

# QUANTITATIVE HOMOGENIZATION OF THE PARABOLIC AND ELLIPTIC GREEN'S FUNCTIONS ON PERCOLATION CLUSTERS

BY PAUL DARIO<sup>1</sup> AND CHENLIN GU<sup>2</sup>

<sup>1</sup>*School of Mathematical Sciences, Tel Aviv University, [pauldario@mail.tau.ac.il](mailto:pauldario@mail.tau.ac.il)*

<sup>2</sup>*DMA, École Normale Supérieure, [chenlin.gu@ens.fr](mailto:chenlin.gu@ens.fr)*

We study the heat kernel and the Green's function on the infinite supercritical percolation cluster in dimension  $d \geq 2$  and prove a quantitative homogenization theorem for these functions with an almost optimal rate of convergence. These results are a quantitative version of the local central limit theorem proved by Barlow and Hambly in (*Electron. J. Probab.* **14** (2009) 1–27). The proof relies on a structure of renormalization for the infinite percolation cluster introduced in (*Comm. Pure Appl. Math.* **71** (2018) 1717–1849), Gaussian bounds on the heat kernel established by Barlow in (*Ann. Probab.* **32** (2004) 3024–3084) and tools of the theory of quantitative stochastic homogenization. An important step in the proof is to establish a  $C^{0,1}$ -large-scale regularity theory for caloric functions on the infinite cluster and is of independent interest.

## CONTENTS

1. Introduction . . . . .	557
1.1. General introduction and main results . . . . .	557
1.2. Strategy of the proof . . . . .	561
1.3. Related results . . . . .	565
1.3.1. Related results about the random conductance model . . . . .	565
1.3.2. Related result about stochastic homogenization . . . . .	566
1.4. Further outlook and conjecture . . . . .	566
1.5. Organization of the article . . . . .	568
1.6. Notation and assumptions . . . . .	568
1.6.1. General notation and assumptions . . . . .	568
1.6.2. Notation of $\mathcal{O}_S$ . . . . .	568
1.6.3. Topology, functions and integration . . . . .	569
1.6.4. Discrete analysis and function spaces . . . . .	570
1.6.5. Convention for constants, exponents and minimal scales/times . . . . .	573
2. Preliminaries . . . . .	573
2.1. Supercritical percolation . . . . .	573
2.1.1. A partition of good cubes . . . . .	574
2.2. Functional inequalities on the infinite cluster . . . . .	577
2.3. Homogenization on percolation clusters . . . . .	578
2.3.1. The first-order corrector . . . . .	579
2.3.2. The centered flux . . . . .	580
2.4. Random walks on graphs . . . . .	580
3. Decay and Lipschitz regularity of the heat kernel . . . . .	581
3.1. Decay of the heat kernel . . . . .	581
3.2. Decay of the gradient of the Green's function . . . . .	583
3.2.1. Homogenization of the parabolic Dirichlet problem . . . . .	583
3.2.2. Large-scale $C^{0,1}$ -regularity estimate . . . . .	594
3.2.3. Decay of the gradient of the heat kernel . . . . .	598

Received September 2019; revised April 2020.  
*MSC2020 subject classifications.* 35B27, 60K37, 60K35.

*Key words and phrases.* Stochastic homogenization, local central limit theorem, parabolic equation, large-scale regularity, supercritical percolation.

4. Quantitative homogenization of the heat kernel . . . . . 599  
 4.1. Two regularization steps . . . . . 600  
 4.2. The two-scale expansion . . . . . 607  
 4.3. Proof of Theorem 1 . . . . . 615  
     Statement  $H_n$  . . . . . 615  
     The base case . . . . . 616  
     The iteration step . . . . . 618  
 5. Quantitative homogenization of the elliptic Green's function . . . . . 619  
 Appendix A: A concentration inequality for the density of the infinite cluster . . . . . 621  
 Appendix B: Quantification of the weak norm of the flux on the infinite cluster . . . . . 627  
 Acknowledgments . . . . . 632  
 References . . . . . 632

**1. Introduction.**

1.1. *General introduction and main results.* In this article, we study the continuous-time random walk on the infinite cluster of the supercritical Bernoulli bond percolation of the Euclidean lattice  $\mathbb{Z}^d$ , in dimension  $d \geq 2$ . The model considered is a specific case of the general random conductance model and can be described as follows. We let  $\mathcal{B}_d$  be the set of bonds of  $\mathbb{Z}^d$ , that is, the set of unordered pairs of nearest neighbors of  $\mathbb{Z}^d$ . We denote by  $\Omega$  the set of functions from  $\mathcal{B}_d$  to the set of nonnegative real numbers  $[0, \infty)$ . A generic element of  $\Omega$  is denoted by  $\mathbf{a}$  and called *an environment*.

For a given environment  $\mathbf{a} \in \Omega$  and a given bond  $e \in \mathcal{B}_d$ , we call the value  $\mathbf{a}(e)$  *the conductance* of the bond  $e$ . We fix an ellipticity parameter  $\lambda \in (0, 1]$  and add some randomness to the model by assuming that the collection of conductances  $\{\mathbf{a}(e)\}_{e \in \mathcal{B}_d}$  is an i.i.d. family of random variables whose law is supported in the set  $\{0\} \cup [\lambda, 1]$ . We define  $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) \neq 0]$  and assume that

$$\mathfrak{p} > \mathfrak{p}_c(d),$$

where  $\mathfrak{p}_c(d)$  is the bond percolation threshold for the lattice  $\mathbb{Z}^d$ . This assumption ensures that, almost surely, there exists a unique infinite connected component of edges with nonzero conductances (or cluster) which we denote by  $\mathcal{C}_\infty$  (see [35]). This cluster has a nonzero density which is given by the probability  $\theta(\mathfrak{p}) := \mathbb{P}[0 \in \mathcal{C}_\infty]$ . The model of continuous-time random walk considered in this article is *the variable speed random walk* (or VSRW) and is defined as follows. Given an environment  $\mathbf{a} \in \Omega$  and a starting point  $y \in \mathcal{C}_\infty$ , we endow each edge  $e \in \mathcal{B}_d$  with a random clock whose law is exponential of parameter  $\mathbf{a}(e)$  and assume that they are mutually independent. We then let  $(X_t)_{t \geq 0}$  be the random walk which starts from  $y$ , that is,  $X_0 = y$ , and, when  $X(t) = x$ , the random walker waits at  $x$  until one of the clocks at an adjacent edge to  $x$  rings, and moves across the edge to the neighboring point instantly. We then restart the clocks. This construction gives rise to a continuous-time Markov process on the infinite cluster  $\mathcal{C}_\infty$  whose generator is the elliptic operator  $\nabla \cdot \mathbf{a} \nabla$  defined by, for each function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$  and each point  $x \in \mathcal{C}_\infty$ ,

$$(1) \quad \nabla \cdot \mathbf{a} \nabla u(x) := \sum_{z \sim x} \mathbf{a}(\{x, z\})(u(z) - u(x)).$$

We denote the transition density of the random walk by

$$p(t, x, y) = p^{\mathbf{a}}(t, x, y) := \mathbb{P}_y^{\mathbf{a}}(X_t = x),$$

and often omit the dependence in the environment  $\mathbf{a}$  in the notation. The transition density can be equivalently defined as the solution of the parabolic equation

$$(2) \quad \begin{cases} \partial_t p(\cdot, \cdot, y) - \nabla \cdot \mathbf{a} \nabla p(\cdot, \cdot, y) = 0 & \text{in } (0, \infty) \times \mathcal{C}_\infty, \\ p(0, \cdot, y) = \delta_y & \text{in } \mathcal{C}_\infty. \end{cases}$$

Due to this characterization, we often refer to the transition density  $p$  as the *heat kernel* or the *parabolic Green's function*.

There are other related models of random walk on supercritical percolation clusters which have been studied in the literature, two of the most common ones are:

1. *The constant speed random walk* (or CSRW), the random walker starts from a point  $y \in \mathcal{C}_\infty$ . When  $X(t) = x$ , it waits for an exponential time of parameter 1 and then jumps to a neighboring point  $z$  according to the transition probability

$$(3) \quad P(x, z) = \frac{\mathbf{a}(\{x, z\})}{\sum_{w \sim x} \mathbf{a}(\{x, w\})}.$$

This construction also gives rise to a continuous-time Markov process whose generator is given by, for each function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$  and each point  $x \in \mathcal{C}_\infty$ ,

$$\frac{1}{\sum_{z \sim x} \mathbf{a}(\{x, z\})} \sum_{z \sim x} \mathbf{a}(\{x, z\})(u(z) - u(x)).$$

2. *The simple random walk* (or SRW), the random walk  $(X_n)_{n \in \mathbb{N}}$  is indexed on the integers, it starts from a point  $y \in \mathcal{C}_\infty$ , when  $X_n = x$ , the value of  $X_{n+1}$  is chosen randomly among all the neighbors of  $x$  following the transition probability (3).

These processes have similar, although not identical, properties and have been the subject of interest in the literature. In the case of the percolation cluster, that is, when the environment  $\mathbf{a}$  is only allowed to take the values 0 or 1, an annealed invariance principle was proved in [42] by De Masi, Ferrari, Goldstein and Wick. In [79], Sidoravicius and Sznitman proved a quenched invariance principle for the simple random walk in dimension  $d \geq 4$ . This result was extended to every dimension  $d \geq 2$  by Berger and Biskup in [28] (for the SRW) and by Mathieu and Piatnitski in [68] (for the CSRW).

For the VSRW, a similar quenched invariance principle holds: there exists a deterministic diffusivity constant  $\bar{\sigma} > 0$  such that, for almost every environment, the following convergence holds in the Skorokhod topology:

$$(4) \quad \varepsilon X \xrightarrow[\varepsilon \rightarrow 0]{\text{(law)}} \bar{\sigma} B.,$$

where  $B.$  is a standard Brownian motion. From a homogenization perspective, the diffusivity  $\bar{\sigma}^2$  of the limiting Brownian motion is related to the homogenized coefficient  $\bar{\mathbf{a}}$  associated to the elliptic and parabolic problems on the percolation cluster by the identity  $\bar{\mathbf{a}} = \frac{1}{2} \theta(\mathbf{p}) \bar{\sigma}^2$  (see the formula (181) of Appendix B).

The properties of the heat kernel  $p$  on the infinite cluster have been investigated in the literature. In [69], Mathieu and Remy proved that, almost surely, the heat kernel decays as fast as  $t^{-d/2}$ . These bounds were extended in [21] by Barlow who established Gaussian lower and upper bounds for this function; we will recall his precise result in Theorem 3.1 below.

In the article [23], Barlow and Hambly proved a parabolic Harnack inequality, a local central limit theorem for the CSRW, and bounds on the elliptic Green's function on the infinite cluster. Their main result can be adapted to the case of the VSRW, and reads as follows: if we define, for each  $t \geq 0$  and  $x \in \mathbb{R}^d$ ,

$$(5) \quad \bar{p}(t, x) := \frac{1}{(2\pi \bar{\sigma}^2 t)^{d/2}} \exp\left(-\frac{|x|^2}{2\bar{\sigma}^2 t}\right),$$

the heat kernel with diffusivity  $\bar{\sigma}$ , then, for each time  $T > 0$ , the following convergence holds,  $\mathbb{P}$ -almost surely on the event  $\{0 \in \mathcal{C}_\infty\}$ :

$$(6) \quad \lim_{n \rightarrow \infty} |n^{d/2} p(nt, g_n^\omega(x), 0) - \theta(\mathbf{p})^{-1} \bar{p}(t, x)| = 0,$$

uniformly in the spatial variable  $x \in \mathbb{R}^d$  and in the time variable  $t \geq T$ , where the notation  $g_n^\omega(x)$  means the closest point to  $\sqrt{n}x$  in the infinite cluster under the environment  $\omega$ .

The main result of this article is a quantitative version of the local central limit theorem for the VSRW and is stated below.

**THEOREM 1.1.** *For each exponent  $\delta > 0$ , there exist a positive constant  $C < \infty$  and an exponent  $s > 0$ , depending only on the parameters  $d, \lambda, \mathfrak{p}$  and  $\delta$ , such that for every  $y \in \mathbb{Z}^d$ , there exists a nonnegative random time  $\mathcal{T}_{\text{par},\delta}(y)$  satisfying the stochastic integrability estimate*

$$\forall T \geq 0, \quad \mathbb{P}(\mathcal{T}_{\text{par},\delta}(y) \geq T) \leq C \exp\left(-\frac{T^s}{C}\right),$$

such that, on the event  $\{y \in \mathcal{C}_\infty\}$ , for every  $x \in \mathcal{C}_\infty$  and every  $t \geq \max(\mathcal{T}_{\text{par},\delta}(y), |x - y|)$ ,

$$(7) \quad |p(t, x, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, x - y)| \leq Ct^{-\frac{d}{2} - (\frac{1}{2} - \delta)} \exp\left(-\frac{|x - y|^2}{Ct}\right).$$

**REMARK 1.** The heat kernel  $p$  does not exactly converge to the heat kernel  $\bar{p}$  and there is an additional normalization constant  $\theta(\mathfrak{p})^{-1}$  in (7). A heuristic reason explaining why such a term is necessary is the following: since  $p(t, \cdot, y)$  is a probability measure on the infinite cluster, one has

$$\sum_{x \in \mathcal{C}_\infty} p(t, x, y) = 1.$$

One also has, by definition of the heat kernel  $\bar{p}$ ,

$$\int_{\mathbb{R}^d} \bar{p}(t, x - y) dx = 1.$$

Since the infinite cluster has density  $\theta(\mathfrak{p})$ , we expect that

$$\sum_{x \in \mathcal{C}_\infty} \bar{p}(t, x - y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{R}^d} \bar{p}(t, x - y) dx = \theta(\mathfrak{p}),$$

and we refer to Proposition 13 for a precise statement. As a consequence, we cannot expect the maps  $p$  and  $\bar{p}$  to be close since they have different mass on the infinite cluster; adding the normalization term  $\theta(\mathfrak{p})^{-1}$  ensures that the mass of  $\theta(\mathfrak{p})^{-1} \bar{p}$  on the infinite cluster is approximately equal to 1.

As an application of this result, we deduce a quantitative homogenization theorem for the elliptic Green's function on the infinite cluster. In dimension  $d \geq 3$ , given an environment  $\mathbf{a} \in \Omega$  and a point  $y \in \mathcal{C}_\infty$ , we define the Green's function  $g(\cdot, y)$  as the solution of the equation

$$-\nabla \cdot \mathbf{a} \nabla g(\cdot, y) = \delta_y \quad \text{in } \mathcal{C}_\infty \text{ such that } g(x, y) \xrightarrow{x \rightarrow \infty} 0.$$

This function exists, is unique almost surely and is related to the transition probability  $p$  through the identity

$$(8) \quad g(x, y) = \int_0^\infty p(t, x, y) dt.$$

In dimension 2, the situation is different since the Green's function is not bounded at infinity, and we define  $g(\cdot, y)$  as the unique function which satisfies

$$-\nabla \cdot \mathbf{a} \nabla g(\cdot, y) = \delta_y \quad \text{in } \mathcal{C}_\infty, \quad \frac{1}{|x|} g(x, y) \xrightarrow{x \rightarrow \infty} 0 \quad \text{and} \quad g(y, y) = 0.$$

This function is related to the transition probability  $p$  through the identity

$$g(x, y) = \int_0^\infty (p(t, x, y) - p(t, y, y)) dt.$$

In the statement below, we denote by  $\bar{g}$  the homogenized Green’s function defined by the formula, for each point  $x \in \mathbb{R}^d \setminus \{0\}$ ,

$$(9) \quad \bar{g}(x) := \begin{cases} -\frac{1}{\pi \bar{\sigma}^2 \theta(\mathfrak{p})} \ln |x| & \text{if } d = 2, \\ \frac{\Gamma(d/2 - 1)}{(2\pi^{d/2} \bar{\sigma}^2 \theta(\mathfrak{p}))} \frac{1}{|x|^{d-2}} & \text{if } d \geq 3, \end{cases}$$

where the symbol  $\Gamma$  denotes the standard Gamma function. Theorem 1.2 describes the asymptotic behavior of the Green’s function  $g$ .

**THEOREM 1.2.** *For each exponent  $\delta > 0$ , there exist a positive constant  $C < \infty$  and an exponent  $s > 0$ , depending only on the parameters  $d, \lambda, \mathfrak{p}$  and  $\delta$ , such that for every  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{ell},\delta}(y)$  satisfying*

$$\forall R \geq 0, \quad \mathbb{P}(\mathcal{M}_{\text{ell},\delta}(y) \geq R) \leq C \exp\left(-\frac{R^s}{C}\right),$$

such that, on the event  $\{y \in \mathcal{C}_\infty\}$ :

1. In dimension  $d \geq 3$ , for every point  $x \in \mathcal{C}_\infty$  satisfying  $|x - y| \geq \mathcal{M}_{\text{ell},\delta}(y)$ ,

$$(10) \quad |g(x, y) - \bar{g}(x - y)| \leq \frac{1}{|x - y|^{1-\delta}} \frac{C}{|x - y|^{d-2}}.$$

2. In dimension 2, the limit

$$K(y) := \lim_{x \rightarrow \infty} (g(x, y) - \bar{g}(x - y)),$$

exists, is finite almost surely and satisfies the stochastic integrability estimate

$$\forall R \geq 0, \quad \mathbb{P}(|K(y)| \geq R) \leq C \exp\left(-\frac{R^s}{C}\right).$$

Moreover, for every point  $x \in \mathcal{C}_\infty$  satisfying  $|x - y| \geq \mathcal{M}_{\text{ell},\delta}(y)$ ,

$$(11) \quad |g(x, y) - \bar{g}(x - y) - K(y)| \leq \frac{C}{|x - y|^{1-\delta}}.$$

**REMARK 2.** In dimension 2, the situation is specific due to the unbounded behavior of the Green’s function, and the theorem identifies the first-order term. The second term in the asymptotic development is of constant order and is random: with the normalization chosen for the Green’s function, the constant  $K$  depends on the geometry of the infinite cluster and cannot be deterministic. We nevertheless expect it not to be too large and prove that it satisfies a stretched exponential stochastic integrability estimate.

We complete this section by mentioning a potential application of these theorems. Theorem 1.1 shows that the law of the VSRW on the infinite percolation cluster converges quantitatively to the one of the Brownian motion  $(\bar{\sigma} B_t)_{t \geq 0}$ . To go one step further in the analysis, one can try to construct a coupling between the random walk  $(X_t)_{t \geq 0}$  and the Brownian motion  $(\bar{\sigma} B_t)_{t \geq 0}$  such that their trajectories are close, that is, such that  $\sup_{0 \leq s \leq t} |X_s - \bar{\sigma} B_s|$  is small. This question is known as the embedding problem: a good error should be at least of order  $o(\sqrt{t})$ . In the case of the simple random walk on  $\mathbb{Z}^d$ , the optimal result is given by the Komlós–Major–Tusnady approximation (see [63, 64]) and gives an error of order  $O(\log t)$ . Adapting this result to the setting considered here requires to take into account the degenerate geometry of the percolation cluster; we believe that Theorem 1.1 can be useful in this regard.

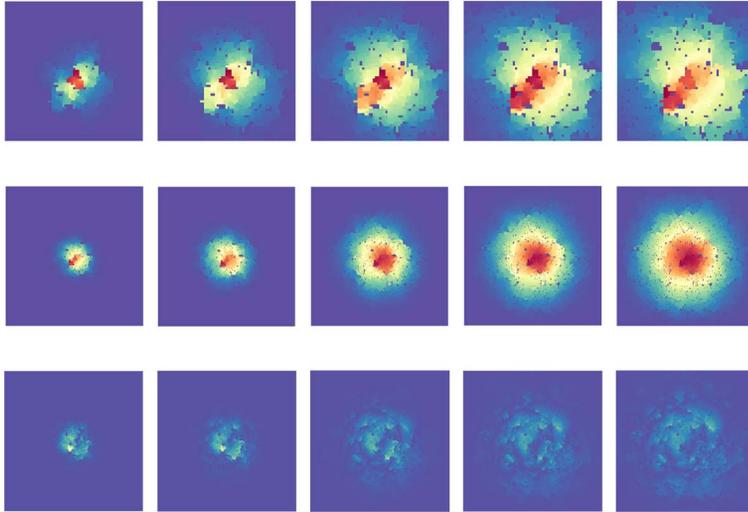


FIG. 1. A simulation to illustrate the convergence of the parabolic Green's function for the VSRW on the infinite cluster with  $\mathfrak{p} = 0.6$ . We use different colors to represent the level sets of the map  $t^{\frac{d}{2}} p(t, \cdot, y)$  in the first two rows. The figures in the first row are drawn for the short times  $t = 100, 200, 300, 400, 500$  in a cube of size  $64 \times 64$  and the level sets of the heat kernel are perturbed by the geometry of the infinite cluster. In the second row, the figures are drawn for the long times  $t = 500, 1000, 2000, 3000, 4000$  and in a cube of size  $256 \times 256$ ; in this case, homogenization happens and the geometry of the level sets of the heat kernel is similar to the one of a Gaussian heat kernel. In the third row, we simulate the function  $t^{\frac{d}{2}} |p(t, \cdot, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, \cdot - y)| \mathbb{1}_{\{x \in \mathcal{C}_\infty\}}$  associated to the figures in the second line and we observe that the errors decay to 0 as the time tends to infinity.

1.2. *Strategy of the proof.* On the supercritical percolation cluster, a qualitative version of Theorem 1.1 is established by Barlow and Hambly in [23], where the strategy implemented is to first prove a parabolic Harnack inequality for the heat equation. From the Harnack inequality, one derives a  $C^{0,\alpha}$ -Hölder regularity estimate (for some small exponent  $\alpha > 0$ ) on the heat kernel. It is then possible to combine this additional regularity with the quenched invariance principle, established on the percolation cluster in [28, 69, 79], to obtain the local central limit theorem.

In the present article, the strategy adopted is different and follows ideas from the theory of stochastic homogenization, more specifically the ones of [15], Chapter 8. A first crucial ingredient in the proof is the first-order corrector, which can be characterized as follows: given a slope  $p \in \mathbb{R}^d$ , the corrector  $\chi_p$  is defined as the unique function (up to a constant) which is a solution of the elliptic equation

$$-\nabla \cdot \mathbf{a}(p + \nabla \chi_p) = 0 \quad \text{in } \mathcal{C}_\infty,$$

and which has sublinear oscillation, that is,

$$\frac{1}{r} \text{osc}_{\mathcal{C}_\infty \cap B_r} \chi_p := \frac{1}{r} \left( \sup_{\mathcal{C}_\infty \cap B_r} \chi_p - \inf_{\mathcal{C}_\infty \cap B_r} \chi_p \right) \xrightarrow{r \rightarrow \infty} 0.$$

The corrector is defined and some of its important properties are presented in Section 2.3. We note that the use of the corrector to study random walk on supercritical percolation cluster is not new: it is a key ingredient in the proofs of the quenched invariance principle (see [28, 69, 79]). Once equipped with this function, the analysis relies on a classical strategy in stochastic homogenization: the two-scale expansion. The general approach relies on the definition of the function

$$(12) \quad h(t, x, y) := \theta(\mathfrak{p})^{-1} \left( \bar{p}(t, x - y) + \sum_{k=1}^d \partial_k \bar{p}(t, x - y) \chi_{e_k}(x) \right),$$

where  $(e_k)_{k=\{1,\dots,d\}}$  denotes the canonical basis of  $\mathbb{R}^d$  and  $\bar{p}$  is the continuous heat kernel defined in (5). The strategy is then to compute the value of

$$(13) \quad \partial_t h - \nabla \cdot \mathbf{a} \nabla h,$$

by using the explicit formula on  $h$  stated in (12) and to prove that it is quantitatively small in the correct functional space (precisely, the parabolic  $\underline{H}^{-1}$  space introduced in (33)). Obtaining this result requires two types of quantitative information on the corrector:

- One needs to have quantitative sublinearity of the corrector, that is,

$$(14) \quad \frac{1}{r^\alpha} \text{osc}_{\mathcal{C}_\infty \cap B_r} \chi_p \xrightarrow{r \rightarrow \infty} 0,$$

for every exponent  $\alpha > 0$ .

- One needs to have a quantitative control on the flux of the corrector in the weak  $\underline{H}^{-1}$  norm,

$$(15) \quad \frac{1}{r^\alpha} \left\| \mathbf{a}(p + \nabla \chi_p) - \frac{1}{2} \bar{\sigma}^2 p \right\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} \xrightarrow{r \rightarrow \infty} 0,$$

for every exponent  $\alpha > 0$ , where  $\bar{\sigma}^2$  is the same diffusivity constant as in the definition (5) of the heat kernel  $\bar{p}$ .

The sublinearity of the corrector in the setting of the percolation cluster is established qualitatively in [28, 69, 79] and quantitatively in [12, 40, 61]. The second property (15) cannot be directly deduced from the results of [12, 40, 61] and Appendix B is devoted to the proof of this result.

Once one has good quantitative control over the  $\underline{H}^{-1}$ -norm of  $\partial_t h - \nabla \cdot \mathbf{a} \nabla h$ , the proof of the result follows from the following two arguments:

1. First, one shows that the function  $h$  is (quantitatively) close to the function  $\theta^{-1}(\mathfrak{p})\bar{p}$ . This is achieved by proving that the second term in the right-hand side of (12) is small and relies on the quantitative sublinearity of the corrector stated in (14).

2. Second, one needs to show that the function  $h$  is (quantitatively) close to the heat kernel  $p$ . To prove this, the strategy is to use that the map  $p$  solves the parabolic equation

$$\partial_t p - \nabla \cdot \mathbf{a} \nabla p = 0,$$

and subtract it from (13) to obtain that  $\partial_t(p - h) - \nabla \cdot \mathbf{a} \nabla(p - h)$  is small in the  $\underline{H}^{-1}$  norm. We then use the function  $(p - h)$  as a test function in the previous equation, to deduce that  $(p - h)$  has to be small in the  $H^1$ -norm.

This strategy is essentially carried out in Section 4.2. Nevertheless, a number of difficulties have to be treated in order to implement it. They are mainly due to three distinct causes which are listed below.

First, the heat kernel  $p$  has an initial condition at time  $t = 0$  which is a Dirac (see equation (2)). It is rather singular and causes serious troubles in the analysis. To fix this issue, one replaces the initial condition in (2) by a function which is smoother, but which is still a good approximation of the Dirac function. The argument is sketched in the following paragraph. We fix a large time  $t > 0$  and want to prove the main estimate (7) for this particular time  $t$ . To this end, we replace the initial condition  $\delta_y$  by the function  $\bar{p}(\tau, \cdot - y)$  for some time  $\tau \ll t$ , and we define

$$(16) \quad \begin{cases} \partial_t q - \nabla \cdot (\mathbf{a} \nabla q) = 0 & \text{in } (\tau, \infty) \times \mathcal{C}_\infty, \\ q(\tau, \cdot, y) = \theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{on } \mathcal{C}_\infty. \end{cases}$$

The strategy is then to make the following compromise: we want to choose the coefficient  $\tau$  small enough (in particular, much smaller than  $t$ ) so that the initial data  $\bar{p}(\tau, \cdot - y)$  is close

to the Dirac function  $\delta_y$ , the objective being that the function  $q(t, \cdot, y)$  is close to  $p(t, \cdot, y)$  (see Lemma 4.1); we also want to choose  $\tau$  large enough so that the initial data  $\bar{p}(\tau, \cdot)$  is smooth enough. Our choice will be  $\tau = t^{1-\kappa}$  for some small exponent  $\kappa > 0$ . This approach is essentially the subject of Section 4.1.

The second difficulty is that the two-scale expansion described at the beginning of the section only yields the result for a small exponent, that is, we obtain a result of the form

$$(17) \quad |p(t, x, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, x - y)| \leq Ct^{-\kappa} t^{-\frac{d}{2}} \exp\left(-\frac{|x - y|^2}{Ct}\right),$$

for a small exponent  $\kappa > 0$ . This result is much weaker than the near-optimal exponent  $\frac{1}{2} - \delta$  stated in Theorem 1.1. The strategy is thus to improve the value of the exponent by a bootstrap argument: by redoing the two-scale expansion and by using the estimate (17) in the proof, we obtain an improved estimate of the form

$$(18) \quad |p(t, x, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, x - y)| \leq Ct^{-\kappa_1} t^{-\frac{d}{2}} \exp\left(-\frac{|x - y|^2}{Ct}\right),$$

where  $\kappa_1$  is a new exponent which is strictly larger than the original exponent  $\kappa$ . We can then redo the proof a second time and use the estimate (18) to obtain the inequality with an exponent  $\kappa_2$  strictly larger than  $\kappa_1$ . An iteration of the argument shows that there exists an increasing sequence  $\kappa_n$  such that, for each  $n \in \mathbb{N}$ , the following estimate holds:

$$(19) \quad |p(t, x, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, x - y)| \leq Ct^{-\kappa_n} t^{-\frac{d}{2}} \exp\left(-\frac{|x - y|^2}{Ct}\right).$$

The sequence  $\kappa_n$  is defined inductively (see the formula (148)) and we can prove that it converges toward the value  $\frac{1}{2}$ ; this is sufficient to prove the near optimal estimate stated in Theorem 1.1.

The third difficulty is the degenerate structure of the environment. It is treated by defining a renormalization structure for the infinite cluster which was first introduced in [12]: building upon standard results in supercritical percolation, we construct a partition of the lattice  $\mathbb{Z}^d$  into cubes of different random sizes which are well connected in the sense of Antal, Penrose and Pisztorá (see [10, 75]), using a Calderón–Zygmund-type stopping time argument. The sizes of the cubes of the partition are random variables which measure how close the geometry of the cluster is from the geometry of the lattice: in the regions where the sizes of the cubes are small, the cluster is well behaved and its geometry is similar to the one of the Euclidean lattice, while in the regions where the sizes of the cubes are large, the geometry of the cluster is ill-behaved (see Figure 3). The probability to have a large cube in the partition is small and stretched exponential integrability estimates are available for these random variables (see Proposition 2(iii) or [75]).

This partition provides a random scale above which the geometry of the infinite cluster is similar to the one of the Euclidean lattice and it allows to adapt the tools of functional analysis needed to perform the two-scale expansion to the percolation cluster. Similar strategies using renormalization techniques were used to study the random walk on the supercritical percolation cluster and we refer for instance to the work of Barlow in [21], who established a Poincaré inequality on the percolation cluster, or to the one of Mathieu and Remy in [69].

The general strategy to study the random walk on the infinite cluster is thus to prove that there exists a random scale above which the geometry of the infinite cluster  $\mathcal{C}_\infty$  is similar to the geometry of the lattice  $\mathbb{Z}^d$ , and to deduce from it that, above a random time which is related to the aforementioned random scale, the random walk has a behavior which is similar to the one of the random walk on  $\mathbb{Z}^d$  (see Figure 1). As a consequence, most of the results described in this article only hold above a random scale (or random time) above which the

infinite cluster has renormalized. Moreover, we need to appeal to a number of random scales (or random times) in the proofs, above which some analytical tools are available: the scale  $\mathcal{M}_{\text{reg}}$  above which a  $C^{0,1}$ -regularity theory is valid (see Theorem 1.3), the time  $\mathcal{T}_{\text{NA}}$  above which a Nash–Aronson estimate for the heat kernel is available (see Theorem 3.1), etc. For all of these random scales and times, stretched exponential integrability estimates are valid.

This strategy describes the proof of Theorem 1.1. Once this result is established, Theorem 1.2, pertaining to the elliptic Green’s function, can be deduced from it thanks to the Duhamel principle stated in (8). This is the subject of Section 5.

We complete this section by describing the content and purposes of Section 3. To perform the analysis described in the previous paragraphs, and in particular to prove that the function  $q$  defined (16) is a good approximation of the heat kernel  $p$ , one needs to have some control over the quantities at stake. In particular, it is useful to have a good control on the heat kernel  $p$  and its gradient  $\nabla p$ . The first one is given by the article of Barlow [21], which provides Gaussian upper and lower bounds for the heat kernel  $p$  (see Theorem 3.1). For the gradient of the heat kernel, we expect to have a behavior similar to the one of the gradient of the heat kernel on  $\mathbb{R}^d$ , that is, a  $C^{0,1}$ -regularity estimate of the form

$$|\nabla_x p(t, x, y)| \leq Ct^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right).$$

Section 3 is devoted to proving a large-scale version of this estimate and is independent of Section 4 and Section 5. The precise statement established in this section is the following.

**THEOREM 1.3.** *There exist an exponent  $s := s(d, \mathfrak{p}, \lambda) > 0$ , a finite positive constant  $C := C(d, \mathfrak{p}, \lambda)$  such that for each point  $x \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{reg}}(x)$  satisfying the stochastic integrability estimate*

$$(20) \quad \forall R \geq 0, \quad \mathbb{P}(\mathcal{M}_{\text{reg}}(x) \geq R) \leq C \exp\left(-\frac{R^s}{C}\right),$$

such that the following statement is valid: for every radius  $r \geq \mathcal{M}_{\text{reg}}(x)$ , every point  $y \in \mathcal{C}_\infty$  and every time  $t \geq \max(4r^2, |x - y|)$ , the following estimate holds:

$$\|\nabla_x p(t, \cdot, y)\|_{\underline{L}^2(B_r(x) \cap \mathcal{C}_\infty)} \leq Ct^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right),$$

where the notation  $\underline{L}^2(B_r(x) \cap \mathcal{C}_\infty)$  denotes the average  $L^2$ -norm over the set  $B_r(x) \cap \mathcal{C}_\infty$  and is defined in (31).

**REMARK 3.** By using the symmetry of the heat kernel, a similar regularity estimate holds for the gradient in the second variable: for each point  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{reg}}(y)$  satisfying the stochastic integrability estimate (20) such that for every radius  $r \geq \mathcal{M}_{\text{reg}}(y)$ , every point  $x \in \mathcal{C}_\infty$  and every time  $t \geq \max(4r^2, |x - y|)$ ,

$$\|\nabla_y p(t, x, \cdot)\|_{\underline{L}^2(B_r(y) \cap \mathcal{C}_\infty)} \leq Ct^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right).$$

The strategy of the proof of this result relies on tools from homogenization theory, in particular the two-scale expansion and the large-scale regularity theory. It is described at the beginning of Section 3.

### 1.3. Related results.

1.3.1. *Related results about the random conductance model.* The random conductance model has been the subject of active research over the recent years, by various authors and under different assumptions over the law of the environment. In the case of uniform ellipticity, that is, when the environment is allowed to take values in  $[\lambda, 1]$ , a quenched invariance principle is proved by Osada in [73] (in the continuous setting) and by Sidoravicius and Sznitman in [79] (in the discrete setting). Gaussian bounds on the heat kernel follow from [43]. This framework is the one of the theory of stochastic homogenization and we refer to Section 1.3.2 for further information.

In the setting when the conductances are only bounded from above, a quenched invariance principle was proved by Mathieu in [67] and by Biskup and Prescott in [32]. In the case when the conductances are bounded from below, a quenched invariance principle and heat kernel bounds are proved in [22] by Barlow and Deuschel. In [3], Andres, Barlow, Deuschel and Hambly established a quenched invariance principle in the general case when the conductances are allowed to take values in  $[0, \infty)$ .

The i.i.d. assumption on the environment can be relaxed: in [5], Andres, Deuschel and Slowik proved a quenched invariance principle for the random walk for general ergodic environment under the moment condition

$$(21) \quad \mathbb{E}[\mathbf{a}(e)^p] + \mathbb{E}[\mathbf{a}(e)^{-q}] < \infty \quad \text{for } p, q \in (1, \infty) \text{ satisfying } \frac{1}{p} + \frac{1}{q} < \frac{2}{d}.$$

We also refer to the works of Chiarni and Deuschel [37], Deuschel, Nguyen and Slowik [44] and Bella and Schäffner [25] for additional quenched invariance principles in degenerate ergodic environment. The case of ergodic, time-dependent, degenerate environment is investigated by Andres, Chiarini, Deuschel and Slowik in [4] where they establish a quenched invariance principle under some moment conditions on the environment. More general models of random walks on percolation clusters with long range correlation, including random interacements and level sets of the Gaussian free field, are studied by Procaccia, Rosenthal and Sapozhnikov in [77], where a quenched invariance principle is established.

The heat kernel has been studied under various assumptions on the environment: a first important property that needs to be investigated is the question of the existence of Gaussian lower and upper bounds. Such estimates are valid in the case of the percolation cluster presented in this article and were originally proved by Barlow in [21]. This result also holds when the conductances are bounded from below and we refer to the works of Mourrat [70] (Theorem 10.1 of the second arxiv version) and of Barlow, Deuschel [22]. It is also known that it cannot hold in full generality: in [29], Berger, Biskup, Hoffman and Kozma established that, when the law of the conductances has a fat tail at 0, the heat kernel can behave anomalistically due to trapping phenomenon (even though a quenched invariance principle still holds by [3]). We refer to the works of Barlow, Boukhadra [31] and Boukhadra [33, 34] for additional results in this direction. Gaussian estimates on the heat kernel for more general graphs were studied by Andres, Deuschel and Slowik in [7] and [9].

The question of Gaussian upper and lower bounds on the heat kernel is related, and in many situations equivalent, to the existence of a parabolic Harnack inequality (see, for instance, Delmotte [43]). On the percolation cluster, the parabolic Harnack inequality is established in [23]. We refer to the article of Andres, Deuschel, Slowik [6] for a proof of elliptic and parabolic Harnack inequalities on general graphs with unbounded weights, to the work of Sapozhnikov [78] for a proof of quenched heat kernel bounds and parabolic Harnack inequality for a general class of percolation models with long-range correlations on  $\mathbb{Z}^d$  and to the articles of Chang [36] and Alves and Sapozhnikov [2] for similar results on loop soup models.

Results on the elliptic Green’s function usually follow from the ones established on the parabolic Green’s function, by an application of the formula (8) in dimension larger than 3. In dimension 2, the situation is different and requires separate considerations; in [8], Andres, Deuschel and Slowik characterize the asymptotics of the Green’s function associated to the random walk killed upon exiting a ball under general assumptions on the environment.

Finally, we refer to [30] for a general review on the random conductance model.

1.3.2. *Related result about stochastic homogenization.* The theory of qualitative stochastic homogenization was developed in the 1980s, with the works of Kozlov [65], Papanicolaou and Varadhan [74] and Yurinskii [81] in the uniformly elliptic setting. Still in the uniformly elliptic setting, a quantitative theory of stochastic homogenization has been developed in the recent years up to the point that it is now well understood thanks to the works of Gloria and Otto in [55–58] and Gloria, Neukamm, Otto [53, 54], building upon the ideas of Naddaf and Spencer in [71]. These results have applications to random walks in random environment, as is explained in [46]. Another approach was initiated by Armstrong and Smart in [17], who extended the techniques of Avellaneda and Lin [19, 20] and the ones of Dal Maso and Modica [38, 39]. These results were then improved in [13, 14], and we refer to the monograph [15] for a detailed review of this approach.

The aforementioned works treated the case of uniformly elliptic environments and the question of the extension of the theory to degenerate environments has drawn some attention over the past few years. A number of results have been achieved and some of them are closely related to the works on the random conductance model presented in the previous section. In [72], Neukamm, Schäffner and Schlömerkemper proved  $\Gamma$ -convergence of the Dirichlet energy associated to some nonconvex energy functionals with degenerate growth. In [66], Lamacz, Neukamm and Otto studied a model of Bernoulli bond percolation, which is modified such that every bond in a fixed direction is declared open. In [48], Fleger, Heida and Slowik proved homogenization results for a degenerate random conductance model with long range jumps. In [24], Bella, Fehrman and Otto studied homogenization of degenerate environment under the moment condition (21) and established a first-order Liouville theorem as well as a large-scale  $C^{1,\alpha}$ -regularity estimate for  $\mathbf{a}$ -harmonic functions. In [51], Giunti, Höfer and Velázquez studied homogenization for the Poisson equation in a randomly perforated domain. In [12], Armstrong and the first author implemented the techniques of [15] to the percolation cluster to obtain quantitative homogenization results as well as a large-scale regularity theory.

1.4. *Further outlook and conjecture.* The results of this article present quantitative rates of convergence for the parabolic and elliptic Green’s functions on the percolation cluster. We do not expect the result to be optimal: the quantitative rate of convergence  $\frac{1}{2} - \delta$  and the stochastic integrability  $s$  in Theorem 1.1 can be improved and so is the case for Theorem 1.2. We expect the following conjecture to hold.

