OPTIMAL CORRECTOR ESTIMATES ON PERCOLATION CLUSTER

BY PAUL DARIO

School of Mathematical Sciences, Tel Aviv University, pauldario@mail.tau.ac.il

We prove optimal quantitative estimates on the first-order correctors on supercritical percolation clusters: we show that they are bounded in dimension larger than 3 and have logarithmic growth in dimension 2 in the sense of stretched exponential moments. The main ingredients are a renormalization scheme of the supercritical percolation cluster, following the works of Pisztora (*Probab. Theory Related Fields* **104** (1996) 427–466); large-scale regularity estimates developed by Armstrong and the author in (*Comm. Pure Appl. Math.* **71** (2018) 1717–1849); and a nonlinear concentration inequality of the Efron–Stein type which is used to transfer quantitative information from the environment to the correctors.

CONTENTS

1. Introduction.

1.1. *Motivation and informal summary of results.* We consider the random conductance model on the supercritical percolation cluster defined as follows. We let $\mathbb{Z}^d$ be the standard hypercubic lattice and $\mathcal{B}_d$ be the set of *bonds* of $\mathbb{Z}^d$. We fix a parameter $\lambda \in (0, 1)$ and we are given a function

$$(1) \qquad\qquad \mathbf{a} : \mathcal{B}_d \to \{0\} \cup [\lambda, 1],$$

the value $\mathbf{a}(e)$ is called the conductance of the bond e and we assume that the collection $(\mathbf{a}(e))_{e \in \mathcal{B}_d}$ is an i.i.d family of random variables. We assume that the probability $\mathfrak{p} := \mathbb{P}(\mathbf{a}(e) \neq 0) > \mathfrak{p}_{\mathfrak{c}}(d)$, where $\mathfrak{p}_{\mathfrak{c}}(d)$ is the bond percolation threshold for the lattice $\mathbb{Z}^d$. It follows that, almost surely, there exists a unique maximal connected component of bonds with nonzero conductance which we denote by $\mathscr{C}_\infty = \mathscr{C}_\infty(\mathbf{a})$. One then wishes to study the continuous time *random walk* X_t in the random environment $\mathbf{a}$ defined as follows. We select an environment $\mathbf{a}$ such that 0 belongs to the infinite cluster $\mathscr{C}_\infty$ and start a random walker at the origin, $X(0) = 0$. Each edge e is equipped with a random clock and rings after exponential waiting time with expectation $\mathbf{a}(e)^{-1}$. When $X(t) = x$, the random walker waits until a clock of an edge adjacent to x rings and then moves instantly across that edge. Note that the

random walker is confined to the infinite cluster $\mathscr{C}_\infty$. This random walk is a Markov process and a common strategy to study it is to consider its generator, which is given by the random discrete elliptic PDE

$$-\nabla \cdot \mathbf{a}\nabla u,$$

where the operator $-\nabla \cdot \mathbf{a}\nabla u$ is defined on functions $u : \mathscr{C}_\infty \to \mathbb{R}$ by, for each point $x \in \mathscr{C}_\infty$,

$$\nabla \cdot \mathbf{a}\nabla u(x) = \sum_{y \sim x} \mathbf{a}(\{x, y\})(u(y) - u(x)).$$

In this article, we wish to study this random elliptic PDE by studying the (random) set of *harmonic functions* for this operator. In [19], it was proved, in the case when $\mathbf{a}$ takes only the two values 0 and 1, that every harmonic function h with prescribed linear growth is close to a linear function: the random vector space of harmonic functions with growth at most linear is finite dimensional; its dimension is equal to $(d + 1)$ almost surely. Moreover, for each harmonic function in this space, there exists a unique vector $p \in \mathbb{R}^d$ such that the difference $\chi_p(x) := h(x) - p \cdot x$ grows sublinearly as x tends to infinity. This result was quantified and extended to the generality presented in this introduction by Armstrong and the author in [7], where it is shown that the corrector is $o(|x|^{1-\delta})$ for some small but strictly positive exponent δ.

The map χ_p is called the corrector and is the central object of this article: our goal is to prove optimal bounds in terms of spatial scaling (and suboptimal with respect to stochastic integrability) on the first-order correctors. We show, in the sense of stretched exponential moments, that the correctors are bounded in dimensions $d \geq 3$, and have increments which grow like the square root of the logarithm of the distance in dimension 2. This result is summarized in the following theorem.

THEOREM 1.1 (Optimal L^∞ estimates for first-order correctors).   *There exist an exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and a constant $C := C(d, \mathfrak{p}, \lambda) < \infty$ such that for each $x, y \in \mathbb{Z}^d$ and each $p \in \mathbb{R}^d$,*

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

where, for a random variable X, we write $X \leq \mathcal{O}_s(K)$, to mean

$$\mathbb{E}\left[\exp\left(\left(\frac{X}{K}\right)^s\right)\right] \leq 2.$$

Obtaining information on the corrector is important and has proved to be useful. For instance, qualitative sublinarity of the corrector can be used to prove invariance principles for the random walker X_t following the general principle described below: if one denotes by $\chi := (\chi_1, \ldots, \chi_d)$ the vector-valued corrector, where χ_i is the corrector such that $e_i \cdot x + \chi_i(x)$ is harmonic, then the process

$$X_t + \chi(X_t) \text{ is a martingale, almost surely with respect to the environment.}$$

The strategy is to apply a standard martingale convergence theorem and then to derive a quenched invariance principle for the rescaled process $\varepsilon X_{t/\varepsilon^2} + \varepsilon\chi(X_{t/\varepsilon^2})$. Using the sublinearity of the corrector allows to prove an invariance principle for the diffusion process X itself. This approach was carried out on the infinite supercritical cluster (in the case when $\mathbf{a}$ takes only the values 0 and 1) first by Sidoravicius and Sznitman in [51] in dimension larger than 4, and a few years later by Mathieu, Piatnitski [42] and Berger, Biskup [20] in

all dimensions $d \geq 2$. Prior to these results, the generator of the random walk was studied by Barlow in [16] and by Mathieu, Remy in [43], who proved heat-kernel bounds for the transition probability.

In the more general setting of i.i.d random conductances, when **a** can a priori take values in $[0, \infty)$, a quenched functional central limit theorem was established by Andres, Barlow, Deuschel and Hambly in [1], provided that there exists an infinite cluster of nonzero conductances, based on the previous works of Mathieu [41], Biskup and Prescott [21], Barlow and Deuschel [17]. More general models of random walks on percolation clusters with long range correlation, including random interlacements and level sets of the Gaussian free field, are studied by Procaccia, Rosenthal and Sapozhnikov in [50].

Tight bounds on the corrector are useful to derive invariance principles but they are also the crucial ingredient for the derivation of optimal error and two-scale expansion estimates for the homogenization of general boundary value problems. They can be used to obtain *a Berry–Essen theorem*, in the spirit of Mourrat [45] in the uniformly elliptic setting (see also Andres and Neukamm [5] for an extension of these results to degenerate and correlated environments) and are also important to obtain precise information on the Green's function for the Laplacian on the infinite cluster as well as on the transition probability for the random walk, as is explained in [10], Chapters 8 and 9. They can also inform the performance of numerical algorithms for the computation of the homogenized diffusivity [46] and of solutions to the heterogeneous equation [8].

The tools developed in this article come from the theory of stochastic homogenization which studies the solutions of the elliptic equation

$$-\nabla \cdot \mathbf{a}\nabla u = 0 \quad \text{in } \mathbb{R}^d,$$

where the environment **a** is a random map from $\mathbb{R}^d$ to the set of symmetric matrices, satisfying some assumptions of ellipticity, stationarity and ergodicity. There have been recent developments in the quantitative homogenization of uniformly elliptic divergence-form equations, which started with the work of Gloria and Otto [36]. In this article, they were able to obtain moments bounds on the corrector with an optimal spatial scaling, by using a spectral gap inequality, which was first introduced into stochastic homogenization by Naddaf and Spencer in [47], to quantify the ergodicity of the coefficient field. This program was then continued by Gloria and Otto in [37–39] and by Neukamm, Gloria and Otto in [33–35] and has implications to random walks as explained in [30].

Another approach was later initiated by Armstrong and Smart in [13], who extended the techniques of Avellaneda and Lin [14, 15] and the ones of Dal Maso and Modica [24, 25], and were able to obtain a large scale $C^{0,1}$-regularity theory under an assumption of finite range dependence on the environment. These results were then generalized by Armstrong, Kuusi and Mourrat to general mixing conditions and to other types of equations [12] and improved to obtain optimal rates of convergence [9, 10].

The theory is now well-understood in the uniformly elliptic setting. Going beyond this setting has been the subject of much research recently in different directions. In [40], Lamacz, Neukamm and Otto were able to extend these results to a model of Bernoulli bond percolation, where the standard model is modified such that all the bonds in a fixed unit direction are always open. Another way of removing the ellipticity assumption can be the following: we define some (scalar) random variables $0 < \lambda \leq \mu < \infty$ according to the formulas

$$\lambda := \inf_{\xi \in \mathbb{R}^d \setminus \{0\}} \frac{\xi \cdot \mathbf{a}\xi}{|\xi|^2} \quad \text{and} \quad \mu := \sup_{\xi \in \mathbb{R}^d \setminus \{0\}} \frac{\xi \cdot \mathbf{a}\xi}{|\xi|^2},$$

and add an assumption on the integrability of λ and μ: there exist $p, q \in [1, \infty]$ such that

$$(3) \qquad\qquad\qquad \mathbb{E}[\lambda^{-p}] + \mathbb{E}[\mu^q] < \infty.$$

This setting was first considered by Andres, Deuschel, Slowik in [3] (see also [4]), and then by Chiarini and Deuschel in [23]. They are able to obtain a quenched invariance principle for the diffusion process under the assumption $1/p + 1/q < 2/d$, which allowed them to perform a Moser iteration. In [18], Bella, Fehrman and Otto, still working under the assumption $1/p + 1/q < 2/d$, were able to obtain a first-order Liouville theorem and a large scale $C^{1,\alpha}$-estimate for **a**-harmonic functions. An extension of these results to the case of time-dependent coefficients has been carried out by [2]. The condition (3) requires the value of the conductances to be nonzero almost surely, an extension of this model in a case when the conductance is allowed to be zero and to be small (under some moments condition) was investigated by Deuschel, Nguyen and Slowik in [29].

The setting considered in this article is different from the models satisfying condition (3): we are working with the i.i.d. random conductance model, and we assume the value of the conductances to be either 0 or larger than some deterministic constant $\lambda > 0$ (see (1)), with the property that $\mathbb{P}(\mathbf{a}(e) \neq 0) > \mathfrak{p}_c(d)$. Despite this difference, the main challenge is essentially the same: adapting the various tools and proofs, available in the uniformly elliptic setting, to the degenerate elliptic environment. To this end, we follow the strategy initiated in the previous paper [7] and appeal to a renormalization structure for the supercritical percolation cluster. The construction is recalled in Section 2, where $\mathbb{Z}^d$ is partitioned into triadic cubes of different random sizes, well-connected in the sense of Penrose an Pisztora [48]. This partition allows to distinguish regions of $\mathbb{Z}^d$ where the infinite cluster is well-behaved, its geometry is similar to the one of the lattice $\mathbb{Z}^d$, from regions where the infinite cluster is badly-behaved. In the first case, it is rather straightforward to adapt the theory developed in the uniformly elliptic setting; problems arise where the infinite cluster is badly-behaved. In this situation the theory cannot be adapted. Fortunately, there are few regions were the cluster is badly-behaved, and the theory of stochastic homogenization in the uniformly elliptic setting is robust enough to be adapted to the supercritical cluster.

Our strategy to prove the optimal scaling estimates for the corrector relies on a concentration inequality (cf. Proposition 8), which gives a convenient way to transfer quantitative information from the coefficient field to the correctors. This idea originates in an unpublished paper from Naddaf and Spencer [47], and was then developed by Gloria and Otto [36, 37] and Gloria, Neukamm and Otto [34] (see also Mourrat [44]) to study stochastic homogenization. More precisely, thanks to this inequality we are able to obtain quantitative estimates on the spatial average of the gradient of the corrector.

We then use the the multiscale Poincaré inequality stated in Proposition 9 to deduce the estimates on the oscillation of the correctors stated in Theorem 1.1 from the bounds on the spatial average of its gradient.

We conclude this introduction by noting that in Theorem 1.1, the spatial scaling is optimal while the stochastic integrability is suboptimal: we only obtain a small exponent $s > 0$ of stochastic integrability. This is due to the degenerate structure of the percolation problem and while our method can provide an explicit value for the exponent s, we do not expect it can be used to derive the optimal exponent. We nevertheless provide a conjecture.

CONJECTURE 1. *In dimension $d = 2$, fix $s < 2$, then there exists a constant $C :=$ $C(s, \mathfrak{p}, \lambda) < \infty$ such that for each $x, y \in \mathbb{Z}^d$ and each $p \in \mathbb{R}^d$,*

$$(4) \qquad |\chi_p(x) - \chi_p(y)|\mathbb{1}_{\{x,y \in \mathscr{C}_\infty\}} \leq \mathcal{O}_{\frac{s(d-1)}{d}}(C|p|\log^{\frac{1}{2}}|x-y|).$$

In dimension 3, there exists a constant $C := C(d, s, \mathfrak{p}, \lambda) < \infty$ such that for each $x, y \in \mathbb{Z}^d$ and each $p \in \mathbb{R}^d$,

$$|\chi_p(x) - \chi_p(y)|\mathbb{1}_{\{x,y \in \mathscr{C}_\infty\}} \leq \mathcal{O}_{\frac{s(d-1)}{d}}(C|p|).$$

The reason behind this conjecture is the following: in the uniformly elliptic setting, it is known that the optimal stochastic integrability is the one provided in the statement of the conjecture without the term $(d-1)/d$, see [10], Theorem 4.1. The additional term $(d-1)/d$ is a surface order large deviation effect which can be heuristically explained by the following argument: in the uniformly elliptic setting and in a given ball, to design a bad environment for which one does not have a good control on the growth of the corrector, it is necessary to have a number of ill-behaved edges of order of the volume of the ball. In the percolation setting, the situation is different and one only needs a number of ill-behaved edges of the order of the surface of the ball to design a bad environment: to illustrate this fact, one can note that, if we let R be the radius of the ball, then it is possible to disconnect the ball into two half-balls with only cR^{d-1} closed edges. This phenomenon should result in a deterioration of the stochastic integrability by a factor $(d-1)/d$.

1.2. *Notation and assumptions.*

1.2.1. *General notation for the probabilistic model.* We denote by $\mathbb{Z}^d$ the standard d-dimensional hypercubic lattice. The set of bonds of $\mathbb{Z}^d$, that is the set of unoriented pairs of nearest neighbors, is denoted by $\mathcal{B}_d := \{\{x, y\} : x, y \in \mathbb{Z}^d, |x - y|_1 = 1\}$. More specifically, given a subset $U \subseteq \mathbb{Z}^d$, we denote by $\mathcal{B}_d(U)$ the set of the bonds of U, that is, $\mathcal{B}_d(U) := \{\{x, y\} : x, y \in U, |x - y|_1 = 1\}$. The canonical basis of $\mathbb{R}^d$ is denoted by $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$. For $x, y \in \mathbb{Z}^d$, we write $x \sim y$ if $\{x, y\} \in \mathcal{B}_d$. For some fixed ellipticity parameter $\lambda \in (0, 1]$, we define the probability space $\Omega := (\{0\} \cup [\lambda, 1])^{\mathcal{B}_d}$ and we equip this probability space with the Borel σ-algebra $\mathcal{F} := \mathcal{B}(\{0\} \cup [\lambda, 1])^{\otimes \mathcal{B}_d}$. Given an edge $e \in \mathcal{B}_d$, we denote by $\mathbf{a}(e)$ the projection

$$\mathbf{a}(e) : \begin{cases} \Omega \to \{0\} \cup [\lambda, 1], \\ (\omega_{e'})_{e' \in \mathcal{B}_d} \mapsto \omega_e. \end{cases}$$

We denote by $\mathbf{a}$ the collection $(\mathbf{a}(e))_{e \in \mathcal{B}_d}$ and we refer to this mapping as the *environment*. For every set $U \subseteq \mathbb{Z}^d$, we denote by $\mathcal{F}(U) \subseteq \mathcal{F}$ the σ-algebra generated by the mappings $(\mathbf{a}(e))_{e \in \mathcal{B}_d(U)}$.

We fix a probability measure $\mathbb{P}_0$ supported in $\{0\} \cup [\lambda, 1]$ satisfying the property

(5) $$\mathfrak{p} := \mathbb{P}_0([\lambda, 1]) > \mathfrak{p}_{\mathfrak{c}}(d),$$

where $\mathfrak{p}_{\mathfrak{c}}(d)$ is the bond percolation threshold for the lattice $\mathbb{Z}^d$. We then equip the measurable space $(\Omega, \mathcal{F})$ with the i.i.d. probability measure $\mathbb{P} = \mathbb{P}_0^{\otimes \mathcal{B}_d}$, so that the sequence $(\mathbf{a}(e))_{e \in \mathcal{B}_d}$ is an i.i.d. collection of random variables of law $\mathbb{P}_0$. The expectation with respect to the probability measure $\mathbb{P}$ is denoted by $\mathbb{E}$.

Given an environment $\mathbf{a}$, we say that a bond $e \in \mathcal{B}_d$ is *open* if $\mathbf{a}(e) > 0$ and *closed* if $\mathbf{a}(e) = 0$. Given two vertices $x, y \in \mathbb{Z}^d$, we say that there is a *path connecting x and y* if there exists a sequence of open edges of the form $\{x, z_1\}, \ldots, \{z_n, z_{n+1}\}, \ldots, \{z_N, y\}$. The two vertices x and y are then said to be *connected*, which we denote by $x \leftrightarrow_{\mathbf{a}} y$, if there exists a path connecting x and y. A *cluster* is a connected subset $\mathscr{C} \subseteq \mathbb{Z}^d$. Thanks to the assumption (5), we know that, $\mathbb{P}$–almost surely, there exists a unique maximal infinite cluster (see [22]). This cluster is denoted by $\mathscr{C}_\infty := \mathscr{C}_\infty(\mathbf{a})$.

We also denote by $E_d := \{(x, y) : x, y \in \mathbb{Z}^d, x \sim y\}$ the set of oriented edges. More generally, we define, for a subset $U \subseteq \mathbb{Z}^d$, $E_d(U) := \{(x, y) : x, y \in U, x \sim y\}$.

For $x \in \mathbb{Z}^d$, we define the translation τ_x on Ω to be the map

$$\tau_x : \begin{cases} \Omega \to \Omega, \\ (\omega_e)_{e \in \mathcal{B}_d} \mapsto (\omega_{e+x})_{e \in \mathcal{B}_d}. \end{cases}$$

Note that the measure $\mathbb{P}$ is stationary with respect to the $\mathbb{Z}^d$-translations: for each $x \in \mathbb{Z}^d$,

$$(6) \qquad (\tau_x)_* \mathbb{P} = \mathbb{P},$$

where $(\tau_x)_* \mathbb{P}$ is the pushforward measure defined by the formula, for each set $A \in \mathcal{F}$, $(\tau_x)_* \mathbb{P}(A) = \mathbb{P}(\tau_x^{-1}(A))$.

1.2.2. *Notation for functions.*

For each vector $p \in \mathbb{R}^d$, we denote by l_p the affine function of slope p, that is, $l_p(x) = p \cdot x$. Given a function u defined on a discrete set $U \subseteq \mathbb{Z}^d$, we define its oscillation by the formula

$$\operatorname*{osc}_U u := \sup_U u - \inf_U u.$$

We define a *vector field* to be a function $G : E_d \to \mathbb{R}$ satisfying the antisymmetry property: for each $(x, y) \in E_d$,

$$G(x, y) = -G(y, x).$$

For a given a function $u : \mathbb{Z}^d \to \mathbb{R}$, we define its gradient ∇u to be the vector field

$$(\nabla u)(x, y) := u(x) - u(y).$$

For a random function defined on a cluster $\mathscr{C}$, $u : \mathscr{C} \to \mathbb{R}$, we define ∇u to be the vector field defined on the edges of $\mathbb{Z}^d$ by the formula

$$(7) \qquad (\nabla u)(x, y) := \begin{cases} u(x) - u(y) & \text{if } x, y \in \mathscr{C} \text{ and } \mathbf{a}(\{x, y\}) \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and $\mathbf{a}\nabla u$ to be the vector field defined by

$$(\mathbf{a}\nabla u)(x, y) := \mathbf{a}(\{x, y\})(\nabla u)(x, y).$$

The cluster $\mathscr{C}$ will frequently be the infinite cluster $\mathscr{C}_\infty$. We may also think of the gradient as a vector-valued operator, as it is commonly the case for continuous functions: we denote by, for any point $x \in \mathscr{C}$ and any function $u : \mathscr{C} \to \mathbb{R}$,

$$(8) \qquad \nabla u(x) := \begin{pmatrix} \nabla u(x + \mathfrak{e}_1, x) \\ \vdots \\ \nabla u(x + \mathfrak{e}_d, x) \end{pmatrix}.$$

For $p \in \mathbb{R}^d$, we denote by p the constant vector field, defined according to the formula

$$p(x, y) := p \cdot (x - y).$$

With these conventions, we have $\nabla l_p = p$. For a given vector field G and a point $x \in \mathbb{Z}^d$, we define

$$(9) \qquad |G|(x) := \left(\sum_{y : (x,y) \in E_d} |G(x, y)|^2 \right)^{\frac{1}{2}}.$$

For a given a subset $U \subseteq \mathbb{Z}^d$, we equip the space of vector fields with a scalar product $\langle \cdot, \cdot \rangle$, defined by

$$\langle F, G \rangle_U := \sum_{(x,y) \in E_d(U)} F(x, y) G(x, y).$$

We will also frequently make use of the following notation, given a vector field G, we define

$$\langle G \rangle_U = \sum_{(x,y) \in E_d(U)} G(x, y)(x - y).$$

The value $\langle G \rangle_U$ belongs to the space $\mathbb{R}^d$. Given an environment $\mathbf{a}$, two functions $u, v : \mathbb{Z}^d \to \mathbb{R}$, and a subset $U \subseteq \mathbb{Z}^d$, the Dirichlet form can be written with the previous notation as

$$\langle \nabla u, \mathbf{a} \nabla v \rangle_U = \sum_{(x,y) \in E_d(U)} (u(x) - u(y)) \mathbf{a}(\{x, y\})(v(x) - v(y)).$$

We define the elliptic operator $-\nabla \cdot \mathbf{a} \nabla$ by, for each function $u : \mathbb{Z}^d \to \mathbb{R}$ and each point $x \in \mathbb{Z}^d$,

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

For a given a subset $U \subseteq \mathbb{Z}^d$, we define the random set of $\mathbf{a}$-harmonic functions in U by

$$\mathcal{A}(U) := \{u : U \to \mathbb{R} : (-\nabla \cdot \mathbf{a} \nabla u)(x) = 0, x \in \mathrm{int}_{\mathbf{a}} U\},$$

where $\mathrm{int}_{\mathbf{a}} U$ is the interior of U with respect to the environment $\mathbf{a}$, defined according to the formula

$$\mathrm{int}_{\mathbf{a}} U := \{x \in U : \forall y \in \mathbb{Z}^d, (y \sim x \text{ and } \mathbf{a}(\{x, y\}) \neq 0) \implies y \in U\}.$$

If U is a finite set, we denote its cardinality by $|U|$.

For vectors of $\mathbb{R}^d$, we denote by $|\cdot|$ the standard infinite norm given by $|x| = \max_{i=1,\dots,d} |x_i|$. We define a pseudometric on the subsets of $\mathbb{Z}^d$ by the formula $\mathrm{dist}(U, V) = \inf_{x \in U, y \in V} |x - y|$.

We also use the notation $B_R(x)$ or $B(x, R)$ to denote the ball centered at $x \in \mathbb{Z}^d$ with radius $R > 0$ with respect to the infinite norm. The ball $B_R(0)$ is simply denoted by B_R.

1.2.3. *Notation for cubes.* A *cube* is a subset of $\mathbb{Z}^d$ of the form

$$\square := (z + (-N, N)^d) \cap \mathbb{Z}^d, \quad N \in \mathbb{N}, z \in \mathbb{Z}^d.$$

We define the *center* and the *size* of the cube $\square$ to be the point $z \in \mathbb{Z}^d$ and the integer $2N - 1$. We denote its size by $\mathrm{size}(\square)$. In particular, with this convention, we have $|\square| = (\mathrm{size}(\square))^d$. For a nonnegative real number $r > 0$ and a cube $\square$, of center $z \in \mathbb{Z}^d$ and size $(2N - 1) \in \mathbb{N}$, we denote by $r\square$ the cube

$$r\square := (z + (-rN, rN)^d) \cap \mathbb{Z}^d.$$

This notation is nonstandard; the multiplication by r only affects the size of the cube but the center of the cube remains unchanged. We introduce a specific category of cubes, namely the *triadic cubes*. A triadic cube is a cube of the form

$$(10) \qquad \square_n(z) := \left(z + \left(-\frac{1}{2}3^n, \frac{1}{2}3^n\right)^d\right) \cap \mathbb{Z}^d, \quad n \in \mathbb{N}, z \in 3^n \mathbb{Z}^d.$$

To simplify the notation, we write $\square_n = \square_n(0)$. This collection of cubes enjoys a number of convenient properties. First, any two triadic cubes (of possibly different sizes) are either disjoint or else one is included in the other. Moreover, for every $m, n \in \mathbb{N}$ with $n \leq m$, the triadic cube $\square_m$ can be uniquely partitioned into $3^{d(m-n)}$ disjoint triadic cubes of size 3^n. We denote by $\mathcal{T}$ the collection of triadic cubes and by $\mathcal{T}_n$ the collection of triadic cubes of size 3^n.

For each integer $n \in \mathbb{N}$ and each cube $\square \in \mathcal{T}_n$, we define the predecessor of $\square$, to be the unique triadic cube $\tilde{\square} \in \mathcal{T}_{n+1}$ such that $\square \subseteq \tilde{\square}$. If $\tilde{\square}$ is the predecessor of $\square$, then we say that $\square$ is a successor $\tilde{\square}$. In particular, a cube of the set $\mathcal{T}_0$ does not have any successor, while a cube of the set $\mathcal{T} \setminus \mathcal{T}_0$ has exactly 3^d successors.

