin{cases} C\ln L \text{ if } d=2,\\ \quad C \text{ if } d\geq 3.\end{cases}\end{aligned}$$

In other words, the question of establishing upper bounds on the fluctuations of the random surface can be reduced to proving the moment condition (1.8) on the moderated environment w. The strategy will thus be to prove (1.8), and the argument is outlined in the following sections.

1.1.3. *A fluctuation estimate for the Langevin dynamic and stochastic integrability of the moderated environment.* In order to prove the moment condition (1.8), we will prove the following fluctuation estimate for the Langevin dynamic: for any $R>0$, there exists a constant C_R depending only on d and R such that, for any time $T\geq 0$ and any edge $e\in E(\mathbb{T}_L)$,

$$\mathbb{P}\left[\forall t\in[0,T],\, |\nabla\phi(t,e)|\leq R\right] \leq C_R\exp\left(-\frac{T}{C_R}\right). \tag{1.10}$$

Combining this result with assumption (1.2) and the definition $\mathbf{a}(t,e) := V''(\nabla\phi(t,e))$ shows that there exists a constant C_V depending only on d and V such that, for any $T\geq 0$ and any edge $e\in E(\mathbb{T}_L)$,

$$\mathbb{P}\left[\forall t\in[0,T],\, \mathbf{a}(t,e)=0\right] \leq C_V\exp\left(-\frac{T}{C_V}\right). \tag{1.11}$$

The estimate (1.11) implies that the environment arising from the Helffer-Sjöstrand representation cannot remain equal to 0 for a long time; it can in fact be generalized (the argument is the one of Proposition 4.4 below) so as to obtain the following stretched exponential stochastic integrability on the moderated environment: there exist an exponent $s>0$ and a constant C_V such that, for any $R\geq 0$,

$$\mathbb{P}\left[w(t,e)\leq \frac{1}{R}\right] \leq C_V\exp\left(-\frac{R^s}{C_V}\right),$$

which then implies the moment condition (1.8).

In the rest of this section, we give an outline of the proof of (1.10) for potentials satisfying (1.2). The argument relies on three observations:

(i) The Langevin dynamic ϕ_L defined in (1.3) can be seen as a *deterministic function* of the initial condition ϕ and the Brownian motions $\{B_t(x) : t\geq 0,\, x\in\mathbb{T}_L\}$.

(ii) For any $x\in\mathbb{T}_L$, the Brownian motion $B_t(x)$ can be decomposed into a sum of independent increments and Brownian bridges as follows: if, for any $n\in\mathbb{Z}_+$ and any $t\in[n,n+1]$, we define

$$X_n(x) := B_{n+1}(x)-B_n(x) \text{ and } W_n(t,x) := B_t(x)-B_n(x)-(t-n)X_n(x), \tag{1.12}$$

then the random variables $\{X_n(x) : n\in\mathbb{Z}_+\}$ form a collection of independent Gaussian random variables (of variance 1), and the stochastic processes $\{W_n(\cdot,x) : n\in\mathbb{Z}_+\}$ form a collection of independent Brownian bridges. Additionally, the increments are independent of the Brownian bridges.

(iii) Since the trajectory of the Brownian motion $B_t(x)$ can be reconstructed from the values of the increments $\{X_n(x) : n \in \mathbb{Z}_+\}$ and the Brownian bridges $\{W_n(\cdot, x) : n \in \mathbb{Z}_+\}$, we can see the Langevin dynamic ϕ_L as a deterministic function of the initial condition, the increments and the Brownian bridges. Using the definition (1.12), we see that the Langevin dynamic solves the system of stochastic differential equations, for any $n \in \mathbb{Z}_+$,

$$\begin{cases} d\phi_L(t,x) = \nabla \cdot V'\left(\nabla \phi_L\right)(t,x) + \sqrt{2} X_n(x) + \sqrt{2} dW_n(t,x) & \text{for } (t,x) \in (n, n+1) \times \mathbb{T}_L, \\ \phi_L(0,x) = \phi(x) & \text{for } x \in \mathbb{T}_L. \end{cases} \tag{1.13}$$

The strategy is then to study the partial derivative of the Langevin dynamic with respect to the increment $X_n(x)$. To this end, we differentiate both sides of (1.13) with respect to the increment $X_n(x)$, and obtain that the partial derivative $w := \partial \phi_L / \partial X_n(x)$ solves the parabolic equation

$$\begin{cases} \partial_t w(t,y) = \nabla \cdot \mathbf{a} \nabla w(t,y) + \sqrt{2} \mathbf{1}_{\{n \le t \le n+1\}} \mathbf{1}_{\{y = x\}} & \text{for } (t,y) \in (n, n+1) \times \mathbb{T}_L, \\ w(0,x) = 0 & \text{for } x \in \mathbb{T}_L, \end{cases}$$

where $\mathbf{a}(t,e) := V''(\nabla \phi_L(t,e))$ is the same environment as the one appearing in the Helffer-Sjöstrand representation formula. The Duhamel's principle then yields the identity (using the definition of the heat kernel (2.3))

$$w(n+1,x) = \sqrt{2} \int_n^{n+1} P_{\mathbf{a}}\left(n+1, x; s, x\right) + \frac{1}{|\mathbb{T}_L|} \, ds.$$

The right-hand side of the previous display can be lower bounded as follows. We first note that $P_{\mathbf{a}}\left(t, x; s, x\right) + \frac{1}{|\mathbb{T}_L|} \in [0,1]$ by the maximum principle. Combining these bounds with the upper bound $\mathbf{a} \le 1$, the identity $\partial_t P_{\mathbf{a}} = \nabla \cdot \mathbf{a} \nabla P_{\mathbf{a}}$ and the definition of the discrete elliptic operator, we deduce that, for any $t, s \in (0, \infty)$ with $t \ge s$,

$$P_{\mathbf{a}}(s,x;s,x) + \frac{1}{|\mathbb{T}_L|} = 1 \text{ and } \left|\partial_t P_{\mathbf{a}}(t,x;s,x)\right| \le 2d, \tag{1.14}$$

which then implies

$$w(n+1,x) \ge \sqrt{2} \int_0^1 \max(1 - 2ds, 0) \, ds = \frac{\sqrt{2}}{4d} > 0.$$

In words, the partial derivative of the value $\phi_L(n+1,x)$ with respect to the increment $X_n(x)$ is lower bounded by $\sqrt{2}/(4d)$ uniformly over all the realizations of the Brownian motions. This implies that $\phi_L(n+1,x)$ is an increasing function of the increment $X_n(x)$, and more specifically, that increasing the value of the increment $X_n(x)$ by a value $X \ge 0$ (while keeping the other increments and the Brownian bridges unchanged), causes the value of $\phi_L(n+1,x)$ to increase by at least $X/(4d)$.

The previous argument can be refined so as to prove that, for any edge $e = (x_0, x) \in E\left(\mathbb{T}_L\right)$, the derivative of the discrete gradient $\nabla \phi_L(n+1,e)$ with respect to the increment $X_n(x)$ is lower bounded by a positive real number uniformly over the realizations of the increments and the Brownian bridges.

This property can then be used to prove the following result: for any $R > 0$, there exists $\varepsilon := \varepsilon(R) > 0$ such that, if we denote by $\mathcal{F}_{n,x}$ the σ-algebra generated by the initial condition ϕ, the Brownian bridges $\{W_m(\cdot, y) : m \in \mathbb{Z}_+, \, y \in \mathbb{T}_L\}$ and the increments $\{X_m(y) : m \neq n \, \text{or} \, x \neq y\}$, then one has the almost sure upper bound on the conditional probability

$$\mathbb{P}\left[\left|\nabla \phi_L(n+1,e)\right| \le R \mid \mathcal{F}_{n,x}\right] \le 1 - \varepsilon. \tag{1.15}$$

In other words, the probability of the event $\{|\nabla \phi_L(n+1,e)| \le R\}$ conditionally on all the randomness except the increment $X_n(x)$ is almost surely smaller than $1 - \varepsilon$.

The inequality (1.15) can then be iterated (making use of the independence between the increments and the Brownian bridges) to prove that, for any $N \in \mathbb{Z}_+$,

$$\mathbb{P}\left[\forall n \in \{1,\dots,N\},\ |\nabla \phi_L(n,e)| \leq R\right] \leq (1-\varepsilon)^N,$$

which implies the exponential decay stated in (1.10).

1.1.4. *Extension of the argument to the potentials satisfying Assumption 1.1.* In the case of potentials satisfying Assumption 1.1, the second derivative of the potential V is unbounded from above, and thus the environment $\mathbf{a}$ appearing in the Helffer-Sjöstrand representation formula can take arbitrarily large values. This implies that the argument written above needs to be modified in two aspects:

- The proof of Mourrat and Otto (2016, Theorem 4.2) is written in infinite volume for degenerate dynamic environments satisfying the upper bound $\mathbf{a} \leq 1$. Their argument needs to be adapted the torus, and to cover a class of environments which may take arbitrarily large values. This is the subject of Section 4.
- In the situation where the environment $\mathbf{a}$ can take arbitrarily large values, the inequality on the time derivative of the heat kernel (1.14) does not hold uniformly over all the realizations of the Brownian motions, and thus the derivative of $\phi_L(n+1,x)$ with respect to the increment $X_n(x)$ cannot be lower bounded (by a strictly positive real number) uniformly over all the realizations of the Brownian motions. This difficulty is handled by first establishing a sharp stochastic integrability estimate on the discrete gradient of a random surface distributed according to $\mu_{\mathbb{T}_L}$ (see Proposition 3.1). Once equipped with this result, we are able to adapt the argument outlined in Section 1.1.3 to this setting (see Proposition 3.3), at the cost of more technicalities and a deterioration of the stochastic integrability in the fluctuation estimate (from exponential rate to the super-polynomial rate in Proposition 3.3).

1.2. *Discussion and background.*

1.2.1. *Random surfaces.* The study of random surfaces was initiated in the 1970s by Brascamp et al. (1975) who obtained sharp localization and delocalization estimates for potentials satisfying $\inf V'' > 0$ and for a class of convex potentials with quadratic growth. Since then, the result of localization and delocalization has been extended to different classes of potentials as mentioned above, and various other aspects of the model have been studied by the mathematical community (see Funaki (2005); Velenik (2006)).

The hydrodynamic limit of the $\nabla\phi$-model for uniformly convex potentials was established by Funaki and Spohn (1997). The result was later extended to various settings: the hydrodynamic limit with Dirichlet boundary condition and with a conservation law was proved by Nishikawa (2003, 2002). More recently, the hydrodynamic limit was established for a class on non-convex potentials by Deuschel et al. (2019).

On the level of fluctuations, it is expected that *the scaling limit* of the $\nabla\phi$-model is a *continuum Gaussian free field* under mild integrability conditions on the potential V. On a rigorous level, a general convergence result has been established for twice continuously differentiable and uniformly convex potentials by Brydges and Yau (1990), Naddaf and Spencer (1997) and Giacomin et al. (2001). In particular, the contributions Naddaf and Spencer (1997); Giacomin et al. (2001) used the Helffer-Sjöstrand representation formula (introduced in Helffer and Sjöstrand (1994)), which has become well-used technique to study the model, and is a central tool in the proof of Theorem 1.2. The scaling limit has then been established in various different settings. A finite-volume version of the result was established by Miller (2011), a local limit theorem was established in two dimensions by Wu (2022) and the scaling limit of the square of the field was identified by Deuschel and Rodriguez (2022). The scaling limit was proved for a class of convex potentials satisfying the assumption

$\inf V'' > 0$ by Andres and Taylor (2021). In the nonconvex setting, it was established in the high temperature regime by Cotar and Deuschel (2012), in the low temperature regime by renormalization group arguments by Hilger (2016, 2020b) (buiding upon the techniques of Adams et al. (2016)), and in the case of non-convex potentials which can be written as a mixture of Gaussian by Biskup and Spohn (2011) and Ye (2019).

Besides the hydrodynamic and scaling limits, other aspects of the model which have been the subject of consideration from the community include: the strict convexity of the surface tension for non-convex potentials by Adams et al. (2016) (the Cauchy-Born rule was also investigated in Adams et al. (2019)), Cotar et al. (2009), its C^2-regularity by Armstrong and Wu (2022), the decay of covariances for the gradient of the field by Delmotte and Deuschel (2005), Cotar and Deuschel (2012), and Hilger (2020a), large deviations by Deuschel et al. (2000), Funaki and Nishikawa (2001), entropic repulsion by Deuschel and Giacomin (2000), the maximum of the field by Belius and Wu (2020) and Wu and Zeitouni (2019), uniqueness (or lack of thereof) of shift-ergodic infinite-volume gradient Gibbs states by Biskup and Kotecký (2007) and Buchholz (2021). A more detailed account of the literature can be found in the review articles Funaki (2005); Velenik (2006).

We complete this section by mentioning some results and recent progress which have been obtained on a related model: the *integer-valued* random surfaces (formally obtained by replacing the Lebesgue measure by the counting measure on $\mathbb{Z}$ in the right-hand side of (1.1)). In this setting, a temperature is usually incorporated in the definition of the model. A different phenomenology is then observed and the model is known to exhibit a phase transition in two-dimensions between a localized regime (at low temperature where the variance of the field remains bounded) and a delocalized regime (at high temperature where the variance of the field grows logarithmically). The existence of this phase transition was originally established in the celebrated article of Fröhlich and Spencer (1981); Kharash and Peled (2017) (we also refer to the work of Wirth (2019) on the maximum of the field based on these techniques), and has been the subject of recent developments in a series of works by Lammers (2022b,a, 2023), van Engelenburg and Lis (2023a,b) and Aizenman et al. (2021). In the high temperature regime and in the case of the discrete Gaussian model (i.e., when $V(x) = x^2/2$), the scaling limit of the model was recently identified by Bauerschmidt et al. (2024+a,+) by implementing a delicate renormalization group argument.

1.2.2. *Parabolic equations with degenerate random coefficients and the random conductance model.* In the uniformly elliptic setting, upper bounds on the heat kernel were obtained in the celebrated work of Nash (1958). Due to the connections between heat kernels and reversible random walks, it has been an active line of research to extend these heat kernel estimates to random degenerate environments, and two cases can be distinguished: the *static* environments and the *dynamic* environments. A typical example of static random degenerate environment is the supercritical Bernoulli (bond) percolation cluster. In that case, the upper bounds on the heat kernel were established by Barlow (2004) and Mathieu and Remy (2004). These bounds (or the ingredient developed to prove it) became one of the ingredients in the proof of the quenched invariance principle for the random walk on the percolation cluster by Sidoravicius and Sznitman (2004), Berger and Biskup (2007), Mathieu and Piatnitski (2007), the parabolic Harnack inequality and the local limit theorem by Barlow, Barlow and Hambly (2009). The existence of heat kernel upper and lower bounds (matching the ones of the lattice) have been established for more general degenerate environments satisfying suitable moments assumptions by Andres et al. (2016, 2019, 2020) and Andres and Halberstam (2021), but this phenomenon is not generic and anomalous heat kernel decay has been proved for some random degenerate environments by Berger et al. (2008), Boukhadra (2010), Biskup and Boukhadra (2012) and Buckley (2013). Besides the question of the behavior of the heat kernel, the invariance principle has been established for degenerate conductances by Biskup and Prescott (2007), Andres et al. (2013), Mathieu (2008), Procaccia et al. (2016) and Bella and Schäffner (2020). We refer to to Biskup (2011) for a survey of the literature on the random conductance model

Significant progress have been achieved in the case of *dynamic* environments (which is the relevant one for the problem considered in this article). In this setting, the invariance principle has been proved under various assumptions on the environment by Boldrighini et al. (1997); Boldrigini et al. (2007), Rassoul-Agha and Seppäläinen (2005), Bandyopadhyay and Zeitouni (2006),Dolgopyat et al. (2008), Avena (2012) and Redig and Völlering (2013), and for general ergodic degenerate conductances with moment conditions by Andres et al. (2018) (a local limit theorem was further established in Andres et al. (2021)).

Finally, heat kernel upper bounds were established for a class of degenerate dynamic environments satisfying a moment assumption on the *moderated* environment introduced above by Mourrat and Otto (2016). The proof of Theorem 1.2 strongly relies on their techniques. Combining and enhancing the techniques of Andres et al. (2018) and Mourrat and Otto (2016), Biskup and Rodriguez (2018) established the quenched invariance principle for random walks evolving in a dynamic degenerate environment satisfying an assumption related to the one used in Mourrat and Otto (2016). In this line of research, we finally mention the recent contribution of Biskup and Pan (2023) which establishes a quenched invariance principle for a class of ergodic degenerate environments in the one-dimensional setting.

1.3. *Further comments and perspective.* It is plausible that the techniques developed in this article can be further developed to obtain more precise information on the behavior of the random surfaces with an interaction potential satisfying Assumption 1.1. It seems for instance reasonable to us that the fluctuation estimate of Proposition 3.3 can be used to prove that the surface tension of the model is strictly convex (i.e., that the eigenvalues of its Hessian are always strictly positive). The strict convexity of the surface tension plays an important role in the proof of the hydrodynamic limit in Funaki and Spohn (1997), and we further believe that this result could be combined with the estimate of Theorem 1.2 to prove a quantitative version of the hydrodynamic limit following the techniques of Armstrong and Dario (2024). Once the quantitative hydrodynamic limit has been established, it should be possible to develop a large-scale regularity theory for the model (see Armstrong and Dario (2024, Theorem 1.5)). This result would then be useful to quantify the ergodicity of the environment appearing in the Helffer-Sjöstrand representation formula and would be helpful to establish a quantitative version of the scaling limit of the model (following the insight of Naddaf and Spencer (1997); Giacomin et al. (2001)). We refer to the introduction of Armstrong and Dario (2024) for a more detailed description of this line of research. We plan to investigate this in a future work. On a qualitative level, we mention that it would be interesting to investigate whether the techniques of Biskup and Rodriguez (2018) can be adapted to the framework considered here to also identify the scaling limit of the model.

1.4. *Organization of the paper.* The rest of the paper is organized as follows. Section 2 collects some notation and preliminary results. In Section 3, we prove a stochastic integrability estimate for the gradient of a random surface distributed according to the periodic Gibbs measure $\mu_{\mathbb{T}_L}$ (Proposition 3.1), and deduce from it a fluctuation estimate for the Langevin dynamic (Proposition 3.3). Section 4 combines the results of Section 3 with the techniques and results of Mourrat and Otto (2016) (essentially adapting their argument to obtain an on-diagonal upper bound for the heat kernel in the case of the torus, and when the environment is not bounded from above but possesses strong stochastic integrability properties), and completes the proof of Theorem 1.2 by using the Helffer-Sjöstrand representation formula.

1.5. *Convention for constants and exponents.* Throughout this article, the symbols C and c denote positive constants which, except if explicitly stated, may vary from line to line, with C increasing and c decreasing. We will always assume that $C \in [1, \infty)$ and $c \in (0, 1]$. These constants may depend on various parameters which will be made explicit in the statements by the following convention: we will write $C := C(d, V)$ to specify that the constant C depends only on d and V.

Acknowledgments: The author is indebted to S. Armstrong, M. Harel, P. Lammers, J.-C. Mourrat, R. Peled, F. Schweiger, O. Zeitouni for encouragement and helpful conversations on the topic of this work, and is specifically grateful to F. Schweiger, O. Zeitouni for suggesting to prove Proposition 3.1, to P. Lammers for explaining a short proof of this result in the case of symmetric potentials (on which the proof below is based), and to J.-C. Mourrat for explaining the arguments of Mourrat and Otto (2016).

2. Notation and preliminary results

2.1. *General notation.* We fix an integer $L \in \mathbb{Z}_+$ with $L \geq 1$, consider the torus $\mathbb{T}_L := (\mathbb{Z}/(2L+1)\mathbb{Z})^d$, and denote by $\pi : \mathbb{Z}^d \to \mathbb{T}_L$ the canonical projection. Given a subset $U \subseteq \mathbb{T}_L$ or $U \subseteq \mathbb{Z}^d$, we let $E(U)$ be the set of *positively oriented* edges of U (for some pre-determined orientation). For $r \in \mathbb{Z}_+$, we let $\Lambda_r := \{-r, \ldots, r\}^d \subseteq \mathbb{Z}^d$ and identify these boxes as subsets of the torus using the canonical embedding π. We note that the canonical embedding $\pi_{|\mathbb{T}_L}$ restricted to the box Λ_L is a bijection, whose inverse will be denoted by $\pi_{|\mathbb{T}_L}^{-1}$. We denote by $|\cdot|$ be the Euclidean norm on $\mathbb{Z}^d$, and, for $x \in \mathbb{T}_L$, we write $|x| := |\pi_{|\mathbb{T}_L}^{-1}(x)|$.

Given an edge $e \in E(\mathbb{T}_L)$, and a vertex $x \in \Lambda_L$, we write $x \in e$ if x is one of the endpoints of e. Given two edges $e, e' \in E(\mathbb{T}_L)$, we write $e \cap e' \neq \emptyset$ if e and e' have at least one endpoint in common.

Given an edge $e \in \mathbb{T}_L$, we write $\sum_{x \in e}$ and $\sum_{e' \cap e \neq \emptyset}$ to respectively sum over the endpoints of e and over the edges which have (at least) one endpoint in common with e. Given a vertex $x \in \mathbb{T}_L$, we write $\sum_{e \ni x}$ to sum over the edges which have x as an endpoint.

Given two real numbers a, b, we denote by $a \wedge b = \min(a, b)$ and by $a \vee b = \max(a, b)$, and by $\lfloor a \rfloor$ and $\lceil a \rceil$ the floor and ceiling of a. We denote by $\mathbf{1}_A$ the indicator function of a set A and, for $x \in \mathbb{T}_L$, we let δ_x be the function defined on the torus by the formula: $\delta_x(y) = 0$ if $y \neq x$ ans $\delta_x(x) = 1$.

For any potential $V : \mathbb{R} \to \mathbb{R}$ satisfying Assumption 1.1, we denote by

$$R_V := 2 \inf \left\{ R \geq 1 \, : \, \inf_{|x| \geq R} V''(x) \geq 1 \right\}. \tag{2.1}$$

Assumption 1.1 guarantees that R_V is a finite nonnegative real number.

