The Annals of Probability
2015, Vol. 43, No. 5, 2332–2373
DOI: 10.1214/14-AOP934

DISORDER, ENTROPY AND HARMONIC FUNCTIONS

BY ITAI BENJAMINI, HUGO DUMINIL-COPIN[1], GADY KOZMA[2] AND ARIEL YADIN

Weizmann Institute of Science, Université de Genève, Weizmann Institute of Science and Ben Gurion University of the Negev

We study harmonic functions on random environments with particular emphasis on the case of the infinite cluster of supercritical percolation on $\mathbb{Z}^d$. We prove that the vector space of harmonic functions growing at most linearly is $(d+1)$-dimensional almost surely. Further, there are no nonconstant sublinear harmonic functions (thus implying the uniqueness of the corrector). A main ingredient of the proof is a quantitative, annealed version of the Avez entropy argument. This also provides bounds on the derivative of the heat kernel, simplifying and generalizing existing results. The argument applies to many different environments; even reversibility is not necessary.

1. Introduction. Since the work of Yau in 1975, where the Liouville property for positive harmonic functions on complete manifolds with nonnegative Ricci curvature was proved [79], the structure of various spaces of harmonic functions has been at the heart of geometric analysis. Some years later, Yau conjectured that the space of polynomial growth harmonic functions of fixed order is always finite dimensional in open manifolds with nonnegative Ricci curvature. Extensive literature has appeared on this conjecture and related problems. Understanding progressed quickly (Yau's conjecture was proved by Colding and Minicozzi [26]) and gave birth to many tools; see [61] for an introduction to the subject.

In the algebraic setting, bounded harmonic functions played a central role since the introduction of the Poisson boundary by Furstenberg [41, 42]; see also the survey [78]. Recently, the geometric approach made a remarkable appearance in the algebraic realm when Kleiner proved that the space of harmonic functions with fixed polynomial growth on the Cayley graph of a group with polynomial volume growth is finite dimensional using the approach of [26]. He used this fact to provide a new proof of Gromov's theorem [55]; see [73] for a quantitative version of this theorem.

Received May 2013; revised March 2014.

[1]Supported by the EU Marie-Curie RTN CODY, the ERC AG CONFRA, as well as by the Swiss FNS and the Weizmann institute.

[2]Supported by the Israel Science Foundation and the Jesselson Foundation.

MSC2010 subject classifications. Primary 60K37; secondary 31A05, 82B43, 37A35, 60B15, 60J10, 20P05.

Key words and phrases. Harmonic functions, percolation, random walk in random environment, stationary graphs, entropy, Avez, Kaimanovich–Vershik, corrector, IIC, UIPQ, planar map, anomalous diffusion.

Another place where harmonic functions have played an important role recently is in the proof of the central limit theorem on random graphs. A central element in the proofs (see, e.g., [16, 43, 66, 74]) is the construction of a harmonic function h on the cluster which is close to linear—the term $\chi(x) = h(x) - \langle x, v\rangle$ is called the corrector and once one shows that $\chi(x) = o(\|x\|)$, the proof may proceed.

The focus of this article is the case of random graphs. Classical tools of geometric analysis do not extend to this context in a straightforward way. Indeed, a random environment is not regular at the microscopic scale. In order to understand harmonic functions, one thus needs to make use only of the control of the macroscopic behavior of the environment. Let us take supercritical percolation as an example; see [44] for background and definitions.

For $p \in (0, 1)$, consider the random graph $G = (V(G), E(G))$ defined by $V(G) = V(\mathbb{Z}^d)$ and $E(G)$ being a random set containing each edge of $\mathbb{Z}^d$ with probability p, independently of the other edges. It is classical that (in dimension $d \geq 2$) there exists $p_c(d) \in (0, 1)$ such that for $p < p_c(d)$, there is almost surely no infinite connected component (also called cluster), while for $p > p_c(d)$, there is a unique infinite cluster. When $p > p_c(d)$, we denote this cluster by ω.

THEOREM 1. *Let $d \geq 2$, and let $p > p_c(d)$. Then with probability* 1, *the infinite cluster ω has no nonconstant sublinear harmonic functions.*

This immediately shows that the corrector χ is unique, as was conjectured by Berger and Biskup [16], Question 3.

In more regular settings, claims of this sort have been proved using the following strategy: try to show that two random walks starting at neighbors will couple before time n with probability bigger than $1 - Cn^{-1/2}$. This fact is classical in the case of the hypercubic lattice $\mathbb{Z}^d$ where an explicit coupling can be exhibited. In the random context it is not clear how to construct an explicit coupling, but a number of approaches in the literature allows one to construct a coupling indirectly. The known Gaussian heat kernel bounds [see (1) below] allow one to construct a coupling that will fail with probability $n^{-\epsilon}$. Using also the central limit theorem already mentioned, one could improve this to $n^{-1/2+o(1)}$. Nevertheless, getting the precise $n^{-1/2}$ seems difficult with these approaches. The approach we will apply below not only gives the precise order $n^{-1/2}$, but the proof is also significantly simpler than those just suggested.

The proof uses an entropy argument similar to Avez [3] who showed that a Cayley graph satisfies the Liouville property if the entropy of the random walk on it is sublinear. In fact the "if" here is an "if and only if" as was shown by Kaimanovich and Vershik [51] and, with a different approach, by Derriennic [33], but we will not need the other direction. Two extensions of this result were known before: it applies to random graphs [11], and it can be quantified [39], Section 5. It turns out that the two generalizations can be applied simultaneously. Further, Theorem 1 is but an example: the techniques work in great generality; even reversibility is not needed. Only stationarity of the walk and some weak (sub-)diffusivity are used. Precise assumptions are detailed below.

The environment as viewed from the particle. To state the full result, we need to define what we mean by "environment." We are interested in environments which are somehow translation invariant. This notion extends the transitivity condition to the random context. Historically, this traces to the works of Papanicolaou and Varadhan [70] and Kozlov [56] who studied random walk in random environments on $\mathbb{Z}^d$ by translating the environment so that the walker remains at $\vec{0}$. In other words, instead of having a walker move around in some environment, the walker stays at the origin, and the environment moves "below" it, hence the name *the environment as viewed from the particle*. When the distribution of the environment stays the same after a single step of this process, the environment is called *stationary*.

The notion was extended beyond $\mathbb{Z}^d$ in [62] who showed a similar phenomenon for Galton–Watson trees: when you do a single step of random walk starting from the root of the tree, the resulting random graph has the same distribution with respect to the new position of the walker.

In such examples the most natural definition of "having the same distribution" uses isomorphisms (in [62] this could be avoided due to the very simple structure of trees, but it appears, e.g., in [1, 11]). The resulting definition looks a little abstract at first, but in fact is very easy to verify in examples. For example, in the $\mathbb{Z}^d$ case, the isomorphisms would be translations, while in the Galton–Watson case, they would be a change of root followed by an arbitrary map. Let us give the details.

Consider a Markov chain $(X_n)_{n\geq 0}$ taking values in some set V. The law of this chain can be encoded by a function $P : V \times V \to [0, 1]$ where $P(x, y)$ denotes the probability to move from x to y. We always assume that our Markov chain is irreducible, that is, that for any $v, w \in V$ there is an n such that $P^n(v, w) > 0$. A *rooted Markov chain* is a triplet (P, V, ρ) where $\rho \in V$ is some vertex that will be called the *root vertex*. Two rooted Markov chains (P, V, ρ) and (P', V', ρ') are considered isomorphic if there is a one-to-one map $\phi : V \to V'$ with $\phi(\rho) = \rho'$ and $P(x, y) = P'(\phi(x), \phi(y))$.

We define an *environment as viewed from the particle*, abbreviated as simply *environment*, to be a random rooted Markov chain. Two environments are considered to have the same law if they are identical as measures on isomorphism classes of rooted Markov chains (alternatively, if they can be coupled in such a way that the resulting rooted Markov chains are isomorphic with probability 1).

DEFINITION 2. An environment (P, V, ρ) is called *stationary* if it has the same law as (P, V, X_1) where X_1 is sampled from $P(\rho, \cdot)$.

As we already remarked, stationary environments are very common, and we provide ten examples in the end of Section 2. Most of these examples are embedded in $\mathbb{Z}^d$, and for these we could have used the definition of [56, 70]. Examples 2.6, 2.8 and 2.10, however, are not embeddable into $\mathbb{Z}^d$, so the isomorphism cannot be taken to be a "translation," though constructing it is still easy.

A very important subset of stationary environments is given by environments V with the structure of a weighted graph [with the weight being a symmetric positive function ν on every edge $(v, w) \in E$, and 0 on every pair $(v, w) \notin E$]. In such case, P is given by

$$P(v, w) = \frac{\nu(v, w)}{\nu(v)} \qquad \text{where } \nu(v) = \sum_x \nu(v, x).$$

These environments will be called *random stationary graphs*. This particular type of Markov chain is also commonly called *reversible*. The reversible case has a rich theory; see, for example, [1, 11] where one can also find many more examples. To clearly distinguish between the reversible and nonreversible case, random stationary graphs will be denoted by (G, ν, ρ) where G is the graph, ν is the weight function and ρ is the root.

The graph distance in G is denoted by $\mathbf{d}^G(\cdot, \cdot)$ and the ball of size r centered at x by $\mathbf{B}_x^G(r)$. We will also consider this distance in nonreversible setting, where it is simply the smallest n such that $P^n(x, y) > 0$ (in this case it may fail to be a metric). Since the distinction between annealed and quenched statements will be clear in the context, we will often drop the dependence on G in the notation. For instance, $\mathbf{P}_x^G$, $\mathbf{d}^G(\cdot, \cdot)$ and $\mathbf{B}_x^G(n)$ will become simply $\mathbf{P}_x$, $\mathbf{d}(\cdot, \cdot)$ and $\mathbf{B}_x(n)$. For the convenience of the reader, we collected the notation and conventions used in this paper in the last section of the introduction (page 2340).

Nonconstant harmonic functions with minimal growth. Let P be a Markov chain with state space V. Then a function $h : V \to \mathbb{R}$ is called harmonic if $h(X_n)$ is a martingale, or in other words, if

$$h(x) = \sum_y P(x, y)h(y) \qquad \forall x.$$

As already mentioned, harmonic functions have had a number of important applications recently. Let us expand on the particular application in Kleiner's proof of Gromov's theorem [55]. It was known since the 1970s that in order to prove Gromov's theorem, it is enough to show that any group with polynomial volume growth has a nontrivial finite-dimensional representation. Kleiner showed that any group has a nontrivial linearly growing harmonic function, and that on groups with polynomial growth, the dimension of the space polynomially growing harmonic functions is finite. Since the group acts on harmonic functions on its Cayley graph by translations, this provides a finite dimensional representation and proves Gromov's theorem. Shalom and Tao [73] showed that a quantitative version of Kleiner's proof can be performed. Further, they characterized the linearly growing harmonic functions (for groups with polynomial volume growth these are the nonconstant harmonic functions with minimal growth [46], Theorem 6.1). They showed (personal communication) that when the group is nilpotent, any such function must be a character of the group (or the sum of a character and a constant), in analogy to the Choquet–Deny theorem [25, 65]. For virtually nilpotent groups this holds *mutatis mutandis*. We plan to analyze harmonic functions with minimal growth in the context of Cayley graphs, especially of wreath products, in a future paper.

We now return to the setting of this paper, that is, of stationary random graphs. Using the entropy of the random walk, it is possible to bound from below the minimal growth of nonconstant harmonic functions in terms of the rate of escape of the random walk. A particularly interesting case is provided by stationary environments with diffusive behavior, for which the bound is often sharp. A stationary environment (P, V, ρ) satisfies *diffusive or subdiffusive behavior* (*DB*) if

(DB) there exists $C > 0$ such that $\mathbb{E}(\mathbf{d}(\rho, X_n)^2) \leq Cn$ for every n.

Here and below $\mathbb{E}$ is the average over both the environment and over the walk (the so-called annealed average). We may now state our main result.

