



Quenched invariance principle for a class of random conductance models with long-range jumps

Marek Biskup¹ · Xin Chen² · Takashi Kumagai³ · Jian Wang⁴

Received: 28 April 2020 / Revised: 15 April 2021 / Accepted: 17 April 2021 / Published online: 13 July 2021
© The Author(s) 2021

Abstract

We study random walks on $\mathbb{Z}^d$ (with $d \geq 2$) among stationary ergodic random conductances $\{C_{x,y} : x, y \in \mathbb{Z}^d\}$ that permit jumps of arbitrary length. Our focus is on the quenched invariance principle (QIP) which we establish by a combination of corrector methods, functional inequalities and heat-kernel technology assuming that the p -th moment of $\sum_{x \in \mathbb{Z}^d} C_{0,x} |x|^2$ and q -th moment of $1/C_{0,x}$ for x neighboring the origin are finite for some $p, q \geq 1$ with $p^{-1} + q^{-1} < 2/d$. In particular, a QIP thus holds for random walks on long-range percolation graphs with connectivity exponents larger than $2d$ in all $d \geq 2$, provided all the nearest-neighbor edges are present. Although still limited by moment conditions, our method of proof is novel in that it avoids proving everywhere-sublinearity of the corrector. This is relevant because we show that, for long-range percolation with exponents between $d + 2$ and $2d$, the corrector

© 2021 M. Biskup, X. Chen, T. Kumagai and J. Wang. Reproduction, by any means, of the entire article for non-commercial purposes is permitted without charge.

✉ Marek Biskup
biskup@math.ucla.edu
<https://www.math.ucla.edu/~biskup/>
Xin Chen
chenxin217@sjtu.edu.cn
Takashi Kumagai
kumagai@kurims.kyoto-u.ac.jp
<https://www.kurims.kyoto-u.ac.jp/~kumagai/>
Jian Wang
jianwang@fjnu.edu.cn

- ¹ Department of Mathematics, UCLA, Los Angeles, CA, USA
- ² Department of Mathematics, Shanghai Jiao Tong University, Shanghai, People's Republic of China
- ³ Research Institute for Mathematical Sciences, Kyoto University, Kyoto, Japan
- ⁴ College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis Applications & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, Fuzhou, People's Republic of China

exists but fails to be sublinear everywhere. Similar examples are constructed also for nearest-neighbor, ergodic conductances in $d \geq 3$ under the conditions complementary to those of the recent work of Bella and Schäffner (Ann Probab 48(1):296–316, 2020). These examples elucidate the limitations of elliptic-regularity techniques that underlie much of the recent progress on these problems.

Mathematics Subject Classification 60F17 · 60J45

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 $P(x, y)$ determined by a collection $\{C_{x,y} : x, y \in \mathbb{Z}^d\}$ of non-negative numbers via

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

where $\pi(x) \in (0, \infty)$ is assumed for all $x \in \mathbb{Z}^d$. The symmetry condition

$$C_{x,y} = C_{y,x}, \quad x, y \in \mathbb{Z}^d, \quad (1.2)$$

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 jumps occur, i.e., those for which $C_{x,y} := 0$ 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 $C_{x,y}$ is set to one for nearest neighbors x and y and zero otherwise, or random walks with α -stable tail when $C_{x,y} := |x - y|^{-(d+\alpha)}$, where $\alpha \in (0, 2)$ and $|x|$ denotes the Euclidean norm of x . Our interest here is in the situation when $\{C_{x,y} : x, y \in \mathbb{Z}^d\}$ 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 $\mathbb{Z}^d$, and
- (2) denoting the origin of $\mathbb{Z}^d$ by “0”, we have $\mathbb{E} C_{0,0} < \infty$,

where we allow $C_{0,0} > 0$ to permit the walk to “hold” in place but require finite mean to ensure that holding does not dominate its behavior. In this framework we 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*) Invariance Principle (AIP) has been known since late 1980s (Kipnis and Varadhan [39], De Masi et al. [33,34]). The adjective “annealed” refers to the convergence taking place for a *joint* law of the chain and the environment. With Assumption 1.1 in force, this convergence was shown under the moment conditions

$$\mathbb{E}\left(\sum_{x \in \mathbb{Z}^d} C_{0,x} |x|^2\right) < \infty \quad \text{and} \quad \mathbb{E}\left(\frac{1}{C_{0,x}}\right) < \infty \quad \text{whenever } |x| = 1. \quad (1.3)$$

These are directly linked to the limiting Brownian motion having finite and positive variance and so, in this sense, can be regarded as optimal.

Much effort in the past 15 years went to derivations of *Individual* or *Quenched* Invariance Principle (QIP) where the convergence to Brownian motion takes place for a.e. sample of the random conductances. The influential initial study by Sidoravicius and Sznitman [49], where a QIP was proved for all uniformly-elliptic nearest-neighbor conductances, elucidated the need for additional ingredients compared to AIP; namely, estimates on the heat kernel. Analyses of the simple random walk on supercritical percolation clusters (Sidoravicius and Sznitman [49], Berger and Biskup [17], Mathieu and Piatnitski [45]) then paved the way to a complete resolution of all i.i.d. nearest-neighbor random conductance models (Mathieu [44], Biskup and Prescott [26], Barlow and Deuschel [12] and Andres et al. [3]).

Compared to the i.i.d. cases, our understanding of general non-uniformly elliptic conductances remains only partial and often restricted to special cases. For nearest-neighbor models satisfying Assumption 1.1, the restriction may come as a limitation on the spatial dimension. Indeed, as shown in Biskup [22, Exercise 4.4 and Theorem 4.7], a QIP holds true in $d = 1, 2$ whenever

$$|x - y| = 1 \quad \Rightarrow \quad \mathbb{E}C_{x,y} < \infty \quad \text{and} \quad \mathbb{E}\left(\frac{1}{C_{0,x}}\right) < \infty. \quad (1.4)$$

These are deemed sharp in light of (1.3) although examples violating (1.4) exist for which QIP fails yet AIP holds (Barlow et al. [11]). Another way to limit the form of the distribution is through decay of correlations. Indeed, Procaccia et al. [48] proved a QIP in correlated percolation models subject to technical conditions on correlation decay.

A third type of restriction comes via moment conditions on individual (nearest-neighbor) conductances. These can be expressed by means of numbers $p, q \geq 1$ such that

$$|x - y| = 1 \quad \Rightarrow \quad C_{x,y} \in L^p(\mathbb{P}) \quad \text{and} \quad \frac{1}{C_{x,y}} \in L^q(\mathbb{P}). \quad (1.5)$$

For these Andres, Deuschel and Slowik [5] proved a QIP under the condition $1/p + 1/q < 2/d$. This extends to a local QIP [6] and also to the control of the heat kernel [7]. Bella and Schäffner [14] recently improved the methods of [5] and gave a proof of a QIP under a slightly weaker condition

$$\frac{1}{p} + \frac{1}{q} < \frac{2}{d-1}. \quad (1.6)$$

The key additional ingredient for the improvement of the moment condition from (1.5) into (1.6) is a local boundedness result for finite difference equations in divergence forms under essentially optimal moment conditions; see [14, Theorem 2]. We will show that (1.6) is, in fact, infinitesimally close to sharp, at least for the method of proof.

The main goal of the present paper is to push the control of a QIP to include models with arbitrarily large jumps. We will work under the following moment assumption:

Assumption 1.2 Assume $d \geq 2$ and that there are $p, q \in (1, \infty)$ satisfying

$$\frac{1}{p} + \frac{1}{q} < \frac{2}{d} \quad (1.7)$$

such that

$$\sum_{x \in \mathbb{Z}^d} C_{0,x} |x|^2 \in L^p(\mathbb{P}) \quad (1.8)$$

and

$$\frac{1}{C_{0,x}} \in L^q(\mathbb{P}) \text{ whenever } |x| = 1. \quad (1.9)$$

In particular, $C_{0,x} > 0$ for all $|x| = 1$ $\mathbb{P}$ -a.s.

Assumption 1.2 is a direct extension of the conditions from the work [5], which is thus subsumed by the present paper (albeit with different proofs). We note that $C_{x,y} > 0$ for all nearest-neighbors x and y ensures that the underlying Markov chain is irreducible.

2 Main results

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

2.1 QIP for general conductances

As noted above, 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 $\mathcal{C}([0, T])$ denote the space of continuous functions on $[0, T]$ endowed with the supremum topology. Given a path $\{Z_n : n \geq 0\}$ of the chain, define

$$B^{(n)}(t) := \frac{1}{\sqrt{n}} \left(Z_{\lfloor tn \rfloor} + (tn - \lfloor tn \rfloor)(Z_{\lfloor tn \rfloor + 1} - Z_{\lfloor tn \rfloor}) \right), \quad t \geq 0. \quad (2.1)$$

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 $\mathcal{C}([0, T])$ by P^0 tends weakly, as $n \rightarrow \infty$, to that of a Brownian motion whose covariance is non-degenerate and constant a.s.

Our main result is then:

Theorem 2.2 *Suppose $d \geq 2$. Then a QIP holds under Assumptions 1.1 and 1.2.*

As noted earlier, for nearest neighbor conductances that are bounded from below, our results degenerate to those of [5]. An interesting corollary arises in the context of 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. While this probability is typically assumed to decay as a power of the distance, our formulation only requires a summability condition.

Corollary 2.3 *Let $d \geq 2$. Given a function $p : \mathbb{Z}^d \rightarrow [0, 1]$ such that*

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

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

To make the setting clear, the conductances in Corollary 2.3 take only values in $\{0, 1\}$ with $\mathbb{P}(C_{x,y} = 1) = p(x - y)$. In each step, the “simple random walk” on the graph selects an available neighbor at random. This is meaningful because the graphs in Corollary 2.3 are automatically connected (since $p(x) = 1$ for $|x| = 1$) and of finite degree at every vertex (as ensured by $\sum_{x \in \mathbb{Z}^d} p(x) < \infty$). Restricting attention to power-law decaying connection probabilities, our results show that if

$$p(x) = |x|^{-s+o(1)}, \quad |x| \rightarrow \infty, \quad (2.2)$$

for some $s > 2d$, then a QIP holds in $d \geq 2$. This is sharp in $d = 2$ but weaker than expected in $d \geq 3$ because, based on (1.3), we expect a QIP to hold for all $s > d + 2$. Note also that, since $C_{xy} \in \{0, 1\}$,

$$\mathbb{E} \left(\left(\sum_{x \in \mathbb{Z}^d} C_{0,x} |x|^2 \right)^p \right) \geq \sum_{x \in \mathbb{Z}^d} p(x) |x|^{2p} \quad (2.3)$$

whenever $p \geq 1$. Condition (3) is thus necessary for Assumption 1.2 to hold.

In the regime $s \in (d, d+2)$, the simple random walk on the long-range percolation graph is supposed to scale to a stable process with index $\alpha := s - d$. This was proved for $\alpha \in (0, 1)$ in all $d \geq 1$ by Crawford and Sly [30,31] under the L^r -space topology for any $r \in [1, \infty)$ (which is weaker than the Skorohod topology). In $d = 1$ the regime when a QIP holds extends to all $\alpha > 1$, i.e., even beyond the summability of $\sum_{x \in \mathbb{Z}^d} |x|^2 p(x)$, cf. [31, Theorem 1.2] (see also Kumagai and Misumi [41, Theorem 2.2] concerning heat kernels). This is due to absence of percolation and the existence of cut-points. A corrector-based approach exists as well (Zhang and Zhang [50]).

We note that, in the regime $s \in (d, 2d)$, the long-range percolation graph is rather different from $\mathbb{Z}^d$. Indeed, as shown by Biskup [20,21] and Biskup and Lin [24], the graph distances grow polylogarithmically with the Euclidean distance and balls in the intrinsic metric thus exhibit stretched-exponential volume growth. When $s = 2d$, the scaling of intrinsic metric relative to Euclidean one is only polynomial, but with exponents strictly less than one (Ding and Sly [36]). Notwithstanding, for $s > d + 2$, these do not seem to affect the asymptotic of the random walk, at least at the level of AIP.

2.2 Lack of everywhere sublinearity

The second set of our results address limitations of the techniques presently used for proofs of the QIP. This requires introducing the basic object of stochastic homogenization, the so-called *corrector* χ . Consider the generator $L := P - \text{id}$ associated with the discrete-time Markov kernel P . Explicitly, L acts on finitely-supported $f: \mathbb{Z}^d \rightarrow \mathbb{R}$ as

$$(Lf)(x) := \frac{1}{\pi(x)} \sum_{y \in \mathbb{Z}^d} C_{x,y} [f(y) - f(x)]. \quad (2.4)$$

Under Assumption 1.1, the conditions (1.3) permit the construction of a random function $\chi: \mathbb{Z}^d \rightarrow \mathbb{R}^d$ which is characterized by the following properties:

- (1) normalization $\chi(0) = 0$,
- (2) stationarity of increments under the shifts of $\mathbb{Z}^d$

$$\{\chi(x) - \chi(y) : x, y \in \mathbb{Z}^d\} \stackrel{\text{law}}{=} \{\chi(x - y) : x, y \in \mathbb{Z}^d\}, \quad (2.5)$$

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

$$\Psi(x) := x + \chi(x) \quad (2.6)$$

in the sense that

$$(L\Psi)(x) = 0, \quad x \in \mathbb{Z}^d. \quad (2.7)$$

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

We refer to, e.g., Biskup [22, Proposition 3.7] for a detailed exposition and proofs of this otherwise completely classical material.

Remark 2.4 In all QIPs discussed in this paper, the covariance matrix $\Sigma = (\Sigma_{ij})$ of the limiting Brownian motion is related to the corrector via

$$\Sigma_{ij} = \frac{1}{\mathbb{E}\pi(0)} \mathbb{E} \left(\sum_{x \in \mathbb{Z}^d} C_{0,x} (x_i + \hat{e}_i \cdot \chi(x)) (x_j + \hat{e}_j \cdot \chi(x)) \right), \quad (2.8)$$

where x_i denotes the i -th Cartesian component of x and $\hat{e}_i$ denotes the unit vector in the i -th coordinate direction. Note that Σ is non-degenerate and finite under (1.3); see Proposition 4.2 for an explicit statement.

It is well known (see [22, Lemma 4.8]) that (1.3) ensures that χ is sublinear along coordinate directions in the sense that $\frac{1}{n} \chi(nx) \rightarrow 0$ $\mathbb{P}$ -a.s. as $n \rightarrow \infty$ for each $x \in \mathbb{Z}^d$. This is what gives a QIP in all $d = 1$ situations. In $d \geq 2$ this implies *sublinearity on average*,

$$\forall \delta > 0: \quad \lim_{n \rightarrow \infty} \frac{1}{n^d} \sum_{|x| \leq n} \mathbb{1}_{\{|\chi(x)| > \delta n\}} = 0, \quad \mathbb{P}\text{-a.s.} \quad (2.9)$$

see [22, Proposition 4.15]. A key step underlying all of the aforementioned approaches to QIP is the proof of *sublinearity everywhere*,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \max_{x: |x| \leq n} |\chi(x)| = 0, \quad \mathbb{P}\text{-a.s.} \quad (2.10)$$

This is known to be sufficient to get a Brownian limit for diffusively-scaled paths of the Markov chain (see, e.g., Biskup [22, Section 4.2], Kumagai [40, Section 8.4] or [5]).