2.1.1. *Functions.* Given a subset $U \subseteq \mathbb{T}_L$ or $U \subseteq \mathbb{Z}^d$, we denote by $|U|$ the cardinality of U. We have in particular $|\mathbb{T}_L| = (2L+1)^d$. For any function $f : U \to \mathbb{R}$, and any exponent $p \geq 1$, we define the L^p-norm and the normalized L^p-norm of f by the formulae

$$\|f\|_{L^p(U)}^p := \sum_{x \in U} f(x)^p \text{ and } \|f\|_{\underline{L}^p(U)}^p := \frac{1}{|U|} \sum_{x \in U} f(x)^p.$$

We denote the discrete gradient of a function $f : \mathbb{T}_L \to \mathbb{R}$ over a positively oriented edge $e = (x, y) \in E(\mathbb{T}_L)$ by the formula

$$\nabla f(e) := f(y) - f(x).$$

We also define $f(e) = (f(x) + f(y))/2$. This definition is motivated by the following identity: for any pair of functions $f, g : \mathbb{T}_L \to \mathbb{R}$ and any $e \in E(\mathbb{T}_L)$,

$$\nabla (fg)(e) = f(e) \nabla g(e) + g(e) \nabla f(e).$$

We extend the definition of L^p-norms to functions defined on edges by writing, for any function $u : E(\mathbb{T}_L) \to \mathbb{R}$,

$$\|u\|^p_{L^p(U)} := \sum_{e \in E(U)} |u(e)|^p \text{ and } \|u\|^p_{\underline{L}^p(U)} := \frac{1}{|U|} \sum_{e \in E(U)} |u(e)|^p.$$

We define the nonlinear elliptic operator $\nabla \cdot V'(\nabla u)$ by the formula

$$\nabla \cdot V'(\nabla u)(x) := \sum_{\substack{e \in E(\mathbb{T}_L) \\ e=(x,y)}} V'(\nabla u(e)) - \sum_{\substack{e \in E(\mathbb{T}_L) \\ e=(y,x)}} V'(\nabla u(e)).$$

This definition takes into account the set $E(\mathbb{T}_L)$ is defined to be the set of positively oriented edges. Making this distinction is useful to cover the case of potentials V which are not symmetric. The main property of this operator is that it satisfies the following discrete integration by parts property: for any pair of functions $u, v : \mathbb{T}_L \to \mathbb{R}$,

$$\sum_{x \in \mathbb{T}_L} \nabla \cdot V'(\nabla u)(x) v(x) = - \sum_{e \in E(\mathbb{T}_L)} V'(\nabla u(e)) \nabla v(e).$$

2.1.2. *Parabolic equations and heat kernel.* An environment is a measurable map $\mathbf{a} : (0, \infty) \times E(\mathbb{T}_L) \to [0, \infty)$. Given an environment, we denote by $\nabla \cdot \mathbf{a} \nabla$ the dynamic elliptic operator defined by the formula: for any map $u : (0, \infty) \times \mathbb{T}_L \to \mathbb{R}$ and any $(t, x) \in (0, \infty) \times \mathbb{T}_L$,

$$\nabla \cdot \mathbf{a} \nabla u(t, x) = \sum_{\substack{e \in E(\mathbb{T}_L) \\ e=(x,y)}} \mathbf{a}(t, e) \nabla u(e) - \sum_{\substack{e \in E(\mathbb{T}_L) \\ e=(y,x)}} \mathbf{a}(t, e) \nabla u(e). \tag{2.2}$$

This operator satisfies the discrete integration by parts property: for any pair of functions $u, v : (0, \infty) \times \mathbb{T}_L \to \mathbb{R}$,

$$\sum_{x \in \mathbb{T}_L} \nabla \cdot \mathbf{a} \nabla u(t, x) v(t, x) = - \sum_{e \in E(\mathbb{T}_L)} \mathbf{a}(t, e) \nabla u(t, e) \nabla v(t, e).$$

For $(s, y) \in [0, \infty) \times \mathbb{T}_L$, we define the heat kernel $P_{\mathbf{a}}(\cdot, \cdot; s, y) : (s, \infty) \times \mathbb{T}_L \to \mathbb{R}$ to be the solution of the parabolic equation

$$\begin{cases} \partial_t P_{\mathbf{a}}(t, x; s, y) - \nabla \cdot \mathbf{a} \nabla P_{\mathbf{a}}(t, x; s, y) = 0 & \text{for } (t, x) \in (0, \infty) \times \mathbb{T}_L, \\ P_{\mathbf{a}}(s, x; s, y) = \delta_y(x) - \dfrac{1}{|\mathbb{T}_L|} & \text{for } x \in \mathbb{T}_L. \end{cases} \tag{2.3}$$

To simplify the notation, we write $P_{\mathbf{a}}(t, x)$ instead of $P_{\mathbf{a}}(t, x; 0, 0)$.

Remark 2.1. The preservation of mass for parabolic equations shows that the sum $\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t, x)$ is constant in time. The normalizing term $1/|\mathbb{T}_L|$ in (2.3) ensures that this sum is equal to 0, and in fact ensures that the heat kernel $P_{\mathbf{a}}$ converges to 0 as the time tends to infinity.

Remark 2.2. The maximum principle for parabolic equation ensures that, for any $(t, x) \in (0, \infty) \times \mathbb{T}_L$,

$$-\frac{1}{|\mathbb{T}_L|} \leq P_{\mathbf{a}}(t, x) \leq 1 - \frac{1}{|\mathbb{T}_L|}. \tag{2.4}$$

Using these inequalities and the identity $\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t, x) = 0$ for any $t \geq 0$, we see that, for any $t \geq 0$,

$$\|P_{\mathbf{a}}(t, \cdot)\|_{L^1(\mathbb{T}_L)} \leq 2. \tag{2.5}$$

2.2. *The Langevin dynamic and the Helffer-Sjöstrand representation formula.* The Gibbs measure $\mu_{\mathbb{T}_L}$ is naturally associated with the Langevin dynamic defined below. In the following definition, we let $L \in \mathbb{Z}_+$ be an integer, consider a collection $\{B_t(x) \, : \, t \geq 0, \, x \in \mathbb{T}_L\}$ of independent Brownian motions, and let $\phi : \mathbb{T}_L \to \mathbb{R}$ be a random surface sampled according to the Gibbs measure $\mu_{\mathbb{T}_L}$ independently of the Brownian motions.

Definition 2.3 (Langevin dynamic in the torus)**.** We define the Langevin dynamic associated with the Gibbs measure $\mu_{\mathbb{T}_L}$ to be the solution $\phi_L : \mathbb{T}_L \to \mathbb{R}$ of the system of stochastic differential equations

$$\begin{cases} d\phi_L(t,x) = \nabla \cdot V' \left(\nabla \phi_L \right) (t,x) + \sqrt{2} dB_t(x) & \text{for } (t,x) \in (0,\infty) \times \mathbb{T}_L, \\ \phi_L(0,x) = \phi(x) & \text{for } x \in \mathbb{T}_L. \end{cases} \tag{2.6}$$

We note that the dynamic ϕ_L can be seen as a *deterministic* function of the initial condition ϕ and of the Brownian motions $\{B_t(x) \, : \, t \geq 0, \, x \in \mathbb{T}_L\}$. To highlight this dependency, we will use the notation

$$\phi_L(t,x) \left(\phi, \{B_t(x) \, : \, t \geq 0, \, x \in \mathbb{T}_L\} \right).$$

This will be useful in Section 3.1.2, as the dynamic can then be differentiated with respect to the increments of the Brownian motions.

The law of the dynamic ϕ_L is not exactly stationnary as the spatially averaged value of the dynamic is not constant: summing the first equation of (2.6) over $x \in \mathbb{T}_L$ (and using a discrete integration by parts on the torus to cancel the term involving the nonlinear elliptic operator) shows the identity

$$\sum_{x \in \mathbb{T}_L} \phi_L(t,x) = \sum_{x \in \mathbb{T}_L} B_t(x).$$

In particular, the law of $\phi_L(t,\cdot)$ is not equal to $\mu_{\mathbb{T}_L}$ (if $t \neq 0$), as the sum would have to be equal to 0. Nevertheless, this is the only obstruction and the process

$$\phi_L(t,\cdot) - \frac{1}{|\mathbb{T}_L|} \sum_{x \in \mathbb{T}_L} \phi_L(t,x)$$

is stationnary both with respect to the space and time variables. It is also reversible. Note that, since the second term in the right-hand side is spatially constant, the discrete gradient $\nabla \phi_L$ is a stationnary process.

We next state the Helffer-Sjöstrand representation which allows to express the variance of linear functionals of a random surface distributed according to $\mu_{\mathbb{T}_L}$ in terms of the solution of a random parabolic equations defined in terms of the Langevin dynamic. The formula was initially introduced in Helffer and Sjöstrand (1994); Naddaf and Spencer (1997); Deuschel et al. (2000); Giacomin et al. (2001) and is stated below in the case of the torus for the specific observable $\phi(0)$.

Proposition 2.4 (Helffer-Sjöstrand representation formula on the torus)**.** *Let $P_{\mathbf{a}}$ be the solution of the parabolic equation in the torus*

$$\begin{cases} \partial_t P_{\mathbf{a}}(t,x) - \nabla \cdot \mathbf{a} \nabla P_{\mathbf{a}}(t,x) = 0 & \text{for } (t,x) \in (0,\infty) \times \mathbb{T}_L, \\ P_{\mathbf{a}}(0,x) = \delta_0(x) - \frac{1}{|\mathbb{T}_L|} & \text{for } x \in \mathbb{T}_L, \end{cases}$$

where $\mathbf{a} : (0,\infty) \times E(\mathbb{T}_L) \to [0,\infty)$ is the random dynamic environment defined by the formula, for any $(t,e) \in (0,\infty) \times E(\mathbb{T}_L)$

$$\mathbf{a}(t,e) := V''(\nabla \phi_L(t,e)).$$

Then one has the identity

$$\mathrm{Var}_{\mathbb{T}_L} \left[\phi(0) \right] = \mathbb{E} \left[\int_0^\infty P_{\mathbf{a}}(t,0) \, dt \right].$$

2.3. *Efron's monotonicity theorem for log concave measures.* In this section, we state the Efron's monotonicity theorem for a pair of independent log-concave random variables due to Efron (1965).

Theorem 2.5 (Efron's monotonicity theorem Efron (1965))**.** *Let (X, Y) be a pair of independent, real-valued and log-concave random variables and let $\Psi : \mathbb{R}^2 \to \mathbb{R}$ be a function which is nondecreasing in each of its arguments, then the conditional expectation*

$$\mathbb{E}\left[\Psi(X,Y) \,|\, X+Y=s\right] \text{ is nondecrasing in } s.$$

2.4. *The discrete Gagliardo-Nirenberg-Sobolev inequality.* We state below the discrete version of the standard Gagliardo-Nirenberg-Sobolev inequality in the torus. The proof can be deduced from the standard Gagliardo-Nirenberg inequality in bounded domain for which we refer to Nirenberg (1959).

Proposition 2.6 (discrete Gagliardo-Nirenberg-Sobolev inequality on the torus)**.** *Fix three exponents $\kappa, \lambda, \mu \in (1,\infty)$ and let $\theta \in [0,1]$ be such that the relation*

$$\frac{1}{\kappa} = \theta\left(\frac{1}{\lambda} - \frac{1}{d}\right) + \frac{1-\theta}{\mu}$$

holds. Then there exists a constant $C := C(d,\kappa,\lambda,\mu,\theta) < \infty$ such that for any $L \geq 1$ and any function $f : \mathbb{T}_L \to \mathbb{R}$,

$$\|f\|_{\underline{L}^\kappa(\mathbb{T}_L)} \leq CL^\theta \|\nabla f\|^\theta_{\underline{L}^\lambda(\mathbb{T}_L)} \|f\|^{1-\theta}_{\underline{L}^\mu(\mathbb{T}_L)} + C\|f\|_{\underline{L}^2(\mathbb{T}_L)}.$$

If $f : \mathbb{T}_L \to \mathbb{R}$ satisfies the additional assumption $\sum_{x\in\mathbb{T}_L} f(x) = 0$, then

$$\|f\|_{\underline{L}^\kappa(\mathbb{T}_L)} \leq CL^\theta \|\nabla f\|^\theta_{\underline{L}^\lambda(\mathbb{T}_L)} \|f\|^{1-\theta}_{\underline{L}^\mu(\mathbb{T}_L)}.$$

In the proofs below, we will apply the discrete Gagliardo-Nirenberg-Sobolev inequality and the Hölder inequality with the following collections of exponents:

$$\lambda_d := \frac{2d+2}{d+2},\ \kappa_d := \frac{d\lambda_d}{d-\lambda_d},\ \sigma_d := \frac{2\kappa_d}{\kappa_d - 2} \text{ and } \tau_d = \frac{2\lambda_d}{2-\lambda_d}, \tag{2.7}$$

and

$$\lambda'_d := \frac{2d+3}{d+2},\ \tau'_d := \frac{2\lambda'_d}{2-\lambda'_d} \text{ and } \theta_d := \frac{2}{3}\frac{2d+3}{2d+2}. \tag{2.8}$$

They are chosen so as to satisfy the following properties:

(1) For any dimension $d \geq 2$, $1 < \lambda_d < 2 < \kappa_d < \infty$, and the Gagliardo-Nirenberg-Sobolev inequality can be applied with the exponents $\kappa = \kappa_d$, $\lambda = \lambda_d$ and $\theta = 1$ (and arbitrary μ).

(2) The pair of exponents (λ_d, τ_d) and (κ_d, σ_d) are chosen so as to satisfy Hölder inequalities, and we have

$$\frac{1}{\tau_d} + \frac{1}{2} = \frac{1}{\lambda_d} \text{ and } \frac{1}{\kappa_d} + \frac{1}{\sigma_d} = \frac{1}{2}.$$

We also note that the following identities hold

$$\frac{1}{\sigma_d} + \frac{1}{\tau_d} = \frac{1}{d} \text{ and } \frac{1}{\tau'_d} + \frac{1}{2} = \frac{1}{\lambda'_d}. \tag{2.9}$$

(3) For any function $f : \mathbb{T}_L \to \mathbb{R}$, one has the inequality

$$\|f\|_{L^{\kappa_d}(\mathbb{T}_L)} \leq CL^{\theta_d} \|\nabla f\|^{\theta_d}_{\underline{L}^{\lambda'_d}(\mathbb{T}_L)} \|f\|^{1-\theta_d}_{\underline{L}^2(\mathbb{T}_L)} + C\|f\|_{\underline{L}^2(\mathbb{T}_L)}.$$

Applying Young's inequality for product, we deduce that, for any $\varepsilon \in (0,1]$,

$$\|f\|_{\underline{L}^{\kappa_d}(\mathbb{T}_L)} \leq \varepsilon L \|\nabla f\|_{\underline{L}^{\lambda'_d}(\mathbb{T}_L)} + C\varepsilon^{-\frac{\theta_d}{1-\theta_d}} \|f\|_{\underline{L}^2(\mathbb{T}_L)}.$$

Remark 2.7. The same inequalities hold on more general subsets than the torus, we will use it below in annuli of the form $A_r := \Lambda_{2r} \setminus \Lambda_r$ with $r \in \{1, \ldots, \frac{L}{2}\}$ which can be seen as a subset of the torus (using the identification mentioned in Section 2.1). In this setting, we have, for any $f : A_r \to \mathbb{R}$ and any $\varepsilon \in (0,1]$,

$$\|f\|_{\underline{L}^{\kappa_d}(A_r)} \leq \varepsilon r \|\nabla f\|_{\underline{L}^{\lambda'_d}(A_r)} + C\varepsilon^{-\frac{\theta_d}{1-\theta_d}} \|f\|_{\underline{L}^2(A_r)}.$$

2.5. *Maximal inequalities.* In this section, we recall some classical properties of maximal functions. We let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\theta_x)_{x\in\mathbb{Z}^d}$ be a measure preserving action of $\mathbb{Z}^d$ on this space. For every measurable function $f : \Omega \to \mathbb{R}$, we define the maximal function

$$M(f) := \sup_{r\in\mathbb{Z}_+} \frac{1}{|\Lambda_r|} \sum_{x\in\Lambda_r} f(\theta_x\omega). \tag{2.10}$$

We next record the L^p maximal inequality, which can be obtained as a consequence of the weak type $(1,1)$ estimate Akcoglu and Krengel (1981, Theorem 3.2) with the Marcinkiewicz interpolation theorem (see Taylor (2006, Appendix D)). The result is stated and used in Mourrat and Otto (2016, Appendix A).

Proposition 2.8 (L^p Maximal inequality)**.** *For any $p \in (1,\infty]$, there exists a constant $C := C(p,d) < \infty$ such that, for any $f \in L^p(\Omega)$,*

$$\|M(f)\|_{L^p(\Omega)} \leq C \|f\|_{L^p(\Omega)}.$$

2.6. *The anchored Nash estimate of Mourrat and Otto.* In this section, we record the anchored Nash estimate proved by Mourrat and Otto (2016, Theorem 2.1). In the statement below and later, we will make use of the notation $|\cdot|_* = |\cdot| + 1$.

Theorem 2.9 (Anchored Nash inequality, Theorem 2.1 of Mourrat and Otto (2016))**.** *Let $p \in (d,\infty)$, $p' \in (d,\infty]$, and $\theta \in [\theta_c, 1]$, where $\theta_c \in [0,1)$ is defined by*

$$\frac{1}{\theta_c} = 1 + \frac{dp+2p}{dp+2d}\left(\frac{p'}{d} - 1\right). \tag{2.11}$$

Define $\alpha, \beta, \gamma \in [0,1)$ by

$$\alpha := (1-\theta)\frac{d}{d+2} + \theta\frac{p}{p+2}, \ \ \beta := (1-\theta)\frac{2}{d+2}, \ \ \text{and } \gamma := \theta\frac{2}{p+2}. \tag{2.12}$$

There exists $C := C(d,p,q,\theta) < \infty$ such that, for any function $f : \mathbb{Z}^d \to \mathbb{R}$, and $w : E(\mathbb{Z}^d) \to (0,\infty)$,

$$\|f\|_{L^2(\mathbb{Z}^d)} \leq C\left(M(w^{-p'})^{\frac{1}{p'}} \|w\nabla f\|_{L^2(\mathbb{Z}^d)}\right)^{\alpha} \|f\|^{\beta}_{L^1(\mathbb{Z}^d)} \||x|_*^{p/2} f\|^{\gamma}_{L^2(\mathbb{Z}^d)}.$$

Remark 2.10. The statement of the maximal function $M(w^{-p'})$ is defined with respect to (Euclidean) balls in Mourrat and Otto (2016, Theorem 2.1) and with boxes in (2.10). The two statements are equivalent, but writing it with boxes will be convenient to state the periodic version of the result in Proposition 4.6 below.

Remark 2.11. By Mourrat and Otto (2016, (3.7)) (or explicit computations), we have $\alpha+\beta+\gamma = 1$ as well as the identity

$$\alpha - \frac{(p-d)\gamma}{2} = \frac{d}{2}(\beta+\gamma) = \frac{d}{2}(1-\alpha) \implies 1 - \frac{(p-d)\gamma}{2\alpha} = -\frac{d}{2}\left(1 - \frac{1}{\alpha}\right). \tag{2.13}$$

2.7. *Stochastic integrability for random variables.* We collect the following elementary property regarding the stochastic integrability stochastic processes.

Lemma 2.12. *Let $(X_t)_{t\geq 0}$ be a continuous stochastic process and assume that there exists two constants $C_0 < \infty$ and $c_0 > 0$ and an exponent $a \geq 1$ such that, for any $t \geq 0$ and any $K \geq 0$,*

$$\mathbb{P}\left(|X_t| \geq K\right) \leq C_0 \exp\left(-c_0 K^a\right). \tag{2.14}$$

Then there exist two constants $c_1 := c_1(C_0, c_0) > 0$ and $C_1 := C_1(c_0, C_0) < \infty$ such that for any nonnegative function $f : (0,\infty) \to \mathbb{R}$ satisfying $\int_0^\infty f(x)\,dx = 1$, and for any $K \geq 0$,

$$\mathbb{P}\left(\int_0^\infty f(t)|X_t| \geq K\right) \leq C_1 \exp\left(-c_1 K^a\right). \tag{2.15}$$

Proof: The proof is based on an application of Jensen inequality. Assumption (2.14) implies that there exists two constants $c_1 := c_1(C_0, c_0) > 0$ and $C_1 := C_1(c_0, C_0) < \infty$ such that, for any $t \geq 0$,

$$\mathbb{E}\left[\exp\left(c_1 |X_t|^a\right)\right] \leq C_1.$$

Using the convexity and monotonicity of the map $x \mapsto \exp(c_1 x^a)$ on $[0,\infty)$, we see that

$$\mathbb{E}\left[\exp\left(c_1 \left(\int_0^\infty f(t)|X_t|\,dt\right)^a\right)\right] \leq \mathbb{E}\left[\int_0^\infty f(t)\exp\left(c_1|X_t|^a\right)\,dt\right] \leq C_1,$$

from which we deduce the bound (2.15). □

3. Fluctuation estimates for the Langevin dynamic

This section is devoted to the proofs of two properties of the Langevin dynamic. The first one provides a stochastic integrability estimate on the gradient of the Langevin dynamic, the second one provides a fluctuation estimate for the Langevin dynamic, arguing that it can only remain contained in a fixed interval for a long time with small probability.

3.1. *Stochastic integrability estimate for the discrete gradient of the field.* In this section, we establish stochastic integrability estimates on the gradient of the Langevin dynamic. We first prove in Section 3.1.1 that the tail of the distribution of the discrete gradient of a random surface distributed according to the measure $\mu_{\mathbb{T}_L}$ decays at least like $K \mapsto \exp(-cK^r)$ (where r is the exponent of Assumption 1.1 encoding the growth of V). We then transfer this stochastic integrability from the Gibbs measure to the Langevin dynamic in Section 3.1.2.

3.1.1. *Stochastic integrability for the Gibbs measure.*

Proposition 3.1. *There exist two constants $c := c(d,V) > 0$ and $C := C(d,V) < \infty$ such that, for any $L \geq 1$ and any edge $e \in E\left(\mathbb{T}_L\right)$, if ϕ is a random surface sampled according to $\mu_{\mathbb{T}_L}$, then*

$$\mathbb{P}\left[|\nabla\phi(e)| > K\right] \leq C\exp\left(-cK^r\right).$$

We present below a proof of this proposition based on the Efron's monotonicity theorem for log-concave measure and a coupling argument (originally due to Funaki and Spohn (1997)) for the Langevin dynamic. We mention that, in the case when the potential V is symmetric (i.e., $V(x) = V(-x)$ for all $x \in \mathbb{R}$), an alternative approach, relying on reflection positivity in the form of the chessboard estimate (following Miłoś and Peled (2015) and Magazinov and Peled (2022, Lemma 3.9)), would yield the same result.

