QUENCHED INVARIANCE PRINCIPLE FOR A CLASS OF RANDOM CONDUCTANCE MODELS WITH LONG-RANGE JUMPS

MAREK BISKUP AND TAKASHI KUMAGAI

Department of Mathematics, UCLA, Los Angeles, California, USA
Research Institute in Mathematical Sciences, Kyoto University, Kyoto, Japan

Abstract: We study random walks on $\mathbb{Z}^d$ among random conductances $\{C_{xy}\colon x,y\in\mathbb{Z}^d\}$ that permit jumps of arbitrary length. Apart from joint ergodicity with respect to spatial shifts, we assume only that the nearest-neighbor conductances are uniformly positive and that $\sum_{x\in\mathbb{Z}^d}C_{0x}|x|^2$ is integrable. Our focus is on the Quenched Invariance Principle (QIP) which we establish in all $d\ge3$ by a combination of corrector methods and heat-kernel technology. In particular, a QIP thus holds for random walks on long-range percolation graphs with exponents larger than $d+2$ in all $d\ge3$, provided all nearest-neighbor edges are present. We then show that, for long-range percolation with exponents between $d+2$ and $2d$, the corrector fails to be sublinear everywhere. Similar examples are constructed also for nearest-neighbor, ergodic conductances in $d\ge4$ under the conditions close to, albeit not exactly, complementary to those of the recent work of S. Andres, M. Slowik and J.-D. Deuschel.

1. Introduction

Random walks among random conductances have seen much interest in recent years. The term "random walk" actually refers to a Markov chain whose states will be confined, for the purpose of the present paper, to the d-dimensional hypercubic lattice $\mathbb{Z}^d$ and the transition probabilities $\mathsf{P}(x,y)$ determined by a collection $\{C_{xy}\colon x,y\in\mathbb{Z}^d\}$ of non-negative numbers via

$$\mathsf{P}(x,y):=\frac{C_{xy}}{\pi(x)},\quad\text{where}\quad \pi(x):=\sum_{y\in\mathbb{Z}^d}C_{xy}. \tag{1.1}$$

(Of course, one needs to assume that $\pi(x)\in(0,\infty)$ for all x in $\mathbb{Z}^d$ to make the above meaningful.) The symmetry condition

$$C_{xy}=C_{yx},\qquad x,y\in\mathbb{Z}^d, \tag{1.2}$$

is imposed and the common value is called the *conductance* of unordered edge $\langle x,y\rangle$. As is easily checked, π is then a reversible measure for the chain. The setting naturally includes the cases when only *nearest-neighbor* conductances are present, i.e., those for which $C_{xy}:=0$ whenever x and y are not nearest neighbors in $\mathbb{Z}^d$ (this includes $x=y$).

Many "ordinary" random walks are naturally covered by the above setting; notably, the simple random walk when C_{xy} is set to one for nearest neighbors x and y and zero otherwise, or random walks with α-stable tail when $C_{xy}:=|x-y|^{-(d+\alpha)}$, where $|x|$ denotes the Euclidean norm of x, etc. Our interest here is in the situation when the family of conductances $\{C_{xy}\colon x,y\in\mathbb{Z}^d\}$ is itself random. Writing $\mathbb{P}$ for the law of the conductances and $\mathbb{E}$ for its expectation, we impose:

Assumption 1.1 Throughout we assume:

(1) $\mathbb{P}$ is stationary and jointly ergodic with respect to the shifts of $\mathbb{Z}^d$.
(2) Denoting the origin of $\mathbb{Z}^d$ by "0", we have $\mathbb{E}C_{00}<\infty$.

In this framework we may then ask what conditions guarantee various properties known for the "ordinary" random walks with symmetric jumps, e.g., lack of speed, recurrence/transience, etc. Here we will focus on the validity of an Invariance Principle, i.e., convergence of the path-law to Brownian motion under the diffusive scaling of space and time.

The so called *Annealed* (or *Averaged*) version of an Invariance Principle has been known since late 1980s (Kipnis and Varadhan [23], De Masi, Ferrari, Goldstein and Wick [20, 21]). The adjective "annealed" revers to the fact that the convergence takes place for a joint law of the paths of the chain and the environment. (A caveat is that the law used for taking average over the environment is not $\mathbb{P}$ but that obtained by weighing $\mathbb{P}$ by $\pi(0)$.) With Assumption 1.1 in force, convergence to a non-degenerate Brownian motion is then obtained under the conditions

$$\mathbb{E}\biggl(\sum_{x\in\mathbb{Z}^d}C_{0x}|x|^2\biggr)<\infty\quad\text{and}\quad\mathbb{E}\Bigl(\frac{1}{C_{0x}}\Bigr)<\infty\text{ whenever }|x|=1.\tag{1.3}$$

(Strictly speaking, the proofs are written only for nearest-neighbor models but requisite modifications are routine.) The role of the second condition is to ensure that the Brownian motion has positive variance and so, in this sense, we can regard these as optimal. Notwithstanding, extensions going beyond this have also been worked out in [20, 21], e.g., that corresponding to the simple random walk on a supercritical percolation cluster.

Much effort in the last 10-15 years went to derivations of *Individual* or *Quenched* Invariance Principles (QIP) where convergence to Brownian motion takes place for a.e. sample of random conductances. The first study in this vain is Sidoravicius and Sznitman [32] where a QIP was proved for all uniformly elliptic nearest-neighbor conductances, i.e., those where C_{xy} is bounded between two positive deterministic constants when $|x-y|=1$ and $C_{xy}=0$ otherwise, still subject to Assumption 1.1(1). A key additional ingredient on top of those used before was the validity of a diffusive (heat-kernel) upper bound for the transition probability $\mathsf{P}^n(x,y)$.

Attention subsequently shifted to degenerate problems such as the simple random walk on supercritical percolation clusters (Sidoravicius and Sznitman [32], Berger and Biskup [12], Mathieu and Piatnitski [28]). Gradual advances made by Mathieu [27], Biskup and Prescott [15], Barlow and Deuschel [8] and Andres, Barlow, Deuschel and Hambly [2] ultimately led to a full resolution of all i.i.d. nearest-neighbor random conductance models subject to the conditions

$$|x-y|=1\quad\Rightarrow\quad\mathbb{E}C_{xy}<\infty\quad\text{and}\quad\mathbb{P}(C_{xy}>0)>p_{\mathrm{c}}(d),\tag{1.4}$$

where $p_{\mathrm{c}}(d)$ denotes the critical threshold for bond percolation on $\mathbb{Z}^d$. (In $d=1$, where $p_{\mathrm{c}}(1)=1$, one needs to replace the second condition by the second condition in (1.3).)

With the case of independent conductances settled, attention moved to general (that is, dependent) non-uniformly elliptic distributions. Our understanding is, at this point, only partial and remains restricted to special cases. Staying in the realm of nearest-neighbor models satisfying

Assumption 1.1(1), the restrictions may come as a limitation on the spatial dimension: a QIP holds true in $d = 1,2$ whenever

$$|x-y| = 1 \quad \Rightarrow \quad \mathbb{E}C_{xy} < \infty \quad \text{and} \quad \mathbb{E}\Big(\frac{1}{C_{0x}}\Big) < \infty, \tag{1.5}$$

see Biskup [14, Lemma 4.8]. This is because in these dimensions one can use an observation of Berger and Biskup [12] that makes it possible to avoid all use of heat-kernel estimates. Another way to limit the form of the distribution is through decay of correlations: Procaccia, Rosenthal and Sapozhnikov [31] recently proved a QIP in correlated percolation models subject to technical conditions on correlation decay. Notably, their setting includes random walks on random interlacements and level sets of the Gaussian Free Field (both in $d \ge 3$).

A third type of restriction considered so far comes via moment conditions on individual conductances. These can be expressed by means of numbers $p,q \ge 1$ such that

$$|x-y| = 1 \quad \Rightarrow \quad C_{xy} \in L^p(\mathbb{P}) \quad \text{and} \quad \frac{1}{C_{xy}} \in L^q(\mathbb{P}). \tag{1.6}$$

For these Andres, Deuschel and Slowik [3] succeeded in proving a QIP under the condition

$$\frac{1}{p} + \frac{1}{q} < \frac{2}{d}. \tag{1.7}$$

A common feature all three approaches above is that they prove that the *corrector* χ, a standard object in stochastic homogenization (see Section 2 for details), is *sublinear everywhere* in the sense that

$$\lim_{n\to\infty} \frac{1}{n} \max_{x\colon |x|\le n} \big|\chi(x)\big| = 0, \qquad \mathbb{P}\text{-a.s.} \tag{1.8}$$

This is known to be sufficient to get a Brownian limit for diffusively-scaled paths of the Markov chain (see, e.g., Biskup [14, Section 4.2], Kumagai [24, Section 8.4] or [3]).

Obviously, conditions (1.6–1.7) are more restrictive than conditions (1.5) that, as we noted, are sufficient for nearest-neighbor AIP in all dimensions and QIP in $d = 1,2$. So, at this time, the range of validity of AIP in $d \ge 3$ is strictly larger than what is known for a QIP. We note that the AIP itself is not enough to guarantee a QIP. Indeed, examples of conductance laws (still nearest-neighbor) have been constructed where the AIP holds but QIP does not (Barlow, Burdzy and Timár [7]). Not surprisingly, both conditions in (1.5) fail in those examples as well.

The main goal of the present paper is to push the control of a QIP further and, in particular, include models with long-range jumps. Although we fall short of reaching the (conjecturally) optimal conditions (1.3), we explain the reasons for the discrepancy between (1.5) and (1.7) in their respective contexts. In particular, our study sheds new light on the principal obstructions that stand in the way of proving QIP under (presumably optimal) conditions (1.3).

2. MAIN RESULTS

We will invariably work with random collections of conductances $\{C_{xy} = C_{yx}\colon x,y \in \mathbb{Z}^d\}$ such that $\pi(x) \in (0,\infty)$ for a.e. sample from $\mathbb{P}$. This ensures that the transition kernel in (1.1) is well defined almost surely in all cases of interest. We will use P^x to denote the law on paths $Z := \{Z_n\colon n \ge 0\}$ of the associated discrete-time Markov chain subject to the initial condition $P^x(Z_0 = x) = 1$.

2.1 QIP for general conductances.

Our main point of interest is the validity of the Quenched Invariance Principle — or QIP for short — which we formalize as follows:

Definition 2.1 *Let $\mathscr{C}([0,T])$ denote the space of continuous functions on $[0,T]$ endowed with the supremum topology. Given a path $\{Z_n\colon n\ge 0\}$ of the chain, define*

$$B^{(n)}(t) := \frac{1}{\sqrt{n}}\Bigl(Z_{\lfloor tn\rfloor} + (tn-\lfloor tn\rfloor)(Z_{\lfloor tn\rfloor+1}-Z_{\lfloor tn\rfloor})\Bigr), \qquad t\ge 0. \tag{2.1}$$

We will say that a QIP holds if for each $T>0$ and $\mathbb{P}$-a.e. realization of the conductances, the law of $B^{(n)}$ induced on $\mathscr{C}([0,T])$ by P^0 tends weakly, as $n\to\infty$, to that of a Brownian motion whose covariance is non-degenerate and constant a.s.

Our main result in this note is then:

Theorem 2.2 *Suppose $d\ge 3$ and assume that, apart from the "basic" conditions in Assumption 1.1, the conductances obey*

$$\mathbb{E}\biggl(\sum_{x\in\mathbb{Z}^d} C_{0x}|x|^2\biggr) < \infty \quad \textit{and} \quad C_{0x}\ge 1 \textit{ whenever } |x|=1. \tag{2.2}$$

Then a QIP holds.

Theorem 2.2 has two simple corollaries that we wish to highlight for potential future reference. The first one concerns nearest-neighbor conductances only:

Corollary 2.3 *Consider any law on nearest-neighbor random conductances that is jointly ergodic with respect to shifts and such that C_{0x} is integrable and bounded below by a deterministic positive constant for all x with $|x|=1$. Then a QIP holds.*

Note that, for models with conductances bounded uniformly from below, this gives better control than the condition (1.7) from [3]. We will explain the reasons for this later.