1.2.4. *The $\mathcal{O}_s$ notation.*   We introduce a series of notation and properties which will be useful to measure the stochastic integrability and sizes of random variables. Given two parameters $s, \theta > 0$ and a nonnegative random variable X, we denote by

$$X \leq \mathcal{O}_s(\theta) \quad \text{if and only if} \quad \mathbb{E}\left[\exp\left(\left(\frac{X}{\theta}\right)^s\right)\right] \leq 2.$$

Note that, by Markov's inequality, the tail of a random variable X satisfying the inequality $X \leq \mathcal{O}_s(\theta)$ decreases stretched exponentially fast: for every $t > 0$,

$$\mathbb{P}[X \geq \theta t] \leq 2 \exp(-t^s).$$

For a given sequence $(Y_i)_{i \in \mathbb{N}}$ of nonnegative random variables and a sequence $(\theta_i)_{i \in \mathbb{N}}$ of nonnegative real numbers, we write

$$X \leq \sum_{i \in \mathbb{N}} Y_i \mathcal{O}_s(\theta_i),$$

to mean that there exists a sequence of nonnegative random variables $(Z_i)_{i \in \mathbb{N}}$ such that for each integer $i \in \mathbb{N}$, $Z_i \leq \mathcal{O}_s(\theta_i)$ and

$$X \leq \sum_{i \in \mathbb{N}} Y_i Z_i.$$

We now record some properties pertaining to this notation. All these properties are proved in [10], Appendix A. The notation is compatible with the addition, meaning that, for any stochastic integrability exponent $s > 0$, there exists a constant C depending only on s, which may be chosen to be 1 if $s \geq 1$, such that

(11) $X_1 \leq \mathcal{O}_s(\theta_1)$ and $X_2 \leq \mathcal{O}_s(\theta_2)$ $\implies$ $X_1 + X_2 \leq \mathcal{O}_s(C(\theta_1 + \theta_2)).$

More generally, for any $s > 0$, there exists a constant $C(s) < \infty$ such that, for every measure space $(X, \mathcal{F}, \mu)$, every jointly measurable family $\{X(x)\}_{x \in E}$ of nonnegative random variables and every measurable function $\theta : E \to \mathbb{R}_+$, we have

(12) $\forall x \in E, \quad X(x) \leq \mathcal{O}_s(\theta(x)) \quad \implies \quad \int_E X(x)\, d\mu(x) \leq \mathcal{O}_s\left(C \int_E \theta(x)\, d\mu(x)\right).$

The constant can be chosen to be

(13) $$\begin{cases} C(s) = \left(\dfrac{1}{s \ln 2}\right)^{\frac{1}{s}} & \text{if } s < 1, \\ C(s) = 1 & \text{if } s \geq 1. \end{cases}$$

From the definition, we have, for each $\lambda \in \mathbb{R}_+$,

$$X \leq \mathcal{O}_s(\theta) \quad \implies \quad \lambda X \leq \mathcal{O}_s(\lambda \theta).$$

This notation is also compatible with the multiplication in the sense that

(14) $|X_1| \leq \mathcal{O}_{s_1}(\theta_1)$ and $|X_2| \leq \mathcal{O}_{s_2}(\theta_2)$ $\implies$ $|XY| \leq \mathcal{O}_{\frac{s_1 s_2}{s_1 + s_2}}(\theta_1 \theta_2).$

It is easy to check from (14) that one can reduce the integrability exponent s, that is, for each $0 < s' < s$, there exists a constant $C := C(s') < \infty$ such that

(15) $X \leq \mathcal{O}_s(\theta_1) \quad \implies \quad X \leq \mathcal{O}_{s'}(C\theta_1).$

1.2.5. *Convention for constants and exponents.* In this article, the symbols c and C denote positive constants which may vary from line to line. These constants depend mainly on three parameters which are fixed through the proofs: the dimension of the space d, the ellipticity λ and the probability $\mathfrak{p} = \mathbb{P}[\mathbf{a}(e) \neq 0]$. Usually, we use C for large constants (whose value is expected to belong to the interval $[1, \infty)$) and c for small constants (whose value is expected to be in the interval $(0, 1]$).

For the stochastic integrability, we use the letter s and will typically have inequalities of the form $X \leq \mathcal{O}_s(C)$. This exponent s depends on the parameters d, λ and $\mathfrak{p}$. Its value can also vary from line to line and is expected to be small.

In Sections 4 and 5, another parameter will be involved in the dependence of the constants and exponents: the spatial integrability $q \in (2, \infty)$ (see Theorem 1.2 below). The dependence in this additional parameter will be displayed thanks to the following convention: we write $C := C(d, \lambda, \mathfrak{p}) < \infty$ (resp. $C := C(d, \lambda, \mathfrak{p}, q) < \infty$) to mean that the constant C depends only on the parameters $d, \lambda, \mathfrak{p}$ (resp. $d, \lambda, \mathfrak{p}, q$) and that its value is expected to be large. For small constants or exponents we use the notation $c := c(d, \lambda, \mathfrak{p}) > 0$, $s := s(d, \lambda, \mathfrak{p}) > 0$ (resp. $c := c(d, \lambda, \mathfrak{p}, q) > 0$, $s := s(d, \lambda, \mathfrak{p}, q) > 0$).

1.3. *Outline of the paper.* The rest of the paper is organized as follows. In Section 2, we recall (mostly without proof) some properties of the infinite cluster which were stated and proved in [7] (and based on [48]) to develop a quantitative homogenization theory on the infinite percolation cluster. In Sections 2.5 and 2.6, we state the concentration inequality and the multiscale Poincaré inequality, which are the two key ingredients in the proof of Theorem 1.2. In Section 3, we use the concentration inequality and the properties of the infinite cluster recorded in Section 2 to obtain an estimate on the spatial averages of the corrector. In Section 4, we use the result established in Section 3 combined with the multiscale Poincaré inequality to prove the optimal L^q-bound on the gradient of the corrector, stated in the following theorem.

THEOREM 1.2 (Optimal L^q estimates for first-order corrector). *For each $q \geq 2$, there exist an exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and a constant $C(d, \mathfrak{p}, \lambda, q) < \infty$ such that for each radius $R \geq 1$ and each $p \in \mathbb{R}^d$,*

$$(16) \quad \left(R^{-d} \sum_{x \in \mathscr{C}_\infty \cap B_R} |\chi_p(x) - (\chi_p)_{\mathscr{C}_\infty \cap B_R}|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(C|p| \log^{\frac{1}{2}} R) & \text{if } d = 2, \\ \mathcal{O}_s(C|p|) & \text{if } d \geq 3. \end{cases}$$

This theorem is strictly weaker than Theorem 1.1; in Section 5 we upgrade the previous L^q bound into the L^∞ bound stated in Theorem 1.1. In Appendix A, we give a proof of the multiscale Poincaré inequality stated in Section 2.6. In Appendix C, we give the proof of a technical lemma used in Section 3.

2. Preliminaries. In this section we record some properties about the infinite percolation cluster in the supercritical regime. Most of these properties were established in [7].

2.1. *The corrector*: *Existence and first properties.* Denote by $\mathcal{A}_1$ the (random) vector space of $\mathbf{a}$-harmonic functions with at most linear growth, that is,

$$\mathcal{A}_1 := \left\{ u : \mathscr{C}_\infty \to \mathbb{R} : \nabla \cdot (\mathbf{a} \nabla u) = 0 \text{ in } \mathscr{C}_\infty \text{ and } \lim_{R \to \infty} \frac{1}{R^2} \|u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_R)} = 0 \right\}.$$

By [19], we know that, $\mathbb{P}$-almost surely, the space $\mathcal{A}_1$ has dimension $(d + 1)$ and that every function $u \in \mathcal{A}_1$ is close to an affine function. More precisely, the space $\mathcal{A}_1$ can be

characterized as follows: there exist a collection of sublinear functions $\{\chi_p\}_{p\in\mathbb{R}^d}$ defined on the infinite cluster and valued in $\mathbb{R}$ such that

$$\mathcal{A}_1 := \{l_p + \chi_p + c : p \in \mathbb{R}^d, c \in \mathbb{R}\}.$$

The functions $\{\chi_p\}_{p\in\mathbb{R}^d}$ are called the correctors. They are defined up to a constant and are unique. To work with these quantities, one has to be careful to only consider quantities which are invariant by adding a constant, such as the oscillation, the gradient, the difference $\chi_p(x) - \chi_p(y)$, etcetera. For later use, we record that the map $p \mapsto \nabla\chi_p$ is linear.

The sublinear growth of the corrector is an important property which was proved qualitatively in [27] and quantitatively in [7]; by [7], (1.22), there exist two exponents $\delta := \delta(d, \mathfrak{p}, \lambda) > 0$, $s := s(d, \mathfrak{p}, \lambda) > 0$ and a constant $C := C(d, \mathfrak{p}, \lambda)$ such that, for each radius $R \geq 1$,

$$(17) \qquad \operatorname*{osc}_{\mathscr{C}_\infty\cap B_R} \chi_p \leq \mathcal{O}_s(C|p|R^{1-\delta}).$$

We may reformulate this property in terms of a minimal scale: by [7], (1.18), there exists a nonnegative random variable $\mathcal{X}$ satisfying $\mathcal{X} \leq \mathcal{O}_s(C)$, such that for each vector $p \in \mathbb{R}^d$ and each radius $R \geq \mathcal{X}$,

$$(18) \qquad \|\chi_p - (\chi_p)_{\mathscr{C}_\infty\cap B_R}\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_R)} \leq C|p|R^{1-\delta}.$$

Moreover, the corrector satisfies the following stationarity property: for each $x, y \in \mathbb{Z}^d$, each $p \in \mathbb{R}^d$ and each $z \in \mathbb{Z}^d$,

$$(19) \qquad \big(\chi_p(x) - \chi_p(y)\big)\mathbb{1}_{\{x,y\in\mathscr{C}_\infty\}}(\mathbf{a}) = \big(\chi_p(x+z) - \chi_p(y+z)\big)\mathbb{1}_{\{z+x,z+y\in\mathscr{C}_\infty\}}(\tau_z\mathbf{a}).$$

2.2. *Triadic partitions of good cubes.*

This section shows how to use the tools developed by Penrose and Pisztora [48] to obtain a renormalization structure of the infinite cluster of supercritical percolation.

2.2.1. *A general scheme for partition of good cubes.*

The construction of the partition is accomplished by a stopping time argument reminiscent of a Calderón–Zygmund–type decomposition. We are given a notion of "good cube" represented by an $\mathcal{F}$-measurable function which maps Ω into the set of all subsets of $\mathcal{T}$. In order words, for each environment $\mathbf{a} \in \Omega$, we are given a subcollection $\mathcal{G}(\mathbf{a}) \subseteq \mathcal{T}$ of triadic cubes. We think of $\square \in \mathcal{T}$ as being a good cube if $\square \in \mathcal{G}(\mathbf{a})$. We frequently drop the dependence in $\mathbf{a}$ and write $\mathcal{G}$ instead of $\mathcal{G}(\mathbf{a})$.

PROPOSITION 1 (Partition of good cubes, Proposition 2.1 of [7]). *Let $\mathcal{G} \subseteq \mathcal{T}$ be a random collection of triadic cubes, as above. Suppose that there exist constants $K, s > 0$ such that*

$$\sup_{z\in3^n\mathbb{Z}^d} \mathbb{P}[z + \square_n \notin \mathcal{G}] \leq K \exp(-K^{-1}3^{ns}).$$

Then, $\mathbb{P}$–almost surely, there exists a partition $\mathcal{S} \subseteq \mathcal{T}$ of $\mathbb{Z}^d$ into triadic cubes with the following properties:

(i) *All predecessors of elements of $\mathcal{S}$ are good*: *for every $\square, \square' \in \mathcal{T}$,*

$$\square' \subseteq \square \quad \text{and} \quad \square' \in \mathcal{S} \quad \Longrightarrow \quad \square \in \mathcal{G}.$$

(ii) *Neighboring elements of $\mathcal{S}$ have comparable sizes*: *for every $\square, \square' \in \mathcal{S}$ such that* $\operatorname{dist}(\square, \square') \leq 1$, *we have*

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

(iii) *Estimate for the coarseness of $\mathcal{S}$: if we denote by $\square_\mathcal{S}(x)$ the unique element of $\mathcal{S}$ containing a point $x \in \mathbb{Z}^d$, then there exists $C(s, K, d) < \infty$ such that*

$$\text{size}(\square_\mathcal{S}(x)) \leq \mathcal{O}_s(C).$$

In addition, if one has the following independence property, for every cube $\square = z + \square_n \in \mathcal{T}$,

(20)   *the event $\{\square \notin \mathcal{G}\}$ is $\mathcal{F}(z + \square_{n+1})$-measurable,*

then one has the following minimal scale property:

(iv) *Minimal scale for $\mathcal{S}$. For each $t \in [1, \infty)$, there exists $C := C(t, s, K, d) < \infty$, an $\mathbb{N}$-valued random variable $\mathcal{M}_t(\mathcal{S})$ and exponent $r := r(t, s, K, d) > 0$ such that*

$$\mathcal{M}_t(\mathcal{S}) \leq \mathcal{O}_r(C)$$

and for each integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}_t(\mathcal{S})$,

$$\frac{1}{|\square_m|} \sum_{x \in \square_m} \text{size}(\square_\mathcal{S}(x))^t \leq C \quad and \quad \sup_{x \in \square_m} \text{size}(\square_\mathcal{S}(x)) \leq 3^{\frac{dm}{d+t}}.$$

2.2.2. *The partition $\mathcal{P}$ of well-connected cubes.* We apply the construction of the previous subsection to obtain a random partition $\mathcal{P}$ of $\mathbb{Z}^d$ which simplifies the geometry of the percolation cluster. This partition plays an important role in the rest of the paper. To obtain bounds on the "good event" which allows us to construct the partition, we use the results of Pisztora [49], Penrose and Pisztora [48] and Antal and Pisztora [6]. We first recall some definitions introduced in those works.

DEFINITION 1 (Crossability and crossing cluster). We say that a cube $\square$ is *crossable* (with respect to an environment $\mathbf{a} \in \Omega$) if each of the d pairs of opposite $(d-1)$-dimensional faces of the cube $\square$ is joined by an open path in $\square$. We say that a cluster $\mathscr{C} \subseteq \square$ is a *crossing cluster for $\square$* if $\mathscr{C}$ intersects each of the $(d-1)$-dimensional faces of $\square$.

DEFINITION 2 (Good cube). We say that a triadic cube $\square \in \mathcal{T}$ is *well-connected* if there exists a crossing cluster $\mathscr{C}$ for the cube $\square$ such that:

(i) Each cube $\square'$ with $\text{size}(\square') \in [\frac{1}{10}\text{size}(\square), \frac{1}{2}\text{size}(\square)]$ and $\square' \cap \frac{3}{4}\square \neq \varnothing$ is crossable.
(ii) Every path $\gamma \subseteq \square'$ with $\text{diam}(\gamma) \geq \frac{1}{10}\text{size}(\square)$ is connected to $\mathscr{C}$ within $\square'$.

We say that $\square \in \mathcal{T}$ is a *good cube* if $\text{size}(\square) \geq 3$, $\square$ is well-connected and each of the 3^d successors of the cube $\square$ are well-connected. We say that $\square \in \mathcal{T}$ is a *bad cube* if it is not a good cube (see Figure 1).

The following estimate on the probability of the cube $\square_n$ being good is a consequence [49], Theorem 3.2, and [48], Theorem 5, as recalled in [6], (2.24).

LEMMA 2.1 ([6], (2.24)). *For each probability $\mathfrak{p} \in (\mathfrak{p}_c, 1]$, there exists a constant $C(d, \mathfrak{p}) < \infty$ such that, for every integer $n \in \mathbb{N}$,*

(21) $$\mathbb{P}[\square_n \text{ is good}] \geq 1 - C \exp(-C^{-1}3^n).$$

It follows from Definition 2 that, for every good cube $\square$, there exists a unique maximal crossing cluster for $\square$ which is contained in $\square$. We denote this cluster by $\mathscr{C}_*(\square)$. In the next lemma, we record the observation that adjacent triadic cubes which have similar sizes and are both good have connected clusters.
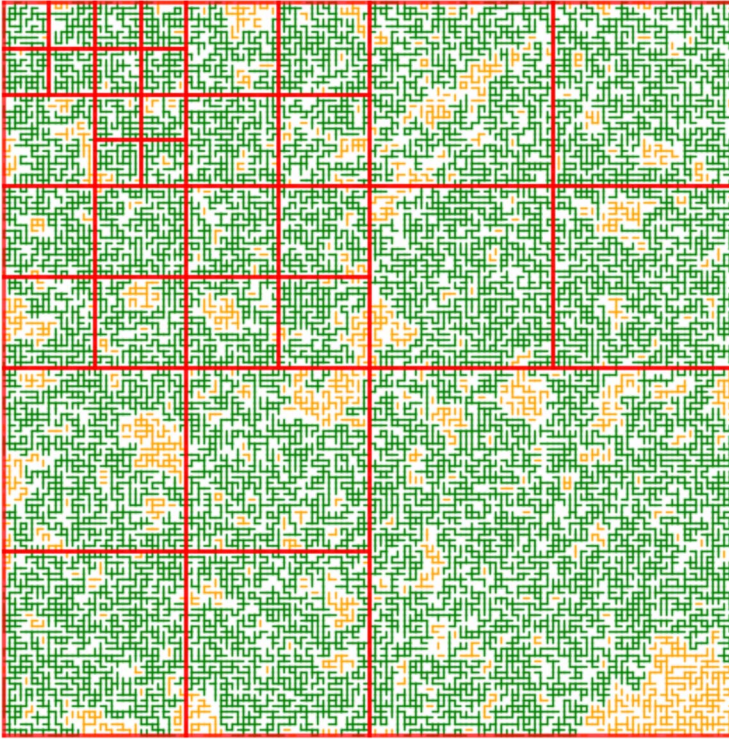
FIG. 1. *A good cube* □. *The cluster* $\mathscr{C}_*(\square)$ *is drawn in green. Simulation by C. Gu.*

LEMMA 2.2 (Lemma 2.8 of [7]). *Let* $n, n' \in \mathbb{N}$ *with* $|n - n'| \leq 1$ *and* $z, z' \in 3^n \mathbb{Z}^d$ *such that*

$$\text{dist}\big(\square_n(z), \square_{n'}(z')\big) \leq 1.$$

Suppose also that $\square_n(z)$ *and* $\square_{n'}(z')$ *are good cubes. Then there exists a cluster* $\mathscr{C}$ *such that*

$$\mathscr{C}_*\big(\square_n(z)\big) \cup \mathscr{C}_*\big(\square_{n'}(z')\big) \subseteq \mathscr{C} \subseteq \square_n(z) \cup \square_{n'}(z').$$

We next define the partition $\mathcal{P}$ of good cubes.

DEFINITION 3. We let $\mathcal{P} \subseteq \mathcal{T}$ be the partition $\mathcal{S}$ of $\mathbb{Z}^d$ obtained by applying Proposition 1 to the collection

$$\mathcal{G} := \{\square \in \mathcal{T} : \square \text{ is good}\}.$$

More generally, for each point $y \in \mathbb{Z}^d$, we let $\mathcal{P}_y \subseteq \mathcal{T}$ be the partition $\mathcal{S}$ of $\mathbb{Z}^d$ obtained by applying Proposition 1 to the collection

$$\mathcal{G} := \{y + \square : \square \in \mathcal{T} \text{ and } y + \square \text{ is good}\}.$$

From the construction of $\mathcal{P}$ and $\mathcal{P}_y$, we also have

$$\mathcal{P}_y = y + \mathcal{P}(\tau_{-y}\mathbf{a}) = \{y + \square : \square \in \mathcal{P}(\tau_{-y}\mathbf{a})\}.$$

We refer to Figure 2 for an illustration of the random partition $\mathcal{P}$.

The (random) partition $\mathcal{P}$ plays an important role throughout the rest of the paper. We denote by $\mathcal{P}_*$ the collection of triadic cubes which contains some elements of $\mathcal{P}$, that is

$$\mathcal{P}_* := \{\square : \square \text{ is a triadic cube and } \square \supseteq \square' \text{ for some } \square' \in \mathcal{P}\}.$$
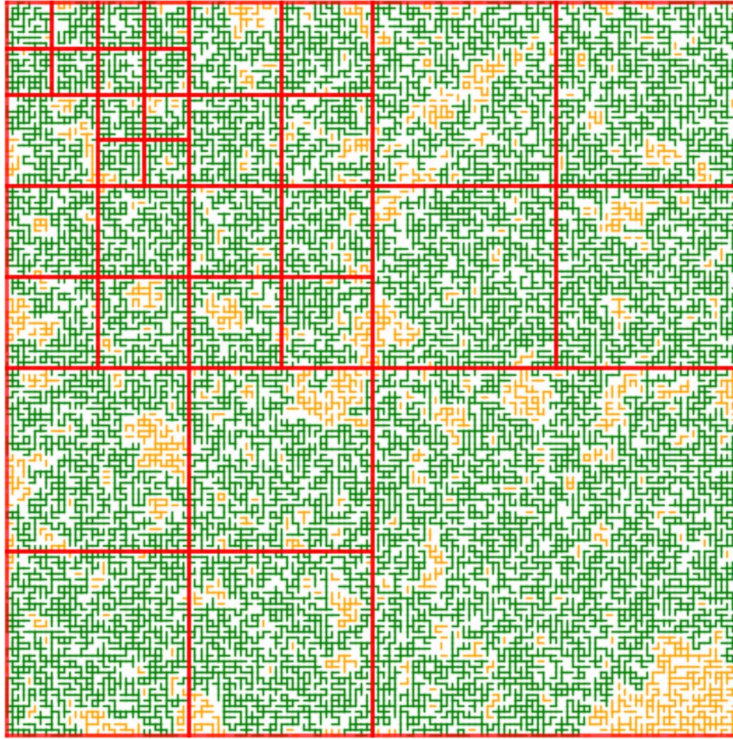
FIG. 2.   *A realization of the partition $\mathcal{P}$ in a cube $\square$. The cluster $\mathscr{C}_*(\square)$ is drawn in green; the clusters in yellow are the small isolated clusters. Simulation by C. Gu.*

Notice that every element of $\mathcal{P}_*$ can be written in a unique way as a disjoint union of elements of $\mathcal{P}$. According to Proposition 1(i), every triadic cube containing an element of $\mathcal{P}$ is good. By Proposition 1(iii) and Lemma 2.1, there exists $C(d, \mathfrak{p}) < \infty$ such that, for every $x \in \mathbb{Z}^d$,

$$(22) \qquad \text{size}(\square_{\mathcal{P}}(x)) \leq \mathcal{O}_1(C).$$

By the properties of the partition $\mathcal{P}$ given in Proposition 1(i) and (ii) and Lemma 2.2, the maximal crossing cluster $\mathscr{C}_*(\square)$ of an element $\square \in \mathcal{P}_*$ must satisfy $\mathscr{C}_*(\square) \subseteq \mathscr{C}_\infty$, since the union of all crossing clusters of the elements of $\mathcal{P}$ is unbounded and connected. Notice also that, although we may not have $\mathscr{C}_*(\square) = \mathscr{C}_\infty \cap \square$, by definition of the partition $\mathcal{P}$ and Property (ii) of Definition 2, we have that, for every cube $\square \in \mathcal{P}$, there exists a cluster $\mathscr{C}$ such that

$$(23) \qquad \mathscr{C}_\infty \cap \square \subseteq \mathscr{C} \subseteq \bigcup_{\square' \in \mathcal{P}, \, \text{dist}(\square, \square') \leq 1} \square'.$$

In other words, for any cube $\square \in \mathcal{P}$ and every pair of points $x, y \in \mathscr{C}_\infty \cap \square$, there exists a path connecting x to y which lies in the cube $\square$ and in its neighbors.

It is also interesting to note that, for $m \in \mathbb{N}$ such that $3^m \geq \mathcal{M}_{2d}(\mathcal{P})$, the sets $\mathscr{C}_*(\square_m)$, $\mathscr{C}_\infty \cap \square_m$ and $\square_m$ have comparable sizes: there exists a constant $C := C(d, \mathfrak{p}) < \infty$ such that

$$(24) \qquad C^{-1}|\square_m| \leq \left|\mathscr{C}_*(\square_m)\right| \leq |\mathscr{C}_\infty \cap \square_m| \leq |\square_m|.$$

This result is a consequence of the Cauchy–Schwarz inequality and the three identities, under the assumption $3^m \geq \mathcal{M}_{2d}(\mathcal{P})$ (which implies in particular that the cube $\square_m$ is good),

$$\sum_{\square \in \mathcal{P}, \, \square \subseteq \square_m} 1 \leq \mathscr{C}_*(\square_m), \qquad \sum_{\square \in \mathcal{P}, \, \square \subseteq \square_m} \text{size}(\square_{\mathcal{P}})^d = |\square_m|$$

and

$$\sum_{\square \in \mathcal{P}, \square \subseteq \square_m} \text{size}(\square_{\mathcal{P}})^{2d} \leq C|\square_m|.$$

The first inequality comes from the fact that each cube of the partition $\mathcal{P}$ contained in the cube $\square_m$ must have nonempty intersection with the cluster $\mathscr{C}_*(\square_m)$, the second one is the preservation of the volume and the third one is a consequence of the assumption $3^m \geq \mathcal{M}_{2d}(\mathcal{P})$.

Given a cube $\square \in \mathcal{P}$, we let $\overline{z}(\square)$ denote the element of the cluster $\mathscr{C}_*(\square)$ which is closest to the point z; if this point is not unique, we break ties by using the lexicographical order.