THEOREM 3. *Let (P, V, ρ) be a stationary environment such that $\mathbb{E}(|\mathbf{B}_\rho(n)|) \leq Cn^d$ for some constants $C, d < \infty$ independent of n. If (P, V, ρ) satisfies* (DB), *then for almost every environment, there are no nonconstant sublinear harmonic functions.*

We say that h is a sublinear function if $h(x) = o(\mathbf{d}(\rho, x))$ as $\mathbf{d}(\rho, x) \to \infty$. Restricting to the case of percolation, it is also quite natural to ask what happens with functions which are sublinear with respect to the Euclidean distance $\|x\|$ (e.g., this is how the question is formulated in [16]). The result of Antal and Pisztora [2] yields that graph and Euclidean distances are comparable on the infinite cluster, and that therefore the previous question follows from Theorem 3.

As already stated, Theorem 3 applies to many different models, some of them significantly less well understood than percolation. See a list of examples at the end of Section 2.

Whether (DB) follows from polynomial growth in the reversible case is an interesting question. The Carne–Varopoulos bound [23, 77] gives that $\mathbf{E}_\rho(d(\rho, X_n)) \leq C\sqrt{n \log n}$, which would give (with the same proof as that of Theorem 3; see Theorem 3′ in Section 2) that any stationary random graph with polynomial volume growth has no nonconstant harmonic functions h with $h(x) \leq C\mathbf{d}(\rho, x)/\sqrt{\log \mathbf{d}(\rho, x)}$. Without stationarity the Carne–Varopoulos bound $\sqrt{n \log n}$ cannot be improved, as was shown by Barlow and Perkins [10]. Kesten gave a beautiful argument that a stationary random graph embedded in $\mathbb{Z}^d$ satisfies (DB); see, for example, [10], Section 2. But it does not seem to apply just assuming polynomial growth.

The relation between entropy, harmonic functions and speed of the random walk holds for more general environments (e.g., with larger growth). We defer to Section 2 for a more complete account of this question.

Polynomially growing functions. As in the case of manifolds, we are interested in the dimension of the space of harmonic functions with prescribed polynomial growth. Of course, one can encounter very different behavior depending on the environment (like in the deterministic case). Hence we will assume that our environments satisfy volume doubling and the Poincaré inequality. Here is the precise formulation of our assumptions on the environment: let (G, ν, ρ) be a rooted weighted graph.

$(VD)_G$. (G, ν, ρ) satisfies the *anchored volume doubling property* $(VD)_G$ if there exists $0 < \mathbf{C}_{\mathrm{VD}} < \infty$ such that the following holds. For every $\lambda < \infty$, there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$, and for every $x \in \mathbf{B}_\rho(\lambda n)$,

$$\nu\big(\mathbf{B}_x(2n)\big) \le \mathbf{C}_{\mathrm{VD}}\nu\big(\mathbf{B}_x(n)\big),$$

where $\nu(\mathbf{B})$ is the total weight of the edges in the ball $\mathbf{B}$.

$(P)_G$. (G, ν, ρ) satisfies the *anchored Poincaré inequality* $(P)_G$ if there exists $\mathbf{C}_{\mathrm{P}} < \infty$ such that the following holds. For every $\lambda < \infty$, there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$, for every $x \in \mathbf{B}_\rho(\lambda n)$ and every $f : \mathbf{B}_x(2n) \to \mathbb{R}$,

$$\sum_{y \in \mathbf{B}_x(n)} \big(f(y) - \overline{f}_{\mathbf{B}_x(n)}\big)^2 \nu(y) \le \mathbf{C}_{\mathrm{P}} n^2 \sum_{(y,z) \in E(\mathbf{B}_x(2n))} \big|f(y) - f(z)\big|^2 \nu(y, z),$$

where

$$\overline{f}_{\mathbf{B}_x(n)} = \frac{1}{\nu(\mathbf{B}_x(n))} \sum_{y \in \mathbf{B}_x(n)} f(y)\nu(y).$$

Similar properties are classical in geometric analysis. They go back to the theory developed by De Giorgi, Nash and Moser [35, 67–69] in the fifties and sixties for uniformly elliptic second-order operators in divergence form. In the classic context, they imply the Harnack principle and Gaussian bounds for the heat kernel. While the definitions above have no randomness in them, they are tailored for the random case: they take into consideration that in most examples of interest these properties do not hold from every point since some unusual points always exist. For this reason, the properties are required to hold for balls which are not too far from our root ρ, relative to their size. This is reminiscent of Barlow's good and very good balls [6], but our requirements are much weaker, we only need the properties to hold for "macroscopic balls," balls whose distance to ρ is proportional to their radius.

Let us remark on the appearance of the number 2 in $\mathbf{B}_x(2n)$ in both properties. For the volume doubling property it is clear that these properties are equivalent for all choices bigger than 1; that is, if one was to define a "3-volume doubling property," then it would be equivalent to the "2-volume doubling property" defined above, though perhaps with different $\mathbf{C}_{\mathrm{VD}}$ and minimal n. The same holds for the Poincaré inequality, under the assumption of volume doubling. This is well known in the standard settings (see, e.g., [49], Section 5), and the proof carries over to the anchored case without any change.

With these definitions we can state the following easy but, we believe, conceptually important theorem. Note that the theorem is for a fixed graph (though the most interesting applications are for random graphs).

THEOREM 4. *Let (G, ν, ρ) be a rooted weighted graph. If (G, ν, ρ) satisfies $(VD)_G$ and $(P)_G$, then for every $k > 0$, the space of harmonic functions with $|h(x)| \leq C\mathbf{d}(\rho, x)^k$ for all x far enough from ρ, is finite dimensional.*

Further, the bound on the dimension depends only on k, $\mathbf{C}_{\mathrm{VD}}$ and $\mathbf{C}_{\mathrm{P}}$, and not on $n_0(\lambda)$.

This theorem represents a discrete anchored version of Yau's conjecture except that the Poincaré inequality must be assumed since it is not automatically satisfied (in Yau's settings every manifold with nonnegative Ricci curvature satisfies a Poincaré inequality [21] while in Kleiner's, every group satisfies an appropriate version of the Poincaré inequality; see, e.g., [71], Lemma 4.1.1). The proof of this theorem follows the existing strategy developed in [26, 31, 55, 73, 75]. Let us stress again that the interesting part is that it requires only *macroscopic* volume growth and Poincaré inequality: the definitions of $(VD)_G$ and $(P)_G$ only examine balls of radius n inside $\mathbf{B}_\rho(\lambda n)$ for some finite λ.

When we apply Theorem 4, the graph G will be random. Since the dimension depends only on $\mathbf{C}_{\mathrm{VD}}$ and $\mathbf{C}_{\mathrm{P}}$, then in particular, if these constants are not random, neither is the bound. Thus, for example, in supercritical percolation there is a constant A (depending only on the dimension d and the probability p) such that $\mathbf{C}_{\mathrm{VD}} \leq A$ and $\mathbf{C}_{\mathrm{P}} \leq A$ almost surely (the minimal n is the only quantity which really changes between configurations). Hence for each k there is a number D_k such that the dimension of harmonic functions of growth at most of order $\mathbf{d}(\rho, x)^k$ is smaller than D_k, almost surely. We discuss a few other examples of random graphs satisfying $\mathbf{C}_{\mathrm{VD}}$ and $\mathbf{C}_{\mathrm{P}}$ in the end of Section 3, but in general one should keep in mind that the Poincaré inequality restricts the behavior of random walk on the graph significantly, so Theorem 4 applies in much less generality than Theorem 3.

Linearly growing functions. In the special case of environments which are modifications of $\mathbb{Z}^d$, we can compare the dimension of harmonic functions with a prescribed growth to the dimension of harmonic functions on $\mathbb{Z}^d$. The simplest perturbation of $\mathbb{Z}^d$ is the supercritical cluster of percolation. We prove the following theorem.

THEOREM 5. *Let $d \geq 2$. For $p > p_c(d)$, let ω be the unique infinite component of percolation on $\mathbb{Z}^d$. Then, the dimension of the vector space of harmonic functions with growth at most linear on ω is equal to $d + 1$ almost surely.*

This theorem must be understood as a first step toward a bigger goal, which would be to compute the dimension of all spaces of harmonic functions with prescribed (polynomial) growth.

The properties of the supercritical percolation cluster used in this proof are quite general: the d-dimensional volume growth and the Poincaré inequality $(P)_\omega$ proved (in stronger form) by Barlow [5] as well as the Gaussian bounds which Barlow concludes from these, and an invariance principle [16, 66, 74]. All these properties witness the close relation between macroscopic properties of the supercritical percolation cluster and $\mathbb{R}^d$. In some sense, it confirms the heuristic that this cluster is an approximation of $\mathbb{Z}^d$.

Heat kernel estimates. Classically [35, 67–69], the kernels of symmetric diffusions are known to have some Hölder regularity. In random environments, few results are known on Hölder behavior: Conlon and Naddaf [27] and Delmotte and Deuschel [32] treated the case of random conductance with a uniform ellipticity condition; see also [43]. The entropy techniques developed for the proof of Theorem 1 allow one to give a very short proof that the space derivative exists. Moreover, it applies in a very general context. We present the case of percolation.

THEOREM 6. *Let $d \geq 2$ and $p > p_c(d)$. Let $\mathbb{P}_p$ be the measure of the infinite cluster of percolation (denoted ω) on $\mathbb{Z}^d$. There exist $C_3, C_4 > 0$ such that for every $n > 0$ and x, x', y at distance less than n of 0, if x and x' are adjacent,*

$$\mathbb{E}_p\big[\big(\mathbf{p}_n(x, y) - \mathbf{p}_{n-1}(x', y)\big)^2 \mathbf{1}_{\{y \in \omega\}} \mathbf{1}_{\{x \text{ and } x' \text{ are adjacent in } \omega\}}\big] \\ \leq \frac{C_3}{n^{d+1}} \exp\big[-C_4 \mathbf{d}(x, y)^2/n\big],$$

where $\mathbf{p}_n(y, x) := \mathbf{P}_y(X_n = x)$ and X_n is the random walk on ω.

Estimates for the heat kernel itself (i.e., not for the derivative) are well understood, and are known as Gaussian estimates (*GE*). Heuristically, Gaussian estimates are bounds of the form

$$\frac{C_1}{n^{d/2}} \exp\big[-C_2 \mathbf{d}(x, y)^2/n\big] \leq \mathbf{P}_x[X_n = y] \leq \frac{C_3}{n^{d/2}} \exp\big[-C_4 \mathbf{d}(x, y)^2/n\big].$$

A few caveats are in place, though. The lower bound cannot hold if there is any kind of periodicity (as in $\mathbb{Z}^d$ or in subgraphs of it, such as supercritical percolation). One should talk about continuous time random walk, lazy random walk, or replace $\mathbf{P}_x[X_n = y]$ with $\mathbf{P}_x[X_n = y] + \mathbf{P}_x[X_{n+1} = y]$. Further, the lower bound does not hold for x and y extremely far away—if $\mathbf{d}(x, y) > n$, then the probability is just zero (in the simple random walk case).

In the case of the infinite cluster of supercritical percolation, these bounds were obtained for continuous time random walk in [6]. They also hold for simple random walk, most of the details are filled in [9]. Again, one should be careful, as (with small probability) the environment in the neighborhood of ρ might be atypical, breaking these estimates for small n. Hence the formulation is as follows. There exist strictly positive constants C_1, C_2, C_3 and C_4 such that for almost every environment ω there exist random variables $n_x(\omega)$, $x \in \mathbb{Z}^d$ so that for every $x, y \in \omega$ and $n > \max\{n_x(\omega), \mathbf{d}(x, y)\}$

$$\begin{aligned}\frac{C_1}{n^{d/2}} \exp[-C_2 \mathbf{d}(x, y)^2/n] &\le \mathbf{P}_x[X_n = y] + \mathbf{P}_x[X_{n+1} = y] \\ &\le \frac{C_3}{n^{d/2}} \exp[-C_4 \mathbf{d}(x, y)^2/n].\end{aligned} \tag{1}$$

Moreover, the random variables $n_x(\omega)$ satisfy a stretched exponential estimate, that is,

$$\mathbb{P}_p(x \in \omega, n_x(\omega) \ge s) \le c e^{-cs^{\varepsilon}} \tag{2}$$

for some $\varepsilon > 0$.

For the proof of Theorem 6 we only need the upper bound in (1). For the proof of Theorem 5 we will also need the lower bound, but only in the regime $|x - y| \approx \sqrt{n}$, that is, in the regime where the probabilities are of order $n^{-d/2}$.