The second corollary concerns random walks on a family of long-range percolation graphs. These graphs are obtained from a "nice" underlying graph, in our case $\mathbb{Z}^d$, by adding edges independently with probability that depends only on the displacement between the endpoints. The most typical situation arises when this probability decays as a power of the distance. However, our formulation only requires a square-summability condition.

Corollary 2.4 *Given a function $\mathfrak{p}\colon \mathbb{Z}^d\to[0,1]$ such that*

(1) $\mathfrak{p}(x)=\mathfrak{p}(-x)$ *for all* $x\in\mathbb{Z}^d$*,*
(2) $\mathfrak{p}(0)=0$ *and* $\mathfrak{p}(x)=1$ *whenever* $|x|=1$*,*
(3) $\sum_{x\in\mathbb{Z}^d}|x|^2\mathfrak{p}(x)<\infty$*,*

consider a random graph with vertices $\mathbb{Z}^d$ and an (unoriented) edge between x and y present with probability $\mathfrak{p}(y-x)$, independently of all other edges. Then a QIP holds for the simple random walk on this graph in all dimensions $d\ge 3$.

Note that our graphs are automatically connected as they contain $\mathbb{Z}^d$ as their subset. Going back to power-law decaying connection probabilities, our results show that if

$$\mathfrak{p}(x) = |x|^{-s+o(1)}, \qquad |x| \to \infty, \tag{2.3}$$

for some $s > d+2$, then a QIP holds in $d \geq 3$. We remark that, in the complementary regime $s \in (d, d+2)$, the scaling limit of the path is supposed to be a stable process with index $\alpha := s - d$. This has been proved for $\alpha \in (0,1)$ by Crawford and Sly [18,19]. Somewhat surprisingly, in $d = 1$ the regime when a QIP holds extends to all $\alpha > 1$, i.e., even beyond the summability condition (3), cf [19, Theorem 1.2] (see also Kumagai and Misumi [25, Theorem 2.2] concerning heat kernels). This is due to absence of percolation and the existence of cut-points. A corrector-based approach exists in this case as well (Zhang and Zhang [33]).

2.2 Lack of sublinearity of the corrector.

In order to give a formula for covariance, and to compare our results with those of Andres, Deuschel and Slowik [3], we have to discuss the relevant properties of the aforementioned corrector. Consider the generator $\mathsf{L} := \mathsf{P} - \mathrm{id}$ associated with the discrete-time Markov kernel P. Explicitly, L acts as

$$(\mathsf{L}f)(x) := \sum_{y \in \mathbb{Z}^d} C_{xy}\big[f(y) - f(x)\big]. \tag{2.4}$$

Under Assumption 1.1(1), the condition on the left of (2.2) then ensures the existence of a random function $\chi \colon \mathbb{Z}^d \to \mathbb{R}^d$, the desired corrector, which is characterized by the following properties:

(1) stationarity of increments under shifts of $\mathbb{Z}^d$,
(2) normalization $\chi(0) = 0$,
(3) weighted square-integrability as in $\mathbb{E}(\sum_{x \in \mathbb{Z}^d} C_{0x} |\chi(x)|^2) < \infty$,
(4) harmonicity of the function

$$\Psi(x) := x + \chi(x) \tag{2.5}$$

in the sense that

$$(\mathsf{L}\Psi)(x) = 0, \qquad x \in \mathbb{Z}^d. \tag{2.6}$$

(For this reason, Ψ is sometimes referred to as "harmonic coordinate.")

We refer to, e.g., Biskup [14, Proposition 3.7] for a detailed exposition and proofs of this otherwise completely classical material.

Remark 2.5 As can be gleaned from the proofs, in all QIPs proved in this paper, the covariance matrix $\Sigma = (\Sigma_{ij})$ of the limiting Brownian motion is related to the corrector via

$$\Sigma_{ij} = \frac{1}{\mathbb{E}\pi(0)} \mathbb{E}\Big(\sum_{x \in \mathbb{Z}^d} C_{0x} \big(x_i + \hat{\mathrm{e}}_i \cdot \chi(x)\big)\big(x_j + \hat{\mathrm{e}}_j \cdot \chi(x)\big)\Big), \tag{2.7}$$

where x_i denotes the i-th Cartesian component of x and $\hat{\mathrm{e}}_i$ denotes the unit vector in the i-th coordinate direction. Note that Σ is non-degenerate and finite due to (2.2) (see Proposition 3.9 for an explicit statement in this vain).

The condition on the right of (2.2) guarantees that χ is sublinear along coordinate directions (Biskup [14, Lemma 4.4]) — this is what immediately handles all $d = 1$ situations — and, in fact, also over a set of full density in $\mathbb{Z}^d$ with $d \geq 2$ (Biskup [14, Proposition 4.15]). This is unfortunately not enough to guarantee sublinearity everywhere in the sense of (1.8). In fact, this would be too much to ask anyway:

Theorem 2.6 *Let $d \geq 3$ and consider the long-range percolation graph obtained from $\mathbb{Z}^d$ as above with $\mathfrak{p}$ having the asymptotic* (2.3) *for some $s \in (d+2, 2d)$. Then the corrector is (well defined yet) not sublinear everywhere.*

Note that this is true despite the fact that a QIP holds. The underlying mechanism for the failure of sublinearity is existence of edges of length n with one endpoint in an $o(n)$ neighborhood of the origin. (An enhanced version of the argument would in fact show that the maximum in (1.8) grows as $n^{1+\varepsilon}$ for some $\varepsilon > 0$.) The restriction on the exponents comes from the simultaneous requirement that such "long" edges exist ($s < 2d$) and yet the corrector can be defined by a standard minimization procedure ($s > d+2$). This naturally confines us to dimensions $d \geq 3$.

A natural question arises whether such examples exist also for nearest-neighbor conductances. This is answered, at least partially, in:

Theorem 2.7 *Suppose $d \geq 4$ and let $p, q \geq 1$ be such that*

$$\frac{1}{p} + \frac{1}{q} > \frac{2}{d-1}. \tag{2.8}$$

Then there is a law $\mathbb{P}$ on nearest-neighbor conductances satisfying Assumption 1.1(1) and

$$|x| = 1 \quad \Rightarrow \quad C_{0x} \in L^p(\mathbb{P}) \quad \textit{and} \quad \frac{1}{C_{0x}} \in L^q(\mathbb{P}) \tag{2.9}$$

for which the corrector is (well defined yet) not sublinear everywhere.

We note that condition (2.8) is nearly complementary, albeit not strictly, to (1.7) under which Andres, Deuschel and Slowik [3] proved that the corrector is sublinear everywhere and thus a QIP holds. Theorem 2.7 thus rules out that the methods of [3] would yield a proof of a QIP in $d \geq 4$ under the (presumably) optimal conditions (1.5).

Of course, this is not the end of the day because, after all, for a QIP it suffices to prove sublinearity just along any typical path of the Markov chain. We believe that the source of the "problem" is that in low dimensions the walk is recurrent, and so its paths explore the entire lattice, while in higher dimensions it it transient, and so its paths have a good chance to avoid places where χ happens to be very large. Unfortunately, we currently do not see how to turn this intuition into a workable argument.

We note that sublinearity along a typical path of the chain is what underlies the known proofs of the AIP under the optimal conditions (1.3). It also plays a crucial role in the recent proof by Ba and Mathieu [4] of a QIP for a continuum diffusion in a periodic random environment.

2.3 Some open problems.

Our methods prove a QIP for long-range percolation graphs with exponents $s > d+2$, in $d \geq 3$, provided these are added to *a priori* connected $\mathbb{Z}^d$. A possible generalization is:

Problem 2.8 *Let $d \geq 1$ and, given an even function $q\colon \mathbb{Z}^d \to [0,1]$ such that* (2.3) *holds with exponent s, consider a random graph with vertex set $\mathbb{Z}^d$ and an edge between x and y present* *with probability $\mathfrak{p}(y-x)$, independently of other edges. Assuming there is an infinite connected component $\mathscr{C}^\infty$ a.s., and $s > d+2$, prove that the simple random walk on $\mathscr{C}^\infty$ obeys a QIP.*

As we shall see very early in Section 3, our method of proof breaks down in this case basically right from the outset. The same applies to two dimensions and so we pose:

Problem 2.9 *Let $d = 2$. Prove a QIP for conductance laws satisfying Assumption 1.1 and* (2.2).

Another open question concerns the violations of everywhere sublinearity of the corrector. As already noted, the condition in Theorem 2.7 does not quite meet the condition in Andres, Deuschel and Slowik [3], although they are "asymptotically close" in the limit $d \to \infty$. As we are quite unclear about what should be the correct statement, we just ask:

Question 2.10 *Let $d \geq 3$ and suppose $p, q \geq 1$ obey $2/(d-1) \geq 1/p + 1/q \geq 2/d$. Is there is a law on nearest-neighbor conductances satisfying Assumption 1.1(1) and* (2.9) *for which the corrector is not sublinear everywhere? Or is the corrector sublinear everywhere as soon as $1/p + 1/q \leq 2/(d-1)$?*

Of course, an ultimate open problem of this whole line of research is whether one can replace the condition of boundedness from below in Corollary 2.4 by a suitable (ideally, first) moment condition on C_{0x}^{-1} when $|x| = 1$. The aforementioned work of Ba and Mathieu [4] may provide some useful leads here.

3. QIP IN DIMENSIONS THREE AND ABOVE

In this section we will prove Theorem 2.2 dealing with a QIP for general conductances satisfying Assumption 1.1 and (2.2). We will invoke some facts from heat-kernel theory. This will in turn force us to work with continuous-time versions of the discrete-time process Z considered so far. A key innovation is the use of

$$\hat{\nu}(x) := \sum_{y \in \mathbb{Z}^d} C_{xy} |x-y|^2 \tag{3.1}$$

as the reference measure for the walk. Note that, by (2.2), $1 \leq \hat{\nu}(x) < \infty$ a.s. for all x.

3.1 Continuous time processes.

Recall that $Z = \{Z_n \colon n \geq 0\}$ denotes the discrete-time process with transition probabilities and associated stationary measure as defined in (1.1). We will consider two continuous-time variants of Z. The first one is the canonical variable-speed chain $X := \{X_t\}$ — the VSRW — obtained from Z by taking jumps at independent exponential times whose parameter at x is $\pi(x)$. As is well known, X is then a continuous-time Markov chain on $\mathbb{Z}^d$ with the generator

$$(\mathsf{L}_X f)(x) := \sum_y C_{xy} \big(f(y) - f(x) \big) \tag{3.2}$$

that, we note, coincides with that in (2.4). The counting measure $\hat{\mu}(x) := 1$ on $\mathbb{Z}^d$ is stationary and reversible for the chain X.

Our second, and more important, continuous-time chain $Y := \{Y_t\}$ will be a time-change of process X defined as follows:

$$Y_t := X_{A_t^{-1}}, \quad \text{where} \quad A_t^{-1} := \inf\{s \geq 0 \colon A_s > t\} \quad \text{for} \quad A_t := \int_0^t \hat{\nu}(X_s)\mathrm{d}s. \tag{3.3}$$

Then Y is a continuous-time Markov chain with generator

$$(\mathsf{L}_Y f)(x) := \frac{1}{\hat{\nu}(x)} \sum_y C_{xy}\bigl(f(y) - f(x)\bigr) \tag{3.4}$$

and Y is thus reversible with respect to $\hat{\nu}$ from (3.1). (Alternatively, Y can be defined directly from Z and independent exponentials that at x have parameter $\pi(x)/\hat{\nu}(x)$.) We will henceforth think of all three chains as defined on the same space and write P^x for the joint law of their paths where (each) chain is at x at time zero a.s.