Proof: We first prove the upper bound: there exists a constant $C := C(d,V) < \infty$ such that, for any $L \in \mathbb{Z}_+$, any $e \in E\left(\mathbb{T}_L\right)$, if we let ϕ be a random surface sampled according to $\mu_{\mathbb{T}_L}$, then

$$\mathbb{E}\left[\left|\nabla\phi(e)\right|^2\right] + \mathbb{E}\left[\left|V'\left(\nabla\phi(e)\right)\right|^2\right] \leq C. \tag{3.1}$$

The proof of the inequality (3.1) is based on the following identity: for any $x \in \mathbb{T}_L$,

$$\mathbb{E}\left[\phi(x)\nabla\cdot V'(\nabla\phi)(x)\right] = -\frac{|\mathbb{T}_L|-1}{|\mathbb{T}_L|}. \tag{3.2}$$

To prove the identity (3.2), we use the following result: for any probability density $f:\mathbb{R}^n \to [0,\infty)$ which is continuously differentiable, such that $|y|f(y)$ tends to 0 at infinity and $y \to (1+|y|)\nabla f(y)$ is integrable, and for any index $i \in \{1,\ldots,n\}$,

$$\int_{\mathbb{R}^n} y_i \frac{df}{dy_i}(y)\,dy = -1.$$

Applying this result when the underlying space is $\Omega^\circ_{\mathbb{T}_L}$ and noting that the function $\delta_x - \frac{1}{|\mathbb{T}_L|} \in \Omega^\circ_{\mathbb{T}_L}$ has an $L^2(\mathbb{T}_L)$-norm equal to $(\frac{|\mathbb{T}_L|-1}{|\mathbb{T}_L|})^{\frac{1}{2}}$, we obtain

$$\frac{|\mathbb{T}_L|}{|\mathbb{T}_L|-1}\int_{\Omega^\circ_{\mathbb{T}_L}} \phi(x)\nabla\cdot V'(\nabla\phi)(x)\mu_{\mathbb{T}_L}(d\phi) = -1,$$

which is the identity (3.2). Summing the inequality (3.2) over the vertices $x \in \mathbb{T}_L$ and performing a discrete integration by parts, we deduce that

$$\mathbb{E}\left[\sum_{e'\in E(\mathbb{T}_L)} V'(\nabla\phi(e'))\nabla\phi(e')\right] = |\mathbb{T}_L| - 1.$$

Using Assumption 1.1 on the potential V, we see that the previous inequality implies

$$\mathbb{E}\left[\sum_{e'\in E(\mathbb{T}_L)} \left|\nabla\phi(e')\right|^2\right] \leq C\left|\mathbb{T}_L\right|.$$

Using that the spatial stationarity of the distribution $\mu_{\mathbb{T}_L}$ (since we consider the Gibbs measure $\mu_{\mathbb{T}_L}$ in the torus), we deduce that, for any edge $e \in \mathbb{T}_L$,

$$\mathbb{E}\left[\left|\nabla\phi(e)\right|^2\right] \leq \frac{C}{\left|\mathbb{T}_L\right|}\mathbb{E}\left[\sum_{e'\in E(\mathbb{T}_L)} \left|\nabla\phi(e')\right|^2\right] \leq C.$$

We next note that, since the Gibbs measure $\mu_{\mathbb{T}_L}$ is log-concave, the Prékopa-Leindler inequality Prékopa (1971, 1973); Leindler (1972) implies that the distribution of the random variable $\nabla\phi(e)$ is also log-concave. This implies that the tail of its distribution decays exponentially fast on the scale of its standard deviation, and thus all the moments of $\nabla\phi(e)$ are bounded uniformly in L. In particular, since the map V' grows at most like a polynomial, we obtain the bound (3.1). We then fix an edge $e \in E(\mathbb{T}_L)$ and introduce the collection of potentials $(V_{e'})_{e'\in E(\mathbb{T}_L)}$

$$V_{e'}(x) := \begin{cases} V(x) \text{ if } e' \neq e, \\ \dfrac{V(x)}{2} \text{ if } e' = e. \end{cases}$$

We then denote by $\phi^e : \mathbb{T}_L \to \mathbb{R}$ be a random surface distributed according to the Gibbs measure

$$\mu^e_{\mathbb{T}_L}(d\phi) := \frac{1}{Z^e_{\mathbb{T}_L}} \exp\left(-\sum_{e'\in E(\mathbb{T}_L)} V_{e'}\left(\nabla\phi(e')\right)\right) d\phi. \tag{3.3}$$

Since the measure (3.3) is log-concave, the random variable $\nabla\phi^e(e)$ is also log-concave. We next prove the following estimate: there exists a constant $C := C(d,V) < \infty$ such that

$$\mathbb{E}\left[\left|\nabla\phi^e(e)\right|^2\right] \leq C. \tag{3.4}$$

The proof of (3.4) is based on a coupling argument for Langevin dynamic. To this end, we introduce the Langevin dynamic associated with the measure $\mu^e_{\mathbb{T}_L}$, i.e.,

$$\begin{cases} d\phi^e_L(t,x) = \nabla \cdot V'_{e'}(\nabla \phi^e_L)(t,x)dt + \sqrt{2}dB_t(x) & \text{for } (t,x) \in [0,\infty] \times \mathbb{T}_L, \\ \phi^e_L(0,x) = \phi^e(x) & \text{for } x \in \mathbb{T}_L, \end{cases} \tag{3.5}$$

where the initial data ϕ^e is distributed according to the measure $\mu^e_{\mathbb{T}_L}$ and is independent of the Brownian motions. We note that, as it was the case for (2.6), the process $\nabla\phi^e_L$ is stationary with respect to the time translations. We next couple the dynamic (3.5) to the one of (2.6) by assuming that they are driven by the same Brownian motions and that the initial conditions ϕ^e and ϕ are independent. Subtracting the two dynamics, we observe that the difference $u := \phi_L - \phi^e_L$ solves the parabolic equation

$$\partial_t u(t,x) - \nabla \cdot \mathbf{a}_e \nabla u(t,x) = \nabla \cdot \left[\left(V'_{e'} - V' \right) (\nabla \phi_L) \right](t,x) \quad \text{for } (t,x) \in [0,\infty] \times \mathbb{T}_L, \tag{3.6}$$

with the definition

$$\mathbf{a}_e(t,e') := \int_0^1 V''_{e'}(s\nabla\phi_L(t,e') + (1-s)\nabla\phi^e_L(t,e'))\, ds.$$

Noting that the potentials $V_{e'}$ and V are only different at the edge e, we may use an energy estimate on the equation (3.6) and obtain, for any $T \geq 0$,

$$\int_0^T \sum_{e' \in E(\mathbb{T}_L)} \mathbf{a}_e(t,e') \left| \nabla u(t,e') \right|^2 dt \leq C \int_0^T \left| V'\left(\nabla\phi_L(t,e)\right) \nabla u(t,e) \right| dt + C \sum_{x\in\mathbb{T}_L} |u(0,x)|^2 . \tag{3.7}$$

The inequality (3.7) implies the following (weaker) estimate

$$\int_0^T \mathbf{a}_e(t,e) \left| \nabla u(t,e) \right|^2 \leq \int_0^T \left| V'\left(\nabla\phi_L(t,e)\right) \nabla u(t,e) \right| dt + \sum_{x\in\mathbb{T}_L} |u(0,x)|^2 .$$

Assumption 1.1 on the potential V implies that there exists a constant $C := C(V) < \infty$ such that

$$\mathbf{a}_e(t,e) \left| \nabla u(t,e) \right|^2 \geq \left| \nabla u(t,e) \right|^2 - C. \tag{3.8}$$

Substituting (3.8) into (3.7) and applying the Cauchy-Schwarz inequality, we deduce that

$$\int_0^T \left| \nabla u(t,e) \right|^2 dt \leq CT + C \int_0^T V'(\nabla\phi_L(t,e))^2\, dt + C \sum_{x\in\mathbb{T}_L} |u(0,x)|^2 .$$

Using the definition $u := \phi_L - \phi^e_L$, we thus obtain

$$\int_0^T \left| \nabla \phi^e_L(t,e) \right|^2 dt \leq CT + C \int_0^T \left(V'(\nabla\phi_L(t,e))^2 + |\nabla\phi_L(t,e)|^2 \right) dt + C \sum_{x\in\mathbb{T}_L} |u(0,x)|^2 .$$

Taking the expectation in both sides of the previous inequality, and using the stationarity of the gradients $\nabla\phi_L$ and $\nabla\phi^e_L$, we deduce that, for any $T > 0$,

$$\mathbb{E}\left[|\nabla\phi^e_L(0,e)|^2 \right] \leq C + C\mathbb{E}\left[V'(\nabla\phi_L(0,e))^2 + |\nabla\phi_L(0,e)|^2 \right] + \frac{C}{T} \sum_{x\in\mathbb{T}_L} \mathbb{E}\left[|u(0,x)|^2 \right].$$

Taking the limit $T \to \infty$ and using the bound (3.1) completes the proof of (3.4).

We next let Y be a real-valued random variable whose law is given by

$$\mu_Y := \frac{1}{Z_Y} \exp\left(-\frac{1}{2} V\left(y\right) \right) dy \quad \text{with} \quad Z_Y := \int_{\mathbb{R}} \exp\left(-\frac{1}{2} V\left(y\right) \right) dy.$$

We couple the random variables Y and ϕ^e by assuming that they are independent. Using that the law of the random variable Y is explicit, the independence of Y and $\nabla \phi^e(e)$ and the bound (3.4), we deduce that there exists a constant $c := c(d, V) > 0$ such that

$$\begin{aligned} \mathbb{P}\left[Y \geq \nabla \phi^e(e)\right] &\geq \mathbb{P}\left[\{Y \geq 2\mathbb{E}\left[|\nabla \phi^e(e)|\right]\} \cap \{\nabla \phi^e(e) \leq 2\mathbb{E}\left[|\nabla \phi^e(e)|\right]\}\right] \\ &= \mathbb{P}\left[\{Y \geq 2\mathbb{E}\left[|\nabla \phi^e(e)|\right]\}\right) \mathbb{P}\left(\{\nabla \phi^e(e) \leq 2\mathbb{E}\left[|\nabla \phi^e(e)|\right]\}\right] \\ &\geq c. \end{aligned} \tag{3.9}$$

We next rely on the observation that the law of $\nabla \phi(e)$ (where ϕ is distributed according to the measure $\mu_{\mathbb{T}_L}$) is equal to the law of the random variable Y conditionally on the event $\{Y - \nabla \phi^e(e) = 0\}$. This property is a consequence of the following observation: if X and Z are two independent real-valued random variables with bounded continuous densities f and g then the law of X conditionally on the event $\{X - Z = 0\}$ has a density proportional to the function fg. In particular, for any non-negative function $F : \mathbb{R} \to [0, \infty)$, one has the identity

$$\mathbb{E}\left[F(\nabla \phi(e))\right] = \mathbb{E}\left[F(Y) \,|\, Y - \nabla \phi^e(e) = 0\right]. \tag{3.10}$$

We then introduce the constant $c_3 := \frac{c_-}{4r(r-1)} > 0$ (where c_- is the constant appearing in Assumption 1.1) and the function

$$F(x) := \begin{cases} 0 \text{ if } x \leq 0, \\ \exp\left(c_3 x^r\right) \text{ if } x \geq 0. \end{cases}$$

Assumption 1.1 on the potential V implies that there exists a constant $C := C(V) < \infty$ such that

$$\mathbb{E}\left[F(Y)\right] = \frac{1}{Z_Y} \int_{\mathbb{R}} F(y) \exp\left(-\frac{1}{2} V(y)\right) \, dy \leq C. \tag{3.11}$$

We then note that the Efron's monotonicity theorem applied to the pair of independent random variables $(Y, \nabla \phi^e)$, the nonnegativity and monotonicity of the function F imply the almost sure inequality

$$\mathbb{E}\left[F(Y) \,|\, Y - \nabla \phi^e(e) = 0\right] \mathbf{1}_{\{Y - \nabla \phi^e(e) \geq 0\}} \leq \mathbb{E}\left[F(Y) \,|\, Y - \nabla \phi^e(e)\right]. \tag{3.12}$$

Combining the bound (3.11) with the identity (3.11), the lower bound (3.9) and the inequality (3.12) yields the existence of a constant $C := C(d, V) < \infty$ such that

$$\begin{aligned} \mathbb{E}\left[F(\nabla \phi(e))\right] &= \mathbb{E}\left[F(Y) \,|\, Y - \nabla \phi^e(e) = 0\right] \\ &\leq \frac{1}{\mathbb{P}\left(Y - \nabla \phi^e(e) \geq 0\right)} \mathbb{E}\left[\mathbb{E}\left[F(Y) \,|\, Y - \nabla \phi^e(e)\right]\right] \\ &\leq \frac{1}{\mathbb{P}\left(Y - \nabla \phi^e(e) \geq 0\right)} \mathbb{E}\left[F(Y)\right] \\ &\leq C. \end{aligned} \tag{3.13}$$

The inequality (3.13) implies that there exist two constants $C := C(d, V) < \infty$ and $c := c(d, V) > 0$ such that, for any $K \geq 1$,

$$\mathbb{P}\left[\nabla \phi(e) > K\right] \leq C \exp\left(-cK^r\right). \tag{3.14}$$

The same argument can be applied with the potential $\tilde{V}(x) := V(-x)$ to obtain the upper bound, for any $K \geq 1$,

$$\mathbb{P}\left[\nabla \phi(e) < -K\right] \leq C \exp\left(-cK^r\right). \tag{3.15}$$

Combining (3.14) and (3.15) completes the proof of Proposition 3.1. $\square$

3.1.2. *Stochastic integrability for the Langevin dynamic.* In this section, we extend the result of the previous section to the Langevin dynamic (using essentially a union bound and the stationarity of the dynamic).

Proposition 3.2. *There exist two constants $c := c(d, V) > 0$ and $C := C(d, V) < \infty$ such that, for any $T \geq 1$ and any $K \geq 1$,*

$$\mathbb{P}\left[\sup_{t\in[0,T]} |\nabla\phi_L(t, e)| \geq K\right] \leq CT \exp\left(-cK^r\right). \tag{3.16}$$

Proof: Fix $K \geq 1$ and let $N := K^r$. We have the inclusion of events

$$\begin{aligned}\left\{\sup_{t\in[0,T]} |\nabla\phi_L(t, e)| \geq K\right\} \subseteq & \left\{\sup_{n\in\{0,\ldots,\lfloor TN\rfloor\}} \left|\nabla\phi_L\left(\frac{n}{N}, e\right)\right| \geq \frac{K}{2}\right\} \\ & \bigcup \left\{\sup_{n\in\{0,\ldots,\lfloor TN\rfloor\}} \sup_{t\in\left[\frac{n}{N}, \frac{n+1}{N}\right]} \left|\nabla\phi_L(t, e) - \nabla\phi_L\left(\frac{n}{N}, e\right)\right| \geq \frac{K}{2}\right\}. \end{aligned} \tag{3.17}$$

We then bound the probabilities of the two terms in the right-hand side separately. For the first one, a union bound, Proposition 3.1, and the identity $N := K^r$ yield

$$\begin{aligned}\mathbb{P}\left[\sup_{n\in\{0,\ldots,\lfloor TN\rfloor\}} \left|\nabla\phi_L\left(\frac{n}{N}, e\right)\right| \geq \frac{K}{2}\right] & \leq \sum_{n=0}^{\lfloor TN\rfloor} \mathbb{P}\left[\left|\nabla\phi_L\left(\frac{n}{N}, e\right)\right| \geq \frac{K}{2}\right] \\ & \leq CK^rT \exp\left(-cK^r\right) \\ & \leq CT \exp\left(-cK^r\right),\end{aligned} \tag{3.18}$$

where we reduced the value of the constant c in the third line to absorb the polynomial factor K^r. For the second term in the right-hand side of (3.17), we first fix $n \in \{0, \ldots, \lfloor TN\rfloor\}$ and use the definition of the Langevin dynamic (2.6) to write

$$\nabla\phi_L(t, e) - \nabla\phi_L\left(\frac{n}{N}, e\right) = \int_{\frac{n}{N}}^{t} \nabla\left(\nabla\cdot V'(\nabla\phi_L)\right)(s, e)\, ds + \nabla B_t(e) - \nabla B_{\frac{n}{N}}(e).$$

This implies

$$\begin{aligned}&\sup_{t\in\left[\frac{n}{N}, \frac{n+1}{N}\right]} \left|\nabla\phi_L(t, e) - \nabla\phi_L\left(\frac{n}{N}, e\right)\right| \\ &\qquad \leq \int_{\frac{n}{N}}^{\frac{n+1}{N}} \left|\nabla\left(\nabla\cdot V'(\nabla\phi_L)\right)(s, e)\right| ds + \sup_{t\in\left[\frac{n}{N}, \frac{n+1}{N}\right]} \left|\nabla B_t(e) - \nabla B_{\frac{n}{N}}(e)\right|.\end{aligned} \tag{3.19}$$

Using the definition of the discrete gradient and Assumption 1.1 on the potential V, we see that

$$\left|\nabla\left(\nabla\cdot V'(\nabla\phi_L)\right)(s, e)\right| \leq \sum_{e'\cap e\neq\emptyset} \left|V'(\nabla\phi_L)(t, e')\right| \leq C + \sum_{e'\cap e\neq\emptyset} \left|\nabla\phi_L(t, e')\right|^{r-1}.$$

Using Lemma 2.12 (with $f = N\mathbf{1}_{\left[\frac{n}{N}, \frac{n+1}{N}\right]}$), we deduce that

$$\begin{aligned}\mathbb{P}\left[\int_{\frac{n}{N}}^{\frac{n+1}{N}} \left|\nabla\left(\nabla\cdot V'(\nabla\phi_L)\right)(s, e)\right| ds \geq \frac{K}{4}\right] & \leq C \exp\left(-c(NK)^{\frac{r}{r-1}}\right) \\ & \leq C \exp\left(-cK^r\right).\end{aligned} \tag{3.20}$$

Additionally, the supremum of the Brownian motions can be estimated by noting that the difference of two independent Brownian motions is equal in law (up to a multiplicative constant equal to $\sqrt{2}$) to a Brownian motion. We obtain

$$\mathbb{P}\left[\sup_{t\in\left[\frac{n}{N},\frac{n+1}{N}\right]}\left|\nabla B_t(e)-\nabla B_{\frac{n}{N}}(e)\right|\geq\frac{K}{4}\right]=\mathbb{P}\left[\sup_{t\in[0,1]}B_t\geq\frac{\sqrt{N}K}{4\sqrt{2}}\right] \tag{3.21}$$
$$\leq C\exp\left(-cNK^2\right)$$
$$\leq C\exp\left(-cK^r\right).$$

Combining (3.19), (3.20) and (3.21), with a union bound, we have obtained

$$\mathbb{P}\left[\sup_{n\in\{0,\dots,\lfloor TN\rfloor\}}\sup_{t\in\left[\frac{n}{N},\frac{n+1}{N}\right]}\left|\nabla\phi_L(t,e)-\nabla\phi_L\left(\frac{n}{N},e\right)\right|\geq\frac{K}{2}\right] \tag{3.22}$$
$$\leq\sum_{n=0}^{\lceil TN\rceil}\mathbb{P}\left[\sup_{t\in\left[\frac{n}{N},\frac{n+1}{N}\right]}\left|\nabla\phi_L(t,e)-\nabla\phi_L\left(\frac{n}{N},e\right)\right|\geq\frac{K}{2}\right]$$
$$\leq CNT\exp\left(-cK^r\right)$$
$$\leq CK^rT\exp\left(-cK^r\right)$$
$$\leq CT\exp\left(-cK^r\right).$$

Combining (3.17), (3.18) and (3.22) completes the proof of (3.16). □

3.2. *A fluctuation estimate for the Langevin dynamic.* Building upon the stochastic integrability estimate for the dynamic established in Proposition 3.2, we prove that the dynamic cannot remain contained in a deterministic interval for a long time. The argument follows the one outline in Section 1.1.3, with additional technicalities to take into account that the second derivative of the potential V is assumed to be unbounded from above. We recall the definition (2.1) of the constant R_V.

Proposition 3.3 (Fluctuation for the Langevin dynamic)**.** *There exist two constants $C := C(d,V) < \infty$ and $c := c(d,V) > 0$ such that, for any $T\geq 1$ and any edge $e\in\mathbb{T}_L$,*

$$\mathbb{P}\left[\forall t\in[0,T],\ |\nabla\phi_L(t,e)|\leq R_V\right]\leq C\exp\left(-c\left(\ln T\right)^{\frac{r}{r-2}}\right). \tag{3.23}$$

Proof: We fix an edge $e\in E(\mathbb{T}_L)$ and will prove the following estimate: there exist two constants $C := C(d,V) < \infty$ and $c := c(d,V) > 0$ and a time $T_0 := T_0(d,V) < \infty$ such that, for any $T\geq T_0$,

$$\mathbb{P}\left[\forall t\in[0,T],\ |\nabla\phi_L(t,e)|\leq R_V\right]\leq C\exp\left(-c\left(\ln T\right)^{\frac{r}{r-2}}\right). \tag{3.24}$$

The bound (3.23) can be deduced from (3.24) by increasing the value of the constant C. Let us fix a time $T\geq 1$ and let $N := (\ln T)/R_V^2$. The definition of the parameter N is motivated by the following inequality: for any T chosen sufficiently large (universally),

$$\begin{aligned}\mathbb{P}\left[\left|B_{1/N}\right| \geq \frac{4}{3} R_V\right] &= \mathbb{P}\left[|B_1| \geq \frac{4}{3}\sqrt{\ln T}\right] \\ &= \frac{2}{\sqrt{2\pi}} \int_{\frac{4}{3}\sqrt{\ln T}}^{\infty} e^{-\frac{x^2}{2}}\, dx \\ &\geq \frac{2}{\sqrt{2\pi}} \int_{\frac{4}{3}\sqrt{\ln T}}^{\frac{4}{3}\sqrt{\ln T}+1} e^{-\frac{x^2}{2}}\, dx \\ &\geq \frac{2}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{4}{3}\sqrt{\ln T}+1\right)^2\right) \\ &\geq \frac{1}{T^{9/10}}.\end{aligned} \tag{3.25}$$

The proof relies on the observation that a Brownian motion can be decomposed into mutually independent Brownian bridges and increments. To be more specific, we introduce the following sets and notation:

- For each $k \in \mathbb{Z}_+$ and each $x \in \mathbb{T}_L$, we let $W_k(\cdot; x)$ be the Brownian bridge defined by the formula
$$\forall t \in \left[0, \frac{1}{N}\right], \quad W_k(t; x) := B_{t+\frac{k}{N}}(x) - B_{\frac{k}{N}}(x) - Nt(B_{\frac{k+1}{N}}(x) - B_{\frac{k}{N}}(x)).$$
We will denote by $\mathcal{W} := \{W_k(\cdot; x) \,:\, k \in \mathbb{Z}_+, x \in \mathbb{T}_L\}$ the collection of Brownian bridges.
- For each $k \in \mathbb{Z}_+$ and each $y \in \mathbb{T}_L$, we denote by $X_k(y)$ the increment
$$X_k(y) := B_{\frac{k+1}{N}}(y) - B_{\frac{k}{N}}(y).$$
We will denote by $\mathcal{X} := \{X_k(x) \,:\, k \in \mathbb{Z}_+,\, x \in \mathbb{T}_L\}$ the set of all the increments. For $(l, y) \in \mathbb{Z}_+ \times \mathbb{T}_L$, the set $\mathcal{X}_{l,y} := \{X_k(x) \,:\, k \in \mathbb{Z}_+,\, x \in \mathbb{T}_L, k \neq l, x \neq y\}$ denotes the collection of all the increments except $X_l(y)$.