Organization of the paper. In the next section, we study the notion of mean entropy of random walks on a stationary random graph to bound the total variation between random walks starting at neighbors. We deduce Theorem 3. Section 3 contains the proof that $(VD)_G$ and $(P)_G$ imply that the space of harmonic functions of prescribed polynomial growth is finite dimensional, that is, Theorem 4. Section 4 deals with the example of the supercritical percolation cluster and analyzes the space of linearly growing harmonic functions. It is completely independent of Section 3. Section 5 contains the proof of Theorem 6. Section 6 regroups some open questions.

Notation. To make the distinction between the reversible and nonreversible case clear, we call the general case "Markov chain" and denote it by (P, V), where V is the space and $P : V \times V \to [0, 1]$ are the transition probabilities, $P(x, y)$ being the probability to move from x to y. We often write P^n which we interpret as a matrix power—of course, $P^n(x, y)$ is also the probability that a random walk starting from x will be at y after n steps.

Any reversible chain can be described as a random walk on a weighted graph. If G is a graph and ν is a function on the edges of G taking values in $[0, \infty)$, then the Markov chain is given by $P(x, y) = \nu(x, y)/\sum_z \nu(x, z)$. Here and below, $\nu(x, y)$ for two vertices x and y is the weight of the edge (x, y). In particular, $\nu(x, y) = \nu(y, x)$, and if (x, y) is not an edge of the graph, then we set $\nu(x, y) = 0$. We will always denote reversible Markov chains by (G, ν). We denote by $E(G)$ the set of edges of the graph G, and for a set of vertices S we denote by $E(S)$ the set of edges between the vertices of S. The notation $x \sim y$ for two vertices will mean that $(x, y) \in E(G)$, that is, that they are neighbors in the graph.

We also consider ν as a measure. For a vertex x, we will denote $\nu(x) = \sum_{y\sim x} \nu(x, y)$ while for a set of vertices S, we will denote $\nu(S) = \sum_{x\in S} \nu(x)$. Note that edges between two vertices of S are counted twice in this sum.

For a fixed graph or Markov chain we denote by $\mathbf{E}$ the expectation with respect to the random walk on that fixed graph. When the starting point of the random walk is specified, we will use subscripts and write for instance $\mathbf{E}_\rho$. The symbol $\mathbb{E}$ is used to denote the expectation with respect to both the environment and the random walk (the "annealed" average). Similarly, bold letters will usually denote "quenched" objects, that is, objects related to an instance G of the environment. The quantity $\mathbf{d}(x, y)$ will denote the graphical distance between two vertices x and y of G, that is, the length of the shortest path in G between x and y, or, in the nonreversible setting, the minimal n such that $P^n(x, y) > 0$. The ball $\{y : \mathbf{d}(x, y) \le r\}$ will be denoted by $\mathbf{B}_x(r)$.

Constants which depend on the environments G are denoted $\mathbf{c}_i$, while constants of the form C_i will refer to constants uniform in the environment. We will occasionally write $\mathbf{c}$ or C for a constant—different appearances of $\mathbf{c}$ or C might be different constants.

The cardinality of a set E will be denoted by $|E|$.

2. The entropy argument. The connection between entropy and random walks was first exhibited by Avez [3] and then made famous in a celebrated paper of Kaimanovich and Vershik [51]; see also Derriennic [33]. For any discrete variable X the entropy is defined by

$$H(X) = \sum_x \phi\bigl(P(X = x)\bigr) \qquad \text{where } \phi(0) = 0 \text{ and } \phi(t) = -t \log t \text{ for any } t > 0.$$

Conditional entropy can be defined by

$$H(X|Y) = \mathbb{E}\bigl[H(X|Y = y)\bigr] = \sum_y P(Y = y) \sum_x \phi\bigl(P(X = x|Y = y)\bigr).$$

It is then quite simple to show that $H(X|Y) = H(X, Y) - H(Y)$ and that $H(X|Y, Z) \le H(X|Y)$ for any three random variables X, Y and Z.

Consider a stationary environment (P, V, ρ) with law $\mathbb{P}$. Conditionally on (P, V, ρ), define the *entropy* of the random walk at times n, m started at ρ by

$$\mathbf{H}_{n,m}(P, V, \rho) = H(X_n, X_m) = \sum_{x,y\in V} \phi\bigl(\mathbf{P}_\rho(X_n = x, X_m = y)\bigr).$$

When $n = m$, we simply denote $\mathbf{H}_{n,n}(P, V, \rho)$ by $\mathbf{H}_n(P, V, \rho)$. In the random context, we define the *mean entropy* (see [11]) by

$$H_{n,m} = \mathbb{E}\bigl[\mathbf{H}_{n,m}(P, V, \rho)\bigr] \quad \text{and} \quad H_n = \mathbb{E}\bigl[\mathbf{H}_n(P, V, \rho)\bigr].$$

There are many ways of measuring the distance between two probability measures μ and ν on some set V, the most standard one being the total variation

$$\|\mu - \nu\|_{\mathrm{TV}} := \frac{1}{2} \sum_{x\in V} \bigl|\mu(x) - \nu(x)\bigr|.$$

In this article, we will use a less standard one. Define $\Delta(\mu, \nu)$ by the formula

$$\Delta(\mu, \nu) := \left[\sum_{x \in V} \frac{(\mu(x) - \nu(x))^2}{\mu(x) + \nu(x)}\right]^{1/2}. \tag{3}$$

Estimating the distance using Δ is stronger than via the total variation: by Cauchy–Schwarz,

$$\begin{aligned} 2\|\mu - \nu\|_{\mathrm{TV}} &= \sum_{x \in V} \left|\mu(x) - \nu(x)\right| = \sum_{x \in V} \sqrt{\mu(x) + \nu(x)} \frac{|\mu(x) - \nu(x)|}{\sqrt{\mu(x) + \nu(x)}} \\ &\le \sqrt{\left(\sum_{x \in V} \mu(x) + \nu(x)\right)\left(\sum_{x \in V} \frac{(\mu(x) - \nu(x))^2}{\mu(x) + \nu(x)}\right)} \qquad (4) \\ &= \sqrt{2}\Delta(\mu, \nu). \end{aligned}$$

This quantity has an advantage compared to the total variation: for any $f : G \to \mathbb{R}$, we have (using Cauchy–Schwarz similarly)

$$\left|\mu(f) - \nu(f)\right| \le \Delta(\mu, \nu)\left(\mu\left(f^2\right) + \nu\left(f^2\right)\right)^{1/2}. \tag{5}$$

With the total variation, one would obtain a similar but weaker inequality with the L^∞-norm in place of the L^2-norm (the former can in principle be much larger than the later). The reasons for using Δ (rather than, say, the total variation distance) will be discussed in more detail on page 2346, but most readers would be better served by reading the paper linearly, that is, first see how Δ is used to prove Theorem 3 and only then take a look at this discussion.

Let us introduce a convenient notation, used only in this section. Let $\mathscr{L}(Z)$ denote the law of a random variable Z, that is, the measure on the space of values of Z induced by it. If $\mathscr{E}$ is some event, then we will denote by $\mathscr{L}(Z|\mathscr{E})$ the law of Z conditioned on $\mathscr{E}$ happening.

With this notation, we are now in a position to state an important lemma, which is a quantitative version of the following well-known fact: for any two random variables X and Y, $H(X, Y) \le H(X) + H(Y)$ with equality holding if and only if X and Y are independent.

LEMMA 7. *For any two random variables X and Y,*

$$\sum_y P(Y = y)\Delta^2\left(\mathscr{L}(X), \mathscr{L}(X|Y = y)\right) \le 2\left(H(X) + H(Y) - H(X, Y)\right). \tag{6}$$

PROOF. We first note that for $t > 0$,

$$2t \log t \ge \frac{(t - 1)^2}{t + 1} + 2t - 2 \tag{7}$$

[this can be seen by Taylor expanding $t\log t$ to the second order at 1, which gives that $t\log t = t - 1 + \frac{(t-1)^2}{2t^*}$ for some t^* in the interval between t and 1, so $t^* \le t+1$]. Denote

$$p(x) = P(X = x), \qquad p(y) = P(Y = y), \qquad p(x,y) = P(X = x, Y = y).$$

Then, the left-hand side of (6) is [recall the definition (3) of Δ]

$$\begin{aligned}\mathrm{LHS} &= \sum_y p(y) \sum_x \frac{(p(x,y)/p(y) - p(x))^2}{p(x,y)/p(y) + p(x)} \\ &= \sum_{y,x} p(x)p(y)\biggl(\frac{(p(x,y)/(p(x)p(y)) - 1)^2}{p(x,y)/(p(x)p(y)) + 1} + \underbrace{2\frac{p(x,y)}{p(x)p(y)} - 2}_{0}\biggr),\end{aligned}$$

where we were allowed to add the expression denoted by $\underbrace{\quad}_{0}$ since summing over x and y makes these terms cancel out (they are both equal to 2). Using (7) this gives

$$\begin{aligned}\mathrm{LHS} &\overset{(7)}{\le} 2\sum_{x,y} p(x)p(y)\biggl(\frac{p(x,y)}{p(x)p(y)} \log \frac{p(x,y)}{p(x)p(y)}\biggr) \\ &= 2\sum_{x,y} p(x,y)\bigl(\log p(x,y) - \log p(x) - \log p(y)\bigr) \\ &= 2\bigl(-H(X,Y) + H(X) + H(Y)\bigr),\end{aligned}$$

where in the last equality we used that $\sum_y p(x,y) = p(x)$ and $\sum_x p(x,y) = p(y)$. $\square$

We will always be interested in the particular case of random walks. In order to lighten the notation, we set

$$\begin{aligned}\Delta_n(x,y) &:= \Delta\bigl(\mathscr{L}(X_n|X_0 = x), \mathscr{L}(X_{n-1}|X_0 = y)\bigr) \\ &= \Delta\bigl(\mathscr{L}(X_n|X_0 = x), \mathscr{L}(X_n|X_1 = y)\bigr),\end{aligned} \tag{8}$$

the last equality following by the Markov property [recall that $\mathscr{L}(X|\mathscr{E})$ denotes the law of X conditioned on $\mathscr{E}$]. Note that the second measure is the law of the random walk after $n-1$ steps, so the definition is not symmetric in x and y.

Lemma 7 is used to proved the following theorem.

THEOREM 8. *Let (P, V, ρ) be a stationary environment. For every $n > 0$, we have*

$$\mathbb{E}\bigl(\Delta_n(\rho, X_1)^2\bigr) \le 2(H_n - H_{n-1}) \tag{9}$$

(*as usual $\mathbb{E}$ is over both the environment and the randomness of X_1*).

Before proving Theorem 8, we state a result from [11] concerning $H_{1,n}$. We isolate it from the rest of the proof because it is the only place where stationarity is used (stationarity replaces transitivity as used in the context of groups).

LEMMA 9. *Let (P, V, ρ) be a stationary environment. For every $n > 0$, we have $H_{1,n} = H_{n-1} + H_1$.*

PROOF. Fix $n > 0$. A simple computation leads to

$$\begin{aligned}
&\mathbf{H}_{1,n}(P, V, \rho) \\
&\qquad = \sum_{x\sim\rho, y\in G} \phi(\mathbf{P}_\rho(X_1 = x, X_n = y)) \\
&\qquad = \sum_{x\sim\rho} \mathbf{P}_\rho(X_1 = x) \sum_{y\in G} \phi(\mathbf{P}_\rho(X_n = y | X_1 = x)) + \sum_{x\sim\rho} \phi(\mathbf{P}_\rho(X_1 = x)),
\end{aligned}$$

which we simplify using the Markov property giving

$$\mathbf{P}_\rho(X_n = y | X_1 = x) = \mathbf{P}_x(X_{n-1} = y).$$

Taking the expectation with respect to the environment we obtain

$$\begin{aligned}
H_{1,n} &= \mathbb{E}\Big[\sum_{x\sim\rho} \mathbf{P}_\rho(X_1 = x) \sum_{y\in G} \phi(\mathbf{P}_x(X_{n-1} = y))\Big] + \mathbb{E}\Big[\sum_{x\sim\rho} \phi(\mathbf{P}_\rho(X_1 = x))\Big] \\
&= \mathbb{E}[\mathbf{H}_{n-1}(P, V, X_1)] + \mathbb{E}[\mathbf{H}_1(P, V, \rho)] = H_{n-1} + H_1,
\end{aligned}$$

where in the last equality we used the fact that (P, V, X_1) has the same law as (P, V, ρ) (this is not a property of entropy, it would hold for any function of the environment). $\square$

Before continuing, let us state one corollary of Lemma 9 which is not necessary for the proof of Theorem 8 but does shed some light on the quantities involved.