CONJECTURE 1. *Fix  $s \in (0, \frac{2(d-1)}{d})$ , there exists a positive constant  $C < \infty$  depending on the parameters  $d, \mathbf{p}, \lambda$  and  $s$ , such that, for each time  $t > 0$  and each pair of points  $x, y \in \mathbb{Z}^d$  such that  $|x - y| \leq t$ , conditionally on the event  $\{x, y \in \mathcal{C}_\infty\}$ ,*

$$|p(t, x, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - y)| \leq \begin{cases} \mathcal{O}_s \left( Ct^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right) \right) & \text{when } d \geq 3, \\ \mathcal{O}_s \left( C \log^{\frac{1}{2}}(1+t) t^{-\frac{3}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right) \right) & \text{when } d = 2, \end{cases}$$

where the notation  $\mathcal{O}_s$  is used to measure the stochastic integrability and is defined in Section 1.6.2. For the elliptic Green's function, a similar result holds:

1. In dimension  $d \geq 3$ , for each  $x, y \in \mathbb{Z}^d$ , conditionally on the event  $\{x, y \in \mathcal{C}_\infty\}$ ,

$$|g(x, y) - \bar{g}(x - y)| \leq \mathcal{O}_s(C|x - y|^{1-d}),$$

where the function  $\bar{g}$  is defined in the equation (9).

2. In dimension 2, for each  $y \in \mathbb{Z}^d$ , conditionally on the event  $\{y \in \mathcal{C}_\infty\}$ , the limit

$$K(y) := \lim_{x \rightarrow \infty} g(x, y) - \bar{g}(x - y)$$

exists, is finite almost surely and satisfies the stochastic integrability estimate

$$|K(y)| \leq \mathcal{O}_s(C).$$

Moreover, for every  $x \in \mathbb{Z}^d$ , conditionally on the event  $\{x, y \in \mathcal{C}_\infty\}$ , one has

$$|g(x, y) - \bar{g}(x - y) - K(y)| \leq \mathcal{O}_s(C \log^{\frac{1}{2}}(1 + |x - y|)|x - y|^{-1}).$$

REMARK 4. This statement cannot be stated with a minimal scale as in Theorems 1.1 and 1.2. This is due to the fact that the estimates scale optimally in time or space; the best possible statements involving a minimal scale are the ones of Theorems 1.1 and 1.2.

This result can be conjectured from the theory of stochastic homogenization in the uniformly elliptic setting (see [15], Theorem 9.11 and Corollary 9.12). There is one main difference between the results in the uniformly elliptic setting and in the percolation setting, which is the stochastic integrability: we expect that the stochastic integrability will be reduced by a factor  $(d - 1)/d$ . This is expected because of a surface order large deviation effect which can be heuristically explained as follows. In the uniformly elliptic setting and in a given ball  $B_R$ , to design a bad environment, that is, an environment on which no good control on the heat kernel is valid, it is necessary to have a number of ill-behaved edges of order of the volume of the ball. In the percolation setting, one can design a bad environment with a number of ill-behaved edges of order of the surface of the ball: given a ball of size  $R$ , it is possible to disconnect it into two half-balls with  $cR^{d-1}$  closed edges. This should result in a deterioration of the stochastic integrability by a factor  $(d - 1)/d$ .

The conjecture improves Theorems 1.1 and 1.2 in two distinct directions: the spatial scaling, where the coefficient  $1/2 - \delta$  is replaced by  $1/2$  for the heat kernel and the coefficient  $1 - \delta$  is replaced by 1 for the elliptic Green's function, and the stochastic integrability, where the exponent  $s$  can take any value in the interval  $(0, \frac{2(d-1)}{d})$ . We believe that the two improvements should follow from different techniques: for the spatial integrability, we think that it should follow by an adaptation of the techniques developed in [15, Chapter 9]. The improvement of the stochastic integrability seems to be a much harder problem which requires separate considerations and should rely on a precise understanding of the geometry of the percolation clusters.

We complete this section by mentioning that the results of this article pertain to the variable speed random walk, but similar results, with similar proofs, should hold for other related models of random walk on the infinite cluster such as the constant speed random walk and the simple random walk. This choice is motivated by the fact that the generator of the VSRW, written in (1), is more convenient to work with than the ones of the CSRW and the SRW, which simplifies the analysis.

1.5. *Organization of the article.* The rest of this article is organized as follows. The remaining section of this [introduction](#) is devoted to the presentation of some useful notation.

In Section 2, we record some preliminary results, including some results from the theory of quantitative stochastic homogenization on the infinite cluster from [12, 40, 61]: the quantitative sublinearity of the corrector and a quantitative estimate to control the  $H^{-1}$ -norm of the centered flux. In Section 3, we recall the Gaussian bounds on the heat kernel which were established by Barlow in [21] and establish a large-scale  $C^{0,1}$ -regularity theory for the heat kernel.

In Section 4, we establish Theorem 1.1. The proof is organized in three subsections: Section 4.1 is devoted to the proofs of three regularization steps, which can be seen as a preparation for the two-scale expansion in Section 4.2. The heart of the proof is Section 4.2, where we perform the two-scale expansion. In Section 4.3, we post-process the result from Section 4.2 and deduce the result of Theorem 1.1.

In Section 5, we use Theorem 1.1 to prove the homogenization of the elliptic Green function, that is, Theorem 1.2.

Appendix A and Appendix B are devoted respectively to two technical estimates: a concentration inequality for the density of the infinite cluster in a cube and the proof of the quantitative estimate of the weak  $H^{-1}$ -norm of the centered flux which is stated in Section 2.3.

1.6. *Notation and assumptions.*

1.6.1. *General notation and assumptions.* We let  $\mathbb{Z}^d$  be the standard  $d$ -dimensional hypercubic lattice and  $\mathcal{B}_d := \{ \{x, y\} : x, y \in \mathbb{Z}^d, |x - y| = 1 \}$  denote the set of bonds. We also denote by  $\vec{\mathcal{B}}_d$  the set of *oriented* bonds, or *edges*, of  $\mathbb{Z}^d$ . We use the notation  $\{x, y\}$  to refer to a bond and  $(x, y)$  to refer to an edge.

We denote the canonical basis of  $\mathbb{R}^d$  by  $\{e_1, \dots, e_d\}$ . For a vector  $p \in \mathbb{R}^d$  and an integer  $i \in \{1, \dots, d\}$ , we denote by  $[p]_i$  its  $i$ th-component, that is,  $p = ([p]_1, \dots, [p]_d)$ . For  $x, y \in \mathbb{Z}^d$ , we write  $x \sim y$  if  $x$  and  $y$  are nearest neighbors. We usually denote a generic edge by  $e$ . We fix an ellipticity parameter  $\lambda \in (0, 1]$  and denote by  $\Omega$  the set of all functions  $\mathbf{a} : \mathcal{B}_d \rightarrow \{0\} \cup [\lambda, 1]$ , that is,  $\Omega = (\{0\} \cup [\lambda, 1])^{\mathcal{B}_d}$  and we denote by  $\mathbf{a}$  a generic element of  $\Omega$ . The Borel  $\sigma$ -algebra on  $\Omega$  is denoted by  $\mathcal{F}$ . For each  $U \subseteq \mathbb{Z}^d$ , we let  $\mathcal{F}(U) \subseteq \mathcal{F}$  denote  $\sigma$ -algebra generated by the projections  $\mathbf{a} \mapsto \mathbf{a}(\{x, y\})$ , for  $x, y \in U$  with  $x \sim y$ .

We fix an i.i.d. probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$ , that is, a measure of the form  $\mathbb{P} = \mathbb{P}_0^{\mathcal{B}_d}$  where  $\mathbb{P}_0$  is a measure of probability supported in the set  $\{0\} \cup [\lambda, 1]$  with the property that, for any fixed bond  $e$ ,

$$\mathfrak{p} := \mathbb{P}_0[\mathbf{a}(e) \neq 0] > \mathfrak{p}_c(d),$$

where  $\mathfrak{p}_c(d)$  is the bond percolation threshold for the lattice  $\mathbb{Z}^d$ . We say that a bond  $e$  is *open* if  $\mathbf{a}(e) \neq 0$  and *closed* if  $\mathbf{a}(e) = 0$ . A connected component of open edges is called a *cluster*. Under the assumption  $\mathfrak{p} > \mathfrak{p}_c(d)$ , there exists almost surely a unique maximal infinite cluster, which is denoted by  $\mathcal{C}_\infty$  and we also note  $\theta(p) := \mathbb{P}[0 \in \mathcal{C}_\infty]$ . From now on, we always consider environments  $\mathbf{a} \in \Omega$  such that there exists a unique infinite cluster of open edges. We denote by  $\mathbb{E}$  the expectation with respect to the measure  $\mathbb{P}$ .

1.6.2. *Notation of  $\mathcal{O}_s$ .* For a parameter  $s > 0$ , we use the notation  $\mathcal{O}_s$  to measure the stochastic integrability of random variables. It is defined as follows, given a random variable  $X$ , we write

$$(22) \quad X \leq \mathcal{O}_s(\theta) \quad \text{if and only if} \quad \mathbb{E}[\exp((\theta^{-1} X)_+^s)] \leq 2,$$

where  $(\theta^{-1}X)_+$  means  $\max\{\theta^{-1}X, 0\}$ . From the inequality (22) and the Markov's inequality, one deduces the following estimate for the tail of the random variable  $X$ : for all  $x > 0$ ,  $\mathbb{P}[X \geq \theta x] \leq 2 \exp(-x^s)$ .

Given a random variable  $X$  satisfying the identity  $X \leq \mathcal{O}_s(\theta)$ , one can check that, for each  $\lambda \in \mathbb{R}^+$ , one has  $\lambda X \leq \mathcal{O}_s(\lambda\theta)$ . Additionally, one can reduce the stochastic integrability parameter  $s$  according to the following statement: for each  $s' \in (0, s]$ , there exists a constant  $0 < C_{s'} < \infty$  such that  $X \leq \mathcal{O}_{s'}(C_{s'}\theta)$ .

To estimate the stochastic integrability of a sum of random variables, we use the following estimate, which can be found in Lemma A.4 of [15]: for each exponent  $s > 0$  there exists a positive constant  $C_s < \infty$  such that for any measure space  $(E, \mathcal{S}, m)$  and any family of random variables  $\{X(z)\}_{z \in E}$ , one has

$$(23) \quad \forall z \in E, \quad X(z) \leq \mathcal{O}_s(\theta(z)) \implies \int_E X(z)m(dz) \leq \mathcal{O}_s\left(C_s \int_E \theta(z)m(dz)\right).$$

The previous statement allows to estimate the stochastic integrability of a sum of random variables: given  $X_1, \dots, X_n$  a collection of nonnegative random variables and  $C_1, \dots, C_n$  a collection of nonnegative constants such that, for any  $i \in \{1, \dots, n\}$ ,  $X_i \leq \mathcal{O}_s(C_i)$ , one has the estimate

$$(24) \quad \sum_{i=1}^n X_i \leq \mathcal{O}_s\left(C_s \sum_{i=1}^n C_i\right).$$

The following lemma is useful to construct minimal scales.

LEMMA 1.4 ([12], Lemme 2.2). *Fix  $K \geq 1, s > 0$  and  $\beta > 0$  and let  $\{X_n\}_{n \in \mathbb{N}}$  be a sequence of nonnegative random variables satisfying the inequality  $X_n \leq \mathcal{O}_s(K3^{-n\beta})$  for every  $n \in \mathbb{N}$ . There exists a positive constant  $C(s, \beta, K) < \infty$  such that the random scale  $M := \sup\{3^n \in \mathbb{N} : X_n \geq 1\}$  satisfies the stochastic integrability estimate  $M \leq \mathcal{O}_{\beta s}(C)$ .*

1.6.3. *Topology, functions and integration.* For every subset  $V \subseteq \mathbb{Z}^d$  and every environment  $\mathbf{a} \in \Omega$ , we consider two sets of bonds  $\mathcal{B}_d(V)$  and  $\mathcal{B}_d^{\mathbf{a}}(V)$ . The first one is inherited from the set of bonds  $\mathcal{B}_d$  of  $\mathbb{Z}^d$ , the second one is inherited from the bonds of nonzero conductance of the environment  $\mathbf{a}$ . They are defined by the formulas

$$\mathcal{B}_d(V) := \{\{x, y\} : x, y \in V, x \sim y\}$$

and

$$\mathcal{B}_d^{\mathbf{a}}(V) := \{\{x, y\} : x, y \in V, x \sim y, \mathbf{a}(\{x, y\}) \neq 0\}.$$

We similarly define the set of edges  $\vec{\mathcal{B}}_d(V)$  and  $\vec{\mathcal{B}}_d^{\mathbf{a}}(V)$ .

The interiors of a set  $V$  with respect to  $\mathcal{B}_d(V)$  and  $\mathcal{B}_d^{\mathbf{a}}(V)$  are defined by the formulas

$$\text{int}(V) := \{x \in V : y \sim x \implies y \in V\}$$

and

$$\text{int}_{\mathbf{a}}(V) := \{x \in V : y \sim x, \mathbf{a}(\{x, y\}) \neq 0 \implies y \in V\}.$$

The boundaries of  $V$  are defined by  $\partial V := V \setminus \text{int}(V)$  and  $\partial_{\mathbf{a}} V := V \setminus \text{int}_{\mathbf{a}}(V)$ . The cardinality of a subset  $V \subseteq \mathbb{Z}^d$  is denoted by  $|V|$  and called the volume of  $V$ . Given two sets  $U, V \subseteq \mathbb{Z}^d$ , we define the distance between  $U$  and  $V$  according to the formula  $\text{dist}(U, V) := \min_{x \in U, y \in V} |x - y|$  and the distance of a point  $x \in \mathbb{Z}^d$  to a set  $V \subseteq \mathbb{Z}^d$  by the notation  $\text{dist}(x, V) := \min_{y \in V} |x - y|$ .

For a subset  $V \subseteq \mathbb{Z}^d$ , the spaces of functions with zero boundary condition are defined by

$$(25) \quad C_0(V) := \{v : V \rightarrow \mathbb{R} : v = 0 \text{ on } \partial V\}, \quad C_0^{\mathbf{a}}(V) := \{v : V \rightarrow \mathbb{R} : v = 0 \text{ on } \partial_{\mathbf{a}} V\}.$$

Given a subset  $U \subseteq \mathbb{Z}^d$  and a function  $u : \mathcal{C}_\infty \cap U \rightarrow \mathbb{R}$  (resp. a function  $F : \mathcal{B}_d^{\mathbf{a}}(\mathcal{C}_\infty \cap U) \rightarrow \mathbb{R}$ ), the integration over the set  $\mathcal{C}_\infty \cap U$  (resp., over  $\mathcal{B}_d^{\mathbf{a}}(\mathcal{C}_\infty \cap U)$ ) is denoted by

$$(26) \quad \int_{\mathcal{C}_\infty \cap U} u := \sum_{x \in \mathcal{C}_\infty \cap U} u(x), \quad \text{resp.} \quad \int_{\mathcal{C}_\infty \cap U} F := \sum_{e \in \mathcal{B}_d^{\mathbf{a}}(\mathcal{C}_\infty \cap U)} F(e),$$

which means that we only integrate on the vertices (resp., open bonds) of the infinite cluster  $\mathcal{C}_\infty$ . We extend this notation to the setting of vector-valued functions  $u : \mathcal{C}_\infty \cap U \rightarrow \mathbb{R}^n$ . We also let  $(u)_V := \frac{1}{|V|} \int_V u$  denote the mean of the function  $u$  over the finite subset  $V \subseteq \mathbb{Z}^d$ .

Given a subset  $V \subseteq \mathbb{Z}^d$ , a vector field is a function  $\vec{F} : \vec{\mathcal{B}}_d(V) \rightarrow \mathbb{R}$  satisfying the anti-symmetry property

$$\vec{F}(\{x, y\}) = -\vec{F}(\{y, x\}).$$

For  $y \in \mathbb{Z}^d$  and  $r > 0$ , we denote by  $B_r(y)$  the discrete Euclidean ball of radius  $r > 0$  and center  $y$ ; we often write  $B_r$  in place of  $B_r(0)$ . A *cube*  $Q$  is a subset of  $\mathbb{Z}^d$  of the form

$$Q := \mathbb{Z}^d \cap (z + [-N, N]^d).$$

We define the *center* and the *size* of the cube given in the previous display above to be the point  $z$  and the integer  $N$ , respectively. The size of the cube  $Q$  is denoted by  $\text{size}(Q)$ . Given an integer  $n \in \mathbb{N}$ , we use the nonstandard convention of denoting by  $nQ$  the cube

$$(27) \quad nQ := \mathbb{Z}^d \cap (z + [-nN, nN]^d).$$

A *triadic cube* is a cube of the form

$$\square_m(z) := z + \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d, \quad z \in 3^m \mathbb{Z}^d, m \in \mathbb{N}.$$

We usually write  $\square_m := \square_m(0)$ . Additionally, we note that  $\text{size}(\square_m) = 3^m$ , denote by  $\mathcal{T}_m := \{z + \square_m : z \in 3^m \mathbb{Z}^d\}$  the set of triadic cubes of size  $3^m$  and by  $\mathcal{T}$  the set of all triadic cubes, that is,  $\mathcal{T} := \bigcup_{m \in \mathbb{N}} \mathcal{T}_m$ .

**1.6.4. Discrete analysis and function spaces.** In this article, we consider two types of objects: functions defined in the continuous space  $\mathbb{R}^d$  and functions defined on the discrete space  $\mathbb{Z}^d$ .

*Notation for discrete functions.* Given a discrete subset  $U \subseteq \mathbb{Z}^d$ , an environment  $\mathbf{a}$  such that there exists an infinite cluster  $\mathcal{C}_\infty$  of open edges, and a function  $u : \mathcal{C}_\infty \cap U \rightarrow \mathbb{R}$ , we define its gradient  $\nabla u$  to be the vector field defined on  $\vec{\mathcal{B}}_d$  by, for each edge  $e = (x, y) \in \vec{\mathcal{B}}_d$ ,

$$(28) \quad \nabla u(e) := \begin{cases} u(y) - u(x) & \text{if } x, y \in \mathcal{C}_\infty \text{ and } \mathbf{a}(\{x, y\}) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We also define the norm of the gradient  $\nabla u$  according to the formula, for each  $x \in \mathcal{C}_\infty$ ,  $|\nabla u|(x) := \sum_{y \sim x} |u(y) - u(x)|$ . We frequently abuse notation and write  $|\nabla u(x)|$  instead of  $|\nabla u|(x)$ .

For a vector field  $\vec{F} : \vec{\mathcal{B}}_d \rightarrow \mathbb{R}$ , we define the discrete divergence operator according to the formula, for each  $x \in \mathbb{Z}^d$ ,

$$\nabla \cdot \vec{F}(x) := \sum_{y \sim x} \vec{F}(x, y).$$

By the discrete integration by parts, one has, for any discrete set  $U \subseteq \mathbb{Z}^d$ , any functions  $v \in C_0^{\mathbf{a}}(\mathcal{C}_\infty \cap U)$  and  $u : \mathcal{C}_\infty \cap U \rightarrow \mathbb{R}$ ,

$$(29) \quad \int_{\mathcal{C}_\infty \cap U} \nabla v \cdot \mathbf{a} \nabla u := \int_{\mathcal{C}_\infty \cap U} \mathbf{a}(e) \nabla v(e) \nabla u(e) = \int_{\mathcal{C}_\infty \cap U} v(-\nabla \cdot \mathbf{a} \nabla u),$$

where the finite difference elliptic operator  $-\nabla \cdot \mathbf{a} \nabla$  is defined in (1).

For  $p \in [1, \infty)$ , we define the  $L^p(\mathcal{C}_\infty \cap U)$ -norm and the normalized  $\underline{L}^p(\mathcal{C}_\infty \cap U)$ -norm by the formulas

$$(30) \quad \begin{aligned} \|u\|_{L^p(\mathcal{C}_\infty \cap U)} &:= \left( \int_{\mathcal{C}_\infty \cap U} |u|^p \right)^{\frac{1}{p}}, \\ \|u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)} &:= \left( \frac{1}{|\mathcal{C}_\infty \cap U|} \int_{\mathcal{C}_\infty \cap U} |u|^p \right)^{\frac{1}{p}}. \end{aligned}$$

We also define the  $L^p(\mathcal{C}_\infty \cap U)$ -norm and the normalized  $\underline{L}^p(\mathcal{C}_\infty \cap U)$ -norm of the gradient of a function  $u : \mathcal{C}_\infty \cap U \rightarrow \mathbb{R}$  by the formulas

$$(31) \quad \begin{aligned} \|\nabla u\|_{L^p(\mathcal{C}_\infty \cap U)} &:= \left( \int_{\mathcal{C}_\infty \cap U} |\nabla u|^p \right)^{\frac{1}{p}}, \\ \|\nabla u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)} &:= \left( \frac{1}{|\mathcal{C}_\infty \cap U|} \int_{\mathcal{C}_\infty \cap U} |\nabla u|^p \right)^{\frac{1}{p}}. \end{aligned}$$

We define the normalized discrete Sobolev norm  $\underline{W}^{1,p}(\mathcal{C}_\infty \cap U)$  by

$$(32) \quad \|u\|_{\underline{W}^{1,p}(\mathcal{C}_\infty \cap U)} := |\mathcal{C}_\infty \cap U|^{-\frac{1}{d}} \|u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)} + \|\nabla u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)},$$

and the dual norm  $\underline{W}^{-1,p}(\mathcal{C}_\infty \cap U)$ ,

$$(33) \quad \|u\|_{\underline{W}^{-1,p}(\mathcal{C}_\infty \cap U)} := \sup_{v \in C_0^{\mathbf{a}}(\mathcal{C}_\infty \cap U), \|v\|_{\underline{W}^{1,p'}(\mathcal{C}_\infty \cap U)} \leq 1} \frac{1}{|\mathcal{C}_\infty \cap U|} \int_{\mathcal{C}_\infty \cap U} uv,$$

with  $\frac{1}{p'} + \frac{1}{p} = 1$ . We use respectively the notation  $\underline{H}^1(\mathcal{C}_\infty \cap U) := \underline{W}^{1,2}(\mathcal{C}_\infty \cap U)$  and  $\underline{H}^{-1}(\mathcal{C}_\infty \cap U) := \underline{W}^{-1,2}(\mathcal{C}_\infty \cap U)$ .

For a function  $u : \mathbb{Z}^d \rightarrow \mathbb{R}$  and a vector  $h \in \mathbb{Z}^d$ , we denote by  $T_h(u) := u(\cdot + h)$  the translation and by  $\mathcal{D}_{e_k}$  the finite difference operator defined by, for any function  $u : \mathbb{Z}^d \rightarrow \mathbb{R}$ ,

$$\mathcal{D}_{e_k} u := \begin{cases} \mathbb{Z}^d \rightarrow \mathbb{R}, \\ x \mapsto T_{e_k}(u)(x) - u(x). \end{cases}$$

We also define the vector-valued finite difference operator  $\mathcal{D}u := (\mathcal{D}_{e_1}u, \mathcal{D}_{e_2}u, \dots, \mathcal{D}_{e_d}u)$ . This definition has two main differences with the gradient on graph defined in (28): it is defined on the vertices of  $\mathbb{Z}^d$  (not on the edges) and it is vector-valued. This second definition of discrete derivative is introduced because it is convenient in the two-scale expansion (see (106)).

Given an environment  $\mathbf{a} \in \Omega$ , and a function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$ , we define the functions  $\mathbf{a} \mathcal{D}u$  and  $\mathbb{1}_{\{\mathbf{a} \neq 0\}} \mathcal{D}u$  by, for each  $x \in \mathcal{C}_\infty$ ,

$$\begin{aligned} \mathbf{a} \mathcal{D}u(x) &:= (\mathbf{a}(\{x, x + e_1\}) \mathcal{D}_{e_1}u(x), \dots, \mathbf{a}(\{x, x + e_d\}) \mathcal{D}_{e_d}u(x)), \\ \mathbb{1}_{\{\mathbf{a} \neq 0\}} \mathcal{D}u(x) &:= (\mathbb{1}_{\{\mathbf{a}(\{x, x + e_1\}) \neq 0\}} \mathcal{D}_{e_1}u(x), \dots, \mathbb{1}_{\{\mathbf{a}(\{x, x + e_d\}) \neq 0\}} \mathcal{D}_{e_d}u(x)). \end{aligned}$$

We extend these functions to the entire space  $\mathbb{Z}^d$  by setting, for each point  $x \in \mathbb{Z}^d \setminus \mathcal{C}_\infty$ ,  $\mathbf{a} \mathcal{D}u(x) = \mathbb{1}_{\{\mathbf{a} \neq 0\}} \mathcal{D}u(x) = 0$ . It is natural to introduce the dual operator  $\mathcal{D}_{e_k}^* u := T_{-e_k}(u) - u$

and the divergence  $\mathcal{D}^*$  defined by, for any vector-valued function  $\tilde{F} : \mathbb{Z}^d \rightarrow \mathbb{R}^d$ ,  $\tilde{F} = (\tilde{F}_1, \tilde{F}_2, \dots, \tilde{F}_d)$ ,

$$\mathcal{D}^* \cdot \tilde{F}(x) = \sum_{k=1}^d \mathcal{D}_{e_k}^* \tilde{F}_k(x).$$

By the discrete integration by parts, one has the equality, for any  $v \in C_0^{\mathbf{a}}(\mathcal{C}_\infty \cap U)$ ,

$$(34) \quad \int_{\mathcal{C}_\infty \cap U} v(\mathcal{D}^* \cdot \mathbf{a} \mathcal{D}u) = \int_{\mathcal{C}_\infty \cap U} \mathcal{D}v \cdot \mathbf{a} \mathcal{D}u.$$

In fact, one can check that the identity  $-\nabla \cdot \mathbf{a} \nabla = \mathcal{D}^* \cdot \mathbf{a} \mathcal{D}$  holds, which allows to interchange the two notation.

Moreover, given a vector  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ , we denote by  $|x| = (\sum_{j=1}^d x_j^2)^{\frac{1}{2}}$  its norm. This allows to extend the definition of the Sobolev norms (30), (32) and (33) to vector-valued functions, and we note that

$$\begin{aligned} c \|u\|_{\underline{W}^{1,p}(\mathcal{C}_\infty \cap U)} &\leq |\mathcal{C}_\infty \cap U|^{-\frac{1}{d}} \|u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)} + \|\mathbb{1}_{\{\mathbf{a} \neq 0\}} \mathcal{D}u\|_{\underline{L}^p(\mathcal{C}_\infty \cap U)} \\ &\leq C \|u\|_{\underline{W}^{1,p}(\mathcal{C}_\infty \cap U)}, \end{aligned}$$

for some constants  $c, C$  which only depend on the dimension  $d$ .

*Notation for continuous functions.* We use the notation  $\partial_k, \nabla, \Delta$  for the standard derivative, gradient and Laplacian on  $\mathbb{R}^d$ , which are only applied to smooth functions. It will always be clear from context whether we refer to the continuous or discrete derivatives. We sometime slightly abuse the notation and denote by  $|\nabla^k \eta|$  the norm of  $k$ th derivatives of the function  $\eta$ .

*Notation for parabolic functions.* For  $r > 0$ , we define the time interval  $I_r := (-r^2, 0]$  and  $I_r(t) := t + (-r^2, 0]$ . We frequently use the parabolic cylinders  $I_r \times B_r$  and  $I_r \times (\mathcal{C}_\infty \cap B_r)$  and define their volumes by

$$|(I_r \times B_r)| = r^2 \times |B_r| \quad \text{and} \quad |(I_r \times (\mathcal{C}_\infty \cap B_r))| = r^2 \times |\mathcal{C}_\infty \cap B_r|.$$

Given a function  $u : I_r \times B_r \rightarrow \mathbb{R}$  (resp.,  $v : I_r \times (\mathcal{C}_\infty \cap B_r) \rightarrow \mathbb{R}$ ), we define the integrals

$$\int_{I_r \times B_r} u := \int_{-r^2}^0 \int_{B_r} u(t, x) dx dt \quad \text{and} \quad \int_{I_r \times (\mathcal{C}_\infty \cap B_r)} v := \int_{-r^2}^0 \int_{\mathcal{C}_\infty \cap B_r} v(t, x) dx dt,$$

and denote the mean of these functions by the notation

$$\begin{aligned} (u)_{I_r \times B_r} &:= \frac{1}{|(I_r \times B_r)|} \int_{-r^2}^0 \int_{B_r} u(t, x) dx dt, \\ (v)_{I_r \times (\mathcal{C}_\infty \cap B_r)} &:= \frac{1}{|(I_r \times (\mathcal{C}_\infty \cap B_r))|} \int_{-r^2}^0 \int_{\mathcal{C}_\infty \cap B_r} v(t, x) dx dt. \end{aligned}$$

Given a finite subset  $V \subseteq \mathbb{Z}^d$  or  $V \subseteq \mathbb{R}^d$ , we denote by  $\partial_\sqcup(I_r \times V)$  the parabolic boundary of the cylinder  $I_r \times V$  defined by the formula

$$\partial_\sqcup(I_r \times V) := (I_r \times \partial V) \cup (\{-r^2\} \times V).$$

Given a real number  $p \geq 1$  and a Lebesgue-measurable function  $u : I_r \times V \rightarrow \mathbb{R}^d$ , we define the norm  $L^p(I_r \times V)$  and the normalized norm  $\underline{L}^p(I_r \times V)$  according to the formulas

$$\|u\|_{L^p(I_r \times V)} := \left( \int_{-r^2}^0 \|u(t, \cdot)\|_{L^p(V)}^p dt \right)^{\frac{1}{p}}$$

and

$$\|u\|_{\underline{L}^p(I_r \times V)} := \left( r^{-2} \int_{-r^2}^0 \|u(t, \cdot)\|_{\underline{L}^p(V)}^p dt \right)^{\frac{1}{p}}.$$

These notation are extended to the gradient of a function  $u : I_r \times V \rightarrow \mathbb{R}^d$  by the formulas

$$\|\nabla u\|_{L^p(I_r \times V)} := \left( \int_{-r^2}^0 \|\nabla u(t, \cdot)\|_{L^p(V)}^p dt \right)^{\frac{1}{p}}$$

and

$$\|\nabla u\|_{\underline{L}^p(I_r \times V)} := \left( r^{-2} \int_{-r^2}^0 \|\nabla u(t, \cdot)\|_{\underline{L}^p(V)}^p dt \right)^{\frac{1}{p}}.$$

Given a real number  $q \geq 1$ , we also define the space  $L^q(I_r; W^{-1,p}(V))$  by

$$L^q(I_r; W^{-1,p}(V)) := \left\{ u : I_r \times V \rightarrow \mathbb{R}^n : \int_{-r^2}^0 \|u\|_{\underline{W}^{-1,p}(V)}^q dt < \infty \right\},$$

and we equip this space with the normalized norm defined by

$$\|u\|_{\underline{L}^q(I_r; \underline{W}^{-1,p}(V))} := \left( r^{-2} \int_{-r^2}^0 \|u(t, \cdot)\|_{\underline{W}^{-1,p}(V)}^q dt \right)^{\frac{1}{q}}.$$

We define the parabolic Sobolev space  $W_{\text{par}}^{1,p}(I_r \times V)$  to be the set of measurable functions  $u : I_r \times V \rightarrow \mathbb{R}$  such that the time derivative  $\partial_t u$ , understood in the sense of distributions, belongs to the space  $W^{-1,p'}(I_r \times V)$  with  $\frac{1}{p} + \frac{1}{p'} = 1$ , that is,

$$W_{\text{par}}^{1,p}(I_r \times V) := \{u \in L^p(I_r \times V) : \partial_t u \in L^{p'}(I_r; W^{-1,p'}(V))\}.$$

We also make use of the notation  $H_{\text{par}}^1(I_r \times V) := W_{\text{par}}^{1,2}(I_r \times V)$  for the  $H^1$  parabolic space and  $\underline{L}^2(I_r; \underline{H}^{-1}(V)) := \underline{L}^2(I_r; \underline{W}^{-1,2}(V))$ .

1.6.5. *Convention for constants, exponents and minimal scales/times.* Throughout this article, the symbols  $c$  and  $C$  denote positive constants which may vary from line to line. These constants may depend only on the dimension  $d$ , the ellipticity  $\lambda$  and the probability  $\mathfrak{p}$ . Similarly, we use the symbols  $\alpha, \beta, \gamma, \delta$  to denote positive exponents which depend only on  $d, \lambda$  and  $\mathfrak{p}$ . Usually, we use the letter  $C$  for large constants (whose value is expected to belong to  $[1, \infty)$ ) and  $c$  for small constants (whose value is expected to be in  $(0, 1]$ ). The values of the exponents  $\alpha, \beta, \gamma, \delta$  are always expected to be small. When the constants and exponents depend on other parameters, we write it explicitly and use the notation  $C := C(d, \mathfrak{p}, t)$  to mean that the constant  $C$  depends on the parameters  $d, \mathfrak{p}$  and  $t$ . We also assume that all the minimal scales and times which appear in this article are larger than 1.

**2. Preliminaries.** In this section, we collect a few results from the theory of supercritical percolation which are important tools in the establishment of Theorems 1.1 and 1.2.

2.1. *Supercritical percolation.*

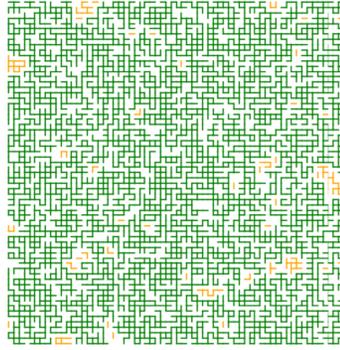


FIG. 2. A pre-good cube, the cluster  $\mathcal{C}_*(\square)$  is drawn in green and the clusters in yellow are the small clusters.

2.1.1. *A partition of good cubes.* An important step to prove results on the behavior of the random walk on the infinite cluster  $\mathcal{C}_\infty$  consists in understanding the geometry of this cluster. A general picture to keep in mind is that the geometry of  $\mathcal{C}_\infty$  is similar, at least on large scales, to the one of the Euclidean lattice  $\mathbb{Z}^d$ . To give a precise mathematical meaning to this statement, the common strategy is to implement a renormalization structure for the infinite cluster. In this article, we use a strategy, which was first introduced by Armstrong and the first author in [12]. It relies on the following geometric definition and lemma which are due to Penrose and Pisztor [75].

DEFINITION 1 (Pre-good cube). We say that a discrete cube  $\square \subseteq \mathbb{Z}^d$  of size  $N$  is *pre-good* if:

1. There exists a cluster of open edges which intersects the  $2d$  faces of the cube  $\square$ . This cluster is denoted by  $\mathcal{C}_*(\square)$ ;
2. The diameter of all the other clusters is smaller than  $\frac{N}{10}$ .

Figure 2 is a simulation of a pre-good cube. We then upgrade this definition into the following definition of good cubes.

DEFINITION 2 (Good cube). We say that a discrete cube  $\square \subseteq \mathbb{Z}^d$  of size  $N$  is *good* if:

1. The cube  $\square$  is pre-good;
2. Every cube  $\square'$  whose size is between  $N/10$  and  $N$  and which has nonempty intersection with  $\square$  is also pre-good.

We note that the event “the cube  $\square$  is good” is  $\mathcal{F}(3\square)$ -measurable. The main reason to use good cubes instead of pre-good cubes is that they satisfy the following connectivity property, which can be obtained from straightforward geometric considerations and whose proof can be found in [12], Lemma 2.8.

LEMMA 2.1 (Connectivity property). Let  $\square_1, \square_2$  be two cubes of  $\mathbb{Z}^d$  which are neighbors, that is, which satisfy

$$\text{dist}(\square_1, \square_2) \leq 1,$$

which have comparable size in the sense that

$$\frac{1}{3} \leq \frac{\text{size}(\square_1)}{\text{size}(\square_2)} \leq 3,$$

and which are both good. Then there exists a cluster  $\mathcal{C}$  such that

$$\mathcal{C}_*(\square_1) \cup \mathcal{C}_*(\square_2) \subseteq \mathcal{C} \subseteq \square_1 \cup \square_2.$$

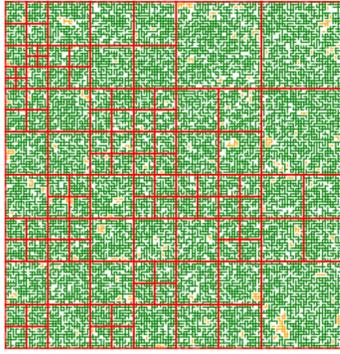


FIG. 3. A realization of the partition  $\mathcal{P}$ , where the cluster in green is the maximal cluster and the cubes in red are elements of  $\mathcal{P}$ .

The main interest in these definitions is that in the supercritical phase  $p > p_c(d)$ , the probability of a cube to be good is exponentially close to 1 in the size of the cube. Such a result is stated in the following proposition and is a direct consequence [76], Theorem 3.2 and [75], Theorem 5.

PROPOSITION 1. Consider a Bernoulli bond percolation of probability  $p \in (p_c(d), 1]$ . Then there exists a positive constant  $C(d, p) < \infty$  such that, for every cube  $\square \subseteq \mathbb{Z}^d$  of size  $N$ ,

$$(35) \quad \mathbb{P}[\square \text{ is good}] \geq 1 - C \exp(-C^{-1}N).$$

The renormalization structure we want to implement relies on the observation that  $\mathbb{Z}^d$  can be partitioned into good cubes of varying sizes. Thanks to the exponential stochastic integrability obtained by Penrose and Pisztora and stated in Proposition 1, we are able to build such a partition. The precise statement is given in the following proposition.

PROPOSITION 2 (Propositions 2.1 and 2.4 of [12]). Under the assumption  $p > p_c$ ,  $\mathbb{P}$ -almost surely, there exists a partition  $\mathcal{P}$  of  $\mathbb{Z}^d$  into triadic cubes with the following properties:

1. All the predecessors of elements of  $\mathcal{P}$  are good cubes, that is, for every pair of triadic cubes  $\square, \square' \in \mathcal{T}$ , one has the property

$$\square' \in \mathcal{P} \quad \text{and} \quad \square' \subseteq \square \quad \implies \quad \square \text{ is good.}$$

2. Neighboring elements of  $\mathcal{P}$  have comparable sizes: for every  $\square, \square' \in \mathcal{P}$  such that  $\text{dist}(\square, \square') \leq 1$ , we have

$$\frac{1}{3} \leq \frac{\text{size}(\square')}{\text{size}(\square)} \leq 3.$$

3. Estimate for the coarseness of  $\mathcal{P}$ : if we denote by  $\square_{\mathcal{P}}(x)$  the unique element of  $\mathcal{P}$  containing a given point  $x \in \mathbb{Z}^d$ , then there exists a constant  $C(p, d) < \infty$  such that

$$(36) \quad \text{size}(\square_{\mathcal{P}}(x)) \leq \mathcal{O}_1(C).$$

4. Minimal scale for  $\mathcal{P}$ . For each  $q \in [1, \infty)$ , there exists a constant  $C := C(d, p, q) < \infty$ , a nonnegative random variable  $\mathcal{M}_q(\mathcal{P})$  and an exponent  $r := r(d, p, q) > 0$  such that

$$(37) \quad \mathcal{M}_q(\mathcal{P}) \leq \mathcal{O}_r(C),$$

and for each radius  $R$  satisfying  $R \geq \mathcal{M}_q(\mathcal{P})$ ,

$$(38) \quad R^{-d} \sum_{x \in \mathbb{Z}^d \cap B_R} \text{size}(\square_{\mathcal{P}}(x))^q \leq C \quad \text{and} \quad \sup_{x \in \mathbb{Z}^d \cap B_R} \text{size}(\square_{\mathcal{P}}(x)) \leq R^{\frac{1}{q}}.$$

REMARK 5. To be precise, this proposition is a consequence of Propositions 2.1 and 2.4 of [12] and of Proposition 1 as is explained in [12], Section 2.2.

Additionally, the precise result stated in [12], Proposition 2.4, is that there exists a minimal scale  $\mathcal{M}_q(\mathcal{P})$  above which  $\sup_{x \in \mathbb{Z}^d \cap B_R} \text{size}(\square_{\mathcal{P}}(x)) \leq R^{\frac{d}{d+q}}$ , with an exponent  $d/(d+q)$  instead of  $1/q$ . Nevertheless, it is straightforward to recover the statement of Proposition 2 from the one of [12], Proposition 2.4.