DEFINITION 4. Given a function $u : \mathscr{C}_\infty \to \mathbb{R}$, we define *the coarsened function with respect to the partition* $\mathcal{P}$ to be

$$[u]_{\mathcal{P}} : \begin{cases} \mathbb{Z}^d \to \mathbb{R}, \\ x \mapsto u\big(\overline{z}(\square_{\mathcal{P}}(x))\big). \end{cases}$$

The reason we use the coarsened function is that it is defined on the entire lattice $\mathbb{Z}^d$ and not on the infinite cluster. This allows to make use of the simpler and more favorable geometric structure of $\mathbb{Z}^d$. The price to pay is the difference between u and $[u]_{\mathcal{P}}$; it depends on the coarseness of the partition $\mathcal{P}$ and the control one has on the gradient of the function u in a way that is made precise in the following proposition. The dependence on the coarseness of $\mathcal{P}$ is present via the size of the cubes of the partition. We recall the notation $|F|(x)$ for a vector field F introduced in (9).

PROPOSITION 2 (Lemma 3.2 of [7]). *For each triadic cube* $\square \in \mathcal{P}_*$, *each exponent* $1 \leq s < \infty$ *and each function* $w : \mathscr{C}_\infty \to \mathbb{R}$,

$$(25) \qquad \sum_{x \in \mathscr{C}_*(\square)} \big|w(x) - [w]_{\mathcal{P}}(x)\big|^s \leq C^s \sum_{x \in \mathscr{C}_*(\square)} \text{size}(\square_{\mathcal{P}}(x))^{sd}|\nabla w|^s(x).$$

More generally, for any family of disjoint cubes $\{\square_i\}_{i \in I} \in (\mathcal{P}_*)^I$, *we have*

$$(26) \qquad \sum_{x \in \mathscr{C}_*(\bigcup_{i \in I} \square_i)} \big|w(x) - [w]_{\mathcal{P}}(x)\big|^s \leq C^s \sum_{x \in \mathscr{C}_*(\bigcup_{i \in I} \square_i)} \text{size}(\square_{\mathcal{P}}(x))^{sd}|\nabla w|^s(x),$$

where $\mathscr{C}_*(\bigcup_{i \in I} \square_i)$ *denotes the union of the maximal clusters of each connected component of the set* $\bigcup_{i \in I} \square_i$.

REMARK 1. We do not have the identity $\mathscr{C}_*(\bigcup_{i \in I} \square_i) = \bigcup_{i \in I} \mathscr{C}_*(\square_i)$. The problem is the same as the one of (23) and thus (26) can not be directly obtained from (25). Nevertheless, we do have the inclusion

$$(27) \qquad \mathscr{C}_\infty \cap \square \subseteq \mathscr{C}\Big(\bigcup_{\square' \in \mathcal{P}, \text{dist}(\square, \square') \leq 1} \square' \Big).$$

Moreover we can control the L^s-norm of the vector field $\nabla[w]_{\mathcal{P}}$ by the L^s-norm of the map ∇w and the coarseness of the partition $\mathcal{P}$ thanks to the following proposition.

PROPOSITION 3 (Lemma 3.3 of [7]). *For each triadic cube* $\square \in \mathcal{P}_*$, *each exponent* $1 \leq s < \infty$ *and each function* $w : \mathscr{C}_\infty \to \mathbb{R}$,

$$(28) \qquad \sum_{x \in \mathscr{C}_*(\square)} |\nabla[w]_{\mathcal{P}}|^s(x) \leq C^s \sum_{x \in \mathscr{C}_*(\square)} \text{size}(\square_{\mathcal{P}}(x))^{sd-1}|\nabla w|^s(x).$$

More generally, for any family of disjoint cubes $\{\square_i\}_{i \in I} \in (\mathcal{P}_*)^I$, *we have*

$$(29) \qquad \sum_{x \in \mathscr{C}_*(\bigcup_{i \in I} \square_i)} |\nabla[w]_{\mathcal{P}}|^s(x) \leq C^s \sum_{x \in \mathscr{C}_*(\bigcup_{i \in I} \square_i)} \text{size}(\square_{\mathcal{P}}(x))^{sd-1}|\nabla w|^s(x).$$

2.3. *Solving the Poisson equation with divergence form source term.* In this section we study the existence and uniqueness of the solution of the equation $-\nabla \cdot \mathbf{a}\nabla u = -\nabla \cdot \xi$ on the infinite cluster $\mathscr{C}_\infty$. We denote by $\sum_{e \subseteq \mathscr{C}_\infty}$ the sum over all the edges of the infinite cluster.

PROPOSITION 4 (Gradient of Green's function). *Given an environment $\mathbf{a} \in \Omega$, we select an edge $e = (x, y) \in E_d$ such that the points x and y belong to the infinite cluster $\mathscr{C}_\infty$. There exist a constant $C := C(d, \lambda) < \infty$ and a function $\nabla G(e, \cdot) : \mathscr{C}_\infty \to \mathbb{R}$, whose gradient with respect to the second variable, denoted by $\nabla\nabla G$, satisfies*

$$(30) \qquad \langle \nabla\nabla G(e, \cdot), \nabla\nabla G(e, \cdot)\rangle_{\mathscr{C}_\infty} \leq C,$$

and is a solution to the equation

$$-\nabla \cdot \mathbf{a}\nabla(\nabla G(e, \cdot)) = \delta_x - \delta_y \quad in \ \mathscr{C}_\infty.$$

Moreover, we have, for each pair of edges e, e' of the infinite cluster,

$$(31) \qquad \nabla\nabla G(e, e') = \nabla\nabla G(e', e).$$

Proposition 4 can be used to solve the equation $-\nabla \cdot \mathbf{a}\nabla w_\xi = -\nabla \cdot \xi$. This is the objective of the following proposition.

PROPOSITION 5. *Let $\xi : E_d \to \mathbb{R}$ be a vector field satisfying*

$$(32) \qquad \xi(x, y) = 0 \quad if \ \mathbf{a}(x, y) = 0 \ or \ x, y \notin \mathscr{C}_\infty.$$

If ξ satisfies $\langle \xi, \xi \rangle_{\mathscr{C}_\infty} < \infty$ then there exists a unique (a.s in the environment and up to a constant) solution w_ξ of the equation

$$-\nabla \cdot \mathbf{a}\nabla w_\xi = -\nabla \cdot \xi \quad in \ \mathscr{C}_\infty.$$

Moreover, we have the following representation

$$(33) \qquad \nabla w_\xi(\cdot) = \sum_{e \subseteq \mathscr{C}_\infty} \xi(e)\nabla\nabla G(e, \cdot).$$

REMARK 2. We extend the definition of Proposition 5 to vector-valued fields $\xi : E_d \to \mathbb{R}^k$. In that case, we will write

$$w_\xi := \begin{cases} \mathscr{C}_\infty \to \mathbb{R}^k \\ z \mapsto \left(w_{\xi_1}(z), \ldots, w_{\xi_1}(z)\right), \end{cases}$$

where $\xi_1, \ldots, \xi_k$ denote the components of the vector ξ; the formula (33) applies in this framework.

PROOF OF PROPOSITIONS 4 AND 5. Let ξ be a vector field satisfying (32) and the inequality $\langle \xi, \xi \rangle_{\mathscr{C}_\infty} < \infty$. We denote by $\dot{H}^1$ the space of functions defined on the infinite cluster whose gradient is in the space $L^2(\mathscr{C}_\infty)$, that is, $\dot{H}^1 := \{u : \mathscr{C}_\infty \to \mathbb{R} : \langle \nabla u, \nabla u \rangle_{\mathscr{C}_\infty} < \infty\}$, and consider the minimization problem

$$\inf_{u \in \dot{H}^1} \frac{1}{2}\langle \nabla u, \mathbf{a}\nabla u \rangle_{\mathscr{C}_\infty} - \langle \xi, \nabla u \rangle_{\mathscr{C}_\infty}.$$

By the standard techniques of the calculus of variations, there exists a unique solution (up to a constant) to this problem denoted by w_ξ. In particular, when ξ is the indicator of an edge e, we obtain the function $\nabla G(e, \cdot)$. To prove the identity (31), we note that

$$\nabla\nabla G(e', e) = \langle \nabla\nabla G(e, \cdot), \mathbf{a}\nabla\nabla G(e', e)\rangle_{\mathscr{C}_\infty} = \nabla\nabla G(e, e').$$

The representation formula (33) follows from standard arguments. $\square$

2.4. *Regularity theory.*   In this subsection, we record a result from the regularity theory established in [7] giving a Lipschitz bound for the gradient of **a**-harmonic functions.

PROPOSITION 6 (Regularity theory on $\mathscr{C}_\infty$, Theorem 2 of [7]).   *There exist a constant $C < \infty$, an exponent $s > 0$ and a random variable $\mathcal{X}$ satisfying*

$$\mathcal{X} \leq \mathcal{O}_s(C), \tag{34}$$

such that for each solution $u : \mathscr{C}_\infty \to \mathbb{R}$ of the equation

$$-\nabla \cdot \mathbf{a}\nabla u = 0 \quad \text{in } \mathscr{C}_\infty \tag{35}$$

and each pair of radii R, r such that $R \geq r \geq \mathcal{X}$, we have

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leq \frac{C}{R}\|u - (u)_{\mathscr{C}_\infty \cap B_R}\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_R)}.$$

We introduce the notation, for each point $x \in \mathbb{Z}^d$,

$$\mathcal{X}(x) := \mathcal{X} \circ \tau_x.$$

This proposition is weaker than Theorem 2 of [7]; it is indeed a consequence of the Caccioppoli inequality and Theorem 2 (iii) of [7] for $k = 0$. As a consequence, we obtain the following Lipschitz bound on the corrector.

PROPOSITION 7 (Lipschitz bound on the corrector).   *There exists a constant $C < \infty$ and an exponent $s > 0$ such that, for each edge $e = (x, y) \in E_d$ and each vector $p \in \mathbb{R}^d$,*

$$|\nabla \chi_p(e)|\mathbb{1}_{\{e \in \mathscr{C}_\infty\}} \leq C|p|\mathcal{X}^{d/2}(x), \tag{36}$$

which implies, by (34),

$$|\nabla \chi_p(e)|\mathbb{1}_{\{e \in \mathscr{C}_\infty\}} \leq \mathcal{O}_s(C|p|) \tag{37}$$

for some smaller exponent s (cf. Section 1.2.5). The same estimate holds for the coarsened corrector

$$|\nabla[\chi_p]_{\mathcal{P}}(e)| \leq \mathcal{O}_s(C|p|). \tag{38}$$

REMARK 3.   The same estimate as (36) would hold with the random variable $\mathcal{X}^{d/2}(y)$ instead of $\mathcal{X}^{d/2}(x)$ in the right-hand side.

PROOF.   By the stationarity of the law of the corrector, we can assume that the edge e touches 0, that is, that $x = 0$. First note that, for each radius $r \geq 1$,

$$|\nabla \chi_p(e)|\mathbb{1}_{\{e \in \mathscr{C}_\infty\}} \leq r^d\|\nabla \chi_p\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)}.$$

By applying Proposition 6 with $r = \mathcal{X}$, and taking the limit $R \to \infty$, we obtain

$$\|p + \nabla \chi_p\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{\mathcal{X}})} \leq C\mathcal{X}^{d/2} \liminf_{R \to \infty} \frac{1}{R}\|l_p + \chi_p - (l_p + \chi_p)_{\mathscr{C}_\infty \cap B_R(x')}\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_R(x'))}$$

$$\leq C\mathcal{X}^{d/2}|p|.$$

A combination of the two previous displays with the integrability estimate (34) yields (37). To prove (38), we combine (37) with Proposition 3 and use the integrability estimate $\text{size}(\Box_{\mathcal{P}}(x)) \leq \mathcal{O}_s(C)$, valid for each point $x \in \mathbb{Z}^d$. This is performed in the following com-

putation: for each edge $e = (x, y) \in \mathbb{R}^d$, we have

$$|\nabla[\chi_p]_{\mathcal{P}}(e)| \leq \sum_{x' \in \mathscr{C}_*(\Box_{\mathcal{P}}(x) \cup \Box_{\mathcal{P}}(y))} |\nabla[\chi_p]_{\mathcal{P}}|(x')$$

$$(39) \qquad \leq C \sum_{x' \in \mathscr{C}_\infty \cap B(x, C\, \mathrm{size}(\Box_{\mathcal{P}}(x)))} \mathrm{size}(\Box_{\mathcal{P}}(x'))^{d-1} |\nabla \chi_p|(x')$$

$$\leq C \sum_{x' \in \mathbb{Z}^d} \mathbb{1}_{\{x' \in \mathscr{C}_\infty \cap B(x, C\, \mathrm{size}(\Box_{\mathcal{P}}(x)))\}} \mathrm{size}(\Box_{\mathcal{P}}(x'))^{d-1} |\nabla \chi_p|(x').$$

Using the estimate, for each point $x \in \mathbb{Z}^d$, $\mathrm{size}(\Box_{\mathcal{P}}(x)) \leq \mathcal{O}_s(C)$, we obtain

$$(40) \qquad \mathbb{1}_{\{x' \in B(x, C\, \mathrm{size}(\Box_{\mathcal{P}}(x)))\}} \leq C \frac{\mathrm{size}(\Box_{\mathcal{P}}(x))^{d+1}}{|x - x'|^{d+1} \vee 1} \leq \frac{\mathcal{O}_s(C)}{|x - x'|^{d+1} \vee 1},$$

where we used the notation $a \vee b := \max(a, b)$. Using the summability of the map $x \mapsto (|x| \vee 1)^{-d-1}$, the properties (12) and (14) on the $\mathcal{O}_s$ notation and the Lipschitz bounds (37) on the corrector, we obtain the result. $\quad\square$

We now present the two main tools to prove Theorem 1.2. The first one is a concentration inequality, thanks to which we obtain quantitative control on the spatial averages of the gradient at scale R (see Proposition 10). We then deduce Theorem 1.2 from Proposition 10 thanks to the multiscale Poincaré inequality (Proposition 9).

2.5. *Concentration inequality for stretched exponential moments.* The following concentration inequality is a key ingredient in the proof of Proposition 10 in the next section; its proof can be found in [11], Proposition 2.2.

PROPOSITION 8 (Proposition 2.2 of [11]). *Fix $\beta \in (0, 2)$. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and set for each bond $e \in \mathcal{B}_d(\mathbb{Z}^d)$,*

$$X'_e = \mathbb{E}[X | \mathcal{F}(\mathcal{B}_d \setminus \{e\})] \quad and \quad \mathbb{V}[X] = \sum_{e \in \mathcal{B}_d} (X - X'_e)^2,$$

then there exists a constant $C := C(d, \beta) < \infty$ such that

$$\mathbb{E}[\exp(|X - \mathbb{E}[X]|^\beta)] \leq C \mathbb{E}[\exp((C\mathbb{V}[X])^{\frac{\beta}{2-\beta}})]^{\frac{2-\beta}{\beta}}.$$

2.6. *Multiscale Poincaré inequality.* The next proposition is a version of the multiscale Poincaré inequality. It controls the oscillations of a function in the L^q-norm (left-hand side of (42)) by the spatial average of the gradient of the function (right-hand side of (42)). We first introduce the discrete heat kernel.

DEFINITION 5 (Discrete heat-kernel). Let $\Phi : [0, \infty) \times \mathbb{Z}^d \to \mathbb{R}$ be the discrete heat kernel on the lattice $\mathbb{Z}^d$, that is, the solution of the parabolic equation

$$\begin{cases} \partial_t \Phi - \Delta \Phi = 0 & \text{in } (0, \infty) \times \mathbb{Z}^d, \\ \Phi(0, \cdot) = \delta_0 & \text{in } \mathbb{Z}^d, \end{cases}$$

where Δ denotes the discrete Laplacian on $\mathbb{Z}^d$. We introduce the notation, for each radius $r > 0$, $\Phi_{r^2} := \Phi(r^2, \cdot)$. It satisfies the estimate, for some constant $C := C(d) < \infty$,

$$(41) \qquad \begin{cases} \Phi(t, x) \leq \dfrac{C}{t^{\frac{d}{2}}} \exp\left(-\dfrac{|x|^2}{Ct}\right) & \text{if } t \geq |x| \qquad \text{(Gaussian regime)}, \\ \Phi(t, x) \leq C \exp\left(-\dfrac{|x|}{C}\left(1 + \ln \dfrac{|x|}{t}\right)\right) & \text{if } t \leq |x| \qquad \text{(Poisson regime)}. \end{cases}$$

We refer to [28] and [26] for a proof of these inequalities.

PROPOSITION 9 (Multiscale Poincaré inequality, heat kernel version).

For each exponent $q \geq 1$, there exists a constant $C := C(d, q) < \infty$ such that for each function $u : \mathbb{Z}^d \to \mathbb{R}$ and each radius $R > 0$,

$$(42) \qquad \|u - (u)_{B_R}\|_{\underline{L}^q(B_R)} \leq C\left(\sum_{x \in \mathbb{Z}^d} R^{-d} e^{-\frac{|x|}{2R}} \left(\int_0^{2R} r |\Phi_{r^2} * \nabla u(x)|^2 \, dr \right)^{\frac{q}{2}} \right)^{\frac{1}{q}},$$

where the operator $*$ is the standard discrete convolution on $\mathbb{Z}^d$ between the heat kernel Φ_{r^2} and the vector valued function ∇u (see the definition (8)). Moreover the dependence in the variable q of the constant C can be quantified as follows, for each exponent $q \geq 2$,

$$C(d, q) \leq A q^{\frac{3}{2}}$$

for some constant $A := A(d) < \infty$.

The proof of this proposition relies on [10], Proposition D.1 and Remark D.6, and is presented in Appendix A.

3. Estimates of the spatial averages of the first-order correctors.   We now have collected all the necessary tools to prove the optimal L^q bounds of the corrector, stated in Theorem 1.2. The strategy is to first prove Proposition 10 thanks to the concentration inequality (Proposition 8). We then deduce the bound on the coarsened corrector thanks to the multiscale Poincaré inequality (Proposition 9) and remove the coarsening thanks to Proposition 2. This eventually yields Theorem 1.2. In this section, we use the notation introduced in (8) and think of the gradient of the coarsened corrector as a vector-valued function.

PROPOSITION 10.   *For each $R \geq 1$, and each $x \in \mathbb{R}^d$, the quantity $(\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(x)$ is well defined and there exist an exponent $s > 0$ and a constant $C < \infty$ such that it satisfies*

$$(43) \qquad \left| (\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(x) \right| \leq \mathcal{O}_s\left(C|p|r^{-\frac{d}{2}} \right).$$

By the stationarity of the gradient of the corrector, it is enough to prove the result when $x = 0$. By linearity of the mapping $p \mapsto \nabla \chi_p$, we may assume $|p| = 1$. We denote by $X = (\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(0)$ and prove

$$|X| \leq \mathcal{O}_s\left(C r^{-\frac{d}{2}} \right).$$

The strategy of the proof is to apply Proposition 8 to the random variable X. We decompose the argument into two lemmas. The first one focuses on the expectation of X.

LEMMA 3.1.   *There exists a constant $C < \infty$ such that*

$$\left| \mathbb{E}[X] \right| \leq C r^{-\frac{d}{2}}.$$

The second one estimates the quantity $\mathbb{V}[X]$.

LEMMA 3.2.   *There exist a constant $C < \infty$ and an exponent $s > 0$, such that*

$$\mathbb{V}[X] \leq \mathcal{O}_s\left(C r^{-d} \right).$$

These lemmas are proved in the following two sections.

3.1. *Estimating the expectation of the spatial averages.* The main objective of this section is to show Lemma 3.1.

PROOF OF LEMMA 3.1. The strategy of the proof is to use the stationarity and the sublinearity of the corrector to prove that the expectation of its gradient is equal to 0. The technical difficulty which arises is that the partition $\mathcal{P}$ is not stationary. This implies that the random variable $\nabla[\chi_p]_{\mathcal{P}}(0)$ is not stationary. To fix this issue we introduce a partition $\mathcal{P}_{\text{stat}}$ which is stationary and equal to $\mathcal{P}$ on a set of large probability. We finally show that the error we make by considering the partition $\mathcal{P}_{\text{stat}}$ instead of $\mathcal{P}$ is small.

For each triplet $x, y, z \in \mathbb{Z}^d$ with $x \sim y$, denote by $\tau_z \mathbf{a}$ the translated environment defined by

$$\tau_z \mathbf{a}(\{x, y\}) = \mathbf{a}(\{x - z, y - z\}).$$

For $k \in \mathbb{N}$, we construct the partition $\mathcal{P}_{\text{stat}}^k$ by applying Proposition 1 to the collection of triadic cubes

$$\mathcal{G}_{\text{stat}}^k := \mathcal{G} \cup \left(\bigcup_{n=k}^{\infty} \mathcal{T}_n \right).$$

Note that this collection is not a set of good cubes in the sense of Definition 2 but it is $3^k \mathbb{Z}^d$-translation invariant. A straightforward consequence is that the partition $\mathcal{P}_{\text{stat}}^k$ is $3^k \mathbb{Z}^d$-stationary: for every environment $\mathbf{a}$, every point $x \in \mathbb{Z}^d$ and $z \in 3^k \mathbb{Z}^d$, one has

(44) $$\text{size}(\square_{\mathcal{P}_{\text{stat}}^k}(x + z))(\tau_z \mathbf{a}) = \text{size}(\square_{\mathcal{P}_{\text{stat}}^k}(x))(\mathbf{a}).$$

With a proof similar to the one of [7], Proposition 2.1 (iv), we obtain

(45) $$\mathbb{P}[\exists x \in \square_k, \square_{\mathcal{P}}(x) \neq \square_{\mathcal{P}_{\text{stat}}^k}(x)] \leq C \exp(-C^{-1} 3^k).$$

For a function $u : \mathscr{C}_\infty \to \mathbb{R}$, we define the coarsened function $[u]_{\mathcal{P}_{\text{stat}}^k}$ with respect to the partition $\mathcal{P}_{\text{stat}}^k$ by the formula

$$[u]_{\mathcal{P}_{\text{stat}}^k} := u\left(\overline{z}_{\text{stat}}(\square_{\mathcal{P}_{\text{stat}}^k}(x))\right)$$

with the notation, for $\square \in \mathcal{T}$,

(46) $$\overline{z}_{\text{stat}}(\square) := \begin{cases} \overline{z}(\square) & \text{if } \overline{z}(\square) \in \mathscr{C}_\infty \text{ and } \square \text{ is a good cube,} \\ \underset{z \in \mathscr{C}_\infty}{\text{argmin}} \, \text{dist}(z, \square) & \text{otherwise.} \end{cases}$$

If there is more than one choice in the argument of the minima, we select the one which is minimal for the lexicographical order. By the stationarity of the gradient of the corrector and (44), we have

(47) $$\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k} \text{ is } 3^k \mathbb{Z}^d\text{-stationnary.}$$

We let $k \in \mathbb{Z}^d$ be the the integer such that $3^k \leq r^{\frac{1}{2}} \leq 3^{k+1}$ and split the proof of Lemma 3.1 into three steps:

(i) In Step 1, we prove

$$\mathbb{E}[|(\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(0) - (\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k})(0)|] \leq C r^{-\frac{d}{2}}.$$

(ii) In Step 2, we prove

$$\mathbb{E}\left[\sum_{x \in \square_k} \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x) \right] = 0.$$

(iii) In Step 3, we use the result obtained in Step 2 to show

$$\left|\mathbb{E}\left[(\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}})(0)\right]\right| \leq C r^{-\frac{d}{2}}.$$

Lemma 3.1 is then a consequence of the main results of Steps 1 and 3.