COROLLARY 10. *$H_n - H_{n-1}$ is decreasing.*

In other words, the sequence H_n is concave.

PROOF. By Lemma 9,

$$H_n - H_{n-1} = H_n - H_{1,n} + H_1 = \mathbb{E}[\mathbf{H}_n - \mathbf{H}_{1,n}] + H_1.$$

The quantity $\mathbf{H}_n - \mathbf{H}_{1,n}$ can be written as the conditioned entropy $-H(X_1|X_n)$ where X_n is the random walk at time n (this statement is quenched). This, however, increases since

$$\mathbf{H}(X_1|X_n) = \mathbf{H}(X_1|X_n, X_{n+1}) \le \mathbf{H}(X_1|X_{n+1}), \tag{10}$$

where the equality is due to the fact that conditioned on X_n, knowing X_{n+1} gives you no information about what happened before time n; that is, by the Markov property, conditional on X_n we have that X_1 is independent of X_{n+1}. The inequality in (10) is a generic fact about entropy—conditioning on more information reduces the relative entropy [namely, $H(X|Y,Z) \le H(X|Y)$ for any three random variables X, Y and Z]. Hence $\mathbf{H}_n - \mathbf{H}_{1,n}$ decreases, and so does its expectation. $\square$

PROOF OF THEOREM 8. This is a direct corollary of Lemmas 9 and 7. Indeed, by Lemma 7,

$$\mathbf{E}\big(\Delta_n(\rho, X_1)^2\big) = \sum_x \mathbf{P}(X_1 = x)\Delta\big(\mathscr{L}(X_n), \mathscr{L}(X_n|X_1 = x)\big)^2 \\ \le 2(\mathbf{H}_1 + \mathbf{H}_n - \mathbf{H}_{1,n}).$$

We now take expectation with respect to the environment and get from Lemma 9 that

$$\mathbb{E}\big(\Delta_n(\rho, X_1)^2\big) \le 2\mathbb{E}(\mathbf{H}_1 + \mathbf{H}_n - \mathbf{H}_{1,n}) = 2(H_n - H_{n-1}). \qquad \square$$

We are now in a position to prove Theorem 3.

PROOF OF THEOREM 3. We only need to prove that for almost every environment, $h(\rho) = h(X_1)$ a.s., for any sublinear harmonic function. Indeed, stationarity would then imply that for almost every P, $h(X_n) = h(X_{n+1})$ a.s. for any sublinear harmonic function. Since the Markov chain is irreducible, (X_n) can visit any vertex, and we deduce that almost surely any sublinear harmonic function is constant.

For any harmonic function h with respect to the environment, we have for all x and n,

$$h(x) = \mathbf{E}_x\big(h(X_n)\big).$$

We use this twice, once for $x = \rho$ and once for an arbitrary x and $n - 1$. We get

$$\big|h(\rho) - h(x)\big| = \big|\mathbf{E}_\rho\big[h(X_n)\big] - \mathbf{E}_x\big[h(X_{n-1})\big]\big| \\ \text{by (5)} \le \Delta_n(\rho, x)\sqrt{\mathbf{E}_\rho\big[h^2(X_n)\big] + \mathbf{E}_x\big[h^2(X_{n-1})\big]}.$$

We use this with $x = X_1$, integrate over X_1 and get

$$\mathbf{E}_\rho\big|h(\rho) - h(X_1)\big| \le \mathbf{E}_\rho\Big[\Delta_n(\rho, X_1)\sqrt{\mathbf{E}_\rho\big[h^2(X_n)\big] + \mathbf{E}_{X_1}\big[h^2(X_{n-1})\big]}\Big] \\ \text{by Cauchy–Schwarz} \le \sqrt{2\mathbf{E}_\rho\big[\Delta_n(\rho, X_1)^2\big]\mathbf{E}_\rho\big[h^2(X_n)\big]}, \tag{11}$$

where in the last line we also used that $\mathbf{E}_\rho[\mathbf{E}_{X_1}[h^2(X_{n-1})]] = \mathbf{E}_\rho[h^2(X_n)]$.

By assumption, the Markov chain has annealed polynomial growth. Therefore, the entropy satisfies

$$H_n \le \mathbb{E}\big[\log\big|\mathbf{B}_\rho(n)\big|\big] \le \log \mathbb{E}\big[\big|\mathbf{B}_\rho(n)\big|\big] \le \log\big[Cn^d\big]$$

and is at most logarithmic (we used the fact that log is concave). Hence $H_n - H_{n-1} \le c/n$ for infinitely many n. Using Theorem 8 and (DB) we get

$$\mathbb{E}\big[n\Delta_n(\rho, X_1)^2\big] + \mathbb{E}\big[n^{-1}\mathbf{d}(X_n, \rho)^2\big] \le C \qquad \text{for infinitely many } n.$$

Hence, by Fatou's lemma, for almost every environment there exists $\mathbf{c}_1 < \infty$ such that

$$\mathbf{E}_\rho\big[n\Delta_n(\rho, X_1)^2\big] + \mathbf{E}_\rho\big[n^{-1}\mathbf{d}(X_n, \rho)^2\big] \le \mathbf{c}_1 \qquad \text{for infinitely many } n, \tag{12}$$

where this time the sequence of n for which it holds depends on the environment, that is, is random.

Now, assume that h has sublinear growth. For any $\varepsilon > 0$, there exists a constant $\mathbf{K}$ such that for all $x \in V$,

$$h^2(x) \le \varepsilon \mathbf{d}(x, \rho)^2 + \mathbf{K}. \tag{13}$$

Putting (13) and (12) in (11), we deduce that for almost every environment, and for every h harmonic and sublinear on it,

$$\mathbf{E}_\rho\big(\big|h(\rho) - h(X_1)\big|\big) \le \mathbf{c}_2\varepsilon^{1/2}.$$

Letting ε go to 0, we deduce that $h(\rho) = h(X_1)$ almost surely for any sublinear harmonic function. $\square$

Inequality (11) relates the entropy to the value of possible harmonic functions at X_n. Its use is not restricted to the case of diffusive environments with polynomial growth. For instance, one can use this inequality to prove a characterization of almost sure Liouville property for stationary random graphs (this was proved in [11] using a more direct generalization of [51]). For completeness, we state the result in [11] here.

COROLLARY 11 ([11]). *Let (P, V, ρ) be a stationary environment. If H_n/n converges to* 0, *then P has the Liouville property* (*i.e., has no nonconstant bounded harmonic functions*) *almost surely.*

We would like to emphasize why we use $\Delta(\mu, \nu)$. Csiszár's inequality [28, 29] relates the total variation between two measures to their relative entropy. In our context, an inequality involving the total variation can also be found, hence giving a bound on the best coupling (in time) between two random walks starting at neighbors. For completeness, we state the inequality here [it is a consequence of (4) applied to (9)]: for a stationary environment (P, V, ρ) and $n > 0$, we have

$$\mathbb{E}\big(\big\|\mathscr{L}(X_n) - \mathscr{L}(X_n|X_1)\big\|_{\mathrm{TV}}^2\big) \le 4(H_n - H_{n-1}).$$

Interestingly, this inequality is not strong enough for our applications, since controlling the probability that two random walks merge before time n says nothing about their behavior when they do not couple.

Other growth rates. The same argument as in Theorem 3 can also be used with growth rates bigger than polynomial. A general statement would be the following.

THEOREM 3′. *Let $\mathbb{P}$ be the measure of a stationary environment (P, V, ρ). For every (nonrandom) sequence $(n_k)_k$ with $n_k \to \infty$, we have that $\mathbb{P}$-a.s. there does not exist a nonconstant harmonic function $h : V \to \mathbb{R}$ such that*

$$\mathbf{E}_\rho\big[h(X_{n_k})^2\big] \cdot (H_{n_k} - H_{n_k-1}) \to 0.$$

In particular this holds for fixed transitive graphs, which is a version of a result of [39], Section 5.

Examples. We finish this section by presenting a collection of examples.

EXAMPLE 2.1 (Random conductance). Consider the graph $\mathbb{Z}^d$, and let ν be given by a shift-invariant law (e.g., i.i.d. positive random variables). We assume that the set of sites connected by edges with positive conductances is infinite. The random walk induces a Markov process on the environment (cf. Kipnis and Varadhan [54]), called the environment as seen from the particle. This process can be made stationary by weighting each configuration proportionally to $\nu(\rho)$.

This model has been studied extensively. Under the assumption of *uniform ellipticity*: $\exists \alpha > 0 : \mathbb{P}[\alpha < \nu(x, y) < 1/\alpha] = 1$, many things are known on the environment. First, the Poincaré inequality is a direct consequence of the $\mathbb{Z}^d$ case. Second, Delmotte proved in [30] that the Poincaré inequality implies that there exist $c_1, c_2 > 0$ such that

$$\mathbf{P}_\rho[X_n = x] < \frac{c_1}{n^{d/2}} e^{-c_2 \mathbf{d}(x,\rho)^2/t}$$

(a corresponding lower bound also holds but is not needed for our purposes). Third, an annealed invariance principle holds in the sense that the law of the paths under the measure integrated over the environment scales to a nondegenerate Brownian motion [54]. In particular, Theorem 3 applies in this case.

Once the assumption of uniform ellipticity is relaxed, matters get more complicated. An example of random conductance models without uniform ellipticity is the infinite cluster of percolation which we will discuss next. For an unusual example of a transitive conductance model, see the work of Disertori, Spencer and Zirnbauer [37] who reduced a supersymmetric hyperbolic sigma model to the study of random walk on a certain (highly correlated) random environment.

EXAMPLE 2.2 (Infinite cluster of percolation). Consider the percolation measure with a parameter p such that there exists an infinite cluster with probability 1.

See [44] for details about percolation. Set $\mathbb{P}_0$ to be the law of the infinite cluster conditioned to contain 0. As in the previous example, the random walk on ω induces a Markov chain on the space Ω of infinite subgraphs of $\mathbb{Z}^d$ containing the origin. When weighting each configuration proportionally to the number of neighbors of the origin we obtain a stationary measure with respect to the shift along the random walk.

Since the infinite cluster of percolation can be seen as a stationary random graph with polynomial volume growth and since the random walk is diffusive [6, 53], Theorem 3 applies, and we get Theorem 1.

EXAMPLE 2.3 (Centered random environments). This is our first nonreversible example. A centered random environment is, roughly speaking, a Markov chain on $\mathbb{Z}^d$ such that the probabilities can be "decomposed" into a sum over cycles. Such environments, even when nonreversible, are still heuristically quite close to reversible, and in particular they have a stationary version which is related to the usual version by an explicit reweighting, like in the reversible case [34], Section 3. See Deuschel and Kösters [34] for a proof of a CLT, which implies (DB)—of course, a CLT is much stronger than (DB). Hence, our results can be applied in this context as well.

EXAMPLE 2.4 (Balanced random environments). This is another nonreversible example, which is "farther" from reversible than the previous one. A balanced random environment is a Markov chain P with state space $\mathbb{Z}^d$ and nearest neighbor movements, such that for every $x \in \mathbb{Z}^d$ and every unit vector e_i, $P(x, x + e_i) = P(x, x - e_i)$. It follows that X_n is a martingale, and hence (DB) is an immediate corollary of the Azuma–Hoeffding inequality. The issue is therefore only stationarity. In the case that the environment μ is uniformly elliptic and stationary and ergodic *to the action of* $\mathbb{Z}^d$ (this is different from our notion of stationarity!), Lawler showed that there exists a stationary measure (in our sense) λ which is mutually absolutely continuous with respect to μ; see [59], Theorem 3. Hence our results apply to λ, and hence also to μ. Guo and Zeitouni weakened the requirement of uniform ellipticity to just ellipticity, at the price of restricting the environment to the i.i.d. case [45]. Berger and Deuschel [18] have removed the requirement of ellipticity altogether in the i.i.d. case.