While our proof of Theorem 2.2 avoids everywhere sublinearity, the conditions we work under are still generally necessary for everywhere sublinearity to hold:

Theorem 2.5 *Let $d \geq 3$ and consider the long-range percolation graph obtained from $\mathbb{Z}^d$ as above with $\mathfrak{p}$ having the asymptotic (2.2) for some $s \in (d + 2, 2d)$. Then the corrector is (well defined yet) not sublinear everywhere.*

Note that this is true despite the conjecture that a QIP holds for all $s > d + 2$. A natural question is whether such examples can be constructed also for nearest-neighbor conductances. This is answered in:

Theorem 2.6 *Suppose $d \geq 3$ and let $p, q \geq 1$ be such that*

$$\frac{1}{p} + \frac{1}{q} > \frac{2}{d-1}. \quad (2.11)$$

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

$$|x| = 1 \quad \Rightarrow \quad C_{0,x} \in L^p(\mathbb{P}) \quad \text{and} \quad \frac{1}{C_{0,x}} \in L^q(\mathbb{P}) \quad (2.12)$$

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

Modulo a boundary case, condition (2.11) is complementary to (1.6) under which Bella and Schäffner [14] proved that the corrector is sublinear everywhere and thus a QIP holds. While Theorem 2.6 presents counterexamples only to the method of proof, rather than the QIP itself, it makes it unlikely that the elliptic-regularity methods underlying [5, 14] would ever yield a proof of a QIP in nearest-neighbor conductance models under the presumably optimal conditions (1.4). It appears that more promising strategies are to either try to infer a QIP directly from the AIP (which is known to hold under (1.4)) or restrict proofs of sublinearity of the corrector to just the set of vertices visited by the random walk. An attempt in the latter direction has been made by Ba and Mathieu [9], albeit for diffusions in periodic environments.

2.3 Main ideas

Although our proofs are based on a combination of the corrector method with heat-kernel technology, our strategy is somewhat different from that used in proofs of QIPs so far. Through the use of functional inequalities we first control the first exit times of the walk from large balls. These are used to prove tightness of diffusively-scaled Markov-chain paths. The proof of a QIP then boils down to the proof of a quenched Central Limit Theorem (CLT). For this we use the corrector method but, since this is “just” a CLT, with everywhere sublinearity replaced by sublinearity on average.

As usual, we work primarily with continuous-time versions of our random walk. A key innovation is the use of

$$\nu(x) := \sum_{y \in \mathbb{Z}^d} C_{x,y} |x - y|^2 \quad (2.13)$$

as the time-change measure for the walk. The need for this particular normalization was discovered in the derivation of off-diagonal heat-kernel estimates using the so called Davies method; cf the proof of Proposition 3.8. Another instance where this measure naturally appears are estimates on Dirichlet forms of spatially-mollified functions; see the proof of Proposition 3.3. Notwithstanding, the use of the time change by ν is purely technical. In particular, a QIP will hold for all time-change measures that have finite expectation under the invariant law from the point of the particle.

Unlike the recent work [5, 6] on QIPs under moment conditions, in order to control the heat kernel and exit probabilities we do not use complicated inductive schemes such as Moser or De Giorgi iterations. Instead, we base our argument on *localization*, which amounts to restricting jumps larger than unity to only a finite “active” ball, and *truncation*, by which we discard jumps larger than a constant (called κ below) multiple

of the “active” ball radius. The localization helps us control “small” jumps using Sobolev inequalities and standard techniques from heat kernel theory. The contribution of “large” jumps is managed with the help of so called Meyer’s construction (see [13,47]) which, to put simply, is a way to control the transition probabilities of a Markov process with jumps by explicating, just as in the integral form of the Kolmogorov backward equation, the first “large” jump of the process.

While the idea to combine localization argument with Meyer’s construction is drawn from a recent paper by some of the authors [28] (see Section 2.2 therein for more details), their extension to the present context requires non-trivial adaptations and generalizations. A key point for us is that localization enables a class of scale dependent Sobolev inequalities and Davies’s method that yield estimates for the heat kernel and exit times (see Propositions 3.3 and 3.8 below). With these in hand, a QIP can be proved from sublinearity on average of the associated corrector alone.

Although we do not address everywhere sublinearity of the corrector in our parameter regime, we suspect that it does hold under Assumption 1.1 and 1.2. The counterexamples in the nearest-neighbor case are strongly inspired by analogous examples of i.i.d. nearest neighbor conductances (Mathieu and Remy [46], Berger, Biskup, Hoffman and Kozma [18], Biskup and Boukhadra [23]) for which the return probabilities exhibit strongly subdiffusive decay while the path distribution still scales to a non-degenerate Brownian motion. The key mechanism there is trapping.

2.4 Open problems

We finish by stating a few open problems that naturally build on the results of the present note. As a starter, we pose:

Problem 2.7 *Under the conditions of Theorem 2.2, prove a local CLT by showing, e.g.,*

$$\lim_{n \rightarrow \infty} \sup_{\substack{x \in (2\mathbb{Z})^d \\ |x| \leq \sqrt{n}}} \left| (2n)^{d/2} P^{2n}(0, x) - k_1(x/\sqrt{n}) \right| = 0, \quad \mathbb{P}\text{-a.s.}, \quad (2.14)$$

where k_1 is the probability density of a centered normal with covariance (2.8).

We believe that this holds under the same conditions as a CLT by analogy with the nearest-neighbor situations in [5,6]. For nearest-neighbor variable speed random walks in ergodic environments, a local CLT is obtained in [6,8,16] under some moment condition whereas a QIP has been established in [3,5,11,12,14]. In particular, according to [35, Proposition 1.5], when $d \geq 4$ and $q = \infty$, the moment condition (1.6) (note that the dimension d of [35] corresponds to $d - 1$ in the present paper) is sufficient and, except for the equality, generally necessary for a local CLT to hold for variable speed random walks in ergodic environments.

Techniques to prove quenched local CLT results are definitely available. Indeed, Andres, Chiarini and Slowik [4] extended the iteration methods underlying [5,6] to non-elliptic situations under a condition slightly weaker than (1.7). Very recently, Bella and Schäffner [15,16] use the regularity theory of weak solutions of elliptic

equations to prove a local CLT for nearest-neighbor variable speed random walks even under the weaker moment condition (1.6). Still, a quenched local CLT in both papers [6, 16] is based on parabolic Harnack inequalities. The problem is that, for models with arbitrarily large jumps, elliptic Harnack inequalities may not hold in general even for large balls and parabolic Harnack inequalities may thus not hold either. See [29, Section 4.2.2] for related discussions on random conductance models with stable-like jumps.

Another extension that we believe should be possible by a reasonably straightforward adaptation of the methods of the present work is the content of:

Problem 2.8 *Let $d \geq 2$ and suppose $\mathbf{p}: \mathbb{Z}^d \rightarrow [0, 1]$ obeys (2.2) with $s > 2d$ and $\mathbf{p}(x) := p$ when $|x| = 1$ for some $p \in [0, 1)$. Assume that the random graph with vertex set $\mathbb{Z}^d$ and an edge between x and y present with probability $\mathbf{p}(y - x)$ independently of other edges, contains an infinite connected component $\mathcal{C}^\infty$ a.s. Prove that the simple random walk on $\mathcal{C}^\infty$ obeys a QIP.*

Here the key challenge is the potential absence (as even $p = 0$ is allowed) of nearest-neighbor edges in the computations involving Dirichlet forms in our proofs. Barlow [10] and the recent work of Flegel et al. [37] provide good possible starting points.

In light of Theorems 2.5 and 2.6, a different strategy than used so far is needed to get a QIP beyond the regime marked by (1.6) or (1.8) and, in particular, for long-range percolation graphs with decay exponents $s \in (d + 2, 2d]$. Here we propose to start with:

Problem 2.9 *Consider the long-range percolation graph with exponents $s \in (d + 2, 2d]$ and $\mathbf{p}(x) = 1$ for $|x| = 1$. Prove a QIP.*

The requirement $\mathbf{p}(x) = 1$ ensures that the underlying graph is connected. A key obstacle is thus the lack of the Sobolev inequalities underlying our proofs. Although the corrector fails to be everywhere sublinear in these cases, this is not an obstacle for our approach, for which sublinearity on average is sufficient. We find it worthwhile to start by addressing the non-percolating regime, i.e., the situations when, upon removal of the nearest-neighbor edges, the graph does not contain an infinite connected component a.s.

A considerably more robust way to go beyond the p, q -conditions would be to prove corrector sublinearity along typical paths of the Markov chain

$$\max_{1 \leq k \leq n} \frac{1}{\sqrt{n}} |\chi(Z_k)| \xrightarrow[n \rightarrow \infty]{P^0} 0, \quad \mathbb{P}\text{-a.s.} \quad (2.15)$$

This is, in fact, what underlies the known proofs of the AIP under the optimal conditions (1.3) and even the present paper goes part of the way along this line. Ba and Mathieu [9] have been able to utilize this strategy to prove a QIP for a continuum diffusion in a random environment subject to a periodicity requirement.

The approach of [9] is based on Dirichlet form theory, time change arguments and new weighted Sobolev-type inequalities for integrable potentials. The main benefit of the periodicity of the environment is that only global weighted Sobolev-type inequalities and on-diagonal heat-kernel upper bounds are required (while here we have to

work with scale-dependent Sobolev inequalities and off-diagonal heat kernel upper bounds). A question relevant from the point of view of our counterexamples is whether the corrector in [9] is everywhere sublinear or not—for if it is, then the periodicity assumption is perhaps too strong to tell us much in our context.

Our last question, which is undoubtedly the one most ambitious, concerns the random walk on one-dimensional long-range percolation graphs (i.e., the setting of Corollary 2.3, (1,2) and (2.2)). Indeed, as noted above, there we get $(s - 1)$ -stable process convergence when $s \in (1, 2)$ and a Brownian limit when $s > 2$.

Problem 2.10 *Prove (quenched or annealed) convergence for suitably scaled random walk on one-dimensional long-range percolation graphs for $s = 2$.*

We conjecture that all the α -stable limits with $\alpha \in (1, 2)$ somehow appear for $s = 2$. If so, we would expect that the index of stability depends on the precise asymptotic of the connection probabilities; i.e., on $\beta := \lim_{|x| \rightarrow \infty} |x|^2 p(x)$. Since the $1/r^2$ -percolation model is known to exhibit multiple phase transitions (cf Aizenman and Newman [1], Imbrie and Newman [38]), the dependence of α on β may even undergo interesting phase transitions as well.

3 Functional inequalities and heat-kernel estimates

We will now move to the exposition of the proofs. In this section, we develop the main technical ingredients underlying the proof of QIP in Theorem 2.2. We start by introducing continuous-time versions of our discrete-time Markov chains.

3.1 Continuous time processes

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

$$(\mathbf{L}_X f)(x) := \sum_{y \in \mathbb{Z}^d} C_{x,y} [f(y) - f(x)]. \quad (3.1)$$

The counting measure $\mu(x) := 1$ on $\mathbb{Z}^d$ is stationary and reversible for X . Hence, the Dirichlet form $(D, \mathcal{F})$ associated with the process X is given by

$$\begin{aligned} D(f, f) &:= \sum_{x,y \in \mathbb{Z}^d} C_{x,y} [f(y) - f(x)]^2, \quad f \in \mathcal{F}, \\ \mathcal{F} &:= \{f \in \ell^2(\mu) : D(f, f) < \infty\}. \end{aligned} \quad (3.2)$$

Here, for any $p \in [1, \infty)$ and any measure λ on $\mathbb{Z}^d$, let $\ell^p(\lambda)$ denote the space of p -integrable functions $f: \mathbb{Z}^d \rightarrow \mathbb{R}$ and denote by $\|f\|_{\ell^p(\lambda)}$ the corresponding ℓ^p -norm.

Our second, and more important, continuous-time chain $Y := \{Y_t : t \geq 0\}$ will be a time change of the process X defined as follows: $Y_t := X_{A_t^{-1}}$ where

$$A_t^{-1} := \inf\{s \geq 0 : A_s > t\} \quad \text{for} \quad A_t := \int_0^t v(X_s) \, ds, \quad (3.3)$$

with $v(x)$ as in (2.13). Then Y is a continuous-time Markov chain on $\mathbb{Z}^d$ with the generator

$$(\mathcal{L}_Y f)(x) := \frac{1}{v(x)} \sum_y C_{x,y} [f(y) - f(x)] \quad (3.4)$$

and Y is thus reversible with respect to v . (Alternatively, Y can be defined directly from Z and independent exponentials that at x have parameter $\pi(x)/v(x)$.) In particular, the Dirichlet form $(\tilde{D}, \tilde{\mathcal{F}})$ associated with the process Y is given by

$$\begin{aligned} \tilde{D}(f, f) &:= \sum_{x,y \in \mathbb{Z}^d} v(x) \frac{C_{x,y}}{v(x)} [f(y) - f(x)]^2 = D(f, f), \quad f \in \tilde{\mathcal{F}}, \\ \tilde{\mathcal{F}} &:= \{f \in \ell^2(v) : \tilde{D}(f, f) < \infty\}. \end{aligned} \quad (3.5)$$

We will henceforth think of the chains Z , X and Y as defined on the same probability space, and write P^x for the joint law of their paths where (each) chain is at x at time zero a.s. We will use E^x to denote expectation with respect to P^x .

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 $\mathbb{Q}_X(d\omega) := \mathbb{P}(d\omega)$,

$$\mathbb{Q}_Y(d\omega) := \frac{v(0)}{\mathbb{E}v(0)} \mathbb{P}(d\omega) \quad \text{and} \quad \mathbb{Q}_Z(d\omega) := \frac{\pi(0)}{\mathbb{E}\pi(0)} \mathbb{P}(d\omega), \quad (3.6)$$

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 $\mathbb{P}$. Moreover, according to Assumption 1.1 and the definitions (1.1) and (2.13), both of the measures π and v satisfy that $\pi(x) = \pi(0) \circ \tau_x$ and $v(x) = v(0) \circ \tau_x$ for all $x \in \mathbb{Z}^d$, where $\{\tau_x\}_{x \in \mathbb{Z}^d}$ are the shifts of $\mathbb{Z}^d$. This structure ensures absence of finite-time blow-ups:

Lemma 3.1 *Suppose Assumptions 1.1 and 1.2 hold. Then both X and Y 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 [43, 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 $N_t := \sup\{n \geq 0 : T_1 + \dots + T_n \leq t\}$ where, conditional on Z , the random times $\{T_k : k \geq 1\}$ are independent exponentials with T_k having parameter $\pi(Z_{k-1})$. Thanks to the 1st and 2nd Borel–Cantelli lemmas,

$$\sum_{k \geq 1} T_k = \infty \quad \text{a.s.} \quad \Leftrightarrow \quad \sum_{k \geq 1} E^x(T_k | Z) = \infty \quad \text{a.s.} \quad (3.7)$$

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

$$\frac{1}{n} \sum_{k=0}^{n-1} 1/\pi(Z_k) \xrightarrow{n \rightarrow \infty} \mathbb{E}_{\mathbb{Q}_Z}(1/\pi(0)) = 1/\mathbb{E}\pi(0) \quad \text{a.s.} \quad (3.8)$$

The limit is positive since $\mathbb{E}\pi(0) < \infty$ by Assumption 1.2. In particular, we have $\sum_{k \geq 0} 1/\pi(Z_k) = \infty$ a.s.

