

Noam Berger<sup>1,2</sup> · Marek Biskup<sup>2</sup>

# Quenched invariance principle for simple random walk on percolation clusters

**Abstract.** We consider the simple random walk on the (unique) infinite cluster of supercritical bond percolation in  $\mathbb{Z}^d$  with  $d \geq 2$ . We prove that, for almost every percolation configuration, the path distribution of the walk converges weakly to that of non-degenerate, isotropic Brownian motion. Our analysis is based on the consideration of a harmonic deformation of the infinite cluster on which the random walk becomes a square-integrable martingale. The size of the deformation, expressed by the so called corrector, is estimated by means of ergodicity arguments.

---

## 1. Introduction

### 1.1. Motivation and model

Consider supercritical bond-percolation on  $\mathbb{Z}^d$ ,  $d \geq 2$ , and the simple random walk on the (unique) infinite cluster. In [38] Sidoravicius and Sznitman asked the following question: Is it true that for a.e. configuration in which the origin belongs to the infinite cluster, the random walk started at the origin exits the infinite symmetric slab  $\{(x_1, \dots, x_d) : |x_d| \leq N\}$  through the “top” side with probability tending to  $1/2$  as  $N \rightarrow \infty$ ? Sidoravicius and Sznitman managed to answer their question affirmatively in dimensions  $d \geq 4$  but dimensions  $d = 2, 3$  remained open. In this paper we extend the desired conclusion to all  $d \geq 2$ . As in [38], we will do so by proving a quenched invariance principle for the paths of the walk.

Random walk on percolation clusters is only one of many instances of “statistical mechanics in random media” that have been recently considered by physicists and mathematicians. Other pertinent examples include, e.g., various diluted spin systems, random copolymers [40], spin glasses [10, 41], random-graph models [9], etc. From this general perspective, the present problem is interesting for at least two reasons: First, a good handle on simple random walk on a given graph is often a prerequisite for the understanding of more complicated processes, e.g., self-avoiding walk or loop-erased random walk. Second, information about the scaling properties of simple random walk on percolation cluster can, in principle, reveal some new important facts about the structure of the infinite cluster and/or its harmonic properties.

---

<sup>1</sup> Department of Mathematics, Caltech, Pasadena, CA 91125, U.S.A.

<sup>2</sup> Department of Mathematics, UCLA, Los Angeles, CA 90095, U.S.A.

Let us begin developing the mathematical layout of the problem. Let  $\mathbb{Z}^d$  be the  $d$ -dimensional hypercubic lattice and let  $\mathbb{B}_d$  be the set of nearest neighbor edges. We will use  $b$  to denote a generic edge,  $\langle x, y \rangle$  to denote the edge between  $x$  and  $y$ , and  $e$  to denote the edges from the origin to its nearest neighbors. Let  $\Omega = \{0, 1\}^{\mathbb{B}_d}$  be the space of all percolation configurations  $\omega = (\omega_b)_{b \in \mathbb{B}_d}$ . Here  $\omega_b = 1$  indicates that the edge  $b$  is occupied and  $\omega_b = 0$  implies that it is vacant. Let  $\mathcal{B}$  be the Borel  $\sigma$ -algebra on  $\Omega$ —defined using the product topology—and let  $\mathbb{P}$  be an i.i.d. measure such that  $\mathbb{P}(\omega_b = 1) = p$  for all  $b \in \mathbb{B}_d$ . If  $x \xleftrightarrow{\omega} \infty$  denotes the event that the site  $x$  belongs to an infinite self-avoiding path using only occupied bonds in  $\omega$ , we write  $\mathcal{C}_\infty = \mathcal{C}_\infty(\omega)$  for the set

$$\mathcal{C}_\infty(\omega) = \{x \in \mathbb{Z}^d : x \xleftrightarrow{\omega} \infty\}. \quad (1.1)$$

By Burton-Keane’s uniqueness theorem [12], the infinite cluster is unique and so  $\mathcal{C}_\infty$  is connected with  $\mathbb{P}$ -probability one.

For each  $x \in \mathbb{Z}^d$ , let  $\tau_x : \Omega \rightarrow \Omega$  be the “shift by  $x$ ” defined by  $(\tau_x \omega)_b = \omega_{x+b}$ . Note that  $\mathbb{P}$  is  $\tau_x$ -invariant for all  $x \in \mathbb{Z}^d$ . Let  $p_c = p_c(d)$  denote the percolation threshold on  $\mathbb{Z}^d$  defined as the infimum of all  $p$ ’s for which  $\mathbb{P}(0 \in \mathcal{C}_\infty) > 0$ . Let  $\Omega_0 = \{0 \in \mathcal{C}_\infty\}$  and, for  $p > p_c$ , define the measure  $\mathbb{P}_0$  by

$$\mathbb{P}_0(A) = \mathbb{P}(A | \Omega_0), \quad A \in \mathcal{B}. \quad (1.2)$$

We will use  $\mathbb{E}_0$  to denote expectation with respect to  $\mathbb{P}_0$ .

For each configuration  $\omega \in \Omega_0$ , let  $(X_n)_{n \geq 0}$  be the simple random walk on  $\mathcal{C}_\infty(\omega)$  started at the origin. Explicitly,  $(X_n)_{n \geq 0}$  is a Markov chain with state space  $\mathbb{Z}^d$ , whose distribution  $P_{0,\omega}$  is defined by the transition probabilities

$$P_{0,\omega}(X_{n+1} = x + e | X_n = x) = \frac{1}{2d} \mathbf{1}_{\{\omega_e=1\}} \circ \tau_x, \quad |e| = 1, \quad (1.3)$$

and

$$P_{0,\omega}(X_{n+1} = x | X_n = x) = \sum_{e: |e|=1} \frac{1}{2d} \mathbf{1}_{\{\omega_e=0\}} \circ \tau_x, \quad (1.4)$$

with the initial condition

$$P_{0,\omega}(X_0 = 0) = 1. \quad (1.5)$$

Thus, at each unit of time, the walk picks a neighbor at random and if the corresponding edge is occupied, the walk moves to this neighbor. If the edge is vacant, the move is suppressed.

## 1.2. Main results

Our main result is that for  $\mathbb{P}_0$ -almost every  $\omega \in \Omega_0$ , the linear interpolation of  $(X_n)$ , properly scaled, converges weakly to Brownian motion. For every  $T > 0$ , let  $(C[0, T], \mathcal{W}_T)$  be the space of continuous functions  $f : [0, T] \rightarrow \mathbb{R}$  equipped with the  $\sigma$ -algebra  $\mathcal{W}_T$  of Borel sets relative to the supremum topology. The precise statement is now as follows:

**Theorem 1.1.** *Let  $d \geq 2$ ,  $p > p_c(d)$  and let  $\omega \in \Omega_0$ . Let  $(X_n)_{n \geq 0}$  be the random walk with law  $P_{0,\omega}$  and let*

$$\tilde{B}_n(t) = \frac{1}{\sqrt{n}}(X_{\lfloor tn \rfloor} + (tn - \lfloor tn \rfloor)(X_{\lfloor tn \rfloor + 1} - X_{\lfloor tn \rfloor})), \quad t \geq 0. \quad (1.6)$$

*Then for all  $T > 0$  and for  $\mathbb{P}_0$ -almost every  $\omega$ , the law of  $(\tilde{B}_n(t) : 0 \leq t \leq T)$  on  $(C[0, T], \mathscr{W}_T)$  converges weakly to the law of an isotropic Brownian motion  $(B_t : 0 \leq t \leq T)$  whose diffusion constant,  $D = E(|B_1|^2) > 0$ , depends only on the percolation parameter  $p$  and the dimension  $d$ .*

The Markov chain  $(X_n)_{n \geq 0}$  represents only one of two natural ways to define a simple random walk on the supercritical percolation cluster. Another possibility is that, at each unit of time, the walk moves to a site chosen uniformly at random from the *accessible* neighbors, i.e., the walk takes no pauses. In order to define this process, let  $(T_k)_{k \geq 0}$  be the sequence of stopping times that mark the moments when the walk  $(X_n)_{n \geq 0}$  made a move. Explicitly,  $T_0 = 0$  and

$$T_{n+1} = \inf\{k > T_n : X_k \neq X_{k-1}\}, \quad n \geq 0. \quad (1.7)$$

Using these stopping times—which are  $P_{0,\omega}$ -almost surely finite for all  $\omega \in \Omega_0$ —we define a new Markov chain  $(X'_n)_{n \geq 0}$  by

$$X'_n = X_{T_n}, \quad n \geq 0. \quad (1.8)$$

It is easy to see that  $(X'_n)_{n \geq 0}$  has the desired distribution. Indeed, the walk starts at the origin and its transition probabilities are given by

$$P_{0,\omega}(X'_n = x + e | X'_n = x) = \frac{\mathbf{1}_{\{\omega_e=1\}} \circ \tau_x}{\sum_{e' : |e'|=1} \mathbf{1}_{\{\omega_{e'}=1\}} \circ \tau_x}, \quad |e| = 1. \quad (1.9)$$

A simple modification of the arguments leading to Theorem 1.1 allows us to establish a functional central limit theorem for this random walk as well:

**Theorem 1.2.** *Let  $d \geq 2$ ,  $p > p_c(d)$  and let  $\omega \in \Omega_0$ . Let  $(X'_n)_{n \geq 0}$  be the random walk defined from  $(X_n)_{n \geq 0}$  as described in (1.8) and let  $\tilde{B}'_n(t)$  be the linear interpolation of  $(X'_k)_{0 \leq k \leq n}$  defined by (1.6) with  $(X_k)$  replaced by  $(X'_k)$ . Then for all  $T > 0$  and for  $\mathbb{P}_0$ -almost every  $\omega$ , the law of  $(\tilde{B}'_n(t) : 0 \leq t \leq T)$  on  $(C[0, T], \mathscr{W}_T)$  converges weakly to the law of an isotropic Brownian motion  $(B_t : 0 \leq t \leq T)$  whose diffusion constant,  $D' = E(|B_1|^2) > 0$ , depends only on the percolation parameter  $p$  and the dimension  $d$ .*

De Gennes [17], who introduced the problem of random walk on percolation cluster to the physics community, thinks of the walk as the motion of “an ant in a labyrinth.” From this perspective, the “lazy” walk  $(X_n)$  corresponds to a “blind” ant, while the “agile” walk  $(X'_n)$  represents a “myopic” ant. While the character of the scaling limit of the two “ants” is the same, there seems to be some distinction in the rate the scaling limit is approached, cf [22] and references therein. As we will see in the proof, the diffusion constants  $D$  and  $D'$  are related via  $D' = D\Theta^2$ , where  $\Theta^{-1}$  is the expected degree of the origin normalized by  $2d$ , cf (6.23).

There is actually yet another way how to “put” simple random walk on  $\mathcal{C}_\infty$ , and that is to use continuous time. Here the corresponding result follows by combining the CLT for the “lazy” walk with an appropriate Renewal Theorem for exponential waiting times.

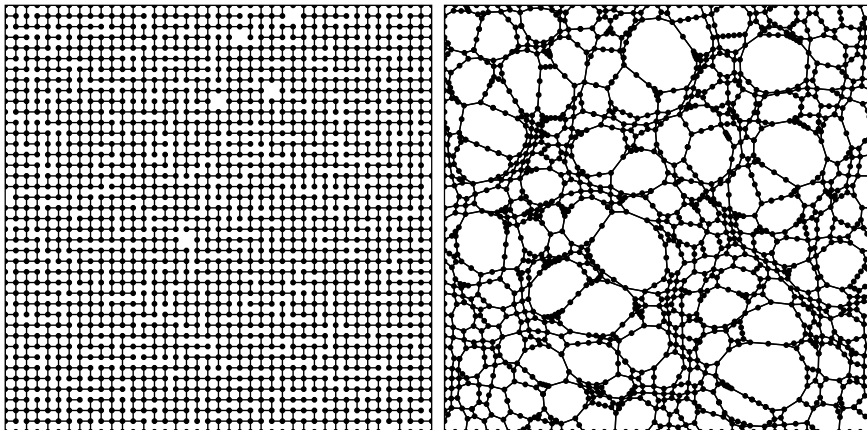
### 1.3. Discussion and related work

The subject of random walk in random environment has a long history; we refer to, e.g., [10, 42] for recent overviews of (certain parts of) this field. On general grounds, each random-media problem comes in two distinct flavors: *quenched*, corresponding to the situations discussed above where the walk is distributed according to an  $\omega$ -dependent measure  $P_{0,\omega}$ , and *annealed*, in which the path distribution of the walk is taken from the averaged measure  $A \mapsto \mathbb{E}_0(P_{0,\omega}(A))$ . Under suitable ergodicity assumptions, the annealed problem typically corresponds to the quenched problem averaged over the starting point. Yet the distinction is clear: In the annealed setting the slab-exit problem from Sect. 1.1 is trivial by the symmetries of the averaged measure, while its answer is *a priori* very environment-sensitive in the quenched measure.

An annealed version of our theorems was proved in the 1980s by De Masi *et al* [13, 14], based on earlier results of Kozlov [28], Kipnis and Varadhan [27] and others in the context of random walk in a field of random conductances. (The results of [13, 14] were primarily two-dimensional but, with the help of [3], they apply to all  $d \geq 2$ ; cf [38].) A number of proofs of quenched invariance principles have appeared in recent years for the cases where an annealed principle was already known. The most relevant paper is that of Sidoravicius and Sznitman [38] which established Theorem 1.2 for random walk among random conductances in all  $d \geq 1$  and, using a very different method, also for random walk on percolation in  $d \geq 4$ . (Thus our main theorem is new only in  $d = 2, 3$ .) The  $d \geq 4$  proof is based on the fact that two independent random walk paths will intersect only very little—something hard to generalize to  $d = 2, 3$ . As this paper shows, the argument for random conductances is somewhat more flexible.

Another paper of relevance is that of Rassoul-Agha and Seppäläinen [37] where a quenched invariance principle was established for *directed* random walks in (space-time) random environments. The directed setting offers the possibility to use independence more efficiently—every time step the walk enters a new environment—but the price to pay for this is the lack of reversibility. The directed nature of the environment also permits consideration of distributions with a drift for which a CLT is not even expected to generally hold in the undirected setting; see [6, 39] for an example of “pathologies” that may arise.

Finally, there have been a number of results dealing with harmonic properties of the simple random walk on percolation clusters. Grimmett, Kesten and Zhang [20] proved via “electrostatic techniques” that this random walk is transient in  $d \geq 3$ ; extensions concerning the existence of various “energy flows” appeared in [1, 24, 26, 29, 33]. A great amount of effort has been spent on deriving estimates on the heat-kernel—i.e., the probability that the walk is at a particular site after  $n$  steps. The first such bounds were obtained by Hecklen and Hoffman [23].



**Fig. 1.** A portion of the infinite cluster  $\mathcal{C}_\infty = \mathcal{C}_\infty(\omega)$  before (left) and after (right) the harmonic deformation  $x \mapsto x + \chi(x, \omega)$ . Here  $p = 0.75$  is already so large that all but a few sites in the entire block belong to  $\mathcal{C}_\infty$ . Upon the deformation, all “holes” (i.e., dual connected components) get considerably stretched and rounded while the “dangling ends” collapse onto the rest of the structure.

Later Mathieu and Remy [31] realized that the right way to approach heat-kernel estimates was through harmonic function theory of the infinite cluster and thus significantly improved the results of [23]. Finally, Barlow [3] obtained, using again harmonic function theory, Gaussian upper and lower bounds for the heat kernel. We refer to [3] for further references concerning this area of research.

*Note:* At the time a preprint version of this paper was first circulated, we learned that Mathieu and Piatnitski had announced a proof of the same result (albeit in continuous-time setting). Their proof, which has in the meantime been posted [30], is close in spirit to that of Theorem 1.1 of [38]; the main tools are Poincaré inequalities, heat-kernel estimates and homogenization theory.

#### 1.4. Outline

Let us outline the main steps of our proof of Theorems 1.1 and 1.2. The principal idea—which permeates in various disguises throughout the work of Papanicolau and Varadhan [35], Kozlov [28], Kipnis and Varadhan [27], De Masi *et al* [13, 14], Sidoravicius and Sznitman [38] and others—is to consider an embedding of  $\mathcal{C}_\infty(\omega)$  into the Euclidean space that makes the corresponding simple random walk a martingale. Formally, this is achieved by finding an  $\mathbb{R}^d$ -valued discrete harmonic function on  $\mathcal{C}_\infty$  with a linear growth at infinity. The distance between the natural position of a site  $x \in \mathcal{C}_\infty$  and its counterpart in this *harmonic embedding* is expressed in terms of the so-called *corrector*  $\chi(x, \omega)$  which is a principal object of study in this paper. See Fig. 1 for an illustration.

It is clear that the corrector can be defined in any finite volume by solving an appropriate discrete Dirichlet problem (this is how Fig. 1 was drawn); the difficult part is to define the corrector in infinite volume while maintaining the natural

(distributional) invariance with respect to shifts of the underlying lattice. Actually, there is an alternative, probabilistic definition of the corrector,

$$\chi(x, \omega) = \lim_{n \rightarrow \infty} (E_{x, \omega}(X_n) - E_{0, \omega}(X_n)). \quad (1.10)$$

However, the only proof we presently have for the existence of such a limit is by following, rather closely, the constructions from Sect. 2.3.