EXAMPLE 2.5 (Random environments with cut points). Under certain conditions, one can prove that a random walk in nonreversible random environments in $\mathbb{Z}^d$, d large enough, has cut points, and deduce from that a CLT and the existence of a stationary environment, hence our techniques apply. See [20] for the details.

Let us give one example which is not embedded in $\mathbb{Z}^d$, and in fact has unbounded degrees.

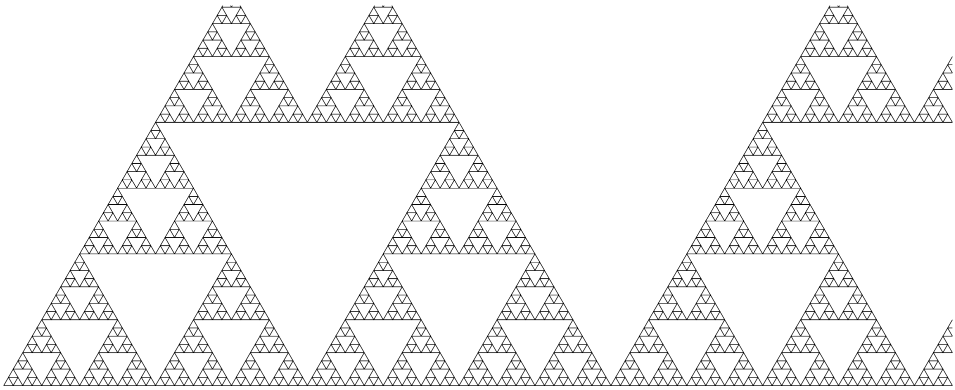

FIG. 1. *A portion of the graphical Sierpinski gasket.*

EXAMPLE 2.6 (Poisson point process). Examine a Poisson point process in $\mathbb{R}^d$. Add the point 0 (this is often called "the Palm process"), and let it be the root. Construct a graph by some process invariant under translations of $\mathbb{R}^d$. For example, connect any two points by an edge with weight which depends on their Euclidean distance [22] or construct the Delauney triangulation [40]. Give each configuration a "probability proportional to the total weight of 0." The resulting process is stationary and diffusive; see, for example, [22], Section 2.1 or [40], Lemma A.1, for stationarity—subdiffusivity can be deduced from [10], Section 2, or from the two previous papers. Hence our theorem applies.

The previous examples dealt with random walks which are diffusive. An interesting situation, which cannot hold in the case of groups, is environments with subdiffusive behavior. We give four examples of these.

EXAMPLE 2.7 (Graphical fractals). A graphical fractal is a graph which is constructed like one of the classical fractals (the Sierpinski gasket, e.g.), but inside out—bigger pieces of the graph are constructed from smaller pieces by connecting them in a repeated fashion; see [4] for precise definitions and main properties. See Figure 1 for an example, the graphical Sierpinski gasket. A graphical fractal always has an invariant measure and is always diffusive or subdiffusive, and in many examples is in fact subdiffusive; see, for example, [5]. Let us remark that a significant part in the remarkable work of Barlow and Bass on the Sierpinski carpet [7] has to do with the construction of a coupling. Therefore, a tool (like the one described in this section) that gives easy proofs that couplings exist should be useful.

EXAMPLE 2.8 (Critical Galton–Watson trees). The critical Galton–Watson tree with any offspring distribution conditioned to survive is stationary (see [50, 62, 64]) and subdiffusive. If the offspring distribution has finite variance, the diffusivity exponent $\frac{1}{3}$ was proved in [53]. Thus Theorem $3'$ applies in this case, and we get that it has no harmonic function of growth $o(\mathbf{d}(\rho, x)^{3/2})$.

This example is not so impressive since (as it is well known) this graph has infinitely many cut-edges between the root and infinity, and therefore the only harmonic functions (without any growth restrictions) are the constants. However, the cut-edges argument fails after even slight variations, while Theorem 3 is robust. Examples include taking a product of a Galton–Watson tree with a finite graph or with itself. The same remark applies to the next example.

EXAMPLE 2.9 (Infinite incipient cluster). Consider critical percolation on $\mathbb{Z}^d$ conditioned on the fact that the origin is connected to infinity [52]. Conditioning on this event, which has probability 0 (proved in $d = 2$ and high d and conjectured in the others), requires some care. Nevertheless, the object can be defined properly using a limit process. For example, one may take $p_c + \epsilon$ percolation, condition on $\vec{0}$ being in the cluster and then take a limit of the resulting measures as $\epsilon \to 0$. Since for each ϵ the measure is stationary (as usual after reweighting the configurations proportionally to the degree of $\vec{0}$), so will be their limit if it exists (or any subsequence limit in general). The limit is known to exist in two dimensions [48, 52] and in high dimensions [47, 76]. It was proved in [53, 57] that the random walk is subdiffusive on this cluster (in high dimension the diffusivity exponent is $\frac{1}{3}$, as on the tree). Since it is embedded in $\mathbb{Z}^d$, it grows no faster than polynomially and the results may be applied in this context.

EXAMPLE 2.10 (Graph limits and UIPQ). Let G_n be fixed or random finite graphs. Take ρ_n to be a random vertex in G_n, selected according to the stationary measure on G_n. Then the limit of (G_n, ρ_n), if it exists, is called the graph limit [15]. This limit is always stationary, [58], Section 1.3.

A particular case is provided by a uniformly chosen planar quadrangulation G_n with n faces. The graph limit is known as the uniform infinite planar quadrangulation. It is well known to be of polynomial growth [24]. In [12], it was proved to be subdiffusive with diffusivity exponent bounded from above by $\frac{1}{3}$. Thus there are no linear growth harmonic functions in this case either.

A remark on connectivity. We assumed throughout that the environment (P, V, ρ) is irreducible, that is, that for any $v, w \in V$ there is some n such that $P^n(v, w) > 0$. This assumption was only used once: we showed that a not-necessarily-irreducible stationary environment satisfies that every harmonic function h has $h(\rho) = h(X_1)$ almost surely, and concluded, using irreducibility, that h is constant. The assumption of irreducibility is of course necessary, as a disconnected graph always has bounded nonconstant harmonic functions, namely functions which are constant on each component, but with different values.

Nevertheless, in the nonreversible case, the assumption of irreducibility can be weakened slightly: we only need to assume that for every v and w there exist n, m and x such that $P^n(v, x) > 0$ and $P^m(w, x) > 0$. The proof is the same—since $h(\rho) = h(X_1)$ almost surely then this gives that $h(v) = h(x) = h(w)$ almost surely and h is constant. The following stationary graph provides a simple example. Take a 3-regular tree T. Choose a height function ℓ (i.e., a function such that each vertex has one neighbor with ℓ bigger by one, and two neighbors with ℓ smaller by one), and orient all edges "up," that is, in the direction of the larger ℓ. Of course, the random walk on the resulting graph is so degenerate it can hardly be called random, as each vertex has only one outgoing edge. But this is irrelevant at this point. This environment is not irreducible in the usual sense, but does satisfy the weaker assumption and hence our results apply (again, in this case it is simple to analyze the harmonic functions directly). Taking the graph product with $\mathbb{Z}$ will yield a slightly less trivial example.

3. Polynomial growth harmonic functions. In this section we prove Theorem 4. The proof boils down to the observation that macroscopic Poincaré inequality and volume growth estimates are sufficient. The strategy follows the lines of Shalom and Tao [73, 75], where a quantitative version of Gromov's theorem on groups of polynomial growth (any group of polynomial growth is virtually nilpotent) is proved. The proof is inspired by an elegant proof of this theorem due to Kleiner [55] utilizing spaces of harmonic functions with polynomial growth in a crucial way. We start with a very general inequality, called the reverse Poincaré inequality, which holds in any graph. For the sake of completeness, we prove it in our context.

PROPOSITION 12 (Reverse Poincaré inequality). *For any weighted graph (G, ν) and any function $h : G \to \mathbb{R}$ harmonic on a ball $\mathbf{B}_x(2n)$,*

$$\sum_{(y,z)\in E(\mathbf{B}_x(n))} \big(h(z) - h(y)\big)^2 \nu(y, z) \le \frac{4}{n^2} \sum_{y\in\mathbf{B}_x(2n)} h(y)^2\nu(y) \tag{14}$$

for every $x \in G$ and $n > 0$.

PROOF. For this proof, we denote the quantity $f(x)$ by f_x. Let $h : G \to \mathbb{R}$ be harmonic on $\mathbf{B}_x(2n)$, and let ϕ be a function such that $\phi_y = 1$ for $y \in \mathbf{B}_x(n)$, $\phi_y = 0$ for $y \notin \mathbf{B}_x(2n - 1)$ and $|\phi_y - \phi_z| \le 1/n$ for all $y \sim z$. For example,

$$\phi_y := \min\left(1, 2 - \frac{\mathbf{d}(y, x)}{n}\right) \qquad \text{for any } y \in \mathbf{B}_x(2n).$$

We have

$$\sum_{E(\mathbf{B}_x(n))} (h_y - h_z)^2\nu(y, z) = \sum_{E(\mathbf{B}_x(n))} \frac{1}{2}(\phi_y^2 + \phi_z^2)(h_y - h_z)^2\nu(y, z). \tag{15}$$

To make the calculation a little shorter we represent the sum on the right-hand side of (15) as a sum of $\frac{1}{2}\phi_y^2(h_y - h_z)^2$ over *directed edges*. Denote by E^* the set of directed edges in $B_x(2n)$, that is, both (y,z) and (z,y) appear in E^* and are different. For an edge $(y,z) \in E^*$, a straightforward (if a little lengthy) computation shows that $\phi_y^2(h_z - h_y)^2$ is equal to the quantity

$$\big(h_z\phi_z^2 - h_y\phi_y^2\big)(h_z - h_y) - h_z(\phi_z - \phi_y)^2(h_z - h_y) - 2h_z\phi_y(\phi_z - \phi_y)(h_z - h_y).$$

We start by dealing with the first term. Rearranging the sum [using the fact that $h\phi^2$ vanishes outside $\mathbf{B}_x(2n-1)$ to add the missing terms on the boundary] gives

$$\sum_{E^*}\big(h_z\phi_z^2 - h_y\phi_y^2\big)(h_z - h_y)\nu(y,z) = 2\sum_{y\in\mathbf{B}_x(2n-1)} h_y\phi_y^2\bigg(\sum_{z\sim y}(h_y - h_z)\nu(z,y)\bigg).$$

Since h is harmonic, this sum equals 0.

For the second term, since $|h_z(h_z - h_y)| \le \frac{3}{2}h_z^2 + \frac{1}{2}h_y^2$ and $|\phi_z - \phi_y| \le 1/n$, we have that each summand is bounded by $(3h_z^2 + h_y^2)/(2n^2)$. When summing over E^* we obtain

$$\bigg|\sum_{E^*} h_z(\phi_z - \phi_y)^2(h_z - h_y)\nu(y,z)\bigg| \le \frac{2}{n^2}\sum_{y\in\mathbf{B}_x(2n)} h_y^2\nu(y).$$

For the third term, note that

$$\big|h_z\phi_y(\phi_z - \phi_y)(h_z - h_y)\big| \le \tfrac{1}{4}(h_y - h_z)^2\phi_y^2 + h_z^2(\phi_z - \phi_y)^2. \tag{16}$$

So,

$$\begin{aligned}
&\sum_{E^*}\big|h_z\phi_y(\phi_z - \phi_y)(h_z - h_y)\big|\nu(z,y)\\
&\qquad\text{by (16)} \le \frac{1}{4}\sum_{E^*}(h_y - h_z)^2\phi_y^2\nu(y,z) + \sum_{E^*} h_z^2(\phi_z - \phi_y)^2\nu(y,z)\\
&\qquad\qquad\quad \le \frac{1}{4}\sum_{E^*}(h_y - h_z)^2\phi_y^2\nu(y,z) + \frac{1}{n^2}\sum_{\mathbf{B}_x(2n)} h_y^2\nu(y)
\end{aligned}$$

using the bound $|\phi_z - \phi_y| \le \frac{1}{n}$ for every $y \sim z$. Putting the bound on the different terms together leads to

$$\sum_{E^*}(h_y - h_z)^2\phi_y^2\nu(y,z) \le \frac{1}{2}\sum_{E^*}(h_y - h_z)^2\phi_y^2\nu(y,z) + \frac{4}{n^2}\sum_{\mathbf{B}_x(2n)} h_y^2\nu(y),$$

which gives

$$\sum_{E(\mathbf{B}_x(n))}(h_z - h_y)^2\nu(y,z) \le \frac{1}{2}\sum_{E^*}(h_z - h_y)^2\phi_y^2\nu(y,z) \le \frac{4}{n^2}\sum_{\mathbf{B}_x(2n)} h_y^2\nu(y). \qquad \square$$

LEMMA 13. *Let (G,ν,ρ) be a rooted graph satisfying the volume doubling condition $(VD)_G$. Then there exists $\mathbf{c}>0$ such that the following holds. For any $\lambda<\infty$, there exist M_λ and n_0 such that for all $n>n_0$, there is a covering of the ball $\mathbf{B}_\rho(\lambda n)$ by $k<M_\lambda$ balls $\mathbf{B}_{y_1}(n),\dots,\mathbf{B}_{y_k}(n)$ satisfying that every point $x\in\mathbf{B}_\rho(n)$ belongs to at most $\mathbf{c}$ balls $\mathbf{B}_{y_i}(2n)$.*

Furthermore, $\mathbf{c}$ depends only on the volume doubling constant $\mathbf{C}_{\mathrm{VD}}$, and M_λ depends only on λ and $\mathbf{C}_{\mathrm{VD}}$.