Step 1. The main result of this step is a consequence of the following computation, by (7) and Proposition 3,

$$\mathbb{E}\left[\left|(\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(0) - (\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}})(0)\right|\right]$$

(48)
$$\leq \mathbb{E}\left[\sum_{x \in B_{r^2}} \left|\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}}(x)\right| \Phi_{r^2}(x) \mathbb{1}_{\{\exists x \in B_{r^2} : \Box_{\mathcal{P}^k_{\text{stat}}}(x) \neq \Box_{\mathcal{P}}(x)\}}\right]$$

$$+ \mathbb{E}\left[\left|\sum_{x \in \mathbb{Z}^d \setminus B_{r^2}} (\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}}(x))\Phi_{r^2}(x)\right|\right].$$

The first term on the right-hand side can be estimated (crudely) the following way. We denote by U_0 the set

$$U_0 := \bigcup_{x \in B_{r^2}} \Box_{\mathcal{P}}(x),$$

we then enlarge this set by adding two additional layers of cubes and define

$$U_1 := \bigcup_{\Box \in \mathcal{P}, \text{dist}(\Box, U_0) \leq 1} \Box \quad \text{and} \quad U := \bigcup_{\Box \in \mathcal{P}, \text{dist}(\Box, U_1) \leq 1} \Box.$$

Note that, by the properties of the partition $\mathcal{P}$ and (12), we have the inequality

(49)
$$|U| = C|U_1| \leq C|U_0| \leq C \sum_{x \in B_{r^2}} \text{size}(\Box_{\mathcal{P}}(x))^d \leq \mathcal{O}_s(Cr^{2d}).$$

Also with these definitions, we have, for each point $x \in B_{r^2}$,

$$\left|\nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}}(x)\right| \leq \sum_{y \in \mathscr{C}_\infty \cap U} |\nabla \chi_p|(y).$$

Similarly, for each point $x \in B_{r^2}$,

$$\left|\nabla[\chi_p]_{\mathcal{P}}(x)\right| \leq \sum_{y \in \mathscr{C}_\infty \cap U} |\nabla \chi_p|(y).$$

This leads to the estimate

$$\left|\sum_{x \in B_{r^2}} (\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}^k_{\text{stat}}}(x))\Phi_{r^2}(x)\right|$$

(50)
$$\leq C\left(\sum_{y \in \mathscr{C}_\infty \cap U} |\nabla \chi_p|(y)\right)\left(\sum_{x \in B_{r^2}} \Phi_{r^2}(x)\right)$$

$$\leq C \sum_{y \in \mathscr{C}_\infty \cap U} |\nabla \chi_p|(y).$$

Using Proposition 37, the estimate on the volume of U given in (49) and a computation similar to the one performed in (39), we obtain

$$\sum_{y \in \mathscr{C}_\infty \cap U} |\nabla \chi_p|(y) \leq \mathcal{O}_s(Cr^{2d}).$$

Then by (45), we also have

$$\mathbb{P}\big[\exists x \in B_{r^2} : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\big] \leq \sum_{z \in 3^k \mathbb{Z}^d \cap B_{r^2}} \mathbb{P}\big[\exists x \in z + \Box_k : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\big]$$

$$\leq \frac{r^{2d}}{3^{dk}} \mathbb{P}\big[\exists x \in \Box_k : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\big]$$

$$\leq \frac{C r^{2d}}{3^{dk}} \exp(-C^{-1} 3^k).$$

In particular, since k has been chosen such that $3^k \leq r^{\frac{1}{2}} < 3^{k+1}$, for each exponent $q > 0$, there exist a constant $C := C(d, \mathfrak{p}, \lambda, q) < \infty$ and an exponent $s := s(d, \mathfrak{p}, \lambda, q) > 0$ such that

$$\mathbb{1}_{\{\exists x \in B_{r^2} : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\}} \leq \mathcal{O}_s(C r^{-q}).$$

Combining the three previous displays with q chosen large enough, the Cauchy–Schwarz inequality and (14), we obtain

$$\left| \sum_{x \in B_{r^2}} (\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)) \Phi_{r^2}(x) \right| \mathbb{1}_{\{\exists x \in B_{r^2} : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\}} \leq \mathcal{O}_s(C r^{-\frac{d}{2}}),$$

which yields in particular

$$\mathbb{E}\left[\left| \sum_{x \in B_{r^2}} (\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)) \Phi_{r^2}(x) \right| \mathbb{1}_{\{\exists x \in B_{r^2} : \Box_{\mathcal{P}_{\text{stat}}^k}(x) \neq \Box_{\mathcal{P}}(x)\}} \right] \leq C r^{-\frac{d}{2}}.$$

We now focus on estimating the second term on the right-hand side of (48). With the same computation as the one we just wrote, one obtains

$$\sum_{x \in B_{r^2}} |\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)| \leq \mathcal{O}_s(C r^{4d}).$$

The proof is identical, we only need to replace the term $\sum_{x \in B_{r^2}} \Phi_{r^2}(x)$ by $C r^{2d}$ in (50). Since this result is valid for any radius $r \geq 1$, we obtain, for each integer $n \in \mathbb{N}$,

$$\sum_{x \in \mathscr{C}_\infty \cap (\Box_{n+1} \setminus \Box_n)} |\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)| \leq \sum_{x \in \mathscr{C}_\infty \cap B_{3^n}} |\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)|$$

$$\leq \mathcal{O}_s(C 3^{4dn}).$$

We then use the estimate (41) on the discrete heat kernel and write

$$\mathbb{E}\left[\sum_{x \in \mathbb{Z}^d \setminus B_{r^2}} |\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)| \Phi_{r^2}(x) \right]$$

$$\leq \sum_{n=\lfloor 2\log_3(r) \rfloor}^{+\infty} \mathbb{E}\left[\exp\left(-\frac{3^n}{r}\right) r^{-d} \sum_{xx \in \mathscr{C}_\infty \cap (\Box_{n+1} \setminus \Box_n)} |\nabla[\chi_p]_{\mathcal{P}}(x) - \nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)| \right]$$

$$\leq \sum_{n=2\log_3(r)}^{+\infty} \exp\left(-\frac{3^n}{r}\right) r^{-d} 3^{4dn}$$

$$\leq C \exp(-C^{-1} r).$$

Combining the estimates of the first and the second terms of the right-hand side completes the proof of Step 1.

REMARK 4.   Most of the estimates of this proof are crude; precise results are not needed. The same argument shows the following (stronger) result: for each exponent $q > 0$, there exists a constant $C := C(d, \mathfrak{p}, \lambda, q) < \infty$ such that for each radius $r \geq 1$ and each integer $k \in \mathbb{N}$ satisfying $3^k \leq r^{\frac{1}{2}} < 3^{k+1}$,

$$\mathbb{E}\big[|\Phi_{r^2} * \nabla[\chi_P]_{\mathcal{P}} - \Phi_{r^2} * \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}|\big] \leq Cr^{-q}.$$

The proof of Lemma 3.1 only requires the result with the value $q = \frac{d}{2}$.

Step 2. We prove the main result of this step by combining the stationarity property (47) with the sublinear growth of the corrector. First notice that by (17), we have, for each radius $r > 1$,

$$\underset{\mathscr{C}_\infty \cap B_r}{\text{osc}} \chi_P \leq \mathcal{O}_s(Cr^{1-\delta}).$$

By the Stokes formula, we have, for each integer $n \in \mathbb{N}$,

$$\left| \sum_{x \in \square_{nk}} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \right| = \left| \sum_{x \in \partial \square_{nk}} [\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x)\mathbf{n}(x) \right| \leq C3^{kn(d-1)} \underset{\mathscr{C}_\infty \cap \square_{nk}}{\text{osc}} \chi_P$$

$$\leq \mathcal{O}_s(C3^{kn(d-\delta)}),$$

where the map $x \mapsto \mathbf{n}(x)$ is the discrete outer normal to the cube $\square_{nk}$. This yields

$$\left| \mathbb{E}\bigg[\sum_{x \in \square_{nk}} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg] \right| \leq C3^{kn(d-\delta)}.$$

We also have, by the stationarity property (47),

$$\mathbb{E}\bigg[\sum_{x \in \square_{nk}} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg] = \sum_{z \in (3^k \mathbb{Z}^d \cap \square_{kn})} \mathbb{E}\bigg[\sum_{x \in z + \square_k} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg]$$

$$= \frac{3^{dkn}}{3^{dk}} \mathbb{E}\bigg[\sum_{x \in \square_k} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg].$$

Combining the two previous results shows

$$\left| \mathbb{E}\bigg[\sum_{x \in \square_k} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg] \right| \leq C3^{dk}3^{-kn\delta}.$$

Sending $n \to \infty$ shows

$$\left| \mathbb{E}\bigg[\sum_{x \in \square_k} \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x) \bigg] \right| = 0.$$

Step 3. First notice that

$$\mathbb{E}\big[(\Phi_{r^2} * \nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k})(0)\big] = (\Phi_{r^2} * \mathbb{E}[\nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}])(0).$$

By (47), the function

$$f := \begin{cases} \mathbb{Z}^d \to \mathbb{R}^d \\ x \mapsto \mathbb{E}\big[\nabla[\chi_P]_{\mathcal{P}_{\text{stat}}^k}(x)\big] \end{cases}$$

is $3^k \mathbb{Z}^d$-periodic. Consequently, there exist complex coefficients $(a_{\mathbf{n}})_{\mathbf{n} \in \square_k}$ such that

$$f(x) = \sum_{\mathbf{n} \in \square_k} a_{\mathbf{n}} \exp\left(\frac{2i\pi \mathbf{n} \cdot x}{3^k} \right).$$

Using that Φ is the solution of the discrete heat equation, which implies that the coefficients of its discrete Fourier transform can be explicitly computed, we obtain the identity

$$(\Phi_{r^2} * \mathbb{E}[\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}}])(0) = \sum_{\mathbf{n} \in \square_k} a_{\mathbf{n}} \exp\left(-r^2 \sum_{i=1}^{d} 2\left(1 - \cos\left(\frac{2\pi \mathbf{n}_i}{k}\right)\right)\right).$$

Notice that the main result of Step 2 is equivalent to the following equality

$$a_0 = 0.$$

Using this identity, the Cauchy–Schwarz inequality and the lower bound $1 - \cos a \geq \frac{a^2}{C}$ for $a \in [-\pi, \pi]$ and a universal constant C, we obtain

$$(51) \qquad |(\Phi_{r^2} * \mathbb{E}[\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}])(0)|^2 \leq C\left(\sum_{\mathbf{n} \in \square_k \setminus \{0\}} |a_{\mathbf{n}}|^2\right)\left(\sum_{\mathbf{n} \in \square_k \setminus \{0\}} \exp\left(-\frac{r^2|\mathbf{n}|^2}{C3^{2k}}\right)\right).$$

Since the integer k was chosen such that $3^k \leq r^{\frac{1}{2}} < 3^{k+1}$, we have

$$(52) \qquad \sum_{\mathbf{n} \in \square_k \setminus \{0\}} \exp\left(-\frac{r^2|\mathbf{n}|^2}{C3^{2k}}\right) \leq C \exp(-C^{-1}r).$$

Moreover, we have

$$\sum_{\mathbf{n} \in \square_k} |a_{\mathbf{n}}|^2 \leq \mathbb{E}\left[\sum_{x \in \square_k} |\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)|^2\right].$$

With the same computation as the one performed in Step 1, we obtain

$$\sum_{x \in \square_k} |\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}(x)|^2 \leq \mathcal{O}_s(C|p|^2 3^{4kd}).$$

Taking the expectation yields

$$\sum_{\mathbf{n} \in \square_k} |a_{\mathbf{n}}|^2 \leq C3^{4kd}.$$

Combining this inequality with (51) and (52), we obtain

$$|(\Phi_{r^2} * \mathbb{E}[\nabla[\chi_p]_{\mathcal{P}_{\text{stat}}^k}])(0)|^2 \leq Cr^{2d} \exp(-C^{-1}r) \leq C \exp(-C^{-1}r),$$

where we increased the value of the constant C in the second inequality to absorb the algebraic growth of the term r^{2d}. This implies in particular the main result of Step 3 and completes the proof of Lemma 3.1. $\square$

3.2. *Estimating the resampling of the spatial averages.*

In this section, we prove Lemma 3.2 which is recalled below.

LEMMA 3.2. *There exist a constant $C < \infty$ and an exponent $s > 0$, such that*

$$\mathbb{V}[X] \leq \mathcal{O}_s(Cr^{-d}).$$

PROOF. We recall Proposition 8 and the notation $X = (\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}})(0)$. Given an environment $\mathbf{a} \in \Omega$ and a bond $e = \{x, y\} \in \mathcal{B}_d$, we want to estimate the term $(X - X'_e)^2$. To this end, one needs to understand how changing the conductance of the bond e can affect the infinite cluster $\mathscr{C}_\infty$ and the partition $\mathcal{P}$. This is studied in the following lemma.

LEMMA 3.3. *There exist two constants $C_0 := C_0(d) < \infty$ and $C := C(d) < \infty$ such that for each bond $e = \{x, y\} \in \mathcal{B}_d$, each pair of environments $\mathbf{a}, \widetilde{\mathbf{a}} \in \Omega$ which are equal on the set $\mathcal{B}_d \setminus \{e\}$ and each point $z \in B(x, C_0 \operatorname{size}(\Box_{\mathcal{P}}(x)))$, one has the estimate*

$$\operatorname{size}(\Box_{\mathcal{P}(\widetilde{\mathbf{a}})}(z)) \leq C \operatorname{size}(\Box_{\mathcal{P}(\mathbf{a})}(x)).$$

Moreover, for each point $z \in \mathbb{Z}^d \setminus B(x, C_0 \operatorname{size}(\Box_{\mathcal{P}}(x)))$, one has the identity

$$\operatorname{size}(\Box_{\mathcal{P}(\widetilde{\mathbf{a}})}(z)) = \operatorname{size}(\Box_{\mathcal{P}(\mathbf{a})}(z)).$$

PROOF. The main ingredients of the proof are listed below:

1. If a good cube $\Box \in \mathcal{P}_*$ is such that $3\Box \cap \{x, y\} = \varnothing$ then $\Box$ is a good cube under the environment $\widetilde{\mathbf{a}}$.

2. By the properties of the partition $\mathcal{P}$, every cube $\Box \in \mathcal{P}$ which does not contain the points x and y is crossable under the environment $\widetilde{\mathbf{a}}$. The predecessors of $\Box_{\mathcal{P}}(x)$ and $\Box_{\mathcal{P}}(y)$ are also crossable under the environment $\widetilde{\mathbf{a}}$.

3. By resampling the conductance of one bond, we cannot create an isolated cluster of size larger than $C \operatorname{size}(\Box_{\mathcal{P}}(x))$, for some constant $C_0 := C_0(d) < \infty$. In particular, there exists a constant $C := C(d) < \infty$ such that every good cube of size larger than $C \operatorname{size}(\Box_{\mathcal{P}}(x))$ under the environment $\mathbf{a}$ satisfies Property (ii) of Definition 2 under the environment $\widetilde{\mathbf{a}}$.

4. There exists a constant $C := C(d) < \infty$ such that every cube of size larger than $C \operatorname{size}(\Box_{\mathcal{P}}(x))$ intersecting the cube $\Box_{\mathcal{P}}(x)$ is crossable by a cluster which does not intersect the cube $\Box_{\mathcal{P}}(x)$.

5. If, for a point $y \in B(x, C_0 \operatorname{size}(\Box_{\mathcal{P}}(x)))$, the size of the cube $\Box_{\mathcal{P}}(y)$ is larger than $C \operatorname{size}(\Box_{\mathcal{P}}(x))$, then the point x belongs to $\Box_{\mathcal{P}}(y)$ or one of its neighbors and thus $\operatorname{size}(\Box_{\mathcal{P}}(y)) \leq C \operatorname{size}(\Box_{\mathcal{P}}(x))$.

Combining these properties shows that every good cube $\Box$ under the environment $\mathbf{a}$ satisfying the estimate $\operatorname{size}(\Box) \geq C \operatorname{size}(\Box_{\mathcal{P}}(x))$ is a good cube under the environment $\widetilde{\mathbf{a}}$. It is then straightforward to see from the previous remarks and the construction of the partition $\mathcal{P}$ in the proof of Proposition 1 that the conclusion of the lemma is valid. $\quad\Box$

To estimate the random variable $(X - X'_e)^2$, we introduce an extended probability space by doubling the variables $(\Omega', \mathcal{F}', \mathbb{P}') = (\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \otimes \mathbb{P})$. Given an environment $(\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d} \in \Omega'$, we denote by pr_1 (resp. pr_2) the first (resp. second) projection, that is, $\operatorname{pr}_1((\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}) = (\mathbf{a}(e'))_{e' \in \mathcal{B}_d}$ (resp. $\operatorname{pr}_2((\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}) = (\widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}$). Any random variable Z defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$ can be seen as a random variable defined on the extended space $(\Omega', \mathcal{F}', \mathbb{P}')$ by the formula $Z = Z \circ \operatorname{pr}_1$, that is,

$$Z((\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}) = Z((\mathbf{a}(e'))_{e' \in \mathcal{B}_d}).$$

Given an enlarged environment $(\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}$, we denote by $\mathbf{a}$ the environment $(\mathbf{a}(e'))_{e' \in \mathcal{B}_d}$ and by $\mathbf{a}^e$ the environment $((\mathbf{a}(e'))_{e' \in \mathcal{B}_d \setminus \{e\}}, \widetilde{\mathbf{a}}(e))$. Similarly, given a random variable Z defined on the space Ω and a bond $e \in \mathcal{B}_d$, we denote by Z^e the random variable defined on the space $(\Omega', \mathcal{F}', \mathbb{P}')$ by the formula

$$(53) \qquad Z^e((\mathbf{a}(e'), \widetilde{\mathbf{a}}(e'))_{e' \in \mathcal{B}_d}) := Z(\mathbf{a}^e).$$

We denote by $\mathcal{P}^e$ and $\mathscr{C}^e_\infty$ the partition of good cubes and the infinite cluster under the environment $\mathbf{a}^e$. It follows from the previous definitions that, for almost every environment $\mathbf{a} \in \Omega$,

$$X(\mathbf{a}) - X'_e(\mathbf{a}) = \int_\Omega \left(\Phi_{r^2} * \left(\nabla [\chi_p]_{\mathcal{P}}(\mathbf{a}, \widetilde{\mathbf{a}}) - \nabla [\chi^e_p]_{\mathcal{P}^e}(\mathbf{a}, \widetilde{\mathbf{a}}) \right) \right)(0) \, d\mathbb{P}(\widetilde{\mathbf{a}}).$$

All the random variables in the proof of this section are considered as random variables on the enlarged probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ unless explicitly stated.

We denote by $\mathbb{E}'$ the expectation with respect to the measure $\mathbb{P}'$. Given a constant $C > 0$, an exponent $s > 0$ and a nonnegative random variable $Z : \Omega' \mapsto \mathbb{R}$, we write

$$Z \le \mathcal{O}_s'(C) \quad \text{if and only if} \quad \mathbb{E}'\left[\exp\left(\left(\frac{Z}{C}\right)^s\right)\right] \le 2.$$

Any random variable Z defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying $Z \le \mathcal{O}_s(C)$ satisfies, as a random variable defined on the extended space $(\Omega', \mathcal{F}', \mathbb{P}')$, the inequality $Z \le \mathcal{O}_s'(C)$. From the definition (53), we see that, for each bond $e \in \mathcal{B}_d$,

(54) $$Z \le \mathcal{O}_s(C) \quad \Longrightarrow \quad Z^e \le \mathcal{O}_s'(C).$$

The estimate (54) is frequently used when the random variable Z is equal to the size of a cube of the partition $\mathcal{P}$ (Proposition 1), the minimal scale $\mathcal{X}$ above which the regularity theory applies (Proposition 6), or the minimal scale $\mathcal{M}_t(\mathcal{P})$ associated to the partition $\mathcal{P}$ (Proposition 1): we have, for each point $x \in \mathbb{Z}^d$ and each bond $e \in \mathcal{B}_d$,

$$\begin{cases} \text{size}(\square_{\mathcal{P}^e}(x)) = (\text{size}(\square_{\mathcal{P}}(x)))^e \le \mathcal{O}_s'(C), \\ \mathcal{X}^e \le \mathcal{O}_s'(C), \\ \mathcal{M}_t(\mathcal{P}^e) = (\mathcal{M}_t(\mathcal{P}))^e \le \mathcal{O}_s'(C). \end{cases}$$

To prove Lemma 3.2, we prove the estimate

(55) $$\sum_{e \in \mathcal{B}_d} |(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)|^2 \le \mathcal{O}_s'(Cr^{-d}).$$

The inequality (55) is sufficient to prove Result 2; indeed with the same argument as in [10], Lemma 2.3, we have

$$\mathbb{E}\left[\exp\left(\left(\frac{\sum_{e \in \mathcal{B}_d}(X - X_e')^2}{Cr^{-d}}\right)^s\right)\right]$$

$$= \int_\Omega \exp\left(\left(\frac{\sum_{e \in \mathcal{B}_d}|\int_\Omega (\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)\,d\mathbb{P}(\tilde{\mathbf{a}})|^2}{Cr^{-d}}\right)^s\right) d\mathbb{P}(\mathbf{a})$$

$$\le \int_\Omega \exp\left(\left(\int_\Omega \frac{\sum_{e \in \mathcal{B}_d}|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)|^2}{Cr^{-d}}\,d\mathbb{P}(\tilde{\mathbf{a}})\right)^s\right) d\mathbb{P}(\mathbf{a})$$

$$\le C \int_\Omega \int_\Omega \exp\left(\left(\frac{\sum_{e \in \mathcal{B}_d}|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)|^2}{Cr^{-d}}\right)^s\right) d\mathbb{P}(\mathbf{a})\,d\mathbb{P}(\tilde{\mathbf{a}})$$

$$\le C\mathbb{E}'\left[\exp\left(\left(\frac{\sum_{e \in \mathcal{B}_d}|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)|^2}{Cr^{-d}}\right)^s\right)\right]$$

$$\le 2C.$$

This yields, after redefinition of the constant C,

$$\sum_{e \in \mathcal{B}_d}(X - X_e')^2 \le \mathcal{O}_s'(Cr^{-d}).$$

Before starting the proof of (55), we select one of the correctors χ_p^e arbitrarily (we recall that they are defined up to a constant). As we are interested in the gradient of the corrector, the value of the constant is not important. We want to give a meaning to the function $[\chi_p^e]_{\mathcal{P}}$ as a random variable defined on the extended probability space Ω'.

Since we do not necessarily have the identity $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, we cannot simply write $[\chi_p^e]_\mathcal{P}(z) = \chi_p^e(z(\square_\mathcal{P}(z)))$. Nevertheless, since the two environments $(\mathbf{a}(e'))_{e' \in \mathcal{B}_d}$ and $((\mathbf{a}(e'))_{e' \in \mathcal{B}_d \setminus \{e\}}, \tilde{\mathbf{a}}(e))$ only differ by one bond, we have either $\mathscr{C}_\infty \subseteq \mathscr{C}_\infty^e$ or $\mathscr{C}_\infty^e \subseteq \mathscr{C}_\infty$. In the former case, we can define $[\chi_p^e]_\mathcal{P}(z) = \chi_p^e(z(\square_\mathcal{P}(z)))$. In the latter case, the cluster $\mathscr{C}_\infty \setminus \mathscr{C}_\infty^e$ is connected to $\mathscr{C}_\infty$ by the bond e. Without loss of generality, we denote it by $e = \{x, y\}$ and assume that $x \in \mathscr{C}_\infty^e$. One can then check that the function

$$(56) \qquad w := \begin{cases} \mathscr{C}_\infty \to \mathbb{R} \\ z \mapsto \chi_p^e(z)\mathbb{1}_{\{z \in \mathscr{C}_\infty^e\}} + (p \cdot (z - x) + \chi_p^e(z))\mathbb{1}_{\{z \notin \mathscr{C}_\infty^e\}} \end{cases}$$

is a solution of the equation

$$-\nabla \cdot \mathbf{a} \nabla(p \cdot x + w) = 0 \quad \text{in } \mathscr{C}_\infty$$

and more precisely that the map $x \mapsto p \cdot x + w(x)$ belongs to the space $\mathcal{A}_1(\mathscr{C}_\infty)$. In particular, this gives the identity $w = \chi_p$. We thus define

$$[\chi_p^e]_\mathcal{P} = [w]_\mathcal{P}.$$

To prove the estimate (55), we use the random variable $[\chi_p^e]_\mathcal{P}$ and split the sum into two terms

$$
\left|\left(\Phi_{r^2} * (\nabla[\chi_p]_\mathcal{P} - \nabla[\chi_p^e]_{\mathcal{P}^e})\right)(0)\right|^2
$$

$$(57) \qquad \leq \underbrace{2\left|\left(\Phi_{r^2} * (\nabla[\chi_p^e]_\mathcal{P} - \nabla[\chi_p^e]_{\mathcal{P}^e})(0)\right)\right|^2}_{(57)(\mathrm{i})} + \underbrace{2\left|\left(\Phi_{r^2} * (\nabla[\chi_p]_\mathcal{P} - \nabla[\chi_p^e]_\mathcal{P})\right)(0)\right|^2}_{(57)(\mathrm{ii})}.$$

We estimate the two terms in the right side in the two steps below.

Step 1. Estimate of the term (57)(i). We use Lemma 3.3 and Proposition 2 with the exponent $s = 1$ to write

$$
\left|\left(\Phi_{r^2} * (\nabla[\chi_p^e]_\mathcal{P} - \nabla[\chi_p^e]_{\mathcal{P}^e})\right)(0)\right|^2
$$

$$
\leq \left(\sum_{z \in \mathbb{Z}^d \cap B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} |\nabla[\chi_p^e]_\mathcal{P}(z)| + |\nabla[\chi_p^e]_{\mathcal{P}^e}(z)|\right)^2
$$

$$
\times \sup_{z \in B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} \Phi_{r^2}^2(z)
$$

$$
\leq C \left(\sum_{z \in \mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} \mathrm{size}(\square_\mathcal{P}(x))^{d-1}\left(|\nabla\chi_p^e|(z) + 1\right)\right)^2
$$

$$
\times \sup_{z \in B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} \Phi_{r^2}^2(z).
$$

The term "$+1$" on the right-hand side comes from the assumption $|p| = 1$ combined with the definition of the map w stated in (56) (in the case $\mathscr{C}_\infty^e \subseteq \mathscr{C}_\infty$). We deduce that

$$(58) \qquad \left|\left(\Phi_{r^2} * (\nabla[\chi_p^e]_\mathcal{P} - \nabla[\chi_p^e]_{\mathcal{P}^e})\right)(0)\right|^2
$$

$$
\leq C\,\mathrm{size}(\square_\mathcal{P}(x))^{3d-2} \sup_{z \in \mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} \left(|\nabla\chi_p^e|^2(z) + 1\right)
$$

$$
\times \sup_{z \in B(x, C\,\mathrm{size}(\square_\mathcal{P}(x)))} \Phi_{r^2}^2(z).
$$

By the heat kernel bound (41) stated in Definition 5, there exists a constant $C(d) < \infty$ such that, for each point $z \in \mathbb{Z}^d$,

$$(59) \qquad \Phi_{r^2} \leq \frac{C}{r^d}\left(\left(\frac{|z|}{r}\right)^{-\frac{d+1}{2}} \wedge 1\right).$$

We denote by $\zeta(z) := (|z|^{-\frac{d+1}{2}} \wedge 1)$ and by $\zeta_r(z) := \frac{1}{r^d}\zeta(\frac{z}{r})$. We use the function ζ instead of the heat kernel Φ_{r^2} to complete the estimate of the term (57)-(i) because it satisfies the inequality

$$\sup_{z \in B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))} \zeta_r^2(z) \leq C\,\mathrm{size}(\square_{\mathcal{P}}(x))^{d+1} \inf_{z \in B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))} \zeta_r^2(z).$$

In particular, the estimate (58) can be rewritten

$$\left|(\Phi_{r^2} * (\nabla[\chi_p^e]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)\right|^2$$
$$\leq C\,\mathrm{size}(\square_{\mathcal{P}}(x))^{4d-1} \sum_{z \in \mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))} \zeta_r(z)^2(|\nabla\chi_p^e|^2(z) + 1).$$

Summing over all the bonds $e \in \mathcal{B}_d$ gives

$$\sum_{e \in \mathcal{B}_d} \left|(\Phi_{r^2} * (\nabla[\chi_p^e]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)\right|^2$$

$$(60) \qquad \begin{aligned} &\leq C \sum_{x \in \mathbb{Z}^d} \mathrm{size}(\square_{\mathcal{P}}(x))^{4d-1} \\ &\quad \times \left(\sum_{z \in \mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))} \zeta_r(z)^2(|\nabla\chi_p^e|^2(z) + 1)\right) \\ &\leq C \sum_{z \in \mathscr{C}_\infty^e} \zeta_r(z)^2(|\nabla\chi_p^e|^2(z) + 1) \\ &\quad \times \left(\sum_{x \in \mathbb{Z}^d} \mathrm{size}(\square_{\mathcal{P}}(x))^{4d-1}\mathbb{1}_{\{z \in B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))\}}\right). \end{aligned}$$

Using the estimate $\mathrm{size}(\square_{\mathcal{P}}(x)) \leq \mathcal{O}_s'(C)$, valid for any point $x \in \mathbb{Z}^d$, we obtain

$$\mathbb{1}_{\{z \in B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))\}} \leq C\frac{\mathrm{size}(\square_{\mathcal{P}}(x))^{d+1}}{(|x - z| \vee 1)^{d+1}} \leq \frac{\mathcal{O}_s'(C)}{(|x - z| \vee 1)^{d+1}}.$$

Since the map $z \mapsto (|z| \vee 1)^{-d-1}$ is summable on $\mathbb{Z}^d$, we use the inequality (12) to obtain

$$(61) \qquad \sum_{x \in \mathbb{Z}^d} \mathrm{size}(\square_{\mathcal{P}}(x))^{4d-1}\mathbb{1}_{\{z \in B(x, C\,\mathrm{size}(\square_{\mathcal{P}}(x)))\}} \leq \mathcal{O}_s'(C).$$

By Proposition 37 and the implication (54), we have the Lipschitz bound on the corrector

$$(62) \qquad |\nabla\chi_p^e(y)|\mathbb{1}_{\{y \in \mathscr{C}_\infty^e\}} \leq \mathcal{O}_s'(C),$$

which implies

$$\sum_{e \in \mathcal{B}_d} \left|(\Phi_{r^2} * (\nabla[\chi_p^e]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)\right|^2 \leq C \sum_{y \in \mathbb{Z}^d} \zeta_r(y)^2 \mathcal{O}_s'(C).$$

We use the estimate (12) and the inequality $\sum_{y \in \mathbb{Z}^d} \zeta_r(y)^2 \leq Cr^{-d}$ to obtain

$$\sum_{e \in \mathcal{B}_d} \left|(\Phi_{r^2} * (\nabla[\chi_p^e]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}^e}))(0)\right|^2 \leq \mathcal{O}_s'\left(\frac{C}{r^d}\right).$$

This completes the proof of the estimate of the first term on the right-hand side of (57).