The random processes X, Y and Z on $\mathbb{Z}^d$ naturally induce corresponding random processes on the space of random environments, via the "point of view of the particle." These are stationary and reversible with respect to the measures $\mathbb{Q}_X$, $\mathbb{Q}_Y$ and $\mathbb{Q}_Z$, respectively, defined by

$$\mathbb{Q}_X(\mathrm{d}\omega) := \mathbb{P}(\mathrm{d}\omega), \quad \mathbb{Q}_Y(\mathrm{d}\omega) := \frac{\hat{\nu}(0)}{\mathbb{E}\hat{\nu}(0)}\mathbb{P}(\mathrm{d}\omega), \quad \mathbb{Q}_Z(\mathrm{d}\omega) := \frac{\pi(0)}{\mathbb{E}\pi(0)}\mathbb{P}(\mathrm{d}\omega), \tag{3.5}$$

where ω denotes a generic element from the sample space carrying the conductance law $\mathbb{P}$. Thanks to our assumptions, all three measures are mutually absolutely continuous with respect to $\mathbb{P}$. This structure ensures absence of finite-time blow-ups:

Lemma 3.1 *Suppose Assumption 1.1 and* (2.2). *Then both* $\{X_t\}$ *and* $\{Y_t\}$ *are conservative under* P^x, *for all* $x \in \mathbb{Z}^d$ *and* $\mathbb{P}$*-a.e. sample of the conductances.*

Proof. We will invoke a standard criterion (see, e.g., Liggett [26, Chapter 2]) plus some stationarity and the fact that X and Y are derived from the discrete-time Markov chain Z. Focusing on X first, we have $X_t = Z_{N_t}$ for $N_t := \sup\{n \geq 0 \colon T_1 + \dots + T_n \leq t\}$ where, conditional on Z, the random times $\{T_k \colon k \geq 1\}$ are independent exponentials with T_k having parameter $\pi(Z_{k-1})$. Thanks to the 1st and 2nd Borel-Cantelli lemmas,

$$\sum_{k\geq 1} T_k = \infty \quad \text{a.s.} \qquad \Leftrightarrow \qquad \sum_{k\geq 1} E(T_k|Z) = \infty \quad \text{a.s.} \tag{3.6}$$

so no blow-ups occur if and only if the sum on the right diverges a.s. Now $E(T_k|Z) = 1/\pi(Z_{k-1})$ and so we need $\sum_{k\geq 0} 1/\pi(Z_k) = \infty$ a.s. The stationarity of $\mathbb{Q}_Z$ for the process on environments induced by Z implies

$$\frac{1}{n}\sum_{k=0}^{n-1} 1/\pi(Z_k) \underset{n\to\infty}{\longrightarrow} E_{\mathbb{Q}_Z}(1/\pi(0)) = 1/\mathbb{E}\pi(0) \quad \text{a.s.} \tag{3.7}$$

The limit is positive since $\mathbb{E}\pi(0) < \infty$ by (2.2). In particular, $\sum_{k\geq 0} 1/\pi(Z_k) = \infty$ a.s. as desired.

The argument for Y process is completely analogous; only that T_{k+1} is now (conditionally on Z) exponential with parameter $\pi(Z_k)/\hat{\nu}(Z_k)$. Here we need $0 < \mathbb{E}\hat{\nu}(0) < \infty$ as implied by (2.2). $\square$

3.2 Diagonal heat-kernel bounds.

We will now move to the derivation of a diagonal bound on the return probability for both X and Y processes. This will elucidate the reason why we have to restrict ourselves to $d \geq 3$. All estimates are deterministic, i.e., uniform in the underlying conductance configuration.

For each $f,g\colon \mathbb{Z}^d \to \mathbb{R}$ with finite support, introduce the Dirichlet form

$$\mathscr{E}(f,g) := \frac{1}{2}\sum_{x,y} C_{xy}\big(f(x)-f(y)\big)\big(g(x)-g(y)\big). \tag{3.8}$$

This can be represented using the canonical inner products in $\ell^2(\hat{\mu})$ and $\ell^2(\hat{\nu})$ as

$$\mathscr{E}(f,g) = -(\mathsf{L}_X f,g)_{\ell^2(\hat{\mu})} = -(\mathsf{L}_Y f,g)_{\ell^2(\hat{\nu})}. \tag{3.9}$$

Our basic observation is then:

Proposition 3.2 (Diffusive on-diagonal estimates) *Let $d \geq 2$. Then there is a constant $c_1 < \infty$ such that for all conductance configurations $\{C_{xy} = C_{yx}\colon x,y \in \mathbb{Z}^d\}$ for which $C_{xy} \geq 1$ holds whenever $|x-y| = 1$ we have*

$$P^x(X_t = x) \leq c_1 t^{-d/2}, \qquad x \in \mathbb{Z}^d,\, t > 0. \tag{3.10}$$

In $d \geq 3$ there is a constant $c_2 < \infty$ such that for all conductance configurations such that $\hat{\nu}(x) < \infty$ at all x, we have

$$P^x(Y_t = x) \leq c_2 t^{-d/2}, \qquad x \in \mathbb{Z}^d,\, t > 0. \tag{3.11}$$

Proof. Let $\|\cdot\|_p$ denote the canonical norm in $\ell^p(\hat{\mu})$. The bound (3.10) is a consequence of the Nash inequality and comparison of Dirichlet forms. Indeed, both X_t and the (constant speed) simple symmetric random walk are reversible with respect to the counting measure $\hat{\mu}$. For the simple symmetric random walk on $\mathbb{Z}^d$, the Nash inequality

$$\|f\|_2^{2+4/d} \leq c_0\, \mathscr{E}^{\rm SRW}(f,f)\, \|f\|_1^{4/d} \tag{3.12}$$

holds for some $c_0 = c_0(d) < \infty$. Since $\mathscr{E}^{\rm SRW}(f,f) \leq \mathscr{E}(f,f)$ due to the fact that the nearest-neighbor conductances are at least one, the same inequality holds also with $\mathscr{E}^{\rm SRW}$ replaced by the Dirichlet form (3.8). From the classic theory (see Carlen, Kusuoka and Stroock [16], the original work of Nash [30] or its review in Kumagai [24, Theorem 3.1.4]) it then follows that the semigroup operator $\mathsf{P}_t := \mathrm{e}^{t\mathsf{L}_X}$ for process X has $\ell^2(\hat{\mu}) \to \ell^\infty(\hat{\mu})$ norm at most $c_0'(d)t^{-d/4}$. Since

$$P^x(X_t = x) = (\delta_x, \mathsf{P}_t\delta_x)_{\ell^2(\hat{\mu})} = \|\mathsf{P}_{t/2}\delta_x\|_2^2 \tag{3.13}$$

holds by the semigroup property and self-adjointness of $\mathsf{P}_{t/2}$ on $\ell^2(\hat{\mu})$, we get (3.10).

In order to convert (3.10) into (3.11), we have to control the time change. Using the monotonicity of $s \mapsto P^x(Y_{2s} = x)$, we get

$$\begin{aligned} P^x(Y_{2t} = x) &\leq \frac{1}{t}E^x\bigg(\int_t^{2t} 1_{\{Y_s = x\}}\mathrm{d}s\bigg) \\ &= \frac{1}{t}E^x\bigg(\int_{A_t}^{A_{2t}} 1_{\{X_u = x\}}\frac{\mathrm{d}u}{A_s'}\bigg) \leq \frac{1}{t}\int_t^\infty P^x(X_u = x)\mathrm{d}u, \end{aligned} \tag{3.14}$$

where we first changed variables using $u := A_s$, then invoked the bound $\hat{\nu}(z) \geq 1$ to estimate $A'_s \geq 1$ and finally applied also $A_t \geq t$ to express the limits of the integral back using the naked t variable. Invoking (3.10) we obtain

$$\frac{1}{t}\int_t^\infty P^x(X_u = x)\mathrm{d}u \leq \frac{c_1}{t}\int_t^\infty u^{-d/2}\mathrm{d}u = c_1 t^{-d/2}\int_1^\infty u^{-d/2}\mathrm{d}u. \tag{3.15}$$

The last integral is finite as soon as $d \geq 3$. □

Remark 3.3 A more careful estimate shows that (3.14) holds with $P^x(X_u = x)$ replaced by $P^x(X_u = x, A_{2t} \geq u)$ in the final integral. Using this in (3.15) we get a bound in terms of a logarithmic moment of A_{2t} in $d = 2$ and a square-root moment of A_{2t} in $d = 1$. Unfortunately, we were not able to find a way how to bound these uniformly in the underlying conductance configuration, which is what we need for the arguments to follow. Hence our restriction to $d \geq 3$ throughout the rest of this section.

3.3 Displacement control.

Our next item of business is the control of displacement of the Y-path away from the starting point. The relevant bounds will be expressed using the quantity

$$C^*_{\mathrm{vol}}(x,R) := \frac{1}{R^d}\sum_{k=1}^\infty \frac{\hat{\nu}\bigl(B(x,2^kR)\bigr)}{2^{kd}}. \tag{3.16}$$

Note that $C^*_{\mathrm{vol}}(x,R) \geq c_0 > 0$ since $\hat{\nu}(x) \geq 1$ for all $x \in \mathbb{Z}^d$. Given $A \subset \mathbb{Z}^d$, we also define

$$\tau_A := \inf\bigl\{t \geq 0\colon Y_t \notin A\bigr\} \tag{3.17}$$

to be the first exit time of Y from A. For any $x \in \mathbb{Z}^d$ and any $r \geq 0$, let $B(x,r)$ denote the set of vertices in $\mathbb{Z}^d$ that are within ℓ^1-distance r from x.

Proposition 3.4 *Let $d \geq 3$. Then we have:*

(i) *There is $c_1 > 0$ such that, for all $x \in \mathbb{Z}^d$, all $R > 0$ and all $t > 0$,*

$$P^x(|Y_t - x| \geq R) \leq c_1 C^*_{\mathrm{vol}}(x,R)\frac{t}{R^2}. \tag{3.18}$$

(ii) *There is $c_2 > 0$ such that, for all $x \in \mathbb{Z}^d$, all $r > 0$ and all $t > 0$,*

$$P^z(\tau_{B(x,3r)} < t) \leq c_2\Bigl(\max_{y \in B(x,5r)} C^*_{\mathrm{vol}}(y,r)\Bigr)\frac{t}{r^2}, \qquad z \in B(x,r). \tag{3.19}$$

We note that, if we were only after (3.18), we could perhaps hope to use the method from Bass [10], based on Nash [30]. However, this argument harbors problems with the use of discrete Gauss-Green formula that we do not know how to overcome — and that have not in fact been treated satisfactorily in all earlier instances either (e.g., Barlow [5, (3.8) and (3.10)] or the formula for Q' in Bass [10]; the issue being the lack of a good *lower* bound on the heat kernel). As (3.19) controls the displacement of a whole path, we have to resort to a different argument anyway.

We will follow Barlow, Grigor'yan and Kumagai [9] (also Barlow, Bass, Chen and Kassmann [6]) and use so-called Meyer decomposition for jump processes (see Ikeda, Nagasawa and Watanabe [22] or Meyer [29]). The idea is to separate the jumps of the Y process into "short" and "long" ones and consider a process Y' that executes only the short jumps; Y is then a combination of strands of (various independent copies of) Y' with an occasional "long" jump stuck in-between.