In particular, the Brownian bridges $\{B_t(x) \,:\, x \in \mathbb{T}_L,\, t \geq 0\}$ are fully determined by the Brownian bridges of $\mathcal{W}$ and the increments of $\mathcal{X}$. This implies, using the discussion of Section 2.2, that the dynamic ϕ_L is fully determined by the initial condition ϕ, the Brownian bridges of $\mathcal{W}$ and the increments of $\mathcal{X}$. We thus introduce the notation

$$\mathcal{R} := (\phi, \mathcal{X}, \mathcal{W}).$$

The set of all possible triplets $\mathcal{R}$ will be denoted by

$$\Omega := \Omega^{\circ}_{\mathbb{T}_L} \times \mathbb{R}^{\mathbb{Z}_+ \times \mathbb{T}_L} \times C\left(\left[0, \frac{1}{N}\right], \mathbb{R}\right)^{\mathbb{Z}_+ \times \mathbb{T}_L}.$$

Since the dynamic $\{\phi_L(t, x) \,:\, t \geq 0,\, x \in \mathbb{T}_L\}$ can interpreted as deterministic functions of $\mathcal{R} \in \Omega$, we will write

$$\phi_L(t, x) := \phi_L(t, x)\,(\mathcal{R}).$$

For $(l, y) \in \mathbb{Z}_+ \times \mathbb{T}_L$, we denote by $\mathcal{R}_{l,y} := (\phi, \mathcal{X}_{l,y}, \mathcal{W})$ and by $\Omega_{l,y}$ the set of possible values for $\mathcal{R}_{l,y}$. We have the identities $\mathcal{R} = (X_l(y), \mathcal{R}_{l,y})$ and $\Omega = \mathbb{R} \times \Omega_{l,y}$. To emphasize the dependency of the dynamic on the increment $X_l(y)$, we will write

$$\phi_L(t, x) = \phi_L(t, x)\,(X_l(y), \mathcal{R}_{l,y}). \tag{3.26}$$

We denote by $\mathcal{F}_{\mathcal{R},l,y}$ the σ-algebra generated by $\mathcal{R}_{l,y}$ and note that the increment $X_l(y)$ is independent of the σ-algebra $\mathcal{F}_{\mathcal{R},l,y}$. For later use, we note that the dynamic $\phi_L(t, x)$ depends only on the increments $X_k(y)$ and the Brownian bridges $W_k(\cdot; y)$ such that $t \geq \frac{k}{N}$. This reflects the fact that the dynamic ϕ_L evaluated at the time t depends only on the realization of the Brownian motions before the time t.

We now fix a positively oriented edge $e \in E(\mathbb{T}_L)$ and let y be the second endpoint of e. For any $l \in \mathbb{Z}_+$, we introduce the following random subset of $\mathbb{R}$ (depending on the collection $\mathcal{R}_{l,y}$),

$$\mathcal{A}_l(\mathcal{R}_{l,y}) := \left\{ X \in \mathbb{R} \, : \, \left| \nabla \phi_L \left(\frac{l+1}{N}, e \right) (X, \mathcal{R}_{l,y}) \right| \leq R_V \right\} \subseteq \mathbb{R}, \tag{3.27}$$

where we used the notation introduced in (3.26). In words, the set $\mathcal{A}_l(\mathcal{R}_{l,y})$ is the set of all possible values for the increment $X_l(y)$ such that the gradient of the dynamic ϕ_L computed at time $(l+1)/N$ at the edge e with initial condition, Brownian bridges and increments given by $\mathcal{R} = (X_l(y), \mathcal{R}_{l,y})$ belongs to the interval $[-R_V, R_V]$.

We next introduce the event $A_l \subseteq \Omega$ defined as follows

$$A_l := \left\{ \mathcal{R} := (X_l(y), \mathcal{R}_{l,y}) \in \Omega \, : \, X_l(y) \in \mathcal{A}_l(\mathcal{R}_{l,y}) \text{ and } \frac{1}{\sqrt{2\pi N}} \int_{\mathcal{A}_l(\mathcal{R}_{l,y})} e^{-\frac{x^2}{2N}} \, dx \leq 1 - \frac{1}{T^{9/10}} \right\}. \tag{3.28}$$

Since the law of the increment $X_l(y)$ is Gaussian of variance $1/N$ and since $X_l(y)$ is independent of the set $\mathcal{R}_{l,y}$, we have the almost sure upper bound

$$\mathbb{E}\left[\mathbf{1}_{A_l} \middle| \mathcal{F}_{\mathcal{R},l,y} \right] \leq 1 - \frac{1}{T^{9/10}}. \tag{3.29}$$

We next estimate the probability for the intersection of all the events A_l for $l \in \{0, \ldots, \lfloor NT \rfloor\}$ and prove the following stretched exponential decay in the time T,

$$\mathbb{P}\left[\bigcap_{l=0}^{\lfloor NT \rfloor} A_l \right] \leq \exp\left(-T^{1/10} \right). \tag{3.30}$$

The proof of (3.30) is obtained by consecutive conditioning. We first note that, since the dynamic $\phi_L(t,x)$ depends only on the increments $X_l(y)$ and the Brownian bridges $W_l(\cdot; y)$ such that $t \geq \frac{l}{N}$, the events $(A_0, \ldots, A_{\lfloor NT \rfloor - 1})$ do not depend on the increment $X_{\lfloor NT \rfloor}(y)$, and are thus measurable with respect to the σ-algebra $\mathcal{F}_{\mathcal{R},\lfloor NT \rfloor,y}$. Combining this observation with the upper bound (3.29), we obtain

$$\begin{aligned}
\mathbb{P}\left[\bigcap_{l=0}^{\lfloor NT \rfloor} A_l \right] &= \mathbb{E}\left[\prod_{l=0}^{\lfloor NT \rfloor} \mathbf{1}_{A_l} \right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{l=0}^{\lfloor NT \rfloor} \mathbf{1}_{A_l} \middle| \mathcal{F}_{\mathcal{R},\lfloor NT \rfloor,y} \right] \right] \\
&= \mathbb{E}\left[\left(\prod_{l=0}^{\lfloor NT \rfloor - 1} \mathbf{1}_{A_l} \right) \times \mathbb{E}\left[\mathbf{1}_{A_{\lfloor NT \rfloor}} | \mathcal{F}_{\mathcal{R},\lfloor NT \rfloor,y} \right] \right] \\
&\leq \left(1 - \frac{1}{T^{9/10}} \right) \mathbb{P}\left[\bigcap_{l=0}^{\lfloor NT \rfloor - 1} A_l \right].
\end{aligned}$$

We may then iterate the previous computation, noting that, for any $l \in \{0, \ldots, \lfloor NT \rfloor - 1\}$, the events $(A_1, \ldots, A_l)$ are measurable with respect to the σ-algebra $\mathcal{F}_{\mathcal{R},l+1,y}$. This leads to the upper bound, for T sufficiently large (depending only on V) so that $\lfloor NT \rfloor + 1 \geq T(\ln T)/R_V^2 \geq T$,

$$\mathbb{P}\left[\bigcap_{l=0}^{\lfloor NT \rfloor} A_l\right] \leq \left(1 - \frac{1}{T^{9/10}}\right)^{\lfloor NT \rfloor + 1} \leq \exp\left(-\frac{\lfloor NT \rfloor + 1}{T^{9/10}}\right) \leq \exp\left(-T^{1/10}\right).$$

We next select a time $T_G := T_G(d, V) < \infty$ and a constant $C_G := C_G(d, V) < \infty$ such that the following implication holds: for any $T \geq T_G$,

$$\sum_{e' \cap e \neq \emptyset} \left|\nabla \phi_L(t, e')\right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{C_G} \implies \sum_{e' \cap e \neq \emptyset} \left|\mathbf{a}(t, e')\right| \leq \frac{N}{2}. \tag{3.31}$$

The identity $N := \ln T / R_V^2$ and Assumption 1.1 ensure that the constant C_G and the time T_G exist and are finite. We then define the interval I_T

$$I_T := \left[-\frac{(\ln T)^{\frac{1}{r-2}}}{(16\sqrt{2}d)C_G}, \frac{(\ln T)^{\frac{1}{r-2}}}{(16\sqrt{2}d)C_G}\right]$$

as well as the good event

$$G_T := \left\{\mathcal{R} \in \Omega \,:\, \sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left|\nabla \phi_L(t, e')(\mathcal{R})\right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{2C_G}\right\} \bigcap \bigcap_{k=0}^{\lfloor NT \rfloor} \left\{X_k(y) \in I_T\right\}.$$

We first show that the probability of the event G_T is close to 1. Using Proposition 3.2, that the law of the increments $\{X_k(y) \,:\, 1 \leq k \leq \lfloor NT \rfloor\}$ is Gaussian of variance $1/N = R_V^2/\ln T$ and a union bound on the complement of the event G_T, we obtain

$$\mathbb{P}\left[G_T^c\right] \leq CT \exp\left(-c\,(\ln T)^{\frac{r}{r-2}}\right) + CNT \exp\left(-c\,(\ln T)^{\frac{r}{r-2}}\right) \leq C \exp\left(-c\,(\ln T)^{\frac{r}{r-2}}\right). \tag{3.32}$$

We will now prove the inclusion of events

$$\left\{\mathcal{R} \in \Omega \,:\, \forall t \in [0,T],\, |\nabla \phi_L(t, e)| \leq R_V\right\} \subseteq \bigcap_{l=0}^{\lfloor NT \rfloor} A_l \cup G_T^c. \tag{3.33}$$

Proposition 3.23 is then obtained by combining (3.30), (3.32), (3.33) and a union bound. The rest of the argument is devoted to the proof of (3.33). As mentioned in Section 1.1.3, we first observe from the definition of the Langevin dynamic that the function

$$(X_l(y), \mathcal{R}_{l,y}) \mapsto \phi_L(t, x)(X_l(y), \mathcal{R}_{l,y})$$

is differentiable with respect to the increment $X_l(y)$, and its derivative can be computed in terms of a solution of a parabolic equation. To be more specific, let us introduce the notation

$$w(t, y)(X_l(y), \mathcal{R}_{l,y}) := \frac{\partial \phi_L(t, y)}{\partial X_l(y)}(X_l(y), \mathcal{R}_{l,y}), \tag{3.34}$$

and note that, for any $l \in \mathbb{Z}_+$ and any $t \in \left[\frac{l}{N}, \frac{l+1}{N}\right]$,

$$B_t(y) = \sum_{k=0}^{l-1} X_k(y) + N\left(t - \frac{k}{N}\right) X_l(y) + W_l\left(t - \frac{l}{N}; y\right),$$

which implies the identity, for any $t \in \left[\frac{l}{N}, \frac{l+1}{N}\right]$,

$$dB_t(y) = N X_l(y) dt + dW_l\left(t - \frac{l}{N}; y\right).$$

Substituting the previous identity in the definition (2.6) of the Langevin dynamic and differentiating both sides of the identity by $X_l(y)$, we see that the function w solves the parabolic equation

$$\begin{cases} \partial_t w(t,x) = \nabla \cdot \mathbf{a} \nabla u(t,x) + \sqrt{2} N \mathbf{1}_{\left[\frac{l}{N}, \frac{l+1}{N}\right]}(t) \delta_y(x) & \text{for } (t,x) \in [0,\infty] \times \mathbb{T}_L, \\ w(0,x) = 0 & \text{for } x \in \mathbb{T}_L, \end{cases}$$

with the environment $\mathbf{a}(t,e') := V''(\nabla \phi_L(t,e'))$. Applying Duhamel's principle with the definition of the heat kernel stated in (2.3), we obtain the identity, for any $t \geq \frac{l}{N}$,

$$w(t,x) = \sqrt{2} N \int_{\frac{l}{N}}^{\min\left(\frac{l+1}{N}, t\right)} \left(P_{\mathbf{a}}(t,x;s,y) + \frac{1}{|\mathbb{T}_L|} \right) ds. \tag{3.35}$$

Additionally, the upper and lower bounds (2.4) imply the following estimate on the gradient of the heat kernel, for any edge $e' \in E(\mathbb{T}_L)$ and any pair of times $(t,s) \in (0,\infty)^2$ with $t \geq s$,

$$\left| \nabla P_{\mathbf{a}} \left(t, e'; s, y\right) \right| \leq 1. \tag{3.36}$$

A combination of the previous displays implies the following bound, for any $\mathcal{R} = (X_l(y), \mathcal{R}_{l,y}) \in \Omega$ and any $(t,e') \in (0,\infty) \times E(\mathbb{T}_L)$,

$$\left| \frac{\partial \nabla \phi_L(t,e')}{\partial X_l(y)} (X_l(y), \mathcal{R}_{l,y}) \right| \leq \sqrt{2}. \tag{3.37}$$

We then fix a realization of the randomness $\mathcal{R} := (X_l(y), \mathcal{R}_{l,y}) \in \Omega$ and assume that $\mathcal{R} \in G_T$. We first claim that, for any increment $X \in I_T$,

$$\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| \nabla \phi_L \left(t, e'\right) (X, \mathcal{R}_{l,y}) \right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{C_G}. \tag{3.38}$$

To prove (3.38), we first use (3.37) and deduce that

$$\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| \nabla \phi_L \left(t, e'\right) (X, \mathcal{R}_{l,y}) - \nabla \phi_L \left(t, e'\right) (X_l(y), \mathcal{R}_{l,y}) \right| \leq \sqrt{2}(4d) \left| X - X_l(y) \right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{2C_G}.$$

By the assumption $(X_l(y), \mathcal{R}_{l,y}) \in G_T$, we have that

$$\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| \nabla \phi_L \left(t, e'\right) (X_l(y), \mathcal{R}_{l,y}) \right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{2C_G}.$$

A combination of the two previous displays with the triangle inequality yields, for any $X \in I_T$,

$$\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| \nabla \phi_L \left(t, e'\right) (X, \mathcal{R}_{l,y}) \right| \leq \frac{(\ln T)^{\frac{1}{r-2}}}{C_G}.$$

Using the definition of the constant C_G and the implication (3.31), we have proved the following result: for any $T \geq T_G$, any $\mathcal{R} := (X_l(y), \mathcal{R}_{l,y}) \in G_T$, any increment $X \in I_T$, one has the upper bound

$$\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| \mathbf{a}(t,e')(X, \mathcal{R}_{l,y}) \right| \leq \frac{N}{2}.$$

The previous upper bound is useful as it can be used to control the derivative in time of the heat kernel. Indeed, using the identity $\partial_t P_{\mathbf{a}} = \nabla \cdot \mathbf{a} \nabla P_{\mathbf{a}}$ together with the bound (3.36), we obtain the estimate, for any pair of times $(s,t) \in [0,\infty)^2$,

$$\left| \partial_t \nabla P_{\mathbf{a}}(t,e;s,y) \right| \leq \sum_{e' \cap e \neq \emptyset} \mathbf{a}(t,e') \left| \nabla P_{\mathbf{a}}(t,e';s,y) \right| \leq \sum_{e' \cap e \neq \emptyset} \mathbf{a}(t,e).$$

Combining the two previous displays with the identity $\nabla P_{\mathbf{a}}(s,e;s,y) = 1$ (since y is the second endpoint of e), we obtain that, for any $\mathcal{R} := (X_l(y), \mathcal{R}_{l,y}) \in G_T$ and any increment $X \in I_T$,

$$\begin{aligned}\frac{\partial \nabla\phi_L\left(\frac{l+1}{N}, e\right)}{\partial X_l(y)}(X, \mathcal{R}_{l,y}) &= \sqrt{2}N \int_{\frac{l}{N}}^{\frac{l+1}{N}} \nabla P_{\mathbf{a}}\left(\frac{l+1}{N}, e; s, y\right)(X, \mathcal{R}_{l,y})\, ds \\ &\geq \sqrt{2}N \int_{\frac{k}{N}}^{\frac{k+1}{N}} 1 - \frac{N}{2}\left(\frac{k+1}{N} - s\right) ds \\ &\geq \frac{3}{4}.\end{aligned}$$

This lower bound on the derivative of the gradient of the dynamic implies that, for any $(X_l(y), \mathcal{R}_{l,y}) \in G_T$, the function

$$X \mapsto \nabla\phi_L\left(\frac{l+1}{N}, e\right)(X, \mathcal{R}_{l,y}) - \frac{3}{4}X \text{ is increasing on the interval } I_T.$$

This implies the following upper bound on the Lebesgue measure of the set $\mathcal{A}_l(\mathcal{R}_{l,y}) \cap I_T$,

$$|\mathcal{A}_l(\mathcal{R}_{l,y}) \cap I_T| \leq \frac{8}{3}R_V,$$

which then yields the estimate, for any T sufficiently large (depending on d and V) so that the computation (3.25) applies

$$\frac{1}{\sqrt{2\pi N}} \int_{\mathcal{A}_l(\mathcal{R}_{l,y})} e^{-\frac{x^2}{2N}}\, dx \leq 1 - \frac{1}{\sqrt{2\pi N}} \int_{I_T \setminus [-\frac{4}{3}R_V, \frac{4}{3}R_V]} e^{-\frac{x^2}{2N}}\, dx \leq 1 - \frac{1}{T^{9/10}}.$$

From the definitions (3.27) and (3.28), the previous inequality implies the identity, for any $l \in \{1, \ldots, \lfloor NT \rfloor\}$,

$$G_T \cap A_l = G_T \cap \left\{\mathcal{R} \in \Omega : \left|\nabla\phi_L\left(\frac{l+1}{N}, e\right)(\mathcal{R})\right| \leq R_V\right\}.$$

Taking the intersection over $l \in \{1, \ldots, \lfloor NT \rfloor\}$ completes the proof of (3.33). □

4. On diagonal upper bound for the heat kernel

In this section, we combine the result of Section 3 with the techniques developed by Mourrat and Otto (2016) to obtain an on-diagonal upper bound for the heat kernel appearing in the Helffer-Sjöstrand representation formula. The section is organized as follows. In Section 4.1, we collect some preliminary definitions and results and state the main technical result of the section (pertaining to the decay rate of the L^2-norm of the heat kernel) in Theorem 4.2. Section 4.2, Section 4.4 and Section 4.5 are devoted to the proof of Theorem 4.2 following the techniques of Mourrat and Otto (2016). The on-diagonal upper bound on the heat kernel is deduced from Theorem 4.2 in Section 4.6. Finally, Section 4.7 completes the proof of Theorem 1.2 by combining the on-diagonal heat kernel estimate with the Helffer-Sjöstrand representation formula.

4.1. *Preliminaries.* We select two exponents $p, p' \in (d, \infty)$ depending only on the dimension d. These exponents will be used to define the moderated environment and apply the anchored Nash inequality, any specific values are admissible (for instance, one can choose $p = p' = d + 1$). We let ϕ_L be the Langevin dynamic in the torus and let $\mathbf{a} := V''(\nabla\phi_L)$ be the environment appearing in the Helffer-Sjöstrand representation formula. Using the stationarity of the gradient of the Langevin dynamic, Proposition 3.1 and the growth condition assumed on the second derivative V'', we know that all the moments of the random environment $\mathbf{a}$ are finite: for any $q \in [1, \infty)$, and any $(t, e) \in (0, \infty) \times E(\mathbb{T}_L)$,

$$\mathbb{E}\left[\mathbf{a}(t,e)^q\right] < \infty. \tag{4.1}$$

Following the insight of Mourrat and Otto (2016), we introduce in this section the following moderated environment w. We first introduce the two functions

$$k_t := \frac{\delta}{(1+t)^{p+3}} \quad \text{and} \quad K_t := k_t + \int_t^\infty s k_s \, ds, \tag{4.2}$$

where $\delta := \delta(d) > 0$ is chosen sufficiently small so that, for any $t, s' \in (0, \infty)$ with $s' \geq t$,

$$\int_t^{s'} K_{s-t} K_{s'-s} \, ds \leq K_{s'-t} \text{ and } \int_0^\infty K_s \, ds \leq 1. \tag{4.3}$$

Using the function k, we define the moderated environment w as follows.

Definition 4.1 (Moderated environment for the Langevin dynamic)**.** We define the *moderated environment* according to the formula, for any $(t, e) \in [0, \infty) \times E(\mathbb{T}_L)$,

$$w(t,e)^2 = \int_t^\infty k_{s-t} \frac{\mathbf{a}(s,e) \wedge 1}{(s-t)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s', e') \vee 1 \, ds'} \, ds. \tag{4.4}$$

Compared to the environment $\mathbf{a}$, the moderated environment w satisfies the property that all the moments of w and of w^{-1} are finite, and we will prove in Proposition 4.4 that, for any $q \in [1, \infty)$, and any $(t, e) \in (0, \infty) \times E(\mathbb{T}_L)$,

$$\mathbb{E}\left[w(t,e)^q\right] + \mathbb{E}\left[w(t,e)^{-q}\right] < \infty. \tag{4.5}$$

This result is proved in Proposition 4.4, and the proof builds upon the fluctuation estimate for the Langevin dynamic proved in Proposition 3.3.

Various functionals of the environments $\mathbf{a}$ and w will appear in the proof of the heat kernel estimate. They are collected below. Their formulae are technical, and we incite the reader to consult as a reference. They all possess the property they have finite moments of all order (see (4.6) and (4.8)).