The argument for Y process is analogous; only that T_{k+1} is now (conditionally on Z) exponential with parameter $\pi(Z_k)/\nu(Z_k)$. Here we also need $0 < \mathbb{E}\nu(0) < \infty$ as implied by Assumption 1.2 as well as $\nu(x) = \nu(0) \circ \tau_x$ for all $x \in \mathbb{Z}^d$ as noted above. $\square$

3.2 Localization and truncation

Our proof focuses on the process Y . The main challenge is to control the contribution of large jumps. As noted earlier, we do this by way of localization, which is a change of the environment that limits all the complexity to a finite ball, and truncation, where we remove jumps larger than a certain cutoff from the environment. We remark that the idea of considering localized modifications of non-local Dirichlet forms has appeared in [28, Section 2.2], but here the construction is more delicate as we need to modify both the conductances and the reference measure.

We start by localization. Denote

$$B(x, R) := x + ([-R, R]^d \cap \mathbb{Z}^d). \quad (3.9)$$

For any integer $R \geq 1$, let

$$C_{x,y}^R := \begin{cases} C_{x,y}, & \text{if } x \in B(0, 2R) \text{ or } y \in B(0, 2R), \\ 1, & \text{if } x \notin B(0, 2R) \text{ and } y \notin B(0, 2R) \text{ and } |x - y| = 1, \\ 0, & \text{otherwise} \end{cases} \quad (3.10)$$

and

$$\nu^R(x) := \begin{cases} \nu(x), & \text{if } x \in B(0, 2R), \\ 1 + \nu(x), & \text{if } x \in B(0, 4R) \setminus B(0, 2R), \\ 1, & \text{if } x \notin B(0, 4R) \end{cases} \quad (3.11)$$

and define a symmetric regular Dirichlet form $(\tilde{D}^R, \tilde{\mathcal{F}}^R)$ by

$$\begin{aligned}\tilde{D}^R(f, f) &:= \sum_{x, y \in \mathbb{Z}^d} C_{x, y}^R [f(y) - f(x)]^2, \quad f \in \tilde{\mathcal{F}}^R, \\ \tilde{\mathcal{F}}^R &:= \{f \in \ell^2(\nu^R) : \tilde{D}^R(f, f) < \infty\}.\end{aligned}\quad (3.12)$$

This form corresponds to the localized version of our process.

Next we move to truncation. Here we first show:

Lemma 3.2 *For all $\kappa \in (0, 1]$ and all $R \geq 1$,*

$$\sup_{x \in \mathbb{Z}^d} \frac{1}{\nu^R(x)} \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x, y}^R |x - y|^2 \leq 1 + 2d \quad (3.13)$$

and

$$\sup_{x \in B(0, 4R)} \frac{1}{\nu^R(x)} \sum_{\substack{y \in \mathbb{Z}^d \\ |y-x| > \kappa R}} C_{x, y}^R \leq \frac{1 + 2d}{\kappa^2 R^2}. \quad (3.14)$$

Proof By (1.9), (3.10) and (3.11), for all $R \geq 1$ and all $x \in B(0, 2R)$,

$$\frac{1}{\nu^R(x)} \sum_{y \in \mathbb{Z}^d} C_{x, y}^R |x - y|^2 = \frac{1}{\nu(x)} \sum_{y \in \mathbb{Z}^d} C_{x, y} |x - y|^2 = 1, \quad (3.15)$$

while for $x \in B(0, 4R) \setminus B(0, 2R)$ we get

$$\begin{aligned}& \frac{1}{\nu^R(x)} \sum_{y \in \mathbb{Z}^d} C_{x, y}^R |x - y|^2 \\ &= \frac{1}{1 + \nu(x)} \left(\sum_{y \in B(0, 2R)} C_{x, y} |x - y|^2 + \sum_{\substack{y \in B(0, 2R)^c \\ |x-y|=1}} 1 \right) \leq 1 + 2d.\end{aligned}\quad (3.16)$$

Hence,

$$\sup_{R \geq 1} \sup_{x \in B(0, 4R)} \frac{1}{\nu^R(x)} \sum_{y \in \mathbb{Z}^d} C_{x, y}^R |x - y|^2 \leq 1 + 2d. \quad (3.17)$$

In particular, for all $\kappa \in (0, 1]$ and all $R \geq 1$,

$$\begin{aligned}& \sup_{x \in B(0, 4R)} \frac{1}{\nu^R(x)} \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x, y}^R |x - y|^2 \\ &\leq \sup_{x \in B(0, 4R)} \frac{1}{\nu^R(x)} \sum_{y \in \mathbb{Z}^d} C_{x, y}^R |x - y|^2 \leq 1 + 2d.\end{aligned}\quad (3.18)$$

On the other hand, (3.10) and (3.11) also give us that for all $R \geq 1$ and all $\kappa \in (0, 1]$,

$$\sup_{x \in B(0, 4R)^c} \frac{1}{v^R(x)} \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x,y}^R |x-y|^2 \leq \sup_{x \in B(0, 4R)^c} \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} 1 = 2d. \quad (3.19)$$

Combining (3.18–3.19), we have (3.13). Noting that, in light of (3.17), the sum in (3.14) is bounded by

$$\frac{1}{\kappa^2 R^2} \sup_{x \in B(0, 4R)} \frac{1}{v^R(x)} \sum_{y \in \mathbb{Z}^d} C_{x,y}^R |x-y|^2 \leq \frac{1+2d}{\kappa^2 R^2}, \quad (3.20)$$

the claim follows. $\square$

For all $\kappa \in (0, 1]$ and all $R \geq 1$ satisfying $\kappa R \geq 1$ we now define a truncated, localized Dirichlet form $(\tilde{D}^{R,\kappa}, \tilde{\mathcal{F}}^R)$ by

$$\tilde{D}^{R,\kappa}(f, f) := \sum_{\substack{x, y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x,y}^R [f(y) - f(x)]^2, \quad f \in \tilde{\mathcal{F}}^R, \quad (3.21)$$

which is well defined by (3.13). A starting point of our derivations is the following Sobolev inequality for $(\tilde{D}^{R,\kappa}, \tilde{\mathcal{F}}^R)$:

Proposition 3.3 *Let $d \geq 2$ and suppose Assumptions 1.1 and 1.2 hold. There are $\varepsilon \in (0, \frac{4}{d-2})$, a constant $c_1 \in (0, \infty)$ and an a.s.-finite random variable $R_0 := R_0(\omega) \geq 1$ such that*

$$\|f\|_{\ell^{2+\varepsilon}(v^R)}^2 \leq c_1 \left(R^{2-\frac{d\varepsilon}{2+\varepsilon}} \tilde{D}^{R,\kappa}(f, f) + R^{-\frac{d\varepsilon}{2+\varepsilon}} \|f\|_{\ell^2(v^R)}^2 \right) \quad (3.22)$$

holds for all $\kappa \in (0, 1]$, all $f \in \ell^2(v^R)$ and all $R \geq R_0$ with $\kappa R \geq 1$.

The proof is based on two lemmas. Consider the Dirichlet form

$$D_1(f, f) := \sum_{\substack{x, y \in \mathbb{Z}^d \\ |x-y|=1}} \bar{C}_{x,y} [f(x) - f(y)]^2 \quad (3.23)$$

associated with an auxiliary collection $\{\bar{C}_{x,y} = \bar{C}_{y,x} : |x-y| = 1\}$ of nearest-neighbor conductances. We then have:

Lemma 3.4 *For all $d \geq 2$, there is $c(d) \in (0, \infty)$ and, for all $p, q \in (\frac{d}{2}, \infty)$ satisfying (1.7) there is $\varepsilon \in (0, \frac{4}{d-2})$ with*

$$\frac{1}{q} + \frac{2}{2+\varepsilon} \frac{1}{p} = \frac{2}{d} - \frac{\varepsilon}{2+\varepsilon} \quad (3.24)$$

such that for all $L \geq 1$, all $f: \mathbb{Z}^d \rightarrow [0, \infty)$ with $\text{supp}(f) \subseteq B(0, L)$, all $\bar{v}: \mathbb{Z}^d \rightarrow [0, \infty)$ and all positive $\{\bar{C}_{x,y} = \bar{C}_{y,x}: |x - y| = 1\}$,

$$\begin{aligned} & \left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \bar{v}(x) \right)^{\frac{2}{2+\varepsilon}} \\ & \leq c(d) \left(\frac{(2+\varepsilon)p}{p-1} \right)^2 \alpha_L^{\frac{2}{p(2+\varepsilon)}} \beta_L^{\frac{1}{q}} L^{2-\frac{d\varepsilon}{2+\varepsilon}} D_1(f, f) \end{aligned} \quad (3.25)$$

holds with

$$\alpha_L := \frac{1}{L^d} \sum_{x \in B(0,L)} \bar{v}(x)^p \quad \text{and} \quad \beta_L := \frac{1}{L^d} \sum_{\substack{x \in B(0,L) \\ y: |y-x|=1}} (\bar{C}_{x,y})^{-q}. \quad (3.26)$$

Proof Let $p, q \in (\frac{d}{2}, \infty)$ obey (1.7). Then (3.24) is solved for ε by

$$\varepsilon := 2 \left(\frac{2}{d} - \frac{1}{p} - \frac{1}{q} \right) \left(\frac{d-2}{d} + \frac{1}{q} \right)^{-1} \in (0, \frac{4}{d-2}). \quad (3.27)$$

Let $s > 1$ be the index Hölder conjugate to p . Then by our restriction on the support of f ,

$$\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \bar{v}(x) \leq \left(\sum_{x \in \mathbb{Z}^d} f(x)^{s(2+\varepsilon)} \right)^{1/s} \alpha_L^{1/p} L^{d/p}. \quad (3.28)$$

Next define r by

$$r \frac{d}{d-1} = s(2+\varepsilon) \quad (3.29)$$

and note that, since $s > 1$ and $\varepsilon > 0$, we have $r > 1$. Using that $|a^\gamma - b^\gamma| \leq \gamma |a - b|(a^{\gamma-1} + b^{\gamma-1})$ holds for all $a, b > 0$ and all $\gamma \geq 1$, the ℓ^1 -Sobolev inequality implies

$$\begin{aligned} \left(\sum_{x \in \mathbb{Z}^d} f(x)^r \right)^{\frac{d-1}{d}} &\leq c \sum_{|x-y|=1} |f(x)^r - f(y)^r| \\ &\leq 2c r \sum_{|x-y|=1} f(x)^{r-1} |f(x) - f(y)|, \end{aligned} \quad (3.30)$$

where c is a d -dependent constant (which is directly related to the isoperimetric constant on $\mathbb{Z}^d$).

Define $\theta \in (0, 1/2)$ by $\theta := \frac{1}{2}(1 - \frac{1}{q})$ and note that, by (3.24),

$$r - 1 = \theta r \frac{d}{d-1}. \quad (3.31)$$

The Hölder inequality and the restriction on the support of f then give

$$\begin{aligned} \sum_{|x-y|=1} f(x)^{r-1} |f(x) - f(y)| &= \sum_{|x-y|=1} f(x)^{r \frac{d}{d-1} \theta} (\bar{C}_{x,y}^{-\frac{1}{1-2\theta}})^{\frac{1}{2}-\theta} \left(\bar{C}_{x,y} |f(x) - f(y)|^2 \right)^{\frac{1}{2}} \\ &\leq \left(2d \sum_{x \in \mathbb{Z}^d} f(x)^{r \frac{d}{d-1}} \right)^{\theta} \left(\sum_{\substack{x \in B(0,L) \\ y: |y-x|=1}} \bar{C}_{x,y}^{-\frac{1}{1-2\theta}} \right)^{\frac{1}{2}-\theta} D_1(f, f)^{\frac{1}{2}}. \end{aligned} \quad (3.32)$$

Since (3.29) and (3.31) show

$$\frac{d-1}{d} - \theta = \frac{1}{s(2+\varepsilon)} \quad (3.33)$$

using $\theta \leq 1/2$ we can combine (3.32) with (3.30) to get

$$\left(\sum_{x \in \mathbb{Z}^d} f(x)^{s(2+\varepsilon)} \right)^{\frac{1}{s(2+\varepsilon)}} \leq 2c(2d)^{\frac{1}{2}} r L^{\frac{d}{2q}} \beta_L^{\frac{1}{2q}} D_1(f, f)^{\frac{1}{2}}. \quad (3.34)$$

Plugging this in (3.28), we obtain

$$\left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \bar{v}(x) \right)^{\frac{2}{2+\varepsilon}} \leq 8dc^2 r^2 L^{\frac{d}{q} + \frac{2d}{p(2+\varepsilon)}} \alpha_L^{\frac{2}{p(2+\varepsilon)}} \beta_L^{1/q} D_1(f, f). \quad (3.35)$$

The claim now follows from (3.24) and (for the r^2 term) (3.29). $\square$

Remark 3.5 As noted by a referee, by (3.27),

$$2 + \varepsilon = 2 \left(1 - \frac{1}{p} \right) \frac{d}{d - 2 + d/q}.$$

The inequality (3.25) in Lemma 3.4 (with an imprecise constant) can be directly deduced from a weighted version of Sobolev inequality given in [5, (26) in Remark 3.6].