To stay close to the notation of the above works, we express the transition rates using the so called *jump kernel*

$$n(x,y) := \frac{C_{xy}}{\hat{\nu}(x)\hat{\nu}(y)}. \tag{3.20}$$

The Dirichlet form can then be expressed as

$$\mathscr{E}(f,g) = \frac{1}{2}\sum_{x,y}(f(x)-f(y))(g(x)-g(y))n(x,y)\hat{\nu}(x)\hat{\nu}(y). \tag{3.21}$$

Given any decomposition $C_{xy} = C'_{xy} + C''_{xy}$ with $C'_{xy}, C''_{xy} \geq 0$, let $n'(x,y)$, resp., $n''(x,y)$ correspond to C'_{xy}, resp., C''_{xy} accordingly. Note that then

$$N(x) := \sum_y n''(x,y)\hat{\nu}(y) \leq \frac{1}{\hat{\nu}(x)}\sum_y C_{xy} \leq 1, \qquad x \in \mathbb{Z}^d. \tag{3.22}$$

In our situation we will also have $N(x) > 0$ for all x so, for each x,

$$q(x,y) := \frac{n''(x,y)\hat{\nu}(y)}{N(x)} \tag{3.23}$$

defines a probability measure on $\mathbb{Z}^d$. (Note that, compared to [9], where $q(x,y)$ is the *density* with respect to a reference measure, we need an extra factor $\hat{\nu}(y)$.)

We will write Y' for the continuous-time Markov process with generator as in (3.4) but with conductances C_{xy} replaced by C'_{xy} while keeping the normalization by $\hat{\nu}(x)$. For each $i \geq 1$ and each $y \in \mathbb{Z}^d$, we now define the following random variables:

(1) $\{Y'_{i,y}(t) \colon t \geq 0\}$, a realization of Y' process with initial value $Y_{i,y}(0) = y$ a.s.,
(2) $W_{i,y}$, a random variable sampled according to $q(y,\cdot)$,
(3) ξ_i, an exponential random variable with parameter 1.

We will regard all these random variables (i.e., for all $i \geq 1$ and all $y \in \mathbb{Z}^d$) as independent. One can then define inductively a collection of stopping times $T_0 := 0 < T_1 < T_2 < \dots$ and a process $\{\widetilde{Y}_t \colon t \geq 0\}$ such that the following holds:

$$\widetilde{Y}_s := Y'_{i,W_{i,\widetilde{Y}_{T_i^-}}}, \qquad T_i \leq s < T_{i+1}. \tag{3.24}$$

where

$$T_i := \inf\{t \geq T_{i-1} : H_t \geq \xi_1 + \cdots + \xi_i\} \qquad i \geq 1, \tag{3.25}$$

for

$$H_t := \int_0^t N(\widetilde{Y}_s)\mathrm{d}s. \tag{3.26}$$

This is possible because, in light of (3.22) we have $H_t \leq t$ and so $T_i \to \infty$ as $i \to \infty$ a.s. In particular, there are only a finite number of jumps in each bounded interval and one can impose the above conditions inductively. The key point, which is proved in Ikeda, Nagasawa and Watanabe [22] or Meyer [29], is that $\widetilde{Y}$ has the same law as Y.

Our use of this theory comes via the following lemma:

Lemma 3.5 *For any $B \subset \mathbb{Z}^d$,*

$$P^x(Y_t \in B) \le P^x(Y_t' \in B) + E^x\biggl(\int_0^t \sum_z n''(Y_s', z)\hat{\nu}(z)P^z(Y_{t-s} \in B)\,\mathrm{d}s\biggr). \tag{3.27}$$

Proof. By adapting Lemma 3.1 of Barlow, Grigor'yan and Kumagai [9] (also [6, Lemma 3.7]) to our present context and notation we readily find out that, for any $B \subset \mathbb{Z}^d$,

$$P^x(Y_t \in B) = P^x(Y_t' \in B,\, T_1 > t) + E^x\biggl(\int_0^t \sum_{z \in B} r_{t-s}(Y_s', z)N(Y_s')\mathrm{e}^{-H_s}\,\mathrm{d}s\biggr), \tag{3.28}$$

where

$$r_t(x,y) := \sum_z q(x,z)P^z(Y_t = y). \tag{3.29}$$

(Note that e^{-H_s} is missing in the integrand of the right hand side of formula (3.3) in [9], which is our (3.28). Note also that $r_t(x,y)$ here corresponds to $r_t(x,y)\hat{\nu}(y)$ in [9].) Dropping $T_1 > t$ from the first probability and e^{-H_s} from the expectation, the desired expression then follows by using the explicit form of $q(\cdot,\cdot)$. □

We will now specialize the above setting to the case at hand. Given a number $K \ge 1$, the scale separating "short" and "long" jumps, we define

$$n'(x,y) := n_K(x,y) := n(x,y)1_{\{|x-y| \le K\}} \quad \text{and} \quad n''(x,y) := n(x,y)1_{\{|x-y| > K\}}. \tag{3.30}$$

Let us denote the heat kernel associated with Y by

$$q_t(x,y) := \frac{P^x(Y_t = y)}{\hat{\nu}(x)} \tag{3.31}$$

and write

$$q_t^{(K)}(x,y) := \frac{P^x(Y_t' = y)}{\hat{\nu}(x)} \tag{3.32}$$

for the heat kernel associated with the process Y'.

Lemma 3.6 *There is a constant $\tilde{c}_1 > 0$ such that for all $K \ge 1$ and all $x_0, y_0 \in \mathbb{Z}^d$ with $|x_0 - y_0| \ge R$, where R is defined by $2R = 3(d+2)K$,*

$$q_t^{(K)}(x_0,y_0) \le \tilde{c}_1(t^{-d/2} \wedge t/R^{d+2}). \tag{3.33}$$

Proof. As $K \ge 1$, both Y and Y' fall into the class of processes treated in Proposition 3.2. Thus, from (3.11) and $\hat{\nu}(x) \ge 1$ we then infer

$$\max\bigl\{q_t(x,y),\, q_t^{(K)}(x,y)\bigr\} \le c_2 t^{-d/2}, \qquad x,y \in \mathbb{Z}^d,\, t > 0. \tag{3.34}$$

In particular, we only have to prove (3.33) for $t \le R^2$.

We will follow closely the proof of Theorem 1.4 in [9]. Let $\|\cdot\|_p$ denote the canonical norm in $\ell^p(\hat{\nu})$. By the general equivalence between heat-kernel bounds and the Nash inequality, cf Carlen, Kusuoka and Stroock [16], from the bound (3.34) for $q_t^{(K)}$ we then get

$$\|f\|_2^{2+4/d} \le c\mathscr{E}_K(f,f)\|f\|_1^{4/d}, \tag{3.35}$$

for some constant $c = c(d) < \infty$, where we let $\mathscr{E}_K$ be the Dirichlet form naturally associated with the (short-range) jump kernel n_K.

Next we will invoke an argument from Carlen, Kusuoka and Stroock [16] (based on an earlier argument of Davies) for obtaining off-diagonal heat-kernel bounds from the Nash inequality (3.35). For that we first introduce the auxiliary objects

$$\begin{aligned} \Gamma_K(\psi)(x) &:= \frac{1}{\hat{\nu}(x)}\sum_y (\mathrm{e}^{\psi(x)-\psi(y)}-1)^2 C_{xy} 1_{\{|x-y|\le K\}},\\ \Lambda(\psi)^2 &:= \big\|\Gamma_K(\psi)\big\|_\infty \vee \big\|\Gamma_K(-\psi)\big\|_\infty,\\ E_K(t,x,y) &:= \sup\big\{|\psi(x)-\psi(y)|-t\Lambda(\psi)^2:\ \Lambda(\psi)<\infty\big\}. \end{aligned} \tag{3.36}$$

Theorem 3.25 in Carlen, Kusuoka and Stroock [16] then shows

$$q_t^{(K)}(x,y) \le ct^{-d/2}\mathrm{e}^{-E_K(2t,x,y)}, \qquad x,y\in\mathbb{Z}^d, \tag{3.37}$$

for some constant $c\in(0,\infty)$. In order to bring (3.37) into a desired from, it thus suffices to supply a good lower bound on $E_K(2t,x,y)$.

We begin by picking a $\lambda>0$ and considering the test function

$$\psi(x) := \lambda\big(R-|x_0-x|\big)_+. \tag{3.38}$$

The triangle inequality gives $|\psi(x)-\psi(y)|\le\lambda|x-y|$ and from $|\mathrm{e}^t-1|^2\le t^2\mathrm{e}^{2|t|}$ and $t^2\mathrm{e}^{-|t|}\le 1$ we then get

$$\begin{aligned} \Gamma_K(\psi)(x) &= \frac{1}{\hat{\nu}(x)}\sum_y (\mathrm{e}^{\psi(x)-\psi(y)}-1)^2 C_{xy} 1_{\{|x-y|\le K\}}\\ &\le \frac{1}{\hat{\nu}(x)}\mathrm{e}^{2\lambda K}\lambda^2\sum_y |x-y|^2 C_{xy} 1_{\{|x-y|\le K\}}\\ &\le c(\lambda K)^2\mathrm{e}^{2\lambda K}K^{-2} \le c\mathrm{e}^{3\lambda K}K^{-2}. \end{aligned} \tag{3.39}$$

Since the same bound applies to $\Gamma_K(-\psi)$ as well, the fact that $\psi(x_0)=R$ while $\psi(y_0)=0$ shows

$$-E_K(2t,x_0,y_0) \le -\lambda R + ct\mathrm{e}^{3\lambda K}K^{-2} \tag{3.40}$$

For the specific choice

$$\lambda := \frac{1}{3K}\log\Big(\frac{K^2}{t}\Big) \tag{3.41}$$

the relation between K and R gives

$$-E_K(2t,x_0,y_0) \le c - \Big(\frac{2+d}{2}\Big)\log\Big(\frac{K^2}{t}\Big). \tag{3.42}$$

Thus

$$q_t^{(K)}(x_0,y_0) \le ct^{-d/2}\mathrm{e}^{-E_K(2t,x_0,y_0)} \le c't^{-d/2}\Big(\frac{t}{K^2}\Big)^{(2+d)/2} = c'\frac{t}{K^{2+d}} = c''\frac{t}{R^{2+d}}, \tag{3.43}$$

where $c,c',c'',c''\in(0,\infty)$ are constants. This is the part of (3.33) we were left to prove. □

We are now ready to prove the first part of Proposition 3.4:

Proof of Proposition 3.4(i). We will apply Lemma 3.5 for properly chosen sets B. For that we need good ways to estimate the terms on the right-hand side of (3.28). Let R be large enough so that K defined by $2R = 3(d+2)K$ obeys $K \geq 1$. First, we note that, for each $x \in \mathbb{Z}^d$,

$$\sum_{y\in B(x,K)^c} n(x,y)\hat{\nu}(y) = \frac{1}{\hat{\nu}(x)} \sum_{y\in B(x,K)^c} C_{xy} \leq \frac{1}{\hat{\nu}(x)} \sum_{y\in B(x,K)^c} C_{xy} \frac{|x-y|^2}{K^2} \leq K^{-2}. \tag{3.44}$$

For any $B \subset \mathbb{Z}^d$, this shows

$$E^x\biggl(\int_0^t \sum_z n''(Y_s,z)\hat{\nu}(z)P^z(Y_{t-s}\in B)\,\mathrm{d}s\biggr) \leq t \sup_{x\in\mathbb{Z}^d} \sum_z n''(x,z)\hat{\nu}(z) \leq tK^{-2}. \tag{3.45}$$

On the other hand, abbreviating $B_r := B(x,r)$, if $R^2 \geq t$ then Lemma 3.6 gives

$$\begin{aligned} P^x\bigl(Y_t' \in B_{2R}\smallsetminus B_R\bigr) &= \sum_{y\in B_{2R}\smallsetminus B_R} q_t^{(K)}(x,y)\hat{\nu}(y) \\ &\leq \tilde{c}_1 \frac{t}{R^{d+2}} \hat{\nu}\bigl(B_{2R}\smallsetminus B_R\bigr). \end{aligned} \tag{3.46}$$