Before stating their definition, we recall the definitions of the exponents introduced in (2.7) and (2.8), let θ_c be the exponent given by (2.11) of Proposition 2.9 (with the values of $p, p' \in (d, \infty)$ selected at the beginning of this section), and let α, β, γ be the exponents defined in (2.12) (with $\theta = \theta_c$). For any time $t \geq 0$, we introduce the six random variables

$$\begin{cases} \mathcal{M}_{p'}(t) := 1 + \left(1 + \|w(t,\cdot)\|_{\underline{L}^{\sigma_d}(\mathbb{T}_L)} \|w^{-1}(t,\cdot)\|_{\underline{L}^{\tau_d}(\mathbb{T}_L)}\right)^2 \sup\limits_{r \in \{1,\ldots,L\}} \|w^{-1}(t,\cdot)\|^2_{\underline{L}^{p'}(\Lambda_r)}, \\ \mathcal{M}_0(t) := 1 + \sup\limits_{r \in \{0,\ldots,L\}} \|\mathbf{a}(t,\cdot)^{1/2}\|^2_{\underline{L}^{\sigma_d}(\Lambda_r)} \left(1 + \|w^{-1}(t,\cdot)\|^2_{\underline{L}^{\tau'_d}(\Lambda_r)}\right), \\ \mathcal{M}_1(t) := 1 + \displaystyle\int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{p}{2(1-\theta_d)}} \, ds, \\ \mathcal{M}_2(t) := 1 + \sup\limits_{x \in \mathbb{T}_L} \dfrac{\sum_{e \ni x} \mathbf{a}(t,e)}{|x|_*^{(p-2)/(p-1)}}, \\ \mathcal{M}_3(t) := 1 + \left(\displaystyle\int_t^\infty K_{s-t} \mathcal{M}_{p'}(s)^{\frac{\alpha}{\beta}} \, ds\right)^\beta, \\ \mathcal{M}_4(t) := 1 + \sup\limits_{s \in [t,t+1]} \left\|w(s,\cdot)^{-1}\right\|^2_{\underline{L}^d(\mathbb{T}_L)}. \end{cases}$$

These six random variables appear at different stages of the proof. The term "+1" is added to the definition to ensure that they are always larger than 1. Their main key property is that they have finite moments of every order: for any $i \in \{0, 1, 2, 3, 4\}$, any $q \in [1, \infty]$ and any $t \geq 0$,

$$\mathbb{E}\left[\mathcal{M}_{p'}(t)^q\right] + \mathbb{E}\left[\mathcal{M}_i(t)^q\right] < \infty. \tag{4.6}$$

The proof of (4.6) is a consequence of the bounds (4.1) and (4.5), the Jensen inequality and the L^p-maximal inequality stated in Proposition 2.8. Building upon these definitions, we consider the maximal functions

$$\begin{cases} \mathscr{M}_1 := \left(\sup_{t \geq 1} \frac{1}{t} \int_0^t \mathcal{M}_1(t)^{\frac{2}{p}} \, dt \right)^{\frac{p}{2}}, \\ \mathscr{M}_2 := \left(\sup_{t \geq 1} \frac{1}{t} \int_0^t \mathcal{M}_2(t) \, dt \right)^{p-1}, \\ \mathscr{M}_3 := \left(\inf_{t \geq 1} \frac{1}{t} \int_0^t \mathcal{M}_3^{-\frac{1}{\alpha}}(s) \, ds \right)^{-\frac{\alpha}{\gamma}}, \\ \mathscr{M}_4 := \left(\inf_{t \geq 1} \frac{1}{t} \int_0^t \mathcal{M}_4^{-1}(s) \, ds \right)^{-1}. \end{cases} \tag{4.7}$$

From the bound (4.6) and the maximal inequality (with respect to the time variable) stated in Proposition 2.8 and the Jensen inequality, we know that all the moments of the random variables listed in (4.7) are finite, i.e., for any $i \in \{1, 2, 3, 4\}$ and any $q \in [1, \infty)$,

$$\mathbb{E}\left[\mathscr{M}_i^q\right] < \infty. \tag{4.8}$$

Finally, building on these definitions, we may define the random variables $\mathscr{M}$ and $\mathscr{M}'$ appearing in the definition of Theorem 4.2 above according to the formulae

$$\mathscr{M} := ((\mathscr{M}_1 + \mathscr{M}_2)\mathscr{M}_3)^{\frac{\gamma}{1-\alpha-\gamma}} \text{ and } \mathscr{M}' := \mathscr{M}_3^{\frac{2\gamma}{(d\beta+p\gamma)} + \frac{\gamma}{\alpha}} \mathscr{M}_4.$$

The inequality (4.8) implies that all the moments of $\mathscr{M}$ and $\mathscr{M}'$ are finite. The main theorem of this section investigates the decay of the $L^2(\mathbb{T}_L)$-norm of the heat kernel. It can be compared to Mourrat and Otto (2016, Theorem 4.2).

Theorem 4.2 (Energy upper bound for dynamic environment)**.** *There exists a constant $C := C(d) < \infty$ such that, for any $t \geq 1$,*

$$\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t, x)^2 \leq \frac{C\mathscr{M}}{(1+t)^{\frac{d}{2}}} \exp\left(-\frac{t}{C\mathscr{M}' L^2} \right).$$

4.2. *Moderation of the environment.* In this section, we adapt the arguments of Mourrat and Otto Mourrat and Otto (2016, Proposition 4.6) to environments which are not bounded from above. Using the terminology introduced in Mourrat and Otto (2016, Definition 3.1), we show that the environment $\mathbf{a}$ is (w, CK)-moderate. The proof of Proposition 4.3 is a notational modification of Mourrat and Otto (2016) and is written below for completeness.

Proposition 4.3 ((w, CK)-moderation)**.** *There exists a constant $C := C(d) > 0$ such that, for every $t \geq 0$ and every solution $u : (0, \infty) \times \mathbb{T}_L \to \mathbb{R}$ of the parabolic equation*

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \text{ in } (0, \infty) \times \mathbb{T}_L,$$

one has the inequality, for any edge $e \in E(\mathbb{T}_L)$,

$$w(t, e)^2 (\nabla u(t, e))^2 \leq C \sum_{e' \cap e \neq \emptyset} \int_t^\infty K_{t-s} \mathbf{a}(s, e') (\nabla u(s, e'))^2 \, ds. \tag{4.9}$$

Proof: Following the proof of Mourrat and Otto (2016, Proposition 4.6), we fix an edge $e \in E(\mathbb{T}_L)$ and first estimate

$$w(t,e)^2(\nabla u(t,e))^2 = \int_t^\infty k_{s-t} \frac{\mathbf{a}(s,e) \wedge 1}{(s-t)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') \vee 1 \, ds'} (\nabla u(t,e))^2 \, ds \tag{4.10}$$
$$\leq 2 \int_t^\infty k_{s-t} \frac{\mathbf{a}(s,e) \wedge 1}{(s-t)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') \vee 1 \, ds'} (\nabla u(s,e))^2 \, ds$$
$$+ 2 \int_t^\infty k_{s-t} \frac{\mathbf{a}(s,e) \wedge 1}{(s-t)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') \vee 1 \, ds'} (\nabla u(s,e) - \nabla u(t,e))^2 \, dt.$$

The first term in the right-hand side can be estimated as follows

$$\int_t^\infty k_{s-t} \frac{\mathbf{a}(s,e) \wedge 1}{(t-s)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') \vee 1 \, ds'} (\nabla u(s,e))^2 \, ds \leq \int_t^\infty k_{s-t} \mathbf{a}(s,e) (\nabla u(s,e))^2 \, ds. \tag{4.11}$$

We next use the identity $\partial_t u = \nabla \cdot \mathbf{a} \nabla u$ and denote by x and y the two endpoints of e. We then write

$$(\nabla u(s,e) - \nabla u(t,e))^2 \leq 2(u(s,x) - u(t,x))^2 + 2(u(s,y) - u(t,y))^2$$
$$\leq 2 \left(\int_t^s \nabla \cdot \mathbf{a} \nabla u(s',x) \, ds' \right)^2 + 2 \left(\int_t^s \nabla \cdot \mathbf{a} \nabla u(s',y) \, ds' \right)^2.$$

We next observe that, by the Cauchy-Schwarz inequality,

$$\left(\int_t^s \nabla \cdot \mathbf{a} \nabla u(s',x) \, ds' \right)^2 \leq C \sum_{e' \ni x} \left(\int_t^s \left| \mathbf{a}(s',e') \nabla u(s',e') \right| \, ds' \right)^2$$
$$\leq C \sum_{e' \ni x} \left(\int_t^s \mathbf{a}(s',e') \, ds' \right) \left(\int_t^s \mathbf{a}(s',e') (\nabla u(s',e'))^2 \, ds' \right).$$

A combination of the two previous displays yields

$$(\nabla u(s,e) - \nabla u(t,e))^2 \leq C \sum_{e' \cap e \neq \emptyset} \left(\int_t^s \mathbf{a}(s',e') \, ds' \right) \left(\int_t^s \mathbf{a}(s',e') (\nabla u(s',e'))^2 \, ds' \right).$$

We thus obtain

$$\frac{\mathbf{a}(s,e) \wedge 1}{(s-t)^{-1} \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') \vee 1 \, ds'} (\nabla u(s,e) - \nabla u(t,e))^2$$
$$\leq C(s-t) \sum_{e' \cap e \neq \emptyset} \int_t^s \mathbf{a}(s',e') (\nabla u(s',e'))^2 \, ds'.$$

Combining the previous estimate with (4.10) and (4.11), we deduce that

$$w(t,e)^2 (\nabla u(t,e))^2 \leq C \int_t^\infty k_{s-t} \mathbf{a}(s,e) (\nabla u(s,e))^2 \, ds$$
$$+ C \sum_{e' \cap e \neq \emptyset} \int_t^\infty k_{s-t} (s-t) \int_t^s \mathbf{a}(s',e') (\nabla u(s',e'))^2 \, ds'$$
$$\leq C \sum_{e' \cap e \neq \emptyset} \int_t^\infty K_{s-t} \mathbf{a}(s,e') (\nabla u(s,e'))^2 \, ds.$$

The proof of Proposition 4.3 is complete. □

4.3. *Stochastic integrability for the moderated environment.* In this section, we establish stochastic integrability estimates for the moderated environment w, and prove that all the moments of w and w^{-1} are finite.

Proposition 4.4 (Stochastic integrability for the moderated environment)**.** *There exist two constants $c := c(d, V) > 0$ and $C := C(d, V) < \infty$ such that, for any $T \geq 1$, any time $t \geq 0$ and any edge $e \in E(\mathbb{T}_L)$,*

$$\mathbb{P}\left[w(t,e) \leq \frac{1}{T}\right] \leq C \exp\left(-c(\ln T)^{\frac{r}{r-2}}\right) \tag{4.12}$$

and

$$\mathbb{P}\left[w(t,e) \geq T\right] \leq C \exp\left(-cT^{\frac{2r}{r-2}}\right). \tag{4.13}$$

Remark 4.5. Proposition 4.4 implies that, for any exponent $q > 0$, any time $t \geq 0$ and any edge $e \in E(\mathbb{T}_L)$,

$$\mathbb{E}\left[w(t,e)^q\right] + \mathbb{E}\left[w(t,e)^{-q}\right] < \infty.$$

Proof: We first prove (4.13). By the stationarity of the gradient of the Langevin dynamic ϕ_L, it is sufficient to prove the result for $t = 0$. We first prove the following inclusion of events: there exists $c := c(d, V) > 0$ such that, for any $T \geq 1$,

$$\begin{aligned}\left\{w(0,e)^2 \leq \frac{c}{T^{p+5}}\right\} \subseteq &\left\{\sup_{t\in[0,T]} \left|\nabla\phi_L(t,e)\right| \leq R_V\right\} \\ &\bigcup \left\{\sup_{t\in[0,T]} V''\left(\nabla\phi_L(t,e)\right) + \sum_{e'\cap e\neq\emptyset} \left|V'\left(\nabla\phi_L\left(t,e'\right)\right)\right| \geq \frac{T}{2}\right\} \\ &\bigcup \left\{\sup_{\substack{t,t'\in[0,T] \\ |t-t'|\leq\frac{1}{T}}} \left|\nabla B_{t'}(e) - \nabla B_t(e)\right| \geq \frac{1}{2}\right\}.\end{aligned} \tag{4.14}$$

The inclusion (4.14) states that, in order for $w(0,e)$ to be small, the dynamic $\nabla\phi_L(\cdot,e)$ has to stay in the interval $[-R_V, R_V]$ for a long time (this behavior is ruled out by Proposition 3.3), or must behave very irregularly, this condition is represented by the second and third events in the right-hand side of (4.14), and can only happen with small probability.

We first prove (4.14). To this end, we will prove the following implication: there exists $c := c(d, V) > 0$ such that, for any $T \geq 1$,

$$\begin{aligned}&\sup_{t\in[0,T]} \left|\nabla\phi_L(t,e)\right| \geq R_V, \ \sup_{t\in[0,T]} V''\left(\nabla\phi_L(t,e)\right) + \sum_{e'\cap e\neq\emptyset} \left|V'\left(\nabla\phi_L\left(t,e'\right)\right)\right| \leq \frac{T}{2} \\ &\qquad \text{and} \sup_{\substack{t,t'\in[0,T] \\ |t-t'|\leq\frac{1}{T}}} \left|\nabla B_{t'}(e) - \nabla B_t(e)\right| \leq \frac{1}{2} \implies w(0,e)^2 \geq \frac{c}{T^{p+5}}.\end{aligned} \tag{4.15}$$

We assume that the event in the left-hand side is satisfied and let $t \in [0, T]$ be such that $|\nabla \phi_L(t,e)| \geq R_V$. Using the definition of the Langevin dynamic (2.6), we see that, for any time $s \in [t - \frac{1}{2T}, t + \frac{1}{2T}]$,

$$\begin{aligned} |\nabla \phi_L(s,e) - \nabla \phi_L(t,e)| &\leq \left| \int_t^s \nabla \left(\nabla \cdot V'(\nabla \phi_L) \right)(s', e) \, ds' \right| + |\nabla B_t(e) - \nabla B_s(e)| \\ &\leq \int_{t - \frac{1}{2T}}^{t + \frac{1}{2T}} \sum_{e' \cap e \neq \emptyset} \left| V' \left(\nabla \phi_L \left(s', e' \right) \right) \right| \, ds' + \frac{1}{2} \\ &\leq 1. \end{aligned}$$

Using the assumption $R_V \geq 2$ which follows from its definition (2.1), we deduce that, for any $s \in \left[t - \frac{1}{2T}, t + \frac{1}{2T}\right]$, $|\nabla \phi_L(s,e)| \geq \frac{R_V}{2}$. This implies, for any $s \in \left[t - \frac{1}{2T}, t + \frac{1}{2T}\right]$,

$$\mathbf{a}(s,e) = V''(\nabla \phi_L(s,e)) \geq 1.$$

The left-hand side of (4.15) yields the upper bound, for any $s \in [0, T]$,

$$\mathbf{a}(s,e) \leq \frac{T}{2}.$$

A combination of the two previous displays with the definition of w stated in (4.4) and the definition of k stated in (4.2) implies, for any $T \geq 1$,

$$\begin{aligned} w(0,e)^2 &= \int_0^\infty k_s \frac{\mathbf{a}(s,e) \wedge 1}{s^{-1} \sum_{e' \cap e \neq \emptyset} \int_0^s \mathbf{a}(s', e') \vee 1 \, ds'} \, ds \\ &\geq \int_{t - \frac{1}{2T}}^{t + \frac{1}{2T}} k_s \frac{\mathbf{a}(s,e) \wedge 1}{s^{-1} \sum_{e' \cap e \neq \emptyset} \int_0^s \mathbf{a}(s', e') \vee 1 \, ds'} \, ds \\ &\geq \frac{2}{T} \int_{t - \frac{1}{2T}}^{t + \frac{1}{2T}} k_s \\ &\geq \frac{c}{T^{p+5}}. \end{aligned}$$

The proof of (4.15), and thus of (4.14) is complete. We next estimate the probabilities of the three events in the right-hand side of (4.14). For the first one, we use Proposition 3.3 and write, for any $T \geq 1$,

$$\mathbb{P} \left(\sup_{t \in [0,T]} |\nabla \phi_L(t,e)| \leq R_V \right) \leq C \exp \left(-c \left(\ln T \right)^{\frac{r}{r-2}} \right).$$

For the second term, we use Assumption 1.1 on the potential V and Proposition 3.2 to obtain that, for any $T \geq 1$,

$$\begin{aligned} &\mathbb{P} \left(\sup_{t \in [0,T]} \sum_{e' \cap e \neq \emptyset} \left| V' \left(\nabla \phi_L \left(t, e' \right) \right) \right| + V'' \left(\nabla \phi_L \left(t, e \right) \right) \geq \frac{T}{2} \right) \\ &\quad \leq CT \exp \left(-cT^{\frac{r}{r-1}} \right) + CT \exp \left(-cT^{\frac{r}{r-2}} \right) \\ &\quad \leq C \exp \left(-cT^{\frac{r}{r-1}} \right). \end{aligned}$$

For the third term, we note that, for any $T \geq 1$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{\substack{t,t' \in [0,T] \\ |t-t'| \leq \frac{1}{T}}} \left|\nabla B_{t'}(e) - \nabla B_t(e)\right| \geq \frac{1}{2}\right) &\leq \sum_{l=0}^{\lceil T^2 \rceil} \mathbb{P}\left(\sup_{t \in \left[\frac{l-1}{T}, \frac{l+1}{T}\right]} \left|\nabla B_t(e) - \nabla B_{\frac{l}{T}}(e)\right| \geq \frac{1}{4}\right) \\
&\leq (T^2+2)\mathbb{P}\left(\sup_{t \in \left[0, \frac{2}{T}\right]} \left|\nabla B_t(e) - \nabla B_{\frac{1}{T}}(e)\right| \geq \frac{1}{4}\right) \\
&\leq (T^2+2)\mathbb{P}\left(\sup_{t \in [0,2]} \left|\nabla B_t(e) - \nabla B_1(e)\right| \geq \frac{\sqrt{T}}{4}\right) \\
&\leq C(T^2+2)\exp(-cT).
\end{aligned}$$

Combining the three previous displays with (4.14) yields, for any $T \geq 1$,

$$\begin{aligned}
\mathbb{P}\left(w(0,e) \leq \frac{c}{T^{p+5}}\right) &\leq C\exp\left(-c(\ln T)^{\frac{r}{r-2}}\right) + C\exp\left(-cT^{\frac{r}{r-1}}\right) + C(T^2+1)\exp(-cT) \\
&\leq C\exp\left(-c(\ln T)^{\frac{r}{r-2}}\right).
\end{aligned} \tag{4.16}$$

This implies (4.12). To prove (4.13), we note that, using the stationarity of the gradient of the Langevin dynamic and Assumption 1.1, for any $t \geq 0$ and any $e \in E(\mathbb{T}_L)$,

$$\mathbb{P}\left[\mathbf{a}(t,e) \geq T\right] \leq C\exp\left(-cT^{\frac{r}{r-2}}\right).$$

We next observe that, from the definitions (4.2) and (4.4),

$$w(t,e)^2 \leq \int_t^\infty k_{s-t}\mathbf{a}(s,e)\,ds \text{ and } \int_t^\infty k_{s-t}\,ds = \int_0^\infty k_s\,ds < \infty.$$

Combining the two previous displays with Lemma 2.12 (and $f(t) = k_t / \int_0^\infty k_s\,ds$) completes the proof of (4.13). $\square$

4.4. *Anchored Nash estimate in the torus.* In this section, we prove a finite-volume version of the anchored Nash estimate of Mourrat and Otto (2016, Theorem 2.1) (see Theorem 2.9). The result is stated below and we emphasize that it only requires a minor adaptation of the proof of Mourrat and Otto (2016, Theorem 2.1).