For the next lemma, let

$$D_0(f, f) := \sum_{\substack{x, y \in \mathbb{Z}^d \\ |x-y|=1}} [f(x) - f(y)]^2 \quad (3.36)$$

denote the Dirichlet form associated with the simple random walk. Recall that μ is the counting measure on $\mathbb{Z}^d$. Then we have:

Lemma 3.6 *For each $d \geq 2$ there is a constant $c(d) \in (0, \infty)$ such that for all $\varepsilon \in (0, \frac{4}{d-2})$ and all $f \in \ell^2(\mu)$,*

$$\begin{aligned} & \left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \right)^{\frac{2}{2+\varepsilon}} \\ & \leq (c(d)(2+\varepsilon))^{\frac{\varepsilon d}{2+\varepsilon}} D_0(f, f)^{\frac{\varepsilon d}{2(2+\varepsilon)}} \left(\sum_{x \in \mathbb{Z}^d} f(x)^2 \right)^{1 - \frac{\varepsilon d}{2(2+\varepsilon)}}. \end{aligned} \quad (3.37)$$

Proof Let $\varepsilon \in (0, \frac{4}{d-2})$ and set

$$r := \max \left\{ 2, (2+\varepsilon) \frac{d-1}{d} \right\}. \quad (3.38)$$

Then $r \frac{d}{d-1} > 2$ and so there exist unique $\beta, \gamma \in \mathbb{R}$ such that

$$2 + \varepsilon = \beta r \frac{d}{d-1} + (1 - \beta)2, \quad (3.39)$$

$$r - 1 = \gamma r \frac{d}{d-1} + \left(\frac{1}{2} - \gamma \right) 2. \quad (3.40)$$

A calculation shows

$$\beta = \varepsilon \left(\frac{rd}{d-1} - 2 \right)^{-1} \quad \text{and} \quad \gamma = (r-2) \left(\frac{rd}{d-1} - 2 \right)^{-1}. \quad (3.41)$$

We claim that $\beta \in (0, 1]$ and $\gamma \in [0, 1/2)$. Indeed, $\beta > 0$ and $\gamma \geq 0$ are immediate from (3.41) and $r \geq 2$. The inequality $\beta \leq 1$ is equivalent $r \geq \frac{d-1}{d}(2+\varepsilon)$, which

holds for our choice of r in (3.38), while $\gamma < 1/2$ is equivalent to $r < 2\frac{d-1}{d-2}$. This requires $2 + \varepsilon < 2\frac{d}{d-2}$, which holds thanks to $\varepsilon < \frac{4}{d-2}$.

We first assume that $f: \mathbb{Z}^d \rightarrow [0, \infty)$ has compact support. Using (3.39) and Hölder's inequality we have

$$\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \leq \left(\sum_{x \in \mathbb{Z}^d} f(x)^{\frac{rd}{d-1}} \right)^\beta \left(\sum_{x \in \mathbb{Z}^d} f(x)^2 \right)^{1-\beta}. \quad (3.42)$$

Furthermore, by the ℓ^1 -Sobolev inequality on $\mathbb{Z}^d$, as in (3.30) we obtain

$$\begin{aligned} \left(\sum_{x \in \mathbb{Z}^d} f(x)^{\frac{rd}{d-1}} \right)^{\frac{d-1}{d}} &\leq cr \sum_{x \in \mathbb{Z}^d} f(x)^{r-1} |f(x) - f(y)| \\ &\leq cr D_0(f, f)^{\frac{1}{2}} \left(\sum_{x \in \mathbb{Z}^d} f(x)^{\frac{rd}{d-1}} \right)^\gamma \left(\sum_{x \in \mathbb{Z}^d} f(x)^2 \right)^{\frac{1}{2}-\gamma}, \end{aligned} \quad (3.43)$$

where $c \in (0, \infty)$ depends only on the spatial dimension d and where we relied on (3.40) to get the second inequality. Hence we get

$$\left(\sum_{x \in \mathbb{Z}^d} f^{\frac{rd}{d-1}}(x) \right)^{\frac{d-1}{d}-\gamma} \leq cr D_0(f, f)^{\frac{1}{2}} \left(\sum_{x \in \mathbb{Z}^d} f^2(x) \right)^{\frac{1}{2}-\gamma}. \quad (3.44)$$

Noting that

$$\frac{d-1}{d} - \gamma = \frac{\beta}{\varepsilon} \left[\frac{d-1}{d} \left(\frac{rd}{d-1} - 2 \right) - (r-2) \right] = \frac{2\beta}{d\varepsilon} \quad (3.45)$$

and, after a short calculation, also

$$1 - \beta + \left(\frac{1}{2} - \gamma \right) \frac{d\varepsilon}{2} = \frac{2 + \varepsilon}{2} - \frac{\varepsilon d}{4}, \quad (3.46)$$

from (3.42) and (3.44) we conclude

$$\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \leq (cr)^{\frac{\varepsilon d}{2}} D_0(f, f)^{\frac{\varepsilon d}{4}} \left(\sum_{x \in \mathbb{Z}^d} f(x)^2 \right)^{\frac{2+\varepsilon}{2} - \frac{\varepsilon d}{4}}. \quad (3.47)$$

Raising both sides to $\frac{2}{2+\varepsilon}$ and using the definition of r , the conclusion for all $f: \mathbb{Z}^d \rightarrow [0, \infty)$ with compact support follows.

Next we suppose that $f \in \ell^2(\mu)$. Choose a sequence $\{f_n\}_{n \geq 1}$ of functions with compact support such that f_n converges to f in $\ell^2(\mu)$ as $n \rightarrow \infty$. As $D_0(f, f) \leq$

$8d\|f\|_{\ell^2(\mu)}^2$ for all $f \in \ell^2(\mu)$, we have $D_0(f_n, f_n) \rightarrow D_0(f, f)$ as $n \rightarrow \infty$. Therefore, the conclusion for $f \in \ell^2(\mu)$ follows by applying f_n into (3.37) first and then letting $n \rightarrow \infty$. $\square$

We remark that an alternative proof of Lemma 3.6 can be devised based on estimates for the transition probabilities of the simple random walk. In particular, (3.37) is true even when $d = 1$ with $\varepsilon \in (0, \infty)$. On the other hand, the proof of Lemmas 3.4 and 3.6 becomes considerably easier in $d \geq 3$ where one can rely on the ℓ^2 -Sobolev inequality.

With the above lemmas in hand, we are ready to give:

Proof of Proposition 3.3 Let $d \geq 2$ and suppose Assumption 1.2 holds. Since the inequality in (1.7) is strict, we may assume that both indices are finite, i.e., $p, q \in (\frac{d}{2}, \infty)$. Under Assumption 1.1, the Spatial Ergodic Theorem yields the existence of a constant $c_0 \in (0, \infty)$ and a random variable $R_0 = R_0(\omega)$ (which may depend on p, q) with $\mathbb{P}(1 \leq R_0 < \infty) = 1$ such that, for all $R \geq R_0$,

$$\sum_{x \in B(0, 16R)} v(x)^p \leq c_0 R^d \quad \text{and} \quad \sum_{\substack{x \in B(0, 16R) \\ y: |y-x|=1}} (C_{x,y})^{-q} \leq c_0 R^d. \quad (3.48)$$

The definitions (3.10–3.11) of $C_{x,y}^R$ and v^R then give, for all $R \geq R_0$,

$$\sup_{x \in \mathbb{Z}^d} \sum_{z \in B(x, 8R)} v^R(z)^p \leq c_1 R^d \quad \text{and} \quad \sum_{\substack{x \in B(0, 8R) \\ y: |y-x|=1}} (C_{x,y}^R)^{-q} \leq c_1 R^d \quad (3.49)$$

for some $c_1 \in (0, \infty)$ that also may depend on p and q .

Let $\varepsilon \in (0, \frac{4}{d-2})$ solve (3.24) and fix $\kappa \in (0, 1]$. Lemma 3.4 with $\bar{v}(x) := v^R(x)$, $\bar{C}_{x,y} := C_{x,y}^R$, $L := 8R$ along with (3.49) shows the existence of a constant $c_2 \in (0, \infty)$ that depends only on d, p, ε and c_1 above such that

$$\left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} v^R(x) \right)^{\frac{2}{2+\varepsilon}} \leq c_2 R^{2-\frac{d\varepsilon}{2+\varepsilon}} \tilde{D}^{R,\kappa}(f, f) \quad (3.50)$$

holds for all $R \geq R_0$ with $\kappa R \geq 1$ and all $f: \mathbb{Z}^d \rightarrow [0, \infty)$ such that $\text{supp}(f) \subseteq B(0, 8R)$. Here we used that $D_1(f, f) \leq \tilde{D}^{R,\kappa}(f, f)$ due to $\kappa R \geq 1$ and the choice $\bar{C}_{x,y} := C_{x,y}^R$.

Next we invoke Lemma 3.6 along with the fact that, for some constant $c > 0$,

$$r^{2-\frac{d\varepsilon}{2+\varepsilon}} a + r^{-\frac{d\varepsilon}{2+\varepsilon}} b \geq ca \left(\frac{b}{a} \right)^{1-\frac{\varepsilon d}{2(2+\varepsilon)}} \quad (3.51)$$

is valid for all $a, b, r > 0$, to get the existence of $c_3 \in (0, \infty)$ such that, for all $f: \mathbb{Z}^d \rightarrow [0, \infty)$ with $\text{supp}(f) \subseteq B(0, 4R)^c$,

$$\begin{aligned} & \left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \nu^R(x) \right)^{\frac{2}{2+\varepsilon}} \\ & \leq c_3 R^{-\frac{d\varepsilon}{2+\varepsilon}} \left(R^2 D_0(f, f) + \sum_{x \in \mathbb{Z}^d} f(x)^2 \right) \\ & \leq c_3 \left(R^{2-\frac{d\varepsilon}{2+\varepsilon}} \tilde{D}^{R,\kappa}(f, f) + R^{-\frac{d\varepsilon}{2+\varepsilon}} \sum_{x \in \mathbb{Z}^d} f(x)^2 \nu^R(x) \right). \end{aligned} \quad (3.52)$$

Here we used $\nu^R(x) = 1$ for all $x \in B(0, 4R)^c$ and the definitions (3.10–3.11) along with $\kappa R \geq 1$ to ensure that the conductance $C_{x,y}^R$ is no smaller than that of the simple random walk whenever x or y is in $\text{supp}(f)$.

Consider a mollifier $\phi_R: \mathbb{Z}^d \rightarrow [0, 1]$ subject to

$$\phi_R(x) \begin{cases} = 1, & x \in B(0, 4R), \\ \in [0, 1], & x \in B(0, 8R) \setminus B(0, 4R), \\ = 0, & x \in B(0, 8R)^c, \end{cases} \quad (3.53)$$

and

$$|\phi_R(x) - \phi_R(y)| \leq \frac{|x - y|}{2R}, \quad x, y \in \mathbb{Z}^d. \quad (3.54)$$

Let $f: \mathbb{Z}^d \rightarrow [0, \infty)$. Since $\text{supp}(f\phi_R) \subseteq B(0, 8R)$ while $\text{supp}(f(1 - \phi_R)) \subseteq B(0, 4R)^c$, the bounds (3.50) and (3.52) show

$$\begin{aligned} & \left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \nu^R(x) \right)^{\frac{2}{2+\varepsilon}} \\ & \leq 2 \left(\sum_{x \in \mathbb{Z}^d} (f(x)\phi_R(x))^{2+\varepsilon} \nu^R(x) \right)^{\frac{2}{2+\varepsilon}} \\ & \quad + 2 \left(\sum_{x \in \mathbb{Z}^d} (f(x)(1 - \phi_R(x)))^{2+\varepsilon} \nu^R(x) \right)^{\frac{2}{2+\varepsilon}} \\ & \leq c_4 R^{2-\frac{d\varepsilon}{2+\varepsilon}} \left[\tilde{D}^{R,\kappa}(f\phi_R, f\phi_R) + \tilde{D}^{R,\kappa}(f(1 - \phi_R), f(1 - \phi_R)) \right] \\ & \quad + c_4 R^{-\frac{d\varepsilon}{2+\varepsilon}} \sum_{x \in \mathbb{Z}^d} f(x)^2 \nu^R(x), \end{aligned} \quad (3.55)$$

where $c_4 := 2 \max\{c_2, c_3\}$. For the sum of the two Dirichlet forms we then get

$$\begin{aligned}
 & \tilde{D}^{R,\kappa}(f\phi_R, f\phi_R) + \tilde{D}^{R,\kappa}(f(1-\phi_R), f(1-\phi_R)) \\
 & \leq 4\tilde{D}^{R,\kappa}(f, f) + 4 \sum_{\substack{x,y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x,y}^R [\phi_R(x) - \phi_R(y)]^2 f(x)^2 \\
 & \leq 4\tilde{D}^{R,\kappa}(f, f) + R^{-2} \sum_{\substack{x,y \in \mathbb{Z}^d \\ |x-y| \leq \kappa R}} C_{x,y}^R |x-y|^2 f(x)^2 \quad (3.56) \\
 & \leq 4\tilde{D}^{R,\kappa}(f, f) + (1+2d)R^{-2} \sum_{x \in \mathbb{Z}^d} f(x)^2 \nu^R(x),
 \end{aligned}$$

where (3.13) was used in the last inequality. Plugging this in (3.55), the claim follows. $\square$

We note that the above proof highlights the need for ν as a reference measure and its modification ν^R .

3.3 Heat-kernel estimates

We will now apply the above functional inequalities to estimates of the heat kernels. Denote by $Y^R := \{Y_t^R : t \geq 0\}$ the Hunt process associated with $(\tilde{D}^R, \tilde{\mathcal{F}}^R)$ and let $p^R(t, x, y)$ be the associated transition probabilities. Similarly, write $Y^{R,\kappa} := \{Y_t^{R,\kappa} : t \geq 0\}$ for the Hunt process associated with $(\tilde{D}^{R,\kappa}, \tilde{\mathcal{F}}^R)$ and let $p^{R,\kappa}(t, x, y)$ be the corresponding transition probabilities.

We start a simple consequence of Proposition 3.3:

Lemma 3.7 *Suppose that Assumptions 1.1 and 1.2 hold, and let $\varepsilon \in (0, \frac{4}{d-2})$ and the random variable $R_0 := R_0(\omega)$ be as in Proposition 3.3. Then there exists a constant $c > 0$ such that*

$$p^{R,\kappa}(t, x, y) \leq cR^{-d} \left(\frac{t}{R^2} \right)^{-\frac{2+\varepsilon}{\varepsilon}} e^{\frac{1}{2}tR^{-2}} \nu^R(y) \quad (3.57)$$

holds for all $\kappa \in (0, 1)$, all $R \geq R_0$ with $\kappa R \geq 1$, all $t > 0$ and all $x, y \in \mathbb{Z}^d$.