Once we have the corrector under control, the proof splits into two parts: (1) proving that the martingale—i.e., the walk on the deformed graph—converges to Brownian motion and (2) proving that the deformation of the path caused by the change of embedding is negligible. The latter part (which is the principal contribution of this work) amounts to a sublinear bound on the corrector  $\chi(x, \omega)$  as a function of  $x$ . Here, somewhat unexpectedly, our level of control is considerably better in  $d = 2$  than in  $d \geq 3$ . In particular, our proof in  $d = 2$  avoids using any of the recent sophisticated discrete-harmonic analyses but, to handle all  $d \geq 2$  uniformly, we need to invoke the main result of Barlow [3]. The proof is actually carried out along these lines only for the setting in Theorem 1.1; Theorem 1.2 follows by noting that the time scales of both walks are comparable.

Here is a summary of the rest of this paper: In Sect. 2 we introduce the aforementioned corrector and prove some of its basic properties. Sect. 3 collects the needed facts about ergodic properties of the Markov chain “on environments.” Both sections are based on previously known material; proofs have been included to make the paper self-contained. The novel parts of the proof—sublinear bounds on the corrector—appear in Sects. 4–5. The actual proofs of our main theorems are carried out in Sect. 6. The Appendix (Sects. A and B) contains the proof of an upper bound for the transition probabilities of our random walk, further discussion and some conjectures.

## 2. Corrector—construction and harmonicity

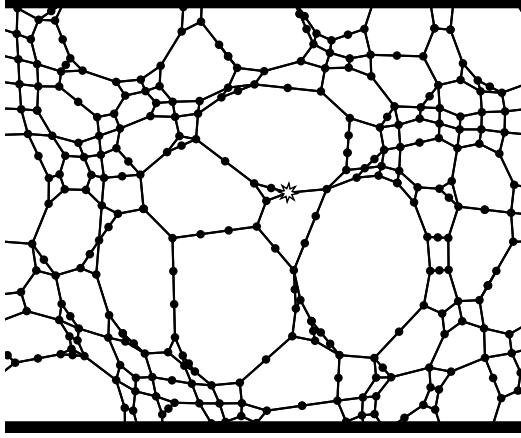
In this section we will define and study the aforementioned corrector which is the basic instrument of our proofs. The main idea is to consider the Markov chain “on environments” (Sect. 2.1). The relevant properties of the corrector are listed in Theorem 2.2 (Sect. 2.2); the proofs are based on spectral calculus (see Sect. 2.3).

### 2.1. Markov chain “on environments”

As is well known, cf Kipnis and Varadhan [27], the Markov chain  $(X_n)_{n \geq 0}$  in (1.3–1.5) induces a Markov chain on  $\Omega_0$ , which can be interpreted as the trajectory of “environments viewed from the perspective of the walk.” The transition probabilities of this chain are given by the kernel  $Q: \Omega_0 \times \mathcal{B} \rightarrow [0, 1]$ ,

$$Q(\omega, A) = \frac{1}{2d} \sum_{e: |e|=1} (\mathbf{1}_{\{\omega_e=1\}} \mathbf{1}_{\{\tau_e \omega \in A\}} + \mathbf{1}_{\{\omega_e=0\}} \mathbf{1}_{\{\omega \in A\}}). \quad (2.1)$$

Our basic observations about the induced Markov chain are as follows:



**Fig. 2.** The harmonic deformation of a percolation configuration in the symmetric slab  $\{(x_1, x_2) \in \mathbb{Z}^2 : |x_2| \leq N\}$ . The star denotes the new position of the origin which in the undeformed configuration was right in the center. The relative vertical shift of the origin corresponds to the deviation of  $P_{0, \omega}$  (top hit before bottom) from one half. The figure also has an interesting electrostatic interpretation: If the bottom and top bars are set to potentials  $-1$  and  $+1$ , respectively, then the site with *deformed* coordinates  $(x_1, x_2)$  has potential  $x_2/N$ .

**Lemma 2.1.** For every bounded measurable  $f : \Omega \rightarrow \mathbb{R}$  and every  $e$  with  $|e| = 1$ ,

$$\mathbb{E}_0(f \circ \tau_e \mathbf{1}_{\{\omega_e=1\}}) = \mathbb{E}_0(f \mathbf{1}_{\{\omega_{-e}=1\}}), \quad (2.2)$$

where  $-e$  is the bond that is opposite to  $e$ . As a consequence,  $\mathbb{P}_0$  is reversible and, in particular, stationary for Markov kernel  $Q$ .

*Proof.* First we will prove (2.2). Neglecting the normalization by  $\mathbb{P}(0 \in \mathcal{C}_\infty)$ , we need that

$$\mathbb{E}(f \circ \tau_e \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_e=1\}}) = \mathbb{E}(f \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_{-e}=1\}}). \quad (2.3)$$

This will follow from  $\mathbf{1}_{\{\omega_e=1\}} = \mathbf{1}_{\{\omega_{-e}=1\}} \circ \tau_e$  and the fact that, on  $\{\omega_e = 1\}$  we have  $\mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} = \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \circ \tau_e$ . Indeed, these observations imply

$$f \circ \tau_e \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_e=1\}} = (f \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_{-e}=1\}}) \circ \tau_e \quad (2.4)$$

and (2.3) then follows by the shift invariance of  $\mathbb{P}$ .

From (2.2) we deduce that for any bounded, measurable  $f, g : \Omega \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_0(f(Qg)) = \mathbb{E}_0(g(Qf)), \quad (2.5)$$

where  $Qf : \Omega \rightarrow \mathbb{R}$  is the function

$$(Qf)(\omega) = \frac{1}{2d} \sum_{e: |e|=1} (\mathbf{1}_{\{\omega_e=1\}} f(\tau_e \omega) + \mathbf{1}_{\{\omega_e=0\}} f(\omega)). \quad (2.6)$$

Indeed, splitting the last sum into two terms, the second part reproduces exactly on both sides of (2.5). For the first part we apply (2.2) and note that averaging over  $e$

allows us to neglect the negative sign in front of  $e$  on the right-hand side. But (2.5) is the *definition* of reversibility and, setting  $f = 1$  and noting that  $Q1 = 1$ , we also get the stationarity of  $\mathbb{P}_0$ .  $\square$

Lemma 2.1 underlines our main reason to work primarily with the “lazy” walk. For the “agile” walk, to get a stationary law on environments, one has to weigh  $\mathbb{P}_0$  by the degree of the origin—a factor that would drag through the entire derivation.

## 2.2. Kipnis-Varadhan construction

Next we will adapt the construction of Kipnis and Varadhan [27] to the present situation. Let  $L^2 = L^2(\Omega, \mathcal{B}, \mathbb{P}_0)$  be the space of all Borel-measurable, square integrable functions on  $\Omega$ . Abusing the notation slightly, we will use “ $L^2$ ” both for  $\mathbb{R}$ -valued functions as well as  $\mathbb{R}^d$ -valued functions. We equip  $L^2$  with the inner product  $(f, g) = \mathbb{E}_0(fg)$ —with “ $fg$ ” interpreted as the dot product of  $f$  and  $g$  when these functions are vector-valued. Let  $Q$  be the operator defined by (2.6). Note that, when applied to a vector-valued function,  $Q$  acts like a scalar, i.e., independently on each component.

From (2.5) we know

$$(f, Qg) = (Qf, g) \quad (2.7)$$

and so  $Q$  is symmetric. An explicit calculation gives us

$$\begin{aligned} |(f, Qf)| &\leq \frac{1}{2d} \sum_{e: |e|=1} \left\{ |(f, \mathbf{1}_{\{\omega_e=1\}} f \circ \tau_e)| + (f, \mathbf{1}_{\{\omega_e=0\}} f) \right\} \\ &\leq \frac{1}{2d} \sum_{e: |e|=1} \left\{ |(f, \mathbf{1}_{\{\omega_e=1\}} f)|^{1/2} |(f, \mathbf{1}_{\{\omega_{-e}=1\}} f)|^{1/2} + (f, \mathbf{1}_{\{\omega_e=0\}} f) \right\} \\ &\leq \frac{1}{2d} \sum_{e: |e|=1} \left\{ (f, \mathbf{1}_{\{\omega_e=1\}} f) + (f, \mathbf{1}_{\{\omega_e=0\}} f) \right\} = (f, f) \end{aligned} \quad (2.8)$$

and so  $\|Q\|_{L^2} \leq 1$ . In particular,  $Q$  is self-adjoint and  $\text{spec}(Q) \subset [-1, 1]$ .

Let  $V: \Omega \rightarrow \mathbb{R}^d$  be the local drift at the origin, i.e.,

$$V(\omega) = \frac{1}{2d} \sum_{e: |e|=1} e \mathbf{1}_{\{\omega_e=1\}}. \quad (2.9)$$

(We will only be interested in  $V(\omega)$  for  $\omega \in \Omega_0$ , but that is of no consequence here.) Clearly, since  $V$  is bounded, we have  $V \in L^2$ . For each  $\epsilon > 0$ , let  $\psi_\epsilon: \Omega \rightarrow \mathbb{R}^d$  be the solution of

$$(1 + \epsilon - Q)\psi_\epsilon = V. \quad (2.10)$$

Since  $1 - Q$  is a non-negative operator,  $\psi_\epsilon$  is well-defined and  $\psi_\epsilon \in L^2$  for all  $\epsilon > 0$ . The following theorem is the core of the whole theory:

**Theorem 2.2.** *There is a function  $\chi: \mathbb{Z}^d \times \Omega_0 \rightarrow \mathbb{R}^d$  such that for every  $x \in \mathbb{Z}^d$ ,*

$$\lim_{\epsilon \downarrow 0} \mathbf{1}_{\{x \in \mathcal{C}_\infty\}} (\psi_\epsilon \circ \tau_x - \psi_\epsilon) = \chi(x, \cdot), \quad \text{in } L^2. \quad (2.11)$$

Moreover, the following properties hold:



(1) (Shift invariance) For  $\mathbb{P}_0$ -almost every  $\omega \in \Omega_0$ ,

$$\chi(x, \omega) - \chi(y, \omega) = \chi(x - y, \tau_y(\omega)) \quad (2.12)$$

holds for all  $x, y \in \mathcal{C}_\infty(\omega)$ .

(2) (Harmonicity) For  $\mathbb{P}_0$ -almost every  $\omega \in \Omega_0$ , the function

$$x \mapsto \chi(x, \omega) + x \quad (2.13)$$

is harmonic with respect to the transition probabilities (1.3–1.4).

(3) (Square integrability) There exists a constant  $C < \infty$  such that

$$\|[\chi(x + e, \cdot) - \chi(x, \cdot)] \mathbf{1}_{\{x \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_e=1\}} \circ \tau_x\|_2 < C \quad (2.14)$$

is true for all  $x \in \mathbb{Z}^d$  and all  $e$  with  $|e| = 1$ .

The rest of this section is spent on proving Theorem 2.2. The proof is based on spectral calculus and it closely follows the corresponding arguments from [27]. Alternative constructions invoke projection arguments, cf [30, 34].

### 2.3. Spectral calculations

Let  $\mu_V$  denote the spectral measure of  $Q: L^2 \rightarrow L^2$  associated with function  $V$ , i.e., for every bounded, continuous  $\Phi: [-1, 1] \rightarrow \mathbb{R}$ , we have

$$(V, \Phi(Q)V) = \int_{-1}^1 \Phi(\lambda) \mu_V(d\lambda). \quad (2.15)$$

(Since  $Q$  acts as a scalar,  $\mu_V$  is the sum of the “usual” spectral measures for the Cartesian components of  $V$ .) In the integral we used that, since  $\text{spec}(Q) \in [-1, 1]$ , the measure  $\mu_V$  is supported entirely in  $[-1, 1]$ . The first observation, made already by Kipnis and Varadhan, is stated as follows:

#### Lemma 2.3.

$$\int_{-1}^1 \frac{1}{1 - \lambda} \mu_V(d\lambda) < \infty. \quad (2.16)$$

*Proof.* With some caution concerning the infinite cluster, the proof is a combination of arguments right before Theorem 1.3 of [27] and those in the proof of Theorem 4.1 of [27]. Let  $f \in L^2$  be a bounded real-valued function and note that, by Lemma 2.1 and the symmetry of the sums,

$$\sum_{e: |e|=1} e \mathbb{E}_0(f \mathbf{1}_{\{\omega_e=1\}}) = \frac{1}{2} \sum_{e: |e|=1} e \mathbb{E}_0((f - f \circ \tau_e) \mathbf{1}_{\{\omega_e=1\}}). \quad (2.17)$$

Hence, for every  $a \in \mathbb{R}^d$  we get

$$\begin{aligned}
(f, a \cdot V) &= \frac{1}{2} \frac{1}{2d} \sum_{e: |e|=1} (e \cdot a) \mathbb{E}_0((f - f \circ \tau_e) \mathbf{1}_{\{\omega_e=1\}}) \\
&\leq \frac{1}{2} \left( \frac{1}{2d} \sum_{e: |e|=1} (e \cdot a)^2 \mathbb{P}(\omega_e = 1) \right)^{1/2} \\
&\quad \times \left( \frac{1}{2d} \sum_{e: |e|=1} \mathbb{E}_0((f - f \circ \tau_e)^2 \mathbf{1}_{\{\omega_e=1\}}) \right)^{1/2}.
\end{aligned} \tag{2.18}$$

The first term on the right-hand side equals a constant times  $|a|$ , while Lemma 2.1 allows us to rewrite the second term into

$$\begin{aligned}
&\frac{1}{2d} \sum_{e: |e|=1} \mathbb{E}_0((f - f \circ \tau_e)^2 \mathbf{1}_{\{\omega_e=1\}}) \\
&= 2 \frac{1}{2d} \sum_{e: |e|=1} \mathbb{E}_0(f(f - f \circ \tau_e) \mathbf{1}_{\{\omega_e=1\}}) = 2(f, (1 - Q)f).
\end{aligned} \tag{2.19}$$

We thus get that there exists a constant  $C_0 < \infty$  such that for all bounded  $f \in L^2$ ,

$$|(f, a \cdot V)| \leq C_0 |a| (f, (1 - Q)f)^{1/2}. \tag{2.20}$$

Applying (2.20) for  $f$  of the form  $f = a \cdot \Psi(Q)V$ , summing  $a$  over coordinate vectors in  $\mathbb{R}^d$  and invoking (2.15), we find that for every bounded continuous  $\Psi: [-1, 1] \rightarrow \mathbb{R}$  and  $C = C_0 \sqrt{d}$ ,

$$\left| \int \Psi(\lambda) \mu_V(d\lambda) \right| \leq C \left( \int (1 - \lambda) \Psi(\lambda)^2 \mu_V(d\lambda) \right)^{1/2}. \tag{2.21}$$

Substituting  $\Psi_\epsilon(\lambda) = (1/\epsilon) \wedge \frac{1}{1-\lambda}$  for  $\Psi$  and noting that  $(1 - \lambda) \Psi_\epsilon(\lambda) \leq 1$ , we get

$$\int \Psi_\epsilon(\lambda) \mu_V(d\lambda) \leq C \left( \int \Psi_\epsilon(\lambda) \mu_V(d\lambda) \right)^{1/2} \tag{2.22}$$

and so

$$\int \Psi_\epsilon(\lambda) \mu_V(d\lambda) \leq C^2. \tag{2.23}$$

The Monotone Convergence Theorem now implies

$$\int \frac{1}{1-\lambda} \mu_V(d\lambda) = \sup_{\epsilon > 0} \int \Psi_\epsilon(\lambda) \mu_V(d\lambda) \leq C^2 < \infty, \tag{2.24}$$

proving the desired claim.  $\square$

Using spectral calculus we will now prove:

**Lemma 2.4.** *Let  $\psi_\epsilon$  be as defined in (2.10). Then*

$$\lim_{\epsilon \downarrow 0} \epsilon \|\psi_\epsilon\|_2^2 = 0. \quad (2.25)$$

Moreover, for  $e$  with  $|e| = 1$  let  $G_e^{(\epsilon)} = \mathbf{1}_{\{0 \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_e=1\}} (\psi_\epsilon \circ \tau_e - \psi_\epsilon)$ . Then for all  $x \in \mathbb{Z}^d$  and all  $e$  with  $|e| = 1$ ,

$$\lim_{\epsilon_1, \epsilon_2 \downarrow 0} \|G_e^{(\epsilon_1)} \circ \tau_x - G_e^{(\epsilon_2)} \circ \tau_x\|_2 = 0 \quad (2.26)$$

*Proof.* The main ideas are again taken more or less directly from the proof of Theorem 1.3 in [27]; some caution is necessary regarding the containment in the infinite cluster in the proof of (2.26). By the definition of  $\psi_\epsilon$ ,

$$\epsilon \|\psi_\epsilon\|_2^2 = \int_{-1}^1 \frac{\epsilon}{(1 + \epsilon - \lambda)^2} \mu_V(d\lambda). \quad (2.27)$$

The integrand is dominated by  $\frac{1}{1-\lambda}$  and tends to zero as  $\epsilon \downarrow 0$  for every  $\lambda$  in the support of  $\mu_V$ . Then (2.25) follows by the Dominated Convergence Theorem.