Combining these estimates we obtain

$$P^x\bigl(Y_t \in B_{2R}\smallsetminus B_R\bigr) \leq ctR^{-2} + \tilde{c}_1 \frac{t}{R^{d+2}} \hat{\nu}\bigl(B_{2R}\smallsetminus B_R\bigr), \qquad R^2 \geq t, \tag{3.47}$$

where c is a numerical constant. This is now free of the cutoff K and so we can sum this over power-of-2 multiples of R to get

$$P^x\bigl(|Y_t - x| > R\bigr) \leq \bigl(2c + \tilde{c}_1 C^*_{\mathrm{vol}}(x,R)\bigr)\frac{t}{R^2} \tag{3.48}$$

whenever $R^2 \geq t$. As $C^*_{\mathrm{vol}}(x,R) \geq c_0$, this has the desired form. For the same reason, once $R^2 \leq t$ the claim holds true trivially as soon as c_1 is chosen large enough. □

Proof of Proposition 3.4(ii). First we note that, for $s \geq 2r$ and $t > 0$,

$$P^x\bigl(Y_{\tau_{B(x,r)}\wedge t} \notin B(x,s)\bigr) \leq 4ts^{-2}. \tag{3.49}$$

This follows from the Lévy system formula (see for example Chen and Kumagai [17, Lemma 4.7] and its application in the end of page 50 of that paper). Indeed,

$$\begin{aligned} P^x\bigl(Y_{\tau_{B(x,r)}\wedge t} \notin B(x,s)\bigr) &= E^x\biggl(\int_0^{\tau_{B(x,r)}\wedge t} \sum_{z\notin B(x,s)} \frac{C_{Y_u,z}}{\hat{\nu}(Y_u)}\,\mathrm{d}u\biggr) \\ &\leq E^x\biggl(\int_0^{\tau_{B(x,r)}\wedge t} \sum_{z\notin B(Y_u,s/2)} \frac{C_{Y_u,z}}{\hat{\nu}(Y_v)}\,\mathrm{d}u\biggr) \\ &\leq 4E^x\Bigl(\frac{\tau_{B(x,r)}\wedge t}{s^2}\Bigr) \leq 4t/s^2. \end{aligned} \tag{3.50}$$

Here the first inequality is because $s \geq 2r$ and the second inequality is due to (3.44).

Now abbreviate $\tau := \tau_{B(z,2r)}$ and $A_i := B(z,2^{i+1}r) \smallsetminus B(z,2^i r)$. Since $\tau \le \tau_{B(x,3r)}$, we have

$$\begin{aligned} P^z(\tau_{B(x,3r)} < t) &\le P^z(\tau < t) \\ &\le P^z\big(|Y_t - z| \ge r\big) + P^z\big(\tau < t, |Y_t - z| < r\big) \\ &\le cC^*_{\mathrm{vol}}(z,r)\frac{t}{r^2} + \sum_{i=1}^\infty E^z\Big(1_{\{\tau<t\}}1_{\{Y_\tau\in A_i\}}P^{Y_\tau}\big(|Y_{t-\tau} - Y_0| \ge r\big)\Big), \end{aligned} \tag{3.51}$$

where (3.18) is used in the second inequality. The $i=1$ term in the sum needs to be bounded separately from others by

$$\begin{aligned} E^z\Big(\max_{y\in B(z,4r)} P^y\big(|Y_{t-\tau} - y| \ge r\big)1_{\{\tau<t\}}\Big) &\le cE^z\Big(\max_{y\in B(z,4r)} C^*_{\mathrm{vol}}(y,r)\frac{t-\tau}{r^2}1_{\{\tau<t\}}\Big) \\ &\le c\Big(\max_{y\in B(z,4r)} C^*_{\mathrm{vol}}(y,r)\Big)\frac{t}{r^2}, \end{aligned} \tag{3.52}$$

where (3.18) is used again. The remainder of the sum is bounded using (3.49) by

$$\sum_{i=2}^\infty P^z\big(Y_{\tau\wedge t} \notin B(z,2^i r)\big) \le \sum_{i=2}^\infty \frac{4t}{(2^i r)^2} \le c'\frac{t}{r^2}. \tag{3.53}$$

Combining these together, we obtain (3.19). $\square$

3.4 Tightness.

The arguments amassed in the previous section permit us to prove tightness for diffusively-scaled process Y. Specifically, we will be able to apply a result of Aldous [1]. Unfortunately, this result if not formulated for the space $\mathscr{C}([0,T])$ but rather for the Skorohod space $\mathscr{D}([0,T])$ of functions $f\colon [0,T] \to \mathbb{R}^d$ that are right continuous on $[0,T)$ and have left limits on $(0,T]$. This space can be endowed with the standard Skorohod topology (see, e.g., Billingsley [13]) that makes it a Polish space. This in turn permits considerations of weak limits of probability measures.

A d-dimensional version of Theorem 1 of Aldous [1] then implies that the sequence $\{Y^{(n)}\}$ of processes is tight in $\mathscr{D}([0,T])$ as soon as the following two conditions are met:

(1) $Y^{(n)}_t$ is tight, as an $\mathbb{R}^d$-valued random variable, for each $t \in [0,T]$ and each $n \ge 1$, and
(2) for any sequence $\{\tau_n\}$, where τ_n is for each $n \ge 1$ a finitely-valued stopping time for the natural filtration of $Y^{(n)}$, and any $\delta_n > 0$ with $\delta_n \to 0$,

$$Y^{(n)}_{(\tau_n+\delta_n)\wedge T} - Y^{(n)}_{\tau_n\wedge T} \underset{n\to\infty}{\longrightarrow} 0 \tag{3.54}$$

in probability.

We will specialize to the choice

$$Y^{(n)}_t := \frac{1}{\sqrt{n}}Y_{nt}, \qquad t \ge 0, \tag{3.55}$$

and plug into this result to prove:

Proposition 3.7 *Let $d \ge 3$. For each $T > 0$ and a.e. realization of the conductances, the laws of $\{Y^{(n)}\colon n \ge 1\}$ induced by P^0 on $\mathscr{D}([0,T])$ are tight.*

We begin by noting:

Lemma 3.8 *Under Assumption 1.1 and* (2.2)*, there is $c_1 > 0$ and, for all $x \in \mathbb{Z}^d$, there is a random variable $R_*(x)$ with $\mathbb{P}(R_*(x) < \infty) = 1$ such that*

$$\max_{y\in B(x,s)} C^*_{\rm vol}(y,r) \le c_1\Bigl(\frac{r+s}{r}\Bigr)^d, \qquad s>0,\ r\ge R_*(x). \tag{3.56}$$

Proof. Under the stated assumptions the Spatial Ergodic Theorem tells us that

$$c := \lim_{k\to\infty} k^{-d}\hat\nu(B(x,k)) \tag{3.57}$$

exists (finitely) and is positive $\mathbb{P}$-a.s.. Hence, there exists $R_*(x)$ with $\mathbb{P}(R_*(x) < \infty) = 1$ such that for all $r \ge R_*(x)$, we have $\hat\nu(B(x,r)) \le 2cr^d$. This implies

$$\begin{aligned}\max_{y\in B(x,s)} C^*_{\rm vol}(y,r) &= \max_{y\in B(x,s)} \frac{1}{r^d}\sum_{k=0}^\infty \frac{\hat\nu(B(y,2^kr))}{2^{dk}}\\ &\le \frac{1}{r^d}\sum_{k=r}^\infty \frac{\hat\nu(B(x,2^k(r+s)))}{2^{dk}} \le 2c\Bigl(\frac{r+s}{r}\Bigr)^d,\end{aligned} \tag{3.58}$$

whenever $r \ge R_*(x)$ and $s > 0$, and so we are done. □

Proof of Proposition 3.7. Let $R_*(x)$ be as in the previous lemma. We will check that the above conditions (1-2) from Aldous [1] hold on the set where $R_*(0) < \infty$. To distinguish various processes, let us write $\tau_B(X)$ for the first exit time of process X from set B.

For (1) we note that, by Proposition 3.4 (ii),

$$P^0\bigl(\tau_{B(0,3r)}(Y^{(n)}) < t\bigr) = P^0\bigl(\tau_{B(0,3r\sqrt{n})}(Y) < nt\bigr) \le c_1 \max_{y\in B(0,5r\sqrt{n})} C^*_{\rm vol}(y,3r\sqrt{n})\frac{nt}{(3r\sqrt{n})^2}. \tag{3.59}$$

Lemma 3.8 now ensures that this is bounded by a constant times t/r^2 as soon as $3r\sqrt{n} \ge R_*(0)$. Whenever $R_*(0) < \infty$ occurs, this implies condition (1) above.

Next, pick $T > 0$ and $\varepsilon > 0$, let τ_n be stopping times bounded by T and choose δ_n with $\delta_n \downarrow 0$. A straightforward use of the strong Markov property and a union bound then show

$$P^0\bigl(|Y^{(n)}_{\tau_n+\delta_n} - Y^{(n)}_{\tau_n}| > \varepsilon\bigr) \le P^0\bigl(\tau_{B(0,3r)}(Y^{(n)}) < T\bigr) + \max_{z\in B(0,3r\sqrt{n})} P^z\bigl(|Y^{(n)}_{\delta_n} - z/\sqrt{n}| > \varepsilon\bigr) \tag{3.60}$$

Using (3.59) and Lemma 3.8, the first quantity on the right is estimated by a constant times T/r^2. For the second quantity we in turn apply Proposition 3.4 (i) along with a simple translation argument to get

$$\max_{z\in B(0,3r\sqrt{n})} P^z\bigl(|Y^{(n)}_{\delta_n} - z/\sqrt{n}| > \varepsilon\bigr) \le c_1 \max_{z\in B(0,3r\sqrt{n})} C^*_{\rm vol}(z,\varepsilon\sqrt{n})\,\frac{\delta_n n}{(\varepsilon\sqrt{n})^2}. \tag{3.61}$$

Lemma 3.8 now shows that this tends to zero as $n \to \infty$. Thus we get

$$\limsup_{n\to\infty} P^0\bigl(|Y^{(n)}_{\tau_n+\delta_n} - Y^{(n)}_{\tau_n}| > \varepsilon\bigr) \le c\,\frac{T}{r^2}, \qquad \text{on } \{R_*(0) < \infty\} \tag{3.62}$$

for some constant $c \in (0,\infty)$ regardless of ε or the choice of stopping times τ_n (as long as $\tau_n \le T$). But the left-hand side does not depend on r and so taking $r \to \infty$, we obtain condition (2) above. Aldous' Theorem 1 then implies tightness of the processes $\{Y^{(n)}\}$ on $\mathscr{D}([0,T])$. □

3.5 Proof of a QIP.

Having proved tightness, our proof of a QIP is now reduced to convergence of finite-dimensional distributions. Let $\{\widetilde{B}_t : t \geq 0\}$ denote a d-dimensional Brownian motion such that

$$E(\widetilde{B}_t) = 0 \quad \text{and} \quad E\big((v \cdot \widetilde{B}_t)^2\big) = t\,\frac{\mathbb{E}\pi(0)}{\mathbb{E}\hat{\nu}(0)}\, v \cdot \Sigma v, \qquad v \in \mathbb{R}^d,\, t \geq 0, \tag{3.63}$$

where Σ is the matrix with entries as in (2.7). Then we have:

Proposition 3.9 *Let $d \geq 3$ and consider the processes $\{Y^{(n)}\}$ from* (3.55) *with law P^0. Then the following holds on a set of conductances of full $\mathbb{P}$-measure: For each $k \geq 1$ and each $t_1, \dots, t_k$ satisfying $0 \leq t_1 < t_2 < \dots < t_k < \infty$,*

$$\big(Y^{(n)}_{t_1}, \dots, Y^{(n)}_{t_k}\big) \xrightarrow[n\to\infty]{\text{law}} (\widetilde{B}_{t_1}, \dots, \widetilde{B}_{t_k}). \tag{3.64}$$

Proof. One of the main issues in the proof is a proper demonstration of the set of conductances of full $\mathbb{P}$-measure on which (3.64) holds for all k-tuples $(t_1, \dots, t_k)$ with the stated properties. We will therefore keep careful track of all requisite events.