Proof Let $\varepsilon \in (0, \frac{4}{d-2})$ and R_0 be as in Proposition 3.3. For $f: \mathbb{Z}^d \rightarrow [0, \infty)$, Hölder's inequality shows

$$\begin{aligned}
 \sum_{x \in \mathbb{Z}^d} f(x)^2 \nu^R(x) &= \sum_{x \in \mathbb{Z}^d} f(x)^{\frac{2+\varepsilon}{1+\varepsilon}} f(x)^{\frac{\varepsilon}{1+\varepsilon}} \nu^R(x) \\
 &\leq \left(\sum_{x \in \mathbb{Z}^d} f(x)^{2+\varepsilon} \nu^R(x) \right)^{\frac{1}{1+\varepsilon}} \left(\sum_{x \in \mathbb{Z}^d} f(x) \nu^R(x) \right)^{\frac{\varepsilon}{1+\varepsilon}}. \quad (3.58)
 \end{aligned}$$

Pick $\kappa \in (0, 1)$ and assume $R \geq R_0$ with $\kappa R \geq 1$. Then (3.22) turns this into the Nash inequality

$$\|f\|_{\ell^2(v^R)}^{4\frac{1+\varepsilon}{2+\varepsilon}} \leq c_1 R^{2-\frac{d\varepsilon}{2+\varepsilon}} \left(\tilde{D}^{R,\kappa}(f, f) + R^{-2} \|f\|_{\ell^2(v^R)}^2 \right) \|f\|_{\ell^1(v^R)}^{\frac{2\varepsilon}{2+\varepsilon}}. \quad (3.59)$$

The general equivalence between heat-kernel bounds and the Nash inequality, cf Carlen et al. [27, Theorem (2.1)], states that the Nash inequality (for any real $n > 0$)

$$\|f\|_2^{2+4/n} \leq A \left(D(f, f) + \delta \|f\|_2^2 \right) \|f\|_1^{4/n} \quad (3.60)$$

leads to a uniform bound on the heat kernel by $(nA/t)^{n/2} e^{\frac{1}{2}\delta t}$ —which reflects the “missing” $1/2$ in our normalization of the Dirichlet form. Applying this to (3.59) with the specific parameter values $n := 2\frac{2+\varepsilon}{\varepsilon}$, $\delta := R^{-2}$ and $A := c_1 R^{2-\frac{d\varepsilon}{2+\varepsilon}}$, we get (3.57). $\square$

The inequality (3.57) is particularly useful when t and R are related by diffusive scaling and it provides a version of a uniform, a.k.a. diagonal, heat-kernel upper bound. For the off-diagonal estimate, we have to work somewhat harder:

Proposition 3.8 *Suppose Assumptions 1.1 and 1.2 hold, and let $\varepsilon \in (0, \frac{4}{d-2})$ and the random variable $R_0 := R_0(\omega)$ be as in Proposition 3.3. For every $\kappa \in (0, 1]$, there is a constant $c \in (0, \infty)$ such that for all $x, y \in \mathbb{Z}^d$, all $R \geq R_0(\omega)$ with $\kappa R \geq 1$ and all $0 < t \leq R^2$,*

$$p^{R,\kappa}(t, x, y) \leq c R^{-d} \left(\frac{t}{R^2} \right)^{-\frac{2+\varepsilon}{\varepsilon}} \exp \left(-\frac{|x-y|}{5\kappa R} \log \left(\frac{R^2}{t} \right) \right) v^R(y). \quad (3.61)$$

Proof We will invoke an argument from Carlen, Kusuoka and Stroock [27] (based on an earlier argument of Davies [32]) for obtaining off-diagonal heat-kernel bounds from the Nash inequality (3.59). For that we first introduce the auxiliary objects

$$\begin{aligned} \Gamma_R(\psi)(x) &:= \frac{1}{v^R(x)} \sum_{y \in \mathbb{Z}^d} (e^{\psi(x)-\psi(y)} - 1)^2 C_{xy}^R \mathbb{1}_{\{|x-y| \leq \kappa R\}}, \\ \Lambda(\psi)^2 &:= \|\Gamma_R(\psi)\|_\infty \vee \|\Gamma_R(-\psi)\|_\infty, \\ E_R(t, x, y) &:= \sup\{|\psi(x) - \psi(y)| - t \Lambda(\psi)^2 : \Lambda(\psi) < \infty\}. \end{aligned} \quad (3.62)$$

Here ψ can be chosen to be any bounded function in the domain of the Dirichlet form $(\tilde{D}^{R,\kappa}, \tilde{\mathcal{F}}^R)$. In particular, we can take ψ with bounded support. Carlen, Kusuoka and Stroock [27, Theorem (3.25)] then shows that there is a constant $c_0 > 0$ such that for all $\kappa \in (0, 1]$, all $R \geq R_0(\omega)$ with $\kappa R \geq 1$, all $t > 0$ and all $x, y \in \mathbb{Z}^d$,

$$p^{R,\kappa}(t, x, y) \leq c_0 R^{-d} \left(\frac{t}{R^2} \right)^{-\frac{2+\varepsilon}{\varepsilon}} e^{\frac{1}{2}tR^{-2}} e^{-E_R(2t,x,y)} v^R(y) \quad (3.63)$$

which, we note, refines the estimate from Lemma 3.7. In order to bring (3.63) into the desired form, it thus suffices to supply a good lower bound on $E_R(2t, x, y)$.

Fix $x_0, y_0 \in \mathbb{Z}^d$, let $\lambda \geq 0$ and consider the test function

$$\psi(x) := \lambda(|x_0 - y_0| - |x_0 - x|)_+. \quad (3.64)$$

The triangle inequality gives $|\psi(x) - \psi(y)| \leq \lambda|x - y|$. According to the elementary inequalities $|e^t - 1|^2 \leq t^2 e^{2|t|}$ and $t^2 e^{-|t|} \leq 2$ for $t \geq 0$, we then get from (3.13) that

$$\begin{aligned} \Gamma_R(\psi)(x) &= \frac{1}{v^R(x)} \sum_{y \in \mathbb{Z}^d} (e^{\psi(x) - \psi(y)} - 1)^2 C_{xy}^R \mathbb{1}_{\{|x - y| \leq \kappa R\}} \\ &\leq \frac{1}{v^R(x)} e^{4\lambda\kappa R} \lambda^2 \sum_{y \in \mathbb{Z}^d} |x - y|^2 C_{xy}^R \mathbb{1}_{\{|x - y| \leq \kappa R\}} \\ &\leq (1 + 2d)\lambda^2 e^{4\lambda\kappa R} \leq 2(1 + 2d)\kappa^{-2} e^{5\lambda\kappa R} R^{-2}. \end{aligned} \quad (3.65)$$

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

$$-E_R(2t, x_0, y_0) \leq -\lambda|x_0 - y_0| + 2(1 + 2d)t\kappa^{-2} e^{5\lambda\kappa R} R^{-2}. \quad (3.66)$$

Suppose that $0 < t \leq R^2$ and set

$$\lambda := \frac{1}{5\kappa R} \log\left(\frac{R^2}{t}\right). \quad (3.67)$$

Then

$$-E_R(2t, x_0, y_0) \leq 2(1 + 2d)\kappa^{-2} - \frac{|x_0 - y_0|}{5\kappa R} \log\left(\frac{R^2}{t}\right). \quad (3.68)$$

Denoting $c := e^{\frac{1}{2} + 2(1+2d)\kappa^{-2}} c_0$, the claim now follows from (3.63). $\square$

3.4 Exit time estimates

The uniform estimate on the transition probabilities of the truncated, localized process $Y^{R,\kappa}$ permits us to control the tails of the exit times thereof. This can then be extended to the process Y as well. Indeed, given $A \subseteq \mathbb{Z}^d$, define the first exit time from A by

$$\tau_A := \inf\{t > 0 : Y_t \notin A\}. \quad (3.69)$$

We then have:

Proposition 3.9 *Under Assumptions 1.1 and 1.2, there is a random variable $R_1 := R_1(\omega)$ with $\mathbb{P}(1 \leq R_1 < \infty) = 1$ and, for each $\delta \in (0, 1]$, a constant $c \in (0, \infty)$ such that for all $t > 0$, all $R \geq \delta^{-1}R_1$ and all $x \in B(0, R)$,*

$$P^x(\tau_{B(x, \delta R)} \leq t) \leq \frac{ct}{R^2}. \quad (3.70)$$

The proof is based on a comparison with the corresponding exit problems for the walks Y^R and $Y^{R, \kappa}$. For all $R \geq 1$, all $\kappa \in (0, 1]$, all $x \in \mathbb{Z}^d$ and all $r \geq 1$, let

$$\tau_{B(x, r)}^{R, \kappa} := \inf\{t \geq 0: Y_t^{R, \kappa} \notin B(x, r)\}. \quad (3.71)$$

We then have:

Lemma 3.10 *Suppose Assumptions 1.1 and 1.2 hold and let $R_0 := R_0(\omega)$ be as in Proposition 3.3. There is $\kappa \in (0, 1]$ and, for each $\delta \in (0, 1]$, also $c > 0$ such that for all $x \in \mathbb{Z}^d$, all $R \geq \max\{16\delta^{-1}R_0, (\kappa\delta)^{-1}\}$ and all $t > 0$,*

$$P^x(\tau_{B(x, \delta R)}^{R, \delta\kappa} \leq t) \leq \frac{ct}{R^2}. \quad (3.72)$$

Proof Let $\varepsilon \in (0, \frac{4}{d-2})$ and R_0 be as in Proposition 3.3 and let $\kappa \in (0, 1]$ be such that

$$\frac{1}{20\kappa} - \frac{2+\varepsilon}{\varepsilon} \geq 1. \quad (3.73)$$

Fix $\delta \in (0, 1]$. Since (3.61) applies with κ replaced by $\kappa\delta$ for all $R \geq R_0$ satisfying $\kappa\delta R \geq 1$, all $0 < t \leq R^2$ and all $x \in \mathbb{Z}^d$, we get

$$\begin{aligned} & P^x(|Y_t^{R, \delta\kappa} - x| \geq \tfrac{1}{2}\delta R) \\ & \leq c_0 \left(\frac{t}{R^2}\right)^{-\frac{2+\varepsilon}{\varepsilon}} R^{-d} \sum_{\substack{y \in \mathbb{Z}^d \\ |y-x| \geq \frac{1}{2}\delta R}} \exp\left(-\frac{|x-y|}{5\delta\kappa R} \log\left(\frac{R^2}{t}\right)\right) \nu^R(y) \\ & = c_0 \left(\frac{t}{R^2}\right)^{-\frac{2+\varepsilon}{\varepsilon}} R^{-d} \sum_{n=1}^{\infty} \sum_{\substack{y \in \mathbb{Z}^d \\ n \leq 2 \frac{|y-x|}{\delta R} < n+1}} \exp\left(-\frac{|x-y|}{5\delta\kappa R} \log\left(\frac{R^2}{t}\right)\right) \nu^R(y), \end{aligned} \quad (3.74)$$

where c_0 depends on κ and δ . Assuming in addition that $t \leq R^2/3$ (which ensures $\log(R^2/t) \geq 1$) and noting that the condition $\delta R \geq 16R_0(\omega)$ enables us to apply (3.48) and (3.49), the two sums on the right are now bounded by

$$\sum_{n=1}^{\infty} \exp\left(-\frac{n}{10\kappa} \log\left(\frac{R^2}{t}\right)\right) \left(\sum_{y \in B(x, \frac{1}{2}(n+1)\delta R)} \nu^R(y) \right)$$

$$\begin{aligned}
&\leq c_1(\delta R)^d \sum_{n=1}^{\infty} n^d \exp\left(-\frac{n}{10\kappa} \log\left(\frac{R^2}{t}\right)\right) \\
&\leq c_2 R^d \exp\left(-\frac{1}{20\kappa} \log\left(\frac{R^2}{t}\right)\right) = c_2 R^d \left(\frac{t}{R^2}\right)^{\frac{1}{20\kappa}}
\end{aligned} \quad (3.75)$$

for some c_2 depending on κ and δ . Combining (3.74–3.75), from (3.73) we obtain

$$P^x(|Y_t^{R,\delta\kappa} - x| \geq \tfrac{1}{2}\delta R) \leq c_3 \frac{t}{R^2}. \quad (3.76)$$

for all $x \in \mathbb{Z}^d$, all $R \geq R_0$ with $\kappa\delta R \geq 1$, $\delta R \geq 16R_0(\omega)$ and all t with $0 < t < R^2/3$.

The strong Markov property at the first exit time from $B(x, \delta R)$ shows

$$\begin{aligned}
P^x(\tau_{B(x,\delta R)}^{R,\delta\kappa} \leq t) &\leq P^x(|Y_{2t}^{R,\delta\kappa} - x| \geq \tfrac{1}{2}\delta R) + P^x(|Y_{2t}^{R,\delta\kappa} - x| \leq \tfrac{1}{2}\delta R, \tau_{B(x,\delta R)}^{R,\delta\kappa} \leq t) \\
&\leq P^x(|Y_{2t}^{R,\delta\kappa} - x| \geq \tfrac{1}{2}\delta R) + \sup_{z \in \mathbb{Z}^d} \sup_{0 < s \leq 2t} P^z(|Y_{2t-s}^{R,\delta\kappa} - z| \geq \tfrac{1}{2}\delta R). \quad (3.77)
\end{aligned}$$

Invoking (3.76), we get the claim for all $t < R^2/6$. Adjusting the constant c if necessary, the claim holds trivially for $t \geq R^2/6$. $\square$

Next we set

$$\tau_A^R := \inf\{t \geq 0: Y_t^R \notin A\}. \quad (3.78)$$

Then we get the following deterministic estimate:

Lemma 3.11 *There is $c > 0$ such that all $t > 0$, all $\kappa \in (0, 1]$, all $R \geq 1$ with $\kappa R \geq 1$, all $r \geq 1$ and all $x \in \mathbb{Z}^d$,*

$$|P^x(\tau_{B(x,r)}^R \leq t) - P^x(\tau_{B(x,r)}^{R,\kappa} \leq t)| \leq ct \sup_{y \in B(x,r)} \frac{1}{v^R(y)} \sum_{\substack{z \in \mathbb{Z}^d \\ |y-z| > \kappa R}} C_{y,z}^R. \quad (3.79)$$

Proof This is proved by following the argument of [28, Lemma 3.1], which is itself based on the Meyer's construction of Y^R (see [13, Section 3.1]). $\square$

We are ready to give:

Proof of Proposition 3.9 Fix $\kappa \in (0, 1]$ in $(\tilde{D}^{R,\kappa}, \tilde{\mathcal{F}}^R)$ to the constant in Lemma 3.10. Since the processes Y and Y^R “see” the same conductances in $B(0, 2R)$, for all $R \geq r \geq 1$, all $t > 0$ and all $x \in B(0, R)$ we have

$$P^x(\tau_{B(x,r)} \leq t) = P^x(\tau_{B(x,r)}^R \leq t). \quad (3.80)$$

Lemma 3.11 along with (3.14) then show

$$\left| P^x(\tau_{B(x,r)} \leq t) - P^x(\tau_{B(x,r)}^{R,\delta\kappa} \leq t) \right| \leq \frac{ct}{R^2} \quad (3.81)$$

for all $R \geq r \geq 1$ with $\delta\kappa R \geq 1$, all $t > 0$ and all $x \in \mathbb{Z}^d$, where c depends on the constant from (3.79), κ and δ . Setting $r := \delta R$, Lemma 3.10 gives the claim with $R_1 := 16R_0/\kappa$. $\square$

Let $p^{B(0,R)}(t, x, y)$ denote the (substochastic) transition probabilities of the process Y killed upon exiting the ball $B(0, R)$. The conclusions for the heat-kernel associated with the truncated, localized process $Y^{R,\kappa}$ can then be transferred to Y :

Proposition 3.12 *Under Assumptions 1.1 and 1.2, and with $\varepsilon \in (0, \frac{4}{d-2})$ and R_0 as in Proposition 3.3, there is a constant $c > 0$ such that for all $R \geq R_0$, all $x, y \in B(0, R)$ and all $0 < t \leq R^2$,*

$$p^{B(0,R)}(t, x, y) \leq c R^{-d} \left(\frac{t}{R^2} \right)^{-\frac{2+\varepsilon}{\varepsilon}} \nu(y). \quad (3.82)$$

Proof Denote by $p^{R,B(0,R)}$ the transition probabilities of the process Y^R killed upon exiting the ball $B(0, R)$. Then for all $x, y \in B(0, R)$ and all $t > 0$,

$$p^{B(0,R)}(t, x, y) = p^{R,B(0,R)}(t, x, y). \quad (3.83)$$

Since, trivially,

$$p^{R,B(0,R)}(t, x, y) \leq p^R(t, x, y) \quad (3.84)$$

it suffices to prove the desired bound for the transition probabilities $p^R(t, x, y)$ of the process Y^R . Here we note that the associated Dirichlet forms obey $\tilde{D}^{R,\kappa}(f, f) \leq \tilde{D}^R(f, f)$ and so the Nash inequality (3.59) applies for the Dirichlet form $(\tilde{D}^R, \mathcal{F}^R)$ as well. Since $\nu^R = \nu$ on $B(0, R)$, the argument from the proof of Lemma 3.7 then gives the claim. $\square$

4 Proof of quenched invariance principle

Having established the needed bounds on the transition probabilities and exit times, we proceed to the proof the quenched invariance principle.