The second part of the claim is proved similarly: First we get rid of the  $x$ -dependence by noting that, due to the fact that  $G_e^{(\epsilon)} \circ \tau_x \neq 0$  enforces  $x \in \mathcal{C}_\infty$ , the translation invariance of  $\mathbb{P}$  implies

$$\|G_e^{(\epsilon_1)} \circ \tau_x - G_e^{(\epsilon_2)} \circ \tau_x\|_2 \leq \|G_e^{(\epsilon_1)} - G_e^{(\epsilon_2)}\|_2. \quad (2.28)$$

Next we square the right-hand side and average over all  $e$ . Using that  $G_e \neq 0$  also enforces  $\omega_e = 1$  and applying (2.17), we thus get

$$\frac{1}{2d} \sum_{e: |e|=1} \|G_e^{(\epsilon_1)} - G_e^{(\epsilon_2)}\|_2^2 = 2(\psi_{\epsilon_1} - \psi_{\epsilon_2}, (1 - Q)(\psi_{\epsilon_1} - \psi_{\epsilon_2})). \quad (2.29)$$

Now we calculate

$$\begin{aligned} & (\psi_{\epsilon_1} - \psi_{\epsilon_2}, (1 - Q)(\psi_{\epsilon_1} - \psi_{\epsilon_2})) \\ &= \int_{-1}^1 \frac{(\epsilon_1 - \epsilon_2)^2 (1 - \lambda)}{(1 + \epsilon_1 - \lambda)^2 (1 + \epsilon_2 - \lambda)^2} \mu_V(d\lambda). \end{aligned} \quad (2.30)$$

The integrand is again bounded by  $\frac{1}{1-\lambda}$ , for all  $\epsilon_1, \epsilon_2 > 0$ , and it tends to zero as  $\epsilon_1, \epsilon_2 \downarrow 0$ . The claim follows by the Dominated Convergence Theorem.  $\square$

Now we are ready to prove Theorem 2.2:

*Proof of Theorem 2.2.* Let  $G_e^{(\epsilon)} \circ \tau_x$  be as in Lemma 2.4. Using (2.26) we know that  $G_e^{(\epsilon)} \circ \tau_x$  converges in  $L^2$  as  $\epsilon \downarrow 0$ . We denote  $G_{x,x+e} = \lim_{\epsilon \downarrow 0} G_e^{(\epsilon)} \circ \tau_x$ . Since  $G_e^{(\epsilon)} \circ \tau_x$  is a gradient field on  $\mathcal{C}_\infty$ , we have  $G_{x,x+e}(\omega) + G_{x+e,x}(\omega) = 0$  and, more generally,  $\sum_{k=0}^n G_{x_k, x_{k+1}} = 0$  whenever  $(x_0, \dots, x_n)$  is a closed loop on  $\mathcal{C}_\infty$ . Thus, we may define

$$\chi(x, \omega) \stackrel{\text{def}}{=} \sum_{k=0}^{n-1} G_{x_k, x_{k+1}}(\omega), \quad (2.31)$$

where  $(x_0, x_1, \dots, x_n)$  is a nearest-neighbor path on  $\mathcal{C}_\infty(\omega)$  connecting  $x_0 = 0$  to  $x_n = x$ . By the above “loop” conditions, the definition is independent of this path for almost every  $\omega \in \{x \in \mathcal{C}_\infty\}$ . The shift invariance (2.12) now follows from this definition and  $G_{x,x+e} = G_{0,e} \circ \tau_x$ .

In light of shift invariance, to prove the harmonicity of  $x \mapsto x + \chi(x, \omega)$  it suffices to show that, almost surely,

$$\frac{1}{2d} \sum_{e: |e|=1} [\chi(0, \cdot) - \chi(e, \cdot)] \mathbf{1}_{\{\omega_e=1\}} = V. \quad (2.32)$$

Since  $\chi(e, \cdot) - \chi(0, \cdot) = G_{0,e}$ , the left hand side is the  $\epsilon \downarrow 0$  limit of

$$\frac{1}{2d} \sum_{e: |e|=1} [\psi_\epsilon - \psi_\epsilon \circ \tau_e] \mathbf{1}_{\{\omega_e=1\}} = (1 - Q)\psi_\epsilon. \quad (2.33)$$

The definition of  $\psi_\epsilon$  tells us that  $(1 - Q)\psi_\epsilon = -\epsilon\psi_\epsilon + V$ . From here we get (2.32) by recalling that  $\epsilon\psi_\epsilon(\omega)$  tends to zero in  $L^2$ .

To prove the square integrability in part (3) we note that, by the construction of the corrector,

$$[\chi(x+e, \cdot) - \chi(x, \cdot)] \mathbf{1}_{\{x \in \mathcal{C}_\infty\}} \mathbf{1}_{\{\omega_e=1\}} \circ \tau_x = G_{x,x+e}. \quad (2.34)$$

But  $G_{x,x+e}$  is the  $L^2$ -limit of  $L^2$ -functions  $G_e^{(\epsilon)} \circ \tau_x$  whose  $L^2$ -norm is bounded by that of  $G_e^{(\epsilon)}$ . Hence (2.14) follows with  $C = \max_{e: |e|=1} \|G_{0,e}\|_2$ .  $\square$

### 3. Ergodic-theory input

Here we will establish some basic claims whose common feature is the use of ergodic theory. Modulo some care for the containment in the infinite cluster, all of these results are quite standard and their proofs (cf Sect. 3.2) may be skipped on a first reading. Readers interested only in the principal conclusions of this section should focus their attention on Theorems 3.1 and 3.2.

#### 3.1. Statements

Our first result concerns the convergence of ergodic averages for the Markov chain on environments. The claim that will suffice for our later needs is as follows:

**Theorem 3.1.** *Let  $f \in L^1(\Omega, \mathcal{B}, \mathbb{P}_0)$ . Then for  $\mathbb{P}_0$ -almost all  $\omega \in \Omega$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau_{X_k}(\omega) = \mathbb{E}_0(f), \quad P_{0,\omega}\text{-almost surely.} \quad (3.1)$$

Similarly, if  $f: \Omega \times \Omega \rightarrow \mathbb{R}$  is measurable with  $\mathbb{E}_0(E_{0,\omega}|f(\omega, \tau_{X_1}\omega)|) < \infty$ , then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\tau_{X_k}\omega, \tau_{X_{k+1}}\omega) = \mathbb{E}_0(E_{0,\omega}(f(\omega, \tau_{X_1}\omega))) \quad (3.2)$$

for  $\mathbb{P}_0$ -almost all  $\omega$  and  $P_{0,\omega}$ -almost all trajectories of  $(X_k)_{k \geq 0}$ .

The next principal result of this section will be the ergodicity of the “induced shift” on  $\Omega_0$ . To define this concept, let  $e$  be a vector with  $|e| = 1$  and, for every  $\omega \in \Omega_0$ , let

$$n(\omega) = \min\{k > 0: ke \in \mathcal{C}_\infty(\omega)\}. \quad (3.3)$$

By Birkhoff’s Ergodic Theorem we know that  $\{k > 0: ke \in \mathcal{C}_\infty\}$  has positive density in  $\mathbb{N}$  and so  $n(\omega) < \infty$  almost surely. Therefore we can define the map  $\sigma_e: \Omega_0 \rightarrow \Omega_0$  by

$$\sigma_e(\omega) = \tau_{n(\omega)e} \omega. \quad (3.4)$$

We call  $\sigma_e$  the *induced shift*. Then we claim:

**Theorem 3.2.** *For every  $e$  with  $|e| = 1$ , the induced shift  $\sigma_e: \Omega_0 \rightarrow \Omega_0$  is  $\mathbb{P}_0$ -preserving and ergodic with respect to  $\mathbb{P}_0$ .*

Both theorems will follow once we establish of ergodicity of the Markov chain on environments (see Proposition 3.5). For finite-state (irreducible) Markov chains the proof of ergodicity is a standard textbook material (cf [36, page 51]), but our state space is somewhat large and so alternative arguments are necessary. Since we could not find appropriate versions of all needed claims in the literature, we include complete proofs.

### 3.2. Proofs

We begin by Theorem 3.2 which will follow from a more general statement, Lemma 3.3, below. Let  $(\mathcal{X}, \mathcal{X}, \mu)$  be a probability space, and let  $T: \mathcal{X} \rightarrow \mathcal{X}$  be invertible, measure preserving and ergodic with respect to  $\mu$ . Let  $A \in \mathcal{X}$  be of positive measure, and define  $n: A \rightarrow \mathbb{N} \cup \{\infty\}$  by

$$n(x) = \min\{k > 0: T^k(x) \in A\}. \quad (3.5)$$

The Poincaré Recurrence Theorem (cf [36, Sect. 2.3]) tells us that  $n(x) < \infty$  almost surely. Therefore we can define, up to a set of measure zero, the map  $S: A \rightarrow A$  by

$$S(x) = T^{n(x)}(x), \quad x \in A. \quad (3.6)$$

Then we have:

**Lemma 3.3.**  *$S$  is measure preserving and ergodic with respect to  $\mu(\cdot|A)$ . It is also almost surely invertible with respect to the same measure.*

*Proof.* (1)  $S$  is measure preserving: For  $j \geq 1$ , let  $A_j = \{x \in A: n(x) = j\}$ . Then the  $A_j$ ’s are disjoint and  $\mu(A \setminus \bigcup_{j \geq 1} A_j) = 0$ . First we show that

$$i \neq j \quad \Rightarrow \quad S(A_i) \cap S(A_j) = \emptyset. \quad (3.7)$$

To do this, we use the fact that  $T$  is invertible. Indeed, if  $x \in S(A_i) \cap S(A_j)$  for  $1 \leq i < j$ , then  $x = T^i(y) = T^j(z)$  for some  $y, z \in A$  with  $n(y) = i$  and  $n(z) = j$ . But the fact that  $T$  is invertible implies that  $y = T^{j-i}(z)$ , which means  $n(z) \leq j - i < j$ , a contradiction. To see that  $S$  is measure preserving, we

note that the restriction of  $S$  to  $A_j$  is  $T^j$ , which is measure preserving. Hence,  $S$  is measure preserving on  $A_j$  and, by (3.7), on the disjoint union  $\bigcup_{j \geq 1} A_j$  as well.

(2)  $S$  is almost surely invertible:  $S^{-1}(\{x\}) \cap \{S \text{ is well defined}\}$  is a one-point set by the fact that  $T$  is itself invertible.

(3)  $S$  is ergodic: Let  $B \in \mathcal{X}$  be such that  $B \subseteq A$  and  $0 < \mu(B) < \mu(A)$ . Assume that  $B$  is  $S$ -invariant. Then  $S^n(x) \notin A \setminus B$  for all  $x \in B$  and all  $n \geq 1$ . This means that for every  $x \in B$  and every  $k \geq 1$  such that  $T^k(x) \in A$ , we have  $T^k(x) \notin A \setminus B$ . It follows that  $C = \bigcup_{k \geq 1} T^k(B)$  is (almost-surely)  $T$ -invariant and  $\mu(C) \in (0, 1)$ , a contradiction with the ergodicity of  $T$ .  $\square$

*Proof of Theorem 3.2.* We know that the shift  $\tau_e$  is invertible, measure preserving and ergodic with respect to  $\mathbb{P}$ . By Lemma 3.3 the induced shift  $\sigma_e: \Omega_0 \rightarrow \Omega_0$  is  $\mathbb{P}_0$ -preserving, almost-surely invertible and ergodic with respect to  $\mathbb{P}_0$ .  $\square$

In the present circumstances, Theorem 3.2 has one important consequence:

**Lemma 3.4.** *Let  $B \in \mathcal{B}$  be a subset of  $\Omega_0$  such that for almost all  $\omega \in B$ ,*

$$P_{0,\omega}(\tau_{X_1}\omega \in B) = 1. \quad (3.8)$$

*Then  $B$  is a zero-one event under  $\mathbb{P}_0$ .*

*Proof.* The Markov property and (3.8) imply that  $P_{0,\omega}(\tau_{X_n}\omega \in B) = 1$  for all  $n \geq 1$  and  $\mathbb{P}_0$ -almost every  $\omega \in B$ . We claim that  $\sigma_e(\omega) \in B$  for  $\mathbb{P}_0$ -almost every  $\omega \in B$ . Indeed, let  $\omega \in B$  be such that  $\tau_{X_n}\omega \in B$  for all  $n \geq 1$ ,  $P_{0,\omega}$ -almost surely. Let  $n(\omega)$  be as in (3.3) and note that we have  $n(\omega)e \in \mathcal{C}_\infty$ . By the uniqueness of the infinite cluster, there is a path of finite length connecting 0 and  $n(\omega)e$ . If  $\ell$  is the length of this path, we have  $P_{0,\omega}(X_\ell = n(\omega)e) > 0$ . This means that  $\sigma_e(\omega) = \tau_{n(\omega)e}(\omega) \in B$ , i.e.,  $B$  is almost surely  $\sigma_e$ -invariant. By the ergodicity of the induced shift,  $B$  is a zero-one event.  $\square$

Our next goal will be to prove that the Markov chain on environments is ergodic. Let  $\mathcal{X} = \Omega^{\mathbb{Z}}$  and define  $\mathcal{X}$  to be the product  $\sigma$ -algebra on  $\mathcal{X}$ ;  $\mathcal{X} = \mathcal{B}^{\otimes \mathbb{Z}}$ . The space  $\mathcal{X}$  is a space of two-sided sequences  $(\dots, \omega_{-1}, \omega_0, \omega_1, \dots)$ —the trajectories of the Markov chain on environments. (Note that the index on  $\omega$  is an index in the sequence which is unrelated to the value of the configuration at a point.) Let  $\mu$  be the measure on  $(\mathcal{X}, \mathcal{X})$  such that for any  $B \in \mathcal{B}^{2n+1}$ ,

$$\begin{aligned} \mu((\omega_{-n}, \dots, \omega_n) \in B) \\ = \int_B \mathbb{P}_0(d\omega_{-n}) Q(\omega_{-n}, d\omega_{-n+1}) \cdots Q(\omega_{n-1}, d\omega_n), \end{aligned} \quad (3.9)$$

where  $Q$  is the Markov kernel defined in (2.1). (Since  $\mathbb{P}_0$  is preserved by  $Q$ , these finite-dimensional measures are consistent and  $\mu$  exists and is unique by Kolmogorov's Theorem.) Clearly,  $(\tau_{X_k}(\omega))_{k \geq 0}$  has the same law in  $\mathbb{E}_0(P_{0,\omega}(\cdot))$  as  $(\omega_0, \omega_1, \dots)$  has in  $\mu$ . Let  $T: \mathcal{X} \rightarrow \mathcal{X}$  be the shift defined by  $(T\omega)_n = \omega_{n+1}$ . Then  $T$  is measure preserving.

**Proposition 3.5.**  *$T$  is ergodic with respect to  $\mu$ .*

*Proof.* Let  $E_\mu$  denote expectation with respect to  $\mu$ . Pick  $A \subseteq \mathcal{X}$  that is measurable and  $T$ -invariant. We need to show that

$$\mu(A) \in \{0, 1\}. \quad (3.10)$$

Let  $f: \Omega \rightarrow \mathbb{R}$  be defined as  $f(\omega_0) = E_\mu(\mathbf{1}_A | \omega_0)$ . First we claim that  $f = \mathbf{1}_A$  almost surely. Indeed, since  $A$  is  $T$ -invariant, there exist  $A_+ \in \sigma(\omega_k: k > 0)$  and  $A_- \in \sigma(\omega_k: k < 0)$  such that  $A$  and  $A_\pm$  differ only by null sets from one another. (This follows by approximation of  $A$  by finite-dimensional events and using the  $T$ -invariance of  $A$ .) Now conditional on  $\omega_0$ , the event  $A_+$  is independent of  $\sigma(\omega_k: k < 0)$  and so Lévy's Martingale Convergence Theorem gives us

$$\begin{aligned} E_\mu(\mathbf{1}_A | \omega_0) &= E_\mu(\mathbf{1}_{A_+} | \omega_0) = E_\mu(\mathbf{1}_{A_+} | \omega_0, \omega_{-1}, \dots, \omega_{-n}) \\ &= E_\mu(\mathbf{1}_{A_-} | \omega_0, \omega_{-1}, \dots, \omega_{-n}) \xrightarrow{n \rightarrow \infty} \mathbf{1}_{A_-} = \mathbf{1}_A, \end{aligned} \quad (3.11)$$

with all equalities valid  $\mu$ -almost surely.

Next let  $B \subset \Omega$  be defined by  $B = \{\omega_0: f(\omega_0) = 1\}$ . Clearly,  $B$  is  $\mathcal{B}$ -measurable and, since the  $\omega_0$ -marginal of  $\mu$  is  $\mathbb{P}_0$ ,

$$\mu(A) = E_\mu(f) = \mathbb{P}_0(B). \quad (3.12)$$

Hence, to prove (3.10), we need to show that

$$\mathbb{P}_0(B) \in \{0, 1\}. \quad (3.13)$$

But  $A$  is  $T$ -invariant and so, up to sets of measure zero, if  $\omega_0 \in B$  then  $\omega_1 \in B$ . This means that  $B$  satisfies condition (3.8) of Lemma 3.4 and so (3.13) holds.  $\square$

Now we can finally prove Theorem 3.1:

*Proof of Theorem 3.1.* Recall that  $(\tau_{X_k}(\omega))_{k \geq 0}$  has the same law in  $\mathbb{E}_0(P_{0,\omega}(\cdot))$  as  $(\omega_0, \omega_1, \dots)$  has in  $\mu$ . Hence, if  $g(\dots, \omega_{-1}, \omega_0, \omega_1, \dots) = f(\omega_0)$  then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{\infty} f \circ \tau_{X_k} \stackrel{\mathcal{D}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{\infty} g \circ T^k. \quad (3.14)$$

The latter limit exists by Birkhoff's Ergodic Theorem and (by Proposition 3.5) equals  $E_\mu(g) = \mathbb{E}_0(f)$  almost surely. The second part is proved analogously.  $\square$

#### 4. Sublinearity along coordinate directions

Equipped with the tools from the previous two sections, we can start addressing the main problem of our proof: the sublinearity of the corrector. Here we will prove the corresponding claim along the coordinate directions in  $\mathbb{Z}^d$ .