Proposition 4.6 (Anchored Nash estimate on the torus)**.** *There exists $C := C(d) < \infty$ such that, for any function $u : \mathbb{T}_L \to \mathbb{R}$ satisfying $\sum_{x \in \mathbb{T}_L} u(x) = 0$ and any time $t \geq 0$,*

$$\|u\|_{L^2(\mathbb{T}_L)} \leq C\left(\mathcal{M}_{p'}(t)^{\frac{1}{2}}\|w(t,\cdot)\nabla u\|_{L^2(\mathbb{T}_L)}\right)^\alpha \|u\|_{L^1(\mathbb{T}_L)}^\beta \||x|_*^{p/2}u\|_{L^2(\mathbb{T}_L)}^\gamma,$$

Proof: In this proof, we fix a time $t \geq 0$, identify the torus $\mathbb{T}_L$ with the box Λ_L and extend the functions u and the moderated environments w periodically to the lattice $\mathbb{Z}^d$. The periodicity of w implies that there exists a constant $c := c(d) > 0$ such that, using the notation (2.10) for the maximal function,

$$cM(w^{-p'}(t,\cdot))^{\frac{1}{p'}} \leq \sup_{r \in \{1,\ldots,L\}} \|w^{-1}(t,\cdot)\|_{\underline{L}^{p'}(\Lambda_r)} \leq M(w^{-p'}(t,\cdot))^{\frac{1}{p'}}. \tag{4.17}$$

We then let $\eta : \mathbb{Z}^d \to \mathbb{R}$ be a cutoff function satisfying

$$\mathbf{1}_{\Lambda_L} \leq \eta \leq \mathbf{1}_{\Lambda_{\frac{4}{3}L}} \text{ and } |\nabla\eta| \leq \frac{C}{L}. \tag{4.18}$$

We next apply Theorem 2.9 with the finitely supported function $\eta u : \mathbb{Z}^d \to \mathbb{R}$ and use the lower bound (4.17) to deduce that

$$\begin{aligned} &\|\eta u\|_{L^2(\mathbb{Z}^d)} \\ &\leq C \left(\left(\sup_{r \in \{1,\ldots,L\}} \|w^{-1}(t,\cdot)\|_{\underline{L}^{p'}(\Lambda_r)} \right) \|w(t,\cdot)\nabla(\eta u)\|_{L^2(\mathbb{Z}^d)} \right)^{\alpha} \|\eta u\|^{\beta}_{L^1(\mathbb{Z}^d)} \||x|_*^{p/2} \eta u\|^{\gamma}_{L^2(\mathbb{Z}^d)}. \end{aligned} \tag{4.19}$$

Using the periodicity of the function u and the definition of the cutoff function η, we have the upper bounds, for some $C := C(d) < \infty$,

$$\|u\|_{L^2(\mathbb{T}_L)} \leq \|\eta u\|_{L^2(\mathbb{Z}^d)}, \ \|\eta u\|_{L^1(\mathbb{Z}^d)} \leq C \|u\|_{L^1(\mathbb{T}_L)}, \text{ and } \||x|_*^{p/2} \eta u\|_{L^2(\mathbb{Z}^d)} \leq C \||x|_*^{p/2} u\|_{L^2(\mathbb{T}_L)}. \tag{4.20}$$

So that there only remains to treat the term $\|w\nabla(\eta u)\|_{L^2(\mathbb{Z}^d)}$. Expanding the discrete gradient and using the properties of the cutoff function η stated in (4.18), we obtain

$$\sum_{e \in E(\mathbb{Z}^d)} w(t,e)^2 \nabla(\eta u)(e)^2 \leq C \sum_{e \in E(\mathbb{Z}^d)} (w(t,e)\eta(e)\nabla u(e))^2 + \frac{C}{L^2} \sum_{e \in E(\Lambda_{\frac{4}{3}L})} (w(t,e)u(e))^2. \tag{4.21}$$

where we recall the notation $\eta(e) = (\eta(x) + \eta(y))/2$ and $u(e) = (u(x) + u(y))/2$ for $e = (x,y) \in E(\mathbb{Z}^d)$. Using the periodicity of the functions u and w, we may rewrite the previous inequality as follows

$$\sum_{e \in E(\mathbb{T}_L)} w(t,e)^2 \nabla(\eta u)(e)^2 \leq C \sum_{e \in E(\mathbb{T}_L)} (w(t,e)\nabla u(e))^2 + \frac{C}{L^2} \sum_{e \in E(\mathbb{T}_L)} (w(t,e)u(e))^2.$$

We then estimate the second term in the right-hand side. To this end, we will use the Hölder inequality and the Gagliardo-Nirenberg-Sobolev inequality. We recall the definitions of the four exponents $\lambda_d, \kappa_d, \sigma_d$ and τ_d introduced in (2.7), and apply first the Hölder inequality and then the Gagliardo-Nirenberg-Sobolev inequality (Proposition 2.6 using that $\sum_{x \in \mathbb{T}_L} u(x) = 0$), then the Hölder inequality. We deduce that

$$\begin{aligned} \left(\frac{1}{|\mathbb{T}_L|} \sum_{e \in E(\mathbb{T}_L)} (w(t,e)u(e))^2 \right)^{\frac{1}{2}} &\leq \|w(t,\cdot)\|_{\underline{L}^{\sigma_d}(\mathbb{T}_L)} \|u\|_{\underline{L}^{\kappa_d}(\mathbb{T}_L)} \\ &\leq CL \|w(t,\cdot)\|_{\underline{L}^{\sigma_d}(\mathbb{T}_L)} \|\nabla u\|_{\underline{L}^{\lambda_d}(\mathbb{T}_L)} \\ &\leq CL \|w(t,\cdot)\|_{\underline{L}^{\sigma_d}(\mathbb{T}_L)} \left\|w^{-1}(t,\cdot)\right\|_{\underline{L}^{\tau_d}(\mathbb{T}_L)} \|w(t,\cdot)\nabla u\|_{\underline{L}^2(\mathbb{T}_L)}. \end{aligned} \tag{4.22}$$

Combining (4.22) with (4.21), we deduce that

$$\|w(t,\cdot)\nabla(\eta u)\|_{L^2(\mathbb{Z}^d)} \leq C(1 + \|w\|_{\underline{L}^{\sigma_d}(\mathbb{T}_L)} \|w^{-1}(t,\cdot)\|_{\underline{L}^{\tau_d}(\mathbb{T}_L)}) \|w(t,\cdot)\nabla u\|_{L^2(\mathbb{T}_L)}.$$

Combining the previous display with (4.19) and (4.20) completes the proof of Proposition 4.6. □

4.5. *Estimate on the L^2-norm of the heat kernel.* Following Mourrat and Otto (2016, Proof of Theorem 3.2), we introduce the notation

$$\mathcal{E}_t := \sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t,x)^2, \ \mathcal{D}_t = \sum_{e \in E(\mathbb{T}_L)} \mathbf{a}(t,e)(\nabla P_{\mathbf{a}}(t,e))^2 \text{ and } \mathcal{N}_t := \sum_{x \in \mathbb{T}_L} |x|_*^p P_{\mathbf{a}}(t,x)^2,$$

as well as the moderated quantities

$$\bar{\mathcal{E}}_t := \int_t^\infty K_{s-t}\mathcal{E}_s \, ds, \ \bar{\mathcal{D}}_t := \int_t^\infty K_{s-t}\mathcal{D}_s \, ds \text{ and } \bar{\mathcal{N}}_t := \int_t^\infty K_{s-t}\mathcal{N}_s \, ds.$$

We note that the following identities hold

$$\partial_t \mathcal{E}_t = -2\mathcal{D}_t,\ \partial_t \bar{\mathcal{E}}_t = -2\bar{\mathcal{D}}_t,\ \text{and}\ \partial_t \bar{\mathcal{N}}_t = \int_t^\infty K_{s-t} \partial_s \mathcal{N}_s \, ds.$$

In particular the maps $\mathcal{E}$ and $\bar{\mathcal{E}}$ are decreasing, since $\mathcal{E}_0 = 1$, we have $\mathcal{E}_t \leq 1$ for any time $t \geq 0$.

4.5.1. *A differential inequality for the weighted L^2-norm of the heat kernel.* Following the proof of Mourrat and Otto (2016), we will need to prove the following lemma which estimates the value of $\bar{\mathcal{N}}_0$ and the derivative $\partial_t \bar{\mathcal{N}}_t$. It closely follows Mourrat and Otto (2016, Proposition 3.3) (written with the function $\mathcal{N}$ instead of $\bar{\mathcal{N}}$), which is itself based on Gloria et al. (2015, (81)). We recall the notation for the maximal quantities $\mathcal{M}_0(t)$, $\mathcal{M}_1(t)$ and $\mathscr{M}_2$ introduced in Section 4.1.

Lemma 4.7. *There exists a constant $C := C(d) < \infty$, such that the following upper bounds hold*

$$\bar{\mathcal{N}}_0 \leq C \mathscr{M}_2 \ \textit{and}\ \partial_t \bar{\mathcal{N}}_t \leq C \mathcal{M}_1(t)^{\frac{2}{p}} (\bar{\mathcal{N}}_t)^{\frac{p-2}{p}} \mathcal{E}_t^{\frac{2}{p}}. \tag{4.23}$$

Proof: We first prove that the term $\mathcal{N}_t$ grows at most polynomially fast in the time t. Specifically, we will prove the upper bound, for any $t \geq 0$,

$$\mathcal{N}_t \leq C \mathscr{M}_2 (1+t)^{p-1}. \tag{4.24}$$

Using the definition of K_t in (4.2) (which implies that it decays asymptotically like $t \mapsto t^{-p-1}$), and integrating the previous inequality, we deduce that

$$\bar{\mathcal{N}}_0 = \int_0^\infty K_t \mathcal{N}_t \, dt \leq C \mathscr{M}_2 \int_0^\infty (1+t)^{-p-1} (1+t)^{p-1} \, dt \leq C \mathscr{M}_2.$$

To prove (4.24), we write, for $x \in \mathbb{T}_L$, $\rho(x) = |x|_*$. We first differentiate the function $\mathcal{N}_t$ and obtain

$$\frac{1}{2} \partial_t \mathcal{N}_t = - \sum_{e \in E(\mathbb{T}_L)} \nabla \left(\rho^p P_{\mathbf{a}} \right) (t,e) \mathbf{a}(t,e) \nabla P_{\mathbf{a}}(t,e).$$

Expanding the discrete gradient, we see that

$$\nabla (\rho^p P_{\mathbf{a}})(t,e) = \left(\nabla \rho^p(e) \right) P_{\mathbf{a}}(t,e) + \rho^p(e) \nabla P_{\mathbf{a}}(t,e).$$

using that there exists a constant $C_0 := C_0(d) < \infty$ (as the exponent p depends only on d) such that $|\nabla \rho^p(e)| \leq C_0 \rho^{p-1}(e)$, we deduce that

$$\frac{1}{2} \partial_t \mathcal{N}_t \leq - \sum_{e \in E(\mathbb{T}_L)} \rho^p(e) \mathbf{a}(t,e) (\nabla P_{\mathbf{a}}(t,e))^2 + C_0 \rho^{p-1}(e) P_{\mathbf{a}}(t,e) \mathbf{a}(t,e) \left| \nabla P_{\mathbf{a}}(t,e) \right|.$$

The second term in the right-hand side can be estimated using Young's inequality

$$\begin{aligned} &\sum_{e \in E(\mathbb{T}_L)} \rho^{p-1}(e) P_{\mathbf{a}}(t,e) \mathbf{a}(t,e) \left| \nabla P_{\mathbf{a}}(t,e) \right| \\ &\qquad \leq \frac{1}{2C_0} \sum_{e \in E(\mathbb{T}_L)} \rho^p(e) \mathbf{a}(t,e) (\nabla P_{\mathbf{a}}(t,e))^2 + \frac{C_0}{2} \sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(t,e) P_{\mathbf{a}}(t,e)^2. \end{aligned}$$

By the Hölder inequality and using that $\mathcal{E}_t \leq 1$, we have that

$$\begin{aligned} \sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(t,e) P_{\mathbf{a}}(t,e)^2 &\leq \left(\sum_{e \in E(\mathbb{T}_L)} \rho^{p-1}(e) \mathbf{a}(t,e)^{\frac{p-1}{p-2}} P_{\mathbf{a}}(t,e)^2 \right)^{\frac{p-2}{p-1}} \left(\sum_{e \in E(\mathbb{T}_L)} P_{\mathbf{a}}(t,e)^2 \right)^{\frac{1}{p-1}} \\ &\leq C \left(\sum_{e \in E(\mathbb{T}_L)} \rho^{p-1}(e) \mathbf{a}(t,e)^{\frac{p-1}{p-2}} P_{\mathbf{a}}(t,e)^2 \right)^{\frac{p-2}{p-1}}. \end{aligned}$$

Using the definition of the random variable $\mathcal{M}_2$, we obtain

$$\begin{aligned}\sum_{e\in E(\mathbb{T}_L)} \rho^{p-2}(e)\mathbf{a}(t,e)P_{\mathbf{a}}(t,e)^2 &\leq \left(\sum_{e\in E(\mathbb{T}_L)} \rho^{p-1}(e)\mathbf{a}(t,e)^{\frac{p-1}{p-2}}P_{\mathbf{a}}(t,e)^2\right)^{\frac{p-2}{p-1}} \\ &\leq C\mathcal{M}_2(t)\left(\sum_{e\in E(\mathbb{T}_L)} \rho^{p}(e)P_{\mathbf{a}}(t,e)^2\right)^{\frac{p-2}{p-1}} \\ &\leq C\mathcal{M}_2(t)\mathcal{N}_t^{\frac{p-2}{p-1}}.\end{aligned}$$

Combining the few previous displays, we obtain

$$\partial_t \mathcal{N}_t \leq C\mathcal{M}_2(t)\mathcal{N}_t^{\frac{p-2}{p-1}}.$$

Integrating the inequality and using $\mathcal{N}_0 = 1$, we obtain

$$\mathcal{N}_t \leq C\mathscr{M}_2(1+t)^{p-1}.$$

The proof of the first part of (4.23) is complete. We next prove the inequality on the derivative of the function $\bar{\mathcal{N}}_t$. To this end, we first compute (similarly as before)

$$\begin{aligned}\frac{1}{2}\partial_t\bar{\mathcal{N}}_t &= \frac{1}{2}\int_t^\infty K_{s-t}\partial_s\mathcal{N}_s\,ds \\ &= \frac{1}{2}\int_t^\infty K_{s-t}\partial_s\mathcal{N}_s\,ds \\ &= -\int_t^\infty K_{s-t}\sum_{e\in E(\mathbb{T}_L)} \nabla\left(\rho^p P_{\mathbf{a}}\right)(s,e)\mathbf{a}(s,e)\nabla P_{\mathbf{a}}(s,e)\,ds.\end{aligned}$$

Expanding the discrete divergence, we deduce that

$$\begin{aligned}\frac{1}{2}\partial_t\bar{\mathcal{N}}_t \leq -\int_t^\infty K_{s-t}\sum_{e\in E(\mathbb{T}_L)} \rho^p(e)\mathbf{a}(s,e)(\nabla P_{\mathbf{a}}(s,e))^2\,ds \\ + C_0\int_t^\infty K_{s-t}\sum_{e\in E(\mathbb{T}_L)} \rho^{p-1}(e)P_{\mathbf{a}}(s,e)\mathbf{a}(s,e)\left|\nabla P_{\mathbf{a}}(s,e)\right|\,ds.\end{aligned} \tag{4.25}$$

The second term in the right-hand side can be estimated using Young's inequality

$$\begin{aligned}&\sum_{e\in E(\mathbb{T}_L)} \rho^{p-1}(e)P_{\mathbf{a}}(s,e)\mathbf{a}(s,e)\left|\nabla P_{\mathbf{a}}(s,e)\right| \\ &\qquad\leq \frac{1}{2C_0}\sum_{e\in E(\mathbb{T}_L)} \rho^p(e)\mathbf{a}(s,e)(\nabla P_{\mathbf{a}}(s,e))^2 + \frac{C_0}{2}\sum_{e\in E(\mathbb{T}_L)} \rho^{p-2}(e)\mathbf{a}(s,e)P_{\mathbf{a}}(s,e)^2.\end{aligned} \tag{4.26}$$

We estimate the second term in the right-hand side. We will use the same technique as in the proof of Proposition 4.6. We first split the sum into dyadic scales. To this end, for $n\in\mathbb{Z}_+$, we denote by $A_n := \Lambda_{2^{n+1}}\setminus\Lambda_{2^n}$ the dyadic annulus and by $\ln_2$ the binary logarithm. We then write

$$\sum_{e\in E(\mathbb{T}_L)} \rho^{p-2}(e)\mathbf{a}(s,e)(P_{\mathbf{a}}(s,e))^2 \leq C\sum_{n=0}^{\lfloor \ln_2 L\rfloor} 2^{(p-2)n}\sum_{e\in E(A_n)} \mathbf{a}(s,e)P_{\mathbf{a}}(s,e)^2. \tag{4.27}$$

For each integer $n \in \{0, \ldots, \lfloor \ln_2 N \rfloor\}$, we apply the same computation as in (4.22) based on the Hölder inequality and the Gagliardo-Nirenberg-Sobolev inequality. We obtain, for any $\varepsilon > 0$,

$$\begin{aligned}
&\left(\frac{1}{|A_n|} \sum_{e \in E(A_n)} \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \right)^{\frac{1}{2}} \qquad (4.28)\\
&\quad \leq \|\mathbf{a}(s,\cdot)^{1/2}\|_{\underline{L}^{\sigma_d}(A_n)} \|P_{\mathbf{a}}(s,\cdot)\|_{\underline{L}^{\kappa_d}(A_n)} \\
&\quad \leq \|\mathbf{a}(s,\cdot)^{1/2}\|_{\underline{L}^{\sigma_d}(A_n)} \left(\varepsilon 2^n \|\nabla P_{\mathbf{a}}(s,\cdot)\|_{\underline{L}^{\lambda_d'}(A_n)} + C\varepsilon^{-\frac{\theta_d}{1-\theta_d}} \|P_{\mathbf{a}}(s,\cdot)\|_{\underline{L}^2(A_n)} \right) \\
&\quad \leq \varepsilon 2^n \|\mathbf{a}(s,\cdot)^{1/2}\|_{\underline{L}^{\sigma_d}(A_n)} \|w^{-1}(s,\cdot)\|_{\underline{L}^{\tau_d'}(A_n)} \|w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\|_{\underline{L}^2(A_n)} \\
&\qquad + C\varepsilon^{-\frac{\theta_d}{1-\theta_d}} \|\mathbf{a}(s,\cdot)^{1/2}\|_{\underline{L}^{\sigma_d}(A_n)} \|P_{\mathbf{a}}(s,\cdot)\|_{\underline{L}^2(A_n)} .
\end{aligned}$$

Using the definition of the maximal function $\mathcal{M}_0$, the inequality (4.28) can be rewritten as follows

$$\sum_{e \in E(A_n)} \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \leq \mathcal{M}_0(s) 2^{2n} \varepsilon^2 \|w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\|_{L^2(A_n)}^2 + C\varepsilon^{-\frac{2\theta_d}{1-\theta_d}} \mathcal{M}_0(s) \|P_{\mathbf{a}}(s,\cdot)\|_{L^2(A_n)}^2 .$$

Using Proposition 4.3, we deduce that

$$\begin{aligned}
\sum_{e \in E(A_n)} \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \leq C\mathcal{M}_0(s) 2^{2n} \varepsilon^2 \int_s^\infty K_{s'-s} \sum_{e \in E(A_n)} \mathbf{a}(s',e) (\nabla P_{\mathbf{a}}(s',e))^2 \, ds' \\
+ C\varepsilon^{-\frac{2\theta_d}{1-\theta_d}} \mathcal{M}_0(s) \sum_{x \in A_n} P_{\mathbf{a}}(s,x)^2.
\end{aligned}$$

Using that $c2^n \leq \rho(x) \leq C2^{n+1}$ for any $x \in A_n$ and summing over the integers $n \in \{0, \ldots, \lfloor \ln_2 L \rfloor\}$, we deduce that

$$\begin{aligned}
\sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \leq C\mathcal{M}_0(s) \varepsilon^2 \int_s^\infty K_{s'-s} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s',e) (\nabla P_{\mathbf{a}}(s',e))^2 \, ds \\
+ C\varepsilon^{-\frac{2\theta_d}{1-\theta_d}} \mathcal{M}_0(s) \sum_{x \in \mathbb{T}_L} \rho^{p-2}(x) P_{\mathbf{a}}(s,x)^2.
\end{aligned}$$

We next choose $\varepsilon = 1/(C_0\sqrt{C\mathcal{M}_0(s)})$ where C_0 is the constant appearing in (4.25) and C is the one appearing in the previous display. We obtain

$$\begin{aligned}
&\sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \\
&\quad \leq \frac{1}{C_0^2} \int_s^\infty K_{s'-s} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s',e) (\nabla P_{\mathbf{a}}(s',e))^2 \, ds + C\mathcal{M}_0(t)^{\frac{1}{1-\theta_d}} \sum_{x \in \mathbb{T}_L} \rho^{p-2}(x) (P_{\mathbf{a}}(s,x))^2.
\end{aligned}$$

Multiplying the previous inequality by the weight function K and integrating over the time variable yields, for any $t \geq 0$,

$$\int_t^\infty K_{s-t} \sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \, ds \\ \leq \frac{1}{C_0^2} \int_t^\infty K_{s-t} \int_s^\infty K_{s'-s} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s',e) (\nabla P_{\mathbf{a}}(s',e))^2 \, ds \, ds' \\ + C \int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{1}{1-\theta_d}} \sum_{x \in \mathbb{T}_L} \rho^{p-2}(x) P_{\mathbf{a}}(s,x)^2 \, ds. \tag{4.29}$$

The first term in the right-hand side can be simplified using the inequality (4.3) as follows

$$\int_t^\infty K_{s-t} \int_s^\infty K_{s'-s} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s',e') (\nabla P_{\mathbf{a}}(s',e))^2 \, ds \, ds' \\ \leq \int_t^\infty K_{s-t} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s,e) (\nabla P_{\mathbf{a}}(s,e))^2 \, ds.$$

The second term in the right-hand side of (4.29) can be estimated using the Hölder inequality as follows

$$\int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{1}{1-\theta_d}} \sum_{x \in \mathbb{T}_L} \rho^{p-2}(x) P_{\mathbf{a}}(s,x)^2 \, ds \\ \leq \left(\int_t^\infty K_{s-t} \sum_{x \in \mathbb{T}_L} \rho^p(x) P_{\mathbf{a}}(s,x)^2 \, ds \right)^{\frac{p-2}{p}} \left(\int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{p}{2(1-\theta_d)}} \sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(s,x)^2 \, ds \right)^{\frac{2}{p}}.$$

Using that the energy $\mathcal{E}_t = \sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t,x)^2$ is decreasing together with the definition of the random variable $\mathcal{M}_1(t)$, we deduce that

$$\int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{p}{2(1-\theta_d)}} \sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(s,x)^2 \, ds \leq \left(\int_t^\infty K_{s-t} \mathcal{M}_0(s)^{\frac{p}{2(1-\theta_d)}} \, ds \right) \mathcal{E}_t \leq \mathcal{M}_1(t) \mathcal{E}_t.$$

A combination of the few previous displays shows that

$$\int_t^\infty K_{s-t} \sum_{e \in E(\mathbb{T}_L)} \rho^{p-2}(e) \mathbf{a}(s,e) P_{\mathbf{a}}(s,e)^2 \, ds \\ \leq \frac{1}{C_0^2} \int_t^\infty K_{s-t} \sum_{e \in E(\mathbb{T}_L)} \rho(e)^p \mathbf{a}(s,e) (\nabla P_{\mathbf{a}}(s,e))^2 \, ds + C \mathcal{M}_1(t)^{\frac{2}{p}} (\bar{\mathcal{N}}_t)^{\frac{p-2}{p}} \mathcal{E}_t^{\frac{2}{p}}.$$

Combining the previous inequality with (4.25) and (4.26) completes the proof of Lemma 4.7. □

4.5.2. *An upper bound on the L^2-norm of the heat kernel.* We next deduce from Lemma 4.7 the energy upper bound for the $L^2(\mathbb{T}_L)$-norm of the heat kernel.