4.1 Tightness

We start with the proof of tightness of diffusively-scaled process Y . Our aim is to apply the criterion for tightness from Aldous [2]. Unfortunately, this result is not formulated for the space $\mathcal{C}([0, T])$ but rather for the Skorohod space $\mathcal{D}([0, T])$ of

functions $f : [0, T] \rightarrow \mathbb{R}^d$ 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 [19]) that makes it a Polish space which in turn permits considerations of weak limits of probability measures.

A d -dimensional version of Aldous [2, Theorem 1] then implies that the sequence $\{Y^{(n)} : n \geq 1\}$ of processes is tight in $\mathcal{D}([0, T])$ when the following two conditions are met:

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

$$Y_{(\tau_n + \delta_n) \wedge T}^{(n)} - Y_{\tau_n \wedge T}^{(n)} \xrightarrow[n \rightarrow \infty]{} 0 \quad (4.1)$$

in probability.

We will apply this to the choice

$$Y_t^{(n)} := \frac{1}{\sqrt{n}} Y_{nt}, \quad t \geq 0, \quad (4.2)$$

to get:

Proposition 4.1 *Let $d \geq 2$. For each $T > 0$ and a.e. realization of the conductances, the laws of $\{Y^{(n)} : n \geq 1\}$ induced by P^0 on $\mathcal{D}([0, T])$ are tight.*

Proof Let $R_1 = R_1(\omega)$ be as in Proposition 3.9. We will check that the above conditions (1-2) from Aldous [2] hold on the set $\{R_1 < \infty\}$. To distinguish various processes, let us write $\tau_B(X)$ for the first exit time of the process X from set B .

For (1) we note that, by (3.70) in Proposition 3.9, when $r\sqrt{n} \geq R_1$,

$$\begin{aligned} P^0(|Y_t^{(n)}| > r) \\ \leq P^0(\tau_{B(0,r)}(Y^{(n)}) \leq t) = P^0(\tau_{B(0,r\sqrt{n})}(Y) \leq nt) \leq c_1 t/r^2 \end{aligned} \quad (4.3)$$

This implies condition (1) above on $\{R_1 < \infty\}$.

Next, pick $T > 0$ and $\eta > 0$, let τ_n be stopping times bounded by T and choose δ_n with $\delta_n \downarrow 0$. For any $r > 0$, the strong Markov property gives

$$\begin{aligned} P^0(|Y_{\tau_n + \delta_n}^{(n)} - Y_{\tau_n}^{(n)}| > \eta) \\ \leq P^0(\tau_{B(0,r)}(Y^{(n)}) \leq T) + \max_{z \in B(0,r\sqrt{n})} P^z(|Y_{\delta_n}^{(n)} - z/\sqrt{n}| > \eta). \end{aligned} \quad (4.4)$$

Using (4.3), the first quantity on the right is at most $c_1 T/r^2$ whenever $r\sqrt{n} \geq R_1$. For the second quantity Proposition 3.9 with $\delta := \eta/r$ gives

$$\max_{z \in B(0,r\sqrt{n})} P^z(|Y_{\delta_n}^{(n)} - z/\sqrt{n}| > \eta)$$

$$\leq \max_{z \in B(0, r\sqrt{n})} P^z(\tau_{B(z, \eta\sqrt{n})}(Y) < n\delta_n) \leq c_2 \delta_n / r^2 \quad (4.5)$$

for $\min\{r\sqrt{n}, \eta\sqrt{n}\} \geq R_1$. While c_2 depends on the ratio η/r , for η and r fixed the right-hand side tends to zero in light of $\delta_n \rightarrow 0$. Thus we get

$$\limsup_{n \rightarrow \infty} P^0(|Y_{\tau_n + \delta_n}^{(n)} - Y_{\tau_n}^{(n)}| > \eta) \leq c_3 T / r^2 \quad \text{on } \{R_1 < \infty\} \quad (4.6)$$

for some constant $c_3 \in (0, \infty)$ regardless of η or the choice of stopping times τ_n (as long as $\tau_n \leq T$). But the left-hand side does not depend on r and so taking $r \rightarrow \infty$, we obtain condition (2) above on $\{R_1 < \infty\}$. Aldous [2, Theorem 1] then implies tightness of the processes $\{Y^{(n)} : n \geq 1\}$ on $\mathcal{D}([0, T])$. $\square$

4.2 Proof of a QIP

Having proved tightness, our proof of a QIP is now reduced to the convergence of finite-dimensional distributions. Let $\{\tilde{B}_t : t \geq 0\}$ denote a d -dimensional Brownian motion such that, for all $v \in \mathbb{R}^d$ and $t \geq 0$,

$$\mathbb{E}(\tilde{B}_t) = 0 \quad \text{and} \quad \mathbb{E}((v \cdot \tilde{B}_t)^2) = t \frac{\mathbb{E}\pi(0)}{\mathbb{E}v(0)} v \cdot \Sigma v, \quad (4.7)$$

where Σ is the matrix with entries as in (2.8). Then we have:

Proposition 4.2 *Let $d \geq 2$ and consider the processes $\{Y^{(n)} : n \geq 1\}$ from (4.2) 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 $0 \leq t_1 < t_2 < \dots < t_k < \infty$,*

$$(Y_{t_1}^{(n)}, \dots, Y_{t_k}^{(n)}) \xrightarrow[n \rightarrow \infty]{\text{law}} (\tilde{B}_{t_1}, \dots, \tilde{B}_{t_k}). \quad (4.8)$$

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 (4.8) 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 (2.6); this is defined (and depends on) conductances in a measurable set Ω_1 with $\mathbb{P}(\Omega_1) = 1$. Given a realization of the conductances and a path $Z := \{Z_n : n \geq 1\}$ of the discrete-time Markov chain, consider the random variables $\{\Psi(Z_n)\}$. A classical argument (cf, e.g., Corollary 3.10 of Biskup [22]) 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

$$t \mapsto \frac{1}{\sqrt{n}} \Psi(Z_{[tn]}) \quad (4.9)$$

induced by P^0 on $\mathcal{D}([0, T])$ —in fact, even on $\mathcal{C}([0, T])$, provided we interpolate values linearly—tends to 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 $\pi(Z_0)/\nu(Z_0), \pi(Z_1)/\nu(Z_1), \dots$, respectively. Then $\{\tilde{Y}_t : t \geq 0\}$, defined by

$$\tilde{Y}_t := Z_{N_t} \quad \text{for } N_t := \max\{k \geq 0 : T_1 + \dots + T_k \leq t\}, \quad (4.10)$$

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

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \frac{\nu(Z_k)}{\pi(Z_k)} = \mathbb{E}_{\mathbb{Q}_Z} \left(\frac{\nu(0)}{\pi(0)} \right), \quad P^0\text{-a.s.} \quad (4.11)$$

and

$$\forall \varepsilon > 0: \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \frac{\nu(Z_k)}{\pi(Z_k)} \mathbb{1}_{\{\nu(Z_k)/\pi(Z_k) > \varepsilon n\}} = 0, \quad P^0\text{-a.s.} \quad (4.12)$$

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 (4.12)) and a renewal argument, we then have

$$\frac{N_t}{t} \xrightarrow[t \rightarrow \infty]{} \frac{\mathbb{E}\pi(0)}{\mathbb{E}\nu(0)} \quad \text{in } P^0\text{-probability} \quad (4.13)$$

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 $\tilde{Y} \stackrel{\text{law}}{=} Y$ now shows that also the law

$$t \mapsto \frac{1}{\sqrt{n}}\Psi(Y_{nt}) \quad (4.14)$$

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

Since convergence on $\mathcal{D}([0, T])$ to a process with continuous paths implies convergence of finite-dimensional distributions, to get (4.8) it now suffices to identify a measurable set $\Omega^* \subseteq \Omega_3$ of conductances with $\mathbb{P}(\Omega^*) = 1$ such that

$$\frac{1}{\sqrt{n}} |\chi(Y_{tn})| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{in } P^0\text{-probability,} \quad (4.15)$$

holds on Ω^* for each $t \geq 0$. For this we argue as follows. For any $\eta > 0$,

$$\begin{aligned} P^0(|\chi(Y_{tn})| \geq \eta\sqrt{n}) &\leq P^0(\tau_{B(0, \eta^{-1}\sqrt{n})}(Y) \leq tn) \\ &+ \sum_{|x| \leq \eta^{-1}\sqrt{n}} \mathbb{1}_{\{|\chi(x)| > \eta\sqrt{n}\}} P^0(\tau_{B(0, \eta^{-1}\sqrt{n})}(Y) > tn, Y_{tn} = x). \end{aligned} \quad (4.16)$$

Assume that η is so small that $t \leq \eta^{-2}$. By Proposition 3.9,

$$P^0(\tau_{B(0, \eta^{-1}\sqrt{n})}(Y) \leq tn) \leq c_1 \eta^2 t \quad \text{whenever} \quad \eta^{-1}\sqrt{n} \geq R_1 \quad (4.17)$$

for some c_1 independent of n , η and t , where $R_1 = R_1(\omega)$ is as in Proposition 3.9. The contribution of this term to (4.16) thus vanishes as $n \rightarrow \infty$ followed by $\eta \downarrow 0$. Let $R_0 = R_0(\omega)$ be as in Proposition 3.3. Proposition 3.12 in turn gives

$$P^0(\tau_{B(0, \eta^{-1}\sqrt{n})}(Y) > tn, Y_{tn} = x) \leq c_2 \eta^d (t\eta^2)^{-\frac{2+\varepsilon}{\varepsilon}} n^{-d/2} \nu(x) \quad (4.18)$$

whenever $\eta^{-1}\sqrt{n} \geq R_0$ and $t \leq \eta^{-2}$, where c_2 is independent of n , x , η and t . Using Hölder's inequality with p as Assumption 1.2, the sum on the right of (4.16) is thus bounded by a constant times

$$\begin{aligned} &(t\eta^2)^{-\frac{2+\varepsilon}{\varepsilon}} \left(\eta^d n^{-d/2} \sum_{|x| \leq \eta^{-1}\sqrt{n}} \nu(x)^p \right)^{1/p} \\ &\times \left(\eta^d n^{-d/2} \sum_{|x| \leq \eta^{-1}\sqrt{n}} \mathbb{1}_{\{|\chi(x)| > \eta\sqrt{n}\}} \right)^{1-1/p}. \end{aligned} \quad (4.19)$$

The p -integrability of ν ensures that the term in the first large parentheses is bounded uniformly in $n \geq 1$, $\mathbb{P}$ -a.s. Thanks to corrector sublinearity on average (2.9), the term in the second large parentheses, and thus the whole expression, tends to zero $\mathbb{P}$ -a.s. as $n \rightarrow \infty$. This proves (4.15) and thus the whole claim. $\square$

Let us now see how the above proposition implies our main result:

Proof of Theorem 2.2 Fix $T > 0$. Proposition 4.1 tells us that the laws of $Y^{(n)}$ are tight on $\mathcal{D}([0, T])$. By Proposition 4.2 we then conclude that $Y^{(n)}$ converges in law to $\tilde{B}$ while the time-change argument in (4.10) and (4.13) then shows that $t \mapsto \frac{1}{\sqrt{n}} Z_{[tn]}$, as an element of $\mathcal{D}([0, T])$, 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 $\mathcal{C}([0, T])$. $\square$

4.3 Assumption 1.2 for long-range percolation

To complete our results concerning quenched invariance principles, it remains to verify the conditions on long-range percolation model that ensure convergence of the random walk to Brownian motion.

Proof of Corollary 2.3 Let $p > \frac{d}{2}$ be as in the statement. Fix any total order $x \preceq y$ on $\mathbb{Z}^d$ and let $\{C_{x,y} : x, y \in \mathbb{Z}^d, x \preceq y\}$ be independent, zero-one valued random variables with $\mathbb{P}(C_{x,y} = 1) = \mathfrak{p}(x - y)$, where $\mathfrak{p}$ is as in the statement. Identify $C_{x,y} = C_{y,x}$ to get symmetric conductances. Given $n_0 \geq 1$ to be determined later, let

$$M_n := \sum_{n_0 \leq |x| \leq n_0+n} |x|^2 (C_{0,x} - \mathbb{E}C_{0,x}) = \sum_{n_0 \leq |x| \leq n_0+n} |x|^2 W(x), \quad (4.20)$$

where

$$W(x) := C_{0,x} - \mathbb{E}C_{0,x}. \quad (4.21)$$

Then $\{M_n : n \geq 1\}$ is a martingale with respect to the filtration $\mathcal{F}_n := \sigma(C_{0,x} : |x| \leq n_0 + n)$ with the variational process

$$\langle M \rangle_n = \sum_{n_0 \leq |x| \leq n_0+n} |x|^4 W(x)^2. \quad (4.22)$$

The Burkholder-Gundy-Davis inequality thus shows, for any $p \geq 1$,

$$\mathbb{E}(|M_n|^p) \leq c_1 \mathbb{E}(\langle M \rangle_n^{p/2}) = c_1 \mathbb{E} \left[\left(\sum_{n_0 \leq |x| \leq n_0+n} |x|^4 W(x)^2 \right)^{p/2} \right]. \quad (4.23)$$

Furthermore, according to [42, Theorem 1], for every $n \geq 1$, the expectation on the right is at most

$$c_2 \inf \left\{ t > 0 : \sum_{n_0 \leq |z| \leq n_0+n} \log \left(\mathbb{E} \left[\left(1 + t^{-1} |z|^4 W(z)^2 \right)^{p/2} \right] \right) \leq p/2 \right\}. \quad (4.24)$$

We now have to estimate the infimum.