Fix  $e$  with  $|e| = 1$  and let  $n(\omega)$  be as defined in (3.3). Define a sequence  $n_k(\omega)$  inductively by  $n_1(\omega) = n(\omega)$  and  $n_{k+1}(\omega) = n_k(\sigma_e(\omega))$ . The numbers  $(n_k)$ , which are well-defined and finite on a set of full  $\mathbb{P}_0$ -measure, represent the successive "arrivals" of  $\mathcal{C}_\infty$  to the positive part of the coordinate axis in direction  $e$ . Let  $\chi$  be the corrector defined in Theorem 2.2. The main goal of this section is to prove the following theorem:

**Theorem 4.1.** For  $\mathbb{P}_0$ -almost all  $\omega \in \Omega_0$ ,

$$\lim_{k \rightarrow \infty} \frac{\chi(n_k(\omega)e, \omega)}{k} = 0. \quad (4.1)$$

The proof is based on the following facts about the moments of  $\chi(n_k(\omega)e, \omega)$ :

**Proposition 4.2.** Abbreviate  $v_e = v_e(\omega) = n_1(\omega)e$ . Then

- (1)  $\mathbb{E}_0(|\chi(v_e, \cdot)|) < \infty$ .
- (2)  $\mathbb{E}_0(\chi(v_e, \cdot)) = 0$ .

The proof of this proposition will in turn be based on a bound on the tails of the length of the shortest path connecting the origin to  $v_e$ . We begin by showing that  $|v_e|$  has exponential tails:

**Lemma 4.3.** For each  $p > p_c$  there exists a constant  $a = a(p) > 0$  such that for all  $e$  with  $|e| = 1$ ,

$$\mathbb{P}_0(|v_e| > n) \leq e^{-an}, \quad n \geq 1. \quad (4.2)$$

*Proof.* The proof uses a different argument in  $d = 2$  and  $d \geq 3$ . In  $d \geq 3$ , we will use the fact that the slab-percolation threshold coincides with  $p_c$ , as was proved by Grimmett and Marstrand [21]. Indeed, given  $p > p_c$ , let  $K \geq 1$  be so large that  $\mathbb{Z}^{d-1} \times \{1, \dots, K\}$  contains an infinite cluster almost surely. By the uniqueness of the percolation cluster in  $\mathbb{Z}^d$ , this slab-cluster is almost surely a subset of  $\mathcal{C}_\infty$ . Our bound in (4.2) is derived as follows: Let  $A_K$  be the event that at least one of the sites in  $\{je: j = 1, \dots, K\}$  is contained in the infinite connected component in  $\mathbb{Z}^{d-1} \times \{1, \dots, K\}$ . Then  $\{|v_e| \geq Kn\} \cap \{0 \in \mathcal{C}_\infty\} \subset \bigcap_{\ell \leq n} \tau_{\ell K e}(A)$ . Since the events  $\tau_{\ell K e}(A)$ ,  $\ell = 1, \dots, n$ , are independent, letting  $p_K = \mathbb{P}(A_K)$  we have

$$\mathbb{P}(|v_e| \geq Kn, 0 \in \mathcal{C}_\infty) \leq (1 - p_K)^n, \quad n \geq 1. \quad (4.3)$$

From here (4.2) follows by choosing  $a$  appropriately.

In dimension  $d = 2$ , we will instead use a duality argument. Let  $\Lambda_n$  be the box  $\{1, \dots, n\} \times \{1, \dots, n\}$ . On  $\{|v_e| \geq n\} \cap \{0 \in \mathcal{C}_\infty\}$ , none of the boundary sites  $\{je: j = 1, \dots, n\}$  are in  $\mathcal{C}_\infty$ . So either at least one of these sites is in a finite component of size larger than  $n$  or there exists a dual crossing of  $\Lambda_n$  in the direction of  $e$ . By the exponential decay of truncated connectivities (Theorem 8.18 of Grimmett [19]) and dual connectivities (Theorem 6.75 of Grimmett [19]), the probability of each of these events decays exponentially with  $n$ .  $\square$

Our next lemma provides the requisite tail bound for the length of the shortest path between the origin and  $v_e$ :

**Lemma 4.4.** Let  $L = L(\omega)$  be the length of the shortest occupied path from 0 to  $v_e$ . Then there exist a constant  $C < \infty$  and  $a > 0$  such that for every  $n \geq 1$ ,

$$\mathbb{P}_0(L > n) < Ce^{-an}. \quad (4.4)$$



*Proof.* Let  $d_\omega(0, x)$  be the length of the shortest path from 0 to  $x$  in configuration  $\omega$ . Pick  $\epsilon > 0$  such that  $\epsilon n$  is an integer. Then

$$\{L > n\} \subset \{|v_e| \geq \epsilon n\} \cup \bigcup_{k=1}^{\epsilon n} \{d_\omega(0, ke) > n; 0, ke \in \mathcal{C}_\infty\}. \quad (4.5)$$

In light of Lemma 4.3, the claim will follow once we show that the probability of all events in the giant union on the right-hand side is bounded by  $e^{-a'n}$  with some  $a' > 0$  (independently of  $k$ ).

We will use the following large-deviation result from Theorem 1.1 of Antal and Pisztora [2]: There exist constants  $a, \rho < \infty$  such that

$$\mathbb{P}(d_\omega(0, x) > \rho|x|) \leq e^{-a|x|} \quad (4.6)$$

once  $|x|$  is sufficiently large. Unfortunately, we cannot use this bound in (4.5) directly, because  $ke$  can be arbitrarily close to 0 (in  $\ell^\infty$  distance on  $\mathbb{Z}^d$ ). To circumvent this problem, let  $w_e$  be the site  $-me$  such that  $m = \min\{m' > \epsilon n : -m'e \in \mathcal{C}_\infty\}$  and let  $A_{x,y} = \{d_\omega(x, y) \geq n/2, x, y \in \mathcal{C}_\infty\}$ . Then, on  $\{d_\omega(0, x) > n\}$ , either  $|w_e| > 2\epsilon n$  or at least one site “between”  $-2\epsilon ne$  and  $-\epsilon ne$  is connected to either 0 or  $ke$  by a path longer than  $n/2$ . Since on  $\{|w_e| > 2\epsilon n\}$  we must have  $|v_{-e} \circ \sigma_{-e}^m| > \epsilon n$  for at least one  $m = 1, \dots, \epsilon n$ , we have

$$\begin{aligned} & \{d_\omega(0, ke) > n; 0, ke \in \mathcal{C}_\infty\} \\ & \subset \left( \bigcup_{m=1}^{\epsilon n} \sigma_{-e}^m(\{|v_{-e}| \geq \epsilon n\}) \cup \bigcup_{\epsilon n \leq \ell \leq 2\epsilon n} (A_{0, -\ell e} \cup A_{ke, -\ell e}) \right). \end{aligned} \quad (4.7)$$

Now all events in the first giant union have the same probability, which is exponentially small by Lemma 4.3. As to the second union, by (4.6) we know that

$$\mathbb{P}_0(A_{0, -\ell e}) \leq e^{-a\ell} \leq e^{-a\epsilon n} \quad (4.8)$$

whenever  $\epsilon$  is so small that  $4\epsilon\rho \leq 1$ , and a similar bound holds for  $A_{ke, -\ell e}$  as well (except that here we need  $6\epsilon\rho \leq 1$ ). The various unions then contribute a linear factor in  $n$ , which is absorbed into the exponential once  $n$  is sufficiently large.  $\square$

It is possible that a proper merge of the arguments in the previous two proofs might yield the same result without relying on Antal and Pisztora’s bound (4.6). (Indeed, the main other “external” ingredient of our proofs is Grimmett and Marstrand’s paper [21] which lies at the core of [2] as well.) However, we find the argument using (4.6) conceptually cleaner and so we are content with the present, even though not necessarily optimal, proof.

Next we state a trivial, but interesting technical lemma:

**Lemma 4.5.** *Let  $p > 1$  and  $r \in [1, p)$ . Suppose that  $X_1, X_2, \dots$  are random variables such that  $\sup_{j \geq 1} \|X_j\|_p < \infty$  and let  $N$  be a random variable taking values in positive integers such that  $N \in L^s$  for some  $s$  satisfying*

$$s > r \frac{1 + 1/p}{1 - r/p}. \quad (4.9)$$

Then  $\sum_{j=1}^N X_j \in L^r$ . Explicitly,

$$\left\| \sum_{j=1}^N X_j \right\|_r \leq C \left( \sup_{j \geq 1} \|X_j\|_p \right) (\|N\|_s)^{s[1/r - 1/p]}, \quad (4.10)$$

where  $C$  is a finite constant depending only on  $p, r$  and  $s$ .

*Proof.* Let us define  $q \in (1, \infty)$  by  $r/p + 1/q = 1$ . From the Hölder inequality and the uniform bound on  $\|X_j\|_p$  we get

$$\begin{aligned} E \left| \sum_{j=1}^N X_j \right|^r &= \sum_{n \geq 1} E \left( \left| \sum_{j=1}^n X_j \right|^r \mathbf{1}_{\{N=n\}} \right) \\ &\leq \sum_{n \geq 1} \left\| \sum_{j=1}^n X_j \right\|_p^r P(N = n)^{1/q} \\ &\leq \left( \sup_{j \geq 1} \|X_j\|_p \right)^r \sum_{n \geq 1} n^r P(N = n)^{1/q}. \end{aligned} \quad (4.11)$$

Under the assumption that  $N$  has  $s$  moments, we get

$$\sum_{n \geq 1} n^r P(N = n)^{1/q} \leq \left( \sum_{n \geq 1} n^{(r-s/q)p/r} \right)^{r/p} (E(N^s))^{1/q} \quad (4.12)$$

by invoking the Hölder inequality one more time. The first term on the right-hand side is finite whenever  $s$  obeys the bound (4.9).  $\square$

*Proof of Proposition 4.2.* Let  $\chi(x, \omega)$  be the corrector. By Theorem 2.2, on the set  $\{x \in \mathcal{C}_\infty\}$ ,  $\chi(x, \cdot)$  is an  $L^2$ -limit of functions  $\chi_\epsilon(x, \cdot) = \psi_\epsilon \circ \tau_x - \psi_\epsilon$ , as  $\epsilon \downarrow 0$ . To prove that  $\chi(v_e, \cdot) \in L^1$ , recall the notation  $G_e^{(\epsilon)}$  from Lemma 2.4 and let—as in Lemma 4.4— $L = L(\omega)$  be the length of the shortest path from 0 to  $v_e$ . Then

$$|\chi_\epsilon(v_e, \omega)| \leq \sum_{x: |x|_\infty \leq L(\omega)} \sum_{e: |e|=1} |G_e^{(\epsilon)} \circ \tau_x(\omega)|. \quad (4.13)$$

But Theorem 2.2 ensures that  $\|G_e^{(\epsilon)} \circ \tau_x\|_2 \leq \|G_e^{(\epsilon)}\|_2 < C$  for all  $x$  and  $e$  and all  $\epsilon > 0$ , while the number of terms in the sum does not exceed  $N(\omega) = 2d(2L(\omega) + 1)^d$ . By Lemma 4.4,  $N$  has all moments and so, by Lemma 4.5,  $\sup_{\epsilon > 0} \|\chi_\epsilon(v_e, \cdot)\|_r < \infty$  for all  $r \in [1, 2)$ . In particular,  $\chi(v_e, \cdot) \in L^1$ .

In order to prove part (2), we first note that a uniform bound on  $L^r$ -norm of  $\chi_\epsilon(v_e, \cdot)$  for some  $r > 1$  implies that the family  $\{\chi_\epsilon(v_e, \cdot)\}_{\epsilon > 0}$  is uniformly integrable. Since  $\chi_\epsilon(v_e, \cdot) \rightarrow \chi(v_e, \cdot)$  in probability,  $\chi_\epsilon(v_e, \cdot) \rightarrow \chi(v_e, \cdot)$  in  $L^1$  and it thus suffices to prove

$$\mathbb{E}_0(\chi_\epsilon(v_e, \cdot)) = 0, \quad \epsilon > 0. \quad (4.14)$$

This is implied by Theorem 3.2 and the fact  $\chi_\epsilon(v_e, \cdot) = \psi_\epsilon \circ \sigma_e - \psi_\epsilon$  with  $\psi_\epsilon$  absolutely integrable.  $\square$

*Proof of Theorem 4.1.* Let  $f(\omega) = \chi(n_1(\omega)e, \omega)$ , and let  $\sigma_e$  be the induced shift in the direction of  $e$ . Then we can write

$$\chi(n_k(\omega)e, \omega) = \sum_{\ell=0}^{k-1} f \circ \sigma_e^\ell(\omega). \quad (4.15)$$

By Proposition 4.2 we have  $f \in L^1$  and  $\mathbb{E}_0(f) = 0$ . Since Theorem 3.2 ensures that  $\sigma_e$  is  $\mathbb{P}_0$ -preserving and ergodic, the claim follows from Birkhoff's Ergodic Theorem.  $\square$

## 5. Sublinearity everywhere

Here we will prove the principal technical estimates of this work. The level of control is different in  $d = 2$  and  $d \geq 3$ , so we treat these cases separately. (Notwithstanding, the  $d \geq 3$  proof applies in  $d = 2$  as well.)

### 5.1. Sublinearity in two dimensions

We begin with an estimate of the corrector in large boxes in  $\mathbb{Z}^2$ :

**Theorem 5.1.** *Let  $d = 2$  and let  $\chi$  be the corrector defined in Theorem 2.2. Then for  $\mathbb{P}_0$ -almost every  $\omega \in \Omega_0$ ,*

$$\lim_{n \rightarrow \infty} \max_{\substack{x \in \mathcal{C}_\infty(\omega) \\ |x|_\infty \leq n}} \frac{|\chi(x, \omega)|}{n} = 0. \quad (5.1)$$

The proof will be based on the following concept:

**Definition 5.2.** *Given  $K > 0$  and  $\epsilon > 0$ , we say that a site  $x \in \mathbb{Z}^d$  is  $K, \epsilon$ -good (or just good) in configuration  $\omega \in \Omega$  if  $x \in \mathcal{C}_\infty(\omega)$  and*

$$|\chi(y, \omega) - \chi(x, \omega)| < K + \epsilon|x - y| \quad (5.2)$$

*holds for every  $y \in \mathcal{C}_\infty(\omega)$  of the form  $y = \ell e$ , where  $\ell \in \mathbb{Z}$  and  $e$  is a unit coordinate vector. We will use  $\mathcal{G}_{K, \epsilon} = \mathcal{G}_{K, \epsilon}(\omega)$  to denote the set of  $K, \epsilon$ -good sites in configuration  $\omega$ .*

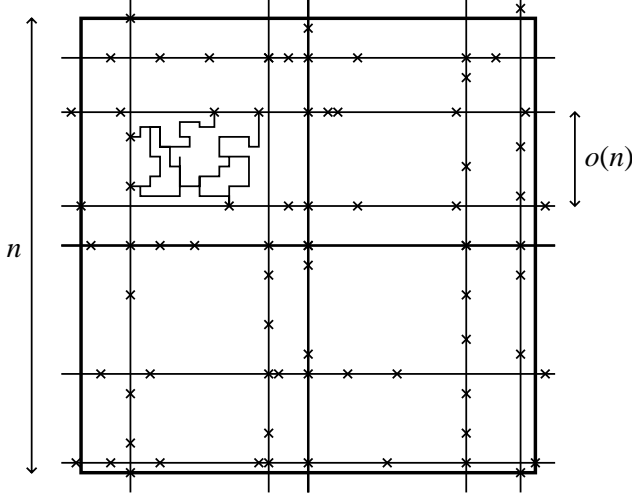
On the basis of Theorem 4.1 it is clear that for each  $\epsilon > 0$  there exists a  $K < \infty$  such that the  $\mathbb{P}_0(0 \in \mathcal{G}_{K, \epsilon}) > 0$ . Our first goal is to estimate the size of the largest interval free of good points in blocks  $[-n, n]$  on the coordinate axes:

**Lemma 5.3.** *Let  $e$  be one of the principal lattice vectors in  $\mathbb{Z}^2$  and, given  $\epsilon > 0$ , let  $K$  be so large that  $\mathbb{P}_0(0 \in \mathcal{G}_{K, \epsilon}) > 0$ . For all  $n \geq 1$  and  $\omega \in \Omega$ , let  $y_0 < \dots < y_r$  be the ordered set of all integers from  $[-n, n]$  such that  $y_i e \in \mathcal{G}_{K, \epsilon}(\omega)$ . Let*

$$\Delta_n(\omega) = \max_{j=1, \dots, r} (y_j - y_{j-1}). \quad (5.3)$$

*(If no such  $y_i$  exists, we define  $\Delta_n(\omega) = \infty$ .) Then*

$$\lim_{n \rightarrow \infty} \frac{\Delta_n}{n} = 0, \quad \mathbb{P}\text{-almost surely.} \quad (5.4)$$



**Fig. 3.** An illustration of the main idea of the proof of Theorem 5.1. Here a square of side  $n$  is intersected by a grid  $\mathbb{G}$  of good lines “emanating” from the good points on the  $x$  and  $y$  axes. The crosses represent the points on these lines which are in  $\mathcal{C}_\infty$ . Along the good lines the corrector grows slower than linear and so anywhere on  $\mathbb{G}$  sublinearity holds. For the part of  $\mathcal{C}_\infty$  that is not on  $\mathbb{G}$ , the maximum principle for  $x \mapsto x + \chi(x, \omega)$  lets us bound the corrector by the values on the parts of the grid that surround it, modulo factors of order  $o(n)$ .