We call a covering with this property *proper*.

PROOF. Let λ and G be as above. Let n be large enough so that $(VD)_G$ holds for 2λ and $n/2$. Given this, we can choose a maximal family of disjoint balls $\mathbf{B}_{y_1}(n/2),\dots,\mathbf{B}_{y_k}(n/2)$ with $y_j\in\mathbf{B}_\rho(\lambda n)$ for all j:

- Since the family $\{\mathbf{B}_{y_j}(n/2)\}$ is maximal, every vertex in $\mathbf{B}_\rho(\lambda n)$ must be within distance $\le n$ from one of the y_j, so $\mathbf{B}_\rho(\lambda n)$ is covered by $\mathbf{B}_{y_1}(n),\dots,\mathbf{B}_{y_k}(n)$.
- For any $x\in\mathbf{B}_\rho(\lambda n)$, if $x\in\mathbf{B}_{y_j}(2n)$, then $\mathbf{B}_{y_j}(n/2)\subset\mathbf{B}_x(3n)$. Using volume doubling we see that $\nu(\mathbf{B}_x(3n))\le\mathbf{C}_{\mathrm{VD}}^4\nu(\mathbf{B}_{y_j}(n/2))$, hence (since these balls are disjoint) we have that the number of y_j such that $x\in\mathbf{B}_{y_j}(2n)$ is at most $\mathbf{C}_{\mathrm{VD}}^4$.
- Using the volume doubling similarly, we get that $\nu(\mathbf{B}_\rho((\lambda+1)n))\le\mathbf{C}\nu(\mathbf{B}_{y_j}(n/2))$ for any j (the constant is $\mathbf{C}_{\mathrm{VD}}^{\lceil\log_2(\lambda+1)\rceil+2}$). Since these balls are all disjoint and fully contained in $\mathbf{B}_\rho((\lambda+1)n)$, we get

$$\begin{aligned}k\min_j\nu\big(\mathbf{B}_{y_j}(n/2)\big)&\le\nu\Big(\bigcup_j\mathbf{B}_{y_j}(n/2)\Big)\le\nu\big(\mathbf{B}_\rho\big((\lambda+1)n\big)\big)\\&\le\mathbf{C}\min_j\nu\big(\mathbf{B}_{y_j}(n/2)\big),\end{aligned}$$

  and we get that the number of balls k is bounded by the same $\mathbf{C}$. $\square$

LEMMA 14. *Let (G,ν,ρ) be a rooted graph satisfying $(P)_G$. Then there exists a $\mathbf{c}>0$ such that for every $\varepsilon>0$ and n large enough, and for every proper covering of $\mathbf{B}_\rho(n)$ by balls of radius εn, if $h:G\to\mathbb{R}$ is harmonic and has 0 mean on all the balls of the covering, then*

$$\sum_{z\in\mathbf{B}_\rho(n)}h(z)^2\nu(z)\le\mathbf{c}\varepsilon^2\sum_{z\in\mathbf{B}_\rho(4n)}h(z)^2\nu(z).\tag{17}$$

Further, $\mathbf{c}$ depends only on $\mathbf{C}_{\mathrm{P}}$, the constant in the Poincaré inequality and on the constants in the definition of a proper cover.

PROOF. Fix n large enough so that $(P)_G$ holds true for $\lambda=1/\varepsilon$ and εn, that is, such that for every $x\in\mathbf{B}_\rho(n)$ and f a map on $\mathbf{B}_\rho(n)$,

$$\sum_{y\in\mathbf{B}_x(\varepsilon n)}\big(f(y)-\overline{f}_{\mathbf{B}_x(\varepsilon n)}\big)^2\nu(y)\le\mathbf{C}_{\mathrm{P}}(\varepsilon n)^2\sum_{(y,z)\in E(\mathbf{B}_x(2\varepsilon n))}|f(y)-f(z)|^2\nu(y,z).$$

Let $h : G \to \mathbb{R}$ be the harmonic function and $\mathbf{B}_{y_1}(\varepsilon n), \dots, \mathbf{B}_{y_k}(\varepsilon n)$ be the proper covering of $\mathbf{B}_\rho(n)$ from the statement of the lemma. The hypothesis asserts that $\overline{h}_{\mathbf{B}_{y_i}(\varepsilon n)} = 0$ for every i, so that Poincaré inequality implies

$$\begin{aligned}\sum_{\mathbf{B}_{y_i}(\varepsilon n)} h^2(z)\nu(z) &= \sum_{\mathbf{B}_{y_i}(\varepsilon n)} \bigl(h(z) - \overline{h}_{\mathbf{B}_{y_i}(\varepsilon n)}\bigr)^2 \nu(z) \\ &\le \mathbf{C}_{\mathrm{P}}\varepsilon^2 n^2 \sum_{E(\mathbf{B}_{y_i}(2\varepsilon n))} \bigl(h(z) - h(t)\bigr)^2 \nu(z, t).\end{aligned}$$

Since the $\mathbf{B}_{y_i}(2\varepsilon n)$ have uniformly bounded overlap (each point belong to at most $\mathbf{c}$ balls), and since $\mathbf{B}_{y_i}(2\varepsilon n) \subset \mathbf{B}_\rho(2n)$, we find

$$\begin{aligned}&\sum_{\mathbf{B}_\rho(n)} h^2(z)\nu(z) \\ &\quad\le \mathbf{c}\mathbf{C}_{\mathrm{P}}\varepsilon^2 n^2 \sum_{E(\mathbf{B}_\rho(2n))} \bigl(h(z) - h(t)\bigr)^2 \nu(z, t).\end{aligned} \tag{18}$$

Using the reverse Poincaré inequality (Proposition 12) for the larger ball, we conclude

$$\sum_{\mathbf{B}_\rho(n)} h^2(z)\nu(z) \le 4\mathbf{c}\mathbf{C}_{\mathrm{P}}\varepsilon^2 \sum_{\mathbf{B}_\rho(4n)} h^2(z)\nu(z), \tag{19}$$

which implies the claim with the constant in the statement of the lemma being $4\mathbf{c}\mathbf{C}_{\mathrm{P}}$. $\square$

PROOF OF THEOREM 4. We aim to prove that the space of harmonic functions u such that $|u(x)| \le \mathbf{C}\mathbf{d}(\rho, x)^k$ for every $x \in G$ is finite dimensional. Consider a rooted graph G satisfying $(VD)_G$ and $(P)_G$. Let $\mathbf{c}$ be large enough so that the two previous lemmas hold true. On the set of harmonic functions on $\mathbf{B}_\rho(n)$, a scalar product between two functions can be defined by

$$\langle f, g\rangle_n = \sum_{\mathbf{B}_\rho(n)} f(x)g(x)\nu(x).$$

Consider d harmonic functions $u_1, \dots, u_d$ on G and set $V = \mathrm{span}(u_1, \dots, u_d)$. Our goal is to compare $\langle\cdot,\cdot\rangle_n$ and $\langle\cdot,\cdot\rangle_{4n}$ for these functions.

Let $\varepsilon > 0$ be some parameter to be fixed later. For n large enough, there exists a proper covering $\mathbf{B}_{y_1}(\varepsilon n), \dots, \mathbf{B}_{y_M}(\varepsilon n)$ of $\mathbf{B}_\rho(n)$ by $M = M_{1/\varepsilon}$ balls. Therefore there is a codimension $d - M$ vector space $V_0 \subset V$ of harmonic functions with mean 0 on each of the balls $\mathbf{B}_{y_i}(\varepsilon n)$. Let $v_1, \dots, v_d$ be an orthogonal basis of V for $\langle\cdot,\cdot\rangle_{4n}$ such that $v_1, \dots, v_{d-M}$ is a basis of V_0. Examine the Gram matrix of

$\{v_i\}$, that is, the $d \times d$ matrix whose entries are $\langle v_i, v_j\rangle_n$. Then

$$\begin{aligned}\det\bigl[\{\langle v_i, v_j\rangle_n\}_{i,j}\bigr] &\le \prod_{i=1}^{d}\langle v_i, v_i\rangle_n \\ &\le \prod_{1}^{d-M} \mathbf{c}\varepsilon^2\langle v_i, v_i\rangle_{4n} \prod_{i=d-M+1}^{d} \langle v_i, v_i\rangle_{4n} \\ &= (\mathbf{c}\varepsilon^2)^{d-M} \det\bigl[\{\langle v_i, v_j\rangle_{4n}\}_{i,j}\bigr],\end{aligned}$$

where in the first line we have used Hadamard's inequality, in the second Lemma 14 and in the last, the fact that (v_i) is orthogonal for $\langle\cdot,\cdot\rangle_{4n}$. Now, the ratio of two Gram determinants is preserved by linear operations on vectors, so we can return from the basis $\{v_i\}$ (which was specific to n) to our "original" basis $\{u_i\}$. We get

$$\det\bigl[\{\langle u_i, u_j\rangle_n\}_{i,j}\bigr] \le (\mathbf{c}\varepsilon^2)^{d-M} \det\bigl[\{\langle u_i, u_j\rangle_{4n}\}_{i,j}\bigr].$$

Iterating the reasoning, we find for every $r > 0$

$$\det\bigl[\{\langle u_i, u_j\rangle_n\}_{i,j}\bigr] \le \bigl[(\mathbf{c}\varepsilon^2)^{d-M}\bigr]^r \det\bigl[\{\langle u_i, u_j\rangle_{4^r n}\}_{i,j}\bigr].$$

The growth of our harmonic functions ensures that every entry of the matrix is smaller than $[\mathbf{C}(4^r n)^k]^2 \nu(\mathbf{B}_\rho(4^r n)) \le \mathbf{C}(4^r n)^{2k+\mathbf{c}}$. Hence we can write

$$\det\bigl[\{\langle u_i, u_j\rangle_n\}_{i,j}\bigr] \le d!n^{(2k+\mathbf{c})d}\mathbf{C}^d\bigl((\mathbf{c}_1\varepsilon^2)^{d-M}4^{(2k+\mathbf{c}_2)d}\bigr)^r.$$

We now fix ε^2 to be $4^{-4k-2\mathbf{c}_2}/\mathbf{c}_1$. If $d > 2M$, this implies

$$(\mathbf{c}_1\varepsilon^2)^{d-M}4^{(2k+\mathbf{c}_2)d} = 4^{(2k+\mathbf{c}_2)(2M-d)} < 1,$$

and the right-hand side would converges to 0. We deduce that $\det[\{\langle u_i, u_j\rangle_n\}_{i,j}] = 0$, and that the u_i restricted to the ball of radius n form a dependent family. Since this is true for every n large enough, we easily deduce that (u_i) is a linearly dependent family. The result holds for any family of d harmonic functions with growth bounded by $\mathbf{Cd}(\cdot,\rho)^k$. It implies that the dimension of the vector space of harmonic functions with such growth is smaller or equal to $2M$. □

EXAMPLE 3.1 (Infinite cluster of percolation). The infinite cluster of percolation satisfies $(VD)_\omega$ and $(P)_\omega$ almost surely [6]. Therefore, spaces of harmonic functions with prescribed polynomial growth are finite dimensional.