Figure 3 (drawn with dyadic cubes instead of triadic cubes to improve readability) illustrates what this partition looks like. It allows to extend functions defined on the infinite cluster to the whole space  $\mathbb{R}^d$ , as is explained below. We consider a function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$ . For each point  $x \in \mathbb{Z}^d$ , we choose a point  $z(x)$  in the cluster  $\mathcal{C}_*(\square_{\mathcal{P}}(x))$  according to some deterministic procedure (for instance, we choose the one which is the closest to the center of the cube and break ties by using the lexicographical order). We then define the coarsened function  $[u]_{\mathcal{P}}$  on  $\mathbb{Z}^d$  according to the formula, for each  $x \in \mathbb{Z}^d$ ,

$$(39) \quad [u]_{\mathcal{P}}(x) := \begin{cases} u(x) & \text{if } x \in \mathcal{C}_\infty, \\ u(z(x)) & \text{otherwise,} \end{cases}$$

and extend it to the whole space  $\mathbb{R}^d$  by setting it to be piecewise constant on the cubes  $[x - \frac{1}{2}, x + \frac{1}{2}]^d$ , for  $x \in \mathbb{Z}^d$ . When the function  $u$  is defined on the parabolic space  $[0, \infty) \times \mathcal{C}_\infty$ , we define its extension to the space  $[0, \infty) \times \mathbb{Z}^d$ , which we also denote by  $[u]_{\mathcal{P}}$ , according to the formula

$$(40) \quad [u]_{\mathcal{P}}(t, x) := \begin{cases} u(t, x) & \text{if } x \in \mathcal{C}_\infty, \\ u(t, z(x)) & \text{otherwise.} \end{cases}$$

For later purposes, we note that, given a function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$ , the  $L^p$ -norm of the function  $\nabla [u]_{\mathcal{P}}$  can be estimated in terms of  $L^p$ -norm of the function  $\nabla u$  and the sizes of the cubes in the partition  $\mathcal{P}$ . Specifically, one has the formula, for any radius  $r \geq \text{size}(\square_{\mathcal{P}}(0))$ ,

$$(41) \quad \|\nabla [u]_{\mathcal{P}}\|_{L^p(\mathbb{Z}^d \cap B_r)}^p \leq C \int_{B_r \cap \mathcal{C}_\infty} \text{size}(\square_{\mathcal{P}}(x))^{pd-1} |\nabla u(x)|^p dx.$$

The proof of this result can be found in [12], Lemma 3.3. Additionally, one can estimate the  $L^p$ -norm of the function  $[u]_{\mathcal{P}}$  in terms of the  $L^p$ -norm of the function  $u$  according to the formula, for any radius  $r \geq \text{size}(\square_{\mathcal{P}}(0))$ ,

$$(42) \quad \|[u]_{\mathcal{P}}\|_{L^p(B_r)}^p \leq \int_{\hat{B}_r \cap \mathcal{C}_\infty} \text{size}(\square_{\mathcal{P}}(x))^d |u(x)|^p dx,$$

where  $\hat{B}_r$  denotes the union of all the cubes in the partition  $\mathcal{P}$  which intersect the ball  $B_r$ , that is,  $\hat{B}_r := \cup\{\square \in \mathcal{P} : \square \cap B_r \neq \emptyset\}$ . This estimate is a consequence of the following argument: by definition of the coarsened function  $[u]_{\mathcal{P}}$ , one has the estimates, for any cube  $\square$  of the partition  $\mathcal{P}$ ,

$$\begin{aligned} \|[u]_{\mathcal{P}}\|_{L^p(\square)}^p &\leq \text{size}(\square)^d \|[u]_{\mathcal{P}}\|_{L^\infty(\square)}^p \\ &\leq \text{size}(\square)^d \|u\|_{L^\infty(\mathcal{C}_\infty \cap \square)}^p \\ &\leq \text{size}(\square)^d \|u\|_{L^p(\mathcal{C}_\infty \cap \square)}^p, \end{aligned}$$

where we used the discrete  $L^\infty - L^p$ -estimate in the third inequality. Summing over all the cubes of the set  $\hat{B}_r$  completes the proof of the estimate (42).

2.2. *Functional inequalities on the infinite cluster.* In this section, we state mostly without proofs, some functional inequalities which are valid on the infinite cluster  $\mathcal{C}_\infty$ . The partition of good cubes presented in Section 2.1.1 allows to prove these estimates and we refer to [12] for the details of the argument. Some of these inequalities were already proved by other renormalization technique: it is in particular the case of the Poincaré inequality which was established by Barlow in [21] (see also Mathieu, Remy [69] and Benjamini, Mossel [27]).

The fact that these bounds are stated on a random graph which has an irregular nature means that they are only valid on balls of size larger than some random minimal scales, denoted by  $\mathcal{M}_{\text{Poinc}}$  and  $\mathcal{M}_{\text{Meyers}}$  in the following statements, which depend on the environment  $\mathbf{a}$  and are large when the environment is ill-behaved.

The first functional inequality we record is the Poincaré inequality, it can be found in [21], Theorem 2.18, for the  $L^2$ -version.

PROPOSITION 3 (Poincaré inequality on  $\mathcal{C}_\infty$ ). *Fix a real number  $p \in [\frac{d}{d-1}, \infty)$ . There exist a constant  $C := C(d, \mathbf{p}, p) < \infty$ , an exponent  $s := s(d, \mathbf{p}, p) > 0$  such that, for any  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{L^p\text{-Poinc}}(y)$  which satisfies the stochastic integrability estimate*

$$(43) \quad \mathcal{M}_{L^p\text{-Poinc}}(y) \leq \mathcal{O}_s(C),$$

such that for each radius  $R \geq \mathcal{M}_{L^p\text{-Poinc}}(y)$  and each function  $u : \mathcal{C}_\infty \cap B_R(y) \rightarrow \mathbb{R}$ ,

$$\|u - (u)_{\mathcal{C}_\infty \cap B_R(y)}\|_{\underline{L}^p(\mathcal{C}_\infty \cap B_R(y))} \leq CR \|\nabla u\|_{\underline{L}^p(\mathcal{C}_\infty \cap B_R(y))}.$$

Moreover, for each function  $u : \mathcal{C}_\infty \cap B_R(y) \rightarrow \mathbb{R}$ , such that  $u = 0$  on the boundary  $\partial(\mathbb{Z}^d \cap B_R(y)) \cap \mathcal{C}_\infty$ ,

$$\|u\|_{\underline{L}^p(\mathcal{C}_\infty \cap B_R(y))} \leq CR \|\nabla u\|_{\underline{L}^p(\mathcal{C}_\infty \cap B_R(y))}.$$

REMARK 6. This inequality is frequently used in the case  $p = 2$ . To shorten the notation, we write  $\mathcal{M}_{\text{Poinc}}(y)$  to refer to the minimal scale  $\mathcal{M}_{L^2\text{-Poinc}}(y)$ .

PROOF. By translation invariance of the model, we can always assume  $y = 0$ . The proof relies on the Sobolev inequality as stated in [12], Proposition 3.4, together with the Hölder inequality by setting  $\mathcal{M}_{L^p\text{-Poinc}}(0) := \mathcal{M}_q(\mathcal{P})$ , for a parameter  $q$  chosen large enough depending only on the dimension  $d$  and the exponent  $p$ .  $\square$

The second estimate we need to record is the parabolic Caccioppoli inequality. This estimate is valid on any subgraph of  $\mathbb{Z}^d$  and is used in Section 3.2.2.

PROPOSITION 4 (Parabolic Caccioppoli inequality on  $\mathcal{C}_\infty$ ). *There exists a positive constant  $C := C(d, \lambda) < \infty$  such that, for each point  $y \in \mathbb{Z}^d$ , each radius  $R \geq 1$  and each function  $u : I_R \times (\mathcal{C}_\infty \cap B_R(y)) \rightarrow \mathbb{R}$  which is  $\mathbf{a}$ -caloric, that is, which is a solution of the parabolic equation*

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \quad \text{in } I_R \times (\mathcal{C}_\infty \cap B_R(y)),$$

one has

$$\|\nabla u\|_{L^2(I_{R/2} \times (\mathcal{C}_\infty \cap B_{R/2}(y)))} \leq \frac{C}{R} \|u - (u)_{I_R \times (\mathcal{C}_\infty \cap B_R(y))}\|_{L^2(I_R \times (\mathcal{C}_\infty \cap B_R(y)))}.$$

PROOF. The proof follows the standard arguments of the Caccioppoli inequality; the fact that the function  $u$  is defined on the infinite cluster does not affect the proof and we omit the details.  $\square$

The third estimate we record is an  $L^\infty L^2$  gradient bound for  $\mathbf{a}$ -caloric functions. The proof of this result can be found in [15], Lemma 8.2, in the uniformly elliptic setting; the extension to the percolation cluster makes no difference in the proof.

LEMMA 2.2. *There exists a positive constant  $C := C(d, \lambda) < \infty$  such that for any radius  $R \geq 1$ , any point  $y \in \mathbb{Z}^d$  and any function  $u : I_R \times (\mathcal{C}_\infty \cap B_R(y)) \rightarrow \mathbb{R}$  which satisfies*

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \quad \text{in } I_R \times (\mathcal{C}_\infty \cap B_R(y)),$$

one has the estimate

$$\sup_{t \in I_{R/2}} \|\nabla u(t, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{R/2}(y))} \leq C \|\nabla u\|_{\underline{L}^2(I_R \times (\mathcal{C}_\infty \cap B_R(y)))}.$$

The last estimate we record in this section is the Meyers estimate for  $\mathbf{a}$ -caloric functions on the percolation cluster. This inequality is a nonconcentration estimate and essentially states that the energy of solutions of a parabolic equation cannot concentrate in small volumes. It is used in Section 3.2.2.

PROPOSITION 5 (Interior Meyers estimate on  $\mathcal{C}_\infty$ ). *There exist a positive constant  $C := C(d, \mathbf{p}, \lambda) < \infty$ , two exponents  $s := s(d, \mathbf{p}, \lambda) > 0$ ,  $\delta_0 := \delta_0(d, \mathbf{p}, \lambda) > 0$  such that, for each  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{Meyers}}(y)$  which satisfies the stochastic integrability estimate*

$$\mathcal{M}_{\text{Meyers}}(y) \leq \mathcal{O}_s(C),$$

such that, for each radius  $R \geq \mathcal{M}_{\text{Meyers}}(y)$  and each function  $u : I_R \times (\mathcal{C}_\infty \cap B_R(y)) \rightarrow \mathbb{R}$  solution of the equation

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \quad \text{in } I_R \times (\mathcal{C}_\infty \cap B_R(y)),$$

one has

$$\|\nabla u\|_{\underline{L}^{2+\delta_0}(I_{R/2} \times (\mathcal{C}_\infty \cap B_{R/2}(y)))} \leq C \|\nabla u\|_{\underline{L}^2(I_R \times (\mathcal{C}_\infty \cap B_R(y)))}.$$

PROOF. The classical proof of the interior Meyers estimate (cf. [49]) is based on an application of the Caccioppoli inequality, the Sobolev inequality and the Gehring’s lemma (cf. [52]). The proof of this result on the percolation cluster for the elliptic problem is written in [12], Proposition 3.8. For the parabolic problem considered here, the proof in the case of uniformly elliptic environments can be found in [11], Appendix B. The argument can be adapted to the percolation cluster following the strategy developed in [12], Section 3. Since the analysis does not contain any new idea regarding the method and the result can be obtained by essentially rewriting the proof, we skip the details.  $\square$

2.3. *Homogenization on percolation clusters.* In this section, we collect some results of stochastic homogenization in supercritical percolation useful in the proof of Theorem 1.1. The proof of this theorem is based on a quantitative two-scale expansion, which relies on two important functions: the first-order corrector and its flux. They are introduced in Sections 2.3.1 and 2.3.2, respectively.

2.3.1. *The first-order corrector.* We let  $\mathcal{A}_1(\mathcal{C}_\infty)$  be the random vector space of  $\mathbf{a}$ -harmonic functions on the infinite cluster  $\mathcal{C}_\infty$  with at most linear growth. This latter condition is expressed in terms of average  $L^2$ -norm and we define

$$\mathcal{A}_1(\mathcal{C}_\infty) := \left\{ u : \mathcal{C}_\infty \rightarrow \mathbb{R} : -\nabla \cdot (\mathbf{a}\nabla u) = 0 \text{ in } \mathcal{C}_\infty \text{ and } \lim_{r \rightarrow \infty} r^{-2} \|u\|_{L^2(\mathcal{C}_\infty \cap B_r)} = 0 \right\}.$$

It is known that this space is almost surely finite-dimensional and that its dimension is equal to  $(d + 1)$  (see [26]). Additionally, every function  $u \in \mathcal{A}_1(\mathcal{C}_\infty)$  can be uniquely written as

$$u(x) = c + p \cdot x + \chi_p(x),$$

where  $c \in \mathbb{R}$ ,  $p \in \mathbb{R}^d$  and  $\chi_p$  is a function called the corrector; it is defined up to a constant and satisfies the quantitative sublinearity property stated in the following proposition.

PROPOSITION 6. *For any exponent  $\alpha > 0$ , there exist an exponent  $s(d, \mathbf{p}, \lambda, \alpha) > 0$  and a positive constant  $C(d, \mathbf{p}, \lambda, \alpha) < \infty$  such that, for any point  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{corr}, \alpha}(y)$  satisfying the stochastic integrability estimate*

$$(44) \quad \mathcal{M}_{\text{corr}, \alpha}(y) \leq \mathcal{O}_s(C),$$

such that for every radius  $r \geq \mathcal{M}_{\text{corr}, \alpha}(y)$ , and every  $p \in \mathbb{R}^d$ ,

$$\text{osc}_{\mathcal{C}_\infty \cap B_r(y)} \chi_p := \left( \sup_{x \in \mathcal{C}_\infty \cap B_r(y)} \chi_p - \inf_{x \in \mathcal{C}_\infty \cap B_r(y)} \chi_p \right) \leq C|p|r^\alpha.$$

PROOF. The proof of this result relies on the optimal scaling estimates for the corrector established in [40]. Indeed by [40], Theorem 1, one has the following result: there exists a constant  $C := C(d, \lambda, \mathbf{p}) < \infty$  and an exponent  $s := s(d, \lambda, \mathbf{p}) < \infty$  such that for each  $x, y \in \mathbb{Z}^d$ , and each  $p \in \mathbb{R}^d$ ,

$$(45) \quad |\chi_p(x) - \chi_p(y)| \mathbb{1}_{\{x, y \in \mathcal{C}_\infty\}} \leq \begin{cases} \mathcal{O}_s(C|p|) & \text{if } d \geq 3, \\ \mathcal{O}_s(C|p| \log^{\frac{1}{2}} |x - y|) & \text{if } d = 2. \end{cases}$$

Proposition 6 is then a consequence of the previous estimate and an application of Lemma 1.4 with the sequence of random variables

$$X_n := 3^{-\alpha n} \sup_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} |\chi_p(x) - \chi_p(y)| \mathbb{1}_{\{x, y \in \mathcal{C}_\infty\}}.$$

To be more precise, we use the estimate (24) to control the maximum of the random variables

$$(46) \quad \begin{aligned} X_n &= 3^{-\alpha n} \sup_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} |\chi_p(x) - \chi_p(y)| \mathbb{1}_{\{x, y \in \mathcal{C}_\infty\}} \\ &\leq 3^{-\alpha n} \left( \sum_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} (|\chi_p(x) - \chi_p(y)| \mathbb{1}_{\{x, y \in \mathcal{C}_\infty\}})^{\frac{2d}{\alpha}} \right)^{\frac{\alpha}{2d}} \\ &\leq \begin{cases} \mathcal{O}_s(C|p|3^{-\frac{\alpha n}{2}}) & \text{if } d \geq 3, \\ \mathcal{O}_s(C|p|\sqrt{n}3^{-\frac{\alpha n}{2}}) & \text{if } d = 2. \end{cases} \end{aligned}$$

Then the sequence  $\{X_n\}_{n \geq 1}$  satisfies the assumption of Lemma 1.4.  $\square$

The fact that the corrector is only defined up to a constant causes some technical difficulties in the proofs, in particular the two-scale expansion stated in (106) and used in the proof of Theorem 1.1 is ill-defined in this setting. To solve this issue, we choose the following (arbitrary) normalization for the corrector: given a point  $y \in \mathbb{Z}^d$  and an environment  $\mathbf{a}$  in the set of probability 1 on which the corrector is well defined, we let  $x \in \mathcal{C}_\infty$  which is the closest to the point  $y$  (and break ties by using the lexicographical order) and normalize the corrector by setting  $\chi_p(x) = 0$ . The choice of the point  $y$  will always be explicitly indicated to avoid confusions. We note that with this normalization, the corrector is not stationary.

2.3.2. *The centered flux.* A second important notion in the implementation of the two-scale expansion is the centered flux; it is defined in the following paragraph.

For a fixed vector  $p = (p_1, \dots, p_d) \in \mathbb{R}^d$ , we consider the mapping  $\mathbf{a}(p + \mathcal{D}\chi_p) : \mathcal{C}_\infty \rightarrow \mathbb{R}^d$  defined by the formula, for each  $x \in \mathbb{Z}^d$ ,

$$\mathbf{a}(p + \mathcal{D}\chi_p)(x) := (\mathbf{a}(\{x, x + e_1\})(p_1 + \mathcal{D}_{e_1}\chi_p(x)), \dots, \mathbf{a}(\{x, x + e_d\})(p_d + \mathcal{D}_{e_d}\chi_p(x))).$$

This function oscillates quickly but it is close to the deterministic slope  $\frac{1}{2}\bar{\sigma}^2 p$  in the  $\underline{H}^{-1}$ -norm on the infinite cluster, where  $\bar{\sigma}$  is the diffusivity of the random walk introduced in (4). This motivates the following definition: for a fixed vector  $p \in \mathbb{R}^d$ , we define the centered flux  $\tilde{\mathbf{g}}_p : \mathcal{C}_\infty \rightarrow \mathbb{R}^d$  according to the formula

$$\tilde{\mathbf{g}}_p := \mathbf{a}(\mathcal{D}\chi_p + p) - \frac{1}{2}\bar{\sigma}^2 p.$$

The following proposition estimates the  $\underline{H}^{-1}$ -norm of the centered flux. It is proved in Appendix B, Proposition 14.

**PROPOSITION 7.** *For any exponent  $\alpha > 0$ , there exist a finite positive constant  $C := C(d, \mathbf{p}, \lambda, \alpha)$  and two exponents  $s := s(d, \mathbf{p}, \lambda, \alpha) > 0$  and  $\alpha := \alpha(d, \mathbf{p}, \lambda) > 0$  such that, for any  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{flux}, \alpha}(y)$  satisfying the stochastic integrability estimate*

$$(47) \quad \mathcal{M}_{\text{flux}, \alpha}(y) \leq \mathcal{O}_s(C),$$

such that for each radius  $r \geq \mathcal{M}_{\text{flux}, \alpha}(y)$ ,

$$(48) \quad \|\tilde{\mathbf{g}}_p\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r(y))} \leq C|p|r^\alpha.$$

**REMARK 7.** We emphasize that, in this article, the previous proposition is not a property of the diffusivity  $\bar{\sigma}^2$  but its definition: building on former result from [12, 40, 61], we prove that there exists a coefficient such that the estimate (48) is satisfied and name this coefficient  $\bar{\sigma}^2$ . Thanks to the estimate (48), we are then able to prove Theorem 1.1 and the invariance principle (4) with the same coefficient  $\bar{\sigma}^2$ . We refer to (181) and Remark 19 for a more detailed discussion.

2.4. *Random walks on graphs.* In this section, we record the Carne–Varopoulos bound pertaining to the transition kernel of the continuous-time random walk which holds on any infinite connected subgraph of  $\mathbb{Z}^d$ . This estimate is not as strong as the ones we are trying to establish (for instance, the ones of Theorem 1.1, or of Theorem 3.1 proved in [21]) but can be applied in greater generality: it applies to any realization of the infinite cluster, that is, to any environment  $\mathbf{a}$  in the set of probability 1 where there exists a unique infinite cluster, without any consideration about its geometry. From a mathematical perspective, this means that there is no minimal scale in the statement of Proposition 8.

**PROPOSITION 8** (Carne–Varopoulos bound, Corollaries 11 and 12 of [41]). *Let  $\mathcal{G}$  be an infinite, connected subgraph of  $\mathbb{Z}^d$  and  $\mathbf{a}$  be a function from the bonds of  $\mathcal{G}$  into  $[\lambda, 1]$ . For  $y \in \mathcal{G}$ , we let  $p(\cdot, \cdot, y)$  be the heat kernel associated to the parabolic equation*

$$\begin{cases} \partial_t p(\cdot, \cdot, y) - \nabla \cdot (\mathbf{a}\nabla p(\cdot, \cdot, y)) = 0 & \text{in } (0, \infty) \times \mathcal{G}, \\ p(0, \cdot, y) = \delta_y & \text{in } \mathcal{G}. \end{cases}$$

Then there exists a positive constant  $C := C(d, \lambda) < \infty$  such that for each point  $x \in \mathcal{G}$ ,

$$(49) \quad p(t, x, y) \leq \begin{cases} C \exp\left(-\frac{|x - y|^2}{Ct}\right) & \text{if } |x - y| \leq t, \\ C \exp\left(-\frac{|x - y|}{C}\left(1 + \ln \frac{|x - y|}{t}\right)\right) & \text{if } |x - y| \geq t. \end{cases}$$

REMARK 8. For later use, we note that when  $|x - y| \geq t$ ,

$$\exp\left(-\frac{|x - y|}{C}\left(1 + \ln \frac{|x - y|}{t}\right)\right) \leq \exp\left(-\frac{t}{C}\right).$$

A consequence of this inequality is that, by increasing the value of the constant  $C$ , one can add a factor  $t^{-d/2}$  in the second line of the right-hand side of (49): for every constant  $0 < C < \infty$ , there exists a finite constant  $C' > C$  such that, when  $|x - y| \geq t$ ,

$$C \exp\left(-\frac{|x - y|}{C}\left(1 + \ln \frac{|x - y|}{t}\right)\right) \leq C' t^{-d/2} \exp\left(-\frac{|x - y|}{C'}\left(1 + \ln \frac{|x - y|}{t}\right)\right).$$

**3. Decay and Lipschitz regularity of the heat kernel.** In this section, we collect and establish some estimates about the decay of the parabolic Green's function. In Section 3.1, we record a result of Barlow in [21], which establishes Gaussian upper bounds on the parabolic Green's function on the infinite cluster. This result is a percolation version of the Nash–Aronson estimate [18], originally proved for uniformly elliptic divergence form diffusions. Building upon the result of Barlow, we then establish estimates on the gradient of the parabolic Green's function on the percolation cluster, stated in Theorem 1.3, thanks to a large-scale  $C^{0,1}$ -regularity estimate. The argument makes use of techniques from stochastic homogenization and follows a classical route which can be decomposed into three steps: we first establish a quantitative homogenization theorem for the parabolic Dirichlet problem (see Section 3.2.1), once this is achieved we prove a large-scale  $C^{0,1}$ -regularity estimate for  $\mathbf{a}$ -caloric functions (see Section 3.2.2). In Section 3.2.3, we use this regularity estimate together with the heat kernel bound of Barlow to obtain the decay of the gradient of the heat kernel stated in Theorem 1.3.

3.1. *Decay of the heat kernel.* In this section, we record the result of Barlow [21], who established Gaussian bounds on the transition kernel. We first introduce the following function.

DEFINITION 3. Given a point  $x \in \mathbb{R}^d$ , a time  $t \in (0, \infty)$  and a constant  $0 < C < \infty$ , we define the function  $\Phi_C$  according to the formula

$$(50) \quad \Phi_C(t, x) := \begin{cases} C t^{-d/2} \exp\left(-\frac{|x|^2}{Ct}\right) & \text{if } |x| \leq t, \\ C t^{-d/2} \exp\left(-\frac{|x|}{C}\left(1 + \ln \frac{|x|}{t}\right)\right) & \text{if } |x| \geq t. \end{cases}$$

We note that this function is radial and increasing in the variable  $C$ . This function corresponds to a discrete heat kernel. For further use, we note that it satisfies the following semigroup property, for each  $t_1, t_2 \in (0, \infty)$  and each  $x, y \in \mathbb{Z}^d$ ,

$$(51) \quad \int_{\mathbb{Z}^d} \Phi_C(t_1, x - z) \Phi_C(t_2, z - y) dz \leq \Phi_{C'}(t_1 + t_2, x - y),$$

for some larger constant  $C' > C$ . This property is proved by an explicit computation or by using the semigroup property of the law of the random walk on  $\mathbb{Z}^d$  (see Remark 11). We define the function  $\Psi_C$  according to the formula

$$(52) \quad \Psi_C(t, r) := \begin{cases} (0, \infty) \times [0, \infty) \rightarrow \mathbb{R}, \\ (t, r) \mapsto -\ln(\Phi_C(t, x)), \quad \text{where } x \in \mathbb{R}^d \text{ satisfies } |x| = r. \end{cases}$$

In particular, one has the identity,

$$\forall x \in \mathbb{R}^d, \quad \Phi_C(t, x) = \exp(-\Psi_C(t, |x|)).$$

The function  $\Psi_C$  satisfies the following properties:

1. It is decreasing in the variable  $C$ , increasing in the variable  $r$  and continuous with respect to both variables;
2. It is convex with respect to the variable  $r$ .

We now record the result of Barlow.

**THEOREM 3.1** (Gaussian upper bound, Theorem 1 and Lemma 1.1 of [21]). *There exist an exponent  $s := s(d, \mathfrak{p}, \lambda) > 0$ , a positive constant  $C := C(d, \mathfrak{p}, \lambda) < \infty$  such that for each  $y \in \mathbb{R}^d$ , there exists a random time  $\mathcal{T}_{\text{NA}}(y)$  satisfying the stochastic integrability estimate*

$$(53) \quad \mathcal{T}_{\text{NA}}(y) \leq \mathcal{O}_s(C),$$

such that, on the event  $\{y \in \mathcal{C}_\infty\}$ , for every time  $t \in (0, \infty)$  satisfying  $t \geq \mathcal{T}_{\text{NA}}(y)$ , and every point  $x \in \mathcal{C}_\infty$ ,

$$(54) \quad p(t, x, y) \leq \Phi_C(t, x - y).$$

**REMARK 9.** The stochastic integrability estimate (53) is not stated in Theorem 1.1 of [21] but is mentioned in its remark equation (0.5) following the theorem.

**REMARK 10.** The estimate in the regime  $t \leq |x - y|$  does not require the assumption that  $t$  is larger than the minimal scale  $\mathcal{T}_{\text{NA}}(y)$  and is in fact a deterministic result: it is a consequence of Proposition 8 (proved in [41]) and Remark 8.

**REMARK 11.** The function  $\Phi_C$  can be used to obtain upper and lower bounds on the law of the random walk on the lattice  $\mathbb{Z}^d$ : there exist constants  $C_1, C_2$  depending only on the dimension  $d$  such that

$$(55) \quad \Phi_{C_1}(t, x - y) \leq p^{\mathbb{Z}^d}(t, x, y) \leq \Phi_{C_2}(t, x - y),$$

where we used the notation  $p^{\mathbb{Z}^d}(t, x, y) := \mathbb{P}_y[X_t = x]$ , and where  $(X_t)_{t \geq 0}$  denotes the VSRW on  $\mathbb{Z}^d$  starting from the point  $y$ . We refer to the work [43] of Delmotte and the work [41] for this result. The estimates (55) can then be used to prove the property (51). Indeed, since the random walk  $(X_t)_{t \geq 0}$  is a Markov process, its transition function  $p^{\mathbb{Z}^d}$  has the semigroup property, and we can write

$$\begin{aligned} \int_{\mathbb{Z}^d} \Phi_{C_1}(t_1, x - z) \Phi_{C_1}(t_2, z - y) dz &\leq \int_{\mathbb{Z}^d} p^{\mathbb{Z}^d}(t_1, x, z) p^{\mathbb{Z}^d}(t_2, z, y) dz \\ &= p^{\mathbb{Z}^d}(t_1 + t_2, x, y) \\ &\leq \Phi_{C_2}(t_1 + t_2, x - y). \end{aligned}$$

This argument gives the estimate (51) in the case  $C = C_1$ , but can be easily extended to any constant  $C > 0$ .

We complete this section by mentioning that the result of Barlow is proved for the heat kernel associated to the constant speed random walk and on the percolation cluster only, that is, when the conductances are only allowed to take the values 0 or 1. The adaptation to the variable speed random walk with uniformly elliptic conductances only requires a typographical change of the proof: all the computations performed in [21] to obtain the upper bound (54) can be adapted to our setting and so is the case of the existing results in the literature which are used in the proof.

3.2. *Decay of the gradient of the Green's function.* The main objective of this section is to prove Theorem 1.3. The proof of this result makes use of techniques from stochastic homogenization and can be split into three distinct steps, which correspond to the three following subsections. The first idea is to prove that the parabolic Green's function is close, on large scales, to a caloric function. This is carried out in Section 3.2.1 and the proof is based on a two-scale expansion. The analysis relies on the sublinearity of the corrector and the estimate on the  $\underline{H}^{-1}$ -norm of the centered flux stated in Section 2.3. This result is only necessary to establish a large-scale regularity theory for which sharp homogenization errors are not needed; we thus do not try to prove an optimal error estimate in the homogenization of the parabolic Dirichlet problem and only prove the result with an algebraic and suboptimal rate of convergence. Then, in Section 3.2.2, we use the homogenization estimate proved in Theorem 3.2.1 to establish a large-scale regularity theory in the spirit of [15], Chapter 3, or [12], Section 7. Finally, in Section 3.2.3, we combine Proposition 9 and the heat kernel bound proved by Barlow and stated in Theorem 3.1 to deduce Theorem 1.3.

3.2.1. *Homogenization of the parabolic Dirichlet problem.* In this section, we prove a quantitative homogenization theorem for the parabolic Cauchy–Dirichlet problem on the infinite cluster. In the following statement, we let  $\eta$  be a smooth, nonnegative function supported in the ball  $B_{\frac{1}{2}}(0)$ , and satisfying the identity  $\int \eta = 1$ . It is used as a smoothing operator in the convolution (59). We also define the set  $\text{Cv}(\mathbb{Z}^d \cap B_r(y))$  to be the convex hull of the set  $\mathbb{Z}^d \cap B_r(y)$ , that is,

$$(56) \quad \begin{aligned} &\text{Cv}(\mathbb{Z}^d \cap B_r(y)) \\ &:= \left\{ z \in \mathbb{R}^d : z = \sum_i \alpha_i x_i, x_i \in \mathbb{Z}^d \cap B_r(y), 0 \leq \alpha_i \leq 1 \text{ and } \sum_i \alpha_i = 1 \right\}. \end{aligned}$$

It is used to define the domain of the homogenized equation so that the boundary condition coincides.

**THEOREM 3.2.** *Fix an exponent  $\delta > 0$ , then there exist a finite positive constant  $C(d, \lambda, p, \delta)$ , two exponents  $s(d, \lambda, p, \delta) > 0$ ,  $\alpha(d, \lambda, p, \delta) > 0$  such that for any point  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{\text{hom},\delta}(y)$  satisfying*

$$\mathcal{M}_{\text{hom},\delta}(y) \leq \mathcal{O}_s(C)$$

*such that, for every  $r > \mathcal{M}_{\text{hom}}(y)$ , and every boundary condition  $f \in W_{\text{par}}^{1,2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r(y)))$ , the following statement is valid. Let  $u$  be the weak solution of the parabolic equation*

$$(57) \quad \begin{cases} (\partial_t - \nabla \cdot \mathbf{a}\nabla)u = 0 & \text{in } I_r \times (\mathcal{C}_\infty \cap B_r(y)), \\ u = f & \text{on } \partial_{\sqcup}(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r(y))) \cap (I_r \times (\mathcal{C}_\infty \cap B_r(y))), \end{cases}$$

*and  $\bar{u}$  be the weak solution of the homogenized, continuous in space, parabolic equation*

$$(58) \quad \begin{cases} \left( \partial_t - \frac{1}{2} \bar{\sigma}^2 \Delta \right) \bar{u} = 0 & \text{in } I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r(y)), \\ \bar{u} = \tilde{f} & \text{on } \partial_{\sqcup}(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r(y))), \end{cases}$$

*where the boundary condition  $\tilde{f}$  is the extension of  $f$  to the continuous parabolic cylinder defined by the formula*

$$(59) \quad \tilde{f} := [f]_{\mathcal{P}} \star \eta,$$

*and the extension  $[f]_{\mathcal{P}}$  is defined in the paragraph following Proposition 2. Then the following estimate holds:*

$$(60) \quad \frac{1}{r} \|u - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r(y)))} \leq Cr^{-\alpha} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r(y)))}.$$

REMARK 12. The equation (57) is discrete in space and continuous in time, while equation (58) is both continuous in space and time. The solution  $u$  and  $\bar{u}$  coincide on the parabolic boundary  $\partial_{\square}(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r(y))) \cap (I_r \times (\mathcal{C}_{\infty} \cap B_r(y)))$  for the equation on the clusters. All the norms in the inequality (60) are discrete in space and continuous in time.

REMARK 13. The reason we define the homogenized limit to be continuous is the following: we need to use a number of results (e.g., regularity theory for the homogenized equation, the Meyers estimate) which are usually stated in the continuous setting. Moreover, one has explicit formulas for the elliptic and parabolic Green’s functions and the continuous object is better behaved regarding scaling properties. On a higher level, the correct limiting object should be the continuous function as, over large-scales, the discrete lattice approximates the continuum.

PROOF OF THEOREM 3.2. By translation invariance of the model, we assume without loss of generality that  $y = 0$  and do some preparation before the proof. We first define the minimal scale  $\mathcal{M}_{\text{hom},\delta}(0)$  to be equal to

$$\mathcal{M}_{\text{hom},\delta}(0) := \max(\mathcal{M}_{\text{Poinc}}(0), \mathcal{M}_q(\mathcal{P}), \mathcal{M}_{\text{corr},\frac{1}{2}}(0), \mathcal{M}_{\text{flux},\frac{1}{2}}(0)),$$

where the parameter  $q$  is assumed to be larger than  $4d$  and will be fixed at the end of the proof. Using the stochastic integrability estimates (37), (43), (44), (47) on the four minimal scales together with the property (24) of the  $\mathcal{O}_s$  notation, one has

$$\mathcal{M}_{\text{hom},\delta}(0) \leq \mathcal{O}_s(C).$$

We record that under the assumption  $r \geq \mathcal{M}_{\text{hom},\delta}(0) \geq \mathcal{M}_{2d}(\mathcal{P})$ , one has

$$(61) \quad cr^d \leq |\mathcal{C}_{\infty} \cap B_r| \leq Cr^d,$$

which allows to compare the number of points of the infinite cluster in the ball  $B_r$  with the volume of the ball  $B_r$ . This estimate can be deduced by an application of the estimate (38) with the Cauchy–Schwarz inequality:

$$\begin{aligned} |B_r|^2 &\leq \left( \sum_{\square \in \mathcal{P}, \square \cap B_r \neq \emptyset} (\text{size}(\square))^d \right)^2 \leq \left( \sum_{\square \in \mathcal{P}, \square \cap B_r \neq \emptyset} 1 \right) \left( \sum_{\square \in \mathcal{P}, \square \cap B_r \neq \emptyset} (\text{size}(\square))^{2d} \right) \\ &\leq C |\mathcal{C}_{\infty} \cap B_r| r^d. \end{aligned}$$

We record the following interior regularity estimate for the homogenized function  $\bar{u}$ , which is standard for solutions of the heat equation (see [47], Theorem 9, Section 2.3): for every pair  $(t, x) \in I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)$ , and every radii  $r_1, r_2 > 0$  such that  $I_{r_1}(t) \times B_{r_2}(x) \subseteq I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)$ , one has the inequality

$$(62) \quad \forall k, l \in \mathbb{N}, \quad |\partial_t^l \nabla^{k+1} \bar{u}|(t, x) \leq C_{k+2l} (r_1)^{-2l} (r_2)^{-k} \|\nabla \bar{u}\|_{\underline{L}^2(I_{r_1}(t) \times B_{r_2}(x))}.$$

We remark that in [47], Theorem 9, Section 2.3, the inequality is stated in the case when  $r_1 = r_2$ ; The estimate (62) can be recovered by a careful investigation of the proof.

We introduce a cut-off function  $\Upsilon$  in the parabolic cylinder  $I_r \times B_r$  constant equal to 1 in the interior of the cylinder and decreasing linearly to 0 in a mesoscopic boundary layer of size  $r' \ll r$ ,

$$(63) \quad \begin{cases} \Upsilon(t, x) \equiv 1 & (t, x) \in I_r \times B_r, \text{dist}(x, \partial B_r) \geq 2r' \text{ and } \text{dist}(t, \partial I_r) \geq 2(r')^2, \\ 0 \leq \Upsilon(t, x) \leq 1 & (t, x) \in I_r \times B_r, \\ \Upsilon(t, x) \equiv 0 & (t, x) \in I_r \times B_r, \text{dist}(x, \partial B_r) \leq r' \text{ or } \text{dist}(t, \partial I_r) \leq (r')^2. \end{cases}$$

The precise value of the parameter  $r'$  is given by the formula  $r' = r^{1-\beta}$  for some small exponent  $\beta$  whose value is decided at the end of the proof. We additionally assume that the function  $\Upsilon$  is smooth and satisfies the estimate, for each pair of integers  $k, l \in \mathbb{N}$ ,

$$(64) \quad |\partial_t^l \nabla^k \Upsilon| \leq C_{k+2l} (r')^{-(k+2l)}.$$

With these quantities, we can prove the following lemma.

LEMMA 3.3. *We have the estimate, for each pair of integers  $k, l \in \mathbb{N}$ ,*

$$(65) \quad \|\partial_t^l \nabla^k (\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)} \leq C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

PROOF. First, by using the inequality (62) and the fact that the map  $\Upsilon$  is supported outside a boundary layer of size  $r'$  in the parabolic cylinder  $I_r \times B_r$ , we obtain the estimate, for each  $(t, x) \in \text{supp}(\Upsilon)$  and each pair of integers  $k, l \in \mathbb{N}$ ,

$$(66) \quad \begin{aligned} |\partial_t^l \nabla^k (\Upsilon \nabla \bar{u})|(t, x) &\leq C_{k+2l} (r')^{-(k+2l)} \|\nabla \bar{u}\|_{\underline{L}^2(I_{r'}(t) \times B_{r'}(x))} \\ &\leq C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla \bar{u}\|_{\underline{L}^2(I_r \times C_V(\mathbb{Z}^d \cap B_r))}. \end{aligned}$$

The inequality (66) implies the  $L^\infty$ -estimate, for each pair of integers  $k, l \in \mathbb{N}$ ,

$$(67) \quad \|\partial_t^l \nabla^k (\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)} \leq C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla \bar{u}\|_{\underline{L}^2(I_r \times C_V(\mathbb{Z}^d \cap B_r))}.$$

We then state the global Meyers estimate for the map  $\bar{u}$ : there exists an exponent  $\delta_0 := \delta_0(d, \lambda, p) > 0$  such that for every  $\delta' \in [0, \delta_0]$ ,

$$(68) \quad \|\nabla \bar{u}\|_{L^{2+\delta'}(I_r \times C_V(\mathbb{Z}^d \cap B_r))} \leq C \|\nabla \tilde{f}\|_{L^{2+\delta'}(I_r \times C_V(\mathbb{Z}^d \cap B_r))}.$$

A proof of this result can be found in [50], Proposition 5.1, where the statement is given for cubes instead of parabolic cylinders (the adaptation to the setting considered here does not affect the proof). Moreover, one can estimate the  $L^p$ -norm of the (continuous) gradient of the function  $\tilde{f}$  in terms of the  $L^p$ -norm of the (discrete) gradient of the maps  $f$  and the sizes of the cubes of the partition. The formula is a consequence of [12], Lemma 3.3, and recalled in (41): for any  $p \geq 1$ , and any radius  $r \geq \text{size}(\square_{\mathcal{P}}(0))$ ,

$$\|\nabla \tilde{f}\|_{L^p(C_V(\mathbb{Z}^d \cap B_r))} \leq \int_{B_r \cap \mathcal{C}_\infty} \text{size}(\square_{\mathcal{P}}(x))^{p d-1} |\nabla f(x)|^p dx.$$

Applying the Hölder inequality to this estimate with  $r \geq \mathcal{M}_{\text{hom},\delta}(0)$ , using the assumption the minimal scale  $\mathcal{M}_{\text{hom},\delta}(0)$  is larger than the minimal scale  $\mathcal{M}_q(\mathcal{P})$  so that Proposition 2 is valid, and choosing the parameter  $q$  to be large enough (larger than the value  $\frac{(4+2\delta)((2+\frac{1}{2}\delta)d-1)}{\delta}$ ), one obtains the following inequality: for any  $p \in [2, 2 + \frac{1}{2}\delta]$ ,

$$(69) \quad \|\nabla \tilde{f}\|_{\underline{L}^p(C_V(\mathbb{Z}^d \cap B_r))} \leq C \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

Together with (68), this shows the inequality, for any exponent  $\delta' \in [0, \min(\delta_0, \frac{1}{2}\delta)]$ ,

$$(70) \quad \|\nabla \bar{u}\|_{\underline{L}^{2+\delta'}(I_r \times C_V(\mathbb{Z}^d \cap B_r))} \leq C \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

Putting the inequality (70) back into the estimate (67) concludes the proof of (65).  $\square$

The key ingredient in the proof of Theorem 3.2 is to use a modified two-scale expansion on the percolation cluster, defined for each  $(t, x) \in I_r \times (\mathcal{C}_\infty \cap B_r)$  by the formula

$$(71) \quad w(t, x) := \bar{u}(t, x) + \Upsilon(t, x) \sum_{k=1}^d \mathcal{D}_{e_k} \bar{u}(t, x) \chi_{e_k}(x),$$

as an intermediate quantity: we prove that the function  $w$  is close to both functions  $u$  and  $\bar{u}$ . Here and in the rest of this section, the map  $\chi_{e_k}$  is the first-order corrector normalized according the procedure described in Section 2.3 around the point  $y = 0$ . The proof of Theorem 3.2 can be decomposed into five steps.