*Proof.* Since  $\mathbb{P}$  is  $\tau_e$  invariant and  $\tau_e$  is ergodic, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n \mathbf{1}_{\{0 \in \mathcal{G}_{K,\epsilon}\}} \circ \tau_e^k = \mathbb{P}(0 \in \mathcal{G}_{K,\epsilon}) \quad (5.5)$$

$\mathbb{P}$ -almost surely. A similar statement applies to the limit  $n \rightarrow -\infty$ . But if  $\Delta_n/n$  does not tend to zero, at least one of these limits would not exist.  $\square$

*Proof of Theorem 5.1.* Fix  $\epsilon \in (0, 1/2)$  and let  $K_0$  be such that  $\mathbb{P}(0 \in \mathcal{G}_{K,\epsilon}) > 0$  for all  $K \geq K_0$  (we are using that  $\mathcal{G}_{K,\epsilon}$  increases with  $K$ ). Let  $\Omega_0^* \subset \Omega_0$  be the set of configurations such that the conclusion of Lemma 5.3 applies for both  $x$  and  $y$ -axes, and that shift-invariance (2.12) holds for all  $x, y$  in the infinite cluster. We will show that for every  $\omega \in \Omega_0^*$  the limsup in (5.1) is less than  $6\epsilon$  almost surely.

Let  $e_1$  and  $e_2$  denote the coordinate vectors in  $\mathbb{Z}^2$ . Fix  $\omega \in \Omega_0^*$  and adjust  $K \geq K_0$  so that  $0 \in \mathcal{G}_{K,\epsilon}$ . (This is possible by the definition of  $\Omega_0^*$ .) Then we define  $(x_k)_{k \in \mathbb{Z}}$  to be the increasing two-sided sequence of all integers such that  $x_k e_1$  exhausts all  $K, \epsilon$ -good points on the  $e_1$ -axis, i.e.,

$$x_k e_1 \in \mathcal{G}_{K,\epsilon}(\omega), \quad k \in \mathbb{Z}. \quad (5.6)$$

If  $\Delta_n$  be the maximal gap between consecutive  $x_j$ 's that lie in  $[-n, n]$ , cf (5.3), we define  $n_1(\omega)$  be the least integer such that  $\Delta_n/n < \epsilon$  for all  $n \geq n_1(\omega)$ . Similarly we identify a two-sided increasing sequence  $(y_n)_{n \in \mathbb{Z}}$  of integers exhausting the sites such that

$$y_k e_2 \in \mathcal{G}_{K,\epsilon}(\omega), \quad k \in \mathbb{Z}, \quad (5.7)$$

and let  $n_2(\omega)$  be the quantity corresponding to  $n_1(\omega)$  in this case.

Let  $n_0 = \max\{n_1, n_2\}$ . We claim that for all  $n \geq n_0(\omega)$ ,

$$\max_{\substack{x \in \mathcal{C}_\infty(\omega) \\ |x|_\infty \leq n}} |\chi(x, \omega)| \leq 2K + 6\epsilon n. \quad (5.8)$$

To prove this, let us consider the grid  $\mathbb{G} = \mathbb{G}(\omega)$  of *good lines*

$$\{x_k e_1 + n e_2 : n \in \mathbb{Z}\}, \quad k \in \mathbb{Z}, \quad (5.9)$$

and

$$\{n e_1 + y_k e_2 : n \in \mathbb{Z}\}, \quad k \in \mathbb{Z}, \quad (5.10)$$

see Fig. 3. As a first step we will use the harmonicity of  $x \mapsto x + \chi(x, \omega)$  to deal with  $x \in \mathcal{C}_\infty \setminus \mathbb{G}$ . Indeed, any such  $x$  is enclosed between two horizontal and two vertical grid lines and every path on  $\mathcal{C}_\infty$  connecting  $x$  to “infinity” necessarily intersects one of these lines at a point which is also in  $\mathcal{C}_\infty$ . Applying the maximum (and minimum) principle for harmonic functions we get

$$\max_{\substack{x \in \mathcal{C}_\infty \setminus \mathbb{G} \\ |x|_\infty \leq n}} |\chi(x, \omega)| \leq 2\epsilon n + \max_{\substack{x \in \mathcal{C}_\infty \cap \mathbb{G} \\ |x|_\infty \leq 2n}} |\chi(x, \omega)|. \quad (5.11)$$

Here we used that the enclosing lines are not more than  $\frac{\epsilon}{1-\epsilon}n \leq 2\epsilon n \leq n$  apart and, in particular, they all intersect the block  $[-2n, 2n] \times [-2n, 2n]$ .

To estimate the maximum on the grid, we pick, say, a horizontal grid line with  $y$ -coordinate  $y_k$  and note that, by (2.12), for every  $x \in \mathcal{C}_\infty$  on this line,

$$\chi(x, \omega) - \chi(y_k e_2, \omega) = \chi(x - y_k e_2, \tau_{y_k e_2} \omega). \quad (5.12)$$

By (5.7) and the fact that  $x - y_k e_2 \in \mathcal{C}_\infty(\tau_{y_k e_2} \omega)$  we have

$$|\chi(x, \omega) - \chi(y_k e_2, \omega)| \leq K + 2\epsilon n \quad (5.13)$$

whenever  $x$  is such that  $|x|_\infty \leq 2n$ . Applying the same argument to the vertical line through the origin, and  $x$  replaced by  $y_k e_2$ , we get

$$|\chi(x, \omega)| \leq 2K + 4\epsilon n \quad (5.14)$$

for every  $x \in \mathcal{C}_\infty \cap \mathbb{G}$  with  $|x|_\infty \leq 2n$ . Combining this with (5.11), the estimate (5.8) and the whole claim are finally proved.  $\square$

Interestingly, a variant of the above strategy for controlling the corrector in  $d = 2$  has independently been developed by Chris Hoffman [25] to control the geodesics in the first-passage percolation on  $\mathbb{Z}^2$ .

### 5.2. Three and higher dimensions

In  $d \geq 3$  we have the following weaker version of Theorem 5.1:

**Theorem 5.4.** *Let  $d \geq 3$ . Then for all  $\epsilon > 0$  and  $\mathbb{P}_0$ -almost all  $\omega$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{(2n+1)^d} \sum_{\substack{x \in \mathcal{C}_\infty(\omega) \\ |x| \leq n}} \mathbf{1}_{\{|\chi(x, \omega)| \geq \epsilon n\}} = 0. \quad (5.15)$$

Here we fix the dimension  $d$  and run an induction over  $\nu$ -dimensional sections of the  $d$ -dimensional box  $\{x \in \mathbb{Z}^d : |x| \leq n\}$ . Specifically, for each  $\nu = 1, \dots, d$ , let  $\Lambda_n^\nu$  be the  $\nu$ -dimensional box

$$\Lambda_n^\nu = \{k_1 e_1 + \dots + k_\nu e_\nu : k_i \in \mathbb{Z}, |k_i| \leq n \ \forall i = 1, \dots, \nu\}. \quad (5.16)$$

The induction eventually gives (5.15) for  $\nu = d$  thus proving the theorem.

Since it is not advantageous to assume that  $0 \in \mathcal{C}_\infty$ , we will carry out the proof for *differences* of the form  $\chi(x, \omega) - \chi(y, \omega)$  with  $x, y \in \mathcal{C}_\infty$ . For each  $\omega \in \Omega$ , we thus consider the (upper) density

$$\varrho_\nu(\omega) = \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \inf_{y \in \mathcal{C}_\infty(\omega) \cap \Lambda_n^1} \frac{1}{|\Lambda_n^\nu|} \sum_{x \in \mathcal{C}_\infty(\omega) \cap \Lambda_n^\nu} \mathbf{1}_{\{|\chi(x, \omega) - \chi(y, \omega)| \geq \epsilon n\}}. \quad (5.17)$$

Note that the infimum is taken only over sites in one-dimensional box  $\Lambda_n^1$ . Our goal is to show by induction that  $\varrho_\nu = 0$  almost surely for all  $\nu = 1, \dots, d$ . The induction step is encapsulated into the following lemma:

**Lemma 5.5.** *Let  $1 \leq \nu < d$ . If  $\varrho_\nu = 0$ ,  $\mathbb{P}$ -almost surely, then also  $\varrho_{\nu+1} = 0$ ,  $\mathbb{P}$ -almost surely.*

Before we start the formal proof, let us discuss its main idea: Suppose that  $\varrho_\nu = 0$  for some  $\nu < d$ ,  $\mathbb{P}$ -almost surely. Pick  $\epsilon > 0$ . Then for  $\mathbb{P}$ -almost every  $\omega$  and all sufficiently large  $n$ , there exists a set of sites  $\Delta \subset \Lambda_n^\nu \cap \mathcal{C}_\infty$  such that

$$|(\Lambda_n^\nu \cap \mathcal{C}_\infty) \setminus \Delta| \leq \epsilon |\Lambda_n^\nu| \quad (5.18)$$

and

$$|\chi(x, \omega) - \chi(y, \omega)| \leq \epsilon n, \quad x, y \in \Delta. \quad (5.19)$$

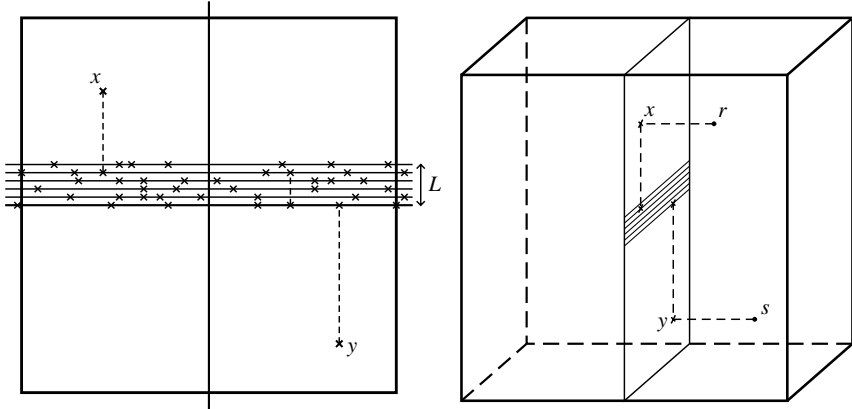
Moreover,  $n$  sufficiently large,  $\Delta$  could be picked so that  $\Delta \cap \Lambda_n^1 \neq \emptyset$  and, assuming  $K \gg 1$ , the non- $K$ ,  $\epsilon$ -good sites could be pitched out with little loss of density to achieve even

$$\Delta \subset \mathcal{G}_{K, \epsilon}. \quad (5.20)$$

(All these claims are direct consequences of the Pointwise Ergodic Theorem and the fact that  $\mathbb{P}(0 \in \mathcal{G}_{K, \epsilon})$  converges to the density of  $\mathcal{C}_\infty$  as  $K \rightarrow \infty$ .)

As a result of this construction we have

$$|\chi(z, \omega) - \chi(x, \omega)| \leq K + \epsilon n \quad (5.21)$$



**Fig. 4.** The main idea underlying the proof of Theorem 5.4. The figure on the left represents an  $n \times n$  square in a two-dimensional plane in  $\mathbb{Z}^3$ ; the crosses now stand for *good sites*; cf Definition 5.2. Here  $L$  is chosen so that  $(1 - \delta)$ -fraction of all vertical lines find a good point on the intersection with one of the  $L$  horizontal lines;  $n$  is assumed so large that every pair of these lines has two good points “above” each other. Any two good points  $x$  and  $y$  in the square are connected by broken-line path that uses at most 4 good points in between. The dashed lines indicate the vertical pieces of one such path. The figure on the right indicates how this is used to control the difference of the corrector at two general points  $r, s \in \mathcal{C}_\infty$  in an  $n \times n \times n$  cube in  $\mathbb{Z}^3$ —with obvious extensions to all  $d \geq 3$ .

for any  $x \in \Delta$  and any  $z \in \Lambda_n^{\nu+1} \cap \mathcal{C}_\infty$  of the form  $x + je_{\nu+1}$ . Thus, if  $r, s \in \mathcal{C}_\infty \cap \Lambda_n^{\nu+1}$  are of the latter form,  $r = x + je_{\nu+1}$  and  $s = y + ke_{\nu+1}$ —see Fig. 4 for an illustration—then (5.21) implies

$$|\chi(r, \omega) - \chi(s, \omega)| \leq |\chi(x, \omega) - \chi(y, \omega)| + 2K + 2\epsilon n. \quad (5.22)$$

Invoking the “induction hypothesis” (5.19), the right-hand side is less than  $2K + 3\epsilon n$ , implying a bound of the type (5.19) but one-dimension higher.

Unfortunately, the above is not sufficient to prove (5.19) for all but a vanishing fraction of all sites in  $\Lambda_n^{\nu+1}$ . The reason is that the  $r$ ’s and  $s$ ’s for which (5.22) holds need to be of the form  $x + je_{\nu+1}$  for some  $x \in \Delta \cap \mathcal{C}_\infty$ . But  $\mathcal{C}_\infty$  will occupy only about  $P_\infty = \mathbb{P}(0 \in \mathcal{C}_\infty)$  fraction of all sites in  $\Lambda_n^\nu$ , and so this argument does not permit us to control more than fraction about  $P_\infty$  of  $\Lambda_n^{\nu+1} \cap \mathcal{C}_\infty$ .

To fix this problem, we will have to work with a “stack” of translates of  $\Lambda_n^\nu$  at the same time. (These correspond to the stack of horizontal lines on the left of Fig. 4.) Explicitly, consider the collection of  $\nu$ -boxes

$$\Lambda_{n,j}^\nu = \tau_{je_{\nu+1}}(\Lambda_n^\nu), \quad j = 1, \dots, L. \quad (5.23)$$

Here  $L$  is a deterministic number chosen so that, for a given  $\delta > 0$ , the set

$$\Delta_0 = \{x \in \Lambda_n^\nu : \exists j \in \{0, \dots, L-1\}, x + je_{\nu+1} \in \Lambda_{n,j}^\nu \cap \mathcal{C}_\infty\} \quad (5.24)$$

is so large that

$$|\Delta_0| \geq (1 - \delta)|\Lambda_n^\nu| \quad (5.25)$$

once  $n$  is sufficiently large. These choices ensure that  $(1 - \delta)$ -fraction of  $\Lambda_n^v$  is now “covered” which by repeating the above argument gives us control over  $\chi(r, \omega)$  for nearly the same fraction of all sites  $r \in \Lambda_n^{v+1} \cap \mathcal{C}_\infty$ .

*Proof of Lemma 5.5.* Let  $v < d$  and suppose that  $\varrho_v = 0$ ,  $\mathbb{P}$ -almost surely. Fix  $\delta$  with  $0 < \delta < \frac{1}{2}P_\infty^2$  and let  $L$  be as defined above. Choose  $\epsilon > 0$  so that

$$L\epsilon + \delta < \frac{1}{2}P_\infty^2. \quad (5.26)$$

For a fixed but large  $K$ , and  $\mathbb{P}$ -almost every  $\omega$  and  $n$  exceeding an  $\omega$ -dependent quantity, for each  $j = 1, \dots, L$ , we can find  $\Delta_j \subset \Lambda_{n,j}^v \cap \mathcal{C}_\infty$  satisfying the properties (5.18–5.20)—with  $\Lambda_n^v$  replaced by  $\Lambda_{n,j}^v$ . Given  $\Delta_1, \dots, \Delta_L$ , let  $\Lambda$  be the set of sites in  $\Lambda_n^{v+1} \cap \mathcal{C}_\infty$  whose projection onto the linear subspace  $\mathbb{H} = \{k_1 e_1 + \dots + k_v e_v : k_i \in \mathbb{Z}\}$  belongs to the corresponding projection of  $\Delta_1 \cup \dots \cup \Delta_L$ . Note that the  $\Delta_j$  could be chosen so that  $\Lambda \cap \Lambda_n^1 \neq \emptyset$ .