EXAMPLE 3.2 (Random conductance). Random conductances with uniform elliptic conditions also satisfy $(VD)_\omega$ and $(P)_\omega$ deterministically. Therefore, spaces of harmonic functions with prescribed polynomial growth are finite dimensional.

EXAMPLE 3.3 (Wedges). Let f be some *slowly varying* function from $[0,\infty)\to[0,\infty)$. Define the wedge with respect to d and f to be

$$W := \{x \in \mathbb{Z}^d : |x_d| \le f(|x_1| + \cdots + |x_{d-1}|)\}.$$

Then it is well-known and not difficult to see that W (with the graph structure inherited from $\mathbb{Z}^d$) satisfies volume doubling and Poincaré inequality. Under some weak conditions on f and d (which we will not detail here, as that would take us too off-topic) so would percolation on W. Hence both W and supercritical percolation on it have a finite dimensional space of harmonic functions.

4. Linearly growing harmonic functions on the infinite cluster of percolation. In this section, we fix $d \ge 2$ and $p > p_c(d)$. As before, we denote the infinite cluster of percolation by ω, and we draw it in $\mathbb{R}^d$ in such a way that ρ coincides with the origin. The graph ω can be thought of as an approximation of $\mathbb{Z}^d$. In particular, macroscopic properties of the cluster are the same as those of $\mathbb{R}^d$. For instance, the random walk satisfies an invariance principle $(CLT)_\omega$ [16, 66, 74]: define

$$\tilde{B}_n(t) := \frac{1}{\sqrt{n}}(X_{tn}),$$

where for noninteger tn we define X_{tn} as the linear interpolation between $X_{\lfloor tn\rfloor}$ and $X_{\lceil tn\rceil}$; that is, $X_{tn} = X_{\lfloor tn\rfloor}(tn - \lfloor tn\rfloor) + X_{\lceil tn\rceil}(\lceil tn\rceil - tn)$. There exists $\sigma(d)$ such that the law of $(\tilde{B}_n(t), 0 < t < \infty)$ converges weakly to the law of a Brownian motion with variance $\sigma(d)$ as $n \to \infty$. The main step in the proof in all three papers [16, 66, 74] is the construction of a d-dimensional space of linearly growing harmonic functions $\{f_v\}_{v\in\mathbb{R}^d}$ such that f_v has slope v, that is, $f_v(x) = \langle v, x\rangle + o(|x|)$. Let us state this as a theorem.

THEOREM 15 ([16, 66, 74]). *Let $d \ge 2$, and $p > p_c(d)$. Let ω be the infinite cluster of percolation on $\mathbb{Z}^d$ with parameter p. Then, there exists $\chi : \omega \to \mathbb{R}^d$ such that $x \mapsto x + \chi(x)$ is harmonic on ω, and*

$$\lim_{n\to\infty} \frac{1}{n} \sup_{x\in\mathbf{B}_\rho(n)} |\chi(x)| = 0 \qquad a.s. \tag{20}$$

This (random) function is called the corrector.

With the constant functions, we get a $(d+1)$-dimensional space of harmonic functions with (sub-)linear growth. Our aim in this section is to prove Theorem 5 from the introduction, namely that there are no other harmonic functions of linear growth.

Proof outline. Let h be a harmonic function with linear growth. Define $h_n : \mathbb{R}^d \to \mathbb{R}$ such that $h_n(x) = h(nx)/n$. In order to prove Theorem 5, we first show that (h_n) forms a precompact family (one can say that h has a scaling limit). The second step is to identify the possible limits. For this, we use the average property at the discrete level and the invariance principle to prove that limits are harmonic on $\mathbb{R}^d$. If the space of limits is at most d-dimensional, one can then use the absence of nonconstant sublinear harmonic functions to show that the space of harmonic functions with linear growth is $(d+1)$-dimensional.

Properties of the supercritical cluster. Recall that the infinite supercritical cluster of percolation ω can be seen as a stationary random graph with polynomial growth. It is well-known that the system is ergodic with respect to the shift by X_1, see, for example, [16], Theorem 3.1. Typical balls have the same growth as in the ambient space $\mathbb{Z}^d$ in the following sense: there exists constants c and C such that for any finite λ, any $n > n_0(\omega)$ sufficiently large and any $x \in \mathbf{B}_\rho(\lambda n)$

$$cn^d \le \nu\bigl(\mathbf{B}_x(n)\bigr) \le Cn^d. \tag{21}$$

Clearly, (21) implies volume doubling $(VD)_\omega$. Moreover, the graph satisfies $(P)_\omega$ almost surely. Both properties were proved by Barlow [6]. Actually, Barlow proved quantitatively stronger versions of (21) and $(P)_\omega$: he obtained the volume growth estimates and the Poincaré inequality for every ball of radius larger than $C \log n$ in $\mathbf{B}_\rho(n)$. These improved versions allow to prove Harnack inequalities and Gaussian estimates (1) on the heat kernel. In [6], these results are stated for continuous time random walk, but they hold also for simple random walk, as was explained in [9], Section 2. We do not need the full force of Gaussian estimates here—in particular we do not need far off-diagonal lower bounds which are particularly difficult—so let us make a list of corollaries from these Gaussian estimates which we will use.

COROLLARY 16. *For every $\lambda < \infty$, every $n \in \mathbb{N}, n > n_0(\omega)$ sufficiently large and every $x \in \mathbf{B}_\rho(\lambda n)$,*

$$\mathbf{P}_x(X_{n^2} = y) \le Cn^{-d} \exp\bigl[-C\mathbf{d}(x,y)^2/n^2\bigr] \qquad \textit{for any } y. \tag{22}$$

In both [6, 9] the results are formulated with $|x-y|$ instead of $\mathbf{d}(x,y)$, but by the results of Antal and Pisztora [2], this is the same. This immediately implies

$$\mathbf{E}_x\bigl[\mathbf{d}(X_{n^2}, x)^2\bigr] \le \mathbf{c}_3 n^2 \tag{23}$$

for some constant $\mathbf{c}_3$ depending on the environment.

The lower bound has some periodicity requirements since $\mathbf{P}_x(X_t = y) = 0$ whenever $t + \sum(x_i - y_i)$ is odd.

COROLLARY 17. *For every $\lambda < \infty$, every $n \in \mathbb{N}$ sufficiently large and every $x \in B(\lambda n)$,*

$$\mathbf{P}_x(X_{n^2} = y) \ge Cn^{-d}$$

for any $y \in \mathbf{B}_x(n)$ such that $n^2 + \sum(x_i - y_i)$ is even.

In the proof we will need in a few places *space ergodicity*. We start with a lemma that encapsulates this for us.

LEMMA 18. *Let $f(x,y,\omega)$ be some positive translation-invariant variable, that is, $f(x+s,y+s,\omega+s)=f(x,y,\omega)$, with $M:=\mathbb{E}f(0,X_1,\omega)<\infty$. Then for every $\lambda>0$ and for almost every environment ω, there exists n_0 such that for all $n>n_0$ and for any $a\in\mathbf{B}_\rho(\lambda n)$,*

$$\sum_{(x,y)\in E(\mathbf{B}_a(n))} f(x,y,\omega)\nu(x,y)\le C\cdot M\cdot n^d.$$

PROOF. We wish to apply the ergodic theorem for $\mathbb{Z}^d$ actions (see, e.g., [72], Theorem 2.6, page 40). We let the probability space be $\{0,1\}^{\mathbb{Z}^d}$ with the product measure and let the probability preserving maps T_i from the statement of the theorem to be translations of coordinates. Clearly the T_i commute. Further, each T_i has only trivial invariant subsets—indeed, if A is invariant under some T_i, then we can ε-approximate A by an event B depending only on finitely many coordinates and then apply T_i^n for n sufficiently large so that B and T_i^nB are independent. We get that $|\mathbb{P}(A)-\mathbb{P}(A)^2|\le 3\varepsilon$. Since ε is arbitrary, $\mathbb{P}(A)$ must be equal to 0 or 1.

Fix v to be one of the d vectors of the standard basis of $\mathbb{Z}^d$ and define a function $F:\{0,1\}^{\mathbb{Z}^d}\to\mathbb{R}$ by

$$F(\xi)=\begin{cases} f\bigl(\rho,v,\omega(\xi)\bigr), & \rho,v\in\omega(\xi),\\ 0, & \text{otherwise,}\end{cases}$$

where $\omega(\xi)$ is the infinite cluster (possibly equal to the empty set). The ergodic theorem implies

$$\lim_{n\to\infty}\frac{1}{(2n)^d}\sum_{-n-1\le i_1,\dots,i_d\le n} F\bigl(T_1^{i_1}\cdots T_d^{i_d}\xi\bigr)=\mathbb{E}(F)\qquad\text{almost surely}$$

(the theorem in [72] is formulated for $\mathbb{Z}_+^d$ actions, but the two-sided version above follows from the one-sided version by applying the one-sided result 2^d times, for each choice of $T_1^{\pm1},\dots,T_d^{\pm1}$, and combining the results). Summing the above over all v in the standard basis element v enables us to go from F to f and to obtain

$$\lim_{n\to\infty}\frac{1}{|E(Q_n)|}\sum_{(x,y)\in E(Q_n))} f(x,y,\omega)\le CM,$$

where $Q_n=[-n-1,n]^d$ and C is some universal constant.

Let us now generalize this to cubes centered around an arbitrary $a\in B_\rho(\lambda n)$. Fix some N, C' sufficiently large such that

$$\mathbb{P}\biggl(\exists n>N,\sum_{(x,y)\in E(Q_n)} f(x,y,\omega)>C'Mn^d\biggr)<\mu,$$

where μ will be defined in the next paragraph (as a function of λ). Define ω to be *good* if the event involved in the previous displayed equation does not happen, and define $b \in \mathbb{Z}^d$ to be *good* if translating ω by $-b$ gives a good configuration. Using ergodicity once again, there exists almost surely $n_0 = n_0(\omega) < \infty$ such that for any $n \geq n_0$,

$$\left|\{\text{good } b \in Q_n\}\right| > (1-2\mu)|Q_n|.$$

In particular, the number of sites in $\mathbf{B}_\rho(\lambda n) \subset Q_{\lambda n}$ which are not good is less than $2\mu|Q_{\lambda n}|$ for $n \geq n_0$. If $\mu = \mu(\lambda)$ is chosen sufficiently small and $n \geq n_0$, then there cannot be any good-free ball of radius n in $\mathbf{B}_\rho(\lambda n)$. Hence, for any $a \in \mathbf{B}_\rho(\lambda n)$, there is a cube $b + Q_{2n} \supset \mathbf{B}_a(n)$ centered around a good point b. This implies that for $n \geq \max\{n_0, N\}$,

$$\sum_{(x,y)\in E(\mathbf{B}_a(n))} f(x,y,\omega) \leq \sum_{(x,y)\in E(b+Q_{2n})} f(x,y,\omega) \leq C'M \cdot (2n)^d.$$

(The assumption $n \geq N$ enables us to use the fact that b is good.) Adding the terms $\nu(x,y)$ only changes the constant. □

Recall the h_n from the proof sketch on page 2356. There we defined $h_n(x) = h(nx)/n$ which is a priori only defined on the contracted infinite cluster. For simplicity let us extend it to all $\mathbb{R}^d$, for example, by extending h to $\mathbb{Z}^d$ by taking the value at the closest point of the infinite cluster, and then to $\mathbb{R}^d$ by defining $h(x) = \sum_{y\in\mathbb{Z}^d} h(y)\phi_y(x)$ where ϕ_y is some partition of unity such that $\operatorname{supp}\phi_y \subset y + [-\frac{2}{3}, \frac{2}{3}]^d$. Once h is extended to all $\mathbb{R}^d$, so is h_n.