Step 2. Estimate of the term (57)(ii). In this step, we prove the inequality

$$\left|\left(\Phi_{r^2} * \left(\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}\right)\right)(0)\right|^2 \leq \mathcal{O}_s'\left(\frac{C}{r^d}\right).$$

To prove this estimate, we distinguish three cases. We recall the two endpoints of the bond e are x and y; they are fixed through the proof.

Case 1. ($x \notin \mathscr{C}_\infty$ and $y \notin \mathscr{C}_\infty$) or $\mathbf{a} = \mathbf{a}^e$. In that case, one has the identities $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ and $\nabla \chi_p = \nabla \chi_p^e$. They imply

$$\left|\left(\Phi_{r^2} * \left(\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}\right)\right)(0)\right|^2 = 0.$$

Case 2. $\mathscr{C}_\infty \neq \mathscr{C}_\infty^e$. In that case, we have $[\chi_p]_{\mathcal{P}} = [\chi_p^e]_{\mathcal{P}}$ by definition of the latter quantity. It implies

$$\left|\left(\Phi_{r^2} * \left(\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}\right)\right)(0)\right|^2 = 0.$$

Case 3. $x, y \in \mathscr{C}_\infty$ and $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ and $\mathbf{a} \neq \mathbf{a}^e$. By definitions of the correctors χ_p and χ_p^e, we have the identity

$$(63) \quad -\nabla \cdot \left(\mathbf{a}\nabla(\chi_p - \chi_p^e)\right) = (\mathbf{a} - \mathbf{a}^e)(\{x, y\})\left(p \cdot (x - y) + \chi_p^e(x) - \chi_p^e(y)\right)(\delta_x - \delta_y) \quad \text{in } \mathscr{C}_\infty.$$

We solve (63) by using Proposition 4 and recall the notation $\nabla G(e, \cdot)$ introduced there. Note that the function

$$\chi_p - \chi_p^e - (\mathbf{a} - \mathbf{a}^e)(x, y)\left(p \cdot (x - y) + \chi_p^e(x) - \chi_p^e(y)\right)\nabla G(e, \cdot) \text{ is } \mathbf{a}\text{-harmonic.}$$

By the sublinear growth of the corrector stated in (18), the $L^2(\mathscr{C}_\infty)$-bound stated in (30) on the gradient of the function $\nabla G(e, \cdot)$ and a version of the Poincaré inequality on the percolation cluster (see for instance the proof of Proposition 6), one can show that the function

$$\chi_p - \chi_p^e - (\mathbf{a} - \mathbf{a}^e)(x, y)\left(p \cdot (x - y) + \chi_p^e(x) - \chi_p^e(y)\right)\nabla G(e, \cdot) \text{ has a sublinear growth.}$$

This implies that this function is constant. In particular, it proves the identity

$$(64) \qquad \nabla \chi_p - \nabla \chi_p^e = (\mathbf{a} - \mathbf{a}^e)(x, y)\left(p \cdot (x - y) + \chi_p^e(x) - \chi_p^e(y)\right)\nabla\nabla G(e, \cdot).$$

We now estimate the right side of (64) thanks to the Lipschitz bound on the corrector stated in Proposition 37 and (62). We distinguish two cases depending on the value of the conductance $\mathbf{a}^e(e)$:

- If $\mathbf{a}^e(e) = \tilde{\mathbf{a}}(e) \neq 0$, then (62) implies the estimate

$$(65) \qquad \left|\chi_p^e(x) - \chi_p^e(y)\right|\mathbb{1}_{\{y \in \mathscr{C}_\infty^e, \tilde{\mathbf{a}}(e) \neq 0\}} \leq \left|\nabla\chi_p^e(y)\right|\mathbb{1}_{\{y \in \mathscr{C}_\infty^e\}} \leq C\left(\mathcal{X}^e(x)\right)^{d/2}.$$

- If $\mathbf{a}^e(e) = \tilde{\mathbf{a}}(e) = 0$, then there exists a path going from x to y which stays in the cube $\Box_{\mathcal{P}^e}(x)$ and its neighbors (its neighbors because we may not have $\Box_{\mathcal{P}^e}(x) = \Box_{\mathcal{P}^e}(y)$ or we may have $x, y \in \mathscr{C}_\infty \setminus \mathscr{C}_*(\Box_{\mathcal{P}^e}(x))$). Combining this remark with Lemma 3.3, we obtain

$$\left|\chi_p^e(x) - \chi_p^e(y)\right| \leq C \sum_{z \in \mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\Box_{\mathcal{P}}(x)))} \left|\nabla\chi_p^e\right|(z)$$

$$\leq C\,\mathrm{size}\left(\Box_{\mathcal{P}}(x)\right)^d \left\|\nabla\chi_p^e\right\|_{\underline{L}^2(\mathscr{C}_\infty^e \cap B(x, C\,\mathrm{size}(\Box_{\mathcal{P}}(x))))}.$$

Using the Lipschitz bounds on the corrector, we deduce that

$$(66) \qquad \left|\chi_p^e(x) - \chi_p^e(y)\right|\mathbb{1}_{\{x, y \in \mathscr{C}_\infty^e, \tilde{\mathbf{a}}(e) = 0\}} \leq \mathrm{size}\left(\Box_{\mathcal{P}}(x)\right)^d \left(\mathcal{X}^e(x)\right)^{d/2}.$$

Combining the estimates (65) and (66), we obtain the inequality

$$
\begin{aligned}
(67) \qquad & |\nabla(\Phi_{r^2} * ([\chi_p]_{\mathcal{P}} - [\chi_p^e]_{\mathcal{P}}))(0)|^2 \\
& \leq |\nabla(\Phi_{r^2} * ([\nabla G(e, \cdot)]_{\mathcal{P}}))(0)|^2 \operatorname{size}(\Box_{\mathcal{P}}(x))^{2d}(\mathcal{X}^e(x))^d.
\end{aligned}
$$

In the next step of the proof, we treat the coarsening in the right-hand side of (67). To this end, we prove that there exist a constant $C := C(d) < \infty$ and a (random) vector field $\gamma_r : E_d \to \mathbb{R}^d$ satisfying the estimate, for each edge $e' = (x', y') \in E_d$,

$$
|\gamma_r(e')| \leq C \operatorname{size}(\Box_{\mathcal{P}}(x'))^{2d} \zeta_r(x')
$$

such that for each function $u : \mathscr{C}_\infty \mapsto \mathbb{R}$ satisfying $\langle \nabla u, \nabla u \rangle_{\mathscr{C}_\infty} < \infty$,

$$
(68) \qquad (\Phi_{r^2} * \nabla[u]_{\mathcal{P}})(0) = \langle \gamma_r, \nabla u \rangle_{\mathscr{C}_\infty}.
$$

We first write

$$
(\Phi_{r^2} * \nabla[u]_{\mathcal{P}})(0) = \sum_{z \in \mathbb{Z}^d} \Phi_{r^2}(z) \nabla[u]_{\mathcal{P}}(z).
$$

Given two neighboring points $z, z' \in \mathbb{Z}^d$, note that the values $[u]_{\mathcal{P}}(z)$ and $[u]_{\mathcal{P}}(z')$ are only different if the points z and z' belong to two different cubes of the partition $\mathcal{P}$. In that case, we have

$$
[u]_{\mathcal{P}}(z) - [u]_{\mathcal{P}}(z') = u(\overline{z}(\Box_{\mathcal{P}}(z))) - u(\overline{z}(\Box_{\mathcal{P}}(z'))).
$$

Recall that there exists a path between $\overline{z}(\Box_{\mathcal{P}}(z))$ and $\overline{z}(\Box_{\mathcal{P}}(z'))$ which lies entirely in the set $\Box_{\mathcal{P}}(z) \cup \Box_{\mathcal{P}}(z')$. We denote this path by $p_{z,z'} \subseteq E_d$. Summing over the edges along this path, we find that

$$
u(\overline{z}(\Box_{\mathcal{P}}(z))) - u(\overline{z}(\Box_{\mathcal{P}}(z'))) = \sum_{e' \in p_{z,z'}} \nabla u(e') = \sum_{e' \in E_d} \nabla u(e') \mathbb{1}_{\{e' \in p_{z,z'}\}}.
$$

If the points z and z' belong to the same cube of the partition $\mathcal{P}$, we keep the same notation with the convention $p_{z,z'} = \varnothing$. Consequently, we have for each pair of neighboring points $z, z' \in \mathbb{Z}^d$,

$$
[u]_{\mathcal{P}}(z) - [u]_{\mathcal{P}}(z') = \sum_{e' \in E_d} \nabla u(e') \mathbb{1}_{\{e' \in p_{z,z'}\}}.
$$

Using this formula, we can rewrite

$$
(\Phi_{r^2} * \nabla[u]_{\mathcal{P}})(0) = \langle \gamma_r, \nabla u \rangle_{\mathscr{C}_\infty},
$$

where γ_r is the vector-valued random field defined by the formula, for each edge $e' \in E_d$,

$$
\gamma_r(e') = \begin{pmatrix} \displaystyle\sum_{z \in \mathbb{Z}^d} \Phi_{r^2}(z) \mathbb{1}_{\{e' \in p_{z,z+\mathfrak{e}_1}\}} \\ \vdots \\ \displaystyle\sum_{z \in \mathbb{Z}^d} \Phi_{r^2}(z) \mathbb{1}_{\{e' \in p_{z,z+\mathfrak{e}_d}\}} \end{pmatrix}.
$$

For each pair of neighboring points $z, z' \in \mathbb{Z}^d$ such that $\Box_{\mathcal{P}}(z) \neq \Box_{\mathcal{P}}(z')$, the path between the points $\overline{z}(\Box_{\mathcal{P}}(z))$ and $\overline{z}(\Box_{\mathcal{P}}(z'))$ lies in the set $\Box_{\mathcal{P}}(z) \cup \Box_{\mathcal{P}}(z')$. In particular an edge $e' = (x', y')$ belongs to the path $p_{z,z'}$ only if $\operatorname{dist}(z, \partial \Box_{\mathcal{P}}(x')) \leq 1$. This argument implies the inequality

$$
(69) \qquad |\gamma_r(e')| \leq C \sum_{z : \operatorname{dist}(z, \partial \Box_{\mathcal{P}}(x')) \leq 1} \Phi_{r^2}(z).
$$

We use the inequality (59) on the discrete heat kernel and note that function ζ satisfies the estimate, for each triadic cube $\square \in \mathcal{T}$,

$$\sup_{\square} \zeta_r \leq C \operatorname{size}(\square)^{\frac{d+1}{2}} \inf_{\square} \zeta_r.$$

As a consequence of the two previous displays, we can rewrite the estimate (69)

$$
\begin{aligned}
|\gamma_r(e')| &\leq \sum_{z:\operatorname{dist}(z,\partial \square_{\mathcal{P}}(x')) \leq 1} \Phi_{r^2}(z) \\
&\leq C \operatorname{size}(\square_{\mathcal{P}}(x'))^{d-1} \sup_{z:\operatorname{dist}(z,\partial \square_{\mathcal{P}}(x')) \leq 1} \zeta_r \\
&\leq C \operatorname{size}(\square_{\mathcal{P}}(x'))^{2d} \zeta_r(x'),
\end{aligned}
$$

(70)

which is the desired inequality. The proof of (68) is complete.

Applying the property (68) to the function $u = \nabla G(e, \cdot)$, the inequality (67) becomes

$$\left|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}))(0)\right|^2 \leq \left|\langle \gamma_r, \nabla\nabla G(e, \cdot)\rangle_{\mathscr{C}_\infty}\right|^2 \operatorname{size}(\square_{\mathcal{P}}(x))^{2d}(\mathcal{X}^e(x))^d.$$

We apply Proposition 5 and denote by $w_{\gamma_r} : \mathscr{C}_\infty \to \mathbb{R}^d$ the solution of the equation

$$-\nabla \cdot \mathbf{a}\nabla w_{\gamma_r} = -\nabla \cdot \gamma_r \quad \text{in } \mathscr{C}_\infty,$$

so that, for each edge e' in the infinite cluster,

$$
\begin{aligned}
\nabla w_{\gamma_r}(e') &= \sum_{e'' \subseteq \mathscr{C}_\infty} \gamma_r(e'') \nabla\nabla G(e'', e') = \sum_{e'' \subseteq \mathscr{C}_\infty} \gamma_r(e'') \nabla\nabla G(e', e'') \\
&= \langle \gamma_r, \nabla\nabla G(e', \cdot)\rangle_{\mathscr{C}_\infty}.
\end{aligned}
$$

This implies the identity

$$\left|w_{\gamma_r}(x) - w_{\gamma_r}(y)\right| = \left|\langle \gamma_r, \nabla\nabla G(e, \cdot)\rangle_{\mathscr{C}_\infty}\right|,$$

and consequently

$$\left|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}))(0)\right|^2 \leq \left|w_{\gamma_r}(x) - w_{\gamma_r}(y)\right|^2 \operatorname{size}(\square_{\mathcal{P}}(x))^{2d}(\mathcal{X}^e(x))^d.$$

We now combine Cases 1, 2 and 3 to obtain the estimate

$$
\begin{aligned}
&\sum_{e \in \mathcal{B}_d} \left|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - [\chi_p^e]_{\mathcal{P}}))(0)\right|^2 \\
&\leq C \sum_{x,y \in \mathscr{C}_\infty, |x-y|_1 = 1} \left|w_{\gamma_r}(x) - w_{\gamma_r}(y)\right|^2 \operatorname{size}(\square_{\mathcal{P}}(x))^{2d} \sum_{e \in \mathcal{B}_d^x} (\mathcal{X}^e(x))^d,
\end{aligned}
$$

where we used the notation $\mathcal{B}_d^x := \{\{x, y\} : y \in \mathbb{Z}^d, y \sim x\}$ to denote the set of bonds connecting the point x to another vertex of $\mathbb{Z}^d$.

Using that for each pair of points $x, y \in \mathscr{C}_\infty$ with $|x - y|_1 = 1$, there exists a path which is contained in the infinite cluster $\mathscr{C}_\infty$, the cube $\square_{\mathcal{P}}(x)$ and its neighbors (the path is simply (x, y) if $\mathbf{a}(\{x, y\}) \neq 0$), we obtain

$$
\begin{aligned}
&\sum_{e \in \mathcal{B}_d} \left|(\Phi_{r^2} * (\nabla[\chi_p]_{\mathcal{P}} - \nabla[\chi_p^e]_{\mathcal{P}}))(0)\right|^2 \\
&\leq C \sum_{z \in \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(z) \operatorname{size}(\square_{\mathcal{P}}(z))^{3d} \left(\sum_{x \in \mathbb{Z}^d, \operatorname{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (\mathcal{X}^e(x))^d \right).
\end{aligned}
$$

(71)

To estimate the term on the right-hand side, we first note that, by definition of the function w_{γ_r} and the inequality (12),

$$\sum_{z \in \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(z) \leq C \sum_{z \in \mathscr{C}_\infty} |\gamma_r(z)|^2 \leq C \sum_{z \in \mathbb{Z}^d} C \operatorname{size}(\square_{\mathcal{P}}(z))^{4d} \zeta_r(z)^2 \leq \mathcal{O}'_s(Cr^{-d}).$$

To complete the proof, we use the two lemmas stated below; their proofs are postponed to Appendices B and C.

PROPOSITION 11 (Meyers estimate).   *There exist a constant $C := C(d, \lambda, \mathfrak{p}) < \infty$, two exponents $s := s(d, \lambda, \mathfrak{p}) > 0$ and $\varepsilon := \varepsilon(d, \lambda, \mathfrak{p}) > 0$ and a random variable $\mathcal{M}_{\text{Meyers}} \leq \mathcal{O}_s(C)$ such that for each integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}_{\text{Meyers}}$, each vector field $\xi : E_d \to \mathbb{R}$ satisfying*

$$\xi(x, y) = 0 \quad \text{if } \mathbf{a}(x, y) = 0 \text{ or } x, y \notin \mathscr{C}_\infty,$$

and each function $v : \mathscr{C}_\infty \mapsto \mathbb{R}$ solution of the equation

$$-\nabla \cdot (\mathbf{a} \nabla v) = -\nabla \cdot \xi \quad \text{in } \mathscr{C}_\infty,$$

one has the estimate

(72)
$$\left(\frac{1}{|\square_m|} \sum_{x \in \square_m \cap \mathscr{C}_\infty} |\nabla v|^{2+\varepsilon}(x) \right)^{\frac{1}{2+\varepsilon}}$$
$$\leq C \left(\frac{1}{|\frac{4}{3}\square_m|} \sum_{x \in \frac{4}{3}\square_m \cap \mathscr{C}_\infty} |\nabla v|^2(x) \right)^{\frac{1}{2}} + C \left(\frac{1}{|\frac{4}{3}\square_m|} \sum_{x \in \frac{4}{3}\square_m \cap \mathscr{C}_\infty} |\xi|^{2+\varepsilon}(x) \right)^{\frac{1}{2+\varepsilon}}.$$

LEMMA 3.4 (Minimal scale).   *There exist a constant $C := C(d, \mathfrak{p}, \lambda) < \infty$, an exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and a random variable $\mathcal{M}_1 \leq \mathcal{O}'_s(C)$ such that for each integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}_1$,*

(73)
$$3^{-dm} \sum_{z \in \square_m} \operatorname{size}(\square_{\mathcal{P}}(z))^{\frac{3d(2+\varepsilon)}{\varepsilon}} \left(\sum_{x \in \mathbb{Z}^d, \operatorname{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}^x_d} (\mathcal{X}^e(x))^d \right)^{\frac{2+\varepsilon}{\varepsilon}} \leq C,$$

where $\varepsilon := \varepsilon(d, \mathfrak{p}, \lambda) > 0$ is the exponent which appears in Proposition 11.

DEFINITION 6 (The partition $\mathcal{U}$).   We define the following family of "good cubes"

$$\mathcal{G} := \{\square \in \mathcal{T} : (72) \text{ and } (73) \text{ hold}\}$$

in which a deterministic Meyers estimate and a minimal scale inequality hold. By Lemma 3.4 and Proposition 11, this collection satisfies the assumption of Proposition 1 (but not the assumption (20)). We denote by $\mathcal{U}$ the partition thus obtained. By the Property (iii) of Proposition 1, one has the inequality

$$\operatorname{size}(\square_{\mathcal{U}}(x)) \leq \mathcal{O}_s(C)$$

for some exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and some constant $C := C(d, \mathfrak{p}, \lambda) < \infty$.

Using the properties of the partition $\mathcal{U}$ and Hölder inequality, one obtains

$$\sum_{z \in \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(z) \operatorname{size}(\square_{\mathcal{P}}(z))^{3d} \left(\sum_{x \in \mathbb{Z}^d, \operatorname{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^d \right)$$

$$= \sum_{\square \in \mathcal{U}} \sum_{z \in \square \cap \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(z)$$

$$\times \operatorname{size}(\square_{\mathcal{P}}(z))^{3d} \left(\sum_{x \in \mathbb{Z}^d, \operatorname{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^d \right)$$

$$\leq \sum_{\square \in \mathcal{U}} |\square| \left(\frac{1}{|\square|} \sum_{z \in \square \cap \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^{2+\varepsilon}(z) \right)^{\frac{2}{2+\varepsilon}}$$

$$\times \left(\frac{1}{|\square|} \sum_{z \in \square} (\operatorname{size}(\square_{\mathcal{P}}(z)))^{\frac{3d(2+\varepsilon)}{\varepsilon}} \left(\sum_{x \in \mathbb{Z}^d, \operatorname{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^d \right)^{\frac{2+\varepsilon}{\varepsilon}} \right)^{\frac{\varepsilon}{2+\varepsilon}}$$

$$\leq C \sum_{\square \in \mathcal{U}} \left(\sum_{x \in \frac{4}{3}\square \cap \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(x) + |\square| \left(\frac{1}{|\frac{4}{3}\square|} \sum_{\frac{4}{3}\square \cap \mathscr{C}_\infty} |\gamma_r|^{2+\varepsilon}(x) \right)^{\frac{2}{2+\varepsilon}} \right).$$

To estimate the term on the right-hand side, we note that the cube $\frac{4}{3}\square$ is included in the set $\bigcup_{\square' \in \mathcal{U}, \operatorname{dist}(\square', \square) \leq 1} \square'$ and the cardinality of the set $\{\square' \in \mathcal{U} : \operatorname{dist}(\square', \square) \leq 1\}$ is bounded by a constant depending only on the dimension d. This leads to the upper bound

$$\sum_{\square \in \mathcal{U}} \sum_{x \in \frac{4}{3}\square \cap \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(x) \leq C \sum_{x \in \mathscr{C}_\infty} |\nabla w_{\gamma_r}|^2(x) \leq \mathcal{O}_s'\left(\frac{C}{r^d} \right).$$

To estimate the second term on the right-hand side, we recall the discrete $l^1 - l^t$-estimate: for any finite sequence of nonnegative real numbers $(b_i)_{0 \leq i \leq n} \in \mathbb{R}_+^{n+1}$ and any exponent $t \geq 1$, $\sum_{i=0}^n b_i^t \leq (\sum_{i=0}^n b_i)^t$. Using this inequality, we obtain

$$\sum_{\square \in \mathcal{U}} |\square| \left(\frac{1}{|\frac{4}{3}\square|} \sum_{x \in \frac{4}{3}\square \cap \mathscr{C}_\infty} |\gamma_r|^{2+\varepsilon}(x) \right)^{\frac{2}{2+\varepsilon}}$$

$$\leq C \sum_{\square \in \mathcal{U}} |\square|^{1 - \frac{2}{2+\varepsilon}} \sum_{x \in \frac{4}{3}\square \cap \mathscr{C}_\infty} \gamma_r^2(x)$$

$$\leq C \sum_{x \in \mathscr{C}_\infty} |\gamma_r|^2(x) \operatorname{size}(\square_{\mathcal{U}}(x))^{d(1 - \frac{2}{2+\varepsilon})}$$

$$\leq C \sum_{x \in \mathbb{Z}^d} \zeta_r(x)^2 \operatorname{size}(\square_{\mathcal{P}}(x))^{4d} \operatorname{size}(\square_{\mathcal{U}}(x))^{d(1 - \frac{2}{2+\varepsilon})}.$$

Using the inequalities $\operatorname{size}(\square_{\mathcal{U}}(x)) \leq \mathcal{O}_s'(C)$, $\operatorname{size}(\square_{\mathcal{P}}(x)) \leq \mathcal{O}_s'(C)$ and (12), we obtain

$$\sum_{\square \in \mathcal{U}} |\square| \left(\frac{1}{|\frac{4}{3}\square|} \sum_{x \in \frac{4}{3}\square \cap \mathscr{C}_\infty} |\gamma_r|^{2+\varepsilon}(x) \right)^{\frac{2}{2+\varepsilon}} \leq \mathcal{O}_s'\left(\frac{C}{r^d} \right).$$

The proof of Result 2, and thus of Proposition 10, is complete. $\square$

4. Optimal L^q estimates for first order corrector. In this section, we show how to obtain L^q optimal scaling estimates on the corrector (Theorem 1.2) from Proposition 10. We first restate the result.

THEOREM 1.2 (Optimal L^q estimates for first order corrector). *There exist two exponents* $s := s(d, \mathfrak{p}, \lambda) > 0$, $k := k(d, \mathfrak{p}, \lambda) < \infty$ *and a constant* $C(d, \mathfrak{p}, \lambda) < \infty$ *such that for each radius* $R \geq 1$, *each exponent* $q \geq 1$ *and each* $p \in \mathbb{R}^d$,

$$
(74) \qquad \left(R^{-d} \sum_{x \in \mathscr{C}_\infty \cap B_R} |\chi_p(x) - (\chi_p)_{\mathscr{C}_\infty \cap B_R}|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(C|p|q^k \log^{\frac{1}{2}} R) & \text{if } d = 2, \\ \mathcal{O}_s(C|p|q^k) & \text{if } d \geq 3. \end{cases}
$$

Before starting the proof, we mention that, in this section, we need to keep track of the dependence on the parameter q of the constants as it will be useful in the next section to obtain the L^∞ bounds on the corrector.