*Step 1: Control over  $\frac{1}{r} \|w - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}$ .* We use the estimate (65) to compute

$$\begin{aligned} \frac{1}{r} \|w - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} &= \frac{1}{r} \left\| \sum_{k=1}^d \Upsilon(\mathcal{D}_{e_k} \bar{u}) \chi_{e_k} \right\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \\ &\leq \frac{1}{r} \|\Upsilon \nabla \bar{u}\|_{L^\infty(I_r \times B_r)} \sum_{k=1}^d \|\chi_{e_k}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r)} \\ &\leq \frac{C}{r} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))} \sum_{k=1}^d \|\chi_{e_k}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r)}. \end{aligned}$$

Using the assumption  $r \geq \mathcal{M}_{\text{hom},\delta}(0) \geq \mathcal{M}_{\text{corr},\frac{1}{2}}(0)$ , we deduce

$$\frac{1}{r} \|w - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq Cr^{-\frac{1}{2}} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

The proof of Step 1 is complete.

*Step 2: Control of the norm  $\frac{1}{r} \|w - u\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}$ .* We first note that the functions  $w$  and  $u$  are equal on the boundary of the parabolic cylinder  $I_r \times (\mathcal{C}_\infty \cap B_r)$ , and use the assumption  $r \geq \mathcal{M}_{\text{hom},\delta}(0) \geq \mathcal{M}_{\text{Poinc}}(0)$  to apply the Poincaré inequality for each fixed time  $t$ , and then integrate over time. This proves

$$\frac{1}{r} \|w - u\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq \|\nabla(w - u)\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

Then we use an integration by part and the uniform ellipticity of the environment on the infinite cluster

$$\begin{aligned} &\|\nabla(w - u)\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}^2 \\ &\leq \frac{1}{\lambda |I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{I_r} \int_{\mathcal{C}_\infty \cap B_r} \nabla(w - u) \cdot \mathbf{a} \nabla(w - u) \\ &= \frac{1}{\lambda |I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{I_r} \int_{\mathcal{C}_\infty \cap B_r} (-\nabla \cdot \mathbf{a} \nabla(w - u))(w - u). \end{aligned}$$

The fact that the functions  $w, u$  have the same initial condition over  $\mathcal{C}_\infty \cap B_r$  implies that the following integral is nonnegative

$$\begin{aligned} &\int_{I_r} \int_{\mathcal{C}_\infty \cap B_r} (\partial_t(w - u))(w - u) \\ &= \frac{1}{2} (\|(w - u)(0, \cdot)\|_{L^2(\mathcal{C}_\infty \cap B_r)}^2 - \|(w - u)(-r^2, \cdot)\|_{L^2(\mathcal{C}_\infty \cap B_r)}^2) \\ &= \frac{1}{2} \|(w - u)(0, \cdot)\|_{L^2(\mathcal{C}_\infty \cap B_r)}^2 \geq 0. \end{aligned}$$

We combine this formula and equation (57) to obtain

$$\begin{aligned} & \|\nabla(w - u)\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}^2 \\ & \leq \frac{1}{\lambda |I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{I_r} \int_{\mathcal{C}_\infty \cap B_r} ((\partial_t - \nabla \cdot \mathbf{a}\nabla)(w - u))(w - u) \\ & \leq \frac{1}{\lambda} \|w - u\|_{\underline{L}^2(I_r; \underline{H}^1(\mathcal{C}_\infty \cap B_r))} \|(\partial_t - \nabla \cdot \mathbf{a}\nabla)w\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))}. \end{aligned}$$

This shows that

$$\frac{1}{r} \|w - u\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r)} \leq C \|(\partial_t - \nabla \cdot \mathbf{a}\nabla)w\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))}.$$

*Step 3: Control over  $\|(\partial_t - \nabla \cdot \mathbf{a}\nabla)w\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))}$ .* In this step, we adopt the finite difference notation and recall the identity  $(\partial_t - \nabla \cdot \mathbf{a}\nabla)w = (\partial_t + \mathcal{D}^* \cdot \mathbf{a}\mathcal{D})w$ . To estimate the  $\underline{H}^{-1}$ -norm of  $(\partial_t - \nabla \cdot \mathbf{a}\nabla)w$ , the idea is to derive an explicit formula for this quantity by using the definition of  $w$  given in (71) and to make a centered flux  $\tilde{\mathbf{g}}_{e_k} = \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k$  appear. We first calculate  $\partial_t w$  and  $\mathcal{D}w$  and obtain the formulas

$$\begin{cases} \partial_t w = \partial_t \bar{u} + \sum_{k=1}^d \partial_t (\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k}, \\ \mathcal{D}w = (1 - \Upsilon) \mathcal{D}\bar{u} + \sum_{k=1}^d (\Upsilon \mathcal{D}_{e_k} \bar{u})(e_k + \mathcal{D}\chi_{e_k}) + \sum_{k=1}^d \mathcal{D}(\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k}. \end{cases}$$

We combine the two equations to calculate  $(\partial_t + \mathcal{D}^* \cdot \mathbf{a}\mathcal{D})w$ ,

$$\begin{aligned} & (\partial_t + \mathcal{D}^* \cdot \mathbf{a}\mathcal{D})w \\ (72) \quad & = \partial_t \bar{u} + \sum_{k=1}^d \partial_t (\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k} + \mathcal{D}^* \cdot ((1 - \Upsilon) \mathbf{a}\mathcal{D}\bar{u}) \\ & \quad + \sum_{k=1}^d \mathcal{D}^* \cdot ((\Upsilon \mathcal{D}_{e_k} \bar{u}) \mathbf{a}(e_k + \mathcal{D}\chi_{e_k})) + \sum_{k=1}^d \mathcal{D}^* \cdot (\mathbf{a}\mathcal{D}(\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k}). \end{aligned}$$

Then we use equation (58) which reads  $\partial_t \bar{u} = \frac{1}{2}\bar{\sigma}^2 \Delta \bar{u}$  to replace the term  $\partial_t \bar{u}$  in the equation above. Notice that here  $\frac{1}{2}\bar{\sigma}^2 \Delta \bar{u}$  refers to the continuous Laplacian, but using the regularity properties on the function  $\bar{u}$  stated in (65), we can replace this term by the discrete Laplacian  $-\frac{1}{2}\bar{\sigma}^2 \mathcal{D}^* \cdot \mathcal{D}\bar{u}$  by paying only a small error. The advantage of this operation is that we can use the two terms  $-\frac{1}{2}\bar{\sigma}^2 \mathcal{D}^* \cdot (\Upsilon \mathcal{D}\bar{u})$  and  $\sum_{k=1}^d \mathcal{D}^* \cdot ((\Upsilon \mathcal{D}_{e_k} \bar{u}) \mathbf{a}(e_k + \mathcal{D}\chi_{e_k}))$  to make the flux appear: we have

$$\begin{aligned} & \sum_{k=1}^d \mathcal{D}^* \cdot ((\Upsilon \mathcal{D}_{e_k} \bar{u}) \mathbf{a}(e_k + \mathcal{D}\chi_{e_k})) - \frac{1}{2}\bar{\sigma}^2 \mathcal{D}^* \cdot (\Upsilon \mathcal{D}\bar{u}) \\ (73) \quad & = \sum_{k=1}^d \mathcal{D}^* \cdot \left( (\Upsilon \mathcal{D}_{e_k} \bar{u}) \left( \mathbf{a}(e_k + \mathcal{D}\chi_{e_k}) - \frac{1}{2}\bar{\sigma}^2 e_k \right) \right) \\ & = \sum_{k=1}^d \mathcal{D}^* (\Upsilon \mathcal{D}_{e_k} \bar{u}) \cdot \tilde{\mathbf{g}}_{e_k}^*, \end{aligned}$$

where  $\tilde{\mathbf{g}}_{e_k}^*$  is a translated version of the flux  $\tilde{\mathbf{g}}_{e_k}$  defined by the formula, for each  $x \in \mathcal{C}_\infty$ ,

$$\tilde{\mathbf{g}}_{e_k}^*(x) := \begin{pmatrix} T_{-e_1} \left[ \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k \right]_1 \\ \vdots \\ T_{-e_d} \left[ \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k \right]_d \end{pmatrix},$$

where we recall the notation  $[\mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k]_i$  introduced in Section 1.6.1 for the  $i$ th-component of the vector  $\mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k$ . In Appendix B, it is proved that the translated flux  $\tilde{\mathbf{g}}_{e_k}^*$  has similar properties as the centered flux  $\tilde{\mathbf{g}}_{e_k}$ . In particular, it is proved in Remark 20 that for every radius  $r \geq \mathcal{M}_{\text{corr}, \frac{1}{2}}(0)$ ,

$$\|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} \leq Cr^{\frac{1}{2}}.$$

Combining the identities (72) and (73), one obtains

$$\begin{aligned} (\partial_t - \mathcal{D}^* \cdot \mathbf{a}\mathcal{D})w &= \underbrace{\frac{1}{2}(\nabla \cdot \bar{\sigma}^2(\Upsilon \nabla \bar{u}) - (-\mathcal{D}^* \cdot \bar{\sigma}^2(\Upsilon \mathcal{D}\bar{u})))}_{(74)\text{-a}} \\ &+ \underbrace{\sum_{k=1}^d \partial_t(\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k}}_{(74)\text{-b}} \\ &+ \underbrace{\mathcal{D}^* \cdot ((1 - \Upsilon)\mathbf{a}\mathcal{D}\bar{u})}_{(74)\text{-c1}} + \underbrace{\frac{1}{2}(\nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u}))}_{(74)\text{-c2}} \\ &+ \underbrace{\sum_{k=1}^d \mathcal{D}^*(\Upsilon \mathcal{D}_{e_k} \bar{u}) \cdot \tilde{\mathbf{g}}_{e_k}^*}_{(74)\text{-d}} \\ &+ \underbrace{\sum_{k=1}^d \mathcal{D}^* \cdot (\mathbf{a}\mathcal{D}(\Upsilon \mathcal{D}_{e_k} \bar{u}) \chi_{e_k})}_{(74)\text{-e}} \end{aligned} \tag{74}$$

There remains to use triangle inequality and estimate the  $\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))$ -norm of each term. The following estimates will be used several times: given two functions  $A : I_r \times (\mathcal{C}_\infty \cap B_r) \rightarrow \mathbb{R}$  and  $B : I_r \times (\mathcal{C}_\infty \cap B_r) \rightarrow \mathbb{R}$ , one has

$$\begin{aligned} &\|AB\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} \\ &= \sup_{\|v\|_{\underline{L}^2(I_r; \underline{H}^1(\mathcal{C}_\infty \cap B_r))} \leq 1} \frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{I_r \times (\mathcal{C}_\infty \cap B_r)} ABv \\ &\leq \|A\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} \sup_{\|v\|_{\underline{L}^2(I_r; \underline{H}^1(\mathcal{C}_\infty \cap B_r))} \leq 1} \|Bv\|_{\underline{L}^2(I_r; \underline{H}^1(\mathcal{C}_\infty \cap B_r))} \\ &\leq \|A\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} (\|B\|_{L^\infty(I_r \times (\mathcal{C}_\infty \cap B_r))} + r\|\nabla B\|_{L^\infty(I_r \times (\mathcal{C}_\infty \cap B_r))}). \end{aligned} \tag{75}$$

From the definition of the  $\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))$ -norm, one also has the estimate

$$\|A\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} \leq r\|A\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq r\|A\|_{L^\infty(I_r \times (\mathcal{C}_\infty \cap B_r))}. \tag{76}$$

The term (74)-a is a difference between a discrete derivative and a continuous derivative; it can be estimated in terms of the third derivative of the function  $\bar{u}$ . Using the estimates (65) and (76) shows

$$\begin{aligned}
 (77) \quad \|(74)\text{-a}\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} &\leq r \|\nabla^2(\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)} \\
 &\leq Cr^{-1} \left(\frac{r}{r'}\right)^{3+\frac{d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.
 \end{aligned}$$

A similar strategy can be used to estimate the term (74)-b

$$\begin{aligned}
 (78) \quad \|(74)\text{-b}\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} &\leq (r \|\partial_t \nabla(\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)} + \|\partial_t(\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)}) \sum_{k=1}^d \|\chi_{e_k}\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} \\
 &\leq C \frac{r}{(r')^3} \left(\frac{r}{r'}\right)^{\frac{2+d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))} \sum_{k=1}^d \|\chi_{e_k}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r)} \\
 &\leq Cr^{-\frac{3}{2}} \left(\frac{r}{r'}\right)^{4+\frac{d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))},
 \end{aligned}$$

where we use the assumption  $r \geq \mathcal{M}_{\text{hom},\delta}(0) \geq \mathcal{M}_{\text{corr},\frac{1}{2}}(0)$  to obtain the sublinearity of the corrector and the regularity estimate (65) to go from the second line to the third line.

To estimate the term (74)-c1, we note that the function  $(1 - \Upsilon)$  is equal to 0 outside a mesoscopic boundary layer of size  $r'$  of the ball  $B_r$ . We thus apply the Meyers estimate (68), with the exponent  $\delta' = \min(\delta_0, \frac{1}{2}\delta)$ , and the Hölder inequality. This shows

$$\begin{aligned}
 (79) \quad \|(74)\text{-c1}\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} &\leq \|(1 - \Upsilon)\mathbf{a}\mathcal{D}\bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \\
 &\leq \|1 - \Upsilon\|_{\underline{L}^{\frac{4+2\delta'}{\delta'}}(I_r \times B_r)} \|\nabla \bar{u}\|_{\underline{L}^{2+\delta'}(I_r \times B_r)} \\
 &\leq C \left(\frac{r'}{r}\right)^{\frac{\delta'}{4+2\delta'}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))},
 \end{aligned}$$

where we used the Hölder inequality to go from the first line to the second line and the Meyers estimate to go from the second line to the third line.

We want to apply a similar technique to treat the term (74)-c2 since it is also a boundary layer term. However, we should notice that here the derivative  $\nabla$  is the continuous gradient defined on  $\mathbb{R}^d$  and there is no conductance  $\mathbf{a}$ , thus we cannot apply a discrete integration by part on the cluster. We will focus on this term later in Step 4.

To estimate the term (74)-d, we apply the inequality (75), the regularity estimate (65) and we use the assumption  $r \geq \mathcal{M}_{\text{hom},\delta}(0) \geq \mathcal{M}_{\text{flux},\frac{1}{2}}(0)$ . We obtain

$$\begin{aligned}
 (80) \quad \|(74)\text{-d}\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} &\leq (\|\nabla(\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)} + r \|\nabla^2(\Upsilon \nabla \bar{u})\|_{L^\infty(I_r \times B_r)}) \sum_{k=1}^d \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} \\
 &\leq Cr^{-\frac{1}{2}} \left(\frac{r}{r'}\right)^{3+\frac{d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.
 \end{aligned}$$

The term (74)-e can be estimated thanks to an integration by part and the regularity estimate (65). This yields

$$\begin{aligned}
 & \| (74)\text{-e} \|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} \\
 (81) \quad & \leq \| \nabla(\Upsilon \nabla \bar{u}) \|_{L^\infty(I_r \times B_r)} \sum_{k=1}^d \| \chi_{e_k} \|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \\
 & \leq C r^{-\frac{1}{2}} \left( \frac{r}{r'} \right)^{2+\frac{d}{2}} \| \nabla f \|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}.
 \end{aligned}$$

*Step 4: Control over the term  $\| \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))}$ .* As was already mentioned, we cannot use a discrete integration by parts to estimate the  $L^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))$ -norm of this term. The strategy relies on the interior regularity estimate (62) which requires careful treatments since it is close to the boundary. We apply the Whitney decomposition on the ball  $B_r$  stated below with a minor adaptation to triadic cubes.

LEMMA 3.4 (Whitney decomposition). *There exists a family of closed triadic cubes  $\{Q_j\}_{j \geq 0}$  such that:*

1.  $B_r = \bigcup_j Q_j$  and the cubes  $Q_j$  have disjoint interiors;
2.  $\sqrt{d} \text{size}(Q_j) \leq \text{dist}(Q_j, \partial B_r) \leq 4\sqrt{d} \text{size}(Q_j)$ ;
3. Two neighboring cubes  $Q_j$  and  $Q_k$  have comparable sizes in the sense that

$$\frac{1}{3} \leq \frac{\text{size}(Q_k)}{\text{size}(Q_j)} \leq 3;$$

4. Each cube  $Q_j$  has at most  $C(d)$  neighbors.

We skip the construction of this partition, refer to [80], Theorem 3, or [59], Appendix J, for the proof and to Figure 4 for an illustration. With the help of this decomposition, we can estimate the norm  $\| \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))}$ . We first relabel the cubes of the decomposition according to their size; we write

$$\{Q_j\}_{j \geq 1} := \bigcup_{n=0}^{\infty} \bigcup_{k=1}^{M_n} \{Q_{n,k}\}, \quad 3^{-(n+1)}r \leq \text{size}(Q_{n,k}) < 3^{-n}r,$$

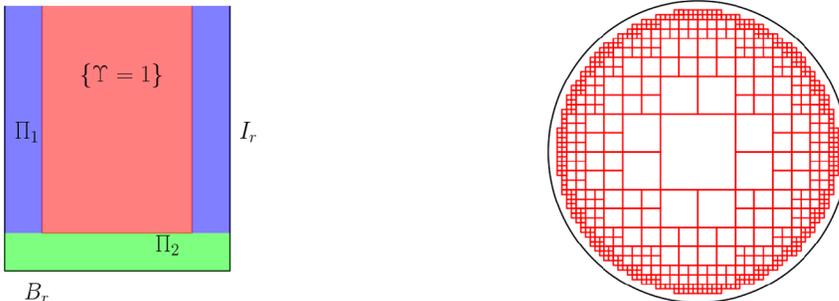


FIG. 4. The figure on the left illustrates the partition of the cylinder  $I_r \times B_r$ , where the domain in blue stands for the set  $\Pi_1$ , the one in green for the set  $\Pi_2$  and the one in red for  $\{(t, x) : \Upsilon(t, x) = 1\}$ . The figure on the right illustrates a Whitney decomposition in the ball  $B_r$ , where we use dyadic cubes to improve the readability.

where  $M_n$  is the number of the cubes whose size belongs to the interval  $[3^{-(n+1)}r, 3^{-n}r)$ . Then we decompose the set  $\text{supp}(1 - \Upsilon)$  into two parts (see Figure 4)

$$\text{supp}(1 - \Upsilon) = \Pi_1 \sqcup \Pi_2,$$

$$\Pi_1 := \{(t, x) \in I_r \times B_r : \text{dist}(x, \partial B_r) \leq 2r', -r^2 + 2(r')^2 \leq t \leq 0\},$$

$$\Pi_2 := \{(t, x) \in I_r \times B_r : -r^2 \leq t \leq -r^2 + 2(r')^2\}.$$

We estimate the weak norm thanks to its definition: we let  $\varphi$  be a function from  $I_r \times (\mathcal{C}_\infty \cap B_r)$  to  $\mathbb{R}$  which satisfies  $\|\varphi\|_{\underline{L}^2(I_r; \underline{H}^1(\mathcal{C}_\infty \cap B_r))} \leq 1$  and is equal to 0 on the boundary  $I_r \times \partial_{\mathbf{a}}(\mathcal{C}_\infty \cap B_r)$ . We split the integral

$$\begin{aligned} & \int_{I_r} \int_{\mathcal{C}_\infty \cap B_r} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\varphi \\ &= \int_{(I_r \times \mathbb{Z}^d) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\varphi \\ & \quad + \int_{(I_r \times \mathbb{Z}^d) \cap \Pi_2} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\varphi, \end{aligned}$$

and treat the two terms separately.

*Step 4.1: Control of the weak norm over  $\Pi_1$ .* For the term involving the set  $\Pi_1$ , we use the Whitney decomposition to integrate on every cube of the partition. We first introduce the time intervals, for  $m, n \in \mathbb{N}$ ,  $I_{1,m} := -m + (-1, 0]$  and  $nI_{1,m} := -m + (-n, 0]$ , and partition the boundary layer  $\Pi_1$  according to the formula

$$\Pi_1 = \bigcup_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \bigcup_{n=0}^{\infty} \bigcup_{k=1}^{M_n} (I_{1,m} \times Q_{n,k}) \cap \Pi_1.$$

We remark that we can restrict our attention to the cubes whose sizes is between 1 and  $r'$  thanks to the properties of the Whitney decomposition. Indeed, the cubes of size larger than  $r'$  remain outside the boundary layer  $\Pi_1$ , since the distance of a cube to the boundary is comparable to its size. On the other hand, the cubes of size smaller than 1 will not contain a point in the lattice  $\text{int}(\text{Cv}(\mathbb{Z}^d \cap B_r))$ , as these cubes are too close to the boundary, and the definition (56) implies that all the points of the lattice in the interior  $\text{int}(\text{Cv}(\mathbb{Z}^d \cap B_r))$  are at distance at least  $\frac{1}{\sqrt{2}}$  from the boundary. We thus have the following partition, if we denote by  $n_0$  and  $n_1$  the integers such that  $3^{-(n_0+1)}r \leq r' < 3^{-n_0}r$ , and  $3^{-(n_1+1)}r \leq 1 < 3^{-n_1}r$ ,

$$\Pi_1 = \bigcup_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \bigcup_{n=n_0}^{n_1} \bigcup_{k=1}^{M_n} (I_{1,m} \times Q_{n,k}) \cap \Pi_1.$$

Using this partition, we can split the integral

$$\begin{aligned} & \int_{(I_r \times \mathcal{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\varphi \\ (82) \quad &= \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \int_{(I_r \times \mathcal{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\varphi. \end{aligned}$$

We fix a cylinder  $I_{1,m} \times Q_{n,k}$ , apply the Cauchy–Schwarz inequality and use the interior regularity estimate (62) of the function  $\bar{u}$  in the cylinder  $I_{1,m} \times Q_{n,k}$  with the property that

the distance  $\text{dist}(Q_{n,k}, \partial B_r)$  is larger than  $\sqrt{d}$   $\text{size}(Q_{n,k})$  and the inclusion  $2Q_{n,k} \subseteq B_r$ . We obtain

$$\begin{aligned} & \left| \int_{(I_r \times \mathcal{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ & \leq \| \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \|_{L^2(I_{1,m} \times Q_{n,k})} \| \varphi \|_{L^2(I_{1,m} \times (\mathcal{C}_\infty \cap Q_{n,k}))} \\ & \leq C(3^{-n}r)^{-1} \| \nabla \bar{u} \|_{L^2(2I_{1,m} \times 2Q_{n,k})} \| \varphi \|_{L^2(I_{1,m} \times (\mathcal{C}_\infty \cap Q_{n,k}))}. \end{aligned}$$

We sum over all the cubes  $\{Q_{n,k}\}_{1 \leq k \leq M_n}$  and apply the Cauchy–Schwarz inequality

$$\begin{aligned} & \sum_{k=1}^{M_n} \left| \int_{(I_r \times \mathcal{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ & \leq \sum_{k=1}^{M_n} C(3^{-n}r)^{-1} \| \nabla \bar{u} \|_{L^2(2I_{1,m} \times 2Q_{n,k})} \| \varphi \|_{L^2(I_{1,m} \times (\mathcal{C}_\infty \cap Q_{n,k}))} \\ & \leq C(3^{-n}r)^{-1} \left( \sum_{k=1}^{M_n} \| \nabla \bar{u} \|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2 \right)^{\frac{1}{2}} \| \varphi \|_{L^2(I_{1,m} \times (\mathcal{C}_\infty \cap (\cup_{k=1}^{M_n} Q_{n,k}))})}. \end{aligned}$$

We then use the following three ingredients:

- Given a discrete set  $A \subseteq \mathbb{Z}^d$ , the  $L^2$ -norm of coarsened function  $[\varphi]_{\mathcal{P}}$  over  $A$  is larger than the one of the function  $\varphi$  over the set  $\mathcal{C}_\infty \cap A$ ;
- We have the inclusion  $\cup_{k=1}^{M_n} Q_{n,k} \subseteq \{x \in B_r : \text{dist}(x, \partial B_r) \leq 5 \times 3^{-n} \sqrt{dr}\}$ ;
- We choose the vertex  $z(\cdot)$  (defined in (39)) to be a point on the boundary  $\partial B_r$  for the cubes of the partition intersecting  $\partial B_r$ . With this convention, the coarsened function  $[\varphi]_{\mathcal{P}}$  is equal to zero on  $\partial B_r$ , so we can apply the Poincaré inequality for  $[\varphi]_{\mathcal{P}}$  in the boundary layer  $\{x \in B_r : \text{dist}(x, \partial B_r) \leq 5 \times 3^{-n} \sqrt{dr}\}$ .

We obtain the estimate

$$\begin{aligned} & (3^{-n}r)^{-1} \| \varphi \|_{L^2(I_{1,m} \times (\mathcal{C}_\infty \cap (\cup_{k=1}^{M_n} Q_{n,k}))})} \\ & \leq C(3^{-n}r)^{-1} \| [\varphi]_{\mathcal{P}} \mathbb{1}_{\{x \in B_r : \text{dist}(x, \partial B_r) \leq 5 \times 3^{-n} \sqrt{dr}\}} \|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ & \leq C \| \nabla [\varphi]_{\mathcal{P}} \mathbb{1}_{\{x \in B_r : \text{dist}(x, \partial B_r) \leq 5 \times 3^{-n} \sqrt{dr}\}} \|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ & \leq C \| \nabla [\varphi]_{\mathcal{P}} \|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))}. \end{aligned}$$

We put these estimates back into (82) and apply once again the Cauchy–Schwarz inequality. We notice that summing over the integers between  $n_0$  and  $n_1$  gives an additional error term of order  $\log^{\frac{1}{2}}(1+r)$ ,

$$\begin{aligned} & \left| \int_{(I_r \times \mathcal{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ (83) \quad & \leq C \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \left( \sum_{k=1}^{M_n} \| \nabla \bar{u} \|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2 \right)^{\frac{1}{2}} \| \nabla [\varphi]_{\mathcal{P}} \|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ & \leq C \log^{\frac{1}{2}}(1+r) \left( \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \| \nabla \bar{u} \|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2 \right)^{\frac{1}{2}} \| \nabla [\varphi]_{\mathcal{P}} \|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))}. \end{aligned}$$

We then estimate the norm  $\|\nabla[\varphi]_{\mathcal{P}}\|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))}$  thanks to the inequalities (41), (38) and the assumption  $r > \mathcal{M}_q(\mathcal{P})$ . We obtain

$$\frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|^{\frac{1}{2}}} \|\nabla[\varphi]_{\mathcal{P}}\|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))} \leq Cr^{\frac{2d-1}{2q}} \|\nabla\varphi\|_{L^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq Cr^{\frac{2d-1}{2q}}.$$

Moreover, by using the properties of the Whitney covering, we see that we can estimate the sum  $\sum_{i=0}^{\lfloor r^2-(r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \|\nabla\bar{u}\|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2$  by the  $L^2$ -norm of the function  $\nabla\bar{u}$  in a boundary layer of size  $6r'$  of the parabolic cylinder  $I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)$  (since every point in the ball  $B_r$  belongs to at most  $C(d)$  cubes of the form  $2Q_j$ ). More specifically, we have the estimate

$$\begin{aligned} & \left( \frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|} \sum_{m=0}^{\lfloor r^2-(r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \|\nabla\bar{u}\|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2 \right)^{\frac{1}{2}} \\ & \leq C \|\nabla\bar{u}\|_{L^2(I_r \times \{x \in B_r : \text{dist}(x, \partial B_r) \leq 6r'\})}. \end{aligned}$$

We then apply the Hölder's inequality and the global Meyers estimate (70) with the exponent  $\delta' = \min(\delta_0, \frac{1}{2}\delta)$ . We obtain

$$\begin{aligned} & \left( \frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|} \sum_{m=0}^{\lfloor r^2-(r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \|\nabla\bar{u}\|_{L^2(2I_{1,m} \times 2Q_{n,k})}^2 \right)^{\frac{1}{2}} \\ & \leq C \|\nabla\bar{u}\|_{L^2(I_r \times \{x \in B_r : \text{dist}(x, \partial B_r) \leq 6r'\})} \\ & \leq C \left(\frac{r'}{r}\right)^{\frac{\delta'}{4+2\delta'}} \|\nabla f\|_{L^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}. \end{aligned}$$

We conclude that

$$\begin{aligned} (84) \quad & \left| \frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{(I_r \times \mathcal{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla\bar{u})\varphi \right| \\ & \leq C \log^{\frac{1}{2}}(1+r)r^{\frac{2d-1}{2q}} \left(\frac{r'}{r}\right)^{\frac{\delta'}{4+2\delta'}} \|\nabla f\|_{L^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}. \end{aligned}$$

*Step 4.2: Control the weak norm over  $\Pi_2$ .* One can repeat all the arguments above to estimate the weak norm over the set  $\Pi_2$ , but we should pay attention to the decomposition over the time interval  $I_r$  since now the support of  $\Pi_2$  is close to the time boundary (see Figure 4). We define the time intervals

$$\forall m \in \mathbb{N}, \quad I_{2,m} := -r^2 + \left(\frac{2}{3}\right)^m (r')^2 + \left(-\frac{1}{3} \times \left(\frac{2}{3}\right)^m (r')^2, 0\right],$$

so that they satisfy  $2I_{2,m} \subseteq I_r$ . We can then apply the same arguments as in the estimates (83) and (84) to obtain the inequality

$$\begin{aligned} & \left| \frac{1}{|I_r \times (\mathcal{C}_\infty \cap B_r)|} \int_{(I_r \times \mathcal{C}_\infty) \cap \Pi_2} \nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla\bar{u})\varphi \right| \\ & \leq C \log^{\frac{1}{2}}(1+r)r^{\frac{2d-1}{2q}} \left(\frac{1}{r^2} \int_{-r^2}^{-r^2+2(r')^2} \|\nabla\bar{u}(t, \cdot)\|_{L^2(\text{Cv}(\mathbb{Z}^d \cap B_r))}^2 dt \right)^{\frac{1}{2}}. \end{aligned}$$

Then we apply Hölder’s inequality in the time variable and the estimate (70) to obtain

$$(85) \quad \begin{aligned} & \left( \frac{1}{r^2} \int_{-r^2}^{-r^2+2(r')^2} \|\nabla \bar{u}(t, \cdot)\|_{\underline{L}^2(\mathbb{C}_v(\mathbb{Z}^d \cap B_r))}^2 dt \right)^{\frac{1}{2}} \\ & \leq \left( \frac{r'}{r} \right)^{\frac{2\delta'}{4+2\delta'}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}. \end{aligned}$$

This gives an estimate for the weak norm of the map  $\nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})$  over the set  $\Pi_2$ . Finally, we combine the estimates (84) and (85) to conclude that

$$(86) \quad \begin{aligned} & \|\nabla \cdot (\bar{\sigma}^2(1 - \Upsilon)\nabla \bar{u})\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathcal{C}_\infty \cap B_r))} \\ & \leq C \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q}} \left( \frac{r'}{r} \right)^{\frac{\delta'}{4+2\delta'}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))}. \end{aligned}$$

*Step 5: Choice of the parameters  $q, \beta$  and conclusion.* We conclude the proof by combing the estimates (77), (78), (79), (80), (81), (86) and by choosing  $r' = r^{1-\beta}$  for some small exponent  $\beta \in (0, 1/2]$  to obtain

$$(87) \quad \frac{1}{r} \|u - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq C \mathcal{E}(r, \beta, q) \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))},$$

where the quantity  $\mathcal{E}(r, \beta, q)$  is defined by the formula

$$(88) \quad \mathcal{E}(r, \beta, q) := r^{-\frac{1}{2} + \beta(3 + \frac{d}{2})} + \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q} - \frac{\beta\delta'}{4+2\delta'}},$$

where we recall that  $\delta' = \min(\delta_0, \frac{1}{2}\delta)$  and that  $\delta_0$  is the exponent given by the Meyers estimate stated in (68).

It remains to select a value for the exponents  $\beta$  and  $q$ . We first choose the value of the exponent  $\beta$  and set  $\beta := \frac{1}{12+2d}$  so that the first term in the right-hand side of (88) is equal to  $r^{-\frac{1}{4}}$ . Then we set  $q := \frac{(12+2d)(2d-1)(4+2\delta')}{\delta'}$  so that  $\frac{2d-1}{2q} = \frac{\beta\delta'}{2(4+2\delta')}$ . With this choice, the second term in the right-hand side of (88) is equal to  $\log^{\frac{1}{2}}(1+r) r^{-\frac{\beta\delta'}{8+4\delta'}}$ .

We obtain that Theorem 3.2 holds with the exponent  $\alpha := \frac{\delta'}{(12+2d)(16+8\delta')} > 0$ . The proof is complete.  $\square$

**3.2.2. Large-scale  $C^{0,1}$ -regularity estimate.** The objective of this section is to prove the following  $C^{0,1}$ -large-scale regularity estimate for  $\mathbf{a}$ -caloric functions on the infinite cluster.

**PROPOSITION 9.** *There exist a constant  $C := C(d, \lambda, \mathbf{p}) < \infty$ , an exponent  $s := s(d, \lambda, \mathbf{p}) > 0$  such that for each point  $y \in \mathbb{Z}^d$ , there exists a nonnegative random variable  $\mathcal{M}_{C^{0,1}\text{-reg}}(y)$  satisfying*

$$(89) \quad \mathcal{M}_{C^{0,1}\text{-reg}}(y) \leq \mathcal{O}_s(C)$$

such that, for every  $r \geq \mathcal{M}_{C^{0,1}\text{-reg}}(y)$ , and every weak solution  $u \in H_{\text{par}}^1(I_r \times (\mathcal{C}_\infty \cap B_R(y)))$  of the equation

$$\partial_t u - \nabla \cdot (\mathbf{a}\nabla u) = 0 \quad \text{in } I_r \times (\mathcal{C}_\infty \cap B_R(y)),$$

one has the estimate, for every radius  $r \in [\mathcal{M}_{C^{0,1}\text{-reg}}(y), R]$ ,

$$(90) \quad \sup_{t \in I_r} \|\nabla u(t, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r(y))} \leq \frac{C}{R} \|[u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_r \times B_R(y)}\|_{\underline{L}^2(I_r \times B_R(y))}.$$

REMARK 14. The right-hand side of the estimate involves the coarsened function  $[u]_{\mathcal{P}}$  and we do not try to remove the coarsening to obtain a result of the form

$$(91) \quad \sup_{t \in I_r} \|\nabla u(t, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r(y))} \leq \frac{C}{R} \|u - (u)_{I_R \times (\mathcal{C}_\infty \cap B_R(y))}\|_{\underline{L}^2(I_R \times (\mathcal{C}_\infty \cap B_R(y)))},$$

even though such a result would be more natural and should be provable. There are two reasons motivating this choice. First, the estimate (90) involving the coarsening is simpler to prove than the inequality (91) and this choice reduces the amount of technicalities in the proof. Second, the objective of this section is to prove the Lipschitz regularity on the heat-kernel stated in Theorem 1.3 and the estimate (90) is sufficient in this regard.

This proposition proves that there exists a large random scale above which one has a good control on the gradient of  $\mathbf{a}$ -caloric functions. Such result belongs to the theory of large-scale regularity which is an important aspect of stochastic homogenization. The result presented above is a percolation version of a known result in the uniformly elliptic setting (see [15], Theorem 8.7) and can be considered a first step toward the establishment of a general large-scale regularity theory for the parabolic problem on the infinite percolation cluster.

We do not establish such a general theory here but we believe that it should follow from similar arguments: in the elliptic setting a general large-scale regularity theory was established in [12] and the generalization to the parabolic setting should be achievable. The reason justifying this choice is that our objective is to prove an estimate on the gradient of the Green's function (Theorem 1.3) and we do not need the full strength of the large-scale regularity theory to prove this result.

The main idea of the proof is that, thanks to Theorem 3.2, an  $\mathbf{a}$ -caloric function is well approximated by a  $\bar{\sigma}^2$ -caloric function. It is then possible to transfer the regularity known for  $\bar{\sigma}^2$ -caloric functions to  $\mathbf{a}$ -caloric functions following the classical ideas of the regularity theory. Such result can only hold when the  $\mathbf{a}$ -caloric function is well approximated by a  $\bar{\sigma}^2$ -caloric function which, according to Theorem 3.2, only holds on large scales.

This strategy has been carried out in [15] and is summarized in the following lemma, for which we refer to [15], Lemma 8.9.

LEMMA 3.5 (Lemma 8.9 of [15]). *Fix an exponent  $\beta \in (0, 1]$ ,  $k \geq 1$  and  $X \geq 1$ . Let  $R \geq 4X$  and  $v \in L^2(I_R \times B_R)$  have the property that, for every  $r \in [X, \frac{1}{4}R]$ , there exists a function  $w \in H_{\text{par}}^1(I_r \times B_r)$  which is a weak solution of*

$$(92) \quad \partial_t w - \frac{\bar{\sigma}^2}{2} \Delta w = 0 \quad \text{in } I_r \times B_r,$$

satisfying

$$\|v - w\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leq Kr^{-\beta} \|v - (v)_{I_{4r} \times B_{4r}}\|_{\underline{L}^2(I_{4r} \times B_{4r})}.$$

Then there exists a constant  $C := C(\beta, K, d, \lambda) < \infty$  such that for every radius  $r \in [X, R]$ ,

$$\frac{1}{r} \|v - (v)_{I_r \times B_r}\|_{\underline{L}^2(I_r \times B_r)} \leq \frac{C}{R} \|v - (v)_{I_R \times B_R}\|_{\underline{L}^2(I_R \times B_R)}.$$

To prove Proposition 9, we apply the previous lemma and combine it with Theorem 3.2. One has to face the following difficulty: we want to apply the previous result in the setting of percolation where the functions are only defined on the infinite cluster and not on  $\mathbb{R}^d$  as in the statement of Lemma 3.5.

To overcome this issue, the idea is to use the partition  $\mathcal{P}$  to extend the function  $u$ , using the definition of the coarsened map  $[u]_{\mathcal{P}}$  stated in (39). The strategy of the proof is then the following:

1. Proving that the function  $[u]_{\mathcal{P}}$  is a good approximation of the function  $u$ . In particular, we wish to prove that if the map  $u$  is well approximated by a  $\bar{\sigma}^2$ -caloric function, then the map  $[u]_{\mathcal{P}}$  is also well approximated by a  $\bar{\sigma}^2$ -caloric function.

2. Apply Lemma 3.5 to the function  $[u]_{\mathcal{P}}$  to obtain a large-scale  $C^{0,1}$ -regularity estimate for this map.

3. Transfer the result from the function  $[u]_{\mathcal{P}}$  to the function  $u$ .

The details are carried out in the following proof.

PROOF. Using the translation invariance of the model, we can assume without loss of generality that  $y = 0$ . We also let  $\delta_0$  be the exponent which appears in the Meyers estimate stated in Proposition 5 and consider the minimal scale  $\mathcal{M}_{\text{Meyers}}(0)$  given by Proposition 5. We consider the minimal scale  $\mathcal{M}_{\text{hom},\delta_0}(0)$  given by Theorem 3.2 and we let  $\alpha$  be the exponent which appears in the estimate (60). We set  $q := \max(\frac{d}{\alpha}, 4d)$  and let  $\mathcal{M}_q(\mathcal{P})$  be a minimal scale provided by Proposition 2. In particular, one has, for each radius  $r \geq \mathcal{M}_q(\mathcal{P})$ ,

$$(93) \quad r^{-d} \sum_{x \in \mathbb{Z}^d \cap B_r} \text{size}(\square_{\mathcal{P}}(x))^q \leq C \quad \text{and} \quad \sup_{x \in \mathbb{Z}^d \cap B_r} \text{size}(\square_{\mathcal{P}}(x)) \leq r^{\frac{1}{q}}.$$

The reasons justifying the choice of the exponents  $q$  will be clear later in the proofs. By Proposition 2, one knows that the minimal scale  $\mathcal{M}_q(\mathcal{P})$  satisfies the stochastic integrability estimate

$$\mathcal{M}_q(\mathcal{P}) \leq \mathcal{O}_s(C).$$

We then let  $\mathcal{M}_{C^{0,1}\text{-reg}}(0)$  be the minimal scale defined by the formula

$$\mathcal{M}_{C^{0,1}\text{-reg}}(0) := \max(\mathcal{M}_q(\mathcal{P}), \mathcal{M}_{\text{Meyers}}(0), \mathcal{M}_{\text{hom},\delta_0}(0)).$$

In the rest of the proof, we assume that the radii  $r$  and  $R$  are always larger than this minimal scale. We also note that, under the assumption  $r \geq \mathcal{M}_{C^{0,1}\text{-reg}}(0)$ , we can compare the volume of the ball  $B_r$  and the cardinality of  $\mathcal{C}_{\infty} \cap B_r$ , and we have the estimate

$$cr^d \leq |\mathcal{C}_{\infty} \cap B_r| \leq Cr^d.$$

This is a consequence of the assumption  $\mathcal{M}_{C^{0,1}\text{-reg}}(0) \geq \mathcal{M}_{\text{hom},\delta_0}(0)$  and the estimate (61).