By their construction, the projections of the  $\Delta_j$ 's,  $j = 1, \dots, L$ , onto  $\mathbb{H}$  “fail to cover” at most  $L\epsilon|\Lambda_n^v|$  sites in  $\Delta_0$ , and so at most  $(\delta + L\epsilon)|\Lambda_n^v|$  sites in  $\Lambda_n^v$  are not of the form  $x + i e_{v+1}$  for some  $x \in \bigcup_j \Delta_j$ . It follows that

$$|(\Lambda_n^{v+1} \cap \mathcal{C}_\infty) \setminus \Lambda| \leq (\delta + L\epsilon)|\Lambda_n^{v+1}|, \quad (5.27)$$

i.e.,  $\Lambda$  contains all except at most  $(L\epsilon + \delta)$ -fraction of all sites in  $\Lambda_n^{v+1}$  that we care about. Next we note that if  $K$  is sufficiently large, then for every  $1 \leq i < j \leq L$ , the set  $\mathbb{H}$  contains at least  $\frac{1}{2}P_\infty^2$ -fraction of sites  $x$  such that

$$z_i \stackrel{\text{def}}{=} x + i e_v \in \mathcal{G}_{K,\epsilon} \quad \text{and} \quad z_j \stackrel{\text{def}}{=} x + j e_v \in \mathcal{G}_{K,\epsilon}. \quad (5.28)$$

Since we assumed (5.26), once  $n \gg 1$ , for each pair  $(i, j)$  with  $1 \leq i < j \leq L$  such  $z_i$  and  $z_j$  can be found so that  $z_i \in \Delta_i$  and  $z_j \in \Delta_j$ . But the  $\Delta_j$ 's were picked to make (5.19) true and so via these pairs of sites we now show that

$$|\chi(y, \omega) - \chi(x, \omega)| \leq K + \epsilon L + 2\epsilon n \quad (5.29)$$

for every  $x, y \in \Delta_1 \cup \dots \cup \Delta_L$ ; see again (the left part of) Fig. 4.

From (5.19) and (5.29) we now conclude that for all  $r, s \in \Lambda$ ,

$$|\chi(r, \omega) - \chi(s, \omega)| \leq 3K + \epsilon L + 4\epsilon n < 5\epsilon n, \quad (5.30)$$

provided that  $\epsilon n > 3K + \epsilon L$ . If  $\varrho_{v,\epsilon}$  denotes the right-hand side of (5.17) before taking  $\epsilon \downarrow 0$ , the bounds (5.27) and (5.30) and  $\Lambda \cap \Lambda_n^1 \neq \emptyset$  yield

$$\varrho_{v+1,5\epsilon}(\omega) \leq \delta + L\epsilon, \quad (5.31)$$

for  $\mathbb{P}$ -almost every  $\omega$ . But the left-hand side of this inequality increases as  $\epsilon \downarrow 0$  while the right-hand side decreases. Thus, taking  $\epsilon \downarrow 0$  and  $\delta \downarrow 0$  proves that  $\rho_{v+1} = 0$  holds  $\mathbb{P}$ -almost surely.  $\square$

*Proof of Theorem 5.4.* The proof is an easy consequence of Lemma 5.5. First, by Theorem 4.1 we know that  $\varrho_1(\omega) = 0$  for  $\mathbb{P}_0$ -almost every  $\omega$ . Invoking appropriate



shifts, the same conclusion applies  $\mathbb{P}$ -almost surely. Using induction on dimension, Lemma 5.5 then tells us that  $\varrho_d(\omega) = 0$  for  $\mathbb{P}_0$ -almost every  $\omega$ . Let  $\omega \in \Omega_0$ . By Theorem 4.1, for each  $\epsilon > 0$  there is  $n_0 = n_0(\omega)$  with  $\mathbb{P}_0(n_0 < \infty) = 1$  such that for all  $n \geq n_0(\omega)$ , we have  $|\chi(x, \omega)| \leq \epsilon n$  for all  $x \in \Lambda_n^1 \cap \mathcal{C}_\infty(\omega)$ . Using this to estimate away the infimum in (5.17), the fact that  $\varrho_d = 0$  now immediately implies (5.15) for all  $\epsilon > 0$ .  $\square$

## 6. Proof of main results

Here we will finally prove our main theorems. First, in Sect. 6.1, we will show the convergence of the “lazy” walk on the deformed graph to Brownian motion and then, in Sect. 6.2, we use our previous results on corrector growth to extend this to the walk on the original graph. This separation will allow us to treat the parts of the proof common for  $d = 2$  and  $d \geq 3$  in a unified way. Theorem 1.2, which concerns the “agile” walk, is proved in Sect. 6.3.

### 6.1. Convergence on deformed graph

We begin with a simple observation that will drive all underlying derivations:

**Lemma 6.1.** *Fix  $\omega \in \Omega_0$  and let  $x \mapsto \chi(x, \omega)$  be the corrector. Given a path of random walk  $(X_n)_{n \geq 0}$  with law  $P_{0, \omega}$ , let*

$$M_n^{(\omega)} = X_n + \chi(X_n, \omega), \quad n \geq 0. \quad (6.1)$$

*Then  $(M_n^{(\omega)})_{n \geq 0}$  is an  $L^2$ -martingale for the filtration  $(\sigma(X_0, \dots, X_n))_{n \geq 0}$ . Moreover, conditional on  $X_{k_0} = x$ , the increments  $(M_{k+k_0}^{(\omega)} - M_{k_0}^{(\omega)})_{k \geq 0}$  have the same law as  $(M_k^{(\tau_x, \omega)})_{k \geq 0}$ .*

*Proof.* Since  $X_n$  is bounded,  $\chi(X_n, \omega)$  is bounded and so  $M_n^{(\omega)}$  is square integrable with respect to  $P_{0, \omega}$ . Since  $x \mapsto x + \chi(x, \omega)$  is harmonic with respect to the transition probabilities of the random walk  $(X_n)$  with law  $P_{0, \omega}$ , we have

$$E_{0, \omega}(M_{n+1}^{(\omega)} | \sigma(X_n)) = M_n^{(\omega)}, \quad n \geq 0, \quad (6.2)$$

$P_{0, \omega}$ -almost surely. Since  $M_n^{(\omega)}$  is  $\sigma(X_n)$ -measurable,  $(M_n^{(\omega)})$  is a martingale. The stated relation between the laws of  $(M_{k+k_0}^{(\omega)} - M_{k_0}^{(\omega)})_{k \geq 0}$  and  $(M_k^{(\tau_x, \omega)})_{k \geq 0}$  is implied by the shift-invariance (2.12) and the fact that  $(M_n^{(\omega)})$  is a simple random walk on the deformed infinite component.  $\square$

Next we will establish the convergence of the above martingale to Brownian motion. The precise statement is as follows:

**Theorem 6.2.** *Let  $d \geq 2$ ,  $p > p_c$  and  $\omega \in \Omega_0$ . Let  $(X_n)_{n \geq 0}$  be the random walk with law  $P_{0, \omega}$  and let  $(M_n^{(\omega)})_{n \geq 0}$  be as defined in (6.1). Let  $(\widehat{B}_n^{(\omega)}(t) : t \geq 0)$  be defined by*

$$\widehat{B}_n^{(\omega)}(t) = \frac{1}{\sqrt{n}}(M_{\lfloor tn \rfloor}^{(\omega)} + (tn - \lfloor tn \rfloor)(M_{\lfloor tn \rfloor + 1}^{(\omega)} - M_{\lfloor tn \rfloor}^{(\omega)})), \quad t \geq 0. \quad (6.3)$$

Then for all  $T > 0$  and  $\mathbb{P}_0$ -almost every  $\omega$ , the law of  $(\widehat{B}_n(t): 0 \leq t \leq T)$  on  $(C[0, T], \mathscr{W}_T)$  converges weakly to the law of an isotropic Brownian motion  $(B_t: 0 \leq t \leq T)$  with diffusion constant  $D$ , i.e.,  $E(B_t^2) = Dt$ , where

$$D = \mathbb{E}_0 \left( E_{0,\omega} |X_1 + \chi(X_1, \omega)|^2 \right) \in (0, \infty). \quad (6.4)$$

*Proof.* Without much loss of generality, we may confine ourselves to the case when  $T = 1$ . Let  $\mathscr{F}_k = \sigma(X_0, X_1, \dots, X_k)$  and fix a vector  $a \in \mathbb{R}^d$ . We will show that (the piece-wise linearization) of  $t \mapsto a \cdot M_{\lfloor tn \rfloor}^{(\omega)}$  scales to one-dimensional Brownian motion. For  $m \leq n$ , consider the random variable

$$V_{n,m}^{(\omega)}(\epsilon) = \frac{1}{n} \sum_{k=0}^m E_{0,\omega} \left( [a \cdot (M_{k+1}^{(\omega)} - M_k^{(\omega)})]^2 \mathbf{1}_{\{|a \cdot (M_{k+1}^{(\omega)} - M_k^{(\omega)})| \geq \epsilon \sqrt{n}\}} \mid \mathscr{F}_k \right). \quad (6.5)$$

In order to apply the Lindeberg-Feller Functional CLT for martingales (Theorem 7.7.3 of Durrett [15]), we need to verify that for  $\mathbb{P}_0$ -almost every  $\omega$ ,

- (1)  $V_{n, \lfloor tn \rfloor}^{(\omega)}(0) \rightarrow Ct$  in  $P_{0,\omega}$ -probability for all  $t \in [0, 1]$  and some  $C \in (0, \infty)$ .
- (2)  $V_{n,n}^{(\omega)}(\epsilon) \rightarrow 0$  in  $P_{0,\omega}$ -probability for all  $\epsilon > 0$ .

Both of these conditions will be implied by Theorem 3.1. Indeed, by the last conclusion of Lemma 6.1 we may write

$$V_{n,m}^{(\omega)}(\epsilon) = \frac{1}{n} \sum_{k=0}^m f_{\epsilon \sqrt{n}} \circ \tau_{X_k}(\omega), \quad (6.6)$$

where

$$f_K(\omega) = E_{0,\omega} \left( [a \cdot M_1^{(\omega)}]^2 \mathbf{1}_{\{|a \cdot M_1^{(\omega)}| \geq K\}} \right). \quad (6.7)$$

Now if  $\epsilon = 0$ , Theorem 3.1 tells us that, for  $\mathbb{P}_0$ -almost every  $\omega$ ,

$$\lim_{n \rightarrow \infty} V_{n,n}^{(\omega)}(0) = \mathbb{E}_0 \left( E_{0,\omega} ([a \cdot M_1^{(\omega)}]^2) \right) = \frac{1}{d} D |a|^2, \quad (6.8)$$

where we used the symmetry of the joint expectations under rotations by  $90^\circ$ . From here condition (1) follows by scaling out the  $t$ -dependence first and working with  $tn$  instead of  $n$ .

On the other hand, when  $\epsilon > 0$ , we have  $f_{\epsilon \sqrt{n}} \leq f_K$  once  $n$  is sufficiently large and so,  $\mathbb{P}_0$ -almost surely,

$$\limsup_{n \rightarrow \infty} V_{n,n}^{(\omega)}(\epsilon) \leq \mathbb{E}_0 \left( E_{0,\omega} ([a \cdot M_1^{(\omega)}]^2 \mathbf{1}_{\{|a \cdot M_1^{(\omega)}| \geq K\}}) \right) \xrightarrow{K \rightarrow \infty} 0, \quad (6.9)$$

where to apply Dominated Convergence we used that  $a \cdot M_1^{(\omega)} \in L^2$ . Hence, the above conditions (1) and (2) hold—in fact, even with limits taken  $P_{0,\omega}$ -almost surely. Applying the Martingale functional CLT and the Cramér-Wold device (Theorem 2.9.2 of [15]), we conclude that, for  $\mathbb{P}_0$ -almost every  $\omega$ , the linear interpolation of the sequence  $(M_k^{(\omega)} / \sqrt{n})_{k=1, \dots, n}$  converges to isotropic Brownian motion with covariance matrix  $\frac{1}{d} D \mathbf{1}$ .

To make the proof complete, we need to show that  $D \in (0, \infty)$ . Here the finiteness is immediate by the square-integrability of  $\chi$ . The positivity can be shown in many ways: either by a direct computation from (6.4) using that  $\mathbb{E}_0(E_{0,\omega}(X_1 \cdot \chi(X_1, \omega))) = 0$  [which in turn is implied by  $\mathbb{E}_0(\chi(e, \omega) 1_{\{\omega_e=1\}}) = 0$  for every coordinate vector  $e$ ] or by invoking the sublinearity of the corrector proved in Theorems 5.1–5.4, or by an appeal to the lower (or, alternatively, upper) bound in [3, Theorem 1].  $\square$

## 6.2. Correction on the corrector

It remains to estimate the influence of the harmonic deformation on the path of the walk. As already mentioned, while our proof in  $d = 2$  is completely self-contained, for  $d \geq 3$  we rely heavily on (a discrete version of) the sophisticated Theorem 1 of Barlow [3].

Let us first dismiss the two-dimensional case of Theorem 1.1:

*Proof of Theorem 1.1 ( $d = 2$ ).* We need to extend the conclusion of Theorem 6.2 to the linear interpolation of  $(X_n)$ . Since the corrector is an additive perturbation of  $M_n^{(\omega)}$ , it clearly suffices to show that, for  $\mathbb{P}_0$ -almost every  $\omega$ ,

$$\max_{1 \leq k \leq n} \frac{|\chi(X_k, \omega)|}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0, \quad \text{in } P_{0,\omega}\text{-probability.} \quad (6.10)$$

By Theorem 5.1 we know that for every  $\epsilon > 0$  there exists a  $K = K(\omega) < \infty$  such that

$$|\chi(x, \omega)| \leq K + \epsilon|x|_\infty, \quad x \in \mathcal{C}_\infty(\omega). \quad (6.11)$$

If  $\epsilon < 1/2$ , then this implies

$$|\chi(X_k, \omega)| \leq 2K + 2\epsilon|M_k^{(\omega)}|. \quad (6.12)$$

But the above CLT for  $(M_k)$  tells us that  $\max_{k \leq n} |M_k^{(\omega)}|/\sqrt{n}$  converges in law to the maximum of a Brownian motion  $B(t)$  over  $t \in [0, 1]$ . Hence, if  $P$  denotes the probability law of the Brownian motion, the Portmanteau Theorem (Theorem 2.1 of [7]) allows us to conclude

$$\limsup_{n \rightarrow \infty} P_{0,\omega}(\max_{k \leq n} |\chi(X_k, \omega)| \geq \delta\sqrt{n}) \leq P\left(\max_{0 \leq t \leq 1} |B(t)| \geq \frac{\delta}{2\epsilon}\right). \quad (6.13)$$

The right-hand side tends to zero as  $\epsilon \downarrow 0$  for all  $\delta > 0$ .  $\square$

In order to prove the same result in  $d \geq 3$ , we will need the following upper bounds on the transition probability of our random walk:

**Theorem 6.3.** (1) *There is a random variable  $C = C(\omega)$  with  $\mathbb{P}_0(C < \infty) = 1$  such that for all  $\omega \in \Omega_0$  and all  $x \in \mathcal{C}_\infty(\omega)$ ,*

$$P_{0,\omega}(X_n = x) \leq \frac{C(\omega)}{n^{d/2}}, \quad n \geq 1. \quad (6.14)$$

(2) There are constants  $c_1, c_2 \in (0, \infty)$  and random variables  $N_x = N_x(\omega)$  such that for all  $\omega \in \Omega_0$ , all  $x \in \mathcal{C}_\infty(\omega)$ , all  $R \geq 1$ , and all  $n \geq N_x(\omega)$ ,

$$P_{x,\omega}(|X_n - x| > R) \leq c_1 \exp\{-c_2 R^2/n\}. \quad (6.15)$$

Moreover, the random variables  $(N_x)$  have stretched-exponential tails, i.e., there exist constants  $c_3 > 0$  and  $\theta > 0$  such that for all  $x \in \mathbb{Z}^d$ ,

$$\mathbb{P}_0(N_x > R) \leq e^{-c_3 R^\theta}, \quad R \geq 1. \quad (6.16)$$

For a continuous-time version of our walk, these bounds are the content of Theorem 1 of Barlow [3]. (In fact, the continuous-time version of the bound (6.14) was obtained already by Mathieu and Remy [31].) Unfortunately, to derive Theorem 6.3 from Barlow's Theorem 1, one needs to invoke various non-trivial facts about percolation and/or mixing of Markov chains. In Appendix A we list these facts and show how to assemble all ingredients together to establish the above upper bounds.

*Proof of Theorem 1.1* ( $d \geq 3$ ). We will adapt (the easier part of) the proof of Theorem 1.1 in Sidoravicius and Sznitman [38]. First we show that the laws of  $(\tilde{B}_n(t) : t \leq T)$  on  $(C[0, T], \mathscr{W}_T)$  are tight. To that end it suffices to show (e.g., by Theorem 8.6 of Ethier-Kurtz [16]) that if  $\mathcal{S}_n$  is the class of all stopping times of the filtration  $(\sigma(\{\tilde{B}_n(s) : s \leq t\}))_{0 \leq t \leq T}$ , then

$$\limsup_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{S}_n} E_{0,\omega}(|\tilde{B}_n(\tau + \epsilon) - \tilde{B}_n(\tau)|^2) = 0. \quad (6.17)$$

As in [38], we replace  $\tau$  by its integer-valued approximation. Explicitly, let  $\hat{\tau} = \lfloor n\tau \rfloor + 1$  and let  $\delta$  be a number such that  $n\delta = \lfloor n\epsilon \rfloor + 1$ . Since  $\hat{\tau}$  differs from  $n\tau$  by a constant of order unity, and similarly for  $\hat{\tau} + n\delta$  and  $n(\tau + \epsilon)$ , we have

$$|\tilde{B}_n(\tau + \epsilon) - \tilde{B}_n(\tau)| \leq \frac{1}{\sqrt{n}} |X_{\hat{\tau}+n\delta} - X_{\hat{\tau}}| + \frac{c_4}{\sqrt{n}} \quad (6.18)$$

for some constant  $c_4 < \infty$ . This allows us to estimate (6.17) by means of the second moment of  $|X_{\hat{\tau}+n\delta} - X_{\hat{\tau}}|$ .