Since $-\mathfrak{p}(x) \leq W(x) \leq \mathbb{1}_{\{C_{0,x}=1\}}$, we have $\mathbb{E}(|W(x)|^r) \leq \mathfrak{p}(x)$ for all $r \geq 1$. By $(1+x)^r \leq (1+ax\mathbb{1}_{\{r \geq 1\}} + bx^r)$ valid with r -dependent $a, b > 0$ for all $x > 0$, we

then get that for $p \geq 1$,

$$\begin{aligned} & \log \left(\mathbb{E} \left[\left(1 + t^{-1} |z|^4 W(z)^2 \right)^{p/2} \right] \right) \\ & \leq \log \left(1 + c_3 t^{-1} |z|^4 \mathbf{p}(z) \mathbb{1}_{\{p/2 \geq 1\}} + c_3 t^{-p/2} |z|^{2p} \mathbf{p}(z) \right) \\ & \leq c_3 t^{-1} |z|^4 \mathbf{p}(z) \mathbb{1}_{\{p/2 \geq 1\}} + c_3 t^{-p/2} |z|^{2p} \mathbf{p}(z) \\ & \leq c_3 (t^{-1} + t^{-p/2}) |z|^{2p} \mathbf{p}(z), \end{aligned} \quad (4.25)$$

where the constant c_3 in the first inequality depends on p , in the second inequality we also used the fact that $\ln(1+x) \leq x$ for all $x > 0$, and the last inequality is due to $|x|^4 \mathbb{1}_{\{p/2 \geq 1\}} \leq |x|^{2p}$ for $|x| \geq 1$. By our assumption, the sum on the right of (4.24) is bounded by a constant times

$$(t^{-1} + t^{-p/2}) \sum_{|x| \geq n_0} |z|^{2p} \mathbf{p}(z) \quad (4.26)$$

uniformly in $n \geq 1$. For a given $t > 0$, say $t := 1$, this can be made smaller than $p/2$ by choosing n_0 sufficiently large. The infimum (4.24) is then bounded by one and so $\sup_{n \geq 1} \mathbb{E}[|M_n|^p] \leq c_1 c_2$. With the help of the Monotone Convergence Theorem we then get $v(0) \in L^p(\mathbb{P})$. The QIP then follows from Theorem 2.2. $\square$

As noted earlier, Corollary 2.3 readily deals with the cases when $\{C_{x,y} = C_{y,x}\}_{x,y \in \mathbb{Z}^d}$ are independent, zero-one valued random variables with (assuming $x \neq y$)

$$\mathbb{P}(C_{x,y} = 1) = \frac{1}{|x - y|^{d+s}}. \quad (4.27)$$

A QIP is then inferred for all $s > d$. Another example is motivated by long range stable-like random conductance models studied in [28]. There one takes (assuming again that $x \neq y$)

$$C_{x,y} := \frac{\xi_{x,y}}{|x - y|^{d+s}} \quad (4.28)$$

where $\{\xi_{x,y} = \xi_{y,x}\}_{x,y \in \mathbb{Z}^d}$ are i.i.d. Bernoulli random variables except for $|x - y| = 1$ where we set $\xi_{x,y} := 1$. In this case the conditions of Corollary 2.3 are met for all $s > 2$.

5 Failures of everywhere sublinearity

In this section we provide the promised counterexamples to everywhere sublinearity of the corrector and thus prove Theorems 2.5 and 2.6. We begin with the counterexample arising in the context of long-range percolation.

5.1 Long-range percolation

Consider long-range percolation with the connection probability $p(x)$ having the asymptotic (2.2) with exponent $s \in (d + 2, 2d)$, which is non-vacuous only when $d \geq 3$. We will assume $p(0) = 0$, $p(x) = 1$ for x with $|x| = 1$ and $p(x) < 1$ for all x with $|x| > 1$. The conductances then obey

- (1) $C_{x,x} = 0$ for all x a.s.,
- (2) $C_{x,y} = 1$ whenever $|x - y| = 1$ a.s.,
- (3) $\mathbb{P}(C_{x,y} = 1) = p(y - x)$ 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 5.1 *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*

$$A(x, y) := \{C_{x,y} = 1\} \cap \{\forall z \in \mathbb{Z}^d \setminus \{x\}: |y - z| > 1 \Rightarrow C_{yz} = 0\}. \quad (5.1)$$

Then

$$A_n := \bigcup_{\substack{x \in \mathbb{Z}^d \\ |x| \leq n^\gamma}} \bigcup_{\substack{y \in \mathbb{Z}^d \\ n < |y| \leq 2n}} A(x, y) \quad (5.2)$$

occurs for infinitely many n a.s.

Proof Instead of (5.1) consider the event

$$\tilde{A}(x, y) := \{C_{x,y} = 1\} \cap \{\forall z \in \mathbb{Z}^d: |y - z| > 1 \ \& \ |z| > n^\gamma \Rightarrow C_{yz} = 0\} \quad (5.3)$$

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 (5.2). Set

$$\tilde{A}_n := \bigcup_{\substack{x \in \mathbb{Z}^d \\ |x| \leq n^\gamma}} \bigcup_{\substack{y \in \mathbb{Z}^d \\ n < |y| \leq 2n}} \tilde{A}(x, y). \quad (5.4)$$

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

$$B_n := \left\{ \begin{array}{l} |x|, |x'| \leq n^\gamma, \ n \leq |y|, |y'| \leq 2n \\ \exists x, x', y, y' \in \mathbb{Z}^d: (x, y) \neq (x', y'), \ C_{x,y} = 1 = C_{x',y'} \\ y \neq y' \Rightarrow C_{y,y'} = 1 \end{array} \right\}, \quad (5.5)$$

we thus have

$$A_n^c \subseteq (\tilde{A}_n^c \cap B_n^c) \cup B_n. \quad (5.6)$$

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 $\mathcal{E}_n$ denote the set of all quadruples (x, x', y, y') that satisfy the geometrical conditions in event B_n . Then, for some constants $c, c' < \infty$,

$$\begin{aligned} \mathbb{P}(B_n) &\leq \sum_{x, x', y, y' \in \mathcal{E}_n} \mathbf{p}(y-x) \mathbf{p}(y'-x') (\delta_{y, y'} + \mathbf{p}(y-y')) \\ &\leq cn^{2d\gamma-2s+o(1)} \sum_{\substack{y, y' \in \mathbb{Z}^d \\ n \leq |y|, |y'| \leq 2n}} (\delta_{y, y'} + \mathbf{p}(y-y')) \leq c'n^{2d\gamma-2s+d+o(1)}, \end{aligned} \quad (5.7)$$

where we first used that both $\mathbf{p}(y-x)$ and $\mathbf{p}(y'-x')$ 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 $z \mapsto \mathbf{p}(z)$ is summable because $s > d$. Noting that, in light of our choice of γ , the final exponent in (5.7) is negative, we get that B_{2n} occurs only for finitely many n , a.s.

Concerning the first event in (5.6), let N denote the number of edges between some x with $|x| \leq n^\gamma$ and some y with $n \leq |y| \leq 2n$ and let $\{(x_i, y_i) : i = 1, \dots, N\}$ list the corresponding pairs of vertices connected by these edges. On $\tilde{A}_n^c \cap B_n^c$ 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 $\mathcal{F}_n := \sigma(C_{x,y} : |x| \leq n^\gamma, n \leq |y| \leq 2n)$, we thus have

$$\mathbb{P}(\tilde{A}_n^c \cap B_n^c | \mathcal{F}_n) \leq \mathbb{1}_{\{N=0\}} + \mathbb{1}_{\{N>0\}} \prod_{j=1}^N \left(1 - \prod_{\substack{y \neq y_1, \dots, y_N \\ |y-y_j| > 1}} (1 - \mathbf{p}(y-y_j)) \right), \quad (5.8)$$

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

$$c := \prod_{|z|>1} (1 - \mathbf{p}(z)) \quad (5.9)$$

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

$$\mathbb{P}(\tilde{A}_n^c \cap B_n^c) \leq \mathbb{P}(N \leq n^\delta) + (1-c)n^{n^\delta} \quad (5.10)$$

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

$$\tilde{q}_n := \min_{|x| \leq n^\gamma} \min_{n \leq |y| \leq 2n} \mathbf{p}(y-x). \quad (5.11)$$

and let V_n be the number of pairs (x, y) with $|x| \leq n^\gamma$ and $n \leq |y| \leq 2n$. 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 $\mathbb{P}(N \leq n^\delta)$ decays, for δ positive but small, exponentially in a power of n . Using this in (5.10), the Borel–Cantelli lemma implies that $\tilde{A}_n^c \cap B_n^c$ 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 2.5 Consider the long-range percolation setting as specified above. The asymptotic (2.2) with $s > d + 2$ implies $\mathbb{E}(\sum_{x \in \mathbb{Z}^d} C_{0,x} |x|^2) < \infty$ and so the corrector can be defined by any of the standard methods (see, e.g., Biskup [22, Section 3] for a discussion of these). In fact, by (2.9) above, the corrector is sublinear on average (cf. [22, Proposition 4.15]), meaning that $\{x: |\chi(x)| > \varepsilon|x|\}$ 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 (2.10) we will assume, for the sake of contradiction, that for each $\varepsilon > 0$ there is a (random) $K < \infty$ such that

$$|\chi(x)| \leq K + \varepsilon|x|, \quad x \in \mathbb{Z}^d. \quad (5.12)$$

(This is equivalent to (2.10).) 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 (2.7) for Ψ from (2.6) at point y then reads

$$x + \chi(x) - (y + \chi(y)) + \sum_{z: |z|=1} (z + \chi(y+z) - \chi(y)) = 0, \quad (5.13)$$

where we noted that $C_{x'y} = 1$ for $x' = x$ and x' being a neighbor of y ; otherwise $C_{x'y} = 0$. Applying (5.12) and the fact that $|x|, |y|, |y+z| \leq 2n+1$ for all z with $|z| = 1$ yield

$$|y-x| \leq (2+4d)K + 2d + \varepsilon(2+4d)(2n+1). \quad (5.14)$$

For ε small this contradicts $|y-x| > n - n^\gamma$. Hence, by Lemma 5.1, (5.12) cannot occur on A_n for n large enough and, since A_n does occur for infinitely many n a.s., (5.12) fails a.s. $\square$

5.2 Nearest-neighbor conductances

Next we move to the context underlying Theorem 2.6. 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 $d \geq 2$ although the ultimate conclusion will be restricted to $d \geq 3$.

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

$$\mathbb{P}(\xi_L(x) = 1) = L^{-d}. \quad (5.15)$$

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

$$\sum_{k \geq 1} L_k^{-d} < \infty \Rightarrow \sup\{k \geq 1 : \xi_{L_k}(x) = 1\} < \infty \text{ a.s. } \forall x \in \mathbb{Z}^d. \quad (5.16)$$

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

$$\Lambda_L := \{j\hat{e}_1 + z : j = -3L, \dots, 3L, z \in \mathbb{Z}^{d-1}, |z| \leq 1\} \setminus \{0\}. \quad (5.17)$$

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

$$\mathbb{P}(\ell(x) \geq j) = 1 - \prod_{k \geq j} (1 - L_k^{-d}) \leq \sum_{k \geq j} L_k^{-d}. \quad (5.18)$$

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

$$\begin{aligned} \sum_{k \geq 1} k L_k^{1-d} &< \infty \\ \Rightarrow \sup\{k \geq 1 : \max_{z \in \Lambda_{L_k}} \ell(x + z) \geq k\} &< \infty \text{ a.s. } \forall x \in \mathbb{Z}^d. \end{aligned} \quad (5.19)$$

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

$$m(x) < \ell(x) \Rightarrow m(x + z) \geq \ell(x) > \ell(x + z), \quad z \in \Lambda_{L_{\ell(x)}}. \quad (5.20)$$

Obviously, the collection $\{(\ell(x), m(x)) : x \in \mathbb{Z}^d\}$ 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 5.2 *Suppose that $\sum_{k \geq 1} k L_k^{1-d} < \infty$. Then there is a constant $c > 0$ such that*

$$c L_k^{-d} \leq \mathbb{P}(m(x) < \ell(x) = k) \leq L_k^{-d} \quad (5.21)$$

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

$$\{\exists x \in \mathbb{Z}^d : L_k \leq |x|_\infty \leq 2L_k, m(x) < \ell(x) = k\} \quad (5.22)$$

occurs for infinitely many k , a.s.

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

$$\mathbb{P}(\ell(x) = k) = L_k^{-d} \prod_{j>k} (1 - L_j^{-d}). \quad (5.23)$$

This yields immediately the upper bound in (5.21). On the other hand, the fact that $\{L_k\}$ is non-decreasing shows

$$\begin{aligned} \mathbb{P}(m(x) < k) &= \left(\prod_{j \geq k} (1 - L_j^{-d}) \right)^{|\Lambda_{L_k}|} \\ &\times \prod_{j>k} \left(\left(\prod_{r \geq j} (1 - L_r^{-d}) \right)^{|\Lambda_{L_j} \setminus \Lambda_{L_{j-1}}|} \right). \end{aligned} \quad (5.24)$$

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 from below by a positive constant uniformly in $k \geq 2$. Since $m(x)$ and $\ell(x)$ are independent we get the lower bound in (5.21) 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

$$G_k(y, j) := \{m(y + j\hat{e}_1) < \ell(y + j\hat{e}_1) = k\}. \quad (5.25)$$

We observe that, by (5.20), we have $G_k(y, j) \cap G_k(y, j') = \emptyset$ as long as $0 < |j - j'| \leq 3L_k$. Hence, invoking also the lower bound in (5.21), we get

$$\mathbb{P}\left(\bigcup_{j=L_k}^{2L_k} G_k(y, j)\right) = \sum_{j=L_k}^{2L_k} \mathbb{P}(m(y + j\hat{e}_1) < \ell(y + j\hat{e}_1) = k) \geq cL_k^{1-d}. \quad (5.26)$$

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

$$\mathbb{P}\left(\bigcup_{\substack{y \in (3\mathbb{Z})^{d-1} \\ L_k \leq |y|_\infty \leq 2L_k}} \bigcup_{j=L_k}^{2L_k} G_k(y, j)\right) \geq c' > 0 \quad (5.27)$$

for some c' independent of k .