We apply Theorem 3.2 to the function  $u$  on the parabolic cylinder  $I_r \times (\mathcal{C}_{\infty} \cap B_r)$ , with the boundary condition  $f = u$  and with the exponent  $\delta_0$ ; this proves that there exists a function  $\bar{u} \in H_{\text{par}}^1(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))$ , which is a solution of the equation (92), such that

$$(94) \quad \frac{1}{r} \|u - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_{\infty} \cap B_r))} \leq Cr^{-\alpha} \|\nabla u\|_{\underline{L}^{2+\delta_0}(I_r \times (\mathcal{C}_{\infty} \cap B_r))}.$$

We split the proof into 4 steps. In the first two steps, we prove that we can apply Lemma 3.5 with the coarsened function  $[u]_{\mathcal{P}}$ , we then post-process the result in Steps 3 and 4 and deduce Proposition 9.

*Step 1.* In this step, we post-process the result of Theorem 3.2: in the statement of the estimate (94), the right-hand side is expressed with an  $L^{2+\delta_0}$ -norm, for some small strictly positive exponent  $\delta_0$ . The goal of this step is to remove this additional assumption. To this end, we use the assumption  $r \geq \mathcal{M}_{C^{0,1}\text{-reg}}(0) \geq \mathcal{M}_{\text{Meyers}}(0)$ , which implies

$$\|\nabla u\|_{\underline{L}^{2+\delta_0}(I_r \times (\mathcal{C}_{\infty} \cap B_r))} \leq C \|\nabla u\|_{\underline{L}^2(I_{2r} \times (\mathcal{C}_{\infty} \cap B_{2r}))}.$$

We then apply the parabolic Caccioppoli inequality, which is stated in Proposition 4, and reads

$$\|\nabla u\|_{\underline{L}^2(I_{2r} \times (\mathcal{C}_{\infty} \cap B_{2r}))} \leq \frac{C}{r} \|u - (u)_{I_{4r} \times (\mathcal{C}_{\infty} \cap B_{4r})}\|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_{\infty} \cap B_{4r}))}.$$

Combining the two previous displays with the inequality (94) shows

$$(95) \quad \|u - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq Cr^{-\alpha} \|u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})}\|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))},$$

and Step 1 is complete.

*Step 2.* The goal of this step is to prove that the  $L^2$ -norm of the difference  $[u]_{\mathcal{P}} - \bar{u}$  on the continuous parabolic cylinder  $I_{r/2} \times B_{r/2}$  is small: we prove that there exists an exponent  $\beta > 0$  such that

$$\|[u]_{\mathcal{P}} - \bar{u}\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leq Cr^{-\beta} \|[u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_{4r} \times B_{4r}}\|_{\underline{L}^2(I_{4r} \times B_{4r})}.$$

The proof of this inequality relies on the estimate (95), which establishes that  $\bar{u}$  is a good approximation of  $u$  on the infinite cluster, together with the following parabolic regularity result: since  $\bar{u}$  is  $\bar{\sigma}^2$ -caloric on the parabolic cylinder  $I_r \times B_r$ , one has the estimate

$$(96) \quad \|\nabla \bar{u}\|_{L^\infty(I_{r/2} \times B_{r/2})} \leq Cr^{-1} \|\bar{u} - (\bar{u})_{I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)}\|_{\underline{L}^2(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))}.$$

We then consider a (continuous) triadic cube  $\square$  of the partition  $\mathcal{P}$  such that  $\square \cap B_{r/2} \neq \emptyset$ . By definition of the coarsening stated in (40), one sees that, for each time  $t \in I_{r/2}$ ,

$$\begin{aligned} & \|[u]_{\mathcal{P}}(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\square)} \\ & \leq \|[u]_{\mathcal{P}}(t, \cdot) - [\bar{u}]_{\mathcal{P}}(t, \cdot)\|_{L^\infty(\square)} + \|[\bar{u}]_{\mathcal{P}}(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\mathcal{C}_\infty \cap \square)} \\ & = \|[u - \bar{u}]_{\mathcal{P}}(t, \cdot)\|_{L^\infty(\square)} + \|[\bar{u}]_{\mathcal{P}}(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\mathcal{C}_\infty \cap \square)}. \end{aligned}$$

We then note that, by definition of the coarsening stated in (40), the  $L^\infty$ -norm of the function  $[u - \bar{u}]_{\mathcal{P}}$  is smaller than the  $L^\infty$ -norm of the function  $u - \bar{u}$ . Combining this observation with the estimate (42), we obtain

$$\|[u]_{\mathcal{P}}(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\square)} \leq \|u(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\mathcal{C}_\infty \cap \square)} + \text{size}(\square) \|\nabla \bar{u}(t, \cdot)\|_{L^\infty(\hat{\square})},$$

where the set  $\hat{\square}$  stands for the union of the cube  $\square$  and all its neighbors in the partition  $\mathcal{P}$ . We use the  $L^\infty - L^2$  estimate, valid in the discrete setting,

$$\|u(t, \cdot) - \bar{u}(t, \cdot)\|_{L^\infty(\mathcal{C}_\infty \cap \square)} \leq \|u(t, \cdot) - \bar{u}(t, \cdot)\|_{L^2(\mathcal{C}_\infty \cap \square)},$$

together with the regularity estimate (96) and the estimate (93) on the sizes of the cubes of the partition  $\mathcal{P}$  to deduce

$$\begin{aligned} \|[u]_{\mathcal{P}} - \bar{u}\|_{L^2(I_{r/2} \times \square)} & \leq \text{size}(\square)^{d/2} \|u - \bar{u}\|_{L^2(I_{r/2} \times (\mathcal{C}_\infty \cap \square))} \\ & \quad + C \text{size}(\square)^{1+d/2} r^{-1} \|\bar{u} - (\bar{u})_{I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)}\|_{\underline{L}^2(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))}. \end{aligned}$$

We then use the bounds (93) to estimate the size of the cube  $\square$  and sum over all the cubes of the partition  $\mathcal{P}$  which intersect the ball  $B_{r/2}$ , and use the estimate  $1 + \frac{d}{2} \leq d$ ,

$$\begin{aligned} \|[u]_{\mathcal{P}} - \bar{u}\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} & \leq Cr^{\frac{\alpha}{2}} \|u - \bar{u}\|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \\ & \quad + Cr^{-1+\alpha} \|\bar{u} - (\bar{u})_{I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)}\|_{\underline{L}^2(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))}. \end{aligned}$$

An estimate on the first term on the right-hand side is provided by the estimate (95). For the second term on the right-hand side, we apply the Poincaré inequality ([11], Corollary 3.4), use the estimate (70) for  $f = u$ , and apply Proposition 4 (the parabolic Caccioppoli inequality) and Proposition 5 (the interior Meyers estimate)

$$\begin{aligned} \|\bar{u} - (\bar{u})_{I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r)}\|_{\underline{L}^2(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))} & \leq r \|\nabla \bar{u}\|_{\underline{L}^2(I_r \times \text{Cv}(\mathbb{Z}^d \cap B_r))} \\ & \leq r \|\nabla u\|_{\underline{L}^{2+\delta}(I_r \times (\mathcal{C}_\infty \cap B_r))} \\ & \leq Cr \|\nabla u\|_{\underline{L}^2(I_{2r} \times (\mathcal{C}_\infty \cap B_{2r}))} \\ & \leq C \|u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})}\|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))}. \end{aligned}$$

Thus, we combine the two estimates and obtain

$$(97) \quad \begin{aligned} \| [u]_{\mathcal{P}} - \bar{u} \|_{\underline{L}^2(I_{r/2} \times B_{r/2})} &\leq r^{\frac{\alpha}{2}} r^{-\alpha} \| u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})} \|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))} \\ &\quad + Cr^{-1+\alpha} \| u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})} \|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))}. \end{aligned}$$

We then set the value  $\beta := \frac{\alpha}{2}$ . Since, by the identity (88), the value of the exponent  $\alpha$  is smaller than  $\frac{1}{2}$ , we have the inequality  $1 - \alpha \geq \frac{\alpha}{2}$ . This implies

$$\| [u]_{\mathcal{P}} - \bar{u} \|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leq Cr^{-\beta} \| u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})} \|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))}.$$

By definition of the coarsened function  $[u]_{\mathcal{P}}$ , we also have

$$(98) \quad \begin{aligned} \| u - (u)_{I_{4r} \times (\mathcal{C}_\infty \cap B_{4r})} \|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))} &\leq \| u - ([u]_{\mathcal{P}})_{I_{4r} \times B_{4r}} \|_{\underline{L}^2(I_{4r} \times (\mathcal{C}_\infty \cap B_{4r}))} \\ &\leq C \| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_{4r} \times B_{4r}} \|_{\underline{L}^2(I_{4r} \times B_{4r})}. \end{aligned}$$

The proof of Step 2 is complete.

*Step 3.* In the two previous steps, we proved that the coarsened function  $[u]_{\mathcal{P}}$  satisfies the assumption of Lemma 3.5, with the choice  $X = \mathcal{M}_{C^{0,1}\text{-reg}}(0)$ . We consequently apply the lemma and obtain that there exists a constant  $C := C(d, \mathfrak{p}, \lambda) < \infty$  such that, for every pair of radii  $r, R$  satisfying  $R \geq r \geq \mathcal{M}_{C^{0,1}\text{-reg}}(0)$ ,

$$(99) \quad \frac{1}{r} \| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_r \times B_r} \|_{\underline{L}^2(I_r \times B_r)} \leq \frac{C}{R} \| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R} \|_{\underline{L}^2(I_R \times B_R)}.$$

Then we apply once again the estimate (98) for the left-hand side of (99) in  $I_r \times (\mathcal{C}_\infty \cap B_r)$  and we obtain

$$(100) \quad \frac{1}{r} \| u - (u)_{I_r \times (\mathcal{C}_\infty \cap B_r)} \|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))} \leq \frac{C}{R} \| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R} \|_{\underline{L}^2(I_R \times B_R)}.$$

*Step 4.* In this final step, we upgrade the large-scale  $C^{0,1}$ -regularity estimate into the estimate (90). The strategy is to use an  $L_t^\infty L_x^2$  regularity estimate which is valid for the  $\mathbf{a}$ -caloric functions since the environment  $\mathbf{a}$  is assumed to be time independent. This result is stated in Lemma 2.2 and we apply it to the function  $u$  to obtain, for each  $r \geq 1$ ,

$$\sup_{t \in I_{r/2}} \| \nabla u(t, \cdot) \|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{r/2})} \leq \frac{C}{r} \| u - (u)_{I_r \times (\mathcal{C}_\infty \cap B_r)} \|_{\underline{L}^2(I_r \times (\mathcal{C}_\infty \cap B_r))}.$$

Combining this result with (100) completes the proof of Proposition 9 with the radius  $\frac{r}{2}$  instead of  $r$ ; this is a minor difference which can be fixed by standard arguments.  $\square$

**3.2.3. Decay of the gradient of the heat kernel.** The objective of this section is to post-process the regularity theory established in Proposition 9 and to apply it to the heat kernel. Together with Theorem 3.1, we deduce Theorem 1.3.

**PROOF OF THEOREM 1.3.** We fix a time  $t \in (0, \infty)$ , two points  $x, y \in \mathbb{Z}^d$  and work on the event  $\{y \in \mathcal{C}_\infty\}$ . We let  $\mathcal{M}_{C^{0,1}\text{-reg}}(x)$  be the minimal scale provided by Proposition 9 and, for  $z \in \mathbb{Z}^d$ , we let  $\mathcal{T}_{\text{NA}}(z)$  be the minimal time provided by Theorem 3.1. We first define the minimal time  $\mathcal{T}'_{\text{NA}}(x)$  by the formula

$$(101) \quad \mathcal{T}'_{\text{NA}}(x) := \sup\{t \in [1, \infty) : \exists z \in B_t(x) \text{ such that } \mathcal{T}_{\text{NA}}(z) \geq t\},$$

so that for every time  $t$  larger than this minimal time, every point  $z \in \mathcal{C}_\infty \cap B_t(x)$ , and every point  $z' \in \mathcal{C}_\infty$ , one has the estimate

$$p(t, z, z') \leq \Phi_C(t, z - z').$$

By using the symmetry of the heat kernel, we also have, for each time  $t \geq \mathcal{T}'_{\text{NA}}(x)$ , for each point  $z \in \mathcal{C}_\infty \cap B_t(x)$ , and each point  $z' \in \mathcal{C}_\infty$ ,

$$(102) \quad p(t, z', z) \leq \Phi_C(t, z - z').$$

Additionally, we claim that this minimal time satisfies the stochastic integrability estimate

$$(103) \quad \mathcal{T}'_{\text{NA}}(x) \leq \mathcal{O}_s(C).$$

The proof of the estimate (103) relies on an application of Lemma 1.4 by choosing

$$X_n := \sup_{z \in \mathbb{Z}^d \cap B_{3^n}(x)} 3^{-n} \mathcal{T}'_{\text{NA}}(z),$$

and we refer to the computation (46) for the details of the argument. We then define the minimal scale

$$(104) \quad \mathcal{M}_{\text{reg}}(x) := \max(\mathcal{M}_{C^{0.1-\text{reg}}}(x), \sqrt{\mathcal{T}'_{\text{NA}}(x)}).$$

Using the definition of the  $\mathcal{O}_s$  notation, the stochastic integrability estimates (89) and (103), one has, by reducing the size of the exponent  $s$  if necessary,

$$\mathcal{M}_{\text{reg}}(x) \leq \mathcal{O}_s(C).$$

In particular the tail of the random variable  $\mathcal{M}_{\text{reg}}(x)$  satisfies the inequality (20). We define, for  $\tau \in [-t, \infty)$  and  $z \in \mathcal{C}_\infty$ ,

$$u(\tau, z) := p(t + 1 + \tau, z, y).$$

We let  $R := \frac{\sqrt{t}}{2}$  and note that the function  $u$  is solution of the parabolic equation on the cylinder  $I_R \times (\mathcal{C}_\infty \cap B_R(x))$ . Applying Proposition 9 with the values  $r, R$  and using the assumption  $R \geq r \geq \mathcal{M}_{\text{reg}}(x)$ , we obtain

$$(105) \quad \sup_{\tau \in I_r} \|\nabla u(\tau, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r(x))} \leq \frac{C}{t^{1/2}} \|[u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R(x)}\|_{\underline{L}^2(I_R \times B_R(x))}.$$

We then note that the assumption  $R \geq \mathcal{M}_{\text{reg}}(x)$  implies  $\frac{1}{4} \geq \mathcal{T}'_{\text{NA}}(x)$  and allows to apply the estimate (102) to bound the right-hand side of the previous display. This shows

$$\|[u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R(x)}\|_{\underline{L}^2(I_R \times B_R(x))} \leq \Phi_C(t, y - x).$$

Combining the previous display with (105), considering the specific value  $\tau = -1$ , and increasing the value of the constant  $C$  shows

$$\|\nabla p(t, \cdot, y)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_r(x))} \leq t^{-1/2} \Phi_C(t, y - x).$$

The proof of Theorem 1.3 is complete.  $\square$

**4. Quantitative homogenization of the heat kernel.** In this section, we establish the theorem of quantitative homogenization of the parabolic Green's function, Theorem 1.1. From now on, we fix a point  $y \in \mathbb{Z}^d$  and only work on the event  $\{y \in \mathcal{C}_\infty\}$ . The proof of this result relies on a two-scale expansion which takes the following form:

$$(106) \quad h(t, x, y) := \theta(p)^{-1} \left( \bar{p}(t, x - y) + \sum_{k=1}^d \mathcal{D}_{e_k} \bar{p}(t, x - y) \chi_{e_k}(x) \right),$$

where the correctors  $\chi_{e_k}$  are normalized around the point  $y$ , following the procedure described after Proposition 6.

As is common for two-scale expansions, the proof relies on two important ingredients:

- The sublinearity of the corrector, stated in Proposition 6;
- The sublinearity of the flux stated in Proposition 7.

The analysis also requires the estimates on the parabolic Green’s function and its gradient provided by Theorem 3.1 and Theorem 1.3. We now present a sketch of the proof of Theorem 1.1. The result can be rewritten by using the  $\Phi_C$  notation: for each exponent  $\delta > 0$  and each  $x \in \mathcal{C}_\infty$ ,

$$(107) \quad |p(t, x, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - y)| \leq t^{-\frac{1}{2} + \delta} \Phi_C(t, x - y), \quad \sqrt{t} \geq \mathcal{T}_{\text{par}, \delta}(y).$$

To prove this result, we first prove a weighted  $L^2$ -estimate (see (52) for the definition of  $\Psi_C$ )

$$(108) \quad \|(p(t, \cdot, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, \cdot - y)) \exp(\Psi_C(t, |\cdot - y|))\|_{L^2(\mathcal{C}_\infty)} \leq Ct^{-\frac{d}{4} - \frac{1}{2} + \delta},$$

and deduce the estimate (107) thanks to the semigroup property, this is proved in Section 4.3. Proving the estimate (108) is the core of the proof. To this end, we need to introduce a mesoscopic time  $1 \ll \tau \ll t$  and a number of intermediate functions which are listed below:

- The two-scale expansion  $h$  defined in (106);
- The function  $q := q(\cdot, \cdot, \tau, y)$  introduced in (109);
- The function  $v := v(\cdot, \cdot, \tau, y)$  introduced in (123);
- The function  $w$  defined by the formula  $w := h - v - q$ .

The idea is to use these functions to split the difference

$$p(t, x, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - y) = (p(t, x, y) - q(t, x, \tau, y)) - v(t, x, \tau, y) - w(t, x, \tau, y) + (h(t, x, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - y)),$$

and then to prove that the  $L^2$ -norm of each of the terms is smaller than  $t^{-\frac{d}{4} - \frac{1}{2} + \delta}$ . More specifically, we organize the proof as follows:

1. in Lemma 4.1, we prove that the term corresponding to the difference  $(p - q)$  is small;
2. in Lemma 4.2, we prove that the term corresponding to the function  $v$  is small;
3. in Proposition 10, we prove that the term corresponding to the function  $w$  is small;
4. the term  $(h(t, x, y) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - y))$  is proved to be small by using the sublinearity of the corrector, the proof is straightforward and not stated in a specific lemma.

The rest of this section is organized as follows. Section 4.1 is devoted to the proof of Lemmas 4.1 and 4.2. Section 4.2 is devoted to the proof of Proposition 10 and is the core of the analysis: we make use of the regularization Lemmas proved in Section 4.1 as well as the various results recorded in the previous sections to perform the two-scale expansion. In Section 4.3, we post-process the results and prove the quantitative convergence of the heat kernel, Theorem 1.1.

4.1. *Two regularization steps.* We now introduce the function  $q$ . For a fixed initial time  $\tau > 0$  and a vertex  $y \in \mathcal{C}_\infty$ , we let  $(t, x) \mapsto q(t, x, \tau, y)$  be the solution of the parabolic problem

$$(109) \quad \begin{cases} \partial_t q - \nabla \cdot (\mathbf{a} \nabla q) = 0 & \text{in } (\tau, \infty) \times \mathcal{C}_\infty, \\ q(\tau, \cdot, \tau, y) = \theta(\mathbf{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{on } \mathcal{C}_\infty. \end{cases}$$

A reason justifying this construction is that the initial condition of the heat kernel  $p$ , which is a Dirac at  $y$ , is too singular and one cannot perform the two-scale expansion due to this lack of regularity. The idea is thus to replace the Dirac by a smoother function, the function  $\bar{p}(\tau, \cdot)$ , and to exploit its more favorable regularity properties to perform the two-scale expansion (see Section 4.2). For this strategy to work, one needs to choose the value of the time  $\tau$  to be both:

- Large enough so that so that the function  $\bar{p}(\tau, \cdot)$  is regular enough: in particular we want  $\tau \gg 1$ ;
- Small enough so that the function  $q$  is a good approximation of the heat kernel  $p$ : we want  $\tau \ll t$ .

The choice of  $\tau$  will be  $\tau := t^{1-\kappa}$ , for some small exponent  $\kappa$  whose value is decided at the end of the proof.

The following lemma states that the function  $q$  is a good approximation of the heat kernel  $p$ , when the coefficient  $\tau$  is chosen such that  $1 \ll \tau \ll t$ . In the following lemma, given an exponent  $\alpha > 0$ , we use the notation  $\mathcal{T}_{\text{dense},\alpha}(y)$  to denote the minimal time introduced in Proposition 13, above which the mass of the homogenized heat kernel  $\bar{p}$  is almost equal to the density of the infinite cluster  $\mathcal{C}_\infty$  up to an error of order  $t^{-\frac{1}{2}+\alpha}$ .

LEMMA 4.1. *For each exponent  $\alpha > 0$  and each vertex  $y \in \mathbb{Z}^d$ , we let  $\mathcal{T}_{\text{approx},\alpha}(y)$  be the minimal time defined by the formula*

$$\mathcal{T}_{\text{approx},\alpha}(y) := \max(\mathcal{M}_{\text{reg}}(y)^2, \mathcal{T}_{\text{dense},\alpha}(y)).$$

*This random variable satisfies the stochastic integrability estimate*

$$\mathcal{T}_{\text{approx},\alpha}(y) \leq \mathcal{O}_s(C),$$

*and the following property: there exists a positive constant  $C := C(d, \mathbf{p}, \lambda, \alpha) < \infty$  such that, on the event  $\{y \in \mathcal{C}_\infty\}$ , for every pair of times  $t, \tau \in (0, \infty)$  such that  $t \geq 3\tau$  and  $\tau \geq \mathcal{T}_{\text{approx},\alpha}(y)$ , and for every  $x \in \mathcal{C}_\infty$ , one has*

$$(110) \quad |q(t, x, \tau, y) - p(t, x, y)| \leq \left( \left( \frac{\tau}{t} \right)^{\frac{1}{2}} + \tau^{-\frac{1}{2}+\alpha} \right) \Phi_C(t, x - y).$$

PROOF. Before starting the proof, we note that the assumptions of the lemma imply the following results:

- The two inequalities  $t \geq 3\tau$  and  $\tau \geq \mathcal{T}_{\text{approx},\alpha}(y)$  imply the estimates  $t \geq \mathcal{T}_{\text{approx},\alpha}(y)$  and  $t - \tau \geq \mathcal{T}_{\text{approx},\alpha}(y)$ ;
- By definition, the minimal time  $\mathcal{T}_{\text{approx},\alpha}(y)$  is larger than the minimal times  $\mathcal{T}_{\text{dense},\alpha}(y)$ ,  $\mathcal{T}'_{NA}(y)$  and the square of the minimal scales  $\mathcal{M}_{\text{Poinc}}(y)$  and  $\mathcal{M}_{\text{reg}}(y)$ . We can thus apply the corresponding results in the proof.

*Step 1: Set up.* We fix a vertex  $y \in \mathbb{Z}^d$  and work on the event  $\{y \in \mathcal{C}_\infty\}$ . We first record the following estimate: under the assumption  $\mathcal{T}_{\text{dense},\alpha}(y)$ , for each radius  $r \geq \sqrt{\tau}$ , we can compare the volume of the ball  $B_r(y)$  and the cardinality of the set  $\mathcal{C}_\infty \cap B_r(y)$ , and we have the estimate

$$cr^d \leq |\mathcal{C}_\infty \cap B_r(y)| \leq Cr^d.$$

We consider a point  $x$  in the infinite cluster  $\mathcal{C}_\infty$  and two times  $t, \tau > 0$  such that  $t \geq \tau \geq \mathcal{T}_{\text{approx},\alpha}(y)$ . By Duhamel's principle, one has

$$(111) \quad q(t, x, \tau, y) - p(t, x, y) = \int_{\mathcal{C}_\infty} (\theta(\mathbf{p})^{-1} \bar{p}(\tau, z - y) - p(\tau, z, y)) p(t - \tau, x, z) dz.$$

Using the inequality  $\tau \geq \mathcal{T}_{\text{dense},\alpha}(y)$  and Proposition 13, we have the inequality

$$\left| \theta(\mathbf{p})^{-1} \int_{\mathcal{C}_\infty} \bar{p}(\tau, z - y) dz - 1 \right| \leq C\tau^{-\frac{1}{2}+\alpha}.$$

Since the mass of the transition kernel  $p(\tau, \cdot, y)$  on the infinite cluster is equal to 1, we deduce that

$$\left| \int_{\mathcal{C}_\infty} (\theta(\mathbf{p})^{-1} \bar{p}(\tau, z - y) - p(\tau, z, y)) dz \right| \leq C\tau^{-\frac{1}{2}+\alpha}.$$

We can thus subtract a constant term equal to  $(p(\tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}$  in the right-hand side of (111) up to a small cost of order  $\tau^{-\frac{1}{2}+\alpha}$ ,

$$\begin{aligned} & |q(t, x, \tau, y) - p(t, x, y)| \\ & \leq \left| \int_{\mathcal{C}_\infty} (\theta(\mathbf{p})^{-1} \bar{p}(\tau, z - y) - p(\tau, z, y))(p(t - \tau, x, z) \right. \\ & \quad \left. - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}) dz \right| \\ & \quad + C\tau^{-\frac{1}{2}+\alpha} |(p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}|. \end{aligned}$$

Using the inequality  $t - \tau \geq \mathcal{T}'_{NA}(y)$  and (101), we apply Theorem 3.1 to obtain

$$\begin{aligned} & |q(t, x, \tau, y) - p(t, x, y)| \\ (112) \quad & \leq \int_{\mathcal{C}_\infty} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ & \quad + \tau^{-\frac{1}{2}+\alpha} \Phi_C(t, x - y). \end{aligned}$$

We then treat the first term on the right-hand side of (112). The strategy is to split the integral into scales: for each integer  $n \geq 1$ , we let  $A_n$  be the dyadic annulus defined by

$$A_n := \{z \in \mathbb{Z}^d : 2^n \sqrt{\tau} \leq |z - y| < 2^{n+1} \sqrt{\tau}\},$$

and we compute

$$\begin{aligned} & \int_{\mathcal{C}_\infty} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ (113) \quad & = \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ & \quad + \sum_{n=0}^\infty \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz. \end{aligned}$$

*Step 2: Multiscale analysis in the ball  $B_{\sqrt{\tau}}(y)$ .* The term pertaining to small scales  $B_{\sqrt{\tau}}(y)$  can be estimated thanks to the estimate  $\Phi_C(\tau, z - y) \leq C\tau^{-d/2}$  and the Poincaré inequality. This latter inequality can be applied since we assumed  $\tau \geq \mathcal{M}_{\text{Poinc}}(y)^2$ . This gives

$$\begin{aligned} & \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ (114) \quad & \leq C \|p(t - \tau, x, \cdot) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y))} \\ & \leq C\sqrt{\tau} \|\nabla_y p(t - \tau, x, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y))}, \end{aligned}$$

where the notation  $\nabla_y$  means that the gradient is on the second spatial variable. We now estimate the term on the right-hand side thanks to Theorem 1.3, or more precisely Remark 3,

which can be applied since we assumed  $t - \tau \geq \mathcal{M}_{\text{reg}}(y)^2$ . We have

$$\begin{aligned} \|\nabla p(t - \tau, x, \cdot)\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y))} &\leq C(t - \tau)^{-1/2} \Phi_C(t - \tau, x - y) \\ &\leq Ct^{-1/2} \Phi_C(t, x - y), \end{aligned}$$

where to go from the first line to the second one we used that  $t - \tau \geq \frac{2}{3}t$  and increased the value of the constant  $C$ .

A combination of the two previous displays shows

$$\begin{aligned} (115) \quad &\int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ &\leq C \left(\frac{\tau}{t}\right)^{1/2} \Phi_C(t, x - y). \end{aligned}$$

This completes the estimate of the term corresponding to the small scales in (113).

*Step 3: Multiscale analysis in the annuli  $A_n$ .* To estimate the terms corresponding to the dyadic annuli, we fix some integer  $n \in \mathbb{N}$  and study the integral in the region  $A_n$ ; thanks to the triangle inequality, we insert a constant term equal to  $(p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}$  in the integral,

$$\begin{aligned} (116) \quad &\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ &\leq \underbrace{\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}| dz}_{(116)\text{-a}} \\ &\quad + \underbrace{\frac{(\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) dz)}{|\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)|} \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}| dz}_{(116)\text{-b}}. \end{aligned}$$

*Step 3.1: Estimate for the term (116)-a in the annuli  $A_n$ .* We estimate (116)-a and distinguish three types of scales:

1. The small scales which are defined as the annuli  $A_n$  such that  $2^{n+2}\sqrt{\tau} \leq \sqrt{t}$ ;
2. The intermediate scales which are defined as the annuli  $A_n$  such that  $t \geq 2^{n+2}\sqrt{\tau} > \sqrt{t}$ ;
3. The large scales which are defined as the annuli  $A_n$  such that  $2^{n+2}\sqrt{\tau} > t$ .

The following estimate for the function  $\Phi_C$  is easy to check by its definition (50) and is used several times in the proof: for any constant  $C' > C$ , there exists a constant  $c > 0$  depending only on the values of  $C$  and  $C'$  such that

$$\left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leq e^{-c2^n\sqrt{\tau}} \left( \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y) \right).$$

This implies that for any positive integer  $k \in \mathbb{N}$ , there exists a constant  $C_k$  depending only on the integer  $k$  and the constant  $C$  such that

$$(117) \quad \left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leq (2^n\sqrt{\tau})^{-k} \left( \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C_k}(\tau, \cdot - y) \right).$$

*Step 3.1.1: Estimate for (116)-a in the small scales  $2^{n+2}\sqrt{\tau} \leq \sqrt{t}$ .* We first focus on the small scales and apply the Poincaré inequality. With a computation similar to (114), one can

estimate the first term in the right-hand side of (116)

$$\begin{aligned} & \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}}| dz \\ & \leq \left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \\ & \quad \times \int_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}}| dz \\ & \leq \left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) C(2^{n+1}\sqrt{\tau})^{\frac{d}{2}+1} \|\nabla_y p(t - \tau, x, \cdot)\|_{L^2(\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)})}. \end{aligned}$$

Using the assumption  $t - \tau \geq \mathcal{M}_{\text{reg}}^2(y)$ , we apply Theorem 1.3. This shows

$$\|\nabla_y p(t - \tau, x, \cdot)\|_{L^2(\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)})} \leq Ct^{-1/2}(2^{n+1}\sqrt{\tau})^{\frac{d}{2}} \Phi_C(t, x - y).$$

Using the explicit formula for the function  $\Phi_C$  stated in (50), one has the estimate

$$\left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) (2^{n+1}\sqrt{\tau})^d \leq C2^{-2n}.$$

Combining the three previous displays shows, for each integer  $n \in \mathbb{N}$  such that  $2^{n+2}\sqrt{\tau} \leq \sqrt{t}$ ,

$$\begin{aligned} & \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}}| dz \\ (118) \quad & \leq C2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y). \end{aligned}$$

*Step 3.1.2: Estimate for (116)-a in the intermediate scales  $t \geq 2^{n+2}\sqrt{\tau} > \sqrt{t}$ .* We now treat the case of the intermediate scales. In this case, we have  $2^{n+2} \geq \sqrt{\frac{t}{\tau}}$ , thus by the estimate (117) for  $k = 2$ , one has

$$(119) \quad \left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leq 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \left( \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C_2}(\tau, \cdot - y) \right).$$

Using the assumptions  $t - \tau \geq \mathcal{T}'_{\text{NA}}(y)$ , we can apply Theorem 3.1, the previous estimate (119) and the convolution property (51) for the map  $\Phi_C$ . We obtain

$$\begin{aligned} & \left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \\ & \quad \times \int_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}}| dz \\ (120) \quad & \leq 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \left( \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C_2}(\tau, \cdot - y) \right) \int_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}} \Phi_C(t - \tau, x - z) dz \\ & \leq 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \int_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}(y)}} \Phi_{C_2}(\tau, y - z) \Phi_C(t - \tau, x - z) dz \\ & \leq 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y), \end{aligned}$$

by increasing the value of the constant  $C$  and using (51) in the last line.

*Step 3.1.3: Estimate for the term (116)-a in the large scales  $2^{n+2}\sqrt{\tau} > t$ .* The computation is similar to the one performed in (120) up to two differences listed below:

1. For these scales, we cannot apply the Gaussian bounds on the heat kernel given by Theorem 3.1. Instead, we thus apply the Carne–Varopoulos bound which is stated in Proposition 8 and can be rewritten with the notation  $\Phi_C$ : for each  $x, z \in \mathcal{C}_\infty$ ,

$$p(t, x, z) \leq t^{\frac{d}{2}} \Phi_C(t, x - z);$$

2. We use the inequality  $2^{n+2}\sqrt{\tau} > t$  and the estimate (117) for  $k = \frac{d}{2} + 2$  to obtain the bound in the annulus  $A_n$

$$\left( \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leq 2^{-n} t^{-\frac{d}{2}} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \left( \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y) \right),$$

for some constant  $C' > C$ .

We can then perform the computation (120) and obtain the estimate

$$\begin{aligned} (121) \quad & \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{B_{2^{n+1}\sqrt{\tau}}(y)}| dz \\ & \leq C 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y). \end{aligned}$$

Combining the estimates (118), (120) and (121), we have obtained, for each integer  $n \in \mathbb{N}$ ,

$$\begin{aligned} (122) \quad & \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{B_{2^{n+1}\sqrt{\tau}}(y)}| dz \\ & \leq C 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y). \end{aligned}$$

*Step 3.2: Estimate for (116)-b in the annuli  $A_n$ .* The second term (116)-b can be estimated thanks to a similar strategy: we first apply the inequality (117) with  $k = d$ ,

$$\begin{aligned} & \frac{(\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) dz)}{|\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)|} \leq \frac{|\mathcal{C}_\infty \cap A_n|}{|\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)|} \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \\ & \leq 2^{dn} \sup_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \\ & \leq \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y), \end{aligned}$$

for some constant  $C' > C$ . Using this estimate, we deduce

$$\begin{aligned} & \frac{(\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) dz)}{|\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)|} \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}| dz \\ & \leq \inf_{\mathcal{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y) \\ & \quad \times \int_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}| dz. \end{aligned}$$

We can then apply the same proof as for the first term in the right-hand side of (116). This proves the inequality

$$\begin{aligned} & \frac{(\int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) dz)}{|\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)|} \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}| dz \\ & \leq C 2^{-n} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y). \end{aligned}$$

Combining the previous inequality with the estimate (122), we obtain, for each integer  $n \in \mathbb{N}$ ,

$$\begin{aligned} & \int_{\mathcal{C}_\infty \cap A_n} \Phi_C(\tau, z - y) |p(t - \tau, x, z) - (p(t - \tau, x, \cdot))_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)}| dz \\ & \leq C 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C(t, x - y). \end{aligned}$$

Combining this inequality with (112), (113), (115) and summing over the integer  $n \in \mathbb{N}$  completes the proof of Lemma 4.1.  $\square$

We now introduce a second intermediate function useful in the proof of Theorem 1.1, the function  $v$  which is defined as follows. For some fixed  $(\tau, y) \in (0, \infty) \times \mathcal{C}_\infty$ , we let  $(t, x) \mapsto v(t, x, \tau, y)$  be the solution of the parabolic equation

$$(123) \quad \begin{cases} \partial_t v(\cdot, \cdot, \tau, y) - \nabla \cdot \mathbf{a} \nabla v(\cdot, \cdot, \tau, y) = 0 & \text{in } (\tau, \infty) \times \mathcal{C}_\infty, \\ v(\tau, \cdot, \tau, y) = h(\tau, \cdot, y) - \theta(\mathbf{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{in } \mathcal{C}_\infty. \end{cases}$$

To define this function, we run the parabolic equation starting from time  $\tau$  and until time  $t$  with the initial condition given by the difference between the two-scale expansion  $h$  defined in (106) and the homogenized heat kernel  $\theta(\mathbf{p})^{-1} \bar{p}(\tau, \cdot - y)$ . By the sublinearity of the corrector stated in Proposition 6, we expect the function  $h(\tau, \cdot, y) - \theta(\mathbf{p})^{-1} \bar{p}(\tau, \cdot - y)$  to be small. The following proposition states that the solution of the parabolic equation with this initial condition remains small (in the sense of the inequality (124)).

LEMMA 4.2. *For any exponent  $\alpha > 0$ , there exists a finite positive constant  $C := C(d, \lambda, \mathbf{p}, \alpha)$  such that for each pair of times  $t, \tau \in (0, \infty)$  satisfying  $t \geq 3\tau$ ,  $(t - \tau) \geq \mathcal{T}'_{\text{NA}}(y)$ , and  $\sqrt{\tau} \geq \mathcal{M}_{\text{corr}, \alpha}(y)$ , the following estimate holds:*

$$(124) \quad |v(t, x, \tau, y)| \leq C \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \Phi_C(t, x - y).$$

PROOF. The proof relies on two main ingredients: the quantitative sublinearity of the corrector and an explicit formula for the function  $v$  in terms of the heat kernel  $p$ .

First, by the definition (123), we have the formula

$$\begin{aligned} v(t, x, \tau, y) &= \int_{\mathcal{C}_\infty} (h(\tau, z, y) - \theta(\mathbf{p})^{-1} \bar{p}(\tau, z - y)) p(t - \tau, x, z) dz \\ &= \int_{\mathcal{C}_\infty} \theta(\mathbf{p})^{-1} \left( \sum_{k=1}^d \mathcal{D}_{e_k} \bar{p}(\tau, z - y) \chi_{e_k}(z) \right) p(t - \tau, x, z) dz. \end{aligned}$$

We then apply the four following estimates:

1. The sublinearity of the corrector: under the assumption  $\sqrt{\tau} \geq \mathcal{M}_{\text{corr}, \alpha}(y)$  and the normalization convention chosen for the corrector in the definition of the two-scale expansion  $h$  stated in (106), one has, for each  $k \in \{1, \dots, d\}$ ,

$$|\chi_{e_k}(z)| \leq \begin{cases} \tau^{\frac{\alpha}{2}} & \text{if } |z - y| \leq \tau^{\frac{1}{2}}, \\ |z - y|^\alpha & \text{if } |z - y| \geq \tau^{\frac{1}{2}}; \end{cases}$$

2. The Gaussian bounds on the transition kernel  $p(t - \tau, x, \cdot)$ , valid under the assumptions  $\tau \geq \mathcal{T}'_{\text{NA}}(y)$  and  $t \geq 3\tau$ : for each  $z \in \mathcal{C}_\infty \cap B_{t-\tau}(y)$ ,

$$p(t - \tau, x, z) \leq \Phi_C(t - \tau, x - z) \leq \Phi_{C'}(t, x - z),$$

where the second inequality follows from the inequality  $t - \tau \geq \frac{2}{3}t$  (by increasing the value of the constant  $C$ );

3. The bound on the transition kernel  $p(t - \tau, x, z)$ : for any point  $z \in \mathcal{C}_\infty \setminus B_{t-\tau}(y)$ ,

$$p(t - \tau, x, z) \leq (t - \tau)^{d/2} \Phi_C(t - \tau, z - x),$$

which is a consequence of the definition of the map  $\Phi_C$  stated in (50), Proposition 8 and the assumption  $t \geq 3\tau$  (by increasing the value of the constant  $C$ );

4. The estimate on the homogenized heat kernel  $\bar{p}$ , which follows from standard results from the regularity theory,

$$|D_{e_k} \bar{p}(\tau, z - y)| \leq C \tau^{-\frac{1}{2}} \Phi_C(\tau, z - y).$$

We obtain the inequality

$$\begin{aligned} & |v(t, x, \tau, y)| \\ & \leq C \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z - y) \Phi_C(t - \tau, x - z) dz \\ (125) \quad & + C \tau^{-\frac{1}{2}} \int_{\mathcal{C}_\infty \cap (B_{t-\tau}(y) \setminus B_{\sqrt{\tau}}(y))} |z - y|^\alpha \Phi_C(\tau, z - y) \Phi_C(t - \tau, x - z) dz \\ & + C \tau^{-\frac{1}{2}} (t - \tau)^{d/2} \int_{\mathcal{C}_\infty \setminus B_{t-\tau}(y)} |z - y|^\alpha \Phi_C(\tau, z - y) \Phi_C(t - \tau, x - z) dz. \end{aligned}$$

We then estimate the three terms in the right-hand side of (125) separately. For the first term, we use the inequality (51) and obtain, for some constant  $C' > C$ ,

$$\begin{aligned} & \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \int_{\mathcal{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z - y) \Phi_C(t - \tau, x - z) dz \\ (126) \quad & \leq \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \int_{\mathbb{Z}^d} \Phi_C(\tau, z - y) \Phi_C(t - \tau, x - z) dz \\ & \leq \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \Phi_{C'}(t, x - y). \end{aligned}$$

To estimate the second term, we use the estimate (117) and deduce that

$$(127) \quad |z - y|^\alpha \Phi_C(\tau, z - y) \leq \tau^{\frac{\alpha}{2}} \Phi_{C'}(\tau, z - y),$$

for some constant  $C' > C$ . To estimate the third term in the right-hand side of (125), we use the estimate (117) again with the value  $k = d/2 + \alpha$  (the estimate applies even though  $k$  is not an integer). We obtain that for some constant  $C' > C$  and for any point  $z \in \mathcal{C}_\infty \setminus B_{t-\tau}(y)$ ,

$$(128) \quad (t - \tau)^{d/2} |z - y|^\alpha \Phi_C(\tau, z - y) \leq |z - y|^{d/2 + \alpha} \Phi_C(\tau, z - y) \leq \Phi_{C'}(\tau, z - y).$$

Finally, combining the identity (125) and the estimates (126), (127), (128), we obtain

$$|v(t, x, \tau, y)| \leq C \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \Phi_{C'}(t, x - y).$$

The proof of Lemma 4.2 is complete.  $\square$

4.2. *The two-scale expansion.* The main objective of this section is to prove that the weighted  $L^2$  norm of the function  $w$  is small in the sense of (129). Before starting the proof, we recall the notation for the function  $\Psi_C$  introduced in (52). We also recall the notation convention for discrete and continuous derivatives:

- In the proof of Proposition 10, the functions  $\bar{p}$ ,  $\psi$  and  $\eta$  are defined on  $\mathbb{R}^d$  and valued in  $\mathbb{R}$ , for these functions, we use the symbols  $\nabla$  and  $\Delta$  to denote respectively the continuous gradient and the continuous Laplacian. To refer to the discrete derivatives, we use the notation  $\mathcal{D}$ ,  $\mathcal{D}^*$ ,  $\mathcal{D}^2$ ,  $\mathcal{D}^3$ , etc.