Proposition 4.8. *There exists a constant $C := C(d) < \infty$ such that, for any $t \geq 0$,*

$$\mathcal{E}_t \leq \frac{C \mathscr{M}}{(1+t)^{\frac{d}{2}}}.$$

Proof: By Lemma 4.7, we have the inequality

$$\partial_t \bar{\mathcal{N}}_t \leq C\mathcal{M}_1(t)^{\frac{2}{p}} (\bar{\mathcal{N}}_t)^{\frac{p-2}{p}} \mathcal{E}_t^{\frac{2}{p}}. \tag{4.30}$$

We define the quantity

$$\Lambda_t := \sup_{s\leq t} (1+s)^{\frac{d}{2}} \mathcal{E}_s,$$

and note that, for any $t \geq 0$, $\Lambda_t \geq 1$. We first observe that (4.30) can be rewritten using the definition Λ_t as follows: for every $t \geq 0$,

$$\partial_t \bar{\mathcal{N}}_t^{\frac{2}{p}} \leq C\mathcal{M}_1(t)^{\frac{2}{p}} \Lambda_t^{\frac{2}{p}} (1+t)^{-\frac{d}{p}}. \tag{4.31}$$

Integrating the previous inequality, recalling the definition of the maximal quantity $\mathscr{M}_1$ and using that Λ_t is increasing in t, we obtain, for any time $t \geq 2$,

$$\begin{aligned} \bar{\mathcal{N}}_t^{\frac{2}{p}} - \bar{\mathcal{N}}_{t/2}^{\frac{2}{p}} &= \int_{t/2}^t \partial_s \bar{\mathcal{N}}_s \, ds \\ &\leq C\Lambda_t^{\frac{2}{p}} \left(1+\frac{t}{2}\right)^{-\frac{d}{p}} \int_{t/2}^t \mathcal{M}_1(s)^{\frac{2}{p}} \, ds \\ &\leq C\Lambda_t^{\frac{2}{p}} (1+t)^{-\frac{d}{p}} \int_0^t \mathcal{M}_1(s)^{\frac{2}{p}} \, ds \\ &\leq C\mathscr{M}_1^{\frac{2}{p}} \Lambda_t^{\frac{2}{p}} (1+t)^{1-\frac{d}{p}}. \end{aligned} \tag{4.32}$$

Iterating the previous inequality (using that the map $t \mapsto \Lambda_t$ is increasing), treating the small values of t (between 0 and 1) using the inequality (4.31), and using the bound on $\bar{\mathcal{N}}_0$ provided by Lemma 4.7, we obtain that, for any $t \geq 0$,

$$\begin{aligned} \bar{\mathcal{N}}_t &\leq C\mathscr{M}_1 \Lambda_t (1+t)^{\frac{p-d}{2}} + C\mathscr{M}_2 \\ &\leq C\left(\mathscr{M}_1 + \mathscr{M}_2\right) \Lambda_t (1+t)^{\frac{p-d}{2}}. \end{aligned} \tag{4.33}$$

Applying the anchored Nash estimate, and using that the L^1-norm of the heat kernel $P_{\mathbf{a}}$ is bounded (see (2.5)), we obtain, for any $t \geq 0$,

$$\mathcal{E}_t \leq C\left(\mathcal{M}_{p'}(t) \left\|w(t,\cdot)\nabla P_{\mathbf{a}}(t,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2\right)^{\alpha} \mathcal{N}_t^{\gamma}.$$

Multiplying the previous inequality by the weight function K and integrating over time, we deduce that, for any $t \geq 0$,

$$\bar{\mathcal{E}}_t \leq \int_t^\infty K_{s-t} \left(\mathcal{M}_{p'}(s) \left\|w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2\right)^{\alpha} \mathcal{N}_s^{\gamma} \, ds. \tag{4.34}$$

Applying the Hölder inequality (recalling that $\alpha+\beta+\gamma = 1$), we deduce that, for any $t \geq 0$,

$$\begin{aligned} &\int_t^\infty K_{s-t} \left(\mathcal{M}_{p'}(t) \left\|w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2\right)^{\alpha} \mathcal{N}_s^{\gamma} \, ds \\ &\quad \leq \left(\int_t^\infty K_{s-t} \mathcal{M}_{p'}(s)^{\frac{\alpha}{\beta}} \, ds\right)^{\beta} \left(\int_t^\infty K_{s-t} \left\|w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds\right)^{\alpha} \left(\int_t^\infty K_{s-t} \mathcal{N}_s \, ds\right)^{\gamma}. \end{aligned} \tag{4.35}$$

The first term in the right-hand side is by definition smaller than $\mathcal{M}_3(t)$. We then estimate the second term in the right-hand side. To this end, we use Proposition 4.3 together with the bound (4.3).

We obtain, for any $t \geq 0$,

$$\begin{aligned}\int_t^\infty K_{s-t} \left\| w(s,\cdot) \nabla P_{\mathbf{a}}(s,\cdot) \right\|_{L^2(\mathbb{T}_L)}^2 \, ds &\leq C \int_t^\infty K_{s-t} \int_s^\infty K_{s'-s} \| \mathbf{a}(s',\cdot)^{1/2} \nabla P_{\mathbf{a}}(s',\cdot) \|_{L^2(\mathbb{T}_L)}^2 \, ds' ds \\ &\leq C \int_t^\infty K_{s-t} \| \mathbf{a}(s,\cdot)^{1/2} \nabla P_{\mathbf{a}}(s,\cdot) \|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq C \bar{\mathcal{D}}_t.\end{aligned}$$

Using the identity $\partial_t \bar{\mathcal{E}}_t = -2\bar{\mathcal{D}}_t$, we may rewrite the inequality (4.35) as follows

$$\int_t^\infty K_{s-t} \left(\mathcal{M}_{p'}(t) \left\| w(s,\cdot) \nabla P_{\mathbf{a}}(s,\cdot) \right\|_{L^2(\mathbb{T}_L)}^2 \right)^\alpha \mathcal{N}_s^\gamma \, ds \leq C \mathcal{M}_3(t) (-\partial_t \bar{\mathcal{E}}_t)^\alpha \bar{\mathcal{N}}_t^\gamma.$$

Combining the previous inequality with (4.34) and using the inequality (4.33), we deduce that

$$\bar{\mathcal{E}}_t \leq C \mathcal{M}_3(t) (-\partial_t \bar{\mathcal{E}}_t)^\alpha \bar{\mathcal{N}}_t^\gamma \leq C \mathcal{M}_3(t) (\mathscr{M}_1 + \mathscr{M}_2)^\gamma (-\partial_t \bar{\mathcal{E}}_t)^\alpha \Lambda_t^\gamma (1+t)^{\frac{(p-d)\gamma}{2}},$$

which can be rewritten as

$$\left(-\partial_t \bar{\mathcal{E}}_t\right) \bar{\mathcal{E}}_t^{-\frac{1}{\alpha}} \geq c \mathcal{M}_3(t)^{-\frac{1}{\alpha}} (\mathscr{M}_1 + \mathscr{M}_2)^{-\frac{\gamma}{\alpha}} \Lambda_t^{-\frac{\gamma}{\alpha}} (1+t)^{-\frac{(p-d)\gamma}{2\alpha}}. \tag{4.36}$$

Using that $t \mapsto \Lambda_t$ is increasing and that $p > d$, we deduce that

$$\bar{\mathcal{E}}_t^{1-\frac{1}{\alpha}} \geq c \Lambda_t^{-\frac{\gamma}{\alpha}} (1+t)^{-\frac{(p-d)\gamma}{2\alpha}} (\mathscr{M}_1 + \mathscr{M}_2)^{-\frac{\gamma}{\alpha}} \int_0^t \mathcal{M}_3^{-\frac{1}{\alpha}}(s) \, ds.$$

Using the identity (2.13) and the definitions of the random variables $\mathscr{M}_3$ and $\mathscr{M}$, we deduce that, for any $t \geq 1$,

$$(1+t)^{\frac{d}{2}} \bar{\mathcal{E}}_t \leq C \Lambda_t^{\frac{\gamma}{1-\alpha}} \mathscr{M}^{1-\frac{\gamma}{1-\alpha}}.$$

Using that $\mathcal{E}_t$ is decreasing in t and the definition of $\bar{\mathcal{E}}_t$, we have the inequality

$$\left(\int_0^1 K_s \, ds \right) \mathcal{E}_{t+1} \leq \bar{\mathcal{E}}_t.$$

Using that $t \mapsto \Lambda_t$ is increasing, we obtain, for any $t \geq 1$,

$$(1+t)^{\frac{d}{2}} \mathcal{E}_{t+1} \leq C \mathscr{M}^{1-\frac{\gamma}{1-\alpha}} \Lambda_{t+1}^{\frac{\gamma}{1-\alpha}}.$$

Combining the previous inequality with the bound $\mathcal{E}_t \leq 1$, the observation that Λ_t is increasing and larger than 1, we deduce that, for any $t \geq 0$,

$$\Lambda_t \leq C \mathscr{M}^{1-\frac{\gamma}{1-\alpha}} \Lambda_t^{\frac{\gamma}{1-\alpha}} \implies \Lambda_t \leq C \mathscr{M}.$$

The proof of Proposition 4.8 is complete. □

4.5.3. *A refined upper bound on L^2-norm on the heat kernel.* This section is devoted to the proof of Theorem 4.2 building upon Proposition 4.8.

Proof of Theorem 4.2: We let $C_2 \geq 1$ be a large constant whose value will be selected later in the argument and shall depend only on d, and define

$$C_1 := C_2^{\frac{2\alpha}{d\beta + p\gamma}} \mathscr{M}_3^{\frac{2\gamma}{(d\beta + p\gamma)}} \text{ and } C_0 := 2 \mathscr{M}_3^{\frac{\gamma}{\alpha}} C_1. \tag{4.37}$$

We then define the three quantities

$$\mathcal{H}_t := e^{\frac{1}{C_0 L^2} \int_0^t \mathcal{M}_4(s)^{-1} \, ds} \mathcal{E}_t, \ \bar{\mathcal{H}}_t := e^{\frac{1}{C_0 L^2} \int_0^t \mathcal{M}_4(s)^{-1} \, ds} \bar{\mathcal{E}}_t \text{ and } \Xi_t := \sup_{s \leq t} (1+s)^{\frac{d}{2}} \mathcal{H}_s.$$

We next prove the following upper bound

$$\bar{\mathcal{E}}_t \leq C \mathcal{M}_4(t) L^2 (-\partial_t \bar{\mathcal{E}}_t). \tag{4.38}$$

To prove the inequality (4.38), we first use that the function $\mathcal{E}_t$ is decreasing and write

$$\begin{aligned}\bar{\mathcal{E}}_t &= \int_t^\infty K_{s-t} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &= \int_t^{t+1} K_{s-t} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds + \int_{t+1}^\infty K_{s-t} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq \int_t^{t+1} K_{s-t} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds + \left\| P_{\mathbf{a}}(t+1,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \int_1^\infty K_s \, ds.\end{aligned}$$

Using a second time that $\mathcal{E}_t$ is decreasing, that the ratio $\int_1^\infty K_s\, ds / \int_0^1 K_s\, ds$ is a finite constant depending only on the parameter d, and the definition of the random variable $\mathcal{M}_4(t)$, we deduce that

$$\begin{aligned}\bar{\mathcal{E}}_t &\leq \left(1 + \frac{\int_1^\infty K_s\, ds}{\int_0^1 K_s\, ds}\right) \int_t^{t+1} K_{s-t} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq C\mathcal{M}_4(t) \int_t^{t+1} K_{s-t} \| w^{-1}(s,\cdot)\|_{\underline{L}^d(\mathbb{T}_L)}^{-2} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq C\mathcal{M}_4(t) \int_t^{\infty} K_{s-t} \| w^{-1}(s,\cdot)\|_{\underline{L}^d(\mathbb{T}_L)}^{-2} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds.\end{aligned}$$

We next use the Gagliardo-Nirenberg-Sobolev inequality (using that $\sum_{x\in\mathbb{T}_L} P_{\mathbf{a}}(s,x) = 0$), then the Hölder inequality. We obtain

$$\begin{aligned}\bar{\mathcal{E}}_t &\leq C\mathcal{M}_4(t) \int_t^{\infty} K_{s-t} \| w^{-1}(s,\cdot)\|_{\underline{L}^d(\mathbb{T}_L)}^{-2} \left\| P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq C\mathcal{M}_4(t) \int_t^{\infty} K_{s-t} \| w^{-1}(s,\cdot)\|_{\underline{L}^d(\mathbb{T}_L)}^{-2} \left\| \nabla P_{\mathbf{a}}(s,\cdot)\right\|_{L^{\frac{2d}{d+2}}(\mathbb{T}_L)}^2 \, ds \\ &\leq C\mathcal{M}_4(t)L^2 \int_t^{\infty} K_{s-t} \left\| w(s,\cdot)\nabla P_{\mathbf{a}}(s,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2 \, ds.\end{aligned}$$

Using Proposition 4.3 together with the bound (4.3), we deduce that

$$\begin{aligned}\bar{\mathcal{E}}_t &\leq C\mathcal{M}_4(t)L^2 \int_t^{\infty} K_{s-t} \| \mathbf{a}(s,\cdot)^{1/2}\nabla P_{\mathbf{a}}(s,\cdot)\|_{L^2(\mathbb{T}_L)}^2 \, ds \\ &\leq C\mathcal{M}_4(t)L^2(-\partial_t\bar{\mathcal{E}}_t).\end{aligned} \tag{4.39}$$

The proof of the inequality (4.38) is complete. We impose here a first condition on the constant C_2 and choose it sufficiently large so that the constant C_2 is larger than $2C$, where C is the constant appearing in the right-hand side of (4.38). We thus deduce that

$$\begin{aligned}\left(-\partial_t\bar{\mathcal{H}}_t\right)\bar{\mathcal{H}}_t^{-\frac{1}{\alpha}} &= e^{\left(1-\frac{1}{\alpha}\right)\frac{1}{C_0L^2}\int_0^t \mathcal{M}_4(s)^{-1}\, ds}\left(-\frac{1}{C_0\mathcal{M}_4(t)L^2}\bar{\mathcal{E}}_t - \partial_t\bar{\mathcal{E}}_t\right)\bar{\mathcal{E}}_t^{-\frac{1}{\alpha}} \\ &\geq \frac{1}{2}e^{\left(1-\frac{1}{\alpha}\right)\frac{1}{C_0L^2}\int_0^t \mathcal{M}_4(s)^{-1}\, ds}\left(-\partial_t\bar{\mathcal{E}}_t\right)\bar{\mathcal{E}}_t^{-\frac{1}{\alpha}}.\end{aligned} \tag{4.40}$$

Applying the anchored Nash inequality, we have

$$\mathcal{E}_t \leq C \left\| P_{\mathbf{a}}(t,\cdot)\right\|_{L^1(\mathbb{T}_L)}^{2\beta} \left(\mathcal{M}_{p'}(t)\left\| w(t,\cdot)\nabla P_{\mathbf{a}}(t,\cdot)\right\|_{L^2(\mathbb{T}_L)}^2\right)^\alpha \mathcal{N}_t^\gamma.$$

We then estimate the first and third terms in the right hand side by using the Cauchy-Schwarz inequality and the bound $\rho(x) \leq CL$, which is valid on the torus since its diameter is of order L. We obtain

$$\left\| P_{\mathbf{a}}(t,\cdot)\right\|_{L^1(\mathbb{T}_L)}^2 \leq CL^d\mathcal{E}_t \text{ and } \mathcal{N}_t \leq CL^p\mathcal{E}_t.$$

Combining the two previous displays, we may write

$$\mathcal{E}_t \leq C \left(L^d \mathcal{E}_t \right)^{\beta} \left(\mathcal{M}_{p'}(t) \left\| w(t,\cdot) \nabla u(t,\cdot) \right\|^2_{L^2(\mathbb{T}_L)} \right)^{\alpha} (L^p \mathcal{E}_t)^{\gamma}.$$

Performing the same computation as in (4.35) and using that the function $\mathcal{E}_t$ is decreasing, we deduce that

$$\bar{\mathcal{E}}_t \leq C \mathcal{M}_3(t) L^{d\beta + p\gamma} \mathcal{E}_t^{\beta+\gamma} (-\partial_t \bar{\mathcal{E}}_t)^{\alpha}.$$

Using the definition of Ξ_t, we further deduce that

$$\bar{\mathcal{E}}_t \leq C \mathcal{M}_3(t) L^{d\beta + p\gamma} (1+t)^{-\frac{(\beta+\gamma)d}{2}} e^{-\frac{\beta+\gamma}{C_0 L^2} \int_0^t \mathcal{M}_4(s)^{-1}\, ds} \Xi_t^{\beta+\gamma} (-\partial_t \bar{\mathcal{E}}_t)^{\alpha}.$$

Rearranging the previous inequality, we obtain

$$(-\partial_t \bar{\mathcal{E}}_t) \bar{\mathcal{E}}_t^{-\frac{1}{\alpha}} \geq c \mathcal{M}_3(t)^{-\frac{1}{\alpha}} L^{-\frac{d\beta+p\gamma}{\alpha}} (1+t)^{\frac{(\beta+\gamma)d}{2\alpha}} e^{\frac{\beta+\gamma}{\alpha C_0 L^2} \int_0^t \mathcal{M}_4(s)^{-1}\, ds} \Xi_t^{-\frac{\beta+\gamma}{\alpha}}.$$

Combining the previous inequality with (4.40) and noting that $1 - \frac{1}{\alpha} = -\frac{\beta+\gamma}{\alpha}$ (since $\alpha+\beta+\gamma = 1$), we can rewrite the previous inequality as follows

$$\begin{aligned} \partial_t \bar{\mathcal{H}}_t^{1-\frac{1}{\alpha}} &\geq c \mathcal{M}_3(t)^{-\frac{1}{\alpha}} L^{-\frac{d\beta+p\gamma}{\alpha}} (1+t)^{\frac{(\beta+\gamma)d}{2\alpha}} \Xi_t^{-\frac{\beta+\gamma}{\alpha}} \\ &\geq c \mathcal{M}_3(t)^{-\frac{1}{\alpha}} \left(\frac{1+t}{L^2} \right)^{\frac{d\beta+p\gamma}{2\alpha}} (1+t)^{-\frac{\gamma(p-d)}{2\alpha}} \Xi_t^{-\frac{\beta+\gamma}{\alpha}}. \end{aligned}$$

Using the definition of the constant C_1, we have, for any $t \geq C_1 L^2$,

$$\partial_t \bar{\mathcal{H}}_t^{1-\frac{1}{\alpha}} \geq c C_2 \mathscr{M}_3^{\frac{\gamma}{\alpha}} \mathcal{M}_3(t)^{-\frac{1}{\alpha}} (1+t)^{-\frac{\gamma(p-d)}{2\alpha}} \Xi_t^{-\frac{\beta+\gamma}{\alpha}}.$$

Integrating the previous inequality and using that Ξ_t is increasing in t, we deduce that, for any $t \geq C_1 L^2$,

$$\bar{\mathcal{H}}_t^{1-\frac{1}{\alpha}} - \bar{\mathcal{H}}_{C_1 L^2}^{1-\frac{1}{\alpha}} \geq c C_2 \Xi_t^{-\frac{\beta+\gamma}{\alpha}} (1+t)^{-\frac{\gamma(p-d)}{2\alpha}} \mathscr{M}_3^{\frac{\gamma}{\alpha}} \int_{C_1 L^2}^t \mathcal{M}_3(s)^{-\frac{1}{\alpha}}\, ds.$$

We next recall the definition of the constant C_0 introduced in (4.37), and lower bound the term in the right-hand side for $t \geq C_0 L^2$. To this end, we use the definition of the random variable $\mathscr{M}_3$ and the lower bound $\mathcal{M}_3 \geq 1$, and obtain, for any $t \geq C_0 L^2$,

$$\begin{aligned} \frac{\mathscr{M}_3^{\frac{\gamma}{\alpha}}}{t} \int_{C_1 L^2}^t \mathcal{M}_3(s)^{-\frac{1}{\alpha}}\, ds &= \frac{\mathscr{M}_3^{\frac{\gamma}{\alpha}}}{t} \int_0^t \mathcal{M}_3(s)^{-\frac{1}{\alpha}}\, ds - \frac{\mathscr{M}_3^{\frac{\gamma}{\alpha}}}{t} \int_0^{C_1 L^2} \mathcal{M}_3(s)^{-\frac{1}{\alpha}}\, ds \\ &\geq 1 - \frac{C_1 L^2 \mathscr{M}_3^{\frac{\gamma}{\alpha}}}{t} \\ &\geq \frac{1}{2}. \end{aligned}$$

A combination of the two previous displays yields, for any $t \geq C_0 L^2$,

$$\bar{\mathcal{H}}_t^{1-\frac{1}{\alpha}} \geq \bar{\mathcal{H}}_t^{1-\frac{1}{\alpha}} - \bar{\mathcal{H}}_{C_1 L^2}^{1-\frac{1}{\alpha}} \geq c C_2 \Xi_t^{-\frac{\beta+\gamma}{\alpha}} (1+t)^{1-\frac{\gamma(p-d)}{2\alpha}} = c C_2 \Xi_t^{1-\frac{1}{\alpha}} (1+t)^{1-\frac{\gamma(p-d)}{2\alpha}}.$$

We finally remove the averaging from the previous inequality. Using that the map $\mathcal{E}_t$ is decreasing, we have the estimate

$$\left(\int_0^1 K_s\, ds \right) \mathcal{E}_{t+1} \leq \bar{\mathcal{E}}_t,$$

and combining the previous inequality with the bound $\mathcal{M}_4^{-1}(t) \leq 1$ (which follows from the definition of $\mathcal{M}_4$), we deduce that

$$\left(\int_0^1 K_s\, ds \right) e^{-1/(C_0 L^2)} \mathcal{H}_{t+1} \leq \bar{\mathcal{H}}_t.$$

Combining the few previous displays and using that $t \mapsto \Xi_t$ is increasing, we deduce that, for any $t \geq C_0 L^2 + 1$,

$$\mathcal{H}_t^{1-\frac{1}{\alpha}} \geq c C_2 \Xi_t^{1-\frac{1}{\alpha}} (1+t)^{1-\frac{\gamma(p-d)}{2\alpha}}.$$

We next impose a second condition on the constant C_2 and assume that $cC_2 \geq 2^{\frac{1}{\alpha}-1}$. This leads to the bound, for any $t \geq C_0 L^2 + 1$,

$$(1+t)^{\frac{d}{2}} \mathcal{H}_t \leq \frac{1}{2} \Xi_t.$$

We next apply Proposition 4.8 and the bound $\mathcal{M}_4(t)^{-1} \leq 1$ and obtain

$$\begin{aligned} \sup_{t \in [0, C_0 L^2 + 1]} (1+t)^{\frac{d}{2}} \mathcal{H}_t &= \sup_{t \in [0, C_0 L^2 + 1]} e^{\frac{1}{C_0 L^2} \int_0^t \mathcal{M}_4(s)^{-1} \, ds} (1+t)^{\frac{d}{2}} \mathcal{E}_t \\ &\leq e^{\frac{C_0 L^2 + 1}{C_0 L^2}} \sup_{t \in [0, C_0 L^2 + 1]} (1+t)^{\frac{d}{2}} \mathcal{E}_t \\ &\leq C \mathscr{M}. \end{aligned}$$

Combining the two previous displays, we deduce that, for any $t \geq 0$,

$$\Xi_t \leq \frac{1}{2} \Xi_t + C \mathscr{M},$$

and thus, for any $t \geq 0$,

$$\Xi_t \leq C \mathscr{M}.$$

The proof of Theorem 4.2 is complete. □

4.6. *On diagonal estimate for the heat kernel.* In this section, we deduce from Theorem 4.2 the on-diagonal upper bound on the heat-kernel $P_{\mathbf{a}}$. In order to state the result, we fix a time $t \in [0, \infty)$ and define the reversed environment

$$\mathbf{a}^{(t)}(t', e) := \mathbf{a}(t - t', e).$$

The environment $\mathbf{a}^{(t)}$ is only defined for the times $t' \in [0, t]$. This is the only relevant property for the statement below; but we note that we may extend its definition to all times so as to make $\mathbf{a}^{(t)}$ a stationary process by for instance extending the definition of the Langevin dynamic to negative times.