PROOF OF THEOREM 1.2. As in the proof of Proposition 10, we assume that $|p| = 1$ to ease the notation. Additionally, note that, by the Jensen inequality, it is enough to prove Theorem 1.2 in the case $q \geq 2$. We consequently make this assumption for the rest of the proof. The proof of this theorem is split into two steps.

- In Step 1, we use Proposition 10 and the multiscale Poincaré inequality (Proposition 9) to show, for each radius $R \geq 1$,

$$
\left(R^{-d} \sum_{x \in B_R} |[\chi_p]_{\mathcal{P}}(x) - ([\chi_p]_{\mathcal{P}})_{B_R}|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(Cq^k \log^{\frac{1}{2}} R) & \text{if } d = 2, \\ \mathcal{O}_s(Cq^k) & \text{if } d \geq 3, \end{cases}
$$

  where the parameters C, k and s depend only on $s, \mathfrak{p}, \lambda$.
- In Step 2, we remove the coarsening, thanks to Proposition 2, to obtain

$$
\left(R^{-d} \sum_{x \in \mathscr{C}_\infty \cap B_R} |\chi_p(x) - (\chi_p)_{\mathscr{C}_\infty \cap B_R}|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(Cq^k \log^{\frac{1}{2}} R) & \text{if } d = 2, \\ \mathcal{O}_s(Cq^k) & \text{if } d \geq 3. \end{cases}
$$

Step 1. We apply Proposition 9 to the function $u = [\chi_p]_{\mathcal{P}}$ and obtain, for each $R \geq 1$,

$$
\|[\chi_p]_{\mathcal{P}} - ([\chi_p]_{\mathcal{P}})_{B_R}\|_{L^q(B_R)} \leq C \left(\sum_{x \in \mathbb{Z}^d} e^{-\frac{|x|}{2R}} \left(\int_0^{2R} r |\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)|^2 \, dr \right)^{\frac{q}{2}} \right)^{\frac{1}{q}}.
$$

To study the term on the right-hand side, we split the integral into two terms

$$
(75) \qquad \begin{aligned} \int_0^{2R} r |\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)|^2 \, dr &= \int_0^1 r |\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)|^2 \, dr \\ &\quad + \int_1^{2R} r |\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)|^2 \, dr. \end{aligned}
$$

By Proposition 10, we know that for each radius $r \geq 1$ and each point $x \in \mathbb{R}^d$,

$$
|\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)| \leq \mathcal{O}_s(Cr^{-\frac{d}{2}}).
$$

This implies

$$
|\Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x)|^2 \leq \mathcal{O}_s(Cr^{-d}).
$$

The second term on the right-hand side can be estimated by using Proposition 10 and the inequality (12). We obtain

$$\int_1^{2R} r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \leq \begin{cases} \mathcal{O}_s(C \log R) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}$$

To estimate the first term on the right-hand side of (75), we use Proposition 7 which reads, for each point $x \in \mathbb{R}^d$,

(76) $$\left| \nabla[\chi_p]_{\mathcal{P}}(x) \right| \leq \mathcal{O}_s(C).$$

Combining the inequality (76) with (12), we obtain

$$\int_0^1 r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \leq \mathcal{O}_s(C).$$

Combining the previous displays shows

$$\int_0^{2R} r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \leq \begin{cases} \mathcal{O}_s(C \log R) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}$$

We obtain

$$\left(\int_0^{2R} r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \right)^{\frac{q}{2}} \leq \begin{cases} \mathcal{O}_{\frac{2s}{q}}(C^{\frac{q}{2}} (\log R)^{\frac{q}{2}}) & \text{if } d = 2, \\ \mathcal{O}_{\frac{2s}{q}}(C^{\frac{q}{2}}) & \text{if } d \geq 3. \end{cases}$$

We apply (12) and keep track of the dependence of the constants in the exponent q thanks to (13). We obtain

$$\sum_{x \in \mathbb{Z}^d} R^{-d} e^{-\frac{|x|}{2R}} \left(\int_0^{2R} r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \right)^{\frac{q}{2}}$$

$$\leq \begin{cases} \mathcal{O}_{\frac{2s}{q}} \left(\left(\dfrac{q}{s \ln(2)} \right)^{\frac{q}{s}} C^{\frac{q}{2}} (\log R)^{\frac{q}{2}} \right) & \text{if } d = 2, \\ \mathcal{O}_{\frac{2s}{q}} \left(\left(\dfrac{q}{s \ln(2)} \right)^{\frac{q}{s}} C^{\frac{q}{2}} \right) & \text{if } d \geq 3. \end{cases}$$

This eventually yields

$$\left(\sum_{x \in \mathbb{Z}^d} R^{-d} e^{-\frac{|x|}{2R}} \left(\int_0^{2R} r \left| \Phi_{r^2} * \nabla[\chi_p]_{\mathcal{P}}(x) \right|^2 dr \right)^{\frac{q}{2}} \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(q^{\frac{1}{s}} C (\log R)^{\frac{1}{2}}) & \text{if } d = 2, \\ \mathcal{O}_s(q^{\frac{1}{s}} C) & \text{if } d \geq 3. \end{cases}$$

We now set $k := \frac{1}{s} + \frac{3}{2}$; this exponent depends only on the parameters $d, \mathfrak{p}, \lambda$. By applying Proposition 9, we obtain

(77) $$\left(R^{-d} \sum_{x \in \mathbb{Z}^d} \left| [\chi_p]_{\mathcal{P}}(x) - ([\chi_p]_{\mathcal{P}})_{B_R} \right|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(C q^k \log^{\frac{1}{2}} R) & \text{if } d = 2, \\ \mathcal{O}_s(C q^k) & \text{if } d \geq 3. \end{cases}$$

The proof of Step 1 is complete.

Step 2. In this step, we remove the coarsening from (77) thanks to Proposition 2. We write

$$
\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - (\chi_p)_{\mathscr{C}_\infty \cap \square_m} \right|^q \right)^{\frac{1}{q}}
$$

$$
\leq 2 \inf_{a \in \mathbb{R}} \left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - a \right|^q \right)^{\frac{1}{q}}
$$

(78)
$$
\leq 2\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - ([\chi_p]_\mathcal{P})_{\square_m} \right|^q \right)^{\frac{1}{q}}
$$

$$
\leq 2\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - [\chi_p]_\mathcal{P}(x) \right|^q \right)^{\frac{1}{q}}
$$

$$
+ 2\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| [\chi_p]_\mathcal{P}(x) - ([\chi_p]_\mathcal{P})_{\square_m} \right|^q \right)^{\frac{1}{q}}.
$$

To estimate the first term on the right-hand side, we first use the inclusion (23) and Proposition 2 to obtain

$$
\sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - [\chi_p]_\mathcal{P}(x) \right|^q \leq \sum_{x \in \mathscr{C}_*(\square_{m+1})} \left| \chi_p(x) - [\chi_p]_\mathcal{P}(x) \right|^q
$$

$$
\leq C \sum_{x \in \mathscr{C}_*(\square_{m+1})} \text{size}(\square_\mathcal{P}(x))^{qd} |\nabla \chi_p|^q(x)
$$

$$
\leq C \sum_{x \in \mathscr{C}_*(\square_{m+1})} \text{size}(\square_\mathcal{P}(x))^{qd} |\nabla \chi_p|^q(x).
$$

By the Lipschitz bounds on the gradient of the corrector and the property of the partition $\mathcal{P}$, we have, for each point $x \in \mathbb{Z}^d$,

$$
\text{size}(\square_\mathcal{P}(x))^{qd} |\nabla \chi_p|^q(x) \mathbb{1}_{\{x \in \mathscr{C}_\infty\}} \leq \mathcal{O}_{\frac{s}{q}}(C^q).
$$

Additionally, for each point $x \in \mathbb{Z}^d \setminus \square_{m+1}$,

$$
\mathbb{1}_{\{x \in \mathscr{C}_*(\square_{m+1})\}} \leq \frac{\text{size}(\square_\mathcal{P}(x))^{d+1}}{|x|^{d+1} \vee 1} \leq \mathcal{O}_s\left(\frac{C}{(|x| \vee 1)^{d+1}}\right).
$$

We then use the inequalities (12) and (13) to obtain

$$
\sum_{x \in \mathscr{C}_*(\square_{m+1})} \text{size}(\square_\mathcal{P}(x))^{qd} |\nabla \chi_p|^q(x) = \sum_{x \in \mathbb{Z}^d} \text{size}(\square_\mathcal{P}(x))^{qd} |\nabla \chi_p|^q(x) \mathbb{1}_{\{x \in \mathscr{C}_*(\square_{m+1})\}}
$$

$$
\leq \mathcal{O}_{\frac{s}{q}}\left(3^{dm}\left(\frac{q}{s \ln(2)}\right)^{\frac{q}{s}} C^q\right)
$$

$$
\leq \mathcal{O}_{\frac{s}{q}}(3^{dm} q^{\frac{q}{s}} C^q).
$$

By (24), we obtain

(79)
$$
\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - [\chi_p]_\mathcal{P}(x) \right|^q \right)^{\frac{1}{q}} \leq \mathcal{O}_s(q^{\frac{1}{s}} C).
$$

To estimate the second term on the right-hand side of (78), we note that

$$
\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| [\chi_p]_\mathcal{P}(x) - ([\chi_p]_\mathcal{P})_{\square_m} \right|^q \right)^{\frac{1}{q}}
$$

(80)

$$
\leq \left(3^{-dm} \sum_{x \in \square_m} \left| [\chi_p]_\mathcal{P}(x) - ([\chi_p]_\mathcal{P})_{\square_m} \right|^q \right)^{\frac{1}{q}}.
$$

We combine (79), (80) and apply (77) to obtain

$$
\left(3^{-dm} \sum_{x \in \mathscr{C}_\infty \cap \square_m} \left| \chi_p(x) - (\chi_p)_{\mathscr{C}_\infty \cap \square_m} \right|^q \right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(q^k C m^{\frac{1}{2}}) & \text{if } d = 2, \\ \mathcal{O}_s(q^k C) & \text{if } d \geq 3, \end{cases}
$$

for some exponents $k := k(d, \mathfrak{p}, \lambda)$, $s := s(d, \mathfrak{p}, \lambda) > 0$ and some constant $C := C(d, \mathfrak{p}, \lambda) < \infty$. The result of Theorem 1.2 requires to prove the previous inequality for a general ball B_R and not a triadic cube. This result is obtained by selecting, for each radius $R \geq 1$, the integer m such that $3^m < R \leq 3^{m+1}$ and by performing a similar analysis. $\quad\square$

5. Optimal L^∞ estimates for the first order corrector. In this section, we prove the L^∞ bound on the corrector, Theorem 1.1.

THEOREM 1.1 (Optimal L^∞ estimates for first order correctors). *There exist an exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and a constant $C := C(d, \mathfrak{p}, \lambda) < \infty$ such that for each $x, y \in \mathbb{Z}^d$ and each $p \in \mathbb{R}^d$,*

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

PROOF. First by the stationarity of the gradient of the corrector, we can assume that $y = 0$. Without loss of generality, we can also assume $|p| = 1$. We want to prove, for each point $x \in \mathbb{Z}^d$,

$$
|\chi_p(x) - \chi_p(0)| \mathbb{1}_{\{0, x \in \mathscr{C}_\infty\}} \leq \begin{cases} \mathcal{O}_s(C \log^{\frac{1}{2}} |x|) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}
$$

We record that, for every exponent $q > 0$ and every point $x \in \mathbb{R}_+$

$$
\exp(x) \geq \frac{x^q}{q^q \exp(-q)}.
$$

This implies, for each triplet $s, q, \theta > 0$,

(81)
$$
X \leq \mathcal{O}_s(\theta) \quad \Longrightarrow \quad \mathbb{E}[X^q] \leq 2\theta^q \left(\frac{q}{s} \right)^{\frac{q}{s}} \exp\left(\frac{q}{s} \right).
$$

We split the proof into six steps.

• In Step 1, we prove that for each exponent $q \geq 1$ and each integer $m \in \mathbb{N}$,

$$
\mathbb{E}\left[\left| [\chi_p]_\mathcal{P}(0) - 3^{-2dm} \sum_{y \in \square_m} \sum_{z \in y + \square_m} [\chi_p]_\mathcal{P}(z) \right|^q \right] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}
$$

- In Step 2, we use the result of Step 1 to prove that for each exponent $q \geq 1$ and each integer $m \in \mathbb{N}$,

$$\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}}(x) - 3^{-2dm} \sum_{y \in \square_m} \sum_{z \in x+y+\square_m} [\chi_p]_{\mathcal{P}}(z)\right|^q\right] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

Note that this statement is not just a consequence of Step 1 and the stationarity of the corrector since the partition $\mathcal{P}$ is not stationary.

- In Step 3, we prove that for each $q \geq 1$ and $m \in \mathbb{N}$, chosen such that $3^m \leq |x| < 3^{m+1}$,

$$\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}}(0) - 3^{-2dm} \sum_{y \in \square_m} \sum_{z \in x+y+\square_m} [\chi_p]_{\mathcal{P}}(z)\right|^q\right] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

- In Step 4, we combine Steps 2 and 3 to obtain, for each exponent $q \geq 1$,

$$(82) \qquad \mathbb{E}[|[\chi_p]_{\mathcal{P}}(x) - [\chi_p]_{\mathcal{P}}(0)|^q] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

- In Step 5, we prove that there exist an exponent $s := s(d, \mathfrak{p}, \lambda) > 0$ and a constant $C := C(d, \mathfrak{p}, \lambda) < \infty$ such that

$$(83) \qquad |[\chi_p]_{\mathcal{P}}(x) - [\chi_p]_{\mathcal{P}}(0)| \leq \begin{cases} \mathcal{O}_s(C \log^{\frac{1}{2}} |x|) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}$$

- In Step 6, we remove the coarsening and show that

$$|\chi_p(x) - \chi_p(0)| \mathbb{1}_{\{0, x \in \mathscr{C}_\infty\}} \leq \begin{cases} \mathcal{O}_s(C \log^{\frac{1}{2}} |x|) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}$$

Step 1. The main tool of this step is the following inequality which was proved in Step 1 of the proof of Theorem 1.2, for each $m \in \mathbb{N}$, and each $q \geq 1$,

$$(84) \qquad \left(3^{-dm} \sum_{y \in \square_m} |[\chi_p]_{\mathcal{P}}(y) - ([\chi_p]_{\mathcal{P}})_{\square_m}|^q\right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(Cq^k \sqrt{m}) & \text{if } d = 2, \\ \mathcal{O}_s(Cq^k) & \text{if } d \geq 3. \end{cases}$$

Note that this implies, by (81),

$$(85) \qquad \mathbb{E}\left[3^{-dm} \sum_{y \in \square_m} |[\chi_p]_{\mathcal{P}}(y) - ([\chi_p]_{\mathcal{P}})_{\square_m}|^q\right] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

For some fixed point $y \in \mathbb{Z}^d$, the stationarity of the corrector stated in (19) implies the identity, for almost every environment $\mathbf{a} \in \Omega$,

$$([\chi_p]_{\mathcal{P}}(-y) - ([\chi_p]_{\mathcal{P}})_{\square_m})(\mathbf{a}) = ([\chi_p]_{\mathcal{P}_y}(0) - ([\chi_p]_{\mathcal{P}_y})_{y+\square_m})(\tau_y \mathbf{a}),$$

where we recall the notation $\mathcal{P}_y = y + \mathcal{P}(\tau_{-y}\mathbf{a})$. Using the stationarity property (6), we obtain, for each exponent $q \geq 1$,

$$\mathbb{E}[|[\chi_p]_{\mathcal{P}_y}(0) - ([\chi_p]_{\mathcal{P}_y})_{y+\square_m}|^q] = \mathbb{E}[|[\chi_p]_{\mathcal{P}}(-y) - ([\chi_p]_{\mathcal{P}})_{\square_m}|^q].$$

Since this result applies for each point $y \in \mathbb{Z}^d$, we can sum over all the points in the cube $\square_m$. We obtain

$$3^{-dm} \sum_{y \in \square_m} \mathbb{E}[|[\chi_p]_{\mathcal{P}_y}(0) - ([\chi_p]_{\mathcal{P}_y})_{y+\square_m}|^q]$$

$$= 3^{-dm} \sum_{y \in \square_m} \mathbb{E}[|[\chi_p]_{\mathcal{P}}(-y) - ([\chi_p]_{\mathcal{P}})_{\square_m}|^q].$$

Thus, by (85),

(86) $$\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}\right|^q\right]\leq\begin{cases}C^q q^{qk}m^{\frac{q}{2}} & \text{if } d=2,\\ C^q q^{qk} & \text{if } d\geq 3.\end{cases}$$

We now remove the translation of the partition and prove, for each point $z\in\mathbb{Z}^d$

(87) $$\left|[\chi_p]_{\mathcal{P}_y}(z)-[\chi_p]_{\mathcal{P}}(z)\right|\leq\mathcal{O}_s(C).$$

To prove this inequality, note that, by definition of the coarsening (4), we have

$$[\chi_p]_{\mathcal{P}_y}(z)-[\chi_p]_{\mathcal{P}}(z)=\chi_p\big(\bar{z}(\square_{\mathcal{P}_y}(z))\big)-\chi_p\big(\bar{z}(\square_{\mathcal{P}}(z))\big),$$

and by definition of the two partitions $\mathcal{P}$ and $\mathcal{P}_y$, there exists a path connecting the cubes $\square_{\mathcal{P}_y}(z)$ and $\square_{\mathcal{P}}(z)$ which lies in the ball $B(z,C\max(\text{size}(\square_{\mathcal{P}_y}(z)),\text{size}(\square_{\mathcal{P}}(z))))$. To simplify the notation, let us denote $R'=C\max(\text{size}(\square_{\mathcal{P}_y}(z)),\text{size}(\square_{\mathcal{P}}(z)))$. We have the estimate

$$\left|[\chi_p]_{\mathcal{P}_y}(z)-[\chi_p]_{\mathcal{P}}(z)\right|\leq\sum_{x\in\mathscr{C}_\infty\cap B_{R'}(z)}|\nabla\chi_p|(x).$$

By Proposition 6, the bounds $R'\leq\mathcal{O}_s(C)$ and $\mathcal{X}(z)\leq\mathcal{O}_s(C)$ and the assumption $|p|=1$, we have

$$\sum_{y\in\mathscr{C}_\infty\cap B'_R(z)}|\nabla\chi_p|(y)\leq\mathcal{O}_s(C).$$

Combining the previous displays completes the proof of (87). To remove the parameter y in (86), we compute

$$\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right]$$

(88)
$$\leq 2^q\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}\right|^q\right]$$

$$+2^q\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}-[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right].$$

By (87) and (12) we have, for each point $y\in\square_m$,

$$\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}-[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|\leq\mathcal{O}_s(C),$$

and thus

$$\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}-[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right]\leq C^q q^{qk}.$$

Summing over the points $y\in\square_m$ yields

$$\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}_y}(0)-([\chi_p]_{\mathcal{P}_y})_{y+\square_m}-[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right]\leq C^q q^{qk}.$$

By the previous display and (86), we have

$$\mathbb{E}\left[3^{-dm}\sum_{y\in\square_m}\left|[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right]\leq\begin{cases}C^q q^{qk}m^{\frac{q}{2}} & \text{if } d=2,\\ C^q q^{qk} & \text{if } d\geq 3.\end{cases}$$

By the Jensen inequality, we obtain

$$\mathbb{E}\left[\left|3^{-dm}\sum_{y\in\square_m}[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}\right|^q\right]\leq\begin{cases}C^q q^{qk}m^{\frac{q}{2}}&\text{if }d=2,\\C^q q^{qk}&\text{if }d\geq 3.\end{cases}$$

Notice that

$$3^{-dm}\sum_{y\in\square_m}[\chi_p]_{\mathcal{P}}(0)-([\chi_p]_{\mathcal{P}})_{y+\square_m}=[\chi_p]_{\mathcal{P}}(0)-3^{-2dm}\sum_{y\in\square_m}\sum_{z\in y+\square_m}[\chi_p]_{\mathcal{P}}(z).$$

Combining the two previous displays completes the proof of Step 1.

Step 2. By the stationarity of the corrector (19), for almost every environment $\mathbf{a}\in\Omega$, every pair of points $y, z\in\mathbb{Z}^d$,

$$[\chi_p]_{\mathcal{P}}(z)(\mathbf{a})=[\chi_p]_{\mathcal{P}_y}(z+y)(\tau_y\mathbf{a}).$$

Using this property, we have

$$\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}_x}(x)-3^{-2dm}\sum_{y\in\square_m}\sum_{z\in x+y+\square_m}[\chi_p]_{\mathcal{P}_x}(z)\right|^q\right]$$

$$=\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}}(0)-3^{-2dm}\sum_{y\in\square_m}\sum_{z\in y+\square_m}[\chi_p]_{\mathcal{P}}(z)\right|^q\right]$$

$$\leq\begin{cases}C^q q^{qk}m^{\frac{q}{2}}&\text{if }d=2,\\C^q q^{qk}&\text{if }d\geq 3.\end{cases}$$

Performing the same computation as in (88), we can replace the partition $\mathcal{P}_x$ by $\mathcal{P}$ in the previous display. This yields

$$\mathbb{E}\left[\left|[\chi_p]_{\mathcal{P}}(x)-3^{-2dm}\sum_{y\in\square_m}\sum_{z\in x+y+\square_m}[\chi_p]_{\mathcal{P}}(z)\right|^q\right]\leq\begin{cases}C^q q^{qk}m^{\frac{q}{2}}&\text{if }d=2,\\C^q q^{qk}&\text{if }d\geq 3.\end{cases}$$

This completes the proof of Step 2.

Step 3. This step is similar to Step 1; the main ingredient is slightly different and presented below. The objective is to prove the following inequality: for $m\in\mathbb{N}$ such that $3^m\leq|x|<3^{m+1}$, and for each $q\geq 1$,

$$\left(3^{-dm}\sum_{y\in\square_m}|[\chi_p]_{\mathcal{P}}(y)-([\chi_p]_{\mathcal{P}})_{x+\square_m}|^q\right)^{\frac{1}{q}}\leq\begin{cases}\mathcal{O}_s(Cq^k\sqrt{m})&\text{if }d=2,\\\mathcal{O}_s(Cq^k)&\text{if }d\geq 3.\end{cases}$$

To prove this result, we note that $x+\square_m\subseteq\square_{m+2}$ and compute

(89)
$$\left(3^{-dm}\sum_{y\in\square_m}|[\chi_p]_{\mathcal{P}}(y)-([\chi_p]_{\mathcal{P}})_{x+\square_m}|^q\right)^{\frac{1}{q}}$$

$$\leq\left(3^{-dm}\sum_{y\in\square_{m+2}}|[\chi_p]_{\mathcal{P}}(y)-([\chi_p]_{\mathcal{P}})_{\square_{m+2}}|^q\right)^{\frac{1}{q}}+|([\chi_p]_{\mathcal{P}})_{\square_{m+2}}-([\chi_p]_{\mathcal{P}})_{x+\square_m}|.$$

The first term is estimated by (85) (replacing the cube $\square_m$ by $\square_{m+2}$). We estimate the second term on the right side of (89) as follows:

$$|([\chi_p]_{\mathcal{P}})_{\square_{m+2}}-([\chi_p]_{\mathcal{P}})_{x+\square_m}|\leq 3^{-dm}\sum_{y\in x+\square_m}|[\chi_p]_{\mathcal{P}}(y)-([\chi_p]_{\mathcal{P}})_{\square_{m+2}}|$$

$$\leq C\left(3^{-dm}\sum_{y\in\square_{m+2}}|[\chi_p]_{\mathcal{P}}(y)-([\chi_p]_{\mathcal{P}})_{\square_{m+2}}|^q\right)^{\frac{1}{q}}.$$

Combining the two previous displays with the inequality (84) shows

$$\left(3^{-dm} \sum_{y\in\square_m} |[\chi_p]_\mathcal{P}(y) - ([\chi_p]_\mathcal{P})_{x+\square_m}|^q\right)^{\frac{1}{q}} \leq \begin{cases} \mathcal{O}_s(Cq^k\sqrt{m}) & \text{if } d = 2, \\ \mathcal{O}_s(Cq^k) & \text{if } d \geq 3. \end{cases}$$

With the same proof as in Step 1, we obtain, for each $q \geq 1$

$$\mathbb{E}\left[\left|3^{-dm} \sum_{y\in\square_m} [\chi_p]_\mathcal{P}(0) - ([\chi_p]_\mathcal{P})_{y+x+\square_m}\right|^q\right] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

But note that

$$[\chi_p]_\mathcal{P}(0) - 3^{-2dm} \sum_{y\in\square_m} \sum_{z\in x+y+\square_m} [\chi_p]_\mathcal{P}(z) = 3^{-dm} \sum_{y\in\square_m} [\chi_p]_\mathcal{P}(0) - ([\chi_p]_\mathcal{P})_{y+x+\square_m}.$$

Combining the two previous displays completes the proof of Step 3.

Step 4. In this step, we first split the integral,

$$\mathbb{E}[|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)|^q] \leq 2^q \mathbb{E}\left[\left|[\chi_p]_\mathcal{P}(0) - 3^{-2dm} \sum_{y\in\square_m} \sum_{z\in x+y+\square_m} [\chi_p]_\mathcal{P}(z)\right|^q\right]$$

$$+ 2^q \mathbb{E}\left[\left|[\chi_p]_\mathcal{P}(x) - 3^{-2dm} \sum_{y\in\square_m} \sum_{z\in x+y+\square_m} [\chi_p]_\mathcal{P}(z)\right|^q\right].$$

Combining the results of Step 2 and Step 3, we have, for $m \in \mathbb{N}$ chosen such that $3^m \leq |x| \leq 3^{m+1}$ and for each $q \geq 1$,

$$\mathbb{E}[|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)|^q] \leq \begin{cases} C^q q^{qk} m^{\frac{q}{2}} & \text{if } d = 2, \\ C^q q^{qk} & \text{if } d \geq 3. \end{cases}$$

Since $m \leq \frac{\log|x|}{\log 3}$, the proof of Step 3 is complete.