Let $\Psi(x)$ denote the “harmonic coordinate” function from (2.5); this is defined (and depends on) conductances in a measurable set Ω_1 with $\mathbb{P}(\Omega_1) = 1$. Given a realization of conductances and a path $Z = \{Z_n\}$ of the discrete-time Markov chain, consider the random variables $\{\Psi(Z_n)\}$. A classical argument (cf, e.g., Corollary 3.10 of Biskup [14]) based on the fact that $\Psi(Z_n)$ is a martingale implies that, under our standing assumptions, there is a measurable set $\Omega_2 \subseteq \Omega_1$ with $\mathbb{P}(\Omega_2) = 1$ such that for each realization of conductances in Ω_2, the law of

$$t \mapsto \frac{1}{\sqrt{n}}\Psi(Z_{\lfloor tn \rfloor}) \tag{3.65}$$

induced by P^0 on $\mathscr{D}([0,T])$ — in fact, even on $\mathscr{C}([0,T])$, provided we interpolate values linearly — tends to a Brownian motion with mean zero and covariance Σ.

Next we will prove a similar statement for $t \mapsto \frac{1}{\sqrt{n}}\Psi(Y_{nt})$ but for that we have to control the time change that takes Z into Y. To that end, conditionally on Z, let $T_0, T_1, \dots$ denote independent exponentials with parameters $\pi(Z_0)/\hat{\nu}(Z_0), \pi(Z_1)/\hat{\nu}(Z_1), \dots$, respectively. Then $\{\widetilde{Y}_t : t \geq 0\}$, defined by

$$\widetilde{Y}_t := Z_{N_t} \quad \text{for} \quad N_t := \max\{k \geq 0\colon T_1 + \dots + T_k \leq t\}, \tag{3.66}$$

has the law of $\{Y_t : t \geq 0\}$. Letting $\Omega_3 \subset \Omega_2$ be the subset of conductances on which

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} \frac{\hat{\nu}(Z_k)}{\pi(Z_k)} = E_{\mathbb{Q}_Z}\Big(\frac{\hat{\nu}(0)}{\pi(0)}\Big), \qquad P^0\text{-a.s.}, \tag{3.67}$$

and

$$\forall \varepsilon > 0\colon \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n} \frac{\hat{\nu}(Z_k)}{\pi(Z_k)} 1_{\{\hat{\nu}(Z_k)/\pi(Z_k) > \varepsilon n\}} = 0, \qquad P^0\text{-a.s.}, \tag{3.68}$$

the stationarity and ergodicity of $\mathbb{Q}_Z$ with respect to the chain on environments induced by Z guarantees (via the Pointwise Ergodic Theorem) that $\mathbb{P}(\Omega_3) = 1$. Invoking the Weak Law of

Large Numbers (with a simple truncation step enabled by (3.68)) and a renewal argument, we then have

$$\frac{N_t}{t} \underset{t\to\infty}{\longrightarrow} \frac{\mathbb{E}\pi(0)}{\mathbb{E}\hat{\nu}(0)}, \qquad \text{in } P^0\text{-probability} \tag{3.69}$$

for all conductances from Ω_3. In light of monotonicity of $t \mapsto N_t$, this gives a locally-uniform closeness of $s \mapsto N_{ts}/t$ to a linear function. By the definition of the Skorohod topology, the identification $\widetilde{Y} \overset{\text{law}}{=} Y$ now shows that also the law

$$t \mapsto \frac{1}{\sqrt{n}}\Psi(Y_{nt}) \tag{3.70}$$

induced by P^0 on $\mathscr{D}([0,T])$ tends to that of a Brownian motion with mean zero and covariance $(\mathbb{E}\pi(0)/\mathbb{E}\hat{\nu}(0))\Sigma$, for every realization of conductances in Ω_3.

Since convergence on $\mathscr{D}([0,T])$ to a process with continuous paths implies convergence of finite-dimensional distributions, to get (3.64) it now suffices to identify a measurable set $\Omega^\star$ of conductances with $\mathbb{P}(\Omega^\star) = 1$ such that

$$\frac{1}{\sqrt{n}}\,|\chi(Y_{tn})| \underset{n\to\infty}{\longrightarrow} 0, \qquad \text{in } P^0\text{-probability}, \tag{3.71}$$

holds on $\Omega^\star$ for each $t \geq 0$. For this we first recall the fact that the corrector is "sublinear on average." More precisely, let $\Omega_4 \subseteq \Omega_3$ be the set of conductances on which, for each $\delta > 0$,

$$\lim_{n\to\infty} \frac{1}{n^d} \sum_{|x|\leq n} 1_{\{|\chi(x)|>\delta n\}} = 0. \tag{3.72}$$

Then $\mathbb{P}(\Omega_4) = 1$; cf Proposition 4.15 of Biskup [14]. We now finally define $\Omega^\star \subset \Omega_4$ to be the subset of conductances such that the quantity from (3.16) satisfies

$$\sup_{r\geq 1} C^*_{\text{vol}}(0,r) < \infty. \tag{3.73}$$

Thanks to the Spatial Ergodic Theorem, $\hat{\nu}(B(0,k))$ grows at most as k^d and so $\mathbb{P}(\Omega^\star) = 1$.

Now consider a conductance configuration from $\Omega^\star$ and pick $t > 0$. A union bound yields

$$P^0\big(|\chi(Y_{tn})| > \varepsilon\sqrt{n}\big) \leq P^0\big(|Y_{tn}| > \varepsilon^{-1}\sqrt{n}\big) + \sum_{|x|\leq \varepsilon^{-1}\sqrt{n}} 1_{\{|\chi(x)|>\varepsilon\sqrt{n}\}} P^0(Y_{tn} = x). \tag{3.74}$$

Proposition 3.4(i) then shows that the first probability is smaller than $c_1\varepsilon^2 t$ times the quantity from (3.73). For the second term we in turn use (3.11) to bound the sum by an absolute constant times $t^{-d/2}n^{-d/2}\sum_{|x|\leq\varepsilon^{-1}\sqrt{n}} 1_{\{|\chi(x)|>\varepsilon n\}}$. Taking $n \to \infty$, invoking (3.72) and then letting $\varepsilon \downarrow 0$, we get (3.71) on $\Omega^\star$ as desired. $\square$

We note that, in the previous proof, we opted for addressing continuous time via a passage through discrete time. This makes it easier to plug into results that already exist in the literature. Let us now see how the above proposition implies our main result:

Proof of Theorem 2.2. Fix $T > 0$. Since by our assumptions $C_{0x} \geq 1$ holds almost surely for all $|x| = 1$, we are permitted to use Proposition 3.7. This tells us that the laws of $Y^{(n)}$ are tight on $\mathscr{D}([0,T])$. By Proposition 3.9 we then conclude that $Y^{(n)}$ converges in law to $\widetilde{B}$ while the time-change argument in (3.66) and (3.69) then show that $t \mapsto \frac{1}{\sqrt{n}}Z_{\lfloor tn\rfloor}$ tends in law to a centered

Brownian motion with covariance Σ. As the limit process has continuous paths, this implies the convergence of the linear interpolation B^n of Z-values from (2.1) in the space $\mathscr{C}([0,T])$. $\square$

It is now easy to verify that Corollaries 2.3 and 2.4 fall under the conditions of Theorem 2.2 and so their statements follow as well.

4. Failures of everywhere sublinearity

In this section we provide the promised counterexamples to everywhere sublinarity of the corrector and thus prove Theorems 2.6 and 2.7. We begin with the counterexample arising in the context of long-range percolation.

4.1 Long-range percolation.

Consider long-range percolation with the connection probability $\mathfrak{p}(x)$ having the asymptotic (2.3) with exponent $s \in (d+2, 2d)$. We will assume $\mathfrak{p}(0) = 0$, $\mathfrak{p}(x) = 1$ for x with $|x| = 1$ and $\mathfrak{p}(x) < 1$ for all x with $|x| > 1$. The conductances then obey

(1) $C_{xx} = 0$ for all x a.s.,
(2) $C_{xy} = 1$ whenever $|x-y| = 1$ a.s.,
(3) $\mathbb{P}(C_{xy} = 1) = \mathfrak{p}(y-x)$ whenever $|x-y| > 1$.

As already mentioned, a key point is the proof of the existence of a "long" edge of length n from $o(n)$-neighborhood of the origin. This would itself be easy to guarantee; what makes it harder is that our arguments also need that the "far away" endpoint of the "long" edge is incident to no other edges than the nearest-neighbor ones. The exact statement is the subject of:

Lemma 4.1 *Noting that $2 < s-d < d$ we may pick $\gamma \in (\frac{s-d}{d}, \frac{2s-d}{2d} \wedge 1)$ and consider the event*

$$A(x,y) := \{C_{xy} = 1\} \cap \bigl\{\forall z \in \mathbb{Z}^d \smallsetminus \{x\} \colon |y-z| > 1 \Rightarrow C_{yz} = 0\bigr\}. \tag{4.1}$$

Then

$$A_n := \bigcup_{\substack{x\in\mathbb{Z}^d\\ |x|\le n^\gamma}} \bigcup_{\substack{y\in\mathbb{Z}^d\\ n<|y|\le 2n}} A(x,y) \tag{4.2}$$

occurs for infinitely many n a.s.

Proof. Instead of (4.1) consider the event

$$\widetilde{A}(x,y) := \{C_{xy} = 1\} \cap \bigl\{\forall z \in \mathbb{Z}^d \colon |y-z| > 1 \;\&\; |z| > n^\gamma \Rightarrow C_{yz} = 0\bigr\} \tag{4.3}$$

whose advantage over $A(x,y)$ is that the two events on the right are now independent as soon as x and y are as in the union in (4.2). Set

$$\widetilde{A}_n := \bigcup_{\substack{x\in\mathbb{Z}^d\\ |x|\le n^\gamma}} \bigcup_{\substack{y\in\mathbb{Z}^d\\ n<|y|\le 2n}} \widetilde{A}(x,y). \tag{4.4}$$

Obviously, $A_n \subset \widetilde{A}_n$. Moreover, for n so large that $n^\gamma < n-1$ (note that $\gamma < 1$), on $A_n^{\mathrm{c}} \smallsetminus \widetilde{A}_n^{\mathrm{c}}$ there is an edge between some x with $|x| \le n^\gamma$ and some y with $n \le |y| \le 2n$ so that y has another edge to some x' with $|x'| \le n^\gamma$. Defining, also for later use,

$$B_n := \left\{ \exists x,x',y,y' \in \mathbb{Z}^d \colon \begin{array}{c} |x|,|x'| \le n^\gamma,\, n \le |y|,|y'| \le 2n \\ (x,y) \ne (x',y'),\, C_{xy} = 1 = C_{x'y'} \\ y \ne y' \,\Rightarrow\, C_{yy'} = 1 \end{array} \right\} \tag{4.5}$$

we thus have

$$A_n^{\mathrm{c}} \subseteq (\widetilde{A}_n^{\mathrm{c}} \cap B_n^{\mathrm{c}}) \cup B_n. \tag{4.6}$$

We will now proceed to estimate probabilities of two events on the right-hand side.

For the probability of B_n, we invoke a straightforward union bound. Let Ξ_n denote the set of all quadruples (x,x',y,y') that satisfy the geometrical conditions in event B_n. Then, for some constants $c,c' < \infty$,

$$\begin{aligned} \mathbb{P}(B_n) &\le \sum_{x,x',y,y' \in \Xi_n} \mathfrak{p}(y-x)\mathfrak{p}(y'-x')\bigl(\delta_{y,y'} + \mathfrak{p}(y-y')\bigr) \\ &\le cn^{2d\gamma-2s+o(1)} \sum_{\substack{y,y' \in \mathbb{Z}^d \\ n \le |y|,|y'| \le 2n}} \bigl(\delta_{y,y'} + \mathfrak{p}(y-y')\bigr) \le c' n^{2d\gamma-2s+d+o(1)}, \end{aligned} \tag{4.7}$$

where we first used that both $\mathfrak{p}(y-x)$ and $\mathfrak{p}(y'-x')$ are at most $n^{-s+o(1)}$, then carried out the sums over x and x' to get a constant times $n^{d\gamma}$ from each and, finally, applied that $z \mapsto \mathfrak{p}(z)$ is summable because $s > d$. Noting that, in light of our choice of γ, the final exponent in (4.7) is negative, we get that B_{2^n} occurs only for finitely many n, a.s.