- All the other functions are defined on the discrete lattice  $\mathbb{Z}^d$  or on the infinite cluster  $\mathcal{C}_\infty$ , for these functions, we use the notation  $\nabla$  to denote the discrete gradient defined on the edges and the notation  $\mathcal{D}$  for the discrete derivative defined on the vertices, following the conventions of Section 1.6.4.

PROPOSITION 10. *For every exponent  $\alpha > 0$ , there exists a positive constant  $C := C(d, p, \lambda, \alpha) < \infty$  such that for every point  $y \in \mathbb{Z}^d$  and every time  $t \in (0, \infty)$  such that  $\sqrt{t} \geq \max(\mathcal{M}_{\text{corr}, \alpha}(y), \mathcal{M}_{\text{flux}, \alpha}(y))$ , one has, on the event  $\{y \in \mathcal{C}_\infty\}$ ,*

$$(129) \quad \|w(t, \cdot, \tau, y) \exp(\Psi_C(t, |\cdot - y|))\|_{L^2(\mathcal{C}_\infty)} \leq C \left(\frac{t}{\tau}\right)^{\frac{1}{2}} \tau^{-\frac{d}{4} - \frac{1}{2} + \frac{\alpha}{2}}.$$

PROOF. The key is to develop a differential inequality for the function  $w$ . The proof is decomposed into five steps and is organized as follows. In Step 1, we use the explicit formula for  $w$  and apply the parabolic operator  $\partial_t - \nabla \cdot \mathbf{a} \nabla$  to the map  $w$  to obtain the formulas (130) and (131). In Step 2, we test the equation obtained in (130) with the function  $\psi w$ , where  $\psi$  is a map which is either equal to the constant 1 or equal to the function  $x \mapsto \exp(\Psi_C(t, |x - y|))$ . In the three remaining steps, we treat the different terms obtained and complete the proof of the estimate (129).

Step 1: *Establishing the equation for  $w$ .* We claim that the function  $w$  satisfies the equation

$$(130) \quad \begin{cases} \partial_t w(\cdot, \cdot, \tau, y) - \nabla \cdot \mathbf{a} \nabla w(\cdot, \cdot, \tau, y) \\ = f(\cdot, \cdot, y) + \mathcal{D}^* \cdot F(\cdot, \cdot, y) + \xi(\cdot, \cdot, y) & \text{in } (\tau, \infty) \times \mathcal{C}_\infty, \\ w(\tau, \cdot, y) = 0 & \text{in } \mathcal{C}_\infty, \end{cases}$$

where the three functions  $f : (0, \infty) \times \mathcal{C}_\infty \times \mathcal{C}_\infty \rightarrow \mathbb{R}$ ,  $F : (0, \infty) \times \mathcal{C}_\infty \times \mathcal{C}_\infty \rightarrow \mathbb{R}^d$  and  $\xi : (0, \infty) \times \mathcal{C}_\infty \times \mathcal{C}_\infty \rightarrow \mathbb{R}$  are defined by the formulas, for each  $(t, y) \in (0, \infty) \times \mathcal{C}_\infty$ ,

$$(131) \quad \begin{cases} f(t, \cdot, y) = \frac{1}{2} \bar{\sigma}^2 (\Delta \bar{p}(t, \cdot - y) - (-\mathcal{D}^* \cdot \mathcal{D} \bar{p}(t, \cdot - y))) \\ \quad + \sum_{k=1}^d (\partial_t \mathcal{D}_{e_k} \bar{p}(t, \cdot - y)) \chi_{e_k}(\cdot), \\ [F]_i(t, \cdot, y) = \sum_{k=1}^d [\mathbf{a} \mathcal{D} \mathcal{D}_{e_k} \bar{p}(t, \cdot - y)]_i T_{e_i}(\chi_{e_k})(\cdot), \quad \forall i \in \{1, \dots, d\}, \\ \xi(t, \cdot, y) = \sum_{k=1}^d \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \cdot \tilde{\mathbf{g}}_{e_k}^*(\cdot), \end{cases}$$

where  $\tilde{\mathbf{g}}_{e_k}^*$  is a translated version of the flux  $\tilde{\mathbf{g}}_{e_k}$  defined by the formula, for each  $x \in \mathcal{C}_\infty$ ,

$$\tilde{\mathbf{g}}_{e_k}^*(x) := \begin{pmatrix} T_{-e_1} \left[ \mathbf{a}(\mathcal{D} \chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k \right]_1 \\ \vdots \\ T_{-e_d} \left[ \mathbf{a}(\mathcal{D} \chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k \right]_d \end{pmatrix},$$

and we recall the notation  $[\mathbf{a}(\mathcal{D} \chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k]_i$  introduced in Section 1.6.1 to denote the  $i$ th-component of the vector  $\mathbf{a}(\mathcal{D} \chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k$ . In Appendix B, it is proved that the translated flux  $\tilde{\mathbf{g}}_{e_k}^*$  has similar properties as the centered flux  $\tilde{\mathbf{g}}_{e_k}$ . In particular, it is proved in Remark 20 that, for every radius  $r \geq \mathcal{M}_{\text{flux}, \alpha}(y)$ ,

$$\|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r(y))} \leq Cr^\alpha.$$

The proof follows from a direct calculation. Since one has the equality

$$(\partial_t - \nabla \cdot \mathbf{a}\nabla)(v + q)(\cdot, \cdot, s, y) = 0,$$

it suffices to focus on the term  $h$  in the definition of  $w$ , the details are left to the reader.

*Step 2: A differential inequality.* In this step and the rest of the proof, we let  $\psi$  be the function from  $(0, \infty) \times \mathbb{R}^d$  to  $\mathbb{R}$  which is either:

- (i) The constant function equal to 1;
- (ii) The function  $\exp(\Psi_C(t, |\cdot - y|))$ , for some large constant  $C > 0$ .

We note that in both cases, the function  $\psi$  is smooth and satisfies the following property:

$$(132) \quad \exists \tilde{C} > 1, \forall |h| \leq 1, \quad |\nabla \psi(\cdot + h)| \leq \tilde{C} |\nabla \psi|.$$

This estimate allows to replace the discrete derivative  $\mathcal{D}\psi$ , by the continuous derivative  $\nabla\psi$  by only paying a constant. The strategy of the proof is then to test the equation (130) with the function  $w\psi^2$  to derive a differential inequality. We first write

$$(133) \quad \underbrace{\int_{\mathcal{C}_\infty} ((\partial_t w - \nabla \cdot \mathbf{a}\nabla w)\psi^2 w)}_{\text{LHS}} = \underbrace{\int_{\mathcal{C}_\infty} f w \psi^2 + \int_{\mathcal{C}_\infty} F \cdot \mathcal{D}(w\psi^2)}_{\text{RHS}} + \int_{\mathcal{C}_\infty} \xi w \psi^2.$$

We then estimate the terms on the left- and right-hand sides separately. Before starting the computation, we mention that in the following paragraphs all the derivatives and integrations are acting on the first spatial variable. For the left-hand side of (133), we have

$$\begin{aligned} \text{LHS} &= \int_{\mathcal{C}_\infty} ((\partial_t w - \nabla \cdot \mathbf{a}\nabla w)\psi^2 w) \\ &= \int_{\mathcal{C}_\infty} \frac{1}{2} \partial_t (\psi^2 w^2) - \int_{\mathcal{C}_\infty} (\partial_t \psi) \psi w^2 + \int_{\mathcal{C}_\infty} \nabla(\psi^2 w) \cdot \mathbf{a}\nabla w. \end{aligned}$$

Using (132) and Young's inequality, we have, for any  $a \in (0, \infty)$ ,

$$\int_{\mathcal{C}_\infty} \nabla(\psi^2 w) \cdot \mathbf{a}\nabla w \geq \lambda \int_{\mathcal{C}_\infty} |\nabla w|^2 \psi^2 - a \int_{\mathcal{C}_\infty} |\nabla w|^2 \psi^2 - \frac{\tilde{C}^2}{a} \int_{\mathcal{C}_\infty} |\nabla \psi|^2 w^2,$$

where, following the notation convention recalled at the beginning of this section, the notation  $\nabla\psi$  denotes the continuous gradient for functions  $\psi$ , while the notation  $\nabla(\psi w)$  refers to the discrete gradient on the infinite cluster since the map  $w$  is only defined on  $\mathcal{C}_\infty$ . By choosing the value  $a = \frac{\lambda}{2}$ , the previous display can be rewritten

$$(134) \quad \text{LHS} \geq \int_{\mathcal{C}_\infty} \frac{1}{2} \partial_t (\psi^2 w^2) + \frac{\lambda}{2} \int_{\mathcal{C}_\infty} |\nabla w|^2 \psi^2 - \int_{\mathcal{C}_\infty} (\partial_t \psi) \psi w^2 - \frac{2\tilde{C}^2}{\lambda} \int_{\mathcal{C}_\infty} |\nabla \psi|^2 w^2.$$

We now focus on terms on the right-hand side of the equality (133). For the first two terms, we use Young's inequality and obtain

$$\begin{aligned} &\int_{\mathcal{C}_\infty} f w \psi^2 + \int_{\mathcal{C}_\infty} F \cdot \mathcal{D}(w\psi^2) \\ &\leq \int_{\mathcal{C}_\infty} t f^2 \psi^2 + \frac{1}{4t} \int_{\mathcal{C}_\infty} w^2 \psi^2 + \frac{\tilde{C}^2}{\lambda} \int_{\mathcal{C}_\infty} |F|^2 \psi^2 + \frac{\lambda}{4} \int_{\mathcal{C}_\infty} |\nabla w|^2 \psi^2 \\ &\quad + \frac{\lambda}{2} \int_{\mathcal{C}_\infty} |F|^2 \psi^2 + \frac{2\tilde{C}^2}{\lambda} \int_{\mathcal{C}_\infty} |\nabla \psi|^2 w^2. \end{aligned}$$

A combination of the two previous displays and the identity (133) shows

$$\begin{aligned}
 & \int_{\mathcal{C}_\infty} \left( \frac{1}{2} \partial_t (\psi^2 w^2) + \frac{\lambda}{4} |\nabla w|^2 \psi^2 \right) \\
 (135) \quad & \leq \int_{\mathcal{C}_\infty} w^2 \left( (\partial_t \psi) \psi + \frac{1}{4t} \psi^2 + \frac{4\tilde{C}^2}{\lambda} |\nabla \psi|^2 \right) \\
 & \quad + 2 \int_{\mathcal{C}_\infty} t f^2 \psi^2 + \left( \frac{\tilde{C}^2}{\lambda} + \frac{\lambda}{2} \right) \int_{\mathcal{C}_\infty} |F|^2 \psi^2 + \int_{\mathcal{C}_\infty} \xi w \psi^2.
 \end{aligned}$$

The value of the constants in the second line of the estimate (135) does not need to be tracked in the proof, we thus rewrite it in the following form:

$$\begin{aligned}
 (136) \quad & \int_{\mathcal{C}_\infty} \left( \frac{1}{2} \partial_t (\psi^2 w^2) + \frac{\lambda}{4} |\nabla w|^2 \psi^2 \right) \leq \int_{\mathcal{C}_\infty} w^2 \left( (\partial_t \psi) \psi + \frac{1}{4t} \psi^2 + \frac{4\tilde{C}^2}{\lambda} |\nabla \psi|^2 \right) \\
 & \quad + C \left( \int_{\mathcal{C}_\infty} t f^2 \psi^2 + |F|^2 \psi^2 + \xi w \psi^2 \right).
 \end{aligned}$$

To complete the proof, we need to prove that the quantities on the second line of the previous display are small:

- One needs to prove that the term  $\int_{\mathcal{C}_\infty} \xi(t, \cdot, y) (\psi^2(t, \cdot, y) w(t, \cdot, s, y))$  is small; this is proved in Step 3;
- One needs to prove that the term  $\int_{\mathcal{C}_\infty} (t f^2(t, \cdot, y) + |F|^2(t, \cdot, y)) \psi^2(t, \cdot, y)$  is small; this is proved in Step 4.

*Step 3: Estimate of the term  $\int_{\mathcal{C}_\infty} \xi(t, \cdot, y) (\psi^2(t, \cdot, y) w(t, \cdot, s, y))$ .* The term  $\xi$  involves the centered flux  $\tilde{\mathbf{g}}_{e_k}^*$ , to prove that this integral is small, the strategy is to use the weak norm estimate on this function stated in Proposition 7 and a multiscale argument. Specifically, the goal of this step is to prove the inequality

$$(137) \quad \int_{\mathcal{C}_\infty} \xi \psi^2 w \leq C t^{-\frac{d}{4}-1+\frac{\alpha}{2}} \|\nabla(w\psi)\|_{L^2(\mathcal{C}_\infty)} + C t^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}} \|w\psi\|_{L^2(\mathcal{C}_\infty)}.$$

As in Lemma 4.1, we need to split the space into scales and we define the dyadic annuli: for each integer  $m \geq 1$ , we let  $A_m$  be the annulus  $A_m := \{z \in \mathbb{Z}^d : 2^{m-1} \sqrt{t} \leq |z - y| < 3 \cdot 2^m \sqrt{t}\}$ , we also let  $A_0 := B_{\sqrt{t}}(y)$ . We then split the proof into two steps:

1. We first prove the estimate

$$(138) \quad \int_{\mathcal{C}_\infty} \xi \psi^2 w \leq C \Xi_1 \|\nabla(w\psi)\|_{L^2(\mathcal{C}_\infty)} + C \Xi_2 \|w\psi\|_{L^2(\mathcal{C}_\infty)},$$

where the two quantities  $\Xi_1, \Xi_2$  are defined by the formulas

$$\begin{aligned}
 (139) \quad \Xi_1 & := \left( \sum_{k=1}^d \sum_{m=0}^\infty (2^m \sqrt{t})^d \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap A_m)}^2 \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)}^2 \right)^{\frac{1}{2}}, \\
 \Xi_2 & := \left( \sum_{k=1}^d \sum_{m=0}^\infty (2^m \sqrt{t})^d \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap A_m)}^2 \right. \\
 & \quad \times \left( \|\mathcal{D}^3 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)}^2 + \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \mathcal{D} \psi\|_{L^\infty(A_m)}^2 \right. \\
 & \quad \left. \left. + (3^m \sqrt{t})^{-2} \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)}^2 \right) \right)^{\frac{1}{2}}.
 \end{aligned}$$

2. We then prove the estimates

$$(140) \quad \Xi_1 \leq Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}} \quad \text{and} \quad \Xi_2 \leq Ct^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}}.$$

The estimate (137) is a consequence of the inequalities (138) and (140).

We now focus on the proof of the inequality (138). The strategy is to use a multiscale analysis. We let  $\eta$  be a smooth cutoff function from  $\mathbb{R}^d$  to  $\mathbb{R}$  satisfying the properties

$$(141) \quad 0 \leq \eta \leq 1, \quad |\nabla\eta| \leq 1, \quad \text{supp}(\eta) \subseteq B_3(y), \quad \eta \equiv 1 \quad \text{in } B_1(y).$$

For an integer  $m$ , we define the rescaled version  $\eta_m$  of  $\eta$  according to the formula  $\eta_m := \eta(\frac{\cdot - y}{2^m \sqrt{t}} + y)$ , we also set the convention  $\eta_{-1} \equiv 0$ . This function satisfies the property: for each  $m \in \mathbb{N}$ ,

$$\begin{aligned} \text{supp}(\eta_m) &\subseteq B_{2^{m+1}\sqrt{t}}(y), & \eta_m &\equiv 1 \quad \text{in } B_{2^m\sqrt{t}}(y), \\ \text{supp}(\eta_m - \eta_{m-1}) &\subseteq A_m, & |\nabla\eta_m| &\leq (2^m \sqrt{t})^{-1}. \end{aligned}$$

We also note that the family of functions  $\eta_m$  can be used as a partition of unity and we have

$$1 = \sum_{m=0}^{\infty} (\eta_m - \eta_{m-1}).$$

With this property, we compute

$$\begin{aligned} \int_{\mathcal{C}_\infty} \xi \psi^2 w &= \sum_{m=0}^{\infty} \int_{\mathcal{C}_\infty} (\eta_m - \eta_{m-1}) \xi \psi^2 w \\ (142) \quad &= \sum_{k=1}^d \sum_{m=0}^{\infty} \int_{\mathcal{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y)} \tilde{\mathbf{g}}_{e_k}^* \cdot \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w \\ &\leq \sum_{k=1}^d \sum_{m=0}^{\infty} (2^m \sqrt{t})^d \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap A_m)} \\ &\quad \times \|\mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w\|_{\underline{H}^1(\mathcal{C}_\infty \cap A_m)}. \end{aligned}$$

Then we calculate the  $\underline{H}^1$ -norm of the term  $\mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w$ . We use the fact that the function  $(\eta_m - \eta_{m-1})$  is supported in the annulus  $A_m$  and write

$$\|\mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w\|_{\underline{H}^1(\mathcal{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y))} \leq C(I_1 + I_2),$$

where the two terms  $I_1$  and  $I_2$  are defined by the formulas

$$\begin{aligned} I_1 &:= \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)} \|\nabla(w\psi)\|_{\underline{L}^2(\mathcal{C}_\infty \cap A_m)} \\ I_2 &:= (\|\mathcal{D}^3 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)} + \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \mathcal{D}\psi\|_{L^\infty(A_m)} \\ &\quad + (3^m \sqrt{t})^{-1} \|\mathcal{D}^2 \bar{p}(t, \cdot - y) \psi\|_{L^\infty(A_m)}) \|w\psi\|_{\underline{L}^2(\mathcal{C}_\infty \cap A_m)}. \end{aligned}$$

We put these equations back into the right-hand side of the estimate (142). This gives

$$\int_{\mathcal{C}_\infty} \xi \psi^2 w \leq C \sum_{k=1}^d \sum_{m=0}^{\infty} (2^m \sqrt{t})^{\frac{d}{2}} \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap A_m)} (I_1 + I_2).$$

We then estimate the two terms on the right-hand side by applying the Cauchy–Schwarz inequality. For the term involving the quantity  $I_1$ , we obtain

$$\begin{aligned} & \sum_{k=1}^d \sum_{m=0}^{\infty} (2^m \sqrt{t})^{\frac{d}{2}} \|\tilde{\mathbf{g}}_{e_k}^* \|_{\underline{H}^{-1}(\mathcal{C}_{\infty} \cap A_m)} I_1 \\ & \leq \left( \sum_{k=1}^d \sum_{m=0}^{\infty} (2^m \sqrt{t})^d \|\tilde{\mathbf{g}}_{e_k}^* \|_{\underline{H}^{-1}(\mathcal{C}_{\infty} \cap B_{2^{m+1}\sqrt{t}(y)})}^2 \| \mathcal{D}^2 \bar{p}(t, \cdot - y) \psi \|_{L^{\infty}(A_m)}^2 \right)^{\frac{1}{2}} \\ & \quad \times \left( \sum_{k=1}^d \sum_{m=0}^{\infty} \|\nabla(w\psi)\|_{L^2(\mathcal{C}_{\infty} \cap A_m)}^2 \right)^{\frac{1}{2}} \\ & \leq C \Xi_1 \|\nabla(w\psi)\|_{L^2(\mathcal{C}_{\infty})}, \end{aligned}$$

where to go from the second to the third line, we used the definition of  $\Xi_1$  given in (139) and the inequality  $\sum_{m=1}^{\infty} \mathbb{1}_{A_m}(y) \leq 4$ . The same argument works for the terms involving the quantities  $I_2$  and  $\Xi_2$ ; this concludes the proof of the estimate (138).

We now prove an estimate on the terms  $\Xi_1, \Xi_2$ ; precisely we prove the inequality (140) which is recalled below

$$(143) \quad \Xi_1 \leq C t^{-\frac{d}{4}-1+\frac{\alpha}{2}} \quad \text{and} \quad \Xi_2 \leq C t^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}}.$$

The proof comes from a direct calculation of the quantities  $\Xi_1$  and  $\Xi_2$ . We recall that the function  $\psi$  is chosen to be either the constant function equal to 1, or the function  $\exp(\Psi_C(t, |\cdot - y|))$ , for some large constant  $C$ .

We first focus on the estimate of the term  $\Xi_1$ ; if the constant  $C$  in the definition of  $\psi$  is chosen large enough, for instance larger than  $8\bar{\sigma}^2$ , then one has the estimate

$$\|\mathcal{D}^2 \bar{p}(t, \cdot - y) \psi \|_{L^{\infty}(A_m(y))} \leq C t^{-\frac{d}{2}-1} \exp\left(-\frac{2^{2m}}{8\bar{\sigma}^2}\right).$$

Thanks to the assumption  $\sqrt{t} > \mathcal{M}_{\text{flux},\alpha}(y)$ , we have the estimate

$$\|\tilde{\mathbf{g}}_{e_k}^* \|_{\underline{H}^{-1}(\mathcal{C}_{\infty} \cap A_m)} \leq \|\tilde{\mathbf{g}}_{e_k}^* \|_{\underline{H}^{-1}(\mathcal{C}_{\infty} \cap B_{2^{m+1}\sqrt{t}(y)})} \leq C (2^m \sqrt{t})^{\alpha}.$$

Combining these two bounds with the definition of  $\Xi_1$  given in (139), we obtain

$$(\Xi_1)^2 \leq C \sum_{k=1}^d \sum_{m=0}^{\infty} (2^m \sqrt{t})^{d+2\alpha} t^{-d-2} \exp\left(-\frac{2^{2m}}{4\bar{\sigma}^2}\right) \leq C t^{-\frac{d}{2}-2+\alpha}.$$

The term  $\Xi_2$  can be estimated thanks to a similar strategy and the details are left to the reader. The proof of the estimate (143), and thus of the inequality (137) is complete.

*Step 4: Quantification of the term  $\int_{\mathcal{C}_{\infty}} (tf^2(t, \cdot, y) + |F|^2(t, \cdot, y))\psi^2(t, \cdot, y)$ .* The goal of this step is to prove the inequality

$$(144) \quad \int_{\mathcal{C}_{\infty}} (tf^2(t, \cdot, y) + |F|^2(t, \cdot, y))\psi^2(t, \cdot, y) \leq C t^{-\frac{d}{2}-2+\alpha}.$$

As was the case in the previous step, the function  $\psi$  is either the constant function equal to 1 or the function  $\exp(\Psi_C(t, |\cdot - y|))$ . In the latter case, we assume that the constant  $C$  is at least larger than  $8\bar{\sigma}^2$ .

We first consider the term involving the function  $f$ . From the definition of this function given in (131), we see that it is the sum of two terms. The first one is the difference of the discrete and the continuous Laplacian of the heat kernel  $\bar{p}$ ; it can be estimated as follows:

$$|(\Delta \bar{p}(t, \cdot - y) - (-\mathcal{D}^* \cdot \mathcal{D} \bar{p}(t, \cdot - y)))\psi| \leq C t^{-\frac{d}{2}-\frac{3}{2}} \exp\left(-\frac{|\cdot - y|^2}{4\bar{\sigma}^2 t}\right).$$

The second term is the quantity  $\sum_{k=1}^d \partial_t \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \chi_{e_k}(\cdot)$ . To estimate it, we split the space into different scales using the functions  $\eta_m$  introduced in Step 3. This gives

$$\begin{aligned} & \int_{\mathcal{C}_\infty} t \sum_{k=1}^d (\partial_t (\mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \chi_{e_k}) \psi)^2 \\ &= \sum_{k=1}^d \sum_{m=0}^\infty \int_{\mathcal{C}_\infty} t (\eta_m - \eta_{m-1}) (\partial_t \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \psi)^2 \chi_{e_k}^2 \\ &\leq \sum_{k=1}^d \sum_{m=0}^\infty t (2^m \sqrt{t})^d \|\chi_{e_k}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y))}^2 \|(\partial_t \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \psi)\|_{L^\infty(A_m)}^2. \end{aligned}$$

We then use the assumption  $\sqrt{t} > \mathcal{M}_{\text{corr},\alpha}(y)$ , which implies  $\|\chi_{e_k}\|_{\underline{L}^2(\mathcal{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y))} \leq C(2^m \sqrt{t})^\alpha$ , and the estimate

$$\|(\partial_t \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \psi)\|_{L^\infty(A_m)} \leq C t^{-\frac{d}{2} - \frac{3}{2}} \exp\left(-\frac{2^{2m}}{8\bar{\sigma}^2}\right),$$

to obtain the inequality

$$\begin{aligned} \int_{\mathcal{C}_\infty} t \sum_{k=1}^d (\partial_t (\mathcal{D}_{e_k} \bar{p}(t, \cdot - y) \chi_{e_k}) \psi)^2 &\leq C \sum_{k=1}^d \sum_{m=0}^\infty t (2^m \sqrt{t})^{d+2\alpha} t^{-d-3} \exp\left(-\frac{2^{2m}}{4\bar{\sigma}^2}\right) \\ &\leq C t^{-\frac{d}{2} - 2 + \alpha}. \end{aligned}$$

The estimate for the term involving the function  $F$  is similar and we skip its proof.

*Step 5: The conclusion.* We collect the results established in the previous steps and complete the proof of Proposition 10. We first consider the inequality (136) in the case  $\psi = 1$ . This gives

$$\int_{\mathcal{C}_\infty} \left(\frac{1}{2} \partial_t w^2 + \frac{\lambda}{4} |\nabla w|^2\right) \leq \frac{1}{4t} \int_{\mathcal{C}_\infty} w^2 + C \left(\int_{\mathcal{C}_\infty} t f^2 + |F|^2 + \xi w\right),$$

since in this case the constant  $\tilde{C}$  introduced in (132) is equal to 1. Applying the main results (137) of Step 3 and (144) of Step 4, we deduce

$$\begin{aligned} & \int_{\mathcal{C}_\infty} \left(\partial_t w^2 + \frac{\lambda}{2} |\nabla w|^2\right) \\ &\leq \frac{1}{2t} \int_{\mathcal{C}_\infty} w^2 + C t^{-\frac{d}{2} - 2 + \alpha} + C t^{-\frac{d}{4} - 1 + \frac{\alpha}{2}} (\|\nabla w\|_{L^2(\mathcal{C}_\infty)} + t^{-\frac{1}{2}} \|w\|_{L^2(\mathcal{C}_\infty)}). \end{aligned}$$

By Young's inequality, the previous display can be simplified

$$\int_{\mathcal{C}_\infty} \left(\partial_t w^2 + \frac{\lambda}{4} |\nabla w|^2\right) \leq \frac{1}{t} \int_{\mathcal{C}_\infty} w^2 + C t^{-\frac{d}{2} - 2 + \alpha},$$

which implies

$$\partial_t \int_{\mathcal{C}_\infty} w^2 \leq \frac{1}{t} \int_{\mathcal{C}_\infty} w^2 + C t^{-\frac{d}{2} - 2 + \alpha}.$$

By integrating over the time interval  $[\tau, t]$ , we obtain that there exists a positive constant  $C := C(d, p, \lambda) < \infty$  such that

$$(145) \quad \int_{\mathcal{C}_\infty} w^2(t, \cdot, s, y) \leq C \left(\frac{t}{\tau}\right) \tau^{-\frac{d}{2} - 1 + \alpha}.$$

We now consider the inequality (136) in the case  $\psi = \exp(\Psi_C(t, \cdot - y))$ . This gives

$$\begin{aligned} & \int_{\mathcal{E}_\infty} \left( \frac{1}{2} \partial_t (\psi^2 w^2) + \frac{\lambda}{4} |\nabla w|^2 \psi^2 \right) \\ & \leq \int_{\mathcal{E}_\infty} w^2 \left( (\partial_t \psi) \psi + \frac{1}{8t} \psi^2 + \frac{4\tilde{C}^2}{\lambda} |\nabla \psi|^2 \right) \\ & \quad + Ct^{-\frac{d}{2}-2+\alpha} + Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}} (\|\nabla(w\psi)\|_{L^2(\mathcal{E}_\infty)} + t^{-\frac{1}{2}} \|w\psi\|_{L^2(\mathcal{E}_\infty)}). \end{aligned}$$

Applying Young’s inequality, the previous display can be simplified and we obtain

$$\begin{aligned} & \int_{\mathcal{E}_\infty} \left( \frac{1}{2} \partial_t (\psi^2 w^2) + \frac{\lambda}{8} |\nabla w|^2 \psi^2 \right) \\ & \leq \int_{\mathcal{E}_\infty} w^2 \left( (\partial_t \psi) \psi + \frac{1}{4t} \psi^2 + \frac{8\tilde{C}}{\lambda} |\nabla \psi|^2 \right) + Ct^{-\frac{d}{2}-2+\alpha}. \end{aligned}$$

We then note that if the constant  $C$  in the definition of  $\psi = \exp(\Psi_C(\cdot, \cdot - y))$  is chosen large enough, then we have

$$(146) \quad (\partial_t \psi) \psi + \frac{1}{4t} \psi^2 + \frac{8\tilde{C}}{\lambda} |\nabla \psi|^2 \leq \frac{C}{t}.$$

A combination of the two previous displays shows the differential inequality

$$\partial_t \int_{\mathcal{E}_\infty} \frac{1}{2} \psi^2 w^2 \leq \frac{C}{t} \int_{\mathcal{E}_\infty} w^2 + Ct^{-\frac{d}{2}-2+\alpha}.$$

We then apply (145) to obtain

$$\partial_t \int_{\mathcal{E}_\infty} \frac{1}{2} \psi^2 w^2 \leq Ct^{-1} \left( \frac{t}{\tau} \right) \tau^{-\frac{d}{2}-1+\alpha} + Ct^{-\frac{d}{2}-2+\alpha}.$$

Integrating with respect to the time  $t$  and recalling that  $w(\tau, \cdot, \tau, y) = 0$ , we obtain

$$\int_{\mathcal{E}_\infty} (w\psi)^2 \leq C \left( \frac{t}{\tau} \right) \tau^{-\frac{d}{2}-1+\alpha},$$

for some constant  $C := C(d, p, \lambda) < \infty$ . This completes the proof of Proposition 10.  $\square$

REMARK 15. The reason why we choose the function  $\psi = \exp(\Psi_C(\cdot, \cdot - y))$ , and why the main result (129) of Proposition 10 is stated with this function can be explained by the inequalities (132) and (146). Indeed, the function  $(t, x) \mapsto \exp(\Psi_C(t, x - y))$  is the one which has the fastest growth as  $x$  tends to infinity such that the inequalities (132) and (146) are satisfied. In particular, there is an important difference between the discrete setting and the continuous setting: in the latter, one does not need the inequality (132) to hold which allows to choose the function  $\psi(t, x) := \exp(\frac{|x-y|^2}{Ct})$ , and to obtain the result with this function (see [15], Lemma 8.22). This observation is consistent with the asymptotic behavior of the discrete heat kernel on the percolation cluster or on  $\mathbb{Z}^d$  which is described by Proposition 8.

We have now collected all the necessary results to prove Theorem 1.1. The following section is devoted to its proof.

4.3. *Proof of Theorem 1.* By translation invariance of the model, it is sufficient to prove the result when  $y = 0 \in \mathcal{C}_\infty$ . We fix an exponent  $\delta > 0$ ; the objective is to apply the results of Proposition 10, Lemma 4.1 and Lemma 4.2 with the mesoscopic time  $\tau = t^{1-\kappa}$  and with following values of exponents

$$\alpha := \frac{\delta}{2} \quad \text{and} \quad \kappa := \frac{\delta}{d+2}.$$

For later use, we note that with these specific choices of exponents, the following estimates hold:

$$(147) \quad (1-\kappa)\left(\frac{1}{2} - \frac{\alpha}{2}\right) > \frac{1}{2} - \delta \quad \text{and} \quad -\kappa + (1-\kappa)\left(\frac{d}{4} + \frac{1}{2} - \frac{\alpha}{2}\right) > \frac{d}{4} + \frac{1}{2} - \delta.$$

The proof relies on an induction argument and we give a setup of the proof. We first define the sequence  $\kappa_n$  of real numbers inductively by the formula

$$(148) \quad \kappa_0 = \frac{\kappa}{2} \quad \text{and} \quad \kappa_{n+1} := \min\left((1-\kappa)\kappa_n + \frac{\kappa}{2}, \frac{1}{2} - \delta\right)$$

This sequence is increasing and is ultimately constant equal to the value  $\frac{1}{2} - \delta$ . We let  $N$  be the integer

$$N := \inf\left\{n \in \mathbb{N} : \kappa_n = \frac{1}{2} - \delta\right\},$$

and we note that this integer only depends on the parameters  $d, p, \lambda$  and  $\delta$ . For each point  $z \in \mathbb{Z}^d$ , we define the random time  $\mathcal{T}_{\text{par}}^0(z)$  according to the formula

$$\mathcal{T}_{\text{par}}^0(z) := 4 \max(\mathcal{T}_{\text{approx},\alpha}(z)^{\frac{1}{1-\kappa}}, \mathcal{M}_{\text{corr},\alpha}(z)^{\frac{2}{1-\kappa}}, \mathcal{M}_{\text{flux},\alpha}(z)^2)$$

so that for any time  $t \geq \mathcal{T}_{\text{par}}^0(z)$ , all the results of Sections 4.1 and 4.2 are valid with the value  $\tau := t^{1-\kappa}$ . We then upgrade the random variable  $\mathcal{T}_{\text{par}}^0(z)$  and define

$$(149) \quad \mathcal{T}_{\text{par}}^1 := \sup\{t \in [1, \infty) : \exists z \in \mathcal{C}_\infty \text{ such that } |z| \leq (N+1)t^{\frac{1}{(1-\kappa)N}} \text{ and } \mathcal{T}_{\text{par}}^0(z) \geq t\},$$

so that for any time  $t \geq \mathcal{T}_{\text{par}}^1$ , and any point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N+1)t^{\frac{1}{(1-\kappa)N}}$ , one has the estimate  $t \geq \mathcal{T}_{\text{par}}^0(z)$ ; this implies that all the results of Sections 4.1 and 4.2 are valid with the value  $\tau := t^{1-\kappa}$  for the heat kernel started from the point  $z$ . This construction is identical to the used to define the minimal time  $\mathcal{T}'_{\text{NA}}(x)$  in (101). As it was the case for the random variable  $\mathcal{T}'_{\text{NA}}(x)$ , an application of Lemma 1.4 shows the stochastic integrability estimate

$$\mathcal{T}_{\text{par}}^1 \leq \mathcal{O}_s(C).$$

For each integer  $n \in \{0, \dots, N\}$ , we let  $H_n$  be the following statement.

*Statement  $H_n$ .* There exists a constant  $C(d, \lambda, n) < \infty$  such that for each time  $t \geq (\mathcal{T}_{\text{par}}^1)^{\frac{1}{(1-\kappa)^n}}$ , each point  $x \in \mathcal{C}_\infty$ , and each point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N-n)t^{\frac{1}{(1-\kappa)^{N-n}}}$ , one has the estimate

$$(150) \quad |p(t, x, z) - \theta(p)^{-1} \bar{p}(t, x - z)| \leq Ct^{-\kappa n} \Phi_C(t, x - z).$$

We prove by induction that the statement  $H_n$  holds for each integer  $n \in \{0, \dots, N\}$ .

*The base case.* We prove that  $H_0$  holds and first prove the  $L^2$ -estimate: for each time  $t \geq \mathcal{T}_{\text{par}}^1$ , and each point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N + 1)t^{\frac{1}{(1-\kappa)^N}}$ ,

$$(151) \quad \left\| (p(t, \cdot, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, \cdot - z)) \exp(\Psi_C(t, |\cdot - z|)) \right\|_{L^2(\mathcal{C}_\infty)} \leq Ct^{-\frac{d}{4} - \frac{\kappa}{2}}.$$

We recall the definitions of the functions  $h, q$  and  $v$  stated in (106), (109), (123), respectively, as well as the definition of  $w$  given by the formula  $w := h - v - q$ . We write

$$(152) \quad \begin{aligned} & p(t, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - z) \\ &= (p(t, x, z) - q(t, x, \tau, z)) - v(t, x, \tau, z) + w(t, x, \tau, z) \\ & \quad + (h(t, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, x, z)). \end{aligned}$$

To prove the estimate (151), we split the  $L^2$ -norm according to the decomposition (152) and estimate each terms thanks to the results established in Sections 4.1 and 4.2:

- The term  $(p(t, x, z) - q(t, x, \tau, z))$  is estimated thanks to Lemma 4.1, this term accounts for an error of order

$$\left( \left( \frac{\tau}{t} \right)^{\frac{1}{2}} + \tau^{-\frac{1}{2} + \alpha} \right) t^{-\frac{d}{4}} = (t^{-\frac{\kappa}{2}} + t^{(1-\kappa)(-\frac{1}{2} + \alpha)}) t^{-\frac{d}{4}} \leq t^{-\frac{d}{4} - \frac{\kappa}{2}};$$

- The term  $w$  is estimated thanks to Proposition 10, this term accounts for an error of order

$$\left( \frac{t}{\tau} \right)^{\frac{1}{2}} \tau^{-\frac{d}{4} - \frac{1}{2} + \frac{\alpha}{2}} \leq t^{-\frac{d}{4} - \frac{1}{2} + \delta},$$

where we used the estimate (147);

- The term  $v(t, x, \tau, z)$  is estimated thanks to Lemma 4.2, this term accounts for an error of order

$$t^{-\frac{d}{4}} \tau^{-\frac{1}{2} + \frac{\alpha}{2}} = t^{-\frac{d}{4} + (1-\kappa)(-\frac{1}{2} + \frac{\alpha}{2})} \leq t^{-\frac{d}{4} - \frac{1}{2} + \delta},$$

where we used the estimate (147);

- The term  $h(t, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, x, z)$  can be estimated as follows. By the definition of  $h$  given in (106), we have

$$h(t, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, x, z) = \sum_{k=1}^d \mathcal{D}_{e_k} \bar{p}(t, x - z) \chi_{e_k}(x).$$

The term can then be estimated by using the sublinearity of the corrector stated in Proposition 6 and the assumption  $\sqrt{t} \geq \mathcal{M}_{\text{corr}, \alpha}(z)$ . The proof is similar to the one of Lemma 4.2 and the details are left to the reader. It accounts for an error of order  $t^{-\frac{d}{4} - \frac{1}{2} + \delta}$ .