Recalling that  $\tau \leq T$ , we may assume that  $\hat{\tau} \leq 2Tn$ . By (6.16) we know that there exists an almost-surely finite random variable  $C' = C'(\omega)$  such that  $\max_{|x| \leq R} N_x \leq C'(\omega)(\log R)^\zeta$  once  $R \geq 2$ , where  $\zeta = 2/\theta$ . Since  $|X_{\hat{\tau}}| \leq 2Tn$ , this implies that  $N_{X_{\hat{\tau}}} \leq C'(\omega)[\log(2Tn)]^\zeta$ . Theorem 6.3(2) and the strong Markov property— $\hat{\tau}$  is a stopping time of the random walk—tell us that, for some constant  $c_5 < \infty$  (depending only on  $c_1, c_2$  and the dimension),

$$E_{0,\omega}(|X_{\hat{\tau}+n\delta} - X_{\hat{\tau}}|^2) \leq c_5 \epsilon n, \quad n \geq n_0(\omega). \quad (6.19)$$

Here we used  $\epsilon - \delta = O(1/n)$  and let  $n_0(\omega)$  be such that  $\delta n \geq C'(\omega)[\log(2Tn)]^\zeta$  for all  $n \geq n_0$ . The bound (6.17) is now proved by combining (6.18–6.19) and taking the required limits.

Once we know that the laws of  $(\tilde{B}_n(t) : t \leq T)$  are tight, it suffices to show the convergence of finite-dimensional distributions. In light of Theorem 6.2 (and the

Markov property of the walk), for that it is enough to prove that for all  $t > 0$  and  $\mathbb{P}_0$ -almost every  $\omega$ ,

$$\frac{\chi(X_{\lfloor tn \rfloor}, \omega)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{in } P_{0,\omega}\text{-probability.} \quad (6.20)$$

Without loss of generality, we need to do this only for  $t = 1$ . By Theorem 6.3, the random variable  $X_n$  lies with probability  $1 - \epsilon$  in the block  $[-M\sqrt{n}, M\sqrt{n}]^d \cap \mathbb{Z}^d$ , provided  $M$  sufficiently large (with “large” depending possibly on  $\omega$ ). Using Theorem 6.3(1) to estimate  $P_{0,\omega}(X_n = x)$  for  $x$  inside this block, we have

$$P_{0,\omega}(|\chi(X_n, \omega)| > \delta\sqrt{n}) \leq \epsilon + C(\omega) \frac{1}{n^{d/2}} \sum_{\substack{x \in \mathcal{C}_\infty(\omega) \\ |x| \leq M\sqrt{n}}} \mathbf{1}_{\{|\chi(x, \omega)| > \delta\sqrt{n}\}}. \quad (6.21)$$

But Theorem 5.4 tells us that, for all  $\delta, M > 0$  and  $\mathbb{P}_0$ -almost every  $\omega$ , the second term tends to zero as  $n \rightarrow \infty$ . This proves (6.20) and the whole claim.  $\square$

### 6.3. Extension to “agile” walk

It remains to prove Theorem 1.2 for the “agile” version of simple random walk on  $\mathcal{C}_\infty$ . Since the proof is based entirely on the statement of Theorem 1.1, we will resume a unified treatment of all  $d \geq 2$ . First we will make the observation that the times of the two walks run proportionally to each other:

**Lemma 6.4.** *Let  $(T_k)_{k \geq 0}$  be the stopping times defined in (1.7). Then for all  $t \geq 0$  and  $\mathbb{P}_0$ -almost every  $\omega$ ,*

$$\frac{T_{\lfloor tn \rfloor}}{n} \xrightarrow[n \rightarrow \infty]{} \Theta t, \quad P_{0,\omega}\text{-almost surely,} \quad (6.22)$$

where

$$\frac{1}{\Theta} = \mathbb{E}_0 \left( \frac{1}{2d} \sum_{e: |e|=1} \mathbf{1}_{\{\omega_e=1\}} \right). \quad (6.23)$$

*Proof.* This is an easy consequence of the second part of Theorem 3.1 and the fact that for  $\mathbb{P}_0$ -almost every  $\omega$  we have  $\tau_x \omega \neq \omega$  once  $x \neq 0$ . Indeed, let  $f(\omega, \omega') = \mathbf{1}_{\{\omega \neq \omega'\}}$ . For  $t = 0$  the statement holds trivially so let us assume that  $t > 0$ . If  $n$  is so large that  $T_{\lfloor nt \rfloor} > 0$ , we have

$$\frac{n}{T_{\lfloor tn \rfloor}} = \frac{1}{T_{\lfloor tn \rfloor}} \sum_{k=1}^{T_{\lfloor tn \rfloor}} f(\tau_{X_{k-1}} \omega, \tau_{X_k} \omega). \quad (6.24)$$

Since  $T_{\lfloor tn \rfloor} \rightarrow \infty$  as  $n \rightarrow \infty$ , by Theorem 3.1 the right hand side converges to the expectation of  $f(\omega, \tau_{X_1} \omega)$  in the annealed measure  $\mathbb{E}_0(P_{0,\omega}(\cdot))$ . A direct calculation shows that this expectation equals  $\Theta$ .  $\square$

*Proof of Theorem 1.2.* The proof is based on a standard approximation argument for stochastic processes. Let  $\tilde{B}_n(t)$  be as in Theorem 1.1 and recall that  $\tilde{B}'_n(t)$  is a

linear interpolation of the values  $\tilde{B}_n(T_k/n)$  for  $k = 0, \dots, n$ . The path-continuity of the processes  $\tilde{B}_n(t)$  as well as the limiting Brownian motion implies that for every  $\epsilon > 0$  there is a  $\delta > 0$  such that

$$P_{0,\omega} \left( \sup_{\substack{t,t' \leq T \\ |t-t'| < \delta}} |\tilde{B}_n(t) - \tilde{B}_n(t')| < \epsilon \right) > 1 - \epsilon \quad (6.25)$$

once  $n$  is sufficiently large. Similarly, Lemma 6.4, the continuity of  $t \mapsto \Theta t$  and the monotonicity of  $k \mapsto T_k$  imply that for  $n$  sufficiently large,

$$P_{0,\omega} \left( \sup_{t \leq T} \left| \frac{T_{\lfloor tn \rfloor}}{n} - \Theta t \right| < \delta \right) > 1 - \epsilon. \quad (6.26)$$

On the intersection of these events, the equality  $\tilde{B}'_n(k/n) = \tilde{B}_n(T_k/n)$  yields

$$\max_{0 \leq k \leq \lfloor Tn \rfloor} |\tilde{B}'_n(k/n) - \tilde{B}_n(\Theta k/n)| < \epsilon. \quad (6.27)$$

In light of piece-wise linearity this shows that, with probability at least  $1 - 2\epsilon$ , the paths  $t \mapsto \tilde{B}'_n(t)$  and  $t \mapsto \tilde{B}_n(\Theta t)$  are within a multiple of  $\epsilon$  in the supremum norm of each other. In particular, if  $B_t$  denotes the weak limit of the process  $(B_n(t) : t \leq T)$ , then  $(\tilde{B}'_n(t) : t \leq T)$  converges in law to  $(B_{\Theta t} : t \leq T)$ . The latter is an isotropic Brownian motion with diffusion constant  $D' = D\Theta^2$ .  $\square$

## A. Heat-kernel upper bounds

Let  $(Z_t)_{t \geq 0}$  denote the continuous-time random walk which attempts a jump to one of its nearest-neighbors at rate one (regardless of the number of accessible neighbors). Let  $q_t^\omega(x, y)$  denote the probability that  $Z_t$  started at  $x$  is at  $y$  at time  $t$ . In his paper [3], Barlow proved the following statement: There exist constants  $C_1, C_2 \in (0, \infty)$  and, for each  $x \in \mathbb{Z}^d$ , a random variable  $S(x) = S(x, \omega) \in (0, \infty)$  such that for all  $x, y \in \mathcal{C}(\omega)$  and all  $t > S(x)$ ,

$$q_t^\omega(x, y) \leq C_1 t^{-d/2} \exp\{-C_2|x - y|^2/t\}. \quad (A.1)$$

Moreover,  $S(x)$  has uniformly stretched-exponential tails, i.e.,

$$\mathbb{P}_0(S(x) > R) \leq e^{-C_3 R^{\theta'}}, \quad R \geq 1. \quad (A.2)$$

Barlow provides also a corresponding, and significantly harder-to-prove lower bound which requires the additional condition  $t > |x - y|$ . However, for (A.1), this condition is redundant.

In the remarks after his Theorem 1, Barlow mentions that appropriate modifications to his arguments yield the corresponding discrete time estimates. Here we present the details of these modifications which are needed to make our proof of the invariance principles in Theorems 1.1 and 1.2 complete. Notice that we do not re-prove Barlow's bounds in their full generality, just the absolute minimum necessary for our purposes.

### A.1. Uniform bound

There will be two kinds of bounds on the heat-kernel as a function of the terminal position of the walk after  $n$  steps: a uniform bound by a constant times  $n^{-d/2}$  and a non-uniform, Gaussian bound on the tails. We begin with the statement of the uniform upper bound:

**Proposition A.1.** *Let  $d \geq 2$  and let  $p > p_c(d)$ . There exists a random variable  $C = C(\omega)$  with  $\mathbb{P}(C < \infty) = 1$  such that for all  $\omega \in \Omega_0$  and all  $x \in \mathcal{C}_\infty$ ,*

$$P_{0,\omega}(X_n = x) \leq \frac{C(\omega)}{n^{d/2}}, \quad n \geq 1. \quad (\text{A.3})$$

The proof will invoke the isoperimetric bound from Barlow [3]:

**Lemma A.2.** *There exists a constant  $c \in (0, \infty)$  such that for  $\mathbb{P}_0$ -almost every  $\omega$  and all  $R$  sufficiently large,*

$$\frac{|\partial\Lambda|}{|\Lambda|} \geq c|\Lambda|^{-1/d} \quad (\text{A.4})$$

for all  $\Lambda \subset \mathcal{C}_\infty \cap [-R, R]^d$  such that  $|\Lambda| > R^{0.01}$ .

*Proof.* This is a consequence of Proposition 2.11 on page 3042, and Lemma 2.13 on page 3045 of Barlow's paper [3].  $\square$

This isoperimetric bound will be combined with the technique of *evolving sets*, developed by Morris and Peres [32], whose salient features we will now recall. Consider a Markov chain on a countable state-space  $V$ , let  $p(x, y)$  be the transition kernel and let  $\pi$  be a stationary measure. Let  $Q(x, y) = \pi(x)p(x, y)$  and for each  $S_1, S_2 \subset V$ , let  $Q(S_1, S_2) = \sum_{x \in S_1} \sum_{y \in S_2} Q(x, y)$ . For each set  $S \subset V$  with finite non-zero total measure  $\pi(S)$  we define the *conductance*  $\Phi_S$  by

$$\Phi_S = \frac{Q(S, S^c)}{\pi(S)}. \quad (\text{A.5})$$

For sufficiently large  $r$ , we also define the function

$$\Phi(r) = \inf\{\Phi_S : \pi(S) \leq r\}. \quad (\text{A.6})$$

The following is the content of Theorem 2 in Morris and Peres [32]: Suppose that  $p(x, x) \geq \gamma$  for some  $\gamma \in (0, 1/2]$  and all  $x \in V$ . Let  $\epsilon > 0$  and  $x, y \in V$ . If  $n$  is so large that

$$n \geq 1 + \frac{(1-\gamma)^2}{\gamma^2} \int_{4[\pi(x) \wedge \pi(y)]}^{4/\epsilon} \frac{4}{u\Phi(u)^2} du, \quad (\text{A.7})$$

then

$$p^n(x, y) \leq \epsilon\pi(y). \quad (\text{A.8})$$

Equipped with this powerful result, we are now ready to complete the proof of Proposition A.1:

*Proof of Proposition A.1.* First we will prove the desired bound for even times. Fix  $\omega \in \Omega$  and let  $Y_n = X_{2n}$  be the random walk on  $\mathcal{C}_\infty(\omega)$  observed only at even times. For each  $x, y \in \mathcal{C}_\infty(\omega)$ , let us use  $p(x, y)$  to denote the transition probability  $P_{x,\omega}(Y_1 = y)$ . Let  $\pi(x)$  denote the degree of  $x$  on  $\mathcal{C}_\infty(\omega)$ . Then  $\pi$  is an invariant measure of this chain. Moreover, by our restriction to even times we have  $p(x, x) \geq (2d)^{-2} > 0$  and so (A.7–A.8) can be applied.

By Lemma A.2 we have that  $\Phi_S \geq c\pi(S)^{-\frac{1}{d}}$  for some  $c > 0$  and all sets  $S$  of the form  $S = \mathcal{C}_\infty \cap [-R, R]^d$  for  $R \gg 1$ . Hence  $\Phi(r) \leq c' r^{-\frac{1}{d}}$  for some finite  $c' = c'(\omega)$ . Plugging into the integral (A.7) and using that  $\pi$  is bounded, we find that if  $n \geq \tilde{c} \epsilon^{-\frac{2}{d}}$ , then (A.8) holds. Here  $\tilde{c}$  is a positive constant that may depend on  $\omega$ . Choosing the minimal  $n$  possible, and applying  $p^n(x, y) = P_{x,\omega}(Y_n = y)$ , the bound (A.8) proves the desired claim for all even times. To extend the result to odd times, we apply the Markov property at time one.  $\square$

## A.2. Gaussian tails

Next we will attend to the Gaussian-tail bound. Given the random variables  $S(x, \omega)$  from (A.1–A.2), define random variables  $N_x = N_x(\omega)$  by

$$N_x = S(x) \vee \sup_{y: y \neq x} \frac{S(y)^2}{|y - x|}. \quad (\text{A.9})$$

Here is a restatement of the corresponding bound from Theorem 6.3:

**Proposition A.3.** *Let  $d \geq 2$  and  $p > p_c(d)$ . There exist constants  $c_1, c_2 \in (0, \infty)$  such that for all  $\omega \in \Omega_0$ , all  $x \in \mathcal{C}_\infty(\omega)$ , all  $R \geq 1$  and all  $n > N_x(\omega)$ ,*

$$\sum_{y: |y-x| > R} P_{x,\omega}(X_n = y) < c_1 \exp\{-c_2 R^2/n\}. \quad (\text{A.10})$$

*Proof.* The proof is an adaptation of Barlow's Theorem 1 to the discrete setting. Let  $(X_n)$  be the discrete time random walk, and let  $(Z_t)_{t \geq 0}$  be the continuous time random walk with jumps occurring at rate 1, both started at  $x$ . We consider the coupling of the two walks such that they make the same moves. We will use  $P$  and  $E$  to denote the coupling measure and the corresponding expectation, respectively.

Let  $n \geq N_x$  and let  $A_n$  be the event that  $|X_n - x| > R$ . Pick  $K > 1$  and let

$$I_n = \int_n^{4n} \mathbf{1}_{\{|Z_t - x| > R/K\}} dt \quad (\text{A.11})$$

be the amount of time in  $[n, 4n]$  that the walk  $(Z_t)$  spends at distance larger than  $R/K$  from  $x$ . By the inequality

$$P(A_n) \leq \frac{E(I_n)}{E(I_n|A_n)}, \quad (\text{A.12})$$

it suffices to derive an appropriate upper bound on  $E(I_n)$  and a matching lower bound on  $E(I_n|A_n)$ . Note that we may assume that  $R \leq n$  because otherwise we have  $P(A_n) = 0$  and there is nothing to prove.



To derive an upper bound on  $E(I_n)$ , we note that for  $t > n$ , our choice  $n \geq N_x$  implies  $t > S(x)$ . The expectation can then be bounded using (A.1):

$$\begin{aligned} E(I_n) &= \int_n^{4n} \sum_{y: |y-x| > R/K} q_t(x, y) dt \\ &\leq C_1 \int_n^{4n} t^{-d/2} \sum_{x: |x| > R/K} e^{-C_2|x|^2/t} dt \leq C_4 n e^{-C_5 \frac{R^2}{n}}, \end{aligned} \quad (\text{A.13})$$

where  $C_4$  and  $C_5$  are constants (possibly depending on  $K$ ).

It thus remains to prove that, for some constant  $C_6 > 0$ ,

$$E(I_n | A_n) \geq C_6 n. \quad (\text{A.14})$$

To derive this inequality, let us recall that the transitions of  $Z_t$  happen at rate one, and they are independent of the path of the walk. Hence, if  $B_n$  is the event that  $Z_t$  attempted at least  $n$  jumps by time  $2n$ , then  $P(B_n | A_n) = P(B_n)$  is bounded away from zero for all  $n \geq 1$ . Therefore, it suffices to prove that  $E(I_n | A_n \cap B_n) \geq C_6 n$ .