Concerning the first event in (4.6), let N denote the number of edges between some x with $|x| \le n^\gamma$ and some y with $n \le |y| \le 2n$ and let $\{(x_i,y_i) \colon i = 1,\dots,N\}$ list the corresponding pairs of vertices connected by these edges. On $\widetilde{A}_n^{\mathrm{c}} \cap B_n^{\mathrm{c}}$ we then know that (once $N > 1$) all y_i are distinct and each y_i must have at least one non-nearest neighbor edge to a vertex z with $|z| > n^\gamma$ and $z \notin \{y_1,\dots,y_N\}$. Conditioning on $\mathscr{F}_n := \sigma(C_{xy} \colon |x| \le n^\gamma, n \le |y| \le 2n)$, we thus have

$$P\bigl(\widetilde{A}_n^{\mathrm{c}} \cap B_n^{\mathrm{c}} \big| \mathscr{F}_n\bigr) \le 1_{\{N=0\}} + 1_{\{N>0\}} \prod_{j=1}^N \Biggl(1 - \prod_{\substack{y \ne y_1,\dots,y_N \\ |y-y_j|>1}} \bigl(1 - \mathfrak{p}(y-y_j)\bigr)\Biggr), \tag{4.8}$$

where N and $(y_1,\dots,y_n)$ are as specified above. The product is bounded from below by

$$c := \prod_{|z|>1} (1 - \mathfrak{p}(z)) \tag{4.9}$$

which is positive by the summability of $\mathfrak{p}$ and our assumption that $\mathfrak{p}(z) < 1$ once $|z| > 1$. Hence,

$$P\bigl(\widetilde{A}_n^{\mathrm{c}} \cap B_n^{\mathrm{c}}\bigr) \le \mathbb{P}(N \le n^\delta) + (1-c)^{n^\delta} \tag{4.10}$$

holds true for any $\delta > 0$. To estimate $\mathbb{P}(N \le n^\delta)$, let

$$\tilde{q}_n := \min_{|x| \le n^\gamma} \min_{n \le |y| \le 2n} \mathfrak{p}(y-x). \tag{4.11}$$

and let V_n be the number of pairs (x,y) with $|x| \le n^\gamma$ and $n \le |y| \le 2n$. Then N is stochastically dominated from below by a binomial random variable with parameters V_n and $\tilde{q}_n$. As $V_n\tilde{q}_n = n^{d(1+\gamma)-s+o(1)}$ with $d(1+\gamma) - s > 0$ by our assumptions about γ, the probability $\mathbb{P}(N \le n^\delta)$

decays, for δ positive but small, exponentially in a power of n. Using this in (4.10), the Borel-Cantelli lemma implies that $\widetilde{A}_n^{\mathrm{c}} \cap B_n^{\mathrm{c}}$ occurs only finitely often a.s. $\square$

With the existence of the desired "long" edge established, we can move to the construction of a counterexample to everywhere sublinearity of the corrector.

Proof of Theorem 2.6. Consider the long-range percolation setting as specified above. The asymptotic (2.3) with $s > d+2$ implies $\mathbb{E}(\sum_{x\in\mathbb{Z}^d} C_{0x}|x|^2) < \infty$ and so the corrector can be defined by any of the standard methods (see, e.g., Biskup [14, Section 3] for a discussion of these). In fact, by (2), the corrector is sublinear on average (cf [14, Proposition 4.15]), meaning that $\{x\colon |\chi(x)| > \varepsilon|x|\}$ is, for each $\varepsilon > 0$, a set of zero density in $\mathbb{Z}^d$.

To show that χ is not sublinear everywhere in the sense of (1.8) we will assume, for the sake of contradiction, that for each $\varepsilon > 0$ there is a (random) $K < \infty$ such that

$$|\chi(x)| \le K + \varepsilon|x|, \qquad x \in \mathbb{Z}^d. \tag{4.12}$$

(This is equivalent to (1.8).) Suppose that A_n occurs and let x and y be the endpoints of an edge that make $A(x,y)$ in the definition of A_n occur. The harmonicity condition (2.6) for Ψ from (2.5) at point y then reads

$$x + \chi(x) - \big(y + \chi(y)\big) + \sum_{z\colon |z|=1} \big(z + \chi(y+z) - \chi(y)\big) = 0, \tag{4.13}$$

where we noted that $C_{x'y} = 1$ for $x' = x$ and x' being a neighbor of y; otherwise $C_{x'y} = 0$. Applying (4.12) and the fact that $|x|, |y|, |y+z| \le 2n+1$ for all z with $|z| = 1$ yields

$$|y - x| \le (2+4d)K + 2d + \varepsilon(2+4d)(2n+1). \tag{4.14}$$

For ε small this contradicts $|y-x| > n - n^\gamma$. Hence (4.12) cannot occur on A_n for n large enough and, since A_n does occurs of infinitely many n a.s., (4.12) fails with probability one. $\square$

4.2 Nearest-neighbor conductances.

Next we move to the context underlying Theorem 2.7. We start by defining some auxiliary processes that will be used later to construct the desired environment law $\mathbb{P}$. As all of these live on the same probability space, we will keep using the same $\mathbb{P}$ throughout. In the construction we assume that $d \ge 2$ although the ultimate conclusion will be restricted to $d \ge 4$.

Let $\{\xi_L(x)\colon L \ge 1, x \in \mathbb{Z}^d\}$ be independent 0-1-valued random variables with

$$\mathbb{P}(\xi_L(x) = 1) = L^{-d}. \tag{4.15}$$

Note that $\xi_1(x) = 1$ a.s. for all x. Consider an increasing sequence $\{L_k\}_{k\ge1}$ of integers with $L_1 = 1$. A simple use of the Borel-Cantelli lemma shows

$$\sum_{k\ge1} L_k^{-d} < \infty \quad\Rightarrow\quad \sup\big\{k \ge 1\colon \xi_{L_k}(x) = 1\big\} < \infty \text{ a.s. } \forall x \in \mathbb{Z}^d. \tag{4.16}$$

(The set on the right is non-empty a.s. as $\xi_1(x) = 1$ a.s.) Thus, assuming henceforth L_k^{-d} to be summable, let $\ell(x)$ denote the maximal k with $\xi_{L_k}(x) = 1$.

Next let $\hat{\mathrm{e}}_1, \dots, \hat{\mathrm{e}}_d$ be the unit vectors in the coordinate directions and let us regard $\mathbb{Z}^{d-1}$ as the integer span of $\{\hat{\mathrm{e}}_2, \dots, \hat{\mathrm{e}}_d\}$. Denote by

$$\Lambda_L := \big\{ j\hat{\mathrm{e}}_1 + z\colon j = -3L, \dots, 3L,\, z \in \mathbb{Z}^{d-1},\, |z| \le 1 \big\} \smallsetminus \{0\}. \tag{4.17}$$

the set consisting of $6L$ vertices in the first coordinate direction and centered at, but not containing, the origin along with all of their nearest neighbors in the other coordinate directions. Note that

$$\mathbb{P}(\ell(x) \ge j) = 1 - \prod_{k \ge j}(1 - L_k^{-d}) \le \sum_{k \ge j} L_k^{-d}. \tag{4.18}$$

By the monotonicity of $k \mapsto L_k$, we have $\sum_{j\ge1} L_j \sum_{k\ge j} L_k^{-d} = \sum_{k\ge1}\sum_{j=1}^k L_j L_k^{-d} \le \sum_{k\ge1} kL_k^{1-d}$, so another use of the Borel-Cantelli lemma gives

$$\sum_{k\ge1} kL_k^{1-d} < \infty \quad \Rightarrow \quad \sup\bigl\{k \ge 1 \colon \max_{z\in\Lambda_{L_k}} \ell(x+z) \ge k\bigr\} < \infty \text{ a.s. } \forall x \in \mathbb{Z}^d. \tag{4.19}$$

Assuming henceforth kL_k^{1-d} to be summable, let $m(x)$ denote the maximal k in this set for the given x. As $\{L_k\}$ is increasing, we get

$$m(x) < \ell(x) \quad \Rightarrow \quad m(x+z) \ge \ell(x) > \ell(x+z), \qquad z \in \Lambda_{L_{\ell(x)}}. \tag{4.20}$$

Obviously, the collection $\{(\ell(x), m(x)) \colon x \in \mathbb{Z}^d\}$ is stationary. Moreover, as Λ_L does not contain the origin, $m(x)$ is independent of $\ell(x)$ for each x. We now observe:

Lemma 4.2 *Suppose that $\sum_{k\ge1} kL_k^{1-d} < \infty$. Then there is a constant $c > 0$ such that*

$$cL_k^{-d} \le \mathbb{P}\bigl(m(x) < \ell(x) = k\bigr) \le L_k^{-d} \tag{4.21}$$

holds true for all $k \ge 2$ and all $x \in \mathbb{Z}^d$. Moreover, if also $L_{k+1} > 2L_k$ for all $k \ge 1$, then

$$\bigl\{\exists x \in \mathbb{Z}^d \colon L_k \le |x|_\infty \le 2L_k,\, m(x) < \ell(x) = k\bigr\} \tag{4.22}$$

occurs for infinitely many k, a.s.

Proof. The definition of $\ell(x)$ gives

$$\mathbb{P}\bigl(\ell(x) = k\bigr) = L_k^{-d} \prod_{j>k}(1 - L_j^{-d}). \tag{4.23}$$

This yields immediately the upper bound in (4.21). On other other hand, the fact that $\{L_k\}$ is non-decreasing shows

$$\mathbb{P}\bigl(m(x) < k\bigr) = \Bigl(\prod_{j\ge k}(1 - L_j^{-d})\Bigr)^{|\Lambda_{L_k}|} \prod_{j>k}\biggl(\Bigl(\prod_{r\ge j}(1 - L_r^{-d})\Bigr)^{|\Lambda_{L_j}\smallsetminus\Lambda_{L_{j-1}}|}\biggr). \tag{4.24}$$

By $|\Lambda_L| = O(L)$, the fact that $L_2 > 1$ and the summability of kL_k^{1-d}, both terms in the parentheses are bounded by a positive constant uniformly in $k \ge 2$. Since $m(x)$ and $\ell(x)$ are independent we get the lower bound in (4.21) as well.

For the second part, recall that we regard $\mathbb{Z}^{d-1}$ as the linear span of $\{\hat{e}_2, \dots, \hat{e}_d\}$ over the ring of integers. Given $y \in \mathbb{Z}^{d-1}$ and $j \in \mathbb{Z}$, define

$$G_k(y, j) := \bigl\{m(y + j\hat{e}_1) < \ell(y + j\hat{e}_1) = k\bigr\}. \tag{4.25}$$

We observe that, by (4.20), we have $G_k(y, j) \cap G_k(y, j') = \emptyset$ as long as $0 < |j - j'| \le 3L_k$. Hence, invoking also the lower bound in (4.21), we get

$$\mathbb{P}\Bigl(\bigcup_{j=L_k}^{2L_k} G_k(y, j)\Bigr) = \sum_{j=L_k}^{2L_k} \mathbb{P}\bigl(m(y + j\hat{e}_1) < \ell(y + j\hat{e}_1) = k\bigr) \ge cL_k^{1-d}. \tag{4.26}$$

Moreover, the giant unions are for distinct $y \in (3\mathbb{Z})^{d-1}$ independent. Hence, we get

$$\mathbb{P}\Bigg(\bigcup_{\substack{y\in(3\mathbb{Z})^{d-1} \\ L_k\le|y|_\infty\le 2L_k}} \bigcup_{j=L_k}^{2L_k} G_k(y,j)\Bigg) \ge c' > 0 \tag{4.27}$$

for some c' independent of k.