Step 5. First we extend the result of Step 4 to the case $0 < q < 1$. By the Jensen inequality, we have, for each $0 < q \leq 1$,

$$\mathbb{E}[|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)|^q] \leq \mathbb{E}[|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)|^2]^{\frac{q}{2}} \leq \begin{cases} C^q \log^{\frac{q}{2}} |x| & \text{if } d = 2, \\ C^q & \text{if } d \geq 3. \end{cases}$$

We now prove the main result of this step. We first deal with the case $d = 2$. Select an exponent $s > 0$ depending only on $d, \mathfrak{p}, \lambda$ such that $s < \frac{1}{k}$, where k is the exponent (depending only on $d, \mathfrak{p}, \lambda$) which appears in (82). We then compute by the Stirling formula. We now set $\sigma := \min(\frac{\log 2}{\log(\sum_{l=0}^{\lfloor\frac{1}{s}\rfloor} \frac{Csl}{l!} + \sum_{l=\lceil\frac{1}{s}\rceil}^{\infty} \frac{Csl(sl)^{skl}}{l!})}, 1) > 0$. Note that σ depends only on $d, \mathfrak{p}, \lambda$. With this value of σ, we have

$$\mathbb{E}\left[\exp\left(\sigma\left(\frac{|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)|}{\log^{\frac{1}{2}} |x|}\right)^s\right)\right] \leq 2.$$

From this computation, we obtain

$$|[\chi_p]_\mathcal{P}(x) - [\chi_p]_\mathcal{P}(0)| \leq \mathcal{O}_s(\sigma^{-\frac{1}{s}} \log^{\frac{1}{2}} |x|).$$

Setting $C := \sigma^{-\frac{1}{s}}$, we obtain (83). The proof in dimension $d \geq 3$ follows the same lines and is even simpler since we do not have the square root of the logarithm.

Step 6. In this step, we remove the coarsening. To this end, we prove, for each point $y \in \mathbb{Z}^d$,

$$|\chi_p(y) - [\chi_p]_\mathcal{P}(y)|\mathbb{1}_{\{y\in\mathscr{C}_\infty\}} \leq \mathcal{O}_s(C).$$

Note that if a point y belongs to the infinite cluster then there exists a path connecting y to $\overline{z}(\Box_{\mathcal{P}}(y))$ which lies in the cube $\Box_{\mathcal{P}}(y)$ and its neighbors. Consequently we have the estimate

$$\left|\chi_p(y) - [\chi_p]_{\mathcal{P}}(y)\right| \mathbb{1}_{\{y \in \mathscr{C}_\infty\}} \leq \sum_{x \in \mathscr{C}_\infty \cap B(y, C\,\mathrm{size}(\Box_{\mathcal{P}}(y)))} |\nabla \chi_p|(x).$$

Applying Proposition 6 gives

(90) $$\left|\chi_p(y) - [\chi_p]_{\mathcal{P}}(y)\right| \mathbb{1}_{\{y \in \mathscr{C}_\infty\}} \leq \mathcal{O}_s(C).$$

We deduce that

$$\left|\chi_p(x) - \chi_p(0)\right| \mathbb{1}_{\{0, x \in \mathscr{C}_\infty\}}$$
$$\leq \left|\chi_p(0) - [\chi_p]_{\mathcal{P}}(0)\right| \mathbb{1}_{\{0 \in \mathscr{C}_\infty\}} + \left|\chi_p(x) - [\chi_p]_{\mathcal{P}}(x)\right| \mathbb{1}_{\{x \in \mathscr{C}_\infty\}} + \left|[\chi_p]_{\mathcal{P}}(x) - [\chi_p]_{\mathcal{P}}(0)\right|.$$

Combining the result of Step 5 with the inequality (90) shows

$$\left|\chi_p(x) - \chi_p(0)\right| \mathbb{1}_{\{0, x \in \mathscr{C}_\infty\}} \leq \begin{cases} \mathcal{O}_s(C \log^{\frac{1}{2}} |x|) & \text{if } d = 2, \\ \mathcal{O}_s(C) & \text{if } d \geq 3. \end{cases}$$

The proof of Step 6 is complete. $\quad\square$

APPENDIX A: PROOF OF THE MULTISCALE POINCARÉ INEQUALITY

In this appendix, we prove the L^q multiscale Poincaré inequality stated in Proposition 9. Contrary to the rest of the article, we prove the result in the continuous setting. This choice is motivated by the two reasons listed below:

1. The argument relies on the statement of Proposition D.1 and Remark D.6 of [10], which are stated in the continuous setting;

2. We need to rescale equations and use results of elliptic regularity; it is thus easier to work in the continuous setting.

The discrete version of the inequality stated in Proposition 9 can be deduced from the continuous one by standard arguments.

We first introduce a few definitions pertaining to the continuous setting. We denote by $C_c^\infty(\mathbb{R}^d, \mathbb{R})$ (resp. $C^\infty(\mathbb{R}^d, \mathbb{R})$) the set of smooth compactly supported (resp. smooth) functions in $\mathbb{R}^d$, by $\mathcal{S}(\mathbb{R}^d)$ the Schwartz space, that is,

$$\mathcal{S}(\mathbb{R}^d) := \left\{ f \in C^\infty(\mathbb{R}^d, \mathbb{R}) : \forall (k, \alpha_1, \ldots, \alpha_d) \in \mathbb{N}^{d+1}, \sup_{x \in \mathbb{R}^d} |x|^k |\partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} f(x)| < \infty \right\}$$

and by $\mathcal{S}'(\mathbb{R}^d)$ its topological dual, the space of tempered distribution. Given a domain $U \subseteq \mathbb{R}^d$, we denote by $C_c^\infty(U, \mathbb{R})$ (resp. $C^\infty(U, \mathbb{R})$) the set of smooth compactly supported (resp. smooth) functions in U.

For $q \in [1, \infty)$, we denote the $L^q(U)$ and normalized $L^q(U)$ norms by For $k \in \mathbb{N}$, we denote by $W^{k,q}(U)$ the Sobolev space, by $W_0^{k,q}(U)$ the closure of $C_c^\infty(U, \mathbb{R})$ in $W^{k,q}(U)$, and by $W_{\mathrm{loc}}^{k,q}(U)$ the space of local Sobolev functions. For $k \in \mathbb{Z}$ with $k < 0$, we denote by $W^{k,q}(U)$ the topological dual of $W_0^{-k,p}(U)$, with $p = \frac{q}{q-1}$.

PROPOSITION 12 (Multiscale Poincaré inequality, heat kernel version). *For each radius* $r > 0$, *we define the continuous heat kernel*

(91) $$\widetilde{\Phi}_{r^2} : \begin{cases} \mathbb{R}^d \to \mathbb{R} \\ x \mapsto r^{-d} \exp\left(-\frac{|x|^2}{r^2}\right). \end{cases}$$

For each exponent $q \geq 2$, there exists a constant $C := C(d, q) < \infty$ such that for each tempered distribution $u \in W^{1,q}_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d)$ and each radius $R > 0$,

$$(92) \qquad \|u - (u)_{B_R}\|_{\underline{L}^q(B_R)} \leq C\left(\int_{\mathbb{R}^d} R^{-d} e^{-\frac{|x|}{R}} \left(\int_0^R r|\widetilde{\Phi}_{r^2} * \nabla u(x)|^2 \, dr\right)^{\frac{q}{2}}\right)^{\frac{1}{q}}.$$

Moreover the dependence on the q variable of the constant C can be estimated as follows, for each $q \geq 2$,

$$C(d, q) \leq Aq^{\frac{3}{2}}$$

for some constant $A := A(d) < \infty$.

Before starting the proof, we need to state the following proposition from [10], Proposition D.1 and Remark D.6, and to record a result from the elliptic regularity theory.

PROPOSITION 13 (Proposition D.1 and Remark D.6 of [10]). *For each $q \geq 2$, there exists a constant $C := C(d, q) < \infty$ such that for every tempered distribution $w \in \mathcal{S}'(\mathbb{R}^d)$,*

$$(93) \qquad \|w\|_{W^{-1,q}(B_1)} \leq C\left(\int_{\mathbb{R}^d} e^{-|x|} \left(\int_0^1 r|\widetilde{\Phi}_{r^2} * w(x)|^2 \, dr\right)^{\frac{q}{2}}\right)^{\frac{1}{q}}.$$

Moreover the constant C satisfies, for each exponent $q \geq 2$, $C(d, q) \leq A\sqrt{q}$, for some constant $A := A(d) < \infty$.

REMARK 5. The statement of [10], Proposition D.1 and Remark D.6, is not identical to the statement of Proposition 13; one needs to perform the change of variables $r^2 := t$ to obtain the estimate (93) from the one of [10].

REMARK 6. The dependence on the q variable of the constant C is not explicit in [10]. It can be recovered from the proof.

We then record a result from the theory of elliptic regularity, it can be found in [31], Lemma 7.12 and Proposition 9.9.

PROPOSITION 14 (Lemma 7.12 and Proposition 9.9 of [31]). *Let $V \subseteq \mathbb{R}^d$ be a bounded domain of $\mathbb{R}^d$. Let $f \in L^p(V)$, $1 < p < \infty$, and let w be the Newtonian potential of f, that is,*

$$w(x) := \int_V \Gamma(x - y) f(y) \, dy,$$

where Γ is the fundamental solution of the Laplace equation, that is,

$$\Gamma(x) := \begin{cases} \dfrac{1}{2\pi} \log |x| & \text{if } d = 2, \\ \dfrac{1}{d(2-d)\omega_d} |x|^{2-d} & \text{if } d \geq 3, \end{cases}$$

where ω_d is the volume of the unit sphere in $\mathbb{R}^d$. Then $w \in W^{2,p}(V)$, $\Delta w = f$ a.e. and

$$\|\nabla^2 w\|_{L^p(V)} \leq C_0 \|f\|_{L^p(V)}$$

and

$$\|w\|_{L^p(V)} + \|\nabla w\|_{L^p(V)} \leq C_1 \|f\|_{L^p(V)}$$

for some constants $C_1 := C_1(d, V) < \infty$ and $C_0 := C_0(d, p, V) < \infty$. Moreover, the dependence on p of the constant C_0 can be explicited:

$$C_0(d, p, V) \leq Ap \quad \text{if } p \geq 2 \quad \text{and} \quad C_0(d, p, V) \leq A\frac{1}{p-1} \quad \text{if } 1 < p \leq 2$$

for some $A := A(d, V) < \infty$.

Before starting the proof, we mention that the dependence on the p variable is not explicit in [31], Proposition 9.9; it can be recovered by keeping track of the constant p in the application of the Marcinkiewicz interpolation theorem. Let us also mention that the case of the logarithmic potential is not considered in [31], Lemma 7.12, (it is useful to obtain the estimate of the L^p norm of w in dimension 2). Nevertheless their proof is general enough to be applied in this setting.

PROOF OF PROPOSITION 9. Let $\psi \in C_c^\infty(B_{\frac{1}{4}}, \mathbb{R})$ and $2 \leq q < \infty$. We denote by p the conjugate exponent of q, that is, $p := \frac{q}{q-1} \in (1, 2]$. We split the proof into 5 steps.

- In Step 1, we show that there exists a constant $C := C(d, \psi) < \infty$ such that, for each function $u \in W^{1,q}(B_1)$,

(94)
$$\|u - \psi * u\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}.$$

- In Step 2, we prove that there exists a constant $C := C(d, \psi) < \infty$ such that, for each function $u \in W^{1,q}(B_1)$,

(95)
$$\|u - \psi * u(0)\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}.$$

- In Step 3, we prove that there exists a constant $C := C(d, q, \psi) < \infty$ such that, for each function $u \in W^{1,q}(B_1)$,

$$\|u\|_{L^q(B_{\frac{1}{2}})} \leq C\|\nabla u\|_{W^{-1,q}(B_{\frac{3}{4}})} + C\|u\|_{W^{-1,q}(B_{\frac{3}{4}})}$$

and that the constant C satisfies $C(d, \psi, q) \leq Aq$ for some $A := A(d, \psi) < \infty$.
- In Step 4, we show that there exists a constant $C := C(d, q, \psi) < \infty$ such that, for each function $u \in W^{1,q}(B_1)$,

$$\|u - (u)_{B_{\frac{1}{2}}}\|_{L^q(B_{\frac{1}{2}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}$$

and that the constant C satisfies $C(d, \psi, q) \leq Aq$ for some constant $A := A(d, \psi) < \infty$.
- In Step 5, we show that for each tempered distribution $u \in W_{\text{loc}}^{1,q}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d)$ and each radius $R > 0$,

$$\|u - (u)_{B_R}\|_{\underline{L}^q(B_R)} \leq C\left(\int_{\mathbb{R}^d} R^{-d} e^{-\frac{|x|}{2R}} \left(\int_0^{2R} r|\widetilde{\Phi}_{r^2} * \nabla u(x)|^2 \, dr\right)^{\frac{q}{2}}\right)^{\frac{1}{q}}$$

and that the constant C satisfies $C(d, q) \leq Aq^{\frac{3}{2}}$ for some $A := A(d) < \infty$.

Step 1. We prove that there exists a constant $C := C(d) < \infty$ such that

$$\|u - u * \psi\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}.$$

Define, for each integer $n \in \mathbb{N}$,

$$\psi_n := 2^{-dn}\psi(2^n \cdot).$$

Since $\psi_n * u \to u$ in $L^q(B_{\frac{3}{4}})$, we can use the triangle inequality to bound

(96)
$$\|u - \psi * u\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq \sum_{n=0}^{\infty} \|\psi_{n+1} * u - \psi_n * u\|_{W^{-1,q}(B_{\frac{3}{4}})}.$$

Since the function $\psi_1 - \psi_0$ is compactly supported in $B_{\frac{1}{4}}$ and of mean 0, we can apply [10], Lemma 5.7, to show that there exists a function $\Psi \in C_c^{\infty}(B_{\frac{1}{4}}, \mathbb{R}^d)$ satisfying

$$\nabla \cdot \Psi = \psi_1 - \psi_0.$$

For each integer $n \in \mathbb{N}$, we denote

$$\Psi_n := 2^{-dn} \Psi(2^n \cdot),$$

by scaling invariance we also have

$$2^{-n} \nabla \cdot \Psi_n = \psi_{n+1} - \psi_n.$$

For each function $g \in W_0^{1,p}(B_{\frac{3}{4}})$, we have

$$\int_{(B_{\frac{3}{4}})} (\psi_{n+1} - \psi_n) * u(x) g(x) \, dx = 2^{-n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \nabla \cdot \Psi_n(x - y) u(y) g(x) \, dx \, dy$$

$$= 2^{-n} \int_{\mathbb{R}^d} \nabla u(y) \cdot \left(\int_{\mathbb{R}^d} \Psi_n(x - y) g(x) \, dx \right) dy.$$

By construction, the function $y \to (\int_{\mathbb{R}^d} \Psi_n(x - y) g(x) \, dx)$ is supported in B_1, we can thus estimate

$$\left| \int_{B_{\frac{3}{4}}} (\psi_{n+1} - \psi_n) * u(x) g(x) \, dx \right| \leq 2^{-n} \left\| \int_{\mathbb{R}^d} \Psi_n(x - \cdot) g(x) \, dx \right\|_{W_0^{1,p}(B_1)} \|\nabla u\|_{W^{-1,q}(B_1)}.$$

Moreover, one can check that there exists a constant $C := C(d, \psi) < \infty$ such that

$$\left\| \int_{\mathbb{R}^d} \Psi_n(x - \cdot) g(x) \, dx \right\|_{W_0^{1,p}(B_1)} \leq C \|g\|_{W_0^{1,p}(B_1)} = C \|g\|_{W_0^{1,p}(B_{\frac{3}{4}})}.$$

Taking the supremum over the functions $g \in W_0^{1,p}(B_{\frac{3}{4}})$ of norm less than 1 and combining this result with (96), we obtain

$$\|u - \psi * u\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq C \|\nabla u\|_{W^{-1,q}(B_1)}$$

for some constant $C := C(d) < \infty$. The proof of Step 1 is complete.

   Step 2. We split the norm

(97) $$\|u - \psi * u(0)\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq \|u - \psi * u\|_{W^{-1,q}(B_{\frac{3}{4}})} + \|\psi * u - \psi * u(0)\|_{W^{-1,q}(B_{\frac{3}{4}})}.$$

But note that, for each point $x \in B_{\frac{3}{4}}$,

(98)
$$|\psi * u(x) - \psi * u(0)| \leq C \|\nabla u\|_{W^{-1,q}(B_1)}.$$

The proof of this estimate is similar to the previous step. By [10], Lemma 5.7, we represent $\psi(\cdot - x) - \psi$ in the form

$$\nabla \cdot \Psi_x = \psi(\cdot - x) - \psi,$$

where $\Psi_x \in C_c^\infty(B_1, \mathbb{R})$ is bounded in $W_0^{1,p}(B_1)$ uniformly in the point $x \in B_{\frac{3}{4}}$. We then prove (98) thanks to an integration by parts. From this argument, we deduce that

$$\|\psi * u - \psi * u(0)\|_{W^{-1,q}(B_{\frac{3}{4}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}.$$

Combining this estimate with (97) and the estimate (94) completes the proof of Step 2.

Step 3. Let $\eta \in C_c^\infty(B_1)$ be a cutoff function satisfying

$$\mathbb{1}_{B_{\frac{1}{2}}} \leq \eta \leq \mathbb{1}_{B_{\frac{3}{4}}} \quad \text{and} \quad |\nabla^2 \eta| + |\nabla \eta| \leq C.$$

For any function $f \in L^p(B_1)$, we denote by w_f the Newtonian potential of f introduced in Proposition 14 with the open set $V = B_1$. We then compute

$$\int_{B_1} \eta(x)u(x)f(x)\,dx$$

$$= \int_{B_1} \eta(x)u(x)\Delta w_f(x)\,dx$$

$$= \int_{B_1} \nabla\eta(x)u(x)\nabla w_f(x) + \eta(x)\nabla u(x)\nabla w_f(x)\,dx$$

$$\leq \|u\|_{W^{-1,q}(B_{\frac{3}{4}})}\|\nabla\eta\nabla w_f\|_{W_0^{1,p}(B_{\frac{3}{4}})} + \|\nabla u\|_{W^{-1,q}(B_{\frac{3}{4}})}\|\eta\nabla w_f\|_{W_0^{1,p}(B_{\frac{3}{4}})}.$$

By the properties of the function η and by Proposition 14, we have

$$\|\nabla\eta\nabla w_f\|_{W_0^{1,p}(B_{\frac{3}{4}})} + \|\eta\nabla w_f\|_{W_0^{1,p}(B_{\frac{3}{4}})} = \|\nabla\eta\nabla w_f\|_{W_0^{1,p}(B_1)} + \|\eta\nabla w_f\|_{W_0^{1,p}(B_1)}$$

$$\leq C\|f\|_{L^p(B_1)}$$

for some constant $C := C(d, p, \eta) < \infty$ satisfying

$$C(d, p, \eta) \leq \frac{A}{p-1},$$

with $A := A(d, \eta) < \infty$. Consequently

$$\|u\|_{L^q(B_{\frac{1}{2}})} \leq \|\eta u\|_{L^q(B_1)} = \sup_{f \in L^p(B_1), \|f\|_{L^p(B_1)}=1} \int_{B_1} \eta(x)u(x)f(x)\,dx$$

$$\leq C\big(\|u\|_{W^{-1,q}(B_{\frac{3}{4}})} + \|\nabla u\|_{W^{-1,q}(B_{\frac{3}{4}})}\big).$$

The proof of Step 3 is complete.

Step 4. Applying the main result of Step 3 to the function $u - \psi * u(0)$, we have

$$\|u - \psi * u(0)\|_{L^q(B_{\frac{1}{2}})} \leq C\big(\|u - \psi * u(0)\|_{W^{-1,q}(B_{\frac{3}{4}})} + \|\nabla u\|_{W^{-1,q}(B_1)}\big).$$

Then by Step 2, we obtain

$$\|u - \psi * u(0)\|_{L^q(B_{\frac{1}{2}})} \leq C\|\nabla u\|_{W^{-1,q}(B_1)}.$$

But we have, for each $a \in \mathbb{R}$

$$\|u - (u)_{B_{\frac{1}{2}}}\|_{L^q(B_{\frac{1}{2}})} \leq 2\|u - a\|_{L^q(B_{\frac{1}{2}})}.$$

Thus

$$\|u - (u)_{B_{\frac{1}{2}}}\|_{L^q(B_{\frac{1}{2}})} \leq 2\inf_{a \in \mathbb{R}} \|u - a\|_{L^q(B_{\frac{1}{2}})} \leq 2\|u - \psi * u(0)\|_{L^q(B_{\frac{1}{2}})}.$$

Combining the previous displays completes the proof of Step 4.

Step 5. Applying the result of Step 4 and Proposition 13, we obtain, for each $q \geq 2$ and each $u \in \mathcal{S}'(\mathbb{R}^d) \cap W^{1,q}_{\text{loc}}(\mathbb{R}^d)$,

$$\|u - (u)_{B_{\frac{1}{2}}}\|_{L^q(B_{\frac{1}{2}})} \leq C\left(\int_{\mathbb{R}^d} e^{-|x|}\left(\int_0^1 r|\widetilde{\Phi}_{r^2} * \nabla u(x)|^2 \, dr\right)^{\frac{q}{2}}\right)^{\frac{1}{q}}$$

for some constant $C := C(d, q)$ satisfying $C(d, q) \leq Aq^{\frac{3}{2}}$. Rescaling the previous estimates eventually shows

$$\|u - (u)_{B_R}\|_{\underline{L}^q(B_R)} \leq C\left(\int_{\mathbb{R}^d} R^{-d} e^{-\frac{|x|}{2R}}\left(\int_0^{2R} r|\widetilde{\Phi}_{r^2} * \nabla u(x)|^2 \, dr\right)^{\frac{q}{2}}\right)^{\frac{1}{q}},$$

and the proof of Proposition 9 is complete. $\square$

APPENDIX B: ELLIPTIC INEQUALITIES ON THE SUPERCRITICAL PERCOLATION CLUSTER

In this section, we record some simple elliptic inequalities, the Caccioppoli inequality and the Meyers estimate. These inequalities were written in [7] for harmonic functions. In our context, we need to apply these results when the right-hand term is not 0 but the divergence of a vector field.

PROPOSITION 15 (Caccioppoli inequality). *Assume that we are given a function u :* $\mathscr{C}_\infty \to \mathbb{R}$ *and a vector field $\xi : E_d \to \mathbb{R}$ satisfying the following condition*:

$$(99) \qquad \xi(x, y) = 0 \quad \text{if } \mathbf{a}(x, y) = 0 \text{ or } x, y \notin \mathscr{C}_\infty.$$

In particular, gradients of functions defined on the infinite cluster satisfy this condition by (7). *Assume that u is solution of the equation,*

$$-\nabla \cdot \mathbf{a}\nabla u = -\nabla \cdot \xi \quad \text{in } \mathscr{C}_\infty.$$

Select two bounded sets $U, V \subseteq \mathbb{Z}^d$ such that $V \subseteq U$ and $\text{dist}(V, \partial U) \geq r \geq 1$. Then there exists a constant $C(\lambda) < \infty$ such that

$$(100) \qquad \int_{\mathscr{C}_\infty \cap V} |\nabla u|^2(x) \, dx \leq \frac{C}{r^2} \int_{\mathscr{C}_\infty \cap (U \setminus V)} |u(x)|^2 \, dx + C \int_{\mathscr{C}_\infty \cap U} |\xi|^2(x) \, dx.$$

PROOF. The strategy of the proof follows the standard technique to prove the Caccioppoli inequality, we select a cutoff function $\eta : \mathbb{Z}^d \to \mathbb{R}$ satisfying

$$\mathbb{1}_V \leq \eta \leq 1, \quad \eta \equiv 0 \text{ on } \mathbb{R}^d \setminus U, \quad \text{and}$$

$$(101)$$

$$\forall x, y \in \mathbb{Z}^d \text{ with } x \sim y, \quad |\eta(x) - \eta(y)|^2 \leq \frac{C(\eta(x) + \eta(y))}{r^2},$$

test the equation satisfied by u with the function ηu and perform standard computations. $\square$

The second elliptic estimate needed in this article is the Meyers estimate. This estimate was also proved in [7] in the case of **a**-harmonic functions.

PROPOSITION 11 (Meyers estimate). *There exist a constant $C := C(d, \lambda, \mathfrak{p}) < \infty$, two exponents $s := s(d, \lambda, \mathfrak{p}) > 0$ and $\varepsilon := \varepsilon(d, \lambda, \mathfrak{p}) > 0$ and a random variable $\mathcal{M}_{\text{Meyers}} \leq$*

$\mathcal{O}_s(C)$ *such that for each* $m \in \mathbb{N}$ *with* $3^m \geq \mathcal{M}_{\text{Meyers}}$, *and each function* $v : \mathscr{C}_\infty \to \mathbb{R}$ *satisfying*

$$-\nabla \cdot \mathbf{a}\nabla v = -\nabla \cdot \xi \quad \text{in } \mathscr{C}_\infty$$

for some vector field $\xi : E_d \to \mathbb{R}$ *satisfying* (99), *the following estimate holds*:

$$
(102) \quad
\begin{aligned}
&\left(\frac{1}{|\square_m|} \int_{\square_m \cap \mathscr{C}_\infty} |\nabla v|^{2+\varepsilon}(x)\, dx \right)^{\frac{1}{2+\varepsilon}} \\
&\leq C \left(\frac{1}{|\frac{4}{3}\square_m|} \int_{\frac{4}{3}\square_m \cap \mathscr{C}_\infty} |\nabla v|^2(x)\, dx \right)^{\frac{1}{2}} + C \left(\frac{1}{|\frac{4}{3}\square_m|} \int_{\frac{4}{3}\square_m \cap \mathscr{C}_\infty} |\xi|^{2+\varepsilon}(x)\, dx \right)^{\frac{1}{2+\varepsilon}}.
\end{aligned}
$$

PROOF. The results of Proposition 3.8 and Definition 3.9 of [7] can be adapted in our context to prove (102). The Meyers estimate is a consequence of the three following ingredients: the Caccioppoli inequality, the Sobolev inequality and the Gehring's lemma. Proposition 15 provides a version of the Caccioppoli inequality well-suited to deal with a divergence from the right-hand side. The Sobolev inequality is valid for any functions. The usual version of the Gehring's Lemma (see for instance Theorem 6.6 and Corollary 6.1 of [32]), is general enough to be applied in this context. □

APPENDIX C: TECHNICAL LEMMA

The objective of this appendix is to prove Lemma 3.4 which is restated below.