There remains to obtain the pointwise estimate (150) in the case  $n = 0$  from the  $L^2$ -estimate (151). To this end, we fix a point  $z \in \mathcal{C}_\infty$  such that  $|z| \leq Nt^{\frac{1}{(1-\kappa)^N}}$ . We may without loss of generality restrict our attention to the points  $x \in \mathcal{C}_\infty$  such that  $|x| \leq (N + 1)t^{\frac{1}{(1-\kappa)^N}}$ , otherwise we necessarily have  $|x - z| \geq t$  and the inequality (150) is satisfied by Proposition 8 and Remark 8. We then use the semigroup property on the heat kernels  $p$  and  $\bar{p}$ : for each  $x, z \in \mathcal{C}_\infty$ , one has

$$(153) \quad \begin{aligned} & p(t, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, x - z) \\ &= \underbrace{\int_{\mathcal{C}_\infty} p\left(\frac{t}{2}, x, y\right) p\left(\frac{t}{2}, y, z\right) - \theta(p)^{-2} \bar{p}\left(\frac{t}{2}, x - y\right) \bar{p}\left(\frac{t}{2}, y - z\right) dy}_{(153)\text{-a}} \end{aligned}$$

$$+ \theta(p)^{-1} \underbrace{\left( \theta(p)^{-1} \int_{\mathcal{C}_\infty} \bar{p}\left(\frac{t}{2}, x - y\right) \bar{p}\left(\frac{t}{2}, y - z\right) dy - \bar{p}(t, x - z) \right)}_{(153)\text{-b}}.$$

We first treat the part (153)-a using the following  $L^2$ -estimate

$$|(153)\text{-a}| \leq (153)\text{-a1} + (153)\text{-a2},$$

where the two terms (153)-a1 and (153)-a2 are defined by the formulas

$$(153)\text{-a1} = \left\| \left( p\left(\frac{t}{2}, x, \cdot\right) - \theta(p)^{-1} \bar{p}\left(\frac{t}{2}, x - \cdot\right) \right) \exp\left(\Psi_C\left(\frac{t}{2}, |x - \cdot|\right)\right) \right\|_{L^2(\mathcal{C}_\infty)} \\ \times \left\| p\left(\frac{t}{2}, \cdot, z\right) \exp\left(-\Psi_C\left(\frac{t}{2}, |x - \cdot|\right)\right) \right\|_{L^2(\mathcal{C}_\infty)}$$

and

$$(153)\text{-a2} = \left\| \left( p\left(\frac{t}{2}, \cdot, z\right) - \theta(p)^{-1} \bar{p}\left(\frac{t}{2}, \cdot - z\right) \right) \exp\left(\Psi_C\left(\frac{t}{2}, |\cdot - z|\right)\right) \right\|_{L^2(\mathcal{C}_\infty)} \\ \times \left\| \theta(p)^{-1} \bar{p}\left(\frac{t}{2}, x - \cdot\right) \exp\left(-\Psi_C\left(\frac{t}{2}, |\cdot - z|\right)\right) \right\|_{L^2(\mathcal{C}_\infty)}.$$

The term (153)-a1 can be estimated by using the three following ingredients:

- The symmetry of the heat kernel  $p$ ;
- The  $L^2$ -estimate (151) applied with the point  $z = x$  which is valid under the assumption  $|x| \leq (N + 1)t^{\frac{1}{(1-\kappa)N}}$ ;
- The upper bound stated in Theorem 3.1, which can be applied since we assumed  $t \geq \mathcal{T}_{\text{par}}^1 \geq 2\mathcal{T}_{\text{NA}}^t(z)$ , and reads, by increasing the value of the constant  $C$  in the right-hand side if necessary,

$$\left\| p\left(\frac{t}{2}, \cdot, z\right) \exp\left(-\Psi_C\left(\frac{t}{2}, |x - \cdot|\right)\right) \right\|_{L^2(\mathcal{C}_\infty)} \leq t^{\frac{d}{4}} \Phi_C(t, x - z).$$

These arguments imply the estimate

$$(153)\text{-a1} \leq t^{-\frac{\kappa}{2}} \Phi_C(t, x - y).$$

The term (153)-a2 can be treated similarly and we omit the details. There remains to estimate the term (153)-b. We note that by an application of the parallelogram law, that is, the identity  $|x - y|^2 + |y - z|^2 = 2(|\frac{x-z}{2}|^2 + |y - \frac{x+z}{2}|^2)$ , the function  $\bar{p}$  satisfies the following property: for each  $t \geq 0$ , and each  $x, y, z \in \mathbb{R}^d$ ,

$$\bar{p}\left(\frac{t}{2}, x - y\right) \bar{p}\left(\frac{t}{2}, y - z\right) = \bar{p}(t, x - z) \bar{p}\left(\frac{t}{4}, y - \frac{x + z}{2}\right).$$

By combining this identity with Proposition 13, we obtain

$$|(153)\text{-b}| = \bar{p}(t, x - z) \left| \int_{\mathcal{C}_\infty} \theta(p)^{-1} \bar{p}\left(\frac{t}{4}, y - \frac{x + z}{2}\right) dy - 1 \right| \leq Ct^{-\frac{1}{2} + \delta} \bar{p}(t, x - z).$$

This completes the proof of the base case.

*The iteration step.* We prove that, for each integer  $n \in \mathbb{N}$ , the statement  $H_{n-1}$  implies the statement  $H_n$ . The strategy follows the one of the base case and we first prove the  $L^2$ -estimate, under the assumption that the statement  $H_{n-1}$  is valid: for each time  $t \geq (\mathcal{T}_{\text{par}}^1)^{\frac{1}{(1-\kappa)^n}}$ , each point  $x \in \mathcal{C}_\infty$ , and each point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N + 1 - n)t^{\frac{1}{(1-\kappa)^{N-n}}}$ ,

$$(154) \quad \|(p(t, \cdot, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, \cdot - z)) \exp(\Psi_C(t, |\cdot - z|))\|_{L^2(\mathcal{C}_\infty)} \leq Ct^{-\frac{d}{4} - \kappa_n}.$$

We use the decomposition (152) with the same value for the mesoscopic time  $\tau = t^{1-\kappa}$ . The error introduced by the terms  $w, v$  and  $h - \theta^{-1}(\mathbf{p})\bar{p}$  are of order  $t^{-\frac{d}{4} - \frac{1}{2} + \delta}$  which is smaller than the value  $t^{-\frac{d}{4} - \kappa_n}$  we want to prove in this step. The limiting factor comes from the term  $(p(t, x, z) - q(t, x, \tau, z))$  which is estimated in Lemma 4.1 and gives an error of order  $t^{-\frac{d}{4} - \frac{\kappa}{2}}$ . The objective of the induction step is to improve this error by using the statement  $H_{n-1}$ .

Under the assumption  $t \geq (\mathcal{T}_{\text{par}}^1)^{\frac{1}{(1-\kappa)^n}}$ , we have  $\tau = t^{1-\kappa} \geq (\mathcal{T}_{\text{par}}^1)^{\frac{1}{(1-\kappa)^{n-1}}}$ . We can thus apply the induction hypothesis  $H_{n-1}$  with time  $\tau$ . This gives the inequality, for each point  $x \in \mathcal{C}_\infty$ , and each point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N + 1 - n)\tau^{\frac{1}{(1-\kappa)^{N+1-n}}}$  (that is to say  $|z| = (N + 1 - n)t^{\frac{1}{(1-\kappa)^{N-n}}}$ ),

$$(155) \quad |p(\tau, x, z) - \theta(\mathbf{p})^{-1} \bar{p}(\tau, x - z)| \leq C\tau^{-\frac{d}{4} - \kappa_{n-1}} \Phi_C(\tau, x - z).$$

This estimate can be used to improve the result of Lemma 4.1 according to the following procedure. We go back to the proof of Lemma 4.1 and in the inequality (112), instead of using the Nash–Aronson estimate stated in Theorem 3.1, we use the homogenization estimate (155). We then proceed with the proof and do not make any other modification. This implies the following improved version of Lemma 4.1:

$$|q(t, x, \tau, z) - p(t, x, z)| \leq \left( \tau^{-\kappa_{n-1}} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} + \tau^{-\frac{1}{2} + \alpha} \right) \Phi_C(t, x - z).$$

Once equipped with this estimate, we can prove the  $L^2$ -estimate (154). The proof is the same as the one presented in the base case, we only use the estimate (154) instead of Lemma 4.1.

We obtain, for any point  $z \in \mathcal{C}_\infty$  satisfying  $|z| \leq (N + 1 - n)\tau^{\frac{1}{(1-\kappa)^{N-n}}}$ ,

$$\begin{aligned} & \|(p(t, \cdot, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, \cdot - z)) \exp(\Psi_C(t, |\cdot - z|))\|_{L^2(\mathcal{C}_\infty)} \\ & \leq t^{-\frac{d}{4}} \tau^{-\kappa_{n-1}} \left( \frac{\tau}{t} \right)^{\frac{1}{2}} + t^{-\frac{d}{4} - \frac{1}{2} + \delta}. \end{aligned}$$

We then use the equality  $\tau = t^{1-\kappa}$  and the inductive definition of the sequence  $\kappa_n$  stated in (148) to deduce the estimate

$$\|(p(t, \cdot, z) - \theta(\mathbf{p})^{-1} \bar{p}(t, \cdot - z)) \exp(\Psi_C(t, |\cdot - z|))\|_{L^2(\mathcal{C}_\infty)} \leq Ct^{-\frac{d}{4} - \kappa_n}.$$

The proof of the pointwise estimate (150) is identical to the proof written for the base case and we omit the details. This completes the proof of the induction step.

We now complete the proof of Theorem 1.1. We then define the minimal time  $\mathcal{T}_{\text{par}, \delta}(0) := (\mathcal{T}_{\text{par}}^1)^{\frac{1}{(1-\kappa)^N}}$ . Since the statement  $H_N$  holds, we have the estimate, for each time  $t \geq \mathcal{T}_{\text{par}, \delta}(0)$ ,

$$|p(t, x, 0) - \theta(\mathbf{p})^{-1} \bar{p}(t, x)| \leq Ct^{-\frac{1}{2} + \delta} \Phi_C(t, x).$$

The proof of Theorem 1.1 is complete in the case  $y = 0$ . The proof in the general case is obtained by using the stationarity of the model.

**5. Quantitative homogenization of the elliptic Green's function.** The objective of this section is to present a theorem of quantitative homogenization for the elliptic Green's function on the infinite cluster, that is, to establish Theorem 1.2. This result is a consequence of the quantitative homogenization theorem for the parabolic Green's function, Theorem 1.1, established in the previous section: in dimension  $d \geq 3$ , it can be essentially obtained by integrating the heat kernel over time since one has the identity, for each  $x, y \in \mathcal{C}_\infty$ ,

$$(156) \quad g(x, y) = \int_0^\infty p(t, x, y) dt.$$

The case of the dimension 2 is more specific and requires some additional attention. In this setting the heat kernel is not integrable as the time  $t$  tends to infinity. This difficulty is related to the recurrence of the random walk on  $\mathbb{Z}^2$  or to the unbounded behavior of the Green's function in dimension 2. To remedy this, we use a corrected version of the formula (156): for each  $x, y \in \mathcal{C}_\infty$ , one has

$$g(x, y) = \int_0^\infty (p(t, x, y) - p(t, y, y)) dt,$$

where  $g$  is the unique elliptic Green's function on the infinite cluster under the environment  $\mathbf{a}$  such that  $g(y, y) = 0$ .

**PROOF OF THEOREM 1.2.** We first treat the case of the dimension  $d \geq 3$ . By the stationarity of the model, we prove the result in the case  $y = 0$ . To simplify the notation, we write  $g(x)$  instead of  $g(x, 0)$ . We let  $\mathcal{T}_{\text{par}, \delta/2}(0)$  be the minimal time provided by Theorem 1.1 with exponent  $\delta/2$  and define the minimal scale  $\mathcal{M}_{\text{ell}, \delta}(0)$  according to the formula

$$\mathcal{M}_{\text{ell}, \delta}(0) := \mathcal{T}_{\text{par}, \delta/2}(0).$$

It is on purpose that we do not respect the parabolic scaling, we need to have  $\mathcal{M}_{\text{ell}, \delta}(0) \gg \sqrt{\mathcal{T}_{\text{par}, \delta/2}(0)}$ . As was mentioned in the introduction of this section, in dimension  $d \geq 3$ , we use the explicit formula (156) and note that Duhamel's principle implies the identity

$$\bar{g}(x) = \theta(\mathbf{p})^{-1} \int_0^\infty \bar{p}(t, x) dt.$$

We obtain

$$|g(x) - \bar{g}(x)| \leq \int_0^\infty |p(t, x, 0) - \theta(p)^{-1} \bar{p}(t, x)| dt.$$

We then split the integral at time  $|x|$ ,

$$(157) \quad \begin{aligned} |g(x) - \bar{g}(x)| &\leq \int_0^{|x|} |p(t, x, 0) - \theta(p)^{-1} \bar{p}(t, x)| dt \\ &\quad + \int_{|x|}^\infty |p(t, x, 0) - \theta(p)^{-1} \bar{p}(t, x)| dt, \end{aligned}$$

and estimate the two terms on the right-hand side separately. The second term is the simplest one, we apply the quantitative estimate (7) provided by Theorem 1.1 and use the assumption  $|x| \geq \mathcal{T}_{\text{par}, \delta/2}(0)$ . This shows

$$(158) \quad \begin{aligned} \int_{|x|}^\infty |p(t, x, 0) - \theta(p)^{-1} \bar{p}(t, x)| dt &\leq C \int_{|x|}^\infty t^{-\frac{d}{2} - \frac{1}{2} + \frac{\delta}{2}} \exp\left(-\frac{|x|^2}{Ct}\right) dt \\ &\leq C \int_0^\infty t^{-\frac{d}{2} - \frac{1}{2} + \frac{\delta}{2}} \exp\left(-\frac{|x|^2}{Ct}\right) dt \\ &\leq C |x|^{-1+\delta} |x|^{2-d}. \end{aligned}$$

To treat the first term in the right-hand side of (157), we use the first estimate (49) of Proposition 8, which is recalled below, for each  $t \in (0, \infty)$ ,  $x \in \mathcal{C}_\infty$  such that  $|x| \geq t$ ,

$$p(t, x, 0) \leq C \exp\left(-C^{-1}|x|\left(1 + \ln \frac{|x|}{t}\right)\right).$$

The same estimate is also valid for the function  $\bar{p}$ . Therefore, the term  $\ln(|x|/t)$  is positive on the interval  $(0, |x|]$  and one has the estimate

$$\begin{aligned} \int_0^{|x|} |p(t, x, 0) - \theta(p)^{-1}\bar{p}(t, x)| dt &\leq C \int_0^{|x|} \exp\left(-C^{-1}|x|\left(1 + \ln \frac{|x|}{t}\right)\right) dt \\ &\leq C \int_0^{|x|} \exp(-C^{-1}|x|) dt \\ &\leq C|x| \exp(-C^{-1}|x|). \end{aligned}$$

By increasing the value of the constant  $C$ , one has

$$\int_0^{|x|} |p(t, x, 0) - \theta(p)^{-1}\bar{p}(t, x)| dt \leq C \exp(-C^{-1}|x|).$$

Combining the previous estimate with (158), we deduce

$$\begin{aligned} |g(x) - \bar{g}(x)| &\leq C|x|^{-1+\delta}|x|^{2-d} + C \exp(-C^{-1}|x|) \\ &\leq C|x|^{-1+\delta}|x|^{2-d}. \end{aligned}$$

This completes the proof of the estimate (10) in dimension larger than 3.

We now focus on the case of the dimension 2. The strategy is similar, but some additional attention is needed due to the fact that the integral (156) is ill-defined in dimension 2. We define the elliptic Green’s function  $g$  on the infinite cluster by the formula

$$(159) \quad g(x) = \int_0^\infty (p(t, x, 0) - p(t, 0, 0)) dt.$$

For the homogenized Green’s function, we cannot use the formula (159) by replacing the transition kernel  $p$  by the homogenized heat kernel  $\bar{p}$ ; indeed the integral

$$(160) \quad \int_0^\infty (\bar{p}(t, x) - \bar{p}(t, 0)) dt,$$

is ill-defined as soon as  $x \neq y$  since the term  $\bar{p}(t, 0)$  is of order  $t^{-1}$  around 0. To overcome this issue, we introduce the notation  $(\bar{p}(t, \cdot))_{B_1} := \frac{1}{|B_1|} \int_{B_1} \bar{p}(t, z) dz$  and note that, for each  $x \in \mathbb{R}^d$ , the integral

$$\int_0^\infty (\bar{p}(t, x) - (\bar{p}(t, \cdot))_{B_1}) dt$$

is well defined. Additionally, the function

$$x \mapsto \theta(p)^{-1} \left( \int_0^\infty \bar{p}(t, x) - (\bar{p}(t, \cdot))_{B_1} dt \right)$$

is equal to  $\bar{g}$  up to a constant (see [45], Chapter 1.8, for detailed discussions). We denote this constant by  $K_1$ , that is, we write, for any  $x \in \mathbb{R}^d \setminus \{0\}$ ,

$$(161) \quad K_1 := \theta(p)^{-1} \left( \int_0^\infty \bar{p}(t, x) - (\bar{p}(t, \cdot))_{B_1} dt \right) - \bar{g}(x).$$

We note that the value  $K_1$  depends only on the diffusivity  $\bar{\sigma}^2$ . Using these two integrals, we have

$$\begin{aligned} g(x) - \bar{g}(x) &= \int_0^\infty (p(t, x, 0) - p(t, 0, 0)) - \theta(p)^{-1}(\bar{p}(t, x) - (\bar{p}(t, \cdot))_{B_1(y)}) dt + K_1 \\ &= \int_0^\infty (p(t, x, 0) - \theta(p)^{-1}\bar{p}(t, x)) dt + K_1 - K_2(0), \end{aligned}$$

where  $K_2$  is defined by the formula

$$(162) \quad K_2(0) := \int_0^\infty p(t, 0, 0) - \theta(p)^{-1}(\bar{p}(t, \cdot))_{B_1} dt.$$

We now prove that this integral is well defined, and that the constant  $K_2$  satisfies the stochastic integrability estimate

$$|K_2(0)| \leq \mathcal{O}_s(C).$$

The proof relies on Theorem 1.1 and on the estimates on the discrete heat kernel  $p(t, 0, 0) \leq 1$  and  $(\bar{p}(t, \cdot))_{B_1} \leq 1$  for all times  $t$ . We compute

$$\begin{aligned} |K_2(0)| &\leq \int_0^\infty |p(t, 0, 0) - \theta(p)^{-1}(\bar{p}(t, \cdot))_{B_1}| dt \\ &\leq \int_0^{\mathcal{T}_{\text{par},\delta}(0)} |p(t, 0, 0) - \theta(p)^{-1}(\bar{p}(t, \cdot))_{B_1}| dt \\ &\quad + \int_{\mathcal{T}_{\text{par},\delta}(0)}^\infty |p(t, 0, 0) - \theta(p)^{-1}\bar{p}(t, 0)| dt \\ &\quad + \theta(p)^{-1} \int_{\mathcal{T}_{\text{par},\delta}(0)}^\infty |\bar{p}(t, 0) - (\bar{p}(t, \cdot))_{B_1}| dt \\ &\leq \int_0^{\mathcal{T}_{\text{par},\delta}(0)} C dt + \int_{\mathcal{T}_{\text{par},\delta}(0)}^\infty Ct^{-\frac{3}{2}+\delta} dt + \int_{\mathcal{T}_{\text{par},\delta}(0)}^\infty Ct^{-\frac{3}{2}} dt \\ &\leq C\mathcal{T}_{\text{par},\delta}(0) + C. \end{aligned}$$

This implies the estimate  $|K_2(0)| \leq \mathcal{O}_s(C)$ . We define  $K(0) := K_1 - K_2(0)$ , and by the previous computation, it satisfies the stochastic integrability estimate  $|K(0)| \leq \mathcal{O}_s(C)$ .

To complete the proof Theorem 1.2 in dimension 2, it is thus sufficient to control the term  $\int_0^\infty (p(t, x, 0) - \theta(p)^{-1}\bar{p}(t, x)) dt$ ; the argument is the same than in dimension larger than 3 and the details are omitted.  $\square$

### APPENDIX A: A CONCENTRATION INEQUALITY FOR THE DENSITY OF THE INFINITE CLUSTER

In this Appendix, we study the density of the infinite cluster in a cube  $\square$ , which is defined as the random variable  $\frac{|\mathcal{C}_\infty \cap \square|}{|\square|}$ . As the size of the cube tends to infinity, an application of the ergodic theorem shows that this random variable converges, almost surely and in  $L^1$ , to the value  $\theta(\mathfrak{p})$ . The objective of the following proposition is to provide a quantitative version of this result.

**PROPOSITION 11.** *There exists a positive constant  $C(d, \mathfrak{p}) < \infty$  such that for any triadic cube  $\square \in \mathcal{T}$  of size  $3^m$ , one has an estimate*

$$(163) \quad \left| \frac{|\mathcal{C}_\infty \cap \square|}{|\square|} - \theta(\mathfrak{p}) \right| \leq \mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}}(C3^{-\frac{dm}{2}}).$$

As a corollary, we obtain that, for any exponent  $\alpha > 0$ , there exist a positive constant  $C(d, \mathbf{p}, \alpha) < \infty$ , an exponent  $s(d, \mathbf{p}, \alpha) > 0$  and a minimal scale  $\mathcal{M}_{\text{dense}, \alpha} \leq \mathcal{O}_s(C)$  such that for every  $3^m \geq \mathcal{M}_{\text{dense}, \delta}(y)$ ,

$$(164) \quad \left| \frac{|\mathcal{C}_\infty \cap \square_m|}{|\square_m|} - \theta(\mathbf{p}) \right| \leq 3^{-(\frac{d}{2} - \alpha)m}.$$

REMARK 16. The stochastic integrability exponent  $\frac{2(d-1)}{3d^2+2d-1}$  in the estimate (163) is suboptimal and we do not try to reach optimality. The spatial scaling is the one of the central limit theorem and is optimal. We note that a result of large deviation for the concentration of the density of the infinite cluster can be found in the article [76], Theorem 1.2, of Pisztora: for any  $\epsilon > 0$  and  $\mathbf{p} > \mathbf{p}_c(d)$ , there exist two constants  $C_1(\mathbf{p}, d, \epsilon) < \infty$ ,  $C_2(\mathbf{p}, d, \epsilon) < \infty$  such that for any cube  $\square$  of size  $3^m$ ,

$$\mathbb{P}\left(\left| \frac{|\mathcal{C}_\infty \cap \square|}{|\square|} - \theta(\mathbf{p}) \right| > \epsilon\right) \leq C_1 \exp(-C_2 3^{(d-1)m}).$$

However, this estimate cannot be used in the setting considered in this article since the dependence of the constants  $C_1$  and  $C_2$  in the variable  $\epsilon$  is not explicit.

We prove Proposition 11 with an exponential version of the Efron–Stein inequality. A proof of this result can be found in Proposition 2.2 of [16]. In the context of supercritical percolation, this inequality was used in Proposition 2.18, Proposition 3.3 of [40] to study the corrector and in Proposition 3.2 of [61] to study the flux. It is stated in the following proposition and we recall the notation introduced in Section 1.6.1: we denote by  $(\Omega, \mathcal{F}, \mathbb{P})$  the probability space and by  $\mathcal{F}(\mathcal{B}_d \setminus \{e\})$  denotes the sigma algebra generated by the collection of random variables  $\{\mathbf{a}(e')\}_{e' \in \mathcal{B}_d \setminus \{e\}}$ .

PROPOSITION 12 (Exponential Efron–Stein inequality, Proposition 2.2 of [16]). *Fix an exponent  $\beta \in (0, 2)$  and let  $X$  be a random variable defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We define the random variables*

$$(165) \quad X_e := \mathbb{E}[X | \mathcal{F}(\mathcal{B}_d \setminus \{e\})], \quad \mathbb{V}[X] := \sum_{e \in \mathcal{B}_d} (X - X_e)^2.$$

There exists a positive constant  $C := C(d, \beta) < \infty$  such that

$$(166) \quad \mathbb{E}[\exp(|X - \mathbb{E}[X]|^\beta)] \leq C \mathbb{E}[\exp((C \mathbb{V}[X])^{\frac{\beta}{2-\beta}})^{\frac{2-\beta}{2}}].$$

We define  $X := \frac{|\mathcal{C}_\infty \cap \square|}{|\square|} - \theta(\mathbf{p})$ . To prove Proposition 11, it suffices to prove the two inequalities

$$(167) \quad \mathbb{E}[X] \leq C_1(\mathbf{p}, d) 3^{-\frac{dm}{2}} \quad \text{and} \quad \mathbb{V}[X] \leq \mathcal{O}_{s'}(C_2(d, \mathbf{p}) 3^{-dm}),$$

and to use the estimate (166) to deduce that  $X \leq \mathcal{O}_s(C)$  with the exponent  $s = \frac{2s'}{1+s'}$ . These two inequalities are natural since they mean that the bias and variance of the random variable satisfy the desired upper bounds. Since the random variable  $X = \frac{1}{|\square|} \sum_{x \in \square} (\mathbb{1}_{\{x \in \mathcal{C}_\infty\}} - \theta(\mathbf{p}))$  is centered, we can focus on the term  $\mathbb{V}[X]$ .

To estimate this term, we consider an independent copy of the environment  $\mathbf{a}$  which we denote by  $\tilde{\mathbf{a}}$  (and enlarge the underlying probability space to achieve this if necessary). Given a bond  $e \in \mathcal{B}_d$ , we define  $\{\mathbf{a}^e(e')\}_{e' \in \mathcal{B}_d}$  “the environment obtained by resampling the conductance at the bond  $e$ ” by the formula

$$\mathbf{a}^e(e') = \begin{cases} \mathbf{a}(e') & \text{if } e' \neq e, \\ \tilde{\mathbf{a}}(e') & \text{if } e' = e. \end{cases}$$

We denote by  $X^e$  the random variable obtained by resampling the bond  $e$ , that is,  $X^e = X(\mathbf{a}^e)$ . We also denote by  $\mathcal{C}_\infty^e$  the infinite cluster under the environment  $\mathbf{a}^e$ . We have the following implication

$$(168) \quad \sum_{e \in \mathcal{B}_d} (X^e - X)^2 \leq \mathcal{O}_{s'}(C3^{-dm}) \implies \mathbb{V}[X] \leq \mathcal{O}_{s'}(C3^{-dm}),$$

whose proof can be found in [61], Lemma 3.1. We note that since the two environments  $\{\mathbf{a}(e')\}_{e' \in \mathcal{B}_d}$  and  $\{\mathbf{a}^e(e')\}_{e' \in \mathcal{B}_d}$  are only different on one bond, the following statement holds  $\mathbb{P}$ -almost surely:

$$\mathcal{C}_\infty^e \subseteq \mathcal{C}_\infty \quad \text{or} \quad \mathcal{C}_\infty \subseteq \mathcal{C}_\infty^e.$$

We have the following identity:

$$|X^e - X| = \frac{1}{|\square|} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|,$$

where  $\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty := (\mathcal{C}_\infty^e \setminus \mathcal{C}_\infty) \cup (\mathcal{C}_\infty \setminus \mathcal{C}_\infty^e)$  denotes the symmetric difference between the two clusters  $\mathcal{C}_\infty$  and  $\mathcal{C}_\infty^e$ . This suggests to study the properties of this quantity and we prove the following lemma.

LEMMA A.1. *The following estimates hold:*

1. *There exists a positive constant  $C(d, p) < \infty$  such that*

$$(169) \quad \forall e \in \mathcal{B}_d, \quad |\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty| \leq \mathcal{O}_{\frac{d-1}{d}}(C).$$

2. *There exists a positive constant  $C(d, p) < \infty$  such that*

$$(170) \quad \forall e \in \mathcal{B}_d \setminus \mathcal{B}_d(3\square), \quad |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2 \leq \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(\frac{C}{\text{dist}(e, \square)^{d+1}}\right),$$

where we recall the notation  $3\square$  introduced in (27). As a corollary, we have that

$$(171) \quad \sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(3\square)} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2 \leq \mathcal{O}_{\frac{d-1}{(3d+1)d}}(C).$$

We first show how to obtain Proposition 11 from Lemma A.1 and then prove Lemma A.1.

PROOF OF PROPOSITION 11. The result is a consequence of the estimate (23) and Lemma A.1. We have

$$\begin{aligned} \sum_{e \in \mathcal{B}_d} (X^e - X)^2 &= \frac{1}{|\square|^2} \sum_{e \in \mathcal{B}_d} (|\mathcal{C}_\infty^e \cap \square| - |\mathcal{C}_\infty \cap \square|)^2 \\ &= 3^{-2dm} \underbrace{\sum_{e \in \mathcal{B}_d(3\square)} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2}_{\leq 3^{d(m+1)} \times \mathcal{O}_{\frac{d-1}{2d}}(C)} \\ &\quad + 3^{-2dm} \underbrace{\sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(3\square)} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2}_{\leq \mathcal{O}_{\frac{d-1}{(3d+1)d}}(C)} \\ &\leq \mathcal{O}_{\frac{d-1}{(3d+1)d}}(C3^{-dm}). \end{aligned}$$

We then complete the proof by applying the implication (168) and Proposition 12 with the exponent  $s = \frac{2(d-1)}{3d^2+2d-1}$ .  $\square$

We now prove Lemma A.1. The argument relies on the upper and lower bounds on the tail of the distribution of the finite clusters in supercritical percolation. The result is stated below, was proved by Kesten and Zhang in [62] for the upper bound and by Aizenman, Delyon and Souillard in [1] for the lower bound. We also refer to the monograph [60], Section 8.6, for related discussions.

**THEOREM A.2** (Subexponential decay of cluster size distribution [1, 62]). *For any supercritical probability  $\mathbf{p} \in (\mathbf{p}_c(d), 1]$ , there exist positive constants  $0 < c_1(d, \mathbf{p}), c_2(d, \mathbf{p}) < \infty$  such that, if we denote by  $\mathcal{C}(0)$  the cluster containing 0 and let  $n$  be a strictly positive integer, then we have the estimate*

$$(172) \quad \forall n \in \mathbb{N}^+, \quad \exp(-c_1 n^{\frac{d-1}{d}}) \leq \mathbb{P}[|\mathcal{C}(0)| = n] \leq \exp(-c_2 n^{\frac{d-1}{d}}).$$

**REMARK 17.** With the notation  $\mathcal{O}_s$ , one can reformulate the upper bound as  $|\mathcal{C}(0)| \leq \mathcal{O}_{\frac{d-1}{d}}(C)$ .

**REMARK 18.** The estimate (172) implies the inequality  $\mathbb{P}[n \leq |\mathcal{C}(0)| < \infty] \leq \exp(-c_3 n^{\frac{d-1}{d}})$ .

**PROOF OF LEMMA A.1.** We first prove the inequality (169). We use the definition of  $\mathcal{O}_s$  notation and prove the estimate

$$\mathbb{E} \left[ \exp \left( \left( \frac{|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty|}{C} \right)^{\frac{d-1}{d}} \right) \right] \leq 2,$$

for some constant  $C(d, \mathbf{p}) < \infty$ . By symmetry, it suffices to consider the case when  $\{\mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0\}$  and we have the identity

$$\mathbb{E} \left[ \exp \left( \left( \frac{|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty|}{C} \right)^{\frac{d-1}{d}} \right) \right] = 1 + 2\mathbb{E} \left[ \exp \left( \left( \frac{|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty|}{C} \right)^{\frac{d-1}{d}} \right) \mathbb{1}_{\{\mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0\}} \right].$$

We then notice that, under the condition  $\{\mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0\}$ , we have the equality  $\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty = \mathcal{C}_\infty^e \setminus \mathcal{C}_\infty$ . We then distinguish two cases:

- Either there exists a finite cluster connected to the bond  $e$  in the environment  $\{\mathbf{a}(e')\}_{e' \in \mathcal{B}_d}$ . In that case, we denote this cluster by  $\mathcal{C}(e)$  and we have the identity

$$\mathcal{C}(e) = \mathcal{C}_\infty^e \Delta \mathcal{C}_\infty;$$

- Or both ends of the bond  $e$  are connected to the infinite cluster  $\mathcal{C}_\infty$  under the environment  $\{\mathbf{a}(e')\}_{e' \in \mathcal{B}_d}$ . In that case, we have the equality  $\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty = \emptyset$ .

We then use Theorem A.2 to estimate the volume of the cluster  $\mathcal{C}(e)$  and we obtain

$$\mathbb{E} \left[ \exp \left( \left( \frac{|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty|}{C} \right)^{\frac{d-1}{d}} \right) \mathbb{1}_{\{\mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0\}} \right] \leq \sum_{n=0}^{\infty} \exp \left( \frac{n^{\frac{d-1}{d}}}{C^{\frac{d-1}{d}}} \right) \exp(-c_1 n^{\frac{d-1}{d}}).$$

Then we can choose a constant  $C$  depending on the parameters  $d$  and  $\mathbf{p}$  such that

$$\mathbb{E} \left[ \exp \left( \left( \frac{|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty|}{C} \right)^{\frac{d-1}{d}} \right) \right] \leq 2.$$

This implies that  $|\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty| \leq \mathcal{O}_{\frac{d-1}{d}}(C)$ .

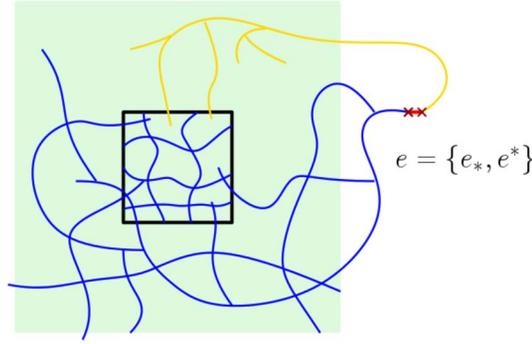


FIG. 5. The figure illustrates the situation when the set  $(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square$  is nonempty under the condition  $\mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0$ . The blue cluster is the infinite cluster  $\mathcal{C}_\infty$  under the environment  $\mathbf{a}$ , and the yellow cluster is the finite cluster connecting the bond  $e$  to the cube  $\square$ . The green square represents the cube  $3\square$ . The probability of the event depicted in the picture becomes exponentially small when the sizes of the cubes are large.

We now prove the estimate (170). It relies on the following observation (see Figure 5): when the bond  $e$  is far away from the cube  $\square$ , the set  $(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square$  is nonempty with exponentially small probability. More precisely, if we denote by  $l = \text{dist}(e, \square)$ , then we have the estimate

$$\begin{aligned}
 & \mathbb{P}[(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square \neq \emptyset] \\
 (173) \quad &= 2\mathbb{P}[\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty \neq \emptyset \text{ and } \mathcal{C}(e) \cap \square \neq \emptyset \text{ and } \mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0] \\
 &\leq 2\mathbb{P}[\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty \neq \emptyset \text{ and } |\mathcal{C}(e)| > l \text{ and } \mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0] \\
 &\leq 2 \exp(-c_3 l^{\frac{d-1}{d}}).
 \end{aligned}$$

We also note that, since the bond  $e$  lies outside the cube  $3\square$ , we have the estimate  $l \geq 3^m$ . This implies the almost sure inequalities

$$|(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square| \leq 3^{dm} \leq l^d.$$

Then we can calculate the expectation

$$\begin{aligned}
 & \mathbb{E} \left[ \exp \left( \left( \frac{l^{d+1} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2}{C} \right)^{\frac{d-1}{(3d+1)d}} \right) \right] \\
 &= \mathbb{E}[\mathbb{1}_{\{(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square = \emptyset\}}] \\
 &\quad + \mathbb{E} \left[ \exp \left( \left( \frac{l^{d+1} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2}{C} \right)^{\frac{d-1}{(3d+1)d}} \right) \mathbb{1}_{\{(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square \neq \emptyset\}} \right] \\
 &\leq 1 + \exp \left( \left( \frac{l^{3d+1}}{C} \right)^{\frac{d-1}{(3d+1)d}} \right) \mathbb{P}[(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square \neq \emptyset] \\
 &\leq 1 + \exp \left( \frac{l^{\frac{d-1}{d}}}{C^{\frac{d-1}{(3d+1)d}}} \right) \mathbb{P}[(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square \neq \emptyset].
 \end{aligned}$$

We use the estimate (173) and select a constant  $C$  large enough such that

$$\mathbb{E} \left[ \exp \left( \left( \frac{l^{d+1} |(\mathcal{C}_\infty^e \Delta \mathcal{C}_\infty) \cap \square|^2}{C} \right)^{\frac{d-1}{(3d+1)d}} \right) \right] \leq 1 + 2 \exp \left( \frac{l^{\frac{d-1}{d}}}{C^{\frac{d-1}{(3d+1)d}}} \right) \exp(-c_4 l^{\frac{d-1}{d}}) \leq 2.$$

This completes the proof of the inequality (170). The estimate (171) is then a consequence of the inequality (23), by noting that the sum  $\sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(3\square)} \text{dist}(e, \square)^{-d-1}$  is finite. Finally, we

define

$$\mathcal{M}_{\text{dense},\alpha} := \sup \left\{ 3^m \in \mathbb{N} : 3^{(\frac{d}{2}-\alpha)m} \left| \frac{|\mathcal{C}_\infty \cap \square_m|}{|\square_m|} - \theta(\mathbf{p}) \right| \geq 1 \right\},$$

and use Lemma 1.4 to obtain that this random variable satisfies the stochastic integrability estimate  $\mathcal{M}_{\text{dense},\alpha} \leq \mathcal{O}_s(C)$ .  $\square$

We complete this section by stating and proving a version of the concentration estimate of Lemma A.1 involving the homogenized heat-kernel. This result is used in Lemma 4.1.