Now observe that the union in (4.27) is a subset of the event (4.22). Also note that, as soon as we have $L_{k+1} > 2L_k$, the unions in (4.27) use, for distinct k's, disjoint sets of underlying coordinates $\{\xi_L(x)\colon L \ge 1, x \in \mathbb{Z}^d\}$ and are thus independent of one another. By the second Borel-Cantelli lemma, the event in (4.22) occurs for infinitely many k a.s. □

Let us introduce the shorthand

$$\eta(x) := 1_{\{m(x)<\ell(x)\}} \tag{4.28}$$

and note that $\{\eta(x)\}$ is a stationary, ergodic process with a positive density of 1's. The following observation will turn out to be quite useful:

Lemma 4.3 *Given $x \in \mathbb{Z}^d$, let $E_L(x)$ denote the set of (nearest-neighbor) edges incident with vertices in $\{x + j\hat{\mathrm{e}}_1\colon\, j = 0,\dots,L\}$. Then*

$$x \ne \tilde{x} \quad \& \quad \eta(x) = 1 = \eta(\tilde{x}) \quad \Rightarrow \quad E_{L_{\ell(x)}}(x) \cap E_{L_{\ell(\tilde{x})}}(\tilde{x}) = \emptyset. \tag{4.29}$$

Proof. If $E_{L_{\ell(x)}}(x) \cap E_{L_{\ell(\tilde{x})}}(\tilde{x}) \ne \emptyset$ and $x \ne \tilde{x}$, then $x \in \tilde{x} + \Lambda_{L_{\ell(\tilde{x})}}$ and $\tilde{x} \in x + \Lambda_{L_{\ell(x)}}$. But then (4.20) and (4.28) yield $m(x) \ge \ell(\tilde{x}) > \ell(x)$ and, similarly, $m(\tilde{x}) \ge \ell(x) > \ell(\tilde{x})$, a contradiction. □

Proof of Theorem 2.7. Let $p, q \ge 1$ be numbers such that (2.8) holds and let $p' > p$ and $q' > q$ be such that we still have

$$\frac{1}{p'} + \frac{1}{q'} > \frac{2}{d-1}. \tag{4.30}$$

(This is where we need to require $d \ge 4$.) Define sequences

$$a_L := L^{-(d-1)/q'} \quad \text{and} \quad b_L := L^{(d-1)/p'}. \tag{4.31}$$

Consider the construction given above with $\{L_k\}$ such that $L_{k+1} > 2L_k$ and $L_1 := 1$ so that all objects $\ell(x)$, $m(x)$ and $\eta(x)$ are well defined. Given an x with $\eta(x) = 1$, denote $k := \ell(x)$ and consider the set of edges $E_{L_k}(x)$ incident with at least one vertex in $\{x + j\hat{\mathrm{e}}_1\colon\, j = 0,\dots,L_k\}$. Set the conductance to b_{L_k} on edges with both endpoints in this set and to a_{L_k} to those with only one endpoint in this set. Thanks to Lemma 4.3, the conductance of each edge is set at most once so no conflict can arise. We set the conductance on edges not in $\bigcup\{E_{L_{\ell(x)}}\colon \eta(x) = 1\}$ to one.

The resulting configuration of conductances is a measurable function of $\{\xi_L(x)\colon L \ge 1, x \in \mathbb{Z}^d\}$ and, since this family is stationary and ergodic with respect to shifts, so is the induced conductance law. Let us check that the integrability conditions (2.9) hold. Fix any x with $|x| = 1$. Noting that $E_{L_k}(z)$ contains L_k edges of conductance b_{L_k} and $R_k := 2 + (2d-2)(L_k+1)$ edges of conductance a_{L_k}, we have

$$\mathbb{E}(C_{0x}^p) \le 1 + \sum_{k\ge1} L_k^{-d}\big(L_k(b_{L_k})^p + R_k\big(a_{L_k})^p\big). \tag{4.32}$$

Plugging in (4.31), invoking that $p' > p$ and $q' > q$ and using that $\{L_k\}$ grows exponentially, we get $C_{0x} \in L^p(\mathbb{P})$ as desired. Similarly,

$$\mathbb{E}(C_{0x}^{-q}) \le 1 + \sum_{k\ge 1} L_k^{-d}\bigl(L_k(b_{L_k})^{-q} + R_k\bigl(a_{L_k}\bigr)^{-q}\bigr), \tag{4.33}$$

which is again finite by (4.31), our choices of p' and q' and the exponential growth of $\{L_k\}$.

Now let us move to the violation of sublinearity of the corrector. Suppose the event (4.22) occurs at some x with $L_k \le |x|_\infty \le 2L_k$. The conductances C_{yz} on edges $\langle y,z\rangle \in E_{L_k}(x)$ then take values a_{L_k} and b_{L_k} as specified above. Denote by $D := \{x + j\hat{\mathrm{e}}_1 \colon\, j = 0,\dots,L_k\}$ the corresponding set of vertices and let

$$\mathscr{E}_D(f) := \sum_{\langle y,z\rangle \in E_{L_k}(x)} C_{yz}\bigl|f(y)-f(z)\bigr|^2 \tag{4.34}$$

be the Dirichlet energy for a ($\mathbb{R}^d$-valued) function f on D. The "harmonic coordinate" Ψ from (2.5) solves the Dirichlet problem on D and so $f := \Psi$ has minimal $\mathscr{E}_D(f)$ among all functions that agree with Ψ on the external boundary ∂D of D.

We now derive bounds on $\mathscr{E}_D(\Psi)$. To get a lower bound, we fix the values at x and $x + L_k\hat{\mathrm{e}}_1$ and set all conductances on edges with only one endpoint in D to zero. Optimizing the remaining values is now a one-dimensional problem whose simple solution yields

$$\mathscr{E}_D(\Psi) \ge b_{L_k}L_k^{-1}\bigl|\Psi(x+L_k\hat{\mathrm{e}}_1)-\Psi(x)\bigr|^2. \tag{4.35}$$

For the upper bound, we take the test function f that equals $\Psi(x)$ everywhere on D. This gives

$$\mathscr{E}_D(\Psi) \le a_{L_k}\sum_{y\in\partial D}\bigl|\Psi(y)-\Psi(x)\bigr|^2. \tag{4.36}$$

Let us now see that this is not compatible with sublinearity of the corrector. Indeed, if (4.12) were true, then the fact that $D\cup\partial D \subset [-3L_k, 3L_k]^d$ yields

$$\bigl|\Psi(y)-\Psi(x)\bigr| \le (1+6\epsilon)L_k + 2K, \qquad y\in\partial D, \tag{4.37}$$

while

$$\bigl|\Psi(x+L_k\hat{\mathrm{e}}_1)-\Psi(x)\bigr| \ge (1-6\epsilon)L_k - 2K. \tag{4.38}$$

But that contradicts the fact, implied by (4.30), that $a_{L_k}L_k^2|\partial D| \ll b_{L_k}L_k$ once k is sufficiently large. Hence we cannot have (4.12) and, at the same time, the event in (4.22) to occur for k large. Lemma 4.2 implies that (4.12) fails for all $\epsilon > 0$ and all $K < \infty$ a.s. □

ACKNOWLEDGMENTS

This work has been partially supported by the grants NSF DMS-1407558, GAČR P201/11/1558 and the Grant-in-Aid for Scientific Research (A) 25247007. The first author would like to thank RIMS at Kyoto University for hospitality that made this project take off the ground.

REFERENCES

[1] D. Aldous. Stopping times and tightness. *Ann. Probab.* **6** (1978), no. 2, 335–340.
[2] S. Andres, M.T. Barlow, J.-D. Deuschel and B.M. Hambly. Invariance principle for the random conductance model. *Probab. Theory Rel. Fields* **156** (2013), no. 3-4, 535–580.

[3] S. Andres, J.-D. Deuschel and M. Slowik. Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.*, to appear.

[4] M. Ba and P. Mathieu. A Sobolev inequality and the individual invariance principle for diffusions in a periodic potential. arXiv:1312.4817 (2013).

[5] M.T. Barlow (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32**, no. 4, 3024–3084.

[6] M.T. Barlow, R.F. Bass, Z.-Q. Chen and M. Kassmann. Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.* **361** (2009), 1963–1999.

[7] M.T. Barlow, K. Burzdy and A. Timár. Comparison of quenched and annealed invariance principles for random conductance model. arXiv:1304.3498 (2013).

[8] M.T. Barlow and J.-D. Deuschel. Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.* **38** (2010), no. 1, 234–276.

[9] M.T. Barlow, A. Grigor'yan and T. Kumagai. Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* ***626*** (2009), 135–157.

[10] R.F. Bass. On Aronson's upper bounds for heat kernels. *Bull. London Math. Soc.* **34** (2002), 415–419.

[11] R.F. Bass and T. Kumagai. Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range. *Trans. Amer. Math. Soc.* **360** (2008), 2041–2075.

[12] N. Berger and M. Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Rel. Fields* **137** (2007), no. 1-2, 83–120.

[13] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York-London-Sydney 1968 xii+253 pp.

[14] M. Biskup. Recent progress on the Random Conductance Model. *Prob. Surveys* **8** (2011) 294–373.

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

[16] E.A. Carlen, S. Kusuoka and D.W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré - Probab. Statist.* **23** (1987), 245-287.

[17] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on d-sets. *Stoch. Proc. Their Appl.* **108** (2003), 27–62.

[18] N. Crawford and A. Sly. Simple random walk on long range percolation clusters I: Heat kernel bounds. *Probab. Theory Rel. Fields* **154** (2012), 753–786.

[19] N. Crawford and A. Sly. Simple Random Walk on Long Range Percolation Clusters II: Scaling Limits. *Ann. Probab.* **41** (2013), no. 2, 445–502.

[20] A. De Masi, P.A. Ferrari, S. Goldstein and W.D. Wick. 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, (1985).

[21] A. De Masi, P.A. Ferrari, S. Goldstein and W.D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Statist. Phys.* **55** (1989), no. 3-4, 787–855.

[22] N. Ikeda, N. Nagasawa, and S. Watanabe. A construction of Markov processes by piecing out. *Proc. Japan Acad.* **42** (1966), 370–375.

[23] C. Kipnis and S.R.S. Varadhan. A central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.* **104** (1986), no. 1, 1–19.

[24] T. Kumagai. *Random Walks on Disordered Media and Their Scaling Limits*. Lecture Notes in Mathematics / Ecole d'Ete de Probabilites de Saint-Flour vol. 2101, Springer-Verlag, 2014.

[25] T. Kumagai and J. Misumi. Heat kernel estimates for strongly recurrent random walk on random media. *J. Theoret. Probab.* **21** (2008), 910–935.

[26] T.M. Liggett. *Continuous time Markov processes. An introduction.* Graduate Studies in Mathematics, vol. 113. American Mathematical Society, Providence, RI, 2010.

[27] P. Mathieu. Quenched invariance principles for random walks with random conductances. *J. Statist. Phys.* **130** (2008), no. 5, 1025–1046.

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

[29] P.-A. Meyer. Renaissance, recollements, mé langes, ralentissement de processus de Markov. *Ann. Inst. Fourier* **25** (1975), no. 3-4, 464–497.

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

[31] E.B Procaccia, R. Rosenthal and A. Sapozhnikov. Quenched invariance principle for simple random walk on clusters in correlated percolation models. arXiv:1310.4764.

[32] V. Sidoravicius and A.-S. Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Rel. Fields*, **129** (2004), no. 2, 219–244.

[33] L. Zhang and Z. Zhang. Scaling limits for one-dimensional long-range percolation: Using the corrector method. *Stat. Probab. Lett*. **83** (2013), no. 11, 2459–2466.