LEMMA 3.4 (Minimal scale). *There exist a constant* $C := C(d, \mathfrak{p}, \lambda) < \infty$, *an exponent* $s := s(d, \mathfrak{p}, \lambda) > 0$ *and a random variable* $\mathcal{M}_1 \leq \mathcal{O}'_s(C)$ *such that for each integer* $m \in \mathbb{N}$ *satisfying* $3^m \geq \mathcal{M}_1$,

$$
3^{-dm} \sum_{z \in \square_m} \text{size}(\square_{\mathcal{P}}(z))^{\frac{3d(2+\varepsilon)}{\varepsilon}} \left(\sum_{x \in \mathbb{Z}^d,\, \text{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^d \right)^{\frac{2+\varepsilon}{\varepsilon}} \leq C,
$$

where $\varepsilon := \varepsilon(d, \mathfrak{p}, \lambda) > 0$ *is the exponent which appears in Proposition* 11.

PROOF. First, notice that one can rewrite

$$
\begin{aligned}
&3^{-dm} \sum_{z \in \square_m} \text{size}(\square_{\mathcal{P}}(z))^{\frac{3d(2+\varepsilon)}{\varepsilon}} \left(\sum_{x \in \mathbb{Z}^d, \text{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^d \right)^{\frac{2+\varepsilon}{\varepsilon}} \\
&\leq C 3^{-dm} \sum_{z \in \square_m} \text{size}(\square_{\mathcal{P}}(z))^{\frac{3d(2+\varepsilon)+2}{\varepsilon}} \sum_{x \in \mathbb{Z}^d, \text{dist}(\square_{\mathcal{P}}(x), \square_{\mathcal{P}}(z)) \leq 1, e \in \mathcal{B}_d^x} (1 + \mathcal{X}^e(x))^{d\frac{2+\varepsilon}{\varepsilon}} \\
&\leq C 3^{-dm} \sum_{x \in \mathbb{Z}^d, \text{dist}(\square_{\mathcal{P}}(x), \square_m) \leq 1, e \in \mathcal{B}_d^x} \text{size}(\square_{\mathcal{P}}(x))^{\frac{3d(2+\varepsilon)+2}{\varepsilon}} (1 + \mathcal{X}^e(x))^{d\frac{2+\varepsilon}{\varepsilon}}.
\end{aligned}
$$

By Property (iv) of Proposition 1 applied with the exponent $t = \frac{6d(2+\varepsilon)+4}{\varepsilon}$, it is clear that for each integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}_t(\mathcal{P})$, we have:

1. The inequality $\sup_{x \in \square_{m+1}} \text{size}(\square_{\mathcal{P}}(x)) \leq 3^{\frac{dm}{d+t}}$, which implies

$$\{x \in \mathbb{Z}^d, \text{dist}(\square_{\mathcal{P}}(x), \square_m) \leq 1\} \subseteq \square_{m+1}.$$

2. The estimate

$$\left(3^{-dm} \sum_{x \in \mathbb{Z}^d, \text{dist}(\square_{\mathcal{P}}(x), \square_m) \leq 1} \text{size}(\square_{\mathcal{P}}(x))^{\frac{6d(2+\varepsilon)+4}{\varepsilon}}\right)^{\frac{1}{2}}$$

$$\leq C\left(3^{-d(m+1)} \sum_{x \in \square_{m+1}} \text{size}(\square_{\mathcal{P}}(x))^{\frac{6d(2+\varepsilon)+4}{\varepsilon}}\right)^{\frac{1}{2}}$$

$$\leq C.$$

Thus by the Cauchy-Schwarz inequality, it is enough to prove that there exists a random variable $\mathcal{M}$ satisfying $\mathcal{M} \leq \mathcal{O}_s'(C)$, such that for each integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}$,

(103) $$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}_d^x} \left(\mathcal{X}^e(x)\right)^{\frac{d(4+2\varepsilon)}{\varepsilon}} \leq C.$$

Unfortunately, we cannot prove this exact statement; we will prove a slightly weaker estimate, Lemma C.1, which is still strong enough to deduce Proposition 10. Define, for each constant $C > 0$, the random variable

$$\mathcal{X}_C := \inf\Big\{r \in [1, \infty) : \forall r', R' \in [r, \infty), \text{ with } r' \leq R', \forall u \in \mathcal{A}(\mathscr{C}_\infty \cap B_{R'})$$

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{r'})} \leq C\frac{r'}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{R'})}\Big\},$$

and, for each point $x \in \mathbb{Z}^d$,

$$\mathcal{X}_C(x) := \mathcal{X}_C \circ \tau_x.$$

Denote by $C_0 := C_0(d, \mathfrak{p}, \lambda) < \infty$ the constant appearing in Proposition 6. By definition we have

$$\mathcal{X}_{C_0} = \mathcal{X}.$$

Note also that the map $C \mapsto \mathcal{X}_C$ is nonincreasing. We have the following lemma.

LEMMA C.1. *For every integrability parameter $t > 0$, there exist a constant $C(d, \mathfrak{p}, \lambda, t) < \infty$, an exponent $s(d, \mathfrak{p}, \lambda, t) > 0$ and a random variable $\mathcal{M}_t^{\mathcal{X}}$ satisfying*

$$\mathcal{M}_t^{\mathcal{X}} \leq \mathcal{O}_s'(C)$$

such that for every integer $m \in \mathbb{N}$ satisfying $3^m \geq \mathcal{M}_t^{\mathcal{X}}$, one has the inequality

$$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}_d^x} |\mathcal{X}_{C_0^2}^e(x)|^t \leq C.$$

REMARK 7.

1. This statement is weaker than (103) since, for each $x \in \mathbb{Z}^d$ and $e \in \mathcal{B}_d^x$,

$$\mathcal{X}_{C_0^2}^e(x) \leq \mathcal{X}_{C_0}^e(x) = \mathcal{X}^e(x).$$

Nevertheless it is enough to prove Result 2, since we only need to replace C_0 by C_0^2 in every computation involving the estimates on the random variables $\chi_p^e(x)$; the result remains unchanged.

2. Applying this result with $t = \frac{d(4+2\varepsilon)}{\varepsilon}$ completes the proof of Lemma 3.4. $\square$

There remains to prove Lemma C.1. Before starting the proof, we introduce a few ingredients and preliminary results. First define, for each radius $R > 0$, and each constant $C > 0$,

(104)
$$\mathcal{X}_{R,C} := \inf\left\{r \in [1, R] : \forall r', R' \in [r, R], \text{ with } r' \leq R', \forall u \in \mathcal{A}(\mathscr{C}_{\max}(B_R) \cap B_{R'})\right.$$
$$\left.\|\nabla u\|_{\underline{L}^2(\mathscr{C}_{\max}(B_R) \cap B_{r'})} \leq C\frac{r'}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}_{\max}(B_R) \cap B_{R'})}\right\},$$

where $\mathscr{C}_{\max}(B_R)$ denotes the largest cluster contained in B_R; if there is more than one candidate, we break ties using a deterministic procedure. Similarly we define, for each point $x \in \mathbb{Z}^d$,

$$\mathcal{X}_C(x) := \mathcal{X}_C \circ \tau_x.$$

This random variable is defined on the enlarged probability space $\Omega \times \Omega$ and is measurable with respect to the σ-algebra $\mathcal{F}(x + B_R) \otimes \{\varnothing, \Omega\}$ (it depends on the edges in the ball $x + B_R$ of the first variable and does not depend on the edges of the second variable). Consequently the random variables $\mathcal{X}_{R,C}(x)$ and $\mathcal{X}_{R,C}(y)$ are independent if $|x - y| > 2R$.

Note also that $\mathcal{X}_{R,C}$ is decreasing in the C variable and, for $R \geq \mathcal{M}_t(\mathcal{P})$, it is increasing in the R variable. We thus denote by, for each $C \geq 1$

$$\mathcal{X}_C := \lim_{R \to \infty} \mathcal{X}_{R,C} = \limsup_{R \geq 1} \mathcal{X}_{R,C} \in [1, \infty].$$

By Proposition 6, we know that there exists a constant $C_0 := C_0(d, \mathfrak{p}, \lambda) < \infty$ such that

(105)
$$\mathcal{X}_{C_0} = \mathcal{X} \leq \mathcal{O}_s'(C),$$

thus, for each radius $R \geq 1$,

$$\mathcal{X}_{R,C_0}\mathbb{1}_{\{R \geq \mathcal{M}_t(\mathcal{P})\}} \leq \mathcal{X}_{C_0} \leq \mathcal{O}_s'(C).$$

Moreover, for each radius $R \geq 1$, we have

$$\mathcal{X}_{R,C_0}\mathbb{1}_{\{R \in [1, \mathcal{M}_t(\mathcal{P})]\}} \leq \mathcal{M}_t(\mathcal{P}) \leq \mathcal{O}_s'(C).$$

Combining the two previous displays yields, for each $R \geq 1$,

$$\mathcal{X}_{R,C_0} \leq \mathcal{O}_s'(C).$$

We now prove the following inequality, for each $R, C > 1$,

(106)
$$\mathcal{X}_{C^2} \leq \mathcal{X}_{R,C} + R\mathbb{1}_{\{R \leq \mathcal{M}_t(\mathcal{P})\}} + \mathcal{X}_C\mathbb{1}_{\{\mathcal{X}_C > R\}}.$$

We split the proof of this inequality into two cases.

Case 1. If $\mathcal{X}_C > R$, then since $C \geq 1$ and the map $C \mapsto \mathcal{X}_C$ is decreasing, the inequality (106) is a consequence of the estimate

$$\mathcal{X}_{C^2} \leq \mathcal{X}_C \leq \mathcal{X}_{R,C} + R\mathbb{1}_{\{R \leq \mathcal{M}_t(\mathcal{P})\}} + \mathcal{X}_C\mathbb{1}_{\{\mathcal{X}_C > R\}}.$$

Case 2. If $\mathcal{X}_C \leq R$ and $R \leq \mathcal{M}_t(\mathcal{P})$, then

$$\mathcal{X}_{C^2} \leq \leq \mathcal{X}_C R\mathbb{1}_{\{R \leq \mathcal{M}_t(\mathcal{P})\}} \leq \mathcal{X}_{R,C} + R\mathbb{1}_{\{R \leq \mathcal{M}_t(\mathcal{P})\}} + \mathcal{X}_C\mathbb{1}_{\{\mathcal{X}_C > R\}}.$$

Case 3. If $\mathcal{X}_C \leq R$ and $R \geq \mathcal{M}_t(\mathcal{P})$ then $\mathscr{C}_{\max}(B_R)$ is equal to the maximal connected component of $\mathscr{C}_\infty \cap B_R$ and we have, for each pair of radii $r, R' > R$ satisfying $R' \geq r$

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leq C\frac{r}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{R'})}.$$

Moreover, for each pair of radii $r, R' \in [\mathcal{X}_{R,C}, R]$ with $R' \geq r$, we have

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leq C\frac{r}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{R'})}.$$

Combining the two previous displays and using $C^2 \geq C$ yields for each pair of radii $r, R' \geq \mathcal{X}_{R,C}$ with $R' \geq r$,

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leq C^2\frac{r}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{R'})}$$

and thus by definition of the random variable $\mathcal{X}_{C^2}$,

$$\mathcal{X}_{C^2} \leq \mathcal{X}_{R,C}$$

and the proof of the inequality (106) is complete.

For $x \in \mathbb{Z}^d$, $e = \{x, y\} \in \mathcal{B}_d$, $C, R \in [1, \infty)$, denote by $\mathcal{X}^e_{R,C}(x)$ the translated and resampled random variable

$$\mathcal{X}^e_{R,C}(x) := \inf\Big\{r \in [1, R] : \forall 1 \leq r' \leq R' \leq R, u \in \mathcal{A}^e\big(\mathscr{C}^e_{\max}(B_R(x)) \cap B_{R'}(x)\big)$$

$$\|\nabla u\|_{\underline{L}^2(\mathscr{C}^e_{\max}(B_R) \cap B_{r'}(x))} \leq C\frac{r'}{R'}\|\nabla u\|_{\underline{L}^2(\mathscr{C}^e_{\max}(B_R(x)) \cap B_{R'}(x))}\Big\}.$$

We also define, for each point $x \in \mathbb{Z}^d$

$$\mathcal{X}^e_C(x) := \lim_{R \to \infty} \mathcal{X}^e_{R,C}(x) = \limsup_{R \geq 1} \mathcal{X}^e_{R,C}(x) \in [1, \infty].$$

The second ingredient in the proof of Lemma C.1 is the following minimal scale lemma. It is an adaptation of [7], Lemma 2.3, and will be used in the proof of Lemma C.1.

LEMMA C.2.   *Fix $K \geq 1$, $s > 0$ and $\beta > 0$ and suppose that $\{X_n\}_{n \in \mathbb{N}}$ is a sequence of random variables satisfying, for every $n \in \mathbb{N}$,*

$$X_n \leq K + \mathcal{O}_s(K3^{-n\beta}).$$

Then there exists $C(s, \beta, K) < \infty$ such that the random scale

$$M := \sup\{3^n \in \mathbb{N} : X_n \geq K + 1\}$$

satisfies the estimate

$$M \leq \mathcal{O}_{s\beta}(C).$$

PROOF.   This result can be deduced by applying [7], Lemma 2.3, to the sequence of random variables $X'_n = \max(X_n - K, 0)$. $\square$

We now turn to the proof of Lemma C.1.

PROOF OF LEMMA C.1.   Fix an exponent $t \in (0, \infty)$ and $m, n \in \mathbb{N}$ with $m > n$. Using the inequality (106), we have

$$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{C^2_0}(x)|^t$$

$$(107) \qquad \leq C3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0}(x)|^t + C3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{C_0}(x)|^t \mathbb{1}_{\{\mathcal{X}^e_{C_0}(x) > 3^n\}}$$

$$+ C3^{-dm} \sum_{x \in \square_m} 3^{tn} \mathbb{1}_{\{3^n \leq \mathcal{M}_t(\mathcal{P}^e)\}} \circ \tau_x.$$

Since $\mathcal{X}^e_{C_0}(x) \leq \mathcal{O}'_s(C)$, for every $t, t' > 0$, there exist an exponent $s'(d, \mathfrak{p}, \lambda, t, t') > 0$ and a constant $C'(d, \mathfrak{p}, \lambda, t, t') < \infty$ such that

$$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{C_0}(x)|^t \mathbb{1}_{\{\mathcal{X}^e_{C_0}(x) > 3^n\}} \leq \mathcal{O}'_{s'}(C' 3^{-nt'})$$

and

$$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} 3^{nt} \mathbb{1}_{\{3^n \geq \mathcal{M}_t(\mathcal{P})\}} \circ \tau_x \leq \mathcal{O}'_{s'}(C' 3^{-nt'}).$$

Combining the previous displays yields

$$3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{C_0^2}(x)|^t \leq C 3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0}(x)|^t + \mathcal{O}'_{s'}(C' 3^{-nt'}).$$

Moreover, notice that by definition of the localized random variable $\mathcal{X}^e_{3^n, C_0}(x)$, we have for each $x \in \mathbb{Z}^d$

$$\sum_{e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0}(x)|^t \leq 2d \times 3^{nt}.$$

The proof of the lemma is then the same as the proof of Steps 1 and 2 of [7], Proposition 2.1, with the random variable $3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{C_0^2}(x)|^t$ instead of $\Lambda_t(z + \square_m, \mathcal{S})$ and the random variable $3^{-dm} \sum_{x \in z + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t$ instead of $\Lambda_t(z' + \square_n, \mathcal{S}_{\text{loc}}(z'))$. We rewrite it for completeness.

We denote by

$$Z := 3^{-dm} \sum_{x \in \square_m, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0}(x)|^t = \frac{|\square_n|}{|\square_m|} \sum_{z \in 3^n \mathbb{Z}^d \cap \square_m} 3^{-dm} \sum_{x \in z + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0}(x)|^t.$$

We first prove that there exists a constant $C := C(d, \mathfrak{p}, \lambda, t) < \infty$ such that

(108) $$Z \leq C + \mathcal{O}'_1(3^{nt - d(m-n)}).$$

To this end, choose an enumeration $\{z^j : 1 \leq j \leq 3^{d(m-n-2)}\}$ of the elements of the set $3^{n+2} \mathbb{Z}^d \cap \square_m$. For each $1 \leq j \leq 3^{d(m-n-2)}$, we let $\{z^{i,j} : 1 \leq i \leq 3^{2d}\}$ be an enumeration of the elements of the set $3^n \mathbb{Z}^d \cap (z^j + \square_{n+2})$, such that, for every $1 \leq j, j' \leq 3^{d(m-n-2)}$ and $1 \leq i \leq 3^{2d}$, $z^j - z^{j'} = z^{i,j} - z^{i,j'}$. The point of this is that, for every $1 \leq i \leq 3^{2d}$ and $1 \leq j < j' \leq 3^{d(m-n-2)}$, we have $\text{dist}(z^{i,j} + \square_n, z^{i,j'} + \square_n) \geq 3^{n+1}$ and therefore, $3^{-dm} \sum_{x \in z^{i,j} + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t$ and $3^{-dm} \sum_{x \in z^{i,j'} + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t$ are independent. Now fix $h > 0$ and compute, using the Hölder inequality and the independence,

$$\log \mathbb{E}[\exp(h 3^{-nt} Z)]$$

$$= \log \mathbb{E}\left[\prod_{i=1}^{3^{2d}} \prod_{j=1}^{3^{d(m-n-2)}} \exp\left(h 3^{-nt - d(m-n)} 3^{-dm} \sum_{x \in z^{i,j} + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t \right) \right]$$

$$\leq 3^{-2d} \sum_{i=1}^{3^{2d}} \log \mathbb{E}\left[\prod_{j=1}^{3^{d(m-n-2)}} \exp\left(h 3^{-nt - d(m-n-2)} 3^{-dm} \sum_{x \in z^{i,j} + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t \right) \right]$$

$$\leq 3^{-2d} \sum_{i=1}^{3^{2d}} \sum_{j=1}^{3^{d(m-n-2)}} \log \mathbb{E}\left[\exp\left(h 3^{-nt - d(m-n-2)} 3^{-dm} \sum_{x \in z^{i,j} + \square_n, e \in \mathcal{B}^x_d} |\mathcal{X}^e_{3^n, C_0^2}(x)|^t \right) \right].$$

This inequality can be rewritten

$$\log \mathbb{E}[\exp(h3^{-nt}Z)]$$

$$\leq 3^{-2d} \sum_{z' \in 3^n \mathbb{Z}^d \cap (z+\Box_m)} \log \mathbb{E}\left[\exp\left(h3^{-nt-d(m-n-2)-dm} \sum_{x \in z'+\Box_n, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0^2}^e(x)|^t\right)\right].$$

We use the inequality

$$\forall y \in [0,1], \quad \exp(y) \leq 1 + 2y$$

to obtain, for every $h \in [0, (2d)^{-t}3^{d(m-n-2)}]$,

$$\exp\left(h3^{-nt-d(m-n-2)} \sum_{x \in z'+\Box_n, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0^2}^e(x)|^t\right)$$

$$\leq 1 + 2h3^{-nt-d(m-n-2)} \sum_{x \in z'+\Box_n, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0^2}^e(x)|^t.$$

Taking the expectation in the previous display and using the elementary inequality

$$\forall y \geq 0, \quad \log(1+y) \leq y,$$

we obtain

$$\log \mathbb{E}[\exp(h3^{-nt}Z)]$$

$$\leq 3^{d(m-n)} \log\left(1 + 2h3^{-nt-d(m-n-1)} \mathbb{E}\left[\sum_{x \in z'+\Box_n, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0^2}^e(x)|^t\right]\right)$$

$$\leq 2h3^{-nt+d} \mathbb{E}\left[\sum_{x \in z'+\Box_n, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0^2}^e(x)|^t\right]$$

$$\leq Ch3^{-nt}.$$

Taking $h := (2d)^{-t}3^{d(m-n-2)}$ yields

$$\mathbb{E}[\exp((2d)^{-t}3^{d(m-n-2)-nt}Z)] \leq \exp(C3^{d(m-n)-nt}).$$

From this and Chebyshev's inequality, we obtain a constant C such that

$$\mathbb{P}[Z \geq C + h] \leq \exp(-hC^{-1}3^{d(m-n)-nt}).$$

This implies (108).

   Step 2. We complete the proof by applying a union bound. Combining (107) and (108) yields

$$\sum_{x \in \Box_m, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0}^e(x)|^t \leq C + \mathcal{O}_1(C3^{nt-d(m-n)}) + \mathcal{O}_{s'}(C3^{-nt'}).$$

We set

$$n := \left\lceil \frac{dm}{d+t+1} \right\rceil \quad \text{and} \quad t' = 1$$

so that the previous line becomes

$$\sum_{x \in \Box_m, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0}^e(x)|^t \leq C + \mathcal{O}_1'(C3^{-\frac{d}{d+t+1}m}) + \mathcal{O}_{s'}'(C3^{-\frac{d}{d+t+1}m}).$$

Thus, by (15) and (11), we obtain the existence of two exponents $s := s(d, \mathfrak{p}, \lambda, t) > 0$, $\beta := \beta(d, \mathfrak{p}, \lambda, t) > 0$ and of a constant $C_0 := C_0(d, \mathfrak{p}, \lambda, t) < \infty$ such that

$$\sum_{x \in \square_m, \, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0}^e(x)|^t \leq C_0 + \mathcal{O}_s'(C_0 3^{-\beta m}).$$

Define

$$\mathcal{M}_t^{\mathcal{X}} := \sup\left\{3^m : \sum_{x \in \square_m, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0}^e(x)|^t \geq C_0 + 1\right\}.$$

We apply Lemma C.2 with $X_n = \sum_{x \in \square_m, e \in \mathcal{B}_d^x} |\mathcal{X}_{3^n, C_0}^e(x)|^t$ and $K = C_0$ to obtain the inequality

$$\mathcal{M}_t^{\mathcal{X}} \leq \mathcal{O}_{s\beta}(C).$$

The proof is complete. □

Acknowledgement. I would like to thank Scott Armstrong, Jean-Christophe Mourrat and Chenlin Gu for helpful discussions and comments.

REFERENCES

[1] 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

[2] 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

[3] 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

[4] ANDRES, S., DEUSCHEL, J.-D. and SLOWIK, M. (2016). Heat kernel estimates for random walks with degenerate weights. *Electron. J. Probab.* **21** Art. ID 33. MR3492937 https://doi.org/10.1214/16-EJP4382

[5] ANDRES, S. and NEUKAMM, S. (2019). Berry–Esseen theorem and quantitative homogenization for the random conductance model with degenerate conductances. *Stoch. Partial Differ. Equ. Anal. Comput.* **7** 240–296. MR3949105 https://doi.org/10.1007/s40072-018-0127-8

[6] 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

[7] 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

[8] ARMSTRONG, S., HANNUKAINEN, A., KUUSI, T. and MOURRAT, J.-C. (2018). An iterative method for elliptic problems with rapidly oscillating coefficients. Preprint. Available at arXiv:1803.03551.

[9] 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

[10] 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

[11] 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

[12] ARMSTRONG, S. N. and MOURRAT, J.-C. (2016). Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.* **219** 255–348. MR3437852 https://doi.org/10.1007/s00205-015-0908-4

[13] 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

[14] 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

[15] 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

[16] BARLOW, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32** 3024–3084. MR2094438 https://doi.org/10.1214/009117904000000748

[17] 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

[18] 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

[19] 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

[20] 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

[21] 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

[22] BURTON, R. M. and KEANE, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505. MR0990777

[23] 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

[24] 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

[25] DAL MASO, G. and MODICA, L. (1986). Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.* **368** 28–42. MR0850613

[26] 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

[27] 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

[28] 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

[29] 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

[30] 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

[31] GILBARG, D. and TRUDINGER, N. S. (2015). *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin.

[32] GIUSTI, E. (2003). *Direct Methods in the Calculus of Variations*. World Scientific, River Edge, NJ. MR1962933 https://doi.org/10.1142/9789812795557

[33] GLORIA, A., NEUKAMM, S. and OTTO, F. (2014). An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *ESAIM Math. Model. Numer. Anal.* **48** 325–346. MR3177848 https://doi.org/10.1051/m2an/2013110

[34] 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

[35] 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

[36] 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

[37] 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

[38] GLORIA, A. and OTTO, F. (2015). The corrector in stochastic homogenization: Near-optimal rates with optimal stochastic integrability. Preprint. Available at arXiv:1510.08290.

[39] 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

[40] LAMACZ, A., NEUKAMM, S. and OTTO, F. (2015). Moment bounds for the corrector in stochastic homogenization of a percolation model. *Electron. J. Probab.* **20** Art. ID 106. MR3418538 https://doi.org/10.1214/EJP.v20-3618

[41] 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

[42] 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

[43] 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

[44] 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

[45] MOURRAT, J.-C. (2012). A quantitative central limit theorem for the random walk among random conductances. *Electron. J. Probab.* **17** Art. ID 97. MR2994845 https://doi.org/10.1214/EJP.v17-2414

[46] MOURRAT, J.-C. (2019). Efficient methods for the estimation of homogenized coefficients. *Found. Comput. Math.* **19** 435–483. MR3937959 https://doi.org/10.1007/s10208-018-9389-9

[47] NADDAF, A. and SPENCER, T. (1998). Estimates on the variance of some homogenization problems. Unpublished preprint.

[48] 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

[49] 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

[50] 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

[51] 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