Since the Langevin dynamic is stationary and reversible with respect to the Gibbs measure $\mu_{\mathbb{T}_L}$, the processes $\mathbf{a}$ and $\mathbf{a}^{(t)}$ have the same law. Let us denote by $\mathscr{M}^{(t)}$ the random variable $\mathscr{M}$ associated with the environment $\mathbf{a}^{(t)}$. Since the processes $\mathbf{a}$ and $\mathbf{a}^{(t)}$ have the same law, the random variables $\mathscr{M}$ and $\mathscr{M}^{(t)}$ also have the same law.

Proposition 4.9 (On-diagonal heat kernel decay)**.** *There exists a constant $C := C(d) < \infty$ such that, for any time $t \geq 0$,*

$$P_{\mathbf{a}}(t, 0) \leq \frac{C \sqrt{\mathscr{M} \mathscr{M}^{(t)}}}{(1+t)^{\frac{d}{2}}} \exp\left(- \frac{t}{C \mathscr{M}' L^2} \right).$$

Proof: Using the convolution property of the heat kernel and the Cauchy-Schwarz inequality, we obtain

$$P_{\mathbf{a}}(t, 0) = \sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t, 0; t/2, x) P_{\mathbf{a}}(t/2, x) \leq \left(\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t, 0; t/2, x)^2 \right)^{\frac{1}{2}} \left(\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t/2, x)^2 \right)^{\frac{1}{2}}. \quad (4.41)$$

To estimate the first term in the right-hand side, we use the following identity between the heat kernel and the heat kernel under the reversed environment (see Mourrat and Otto (2016, Lemma 4.5))

$$P_{\mathbf{a}}(t,0;t/2,x) = P_{\mathbf{a}^{(t)}}\left(t/2,x\right).$$

Applying Proposition 4.8 with the environment $\mathbf{a}^{(t)}$ (and thus the random variable $\mathscr{M}^{(t)}$), we deduce that

$$\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t,0;t/2,x)^2 \leq \frac{C\mathscr{M}^{(t)}}{(1+t)^{\frac{d}{2}}}.$$

The second term in the right-hand side of (4.41) can be estimated using Theorem 4.2. We obtain

$$\sum_{x \in \mathbb{T}_L} P_{\mathbf{a}}(t/2,x)^2 \leq \frac{C\mathscr{M}}{(1+t)^{\frac{d}{2}}} \exp\left(-\frac{t}{C\mathscr{M}' L^2}\right).$$

Combining the two previous displays with (4.41) completes the proof of Proposition 4.9.

□

4.7. *Helffer-Sjöstrand representation formula and proof of Theorem 1.2.* We are then able to complete the proof of the localization and delocalization estimate for the random surface stated in Theorem 1.2 by combining Proposition 4.9 and the Helffer-Sjöstrand representation formula.

Proof of Theorem 1.2: By the Helffer-Sjöstrand representation formula, we have the identity

$$\mathrm{Var}_{\mathbb{T}_L}\left[\phi(0)\right] = \mathbb{E}\left[\int_0^\infty P_{\mathbf{a}}\left(t,0\right)\, dt\right].$$

Applying Proposition 4.9 and using the inequality $\exp(-t) \leq 1/t$ for $t > 0$, we see that

$$\begin{aligned}\int_0^\infty P_{\mathbf{a}}\left(t,0\right)\, dt &\leq \int_0^\infty \frac{C\sqrt{\mathscr{M}\mathscr{M}^{(\frac{t}{2})}}}{(1+t)^{\frac{d}{2}}} \exp\left(-\frac{t}{C\mathscr{M}' L^2}\right) dt \\ &\leq \int_0^{L^2} \frac{C\sqrt{\mathscr{M}\mathscr{M}^{(\frac{t}{2})}}}{(1+t)^{\frac{d}{2}}} dt + \int_{L^2}^\infty \frac{C\sqrt{\mathscr{M}\mathscr{M}^{(\frac{t}{2})}}}{(1+t)^{\frac{d}{2}}} \frac{\mathscr{M}' L^2}{t}\, dt.\end{aligned}$$

Taking the expectation in the previous inequality, and using that all the moments of the random variables $\mathscr{M}$, $\mathscr{M}^{(t)}$ and $\mathscr{M}'$ are finite (in particular the random variables $\sqrt{\mathscr{M}\mathscr{M}^{(\frac{t}{2})}}$ and $\sqrt{\mathscr{M}\mathscr{M}^{(\frac{t}{2})}}\mathscr{M}'$ have a finite expectation whose value can be bounded uniformly in t) completes the proof of Theorem 1.2. □

References

Adams, S., Buchholz, S., Kotecký, R., and Müller, S. Cauchy-Born rule from microscopic models with non-convex potentials. *ArXiv Mathematics e-prints* (2019). arXiv: 1910.13564.

Adams, S. and Koller, A. The Hessian of surface tension characterises scaling limit of gradient models with non-convex energy. *ArXiv Mathematics e-prints* (2023). arXiv: 2306.12226.

Adams, S., Kotecký, R., and Müller, S. Strict convexity of the surface tension for non-convex potentials. *ArXiv Mathematics e-prints* (2016). arXiv: 1606.09541.

Aizenman, M., Harel, M., Peled, R., and Shapiro, J. Depinning in integer-restricted Gaussian Fields and BKT phases of two-component spin models. *ArXiv Mathematics e-prints* (2021). arXiv: 2110.09498.

Akcoglu, M. A. and Krengel, U. Ergodic theorems for superadditive processes. *J. Reine Angew. Math.*, **323**, 53–67 (1981). MR611442.

Andres, S., Barlow, M. T., Deuschel, J.-D., and Hambly, B. M. Invariance principle for the random conductance model. *Probab. Theory Related Fields*, **156** (3-4), 535–580 (2013). MR3078279.

Andres, S., Chiarini, A., Deuschel, J.-D., and Slowik, M. Quenched invariance principle for random walks with time-dependent ergodic degenerate weights. *Ann. Probab.*, **46** (1), 302–336 (2018). MR3758732.

Andres, S., Chiarini, A., and Slowik, M. Quenched local limit theorem for random walks among time-dependent ergodic degenerate weights. *Probab. Theory Related Fields*, **179** (3-4), 1145–1181 (2021). MR4242632.

Andres, S., Deuschel, J.-D., and Slowik, M. Heat kernel estimates for random walks with degenerate weights. *Electron. J. Probab.*, **21**, Paper No. 33, 21 (2016). MR3492937.

Andres, S., Deuschel, J.-D., and Slowik, M. Heat kernel estimates and intrinsic metric for random walks with general speed measure under degenerate conductances. *Electron. Commun. Probab.*, **24**, Paper No. 5, 17 (2019). MR3916337.

Andres, S., Deuschel, J.-D., and Slowik, M. Green kernel asymptotics for two-dimensional random walks under random conductances. *Electron. Commun. Probab.*, **25**, Paper No. 58, 14 (2020). MR4137943.

Andres, S. and Halberstam, N. Lower Gaussian heat kernel bounds for the random conductance model in a degenerate ergodic environment. *Stochastic Process. Appl.*, **139**, 212–228 (2021). MR4271911.

Andres, S. and Taylor, P. A. Local limit theorems for the random conductance model and applications to the Ginzburg-Landau $\nabla\phi$ interface model. *J. Stat. Phys.*, **182** (2), Paper No. 35, 35 (2021). MR4212193.

Armstrong, S. and Dario, P. Quantitative hydrodynamic limits of the Langevin dynamics for gradient interface models. *Electron. J. Probab.*, **29**, Paper no. 1 (2024). MR4689353.

Armstrong, S. and Wu, W. C^2 regularity of the surface tension for the $\nabla\phi$ interface model. *Comm. Pure Appl. Math.*, **75** (2), 349–421 (2022). MR4373172.

Avena, L. Symmetric exclusion as a model of non-elliptic dynamical random conductances. *Electron. Commun. Probab.*, **17**, no. 44, 8 (2012). MR2988390.

Bandyopadhyay, A. and Zeitouni, O. Random walk in dynamic Markovian random environment. *ALEA Lat. Am. J. Probab. Math. Stat.*, **1**, 205–224 (2006). MR2249655.

Barlow, M. T. Random walks on supercritical percolation clusters. *Ann. Probab.*, **32** (4), 3024–3084 (2004). MR2094438.

Barlow, M. T. and Hambly, B. M. Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab.*, **14**, no. 1, 1–27 (2009). MR2471657.

Bauerschmidt, R., Park, J., and Rodriguez, P.-F. The Discrete Gaussian model, I. Renormalisation group flow at high temperature (2024+a). To appear in *Ann. Probab.*. arXiv: 2202.02286.

Bauerschmidt, R., Park, J., and Rodriguez, P.-F. The Discrete Gaussian model, II. Infinite-volume scaling limit at high temperature (2024+b). To appear in *Ann. Probab.*. arXiv: 2202.02287.

Belius, D. and Wu, W. Maximum of the Ginzburg-Landau fields. *Ann. Probab.*, **48** (6), 2647–2679 (2020). MR4164451.

Bella, P. and Schäffner, M. Quenched invariance principle for random walks among random degenerate conductances. *Ann. Probab.*, **48** (1), 296–316 (2020). MR4079437.

Berger, N. and Biskup, M. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields*, **137** (1-2), 83–120 (2007). MR2278453.

Berger, N., Biskup, M., Hoffman, C. E., and Kozma, G. Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.*, **44** (2), 374–392 (2008). MR2446329.

Biskup, M. Recent progress on the random conductance model. *Probab. Surv.*, **8**, 294–373 (2011). MR2861133.

Biskup, M. and Boukhadra, O. Subdiffusive heat-kernel decay in four-dimensional i.i.d. random conductance models. *J. Lond. Math. Soc. (2)*, **86** (2), 455–481 (2012). MR2980920.

Biskup, M. and Kotecký, R. Phase coexistence of gradient Gibbs states. *Probab. Theory Related Fields*, **139** (1-2), 1–39 (2007). MR2322690.

Biskup, M. and Pan, M. An invariance principle for one-dimensional random walks in degenerate dynamical random environments. *Electron. J. Probab.*, **28**, Paper No. 153, 18 (2023). MR4669755.

Biskup, M. and Prescott, T. M. Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.*, **12**, no. 49, 1323–1348 (2007). MR2354160.

Biskup, M. and Rodriguez, P.-F. Limit theory for random walks in degenerate time-dependent random environments. *J. Funct. Anal.*, **274** (4), 985–1046 (2018). MR3743189.

Biskup, M. and Spohn, H. Scaling limit for a class of gradient fields with nonconvex potentials. *Ann. Probab.*, **39** (1), 224–251 (2011). MR2778801.

Boldrighini, C., Minlos, R. A., and Pellegrinotti, A. Almost-sure central limit theorem for a Markov model of random walk in dynamical random environment. *Probab. Theory Related Fields*, **109** (2), 245–273 (1997). MR1477651.

Boldrigini, K., Minlos, R. A., and Pellegrinotti, A. Random walks in a random (fluctuating) environment. *Russ. Math. Surv.*, **62** (4), 663 (2007). MR2358736.

Boukhadra, O. Heat-kernel estimates for random walk among random conductances with heavy tail. *Stochastic Process. Appl.*, **120** (2), 182–194 (2010). MR2576886.

Brascamp, H. J., Lieb, E. H., and Lebowitz, J. L. The statistical mechanics of anharmonic lattices. *Bull. Inst. Internat. Statist.*, **46** (1), 393–404 (1976) (1975). MR676341.

Bricmont, J., Fontaine, J.-R., and Lebowitz, J. L. Surface tension, percolation, and roughening. *J. Statist. Phys.*, **29** (2), 193–203 (1982). MR684948.

Brydges, D. and Spencer, T. Fluctuation estimates for sub-quadratic gradient field actions. *J. Math. Phys.*, **53** (9), 095216, 5 (2012). MR2905798.

Brydges, D. and Yau, H.-T. Grad ϕ perturbations of massless Gaussian fields. *Comm. Math. Phys.*, **129** (2), 351–392 (1990). MR1048698.

Buchholz, S. Phase transitions for a class of gradient fields. *Probab. Theory Related Fields*, **179** (3-4), 969–1022 (2021). MR4242629.

Buckley, S. Anomalous heat kernel behaviour for the dynamic random conductance model. *Electron. Commun. Probab.*, **18**, no. 1, 11 (2013). MR3011528.

Cotar, C. and Deuschel, J.-D. Decay of covariances, uniqueness of ergodic component and scaling limit for a class of $\nabla\phi$ systems with non-convex potential. *Ann. Inst. Henri Poincaré Probab. Stat.*, **48** (3), 819–853 (2012). MR2976565.

Cotar, C., Deuschel, J.-D., and Müller, S. Strict convexity of the free energy for a class of non-convex gradient models. *Comm. Math. Phys.*, **286** (1), 359–376 (2009). MR2470934.

Delmotte, T. and Deuschel, J.-D. On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab. Theory Related Fields*, **133** (3), 358–390 (2005). MR2198017.

Deuschel, J.-D. and Giacomin, G. Entropic repulsion for massless fields. *Stochastic Process. Appl.*, **89** (2), 333–354 (2000). MR1780295.

Deuschel, J.-D., Giacomin, G., and Ioffe, D. Large deviations and concentration properties for $\nabla\phi$ interface models. *Probab. Theory Related Fields*, **117** (1), 49–111 (2000). MR1759509.

Deuschel, J.-D., Nishikawa, T., and Vignaud, Y. Hydrodynamic limit for the Ginzburg-Landau $\nabla\phi$ interface model with non-convex potential. *Stochastic Process. Appl.*, **129** (3), 924–953 (2019). MR3913274.

Deuschel, J.-D. and Rodriguez, P.-F. An isomorphism theorem for Ginzburg-Landau interface models and scaling limits. *ArXiv Mathematics e-prints* (2022). arXiv: 2206.14805.

Dobrushin, R. L. and Shlosman, S. B. Nonexistence of one- and two-dimensional Gibbs fields with noncompact group of continuous symmetries. In *Multicomponent random systems*, volume 6 of

Adv. Probab. Related Topics, pp. 199–210. Dekker, New York (1980). MR599536.

Dolgopyat, D., Keller, G., and Liverani, C. Random walk in Markovian environment. *Ann. Probab.*, **36** (5), 1676–1710 (2008). MR2440920.

Efron, B. Increasing properties of Pólya frequency functions. *Ann. Math. Statist.*, **36**, 272–279 (1965). MR171335.

Fröhlich, J. and Pfister, C. On the absence of spontaneous symmetry breaking and of crystalline ordering in two-dimensional systems. *Comm. Math. Phys.*, **81** (2), 277–298 (1981). MR632763.

Fröhlich, J. and Spencer, T. The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Comm. Math. Phys.*, **81** (4), 527–602 (1981). MR634447.

Funaki, T. Stochastic interface models. In *Lectures on probability theory and statistics*, volume 1869 of *Lecture Notes in Math.*, pp. 103–274. Springer, Berlin (2005). MR2228384.

Funaki, T. and Nishikawa, T. Large deviations for the Ginzburg-Landau $\nabla\phi$ interface model. *Probab. Theory Related Fields*, **120** (4), 535–568 (2001). MR1853482.

Funaki, T. and Spohn, H. Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model. *Comm. Math. Phys.*, **185** (1), 1–36 (1997). MR1463032.

Giacomin, G., Olla, S., and Spohn, H. Equilibrium fluctuations for $\nabla\phi$ interface model. *Ann. Probab.*, **29** (3), 1138–1172 (2001). MR1872740.

Gloria, A., Neukamm, S., and Otto, F. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, **199** (2), 455–515 (2015). MR3302119.

Helffer, B. and Sjöstrand, J. On the correlation for Kac-like models in the convex case. *J. Statist. Phys.*, **74** (1-2), 349–409 (1994). MR1257821.

Hilger, S. Scaling limit and convergence of smoothed covariance for gradient models with non-convex potential. *ArXiv Mathematics e-prints* (2016). arXiv: 1603.04703.

Hilger, S. Decay of covariance for gradient models with non-convex potential. *ArXiv Mathematics e-prints* (2020a). arXiv: 2007.10869.

Hilger, S. Scaling limit and strict convexity of free energy for gradient models with non-convex potential. *ArXiv Mathematics e-prints* (2020b). arXiv: 2005.12973.

Ioffe, D., Shlosman, S., and Velenik, Y. 2D models of statistical physics with continuous symmetry: the case of singular interactions. *Comm. Math. Phys.*, **226** (2), 433–454 (2002). MR1892461.

Kharash, V. and Peled, R. The Fröhlich-Spencer Proof of the Berezinskii-Kosterlitz-Thouless Transition. *ArXiv Mathematics e-prints* (2017). arXiv: 1711.04720.

Lammers, P. A dichotomy theory for height functions. *ArXiv Mathematics e-prints* (2022a). arXiv: 2211.14365.

Lammers, P. Height function delocalisation on cubic planar graphs. *Probab. Theory Related Fields*, **182** (1-2), 531–550 (2022b). MR4367953.

Lammers, P. Bijecting the BKT transition. *ArXiv Mathematics e-prints* (2023). arXiv: 2301.06905.

Leindler, L. On a certain converse of Hölder's inequality. In *Linear operators and approximation (Proc. Conf., Math. Res. Inst., Oberwolfach, 1971)*, Internat. Ser. Numer. Math., Vol. 20, pp. 182–184. Birkhäuser Verlag, Basel-Stuttgart (1972). MR430188.

Magazinov, A. and Peled, R. Concentration inequalities for log-concave distributions with applications to random surface fluctuations. *Ann. Probab.*, **50** (2), 735–770 (2022). MR4399162.

Mathieu, P. Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.*, **130** (5), 1025–1046 (2008). MR2384074.

Mathieu, P. and Piatnitski, A. Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **463** (2085), 2287–2307 (2007). MR2345229.

Mathieu, P. and Remy, E. Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.*, **32** (1A), 100–128 (2004). MR2040777.

Miller, J. Fluctuations for the Ginzburg-Landau $\nabla\phi$ interface model on a bounded domain. *Comm. Math. Phys.*, **308** (3), 591–639 (2011). MR2855536.

Miłoś, P. and Peled, R. Delocalization of two-dimensional random surfaces with hard-core constraints. *Comm. Math. Phys.*, **340** (1), 1–46 (2015). MR3395146.

Mourrat, J.-C. and Otto, F. Anchored Nash inequalities and heat kernel bounds for static and dynamic degenerate environments. *J. Funct. Anal.*, **270** (1), 201–228 (2016). MR3419760.

Naddaf, A. and Spencer, T. On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, **183** (1), 55–84 (1997). MR1461951.

Nash, J. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, **80**, 931–954 (1958). MR100158.

Nirenberg, L. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, **13**, 115–162 (1959). MR109940.

Nishikawa, T. Hydrodynamic limit for the Ginzburg-Landau $\nabla\phi$ interface model with a conservation law. *J. Math. Sci. Univ. Tokyo*, **9** (3), 481–519 (2002). MR1930416.

Nishikawa, T. Hydrodynamic limit for the Ginzburg-Landau $\nabla\phi$ interface model with boundary conditions. *Probab. Theory Related Fields*, **127** (2), 205–227 (2003). MR2013982.

Prékopa, A. Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math. (Szeged)*, **32**, 301–316 (1971). MR0315079.

Prékopa, A. On logarithmic concave measures and functions. *Acta Sci. Math. (Szeged)*, **34**, 335–343 (1973). MR0404557.

Procaccia, E. B., Rosenthal, R., and Sapozhnikov, A. Quenched invariance principle for simple random walk on clusters in correlated percolation models. *Probab. Theory Related Fields*, **166** (3-4), 619–657 (2016). MR3568036.

Rassoul-Agha, F. and Seppäläinen, T. An almost sure invariance principle for random walks in a space-time random environment. *Probab. Theory Related Fields*, **133** (3), 299–314 (2005). MR2198014.

Redig, F. and Völlering, F. Random walks in dynamic random environments: a transference principle. *Ann. Probab.*, **41** (5), 3157–3180 (2013). MR3127878.

Sidoravicius, V. and Sznitman, A.-S. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields*, **129** (2), 219–244 (2004). MR2063376.

Taylor, M. E. *Measure theory and integration*, volume 76 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2006). ISBN 978-0-8218-4180-8; 0-8218-4180-7. MR2245472.

van Engelenburg, D. and Lis, M. An elementary proof of phase transition in the planar XY model. *Comm. Math. Phys.*, **399** (1), 85–104 (2023a). MR4567369.

van Engelenburg, D. and Lis, M. On the duality between height functions and continuous spin models. *ArXiv Mathematics e-prints* (2023b). arXiv: 2303.08596.

Velenik, Y. Localization and delocalization of random interfaces. *Probab. Surv.*, **3**, 112–169 (2006). MR2216964.

Wirth, M. Maximum of the integer-valued Gaussian free field. *ArXiv Mathematics e-prints* (2019). arXiv: 1907.08868.

Wu, W. Local central limit theorem for gradient field models. *ArXiv Mathematics e-prints* (2022). arXiv: 2202.13578.

Wu, W. and Zeitouni, O. Subsequential tightness of the maximum of two dimensional Ginzburg-Landau fields. *Electron. Commun. Probab.*, **24**, Paper No. 19, 12 (2019). MR3933043.

Ye, Z. Models of gradient type with sub-quadratic actions. *J. Math. Phys.*, **60** (7), 073304, 26 (2019). MR3982951.