Let  $T$  be the first time when the walk ( $Z_t$ ) is farther from  $x$  than  $R$ . On  $A_n \cap B_n$ , this happens before time  $2n$ , i.e.,  $T \leq 2n$ . Let  $Q_R = [-R, R]^d \cap \mathbb{Z}^d$  and  $Q_{R/K} = [-R/K, R/K]^d \cap \mathbb{Z}^d$ . Then for values  $z$  on the external boundary of  $Q_R$ —which are those that  $Z_T$  can take—the bound (A.1) tells us

$$\sum_{y \in Q_{R/K}} q_t^o(z, y) \leq C_1 \left( \frac{2R}{K} \right)^d \max_{s>0} \left\{ s^{-d/2} \exp\left(-\frac{1}{4} C_2 R^2/s\right) \right\} \leq C_7 K^{-d}, \quad (\text{A.15})$$

provided that  $t > S(z)$ . But our assumptions  $n \geq N_x$  and  $R \leq n$  imply  $n \geq S(z)$ , and so in light of the fact that  $T \leq 2n$  on  $A_n \cap B_n$ , (A.15) actually holds for all  $t$  such that  $T + t \in [3n, 4n]$ . Plugging  $Z_T$  for  $z$  on the left-hand side and taking expectation gets us an upper bound on  $P(Z_t \in Q_{R/K} | A_n \cap B_n)$ —with  $t$  now playing the role of  $T + t$ . Hence,

$$E(I_n | A_n \cap B_n) \geq \int_{3n}^{4n} P(Z_t \notin Q_{R/K} | A_n \cap B_n) \geq n(1 - C_7 K^{-d}). \quad (\text{A.16})$$

Choosing  $K$  sufficiently large, the right-hand side grows linearly in  $n$ .  $\square$

*Proof of Theorem 6.3.* Part (1) is a direct consequence of Proposition A.1, while part (2) follows from Proposition A.3 and the fact that if the  $S(x)$  have stretched exponential tails (uniformly in  $x$ ), then so do the  $N_x$ 's.  $\square$

## B. Some questions and conjectures

While our control of the corrector in  $d \geq 3$  is sufficient to push the proof of the functional CLT through, it is not sufficient to provide the *conceptually correct* proof of the kind we have constructed for  $d = 2$ . However, we do not see any reason why  $d \geq 3$  should be different from  $d = 2$ , so our first conjecture is:

**Conjecture 1.** *Theorem 5.1 is true in all  $d \geq 2$ .*

Our proof of Theorem 5.1 in  $d = 2$  hinged on the fact that the corrector plus the position is a harmonic function on the percolation cluster. Of interest is the question whether harmonicity is an essential ingredient or just mere convenience. Yuval Peres suggested the following generalization of Conjecture 1:

**Question 2.** *Let  $f: \mathbb{Z}^d \rightarrow \mathbb{R}$  be a shift invariant, ergodic process on  $\mathbb{Z}^d$  whose gradients are in  $L^1$  and have expectation zero. Is it true that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \max_{x \in \mathbb{Z}^d \cap [-n, n]^d} |f(x)| = 0 \quad (\text{B.1})$$

*almost surely?*

*Update:* The above question, while obviously true in  $d = 1$ , has a negative answer in all  $d \geq 2$ . The first counterexample, based on constructions in [43] and [11], was provided to us by Martin Zerner. Later Tom Liggett pointed out the following, embarrassingly simple, counterexample: Let  $f(x)$  be i.i.d. with distribution function  $P(f(x) > u) = u^{-d}$  for  $u \geq 1$ . Then  $(f(x))_{x \in \mathbb{Z}^d}$  is shift-invariant, ergodic, with  $f \in L^1$  and the gradients of  $f$  having zero mean, yet  $n^{-1} \max_{|x| \leq n} |f(x)|$  has a non-trivial distributional limit as  $n \rightarrow \infty$ .

The harmonic embedding of  $\mathcal{C}_\infty$  has been indispensable for our proofs, but it also appears to be a very interesting object in its own right. This motivates many questions about the corrector  $\chi(x, \omega)$ . Unfortunately, at the moment it is not even clear what properties make the corrector unique. The following question has been asked by Scott Sheffield:

**Question 3.** *Is it true that, for a.e.  $\omega \in \Omega_0$ , there exists only one vector-valued function  $x \mapsto \chi(x, \omega)$  on  $\mathcal{C}_\infty(\omega)$  such that  $x \mapsto x + \chi(x, \omega)$  is harmonic on  $\mathcal{C}_\infty(\omega)$ ,  $\chi(0, \omega) = 0$  and  $\chi(x, \omega)/|x| \rightarrow 0$  as  $|x| \rightarrow \infty$ ?*

If this question is answered in the affirmative, we could generate the corrector by its finite-volume approximations (this would also fully justify Fig. 1). If we restrict ourselves to functions that have the shift-invariance property (2.12), uniqueness can presumably be shown using the “electrostatic methods” from, e.g., [18]. However, it is not clear whether (2.12) holds for the corrector defined by the thermodynamic limit from finite boxes.

As to the more detailed properties of the corrector, for the purposes of the present work one would like to know how  $\chi(x, \omega)$  scales with  $x$  and whether it has a well-defined scaling limit. We believe that, in sufficiently high dimension, the corrector is actually tight:

**Conjecture 4.** *Let  $d \gg 1$ . Then for each  $\epsilon > 0$  there exists  $K < \infty$  such that  $\mathbb{P}_0(|\chi(x, \omega)| > K | x \in \mathcal{C}_\infty) < \epsilon$  for all  $x \in \mathbb{Z}^d$ .*

It appears that one might be able to prove Conjecture 4 by using Barlow’s heat-kernel estimates. To capture the behavior in low dimensions, we make a somewhat wilder guess:

**Conjecture 5.** *Let  $d \geq 1$ . Then the law of  $x \mapsto \epsilon^{\frac{2-d}{2}} \chi(\lfloor x/\epsilon \rfloor)$  on compact subsets of  $\mathbb{R}^d$  converges weakly (as  $\epsilon \downarrow 0$ ) to Gaussian Free Field, i.e., a multivariate Gaussian field with covariance proportional to  $\Delta^{-1} \mathbf{1}$ , where  $\Delta$  is the Dirichlet Laplacian on  $\mathbb{R}^d$  and  $\mathbf{1}$  is the  $d$ -dimensional unit matrix.*

Here is a heuristic reasoning that led us to these conjectures: Consider the problem of random conductances to avoid problems with conditioning on containment in the infinite cluster. To show the above convergence, we need that for any smooth  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  with compact support,

$$\epsilon^d \sum_{x \in \mathbb{Z}^d} (\Delta f)(x\epsilon) \epsilon^{\frac{2-d}{2}} \chi(x) \xrightarrow[\epsilon \downarrow 0]{\mathcal{D}} \mathcal{N}(0, \sigma^2 \|\nabla f\|_2^2 \mathbf{1}), \quad (\text{B.2})$$

where  $\nabla$  and  $\Delta$  denote the (continuous) gradient and Laplacian, respectively, and where  $\mathcal{N}(0, C)$  is a mean-zero, covariance- $C$  multivariate normal random variable. Next we note that the corrector is defined, more or less, as the solution to the equation  $\Delta_d \chi = -V$ , where  $V$  is the local drift and  $\Delta_d$  is the relevant generator, which is basically a discrete Laplacian on  $\mathbb{Z}^d$ . Thus, if  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  is smooth with compact support and  $g_\epsilon(x) = g(x\epsilon)$ , then

$$\begin{aligned} \epsilon^{\frac{d+2}{2}} \sum_{x \in \mathbb{Z}^d} \epsilon^{-2} (\Delta_d g_\epsilon)(x) \chi(x) &= -\epsilon^{\frac{d-2}{2}} \sum_{x \in \mathbb{Z}^d} g_\epsilon(x) V(x) \\ &= \epsilon^{d/2} \sum_{e: |e|=1} \sum_{x \in \mathbb{Z}^d} \frac{g_\epsilon(x+e) - g_\epsilon(x)}{\epsilon} e \omega_{(x, x+e)} \\ &\xrightarrow[\epsilon \downarrow 0]{\mathcal{D}} \mathcal{N}(0, \|\nabla g\|_2^2 \mathbf{1}) \end{aligned} \quad (\text{B.3})$$

The convergence statement (B.2) would then follow from (B.3) provided we can replace the “discretized” Laplacian  $\epsilon^{-2} \Delta_d g_\epsilon$  by its continuous counterpart  $\Delta g$ .

Note that for  $d = 1$  and conductances bounded away from zero, Conjecture 5 is actually a theorem. Indeed, the corrector is a random walk with increments given by reciprocal conductances and so the convergence follows by the invariance principle for random walks. Conjecture 5 suggests that Conjecture 4 applies for  $d \geq 3$ .

Despite the emphasis on the harmonic embedding of  $\mathcal{C}_\infty$ , our proofs used, quite significantly, the underlying group structure of  $\mathbb{Z}^d$ ; e.g., in Sect. 4. Presumably this will not prevent application of our method to other regular lattices, but for more irregular graphs, e.g., Voronoi percolation in  $\mathbb{R}^d$ , significant changes may be necessary. A similar discussion applies to various natural subdomains of  $\mathbb{Z}^d$ ; for instance, it is not clear how to adapt our proof to random walk on the infinite percolation cluster in the half-space  $\{x \in \mathbb{Z}^d: x_d \geq 0\}$ .

A different direction of generalizations are the models of *long-range* percolation with power-law decay of bond probabilities. Here we conjecture:

**Conjecture 6.** *Let  $d \geq 1$  and consider long-range percolation obtained by adding to  $\mathbb{Z}^d$  a bond between every two distinct sites  $x, y \in \mathbb{Z}^d$  independently with probability proportional to  $|x - y|^{-(d+\alpha)}$ . If  $\alpha \in (0, 2)$ , then the corresponding random walk scales to a symmetric  $\alpha$ -stable Levy process in  $\mathbb{R}^d$ .*

Note that, according to this conjecture, in  $d = 1$ , the interval  $\alpha \in (0, 2)$  of “interesting” exponents is larger than the interval for which an infinite connected component may occur even without the “help” of nearest neighbor connections. On the other hand, in dimensions  $d \geq 3$ , the interval conjectured for stable convergence is strictly smaller than that of “genuine” long-range percolation behavior, as defined, e.g., in terms of the scaling of graph distance with Euclidean distance; cf [4, 5, 8].

## Acknowledgments

The research of M.B. was supported by the NSF grant DMS-0306167. Part of the research was performed while N.B. visited ETH-FIM and M.B. visited Microsoft Research in Redmond. We wish to thank these institutions for their hospitality and financial support. We are also grateful to G.Y. Amir, A. De Masi, A. Dembo, P. Ferrari, T. Liggett, S. Olla, Y. Peres, O. Schramm, S. Sheffield, V. Sidoravicius, A.-S. Sznitman and M.P.W. Zerner for interesting and useful discussions at various stages of this project.

## References

1. O. Angel, I. Benjamini, N. Berger and Y. Peres (2000). Transience of percolation clusters on wedges. *Electron. J. Probab.* (to appear).
2. P. Antal and A. Pisztor (1996). On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24**, no. 2, 1036–1048.
3. M.T. Barlow (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32**, no. 4, 3024–3084.
4. I. Benjamini and N. Berger (2001). The diameter of long-range percolation clusters on finite cycles. *Random Struct. Alg.* **19**, no. 2, 102–111.
5. N. Berger (2002). Transience, recurrence and critical behavior for long-range percolation. *Commun. Math. Phys.* **226** 531–558.
6. N. Berger, N. Gantert and Y. Peres (2003). The speed of biased random walk on percolation clusters, *Probab. Theory Related Fields* **126**, no. 2, 221–242.
7. P. Billingsley (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York-London-Sydney.
8. M. Biskup (2004). On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.* **32**, no. 4, 2938–2977.
9. B. Bollobás (2001). *Random Graphs* (Second edition), Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge.
10. E. Bolthausen and A.-S. Sznitman (2002). *Ten Lectures on Random Media*, DMV Seminar, vol. 32, Birkhäuser Verlag, Basel.
11. M. Bramson, O. Zeitouni and M.P.W. Zerner (2005) Shortest Spanning trees and a counter example for random walks in random environment. Preprint (arxiv:math.PR/0501533).
12. R.M. Burton and M. Keane (1989). Density and uniqueness in percolation. *Commun. Math. Phys.* **121**, no. 3, 501–505.
13. A. De Masi, P.A. Ferrari, S. Goldstein and W.D. Wick (1985). Invariance principle for reversible Markov processes with application to diffusion in the percolation regime. In: *Particle Systems, Random Media and Large Deviations (Brunswick, Maine)*, pp. 71–85, *Contemp. Math.*, **41**, Amer. Math. Soc., Providence, RI.
14. A. De Masi, P.A. Ferrari, S. Goldstein and W.D. Wick (1989). An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Statist. Phys.* **55**, no. 3-4, 787–855.

15. R. Durrett (2005). *Probability: Theory and Examples* (Third edition), Brooks/Cole—Thomson Learning, Belmont, CA.
16. S.N. Ethier and T.G. Kurtz (1986). *Markov Processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.
17. P.G. de Gennes (1976). La percolation: Un concept unificateur. *La Recherche* **7**, 919–927.
18. K. Golden and G. Papanicolaou (1983). Bounds for effective parameters of heterogeneous media by analytic continuation. *Commun. Math. Phys.* **90**, no. 4, 473–491.
19. G.R. Grimmett (1999). *Percolation* (Second edition), Grundlehren der Mathematischen Wissenschaften, vol. 321. Springer-Verlag, Berlin.
20. G.R. Grimmett, H. Kesten, and Y. Zhang (1993). Random walk on the infinite cluster of the percolation model. *Probab. Theory Related Fields* **96**, no. 1, 33–44.
21. G.R. Grimmett and J.M. Marstrand (1990). The supercritical phase of percolation is well behaved, *Proc. Roy. Soc. London Ser. A* **430**, no. 1879, 439–457.
22. A.B. Harris, Y. Meir and A. Aharony (1987). Diffusion on percolating clusters, *Phys. Rev. B* (3) **36**, no. 16, 8752–8764.
23. D. Heicklen and C. Hoffman (2005). Return probabilities of a simple random walk on percolation clusters. *Electron. J. Probab.* **10**, no. 8, 250–302 (electronic).
24. C. Hoffman (2000). Energy of flows on  $Z^2$  percolation clusters. *Random Structures Algorithms* **16**, no. 2, 143–155.
25. C. Hoffman (2005). Geodesics in first passage percolation. Preprint (arxiv:math.PR/0508114).
26. C. Hoffman and E. Mossel (2001). Energy of flows on percolation clusters. *Potential Anal.* **14**, no. 4, 375–385.
27. C. Kipnis, and S.R.S Varadhan (1986). A central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.* **104**, no. 1, 1–19.
28. S.M. Kozlov (1985). The method of averaging and walks in inhomogeneous environments. *Russian Math. Surveys* **40**, no. 2, 73–145.
29. D. Levin and Y. Peres (1999). Energy and cutsets in infinite percolation clusters. In: *Random walks and discrete potential theory (Cortona, 1997)*, pp. 265–278, Sympos. Math., XXXIX, Cambridge Univ. Press, Cambridge.
30. P. Mathieu and A.L. Piatnitski (2005). Quenched invariance principles for random walks on percolation clusters. Preprint (arxiv:math.PR/0505672).
31. P. Mathieu and E. Remy (2004). Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.* **32**, no. 1A, 100–128.
32. B. Morris and Y. Peres (2003). Evolving sets, mixing and heat kernel bounds. *Probab. Theory Rel. Fields* (to appear).
33. E. Mossel and O. Häggström (1998). Nearest-neighbor walks with low predictability profile and percolation in  $2 + \epsilon$  dimensions. *Ann. Probab.* **26**, no. 3, 1212–1231.
34. S. Olla (1994). *Homogenization of Diffusion Processes in Random Fields*. Manuscript of Centre de Mathématiques Appliquées (available at <http://www.ceremade.dauphine.fr/~olla/lho.ps>).
35. G. Papanicolaou and S.R.S. Varadhan (1979). *Boundary value problems with rapidly oscillating random coefficients*. Colloquia Mathematica Societatis János Bolyai, vol. 27. Random Fields, Esztergom (Hungary), pp. 835–873.
36. K. Petersen (1989). *Ergodic Theory*. Corrected reprint of the 1983 original. Cambridge Studies in Advanced Mathematics, vol 2. Cambridge University Press, Cambridge.
37. F. Rassoul-Agha and T. Seppäläinen (2004). An almost sure invariance principle for random walks in a space-time random environment. *Probab. Theory Rel. Fields* (to appear).
38. V. Sidoravicius and A.-S. Sznitman (2004). Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Rel. Fields* **129**, no. 2, 219–244.

39. A.-S. Sznitman (2003). On the anisotropic walk on the supercritical percolation cluster. *Comm. Math. Phys.* **240**, no. 1-2, 123–148.
40. C.E. Soteris and S.G. Whittington (2004). The statistical mechanics of random copolymers, *J. Phys. A* **37**, no. 41, R279–R325.
41. M. Talagrand (2003). *Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models*, A Series of Modern Surveys in Mathematics, vol. 46, Springer-Verlag, Berlin.
42. O. Zeitouni (2004). *Random walks in random environment*. Lectures on probability theory and statistics, pp. 189–312, Lecture Notes in Math., 1837, Springer, Berlin.
43. M.P.W. Zerner and F. Merkl (2001). A zero-one law for planar random walk in random environment. *Ann. Probab.* **29**, no. 4, 1716–1732.