Now observe that the union in (5.27) is a subset of the event (5.22). Also note that, as soon as we have $L_{k+1} > 2L_k$, the unions in (5.27) use, for distinct k 's, disjoint sets

of underlying coordinates $\{\xi_L(x) : L \geq 1, x \in \mathbb{Z}^d\}$ and are thus independent of one another. By the second Borel–Cantelli lemma, the event in (5.22) occurs for infinitely many k a.s. $\square$

Let us introduce the shorthand

$$\kappa(x) := \mathbb{1}_{\{m(x) < \ell(x)\}} \quad (5.28)$$

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

Lemma 5.3 *Given $x \in \mathbb{Z}^d$, let $E_L(x)$ denote the set of (nearest-neighbor) edges incident with vertices in $\{x + j\hat{e}_1 : j = 0, \dots, L\}$. Then*

$$x \neq \tilde{x} \ \& \ \kappa(x) = 1 = \kappa(\tilde{x}) \ \Rightarrow \ E_{L_{\ell(x)}}(x) \cap E_{L_{\ell(\tilde{x})}}(\tilde{x}) = \emptyset. \quad (5.29)$$

Proof If $E_{L_{\ell(x)}}(x) \cap E_{L_{\ell(\tilde{x})}}(\tilde{x}) \neq \emptyset$ and $x \neq \tilde{x}$, then $x \in \tilde{x} + \Lambda_{L_{\ell(\tilde{x})}}$ and $\tilde{x} \in x + \Lambda_{L_{\ell(x)}}$. But then (5.20) and (5.28) yield $m(x) \geq \ell(\tilde{x}) > \ell(x)$ and, similarly, $m(\tilde{x}) \geq \ell(x) > \ell(\tilde{x})$, a contradiction. $\square$

Proof of Theorem 2.6 Let $p, q \geq 1$ be numbers such that (2.11) holds and let $p' > p$ and $q' > q$ be such that we still have

$$\frac{1}{p'} + \frac{1}{q'} > \frac{2}{d-1}. \quad (5.30)$$

(This is where we need to require $d \geq 3$.) Define sequences

$$a_L := L^{-(d-1)/q'} \quad \text{and} \quad b_L := L^{(d-1)/p'}. \quad (5.31)$$

Consider the construction given above with $\{L_k : k \geq 1\}$ such that $L_{k+1} > 2L_k$ and $L_1 := 1$ so that all objects $\ell(x)$, $m(x)$ and $\kappa(x)$ are well defined. Given an x with $\kappa(x) = 1$, denote $k := \ell(x)$ and consider the set of edges $E_{L_k}(x)$ incident with at least one vertex in $\{x + j\hat{e}_1 : j = 0, \dots, L_k\}$. 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 5.3, the conductance of each edge is set at most once so no conflict can arise. We set the conductance on edges not in $\bigcup\{E_{L_{\ell(x)}} : \kappa(x) = 1\}$ to one.

The resulting configuration of conductances is a measurable function of random variables $\{\xi_L(x) : L \geq 1, x \in \mathbb{Z}^d\}$ 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 (2.12) 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

$$\mathbb{E}(C_{0,x}^p) \leq 1 + \sum_{k \geq 1} L_k^{-d} (L_k (b_{L_k})^p + R_k (a_{L_k})^p). \quad (5.32)$$

Plugging in (5.31), invoking that $p' > p$ and $q' > q$ and using that $\{L_k\}$ grows exponentially, we get $C_{0,x} \in L^p(\mathbb{P})$ as desired. Similarly,

$$\mathbb{E}(C_{0,x}^{-q}) \leq 1 + \sum_{k \geq 1} L_k^{-d} (L_k(b_{L_k})^{-q} + R_k(a_{L_k})^{-q}), \quad (5.33)$$

which is again finite by (5.31), our choices of p' and q' and the exponential growth of the sequence $\{L_k\}$.

Now let us move to the violation of sublinearity of the corrector. Suppose the event (5.22) occurs at some x with $L_k \leq |x|_\infty \leq 2L_k$. 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 $D := \{x + j\hat{e}_1 : j = 0, \dots, L_k\}$ the corresponding set of vertices (which depends on x) and let

$$\mathcal{E}_D(f) := \sum_{\langle y, z \rangle \in E_{L_k}(x)} C_{yz} |f(y) - f(z)|^2 \quad (5.34)$$

be the Dirichlet energy for a $(\mathbb{R}^d)$ -valued function f on D . The “harmonic coordinate” Ψ from (2.6) solves the Dirichlet problem on D and so $f := \Psi$ has minimal $\mathcal{E}_D(f)$ among all functions that agree with Ψ on the external boundary ∂D of D .

We now derive bounds on $\mathcal{E}_D(\Psi)$. To get a lower bound, we fix the values at x and $x + L_k\hat{e}_1$ 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

$$\mathcal{E}_D(\Psi) \geq b_{L_k} L_k^{-1} |\Psi(x + L_k\hat{e}_1) - \Psi(x)|^2. \quad (5.35)$$

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

$$\mathcal{E}_D(\Psi) \leq a_{L_k} \sum_{y \in \partial D} |\Psi(y) - \Psi(x)|^2. \quad (5.36)$$

Let us now see that this is not compatible with sublinearity of the corrector. Indeed, if (5.12) were true, then the fact that $D \cup \partial D \subset [-3L_k, 3L_k]^d$ yields

$$|\Psi(y) - \Psi(x)| \leq (1 + 6\varepsilon)L_k + 2K, \quad y \in \partial D, \quad (5.37)$$

while

$$|\Psi(x + L_k\hat{e}_1) - \Psi(x)| \geq (1 - 6\varepsilon)L_k - 2K. \quad (5.38)$$

But that contradicts the fact, implied by (5.30), that $a_{L_k} L_k^2 |\partial D| \ll b_{L_k} L_k$ once k is sufficiently large. Hence we cannot have (5.12) and, at the same time, the event in

(5.22) to occur for k large. Lemma 5.2 implies that (5.12) fails for all $\varepsilon > 0$ and all $K < \infty$ a.s. $\square$

Acknowledgements We thank anonymous referees for their helpful comments and corrections. This research has been supported by National Science Foundation (US) Awards DMS-1712632 and DMS-1954343, the National Natural Science Foundation of China (No. 11871338), JSPS KAKENHI Grant Number JP17H01093, the Alexander von Humboldt Foundation, the National Natural Science Foundation of China (Nos. 11831014 and 12071076), the Program for Probability and Statistics: Theory and Application (No. IRTL1704) and the Program for Innovative Research Team in Science and Technology in Fujian Province University (IRTSTFJ). The non-Kyoto based authors would like to thank RIMS at Kyoto University for hospitality that made this project possible. A version of this paper by two of the present authors was previously posted on arXiv [25] and was subsequently withdrawn due to an error in one of the key arguments. The present paper subsumes the parts of [25] that are worth saving.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Aizenman, M., Newman, C.: Discontinuity of the percolation density in one-dimensional $1/|x - y|^2$ percolation models. *Commun. Math. Phys.* **107**, 611–647 (1986)
2. Aldous, D.: Stopping times and tightness. *Ann. Probab.* **6**(2), 335–340 (1978)
3. Andres, S., Barlow, M.T., Deuschel, J.-D., Hambly, B.M.: Invariance principle for the random conductance model. *Probab. Theory Rel. Fields* **156**(3–4), 535–580 (2013)
4. Andres, S., Chiarini, A., Slowik, M.: Quenched local limit theorem for random walks among time-dependent ergodic degenerate weights. *Probab. Theory Rel. Fields* (**to appear**)
5. Andres, S., Deuschel, J.-D., Slowik, M.: Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.* **43**(4), 1866–1891 (2015)
6. Andres, S., Deuschel, J.-D., Slowik, M.: Harnack inequalities on weighted graphs and some applications to the random conductance model. *Probab. Theory Rel. Fields* **164**(3–4), 931–977 (2016)
7. Andres, S., Deuschel, J.-D., Slowik, M.: Heat kernel estimates and intrinsic metric for random walks with general speed measure under degenerate conductances. *Electron. Commun. Probab.* **24** (2019), paper no. 5
8. Andres, S., Taylor, P.A.: Local limit theorems for the random conductance model and applications to the Ginzburg–Landau $\nabla\phi$ interface model. *J. Stat. Phys.* **182** (2021), paper no. 35
9. Ba, M., Mathieu, P.: A Sobolev inequality and the individual invariance principle for diffusions in a periodic potential. *SIAM J. Math. Anal.* **47**(3), 2022–2043 (2015)
10. Barlow, M.T.: Random walks on supercritical percolation clusters. *Ann. Probab.* **32**(4), 3024–3084 (2004)
11. Barlow, M.T., Burzdy, K., Timár, A.: Comparison of quenched and annealed invariance principles for random conductance model. *Probab. Theory Rel. Fields* **164**, 741–770 (2016)
12. Barlow, M.T., Deuschel, J.-D.: Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.* **38**(1), 234–276 (2010)
13. Barlow, M.T., Grigor’yan, A., Kumagai, T.: Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626**, 135–157 (2009)
14. Bella, P., Schäffner, M.: Quenched invariance principle for random walks among random degenerate conductances. *Ann. Probab.* **48**(1), 296–316 (2020)
15. Bella, P., Schäffner, M.: Local boundedness and Harnack inequality for solutions of linear nonuniformly elliptic equations. *Commun. Pure Appl. Math.* **74**(3), 453–477 (2021)
16. Bella, P., Schäffner, M.: Non-uniformly parabolic equations and applications to the random conductance model. [arXiv:2009.11535](https://arxiv.org/abs/2009.11535)
17. Berger, N., Biskup, M.: Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Rel. Fields* **137**(1–2), 83–120 (2007)

18. Berger, N., Biskup, M., Hoffman, C.E., Kozma, G.: Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré* **27**(2), 374–392 (2008)
19. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
20. Biskup, M.: On the scaling of the chemical distance in long range percolation models. *Ann. Probab.* **32**(4), 2938–2977 (2004)
21. Biskup, M.: Graph diameter in long-range percolation. *Rand. Struct. Algorithm* **39**(2), 210–227 (2011)
22. Biskup, M.: Recent progress on the random conductance model. *Prob. Surv.* **8**, 294–373 (2011)
23. Biskup, M., Boukhadra, O.: Subdiffusive heat-kernel decay in four-dimensional i.i.d. random conductance models. *J. Lond. Math. Soc.* **86**(2), 455–481 (2012)
24. Biskup, M., Lin, J.: Sharp asymptotic for the chemical distance in long-range percolation. *Rand. Struct. Algorithm* **55**, 560–583 (2019)
25. Biskup, M., Kumagai, T.: Quenched invariance principle for a class of random conductance models with long-range jumps. [arXiv:1412.0175](https://arxiv.org/abs/1412.0175)
26. Biskup, M., Prescott, T.M.: Functional CLT for random walk among bounded conductances. *Electron. J. Probab.* **12**, 1323–1348 (2007). paper no. 49
27. Carlen, E.A., Kusuoka, S., Stroock, D.W.: Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré Probab. Statist.* **23**, 245–287 (1987)
28. Chen, X., Kumagai, T., Wang, J.: Random conductance models with stable-like jumps: quenched invariance principle. *Ann. Appl. Probab.* **(to appear)**
29. Chen, X., Kumagai, T., Wang, J.: Random conductance models with stable-like jumps: Heat kernel estimates and Harnack inequalities. *J. Funct. Anal.* **279**, paper no. 108656 (2020)
30. Crawford, N., Sly, A.: Simple random walk on long range percolation clusters I: heat kernel bounds. *Probab. Theory Rel. Fields* **154**, 753–786 (2012)
31. Crawford, N., Sly, A.: Simple random walk on long range percolation clusters II: scaling limits. *Ann. Probab.* **41**(2), 445–502 (2013)
32. Davies, E.B.: Explicit constants for Gaussian upper bounds on heat kernels. *Am. J. Math.* **109**(2), 319–333 (1987)
33. De Masi, A., Ferrari, P.A., Goldstein, S., Wick, W.D.: Invariance principle for reversible Markov processes with application to diffusion in the percolation regime. In: *Particle Systems, Random Media and Large Deviations* (Brunswick, Maine), *Contemporary Mathematics*, vol. 41, pp. 71–85. American Mathematical Society, Providence, RI (1985)
34. De Masi, A., Ferrari, P.A., Goldstein, S., Wick, W.D.: An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55**(3–4), 787–855 (1989)
35. Deuschel, J.-D., Fukushima, R.: Quenched tail estimate for the random walk in random scenery and in random layered conductance II. *Electron. J. Probab.* **25**, paper no. 75 (2020)
36. Ding, J., Sly, A.: Distances in critical long range percolation. [arXiv:1303.3995](https://arxiv.org/abs/1303.3995)
37. Flegel, F., Heida, M., Slowik, M.: Homogenization theory for the random conductance model with degenerate ergodic weights and unbounded-range jumps. *Ann. Inst. Henri. Poincaré Probab. Stat.* **55**, 1226–1257 (2019)
38. Imbrie, J., Newman, C.: An intermediate phase with slow decay of correlations in one-dimensional $1/|x - y|^2$ percolation, Ising and Potts models. *Commun. Math. Phys.* **118**(2), 303–336 (1988)
39. Kipnis, C., Varadhan, S.R.S.: A central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.* **104**(1), 1–19 (1986)
40. Kumagai, T.: *Random Walks on Disordered Media and Their Scaling Limits*. *Lecture Notes in Mathematics/Ecole d’Ete de Probabilites de Saint-Flour*, vol. 2101. Springer, Berlin (2014)
41. Kumagai, T., Misumi, J.: Heat kernel estimates for strongly recurrent random walk on random media. *J. Theor. Probab.* **21**, 910–935 (2008)
42. Latala, R.: Estimation of moments of sums of independent real random variables. *Ann. Probab.* **25**, 1502–1513 (1997)
43. Liggett, T.M.: *Continuous Time Markov Processes. An Introduction*. *Graduate Studies in Mathematics*, vol. 113. American Mathematical Society, Providence, RI (2010)
44. Mathieu, P.: Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.* **130**(5), 1025–1046 (2008)
45. Mathieu, P., Piatnitski, A.L.: Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463**, 2287–2307 (2007)

46. Mathieu, P., Remy, E.: Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.* **32**(1A), 100–128 (2004)
47. Meyer, P.-A.: Renaissance, recollements, mélanges, ralentissement de processus de Markov. *Ann. Inst. Fourier* **25**, 464–497 (1975)
48. Procaccia, E.B., Rosenthal, R., Sapozhnikov, A.: Quenched invariance principle for simple random walk on clusters in correlated percolation models. *Probab. Theory Rel. Fields* **166**, 619–657 (2016)
49. Sidoravicius, V., Sznitman, A.-S.: Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Rel. Fields* **129**(2), 219–244 (2004)
50. Zhang, L., Zhang, Z.: Scaling limits for one-dimensional long-range percolation: using the corrector method. *Stat. Probab. Lett.* **83**(11), 2459–2466 (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.