PROPOSITION 13. *There exists a positive constant  $C(d, \mathbf{p}) < \infty$  such that, for any time  $t > 0$ , and any vertex  $y \in \mathbb{Z}^d$ , one has the estimate*

$$(174) \quad \left| \int_{\mathcal{C}_\infty} \bar{p}(t, x - y) dx - \theta(\mathbf{p}) \right| \leq \mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}}(Ct^{-\frac{1}{2}}).$$

As a corollary, for any  $\alpha > 0$  and  $y \in \mathbb{Z}^d$ , there exist a positive constant  $C(d, \mathbf{p}, \alpha) < \infty$ , an exponent  $s(d, \mathbf{p}, \alpha) > 0$  and a minimal time  $\mathcal{T}_{\text{dense},\alpha}(y) \leq \mathcal{O}_s(C)$  such that, for every time  $t > \mathcal{T}_{\text{dense},\alpha}(y)$ , we have

$$(175) \quad \left| \int_{\mathcal{C}_\infty} \bar{p}(t, x - y) dx - \theta(\mathbf{p}) \right| \leq Ct^{-(\frac{1}{2}-\alpha)}.$$

PROOF. Without loss of generality, we suppose that  $y = 0$ . The strategy is similar to the one of the proof of Proposition 11. We denote by  $Y := \int_{\mathcal{C}_\infty} \bar{p}(t, x) dx - \theta(\mathbf{p})$ , apply the concentration inequality stated in Proposition 12 and verify the two conditions (167) and (168). For the term involving the expectation, we have

$$\begin{aligned} |\mathbb{E}[Y]| &= \left| \int_{\mathbb{Z}^d} \bar{p}(t, x) \mathbb{1}_{\{x \in \mathcal{C}_\infty\}} dx - \theta(\mathbf{p}) \right| \\ &= \theta(\mathbf{p}) \left| \int_{\mathbb{Z}^d} \bar{p}(t, x) dx - \int_{\mathbb{R}^d} \bar{p}(t, x) dx \right| \leq \frac{C}{\sqrt{t}}, \end{aligned}$$

by the estimate on the gradient of the heat kernel. We then focus on the variation, that is,

$$\sum_{e \in \mathcal{B}_d} (Y^e - Y)^2 = \sum_{e \in \mathcal{B}_d} \left( \int_{\mathbb{Z}^d} \bar{p}(t, x) (\mathbb{1}_{\{x \in \mathcal{C}_\infty^e\}} - \mathbb{1}_{\{x \in \mathcal{C}_\infty\}}) dx \right)^2.$$

We apply a multiscale analysis: we define the balls and annuli

$$\begin{aligned} B_{-1} &:= \emptyset, \quad \forall n \geq 1, \\ B_n &:= \{x \in \mathbb{Z}^d : |x| \leq 3^n \sqrt{t}\}, \quad \forall n \geq 0, \\ A_n &:= B_n \setminus B_{n-1}. \end{aligned}$$

We also define, for any subset  $A \subseteq \mathbb{Z}^d$ ,  $I_A^e := \int_A \bar{p}(t, x) (\mathbb{1}_{\{x \in \mathcal{C}_\infty^e\}} - \mathbb{1}_{\{x \in \mathcal{C}_\infty\}}) dx$ . This notation is useful to localize the random variables  $(Y^e - Y)$ . We write

$$\sum_{e \in \mathcal{B}_d} (Y^e - Y)^2 = \sum_{e \in \mathcal{B}_d} (I_{\mathbb{Z}^d}^e)^2 = \sum_{e \in \mathcal{B}_d} \left( \sum_{n=0}^\infty I_{A_n}^e \right)^2.$$

Then we use the Cauchy–Schwarz inequality to factorize the sum

$$\left( \sum_{n=0}^\infty I_{A_n}^e \right)^2 = \left( \sum_{n=0}^\infty I_{A_n}^e 3^n \times 3^{-n} \right)^2 \leq \left( \sum_{n=0}^\infty 3^{2n} (I_{A_n}^e)^2 \right) \left( \sum_{n=0}^\infty 3^{-2n} \right) \leq 2 \sum_{n=0}^\infty 3^{2n} (I_{A_n}^e)^2.$$

With Fubini's theorem, we obtain

$$(176) \quad \sum_{e \in \mathcal{B}_d} (Y^e - Y)^2 \leq 2 \sum_{n=0}^{\infty} 3^{2n} \sum_{e \in \mathcal{B}_d} (I_{A_n}^e)^2.$$

We fix an integer  $n \in \mathbb{N}$  and estimate the quantity  $\sum_{e \in \mathcal{B}_d} (I_{A_n}^e)^2$ . The strategy is similar to the proof of Proposition 11 and we adapt the proof of Lemma A.1 from the case of cubes to the case of balls

$$\begin{aligned} \sum_{e \in \mathcal{B}_d} (I_{A_n}^e)^2 &= \sum_{e \in \mathcal{B}_d(B_{n+1})} (I_{A_n}^e)^2 + \sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(B_{n+1})} (I_{A_n}^e)^2 \\ &\leq \left( \max_{A_n} \bar{p}(t, x) \right)^2 \\ &\quad \times \left( \sum_{e \in \mathcal{B}_d(B_{n+1})} |(\mathcal{C}_{\infty}^e \Delta \mathcal{C}_{\infty}) \cap A_n|^2 + \sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(B_{n+1})} |(\mathcal{C}_{\infty}^e \Delta \mathcal{C}_{\infty}) \cap A_n|^2 \right) \\ &\leq \frac{1}{(2\pi t \bar{\sigma}^2)^{d/2}} \exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right) \\ &\quad \times \left( \sum_{e \in \mathcal{B}_d(B_{n+1})} |(\mathcal{C}_{\infty}^e \Delta \mathcal{C}_{\infty}) \cap B_n|^2 + \sum_{e \in \mathcal{B}_d \setminus \mathcal{B}_d(B_{n+1})} |(\mathcal{C}_{\infty}^e \Delta \mathcal{C}_{\infty}) \cap B_n|^2 \right) \\ &\leq \mathcal{O}_{\frac{d-1}{(3d+1)d}} \left( C 3^{dn} t^{-\frac{d}{2}} \exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right) \right). \end{aligned}$$

We put this inequality back in equation (176) and use the estimate (23) to conclude

$$\sum_{e \in \mathcal{B}_d} (Y^e - Y)^2 \leq \mathcal{O}_{\frac{d-1}{(3d+1)d}} \left( C t^{-\frac{d}{2}} \left( \sum_{n=0}^{\infty} 3^{(d+2)n} \exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right) \right) \right) \leq \mathcal{O}_{\frac{d-1}{(3d+1)d}} (C t^{-\frac{d}{2}}).$$

Finally, for any exponent  $\alpha > 0$ , we define

$$\mathcal{T}_{\text{dense}, \alpha}(0) := \sup \left\{ t \in (0, \infty) : t^{(\frac{1}{2}-\alpha)} \left| \int_{\mathcal{C}_{\infty}} \bar{p}(t, x) dx - \theta(\mathbf{p}) \right| \geq 1 \right\},$$

and apply Lemma 1.4 to conclude Proposition 13.  $\square$

### APPENDIX B: QUANTIFICATION OF THE WEAK NORM OF THE FLUX ON THE INFINITE CLUSTER

In this Appendix, we prove a quantification of the  $\underline{H}^{-1}$ -norm of the flux on the cluster. We recall that the flux on the cluster associated to the direction  $e_k$  is defined by

$$(177) \quad \tilde{\mathbf{g}}_{e_k} : \mathcal{C}_{\infty} \rightarrow \mathbb{R}^d, \quad \tilde{\mathbf{g}}_{e_k} = \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k.$$

The main estimate is stated in the following proposition.

**PROPOSITION 14.** *Fix a point  $y \in \mathbb{Z}^d$ , for each exponent  $\alpha > 0$ , there exist a positive constant  $C := C(\lambda, d, \mathbf{p}, \alpha) < \infty$ , an exponent  $s := s(\lambda, d, \mathbf{p}, \alpha) > 0$ , and a random variable  $\mathcal{M}_{\text{flux}, \alpha}(y)$  satisfying the stochastic integrability estimate*

$$\mathcal{M}_{\text{flux}, \alpha}(y) \leq \mathcal{O}_s(C),$$

such that, for every radius  $r \geq \mathcal{M}_{\text{flux},\alpha}(y)$ , one has

$$(178) \quad \sum_{k=1}^d \|\tilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r(y))} \leq Cr^\alpha.$$

Without loss of generality, we assume  $y = 0$ . The strategy of the proof is to make use of another centered flux defined on the entire space  $\mathbb{Z}^d$ ,

$$\mathbf{g}_{e_k} : \mathbb{Z}^d \rightarrow \mathbb{R}^d, \quad \mathbf{g}_{e_k} = \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \bar{\mathbf{a}}e_k.$$

The homogenized conductance  $\bar{\mathbf{a}}$  is defined in [12], Definition 5.1, by the formula: for each  $p \in \mathbb{R}^d$ ,

$$\frac{1}{2}p \cdot \bar{\mathbf{a}}p := \lim_{n \rightarrow \infty} \mathbb{E}[v(\square_m, p)],$$

where the energy  $v(\square_m, p)$  is defined by

$$(179) \quad v(\square_m, p) := \inf_{u \in l_p + C_0(\mathcal{C}_\infty \cap \square_m)} \frac{1}{2|\square_m|} \int_{\square_m} \nabla u \cdot \mathbf{a} \nabla u,$$

where the notation  $l_p$  denotes the affine function of slope  $p$  (i.e., for each point  $x \in \mathbb{Z}^d$ ,  $l_p(x) = p \cdot x$ ) and the symbol  $C_0(\mathcal{C}_\infty \cap \square_m)$  denotes the set of functions defined on the set  $\mathcal{C}_\infty \cap \square_m$ , valued in  $\mathbb{R}$ , which are equal to 0 on the boundary  $\mathcal{C}_\infty \cap \partial \square_m$ . The reason we introduce this quantity is that, building upon the results of [12], we can prove the following  $\underline{H}^{-1}$ -estimate: there exists a nonnegative random variable  $\mathcal{M}_{\text{flux}-\mathbb{Z}^d,\alpha}$  satisfying the stochastic integrability estimate  $\mathcal{M}_{\text{flux}-\mathbb{Z}^d,\alpha} \leq \mathcal{O}_s(C)$  such that, for every  $r \geq \mathcal{M}_{\text{flux}-\mathbb{Z}^d,\alpha}$ ,

$$(180) \quad \sum_{k=1}^d \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leq Cr^\alpha,$$

where the  $\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)$ -norm is defined by the formula

$$\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} = \sup_{\|\varphi\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} \leq 1} \frac{1}{|\mathbb{Z}^d \cap B_r|} \int_{\mathbb{Z}^d \cap B_r} \varphi \mathbf{g}_{e_k},$$

and where the  $\underline{H}^1(\mathbb{Z}^d \cap B_r)$ -norm of a function  $\varphi : \mathbb{Z}^d \cap B_r \rightarrow \mathbb{R}$ , denotes the discrete normalized Sobolev norm defined by the formula

$$\|\varphi\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)}^2 := r^{-1} \|\varphi\|_{\underline{L}^2(\mathbb{Z}^d \cap B_r)} + \|\nabla \varphi\|_{\underline{L}^2(\mathbb{Z}^d \cap B_r)}.$$

Once this result is established, we set the value

$$(181) \quad \bar{\sigma}^2 := 2\theta(\mathfrak{p})^{-1} \bar{\mathbf{a}},$$

and deduce Proposition 14 from the estimate. The main difference between the estimates (178) and (180) is that in the former estimate, the  $\underline{H}^{-1}$ -norm is computed on the ball  $B_r$  while in the latter is computed on the intersection  $\mathcal{C}_\infty \cap B_r$ . This makes an important difference and motivates the introduction of the diffusivity  $\bar{\sigma}^2$  in (181) and of the new flux  $\tilde{\mathbf{g}}_{e_k}$  in (177). In the following paragraph, we give a heuristic argument explaining why we expect Proposition 14 to hold assuming that the estimate (180) is valid.

We start by using the constant test function equal to 1 in the definition of the  $\underline{H}^{-1}(B_r)$  norm in the estimate (180) shows

$$(182) \quad \int_{\mathbb{Z}^d \cap B_r} \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) \simeq \int_{\mathbb{Z}^d \cap B_r} \bar{\mathbf{a}}e_k,$$

where the symbol  $\simeq$  means that the two quantities on the left and right hand sides differ by a small term, which by (180) is of order  $r^{-(1-\alpha)}$ . Since the function  $\mathbf{a}(\mathcal{D}\chi_{e_k} + e_k)$  is defined to be equal to 0 outside the infinite cluster, the left-hand side of (182) can be rewritten

$$\int_{\mathbb{Z}^d \cap B_r} \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) = \int_{\mathcal{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k).$$

For the right-hand side of (182), using that the density of the cluster has density  $\theta(\mathbf{p})$ , one expects

$$\int_{\mathbb{Z}^d \cap B_r} \bar{\mathbf{a}}e_k \simeq \int_{\mathcal{C}_\infty \cap B_r} \theta(\mathbf{p})^{-1} \bar{\mathbf{a}}e_k.$$

This shows

$$\int_{\mathcal{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) \simeq \int_{\mathcal{C}_\infty \cap B_r} \theta(\mathbf{p})^{-1} \bar{\mathbf{a}}e_k.$$

Thus if we want the estimate (178) to hold, the only admissible value for the coefficient  $\bar{\sigma}^2$  is  $2\theta(\mathbf{p})^{-1}\bar{\mathbf{a}}$ , indeed testing the constant function equal to 1 in the definition of the  $\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)$ -norm of the estimate (178) shows

$$\int_{\mathcal{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) \simeq \int_{\mathcal{C}_\infty \cap B_r} \frac{1}{2} \bar{\sigma}^2 e_k.$$

REMARK 19. We note that the identity (181) is the definition of the diffusivity  $\bar{\sigma}^2$  used in this article: thanks to this definition and the result of Proposition 14, we are able to prove Theorem 1.1, and then to recover the invariance principle stated in (4).

The rest of this section is organized as follows. We first explain how to prove the estimate (180) by using the results of [61] and the strategies of stochastic homogenization in the uniformly elliptic setting presented in [15]. We then show how to deduce Proposition 14 from the inequality (180).

PROOF OF THE ESTIMATE (180). We first extend the function  $\mathbf{g}_{e_k}$  from  $\mathbb{Z}^d$  to  $\mathbb{R}^d$  and let  $[\mathbf{g}_{e_k}]$  be the function defined on  $\mathbb{R}^d$ , which is equal to  $\mathbf{g}_{e_k}$  on  $\mathbb{Z}^d$  and which is piecewise constant on the unit cubes  $z + [-\frac{1}{2}, \frac{1}{2}]^d$ . We have the identity  $\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \simeq C \|[\mathbf{g}_{e_k}]\|_{\underline{H}^{-1}(B_r)}$  up to a constant  $C$  depending only on the dimension, where  $\underline{H}^{-1}(B_r)$  is the standard Sobolev norm. We then want to control the continuous  $\underline{H}^{-1}(B_r)$  norm of  $[\mathbf{g}_{e_k}]$ . The strategy is to apply the multiscale Poincaré inequality stated in [15], Remark D.6, equation (D.28). Its rescaled version reads

$$(183) \quad \|[\mathbf{g}_{e_k}]\|_{\underline{H}^{-1}(B_r)} \leq Cr \left( \int_0^1 \left( \int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2t} \star [\mathbf{g}_{e_k}]|^2(x) dx \right) dt \right)^{\frac{1}{2}},$$

where the function  $\Phi_t$  is the standard heat kernel defined by  $\Phi_t := \frac{1}{(2\pi t)^{\frac{d}{2}}} \exp(-\frac{|x|^2}{2t})$  and the operator  $\star$  is the standard convolution on  $\mathbb{R}^d$ . We then apply the following results:

- The spatial average of the flux decays: one has the estimate, for each  $t > 0$ ,

$$(184) \quad |\Phi_t \star [\mathbf{g}_{e_k}]| \leq \mathcal{O}_s(Ct^{-\frac{d}{4}}).$$

A proof of this result can be found in [61], Section 3.1, Proposition 1.1.

- The flux is essentially bounded: one has the estimate, for each  $t > 0$ ,

$$(185) \quad |\Phi_t \star [\mathbf{g}_{e_k}]| \leq \mathcal{O}_s(C).$$

To prove this estimate, we first note that the bound on the corrector stated in (45) implies the following Lipschitz estimate on the corrector (by choosing  $x$  and  $y$  to be two neighboring points): for each vector  $p \in B_1$ , and each edge  $e \in \mathcal{B}_d$ ,

$$(186) \quad |\nabla \chi_p(e)| \leq \mathcal{O}_s(C|p|).$$

This estimate is also stated in [12], Remark 1.1. The inequality (185) is then a consequence of the estimate (186) and the property (23) of the  $\mathcal{O}_s$  notation.

We then truncate the integral in the right-hand side of (183) at the value  $t = r^{-2}$  and obtain

$$\begin{aligned} \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} &\leq Cr \left( \int_0^{r^{-2}} \left( \int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2t} \star [\mathbf{g}_{e_k}]|^2(x) dx \right) dt \right)^{\frac{1}{2}} \\ &\quad + Cr \left( \int_{r^{-2}}^1 \left( \int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2t} \star [\mathbf{g}_{e_k}]|^2(x) dx \right) dt \right)^{\frac{1}{2}}. \end{aligned}$$

To estimate the first term in the right-hand side, we apply the estimate (185) and, to estimate the second term, we apply the estimate (184). Together with the property (23) of the  $\mathcal{O}_s$  notation, this gives

$$(187) \quad \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leq \begin{cases} \mathcal{O}_s(C) & \text{if } d \geq 3, \\ \mathcal{O}_s(\log^{\frac{1}{2}}(1+r)) & \text{if } d = 2. \end{cases}$$

Finally, for every exponent  $\alpha > 0$ , we set

$$\mathcal{M}_{\text{flux}-\mathbb{Z}^d, \alpha} := \sup \left\{ r \in \mathbb{R}^+ : r^{-\alpha} \sum_{k=1}^d \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \geq 1 \right\},$$

and apply Lemma 1.4. This completes the proof of the estimate (180).  $\square$

**PROOF OF PROPOSITION 14.** We fix an exponent  $\alpha > 0$ . We define the exponent  $q := \max(\frac{(2d-1)}{\alpha}, 2d)$  and split the proof into 3 steps.

*Step 1.* In this step, we establish the inequality, for any radius  $r \geq \mathcal{M}_q(\mathcal{P})$ ,

$$(188) \quad \|\tilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} \leq Cr^{\frac{\alpha}{2}} \left( \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} + \left\| \frac{1}{2} \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathbf{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \right),$$

where  $C$  is a constant depending only on the parameters  $\lambda, d, \mathbf{p}$ . We recall the definition of the  $\underline{H}^{-1}$ -norm on the infinite cluster

$$\|\tilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} = \sup_{\|\varphi\|_{\underline{H}^1(\mathcal{C}_\infty \cap B_r)} \leq 1} \frac{1}{|\mathcal{C}_\infty \cap B_r|} \int_{\mathcal{C}_\infty \cap B_r} \varphi \tilde{\mathbf{g}}_{e_k}.$$

We fix a function  $\varphi : \mathcal{C}_\infty \cap B_r \rightarrow \mathbb{R}$  such that  $\|\varphi\|_{\underline{H}^1(\mathcal{C}_\infty \cap B_r)} \leq 1$ . The main idea is to extend the function  $\varphi$  from the infinite cluster to  $\mathbb{Z}^d$ . To this end, we use the coarsened function  $[\varphi]_{\mathcal{P}}$  introduced in Section 2.1.1. We extend the function  $\tilde{\mathbf{g}}_{e_k}$  by 0 outside the infinite cluster so that we have

$$\int_{\mathcal{C}_\infty \cap B_r} \varphi \tilde{\mathbf{g}}_{e_k} = \int_{\mathbb{Z}^d \cap B_r} [\varphi]_{\mathcal{P}} \tilde{\mathbf{g}}_{e_k}.$$

Since the radius  $r$  is assumed to be larger than the minimal scale  $\mathcal{M}_q(\mathcal{P})$ , the ratio  $\frac{|\mathbb{Z}^d \cap B_r|}{|\mathcal{C}_\infty \cap B_r|}$  is bounded from above by a constant  $C(d, \mathfrak{p})$ . Then we compute

$$\begin{aligned}
 & \frac{1}{|\mathcal{C}_\infty \cap B_r|} \int_{\mathcal{C}_\infty \cap B_r} \varphi \tilde{\mathbf{g}}_{e_k} \\
 &= \frac{1}{|\mathcal{C}_\infty \cap B_r|} \int_{\mathbb{Z}^d \cap B_r} [\varphi]_{\mathcal{P}} (\tilde{\mathbf{g}}_{e_k} - \mathbf{g}_{e_k}) + \frac{1}{|\mathcal{C}_\infty \cap B_r|} \int_{\mathbb{Z}^d \cap B_r} [\varphi]_{\mathcal{P}} \mathbf{g}_{e_k} \\
 (189) \quad &\leq \left( \frac{|\mathbb{Z}^d \cap B_r|}{|\mathcal{C}_\infty \cap B_r|} \right) \left\| [\varphi]_{\mathcal{P}} \right\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} \left( \|\tilde{\mathbf{g}}_{e_k} - \mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} + \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \right) \\
 &\leq C(d, \mathfrak{p}) \left\| [\varphi]_{\mathcal{P}} \right\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} \\
 &\quad \times \left( \left\| \frac{1}{2} \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} + \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \right),
 \end{aligned}$$

where we used the equation  $\bar{\mathbf{a}} = \frac{1}{2} \theta(\mathfrak{p}) \bar{\sigma}^2$  to go from the second line to the third line. We then use the estimates (41) and (42) to estimate the  $\underline{H}^1$ -norm of the coarsened function  $\varphi$  in terms of the  $\underline{H}^1$ -norm of the function  $\varphi$ , and the assumption  $r \geq \mathcal{M}_q(\mathcal{P})$  to estimate the size of the cubes of the partition. This gives

$$\left\| [\varphi]_{\mathcal{P}} \right\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} \leq Cr^{\frac{\alpha}{2}} \|\varphi\|_{\underline{H}^1(\mathcal{C}_\infty \cap B_r)}.$$

Combining the previous estimate with the inequality (189) completes the proof of Step 1.

*Step 2: Control over the quantity  $\|\bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p}))\|_{\underline{H}^{-1}(B_r)}$ .* We let  $\square_m$  be the triadic cube such that  $\square_{m-1} \subseteq B_r \subseteq \square_m$ . We note that

$$\left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leq C \left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)},$$

where the constant  $C$  depends only on the dimension  $d$ . We apply another version of the multiscale Poincaré inequality, which is stated in [15], Proposition 1.7, (in the continuous setting, the extension to the discrete setting considered here does not affect the proof) and reads

$$\begin{aligned}
 & \left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)} \\
 & \leq C \sum_{n=0}^{m-1} 3^n \left( \frac{1}{|3^n \mathbb{Z}^d \cap \square_m|} \sum_{y \in 3^n \mathbb{Z}^d \cap \square_m} \bar{\sigma}^4 (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p}))_{y+\square_n}^2 \right)^{\frac{1}{2}},
 \end{aligned}$$

where we recall the notation  $(f)_{y+\square_n} = \frac{1}{|\square_n|} \sum_{x \in y+\square_n} f(x)$ . We apply Proposition 11

$$(\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p}))_{y+\square_n} \leq \mathcal{O}_s(C3^{-\frac{dn}{2}}).$$

Using that the dimension is larger than 2 and the property (23) of the  $\mathcal{O}_s$  notation, we obtain

$$\left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)} \begin{cases} \mathcal{O}_s(C) & \text{if } d \geq 3, \\ \mathcal{O}_s(Cm) & \text{if } d = 2. \end{cases}$$

We then apply Lemma 1.4 to the collection of random variables

$$X_m := 3^{-\frac{\alpha m}{2}} \left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)},$$

to construct a minimal scale  $\mathcal{M}_{\text{cluster}, \frac{\alpha}{2}}$  such that, for any radius  $r \geq \mathcal{M}_{\text{cluster}, \frac{\alpha}{2}}$ ,

$$(190) \quad \left\| \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leq Cr^{\frac{\alpha}{2}}.$$

*Step 3: The conclusion.* We let  $\mathcal{M}_{\text{flux}-\mathbb{Z}^d, \frac{\alpha}{2}}$  be the minimal scale provided by equation (180) with the exponent  $\frac{\alpha}{2}$ . We define the random variable  $\mathcal{M}_{\text{flux}, \alpha}(0)$  according to the formula

$$\mathcal{M}_{\text{flux}, \alpha}(0) := \max(\mathcal{M}_{\text{cluster}, \frac{\alpha}{2}}, \mathcal{M}_{\text{flux}-\mathbb{Z}^d, \frac{\alpha}{2}}, \mathcal{M}_q(\mathcal{P})).$$

Combining the main results (188) of Step 1 and (190) of Step 2 shows, for any  $r \geq \mathcal{M}_{\text{flux}, \alpha}(0)$ ,

$$\begin{aligned} \|\tilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r)} &\leq Cr^{\frac{\alpha}{2}} \left( \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} + \left\| \frac{1}{2} \bar{\sigma}^2 e_k (\mathbb{1}_{\mathcal{C}_\infty} - \theta(\mathbf{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \right) \\ &\leq Cr^{\frac{\alpha}{2}} (r^{\frac{\alpha}{2}} + r^{\frac{\alpha}{2}}) \\ &\leq Cr^\alpha. \end{aligned}$$

Thus, we obtain the main result (178) of Proposition 14.  $\square$

REMARK 20. One result used in this article is a variation of Proposition 14. We are interested in another function  $\tilde{\mathbf{g}}_{e_k}^* : \mathcal{C}_\infty \rightarrow \mathbb{R}^d$ , satisfying the identity, for each function  $u : \mathcal{C}_\infty \rightarrow \mathbb{R}$ ,

$$(191) \quad \mathcal{D}^* \cdot \left( u \left( \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k \right) \right) = \mathcal{D}^* u \cdot \tilde{\mathbf{g}}_{e_k}^*.$$

One can check that the quantity  $\tilde{\mathbf{g}}_{e_k}^*$  is different from  $\tilde{\mathbf{g}}_{e_k}$  with an exact formula

$$\tilde{\mathbf{g}}_{e_k}^* := \begin{pmatrix} T_{-e_1} \left[ \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k \right]_1 \\ \vdots \\ T_{-e_d} \left[ \mathbf{a}(\mathcal{D}\chi_{e_k} + e_k) - \frac{1}{2} \bar{\sigma}^2 e_k \right]_d \end{pmatrix},$$

where the (minor) difference comes from the translation when applying the finite difference operator. The  $\underline{H}^{-1}$ -norm of the function  $\tilde{\mathbf{g}}_{e_k}^*$  can also be controlled and one has the following property: for any exponent  $\alpha > 0$ , any vertex  $y \in \mathbb{Z}^d$ , and any radius  $r \geq \mathcal{M}_{\text{flux}, \alpha}(y)$ , one has the estimate

$$\sum_{k=1}^d \|\tilde{\mathbf{g}}_{e_k}^*\|_{\underline{H}^{-1}(\mathcal{C}_\infty \cap B_r(y))} \leq Cr^\alpha.$$

The proof is identical to the proof of Proposition 14 and the details are left to the reader.

**Acknowledgments.** We would like to thank Jean-Christophe Mourrat and Scott Armstrong for helpful discussions and comments. PD is supported by the Israel Science Foundation grants 861/15 and 1971/19 and by the European Research Council starting grant 678520 (LocalOrder). CG is supported by the PhD scholarship from École Polytechnique.

REFERENCES

[1] AIZENMAN, M., DELYON, F. and SOUILLARD, B. (1980). Lower bounds on the cluster size distribution. *J. Stat. Phys.* **23** 267–280. MR0594824 <https://doi.org/10.1007/BF01011369>

[2] ALVES, C. and SAPOZHNIKOV, A. (2019). Decoupling inequalities and supercritical percolation for the vacant set of random walk loop soup. *Electron. J. Probab.* **24** 110. MR4017128 <https://doi.org/10.1214/19-ejp360>

[3] ANDRES, S., BARLOW, M. T., DEUSCHEL, J.-D. and HAMBLY, B. M. (2013). Invariance principle for the random conductance model. *Probab. Theory Related Fields* **156** 535–580. MR3078279 <https://doi.org/10.1007/s00440-012-0435-2>

- [4] ANDRES, S., CHIARINI, A., DEUSCHEL, J.-D. and SLOWIK, M. (2018). Quenched invariance principle for random walks with time-dependent ergodic degenerate weights. *Ann. Probab.* **46** 302–336. MR3758732 <https://doi.org/10.1214/17-AOP1186>
- [5] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2015). Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.* **43** 1866–1891. MR3353817 <https://doi.org/10.1214/14-AOP921>
- [6] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2016). Harnack inequalities on weighted graphs and some applications to the random conductance model. *Probab. Theory Related Fields* **164** 931–977. MR3477784 <https://doi.org/10.1007/s00440-015-0623-y>
- [7] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2016). Heat kernel estimates for random walks with degenerate weights. *Electron. J. Probab.* **21** 33. MR3492937 <https://doi.org/10.1214/16-EJP4382>
- [8] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2020). Green kernel asymptotics for two-dimensional random walks under random conductances. *Electron. Commun. Probab.* **25** 58.
- [9] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2019). Heat kernel estimates and intrinsic metric for random walks with general speed measure under degenerate conductances. *Electron. Commun. Probab.* **24** 5. MR3916337 <https://doi.org/10.1214/18-ECP207>
- [10] ANTAL, P. and PISZTORA, A. (1996). On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** 1036–1048. MR1404543 <https://doi.org/10.1214/aop/1039639377>
- [11] ARMSTRONG, S., BORDAS, A. and MOURRAT, J.-C. (2018). Quantitative stochastic homogenization and regularity theory of parabolic equations. *Anal. PDE* **11** 1945–2014. MR3812862 <https://doi.org/10.2140/apde.2018.11.1945>
- [12] ARMSTRONG, S. and DARIO, P. (2018). Elliptic regularity and quantitative homogenization on percolation clusters. *Comm. Pure Appl. Math.* **71** 1717–1849. MR3847767 <https://doi.org/10.1002/cpa.21726>
- [13] ARMSTRONG, S., KUUSI, T. and MOURRAT, J.-C. (2016). Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.* **347** 315–361. MR3545509 <https://doi.org/10.1007/s00220-016-2663-2>
- [14] ARMSTRONG, S., KUUSI, T. and MOURRAT, J.-C. (2017). The additive structure of elliptic homogenization. *Invent. Math.* **208** 999–1154. MR3648977 <https://doi.org/10.1007/s00222-016-0702-4>
- [15] ARMSTRONG, S., KUUSI, T. and MOURRAT, J.-C. (2019). *Quantitative Stochastic Homogenization and Large-Scale Regularity. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **352**. Springer, Cham. MR3932093 <https://doi.org/10.1007/978-3-030-15545-2>
- [16] ARMSTRONG, S. and LIN, J. (2017). Optimal quantitative estimates in stochastic homogenization for elliptic equations in nondivergence form. *Arch. Ration. Mech. Anal.* **225** 937–991. MR3665674 <https://doi.org/10.1007/s00205-017-1118-z>
- [17] ARMSTRONG, S. N. and SMART, C. K. (2016). Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)* **49** 423–481. MR3481355 <https://doi.org/10.24033/asens.2287>
- [18] ARONSON, D. G. (1967). Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.* **73** 890–896. MR0217444 <https://doi.org/10.1090/S0002-9904-1967-11830-5>
- [19] AVELLANEDA, M. and LIN, F.-H. (1987). Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.* **40** 803–847. MR0910954 <https://doi.org/10.1002/cpa.3160400607>
- [20] AVELLANEDA, M. and LIN, F.-H. (1991).  $L^p$  bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.* **44** 897–910. MR1127038 <https://doi.org/10.1002/cpa.3160440805>
- [21] BARLOW, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32** 3024–3084. MR2094438 <https://doi.org/10.1214/009117904000000748>
- [22] BARLOW, M. T. and DEUSCHEL, J.-D. (2010). Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.* **38** 234–276. MR2599199 <https://doi.org/10.1214/09-AOP481>
- [23] BARLOW, M. T. and HAMBLY, B. M. (2009). Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab.* **14** 1–27. MR2471657 <https://doi.org/10.1214/ejp.v14-587>
- [24] BELLA, P., FEHRMAN, B. and OTTO, F. (2018). A Liouville theorem for elliptic systems with degenerate ergodic coefficients. *Ann. Appl. Probab.* **28** 1379–1422. MR3809467 <https://doi.org/10.1214/17-AAP1332>
- [25] BELLA, P. and SCHÄFFNER, M. (2020). Quenched invariance principle for random walks among random degenerate conductances. *Ann. Probab.* **48** 296–316. MR4079437 <https://doi.org/10.1214/19-AOP1361>
- [26] BENJAMINI, I., DUMINIL-COPIN, H., KOZMA, G. and YADIN, A. (2015). Disorder, entropy and harmonic functions. *Ann. Probab.* **43** 2332–2373. MR3395463 <https://doi.org/10.1214/14-AOP934>

- [27] BENJAMINI, I. and MOSSEL, E. (2003). On the mixing time of a simple random walk on the super critical percolation cluster. *Probab. Theory Related Fields* **125** 408–420. MR1967022 <https://doi.org/10.1007/s00440-002-0246-y>
- [28] BERGER, N. and BISKUP, M. (2007). Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields* **137** 83–120. MR2278453 <https://doi.org/10.1007/s00440-006-0498-z>
- [29] BERGER, N., BISKUP, M., HOFFMAN, C. E. and KOZMA, G. (2008). Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 374–392. MR2446329 <https://doi.org/10.1214/07-AIHP126>
- [30] BISKUP, M. (2011). Recent progress on the random conductance model. *Probab. Surv.* **8** 294–373. MR2861133 <https://doi.org/10.1214/11-PS190>
- [31] BISKUP, M. and BOUKHADRA, O. (2012). Subdiffusive heat-kernel decay in four-dimensional i.i.d. random conductance models. *J. Lond. Math. Soc.* (2) **86** 455–481. MR2980920 <https://doi.org/10.1112/jlms/jds012>
- [32] BISKUP, M. and PRESCOTT, T. M. (2007). Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.* **12** 1323–1348. MR2354160 <https://doi.org/10.1214/EJP.v12-456>
- [33] BOUKHADRA, O. (2010). Heat-kernel estimates for random walk among random conductances with heavy tail. *Stochastic Process. Appl.* **120** 182–194. MR2576886 <https://doi.org/10.1016/j.spa.2009.11.001>
- [34] BOUKHADRA, O. (2010). Standard spectral dimension for the polynomial lower tail random conductances model. *Electron. J. Probab.* **15** 2069–2086. MR2745726 <https://doi.org/10.1214/EJP.v15-839>
- [35] BURTON, R. M. and KEANE, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505. MR0990777
- [36] CHANG, Y. (2017). Supercritical loop percolation on  $\mathbb{Z}^d$  for  $d \geq 3$ . *Stochastic Process. Appl.* **127** 3159–3186. MR3692311 <https://doi.org/10.1016/j.spa.2017.02.003>
- [37] CHIARINI, A. and DEUSCHEL, J.-D. (2016). Invariance principle for symmetric diffusions in a degenerate and unbounded stationary and ergodic random medium. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1535–1563. MR3573286 <https://doi.org/10.1214/15-AIHP688>
- [38] DAL MASO, G. and MODICA, L. (1986). Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl.* (4) **144** 347–389. MR0870884 <https://doi.org/10.1007/BF01760826>
- [39] DAL MASO, G. and MODICA, L. (1986). Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.* **368** 28–42. MR0850613
- [40] DARIO, P. (2018). Optimal corrector estimates on percolation clusters. arXiv preprint. Available at [arXiv:1805.00902](https://arxiv.org/abs/1805.00902).
- [41] DAVIES, E. B. (1993). Large deviations for heat kernels on graphs. *J. Lond. Math. Soc.* (2) **47** 65–72. MR1200978 <https://doi.org/10.1112/jlms/s2-47.1.65>
- [42] DE MASI, A., FERRARI, P. A., GOLDSTEIN, S. and WICK, W. D. (1989). An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55** 787–855. MR1003538 <https://doi.org/10.1007/BF01041608>
- [43] DELMOTTE, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoam.* **15** 181–232. MR1681641 <https://doi.org/10.4171/RMI/254>
- [44] DEUSCHEL, J.-D., NGUYEN, T. A. and SLOWIK, M. (2018). Quenched invariance principles for the random conductance model on a random graph with degenerate ergodic weights. *Probab. Theory Related Fields* **170** 363–386. MR3748327 <https://doi.org/10.1007/s00440-017-0759-z>
- [45] DURRETT, R. (1984). *Brownian Motion and Martingales in Analysis*. Wadsworth Mathematics Series. Wadsworth International Group, Belmont, CA. MR0750829
- [46] EGLOFFE, A.-C., GLORIA, A., MOURRAT, J.-C. and NGUYEN, T. N. (2015). Random walk in random environment, corrector equation and homogenized coefficients: From theory to numerics, back and forth. *IMA J. Numer. Anal.* **35** 499–545. MR3335214 <https://doi.org/10.1093/imanum/dru010>
- [47] EVANS, L. C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Amer. Math. Soc., Providence, RI. MR2597943 <https://doi.org/10.1090/gsm/019>
- [48] FLEGEL, F., HEIDA, M. and SLOWIK, M. (2019). Homogenization theory for the random conductance model with degenerate ergodic weights and unbounded-range jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 1226–1257. MR4010934 <https://doi.org/10.1214/18-aihp917>
- [49] GIAQUINTA, M. and GIUSTI, E. (1982). On the regularity of the minima of variational integrals. *Acta Math.* **148** 31–46. MR0666107 <https://doi.org/10.1007/BF02392725>
- [50] GIAQUINTA, M. and MODICA, G. (1979). Regularity results for some classes of higher order nonlinear elliptic systems. *J. Reine Angew. Math.* **311/312** 145–169. MR0549962
- [51] GIUNTI, A., HÖFER, R. and VELÁZQUEZ, J. J. L. (2018). Homogenization for the Poisson equation in randomly perforated domains under minimal assumptions on the size of the holes. *Comm. Partial Differential Equations* **43** 1377–1412. MR3915491 <https://doi.org/10.1080/03605302.2018.1531425>

- [52] GIUSTI, E. (2003). *Direct Methods in the Calculus of Variations*. World Scientific Co., River Edge, NJ. MR1962933 <https://doi.org/10.1142/9789812795557>
- [53] GLORIA, A., NEUKAMM, S. and OTTO, F. (2015). Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199** 455–515. MR3302119 <https://doi.org/10.1007/s00222-014-0518-z>
- [54] GLORIA, A., NEUKAMM, S. and OTTO, F. (2020). A regularity theory for random elliptic operators. *Milan J. Math.* **88** 99–170. MR4103433 <https://doi.org/10.1007/s00032-020-00309-4>
- [55] GLORIA, A. and OTTO, F. (2011). An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.* **39** 779–856. MR2789576 <https://doi.org/10.1214/10-AOP571>
- [56] GLORIA, A. and OTTO, F. (2012). An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.* **22** 1–28. MR2932541 <https://doi.org/10.1214/10-AAP745>
- [57] GLORIA, A. and OTTO, F. (2015). The corrector in stochastic homogenization: Optimal rates, stochastic integrability, and fluctuations. arXiv preprint. Available at arXiv:1510.08290.
- [58] GLORIA, A. and OTTO, F. (2017). Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc. (JEMS)* **19** 3489–3548. MR3713047 <https://doi.org/10.4171/JEMS/745>
- [59] GRAFAKOS, L. (2004). *Classical and Modern Fourier Analysis*. Pearson Education, Inc., Upper Saddle River, NJ. MR2449250
- [60] GRIMMETT, G. (1999). *Percolation*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **321**. Springer, Berlin. MR1707339 <https://doi.org/10.1007/978-3-662-03981-6>
- [61] GU, C. (2019). An efficient algorithm for solving elliptic problems on percolation clusters. arXiv preprint. Available at arXiv:1907.13571.
- [62] KESTEN, H. and ZHANG, Y. (1990). The probability of a large finite cluster in supercritical Bernoulli percolation. *Ann. Probab.* **18** 537–555. MR1055419
- [63] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131. MR0375412 <https://doi.org/10.1007/BF00533093>
- [64] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrsch. Verw. Gebiete* **34** 33–58. MR0402883 <https://doi.org/10.1007/BF00532688>
- [65] KOZLOV, S. M. (1978). Averaging of differential operators with almost periodic rapidly oscillating coefficients. *Mat. Sb. (N.S.)* **107** 199–217, 317. MR0512007
- [66] LAMACZ, A., NEUKAMM, S. and OTTO, F. (2015). Moment bounds for the corrector in stochastic homogenization of a percolation model. *Electron. J. Probab.* **20** 106. MR3418538 <https://doi.org/10.1214/EJP.v20-3618>
- [67] MATHIEU, P. (2008). Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.* **130** 1025–1046. MR2384074 <https://doi.org/10.1007/s10955-007-9465-z>
- [68] MATHIEU, P. and PIATNITSKI, A. (2007). Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463** 2287–2307. MR2345229 <https://doi.org/10.1098/rspa.2007.1876>
- [69] MATHIEU, P. and REMY, E. (2004). Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.* **32** 100–128. MR2040777 <https://doi.org/10.1214/aop/1078415830>
- [70] MOURRAT, J.-C. (2011). Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 294–327. MR2779406 <https://doi.org/10.1214/10-AIHP375>
- [71] NADDAF, A. and SPENCER, T. (1998). Estimates on the variance of some homogenization problems. Unpublished manuscript.
- [72] NEUKAMM, S., SCHÄFFNER, M. and SCHLÖMERKEMPER, A. (2017). Stochastic homogenization of non-convex discrete energies with degenerate growth. *SIAM J. Math. Anal.* **49** 1761–1809. MR3650427 <https://doi.org/10.1137/16M1097705>
- [73] OSADA, H. (1983). Homogenization of diffusion processes with random stationary coefficients. In *Probability Theory and Mathematical Statistics (Tbilisi, 1982)*. *Lecture Notes in Math.* **1021** 507–517. Springer, Berlin. MR0736016 <https://doi.org/10.1007/BFb0072946>
- [74] PAPANICOLAOU, G. C. and VARADHAN, S. R. S. (1981). Boundary value problems with rapidly oscillating random coefficients. In *Random Fields, Vol. I, II (Esztergom, 1979)*. *Colloquia Mathematica Societatis János Bolyai* **27** 835–873. North-Holland, Amsterdam. MR0712714
- [75] PENROSE, M. D. and PISZTORA, A. (1996). Large deviations for discrete and continuous percolation. *Adv. in Appl. Probab.* **28** 29–52. MR1372330 <https://doi.org/10.2307/1427912>
- [76] PISZTORA, A. (1996). Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields* **104** 427–466. MR1384040 <https://doi.org/10.1007/BF01198161>

- [77] PROCACCIA, E. B., ROSENTHAL, R. and SAPOZHNIKOV, A. (2016). Quenched invariance principle for simple random walk on clusters in correlated percolation models. *Probab. Theory Related Fields* **166** 619–657. [MR3568036](#) <https://doi.org/10.1007/s00440-015-0668-y>
- [78] SAPOZHNIKOV, A. (2017). Random walks on infinite percolation clusters in models with long-range correlations. *Ann. Probab.* **45** 1842–1898. [MR3650417](#) <https://doi.org/10.1214/16-AOP1103>
- [79] SIDORAVICIUS, V. and SZNITMAN, A.-S. (2004). Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields* **129** 219–244. [MR2063376](#) <https://doi.org/10.1007/s00440-004-0336-0>
- [80] STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series*. **30**. Princeton Univ. Press, Princeton, NJ. [MR0290095](#)
- [81] YURINSKIĪ, V. V. (1986). Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.* **27** 167–180, 215. [MR0867870](#)