PROPOSITION 19. *For almost every environment ω, any harmonic function h on ω with linear growth satisfies that for every compact $K \subset \mathbb{R}^d$, the sequence $(h_n)|_K$ is uniformly bounded and equicontinuous.*

PROOF. Fix a harmonic map h with (at most) linear growth on an environment ω. There exists $A > 0$ such that $|h(x)| \leq A|x|$. We only need to prove equicontinuity on the ball, as this property passes to subsets. To do so, we prove that for any $\eta > 0$, there exists $\delta > 0$ such that $(h(a) - h(b))^2 \leq \eta n^2$ for any two points $a, b \in \mathbf{B}_\rho(n)$ at distance δn of each other, when n is large enough (it is easy to see that our procedure for extending h to $\mathbb{R}^d$ allows to prove the needed estimates only for a and b in ω). For this reason, we will always assume that n is large enough so that the Poincaré inequality $(P)_\omega$ and the d-dimensional volume growth (21) hold true for $\lambda = 2$.

Let $\delta, \varepsilon > 0$ to be fixed later (think of $\varepsilon \ll \delta$) and $a, b \in \mathbf{B}_\rho(n)$ with $\mathbf{d}(a,b) \leq \delta n$. Let B be some ball of radius $2\delta n$ containing both $\mathbf{B}_a(\delta n)$ and $\mathbf{B}_b(\delta n)$—for example, around the middle point of $[ab]$. Let $\overline{h}$ be the average $\frac{1}{\nu(B)}\sum_{x\in B} h(x)\nu(x)$. Since $|h(a) - h(b)| \leq |h(a) - \overline{h}| + |h(b) - \overline{h}|$, it is enough to estimate these terms. Let us focus on estimating $|h(a) - \overline{h}|$ (the other term is symmetric).

Set $\mathscr{E}$ to be the event that $|X_{(\varepsilon n)^2} - a| \geq \delta n$. Note that

$$\mathbf{P}_a(\mathscr{E}) \leq \frac{\mathbf{E}_a(|X_{(\varepsilon n)^2} - a|^2)}{(\delta n)^2} \overset{(23)}{\leq} \frac{\mathbf{c}_3(\varepsilon n)^2}{(\delta n)^2} = \mathbf{c}_3(\varepsilon/\delta)^2,$$

where the Markov inequality was used in the first inequality and the quenched diffusive behavior (23) in the second.

Now, we have

$$\begin{aligned}|h(a) - \overline{h}|^2 &\leq \big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\big]\big)^2 \\ &\leq 2\big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}}\big]\big)^2 + 2\big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}^c}\big]\big)^2.\end{aligned}$$

We first deal with the first term on the right:

$$\begin{aligned}\big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}}\big]\big)^2 &\leq \mathbf{E}_a\big[\big(|h(X_{(\varepsilon n)^2})| + |\overline{h}|\big)^2\big] \cdot \mathbf{P}_a(\mathscr{E}) \\ &\leq \big(2\mathbf{E}_a\big[h(X_{(\varepsilon n)^2})^2\big] + 2\overline{h}^2\big) \cdot \mathbf{P}_a(\mathscr{E}) \\ \text{since } h(x) \leq A|x| \qquad &\leq 2A\big(\mathbf{E}_a\big[|X_{(\varepsilon n)^2}|^2\big] + (1+2\delta)^2 n^2\big) \cdot \mathbf{P}_a(\mathscr{E}) \\ \text{by (23)} \qquad &\leq 2A\big(\mathbf{c}_3\big(1+\varepsilon^2\big) + (1+2\delta)^2\big)n^2 \cdot \mathbf{c}_3\frac{\varepsilon^2}{\delta^2} = \mathbf{c}_5 n^2 \frac{\varepsilon^2}{\delta^2},\end{aligned}$$

where Cauchy–Schwarz was used in the first inequality.

For the second term, the heat kernel upper bound (22) shows that $\mathbf{P}_a(X_{(\varepsilon n)^2} = x) \leq C_6/(\varepsilon n)^d$ for any $x \in \mathbf{B}_\rho(n)$ and n large enough. Therefore,

$$\begin{aligned}\big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}^c}\big]\big)^2 &\leq \mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|^2\mathbf{1}_{\mathscr{E}^c}\big] \\ &\leq \frac{C_6}{(\varepsilon n)^d} \sum_{x \in \mathbf{B}_a(\delta n)} |h(x) - \overline{h}|^2 \nu(x) \\ &\leq \frac{C_6}{(\varepsilon n)^d} \sum_{x \in B} |h(x) - \overline{h}|^2 \nu(x).\end{aligned}$$

Poincaré's inequality implies

$$\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}^c}\big]^2 \leq \frac{\mathbf{C}_{\mathrm{P}}C_6}{(\varepsilon n)^d}(2\delta n)^2 \sum_{(x,y) \in E(B')} |h(x) - h(y)|^2 \nu(x,y),$$

where B' is the ball with same center as B and radius $4\delta n$.

Now, the quantity Δ_n introduced in (9) controls the gradient of a harmonic function. Indeed, the same reasoning as the one used to derive (11) implies that

$$|h(x) - h(y)|^2 \leq \big(\mathbf{E}_x\big[|h(X_n)|^2\big] + \mathbf{E}_y\big[|h(X_{n-1})|^2\big]\big)\Delta_n(x,y)^2$$

for every n. Using the bound $|h(z)| \leq A|z|$, diffusivity and taking the liminf, we obtain

$$|h(x) - h(y)|^2 \leq \mathbf{c}_7 \liminf_{n \to \infty} n\Delta_n(x,y)^2,$$

where $\mathbf{c}_7$ does not depend on the points x, y (though it does depend on h through A). Denote this lim inf by $\Delta_\infty(x, y)^2$. We get

$$\big(\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}^c}\big]\big)^2 \le \frac{\delta^2}{\varepsilon^d}\frac{\mathbf{c}_8}{n^{d-2}} \sum_{(x,y)\in E(B)} \Delta_\infty(x, y)^2 \nu(x, y).$$

We next note that $\mathbb{E}\Delta_\infty(\rho, X_1)^2 < \infty$. Indeed, the infinite cluster of percolation is a subgraph of $\mathbb{Z}^d$, it has uniform polynomial growth and $H_n \le C_1 \log n$ for every n. Theorem 8 implies that $\mathbb{E}[\Delta_n(\rho, X_1)^2] \le C_2/n$ for an infinite number of n. Using Fatou's lemma, we obtain that $\mathbb{E}[\Delta_\infty(\rho, X_1)^2] < \infty$. Thus we may use Lemma 18 for the function $f = \Delta_\infty^2$ and get (with the fact that B' has radius $4\delta n$),

$$\mathbf{E}_a\big[|h(X_{(\varepsilon n)^2}) - \overline{h}|\mathbf{1}_{\mathscr{E}^c}\big]^2 \le \mathbf{c}_9 \frac{\delta^{d+2}}{\varepsilon^d} n^2.$$

Putting together the estimates for the two terms, we obtain

$$\big(h(a) - h(b)\big)^2 \le n^2\left(\mathbf{c}_3 \frac{\varepsilon^2}{\delta^2} + \mathbf{c}_9 \frac{\delta^{d+2}}{\varepsilon^d}\right),$$

which implies the claim provided $\delta = \varepsilon^{(d+1)/(d+2)}$. $\square$

LEMMA 20. *For almost every environment ω, for any harmonic function h on ω with linear growth, any subsequential limit of h_n is linear.*

PROOF. Let n_k be a sequence such that h_{n_k} converges uniformly on compact subsets of $\mathbb{R}^d$, and denote the limit by ℓ. Let now B_t be a Brownian motion with variance $\sigma(d)$, where $\sigma(d)$ comes from the invariance principle for random walk on ω, see page 2356. Our first goal is to derive a mean-value property anchored at the origin. Namely, we wish to prove that

$$\mathbf{E}_0\big[\ell(B_t)\big] = \ell(0) \qquad \text{for any } t > 0. \tag{24}$$

To see (24) note that h is harmonic and hence $\mathbf{E}_\rho[h(X_t)] = h(\rho)$ or equivalently

$$\mathbf{E}_0\big[h_n(X_{n^2 t}/n)\big] = h_n(0).$$

The central limit theorem (Theorem 15) allows to control $h(X_t)$ in a ball of radius $\approx \sqrt{t}$. Namely, because $X_{n^2 t}/n$ converges weakly to B_t, and because ℓ is continuous (as a locally uniform limit of the h_{n_k}), for any $K > 0$,

$$\big|\mathbf{E}_0\big[\ell(X_{n^2 t}/n) \cdot \mathbf{1}_{\{|X_{n^2 t}/n| < K\}}\big] - \mathbf{E}_0\big[\ell(B_t) \cdot \mathbf{1}_{\{|B_t| < K\}}\big]\big| \to 0,$$

where the convergence is as $n \to \infty$. The Gaussian bounds (22) and the linear bounds on h_n and ℓ allow to control $h(X_t)$ outside that ball,

$$\big|\mathbf{E}_0\big[h_n(X_{n^2 t}/n) \cdot \mathbf{1}_{\{|X_{n^2 t}/n| \ge K\}}\big]\big| \le \varepsilon(K) \qquad \text{for any } n \text{ sufficiently large},$$

where $\varepsilon(K) \to 0$ as $K \to \infty$. A similar estimate holds for $\ell(B_t)$. This shows (24).

We now extend (24) from 0 to all points u using Lemma 18. Let us recall that weak convergence is metrizable. For example, it is equivalent to convergence in the Lévy–Prokhorov distance metric (e.g., [38], Section 11.3), especially Theorem 11.3.3. We will not need any property of the Lévy–Prokhorov metric except that it is equivalent to weak convergence.

Now fix t, and fix also some ε and some n_0. Consider a vertex x in the cluster to be good if the Gaussian estimates (22) hold for all $n > n_0$ and if the Lévy–Prokhorov distance between $X_{n^2 t}/n$ (started from x) and B_t (started from x/n) is smaller than ε, again for all $n > n_0$. If n_0 is sufficiently large (depending on t and ε), the probability of x being good will be larger than $1 - \varepsilon$. For the Gaussian estimates this follows directly from (22) while for the Lévy–Prokhorov distance this follows from the equivalence of Lévy–Prokhorov convergence and weak convergence. Fix therefore n_0 to satisfy this property.

Now use Lemma 18 with the function f being $f(x, y) = \mathbf{1}_{\{x \text{ is bad}\}}$ (the y variable is simply ignored) and with some arbitrary λ. We get that for sufficiently large n, the number of bad x in $\mathbf{B}_\rho(\lambda n)$ is bounded by $C(\lambda) n^d \mathbb{P}(0 \text{ is bad}) \le C(\lambda)\varepsilon n^d$. Define

$$B_n := \{u \in \mathbb{R}^d : |u| \le \lambda, un \text{ is bad}\}$$

(where as usual we in fact take the point of the infinite cluster closest to un and check whether it is bad). Since the measure of B_n is smaller than $C\varepsilon$, we see that, except for a set of measure smaller than $C\varepsilon$, every $u \in \mathbb{R}^d$ with $|u| \le \lambda$ satisfies that un_k is good for infinitely many n_k (it does not matter that n_k itself depends on the environment here). But ε (both for the error and for the measure of the bad set) was arbitrary. Taking $\varepsilon \to 0$ and then $\lambda \to \infty$ we see that for almost every $u \in \mathbb{R}^d$ there is a sequence $n'_k = n'_k(u)$ (a subsequence of n_k) such that:

1. The Gaussian estimates hold for $X_{n'_k}$ started from un'_k.
2. The Lévy–Prokhorov distance between $X_{(n'_k)^2 t}/n'_k$ started from un'_k and Brownian motion started from u goes to zero.

Using again the equivalence of Lévy–Prokhorov convergence and weak convergence we get that random walk started from un'_k converges to Brownian motion started from u. We can now repeat the argument that led to (24) literally and get

$$\mathbf{E}_u[\ell(B_t)] = \ell(u)$$

for almost every u. Since ℓ is continuous, this in fact holds everywhere. Since t was arbitrary, $\ell(B_t)$ is a continuous martingale, from any starting point.

The lemma is now proved. Using the strong Markov property we get that $\ell(u)$ is equal to its average over a sphere of arbitrary radius around u, in other words, we have established the mean-value property hence ℓ is (continuously) harmonic and has a linear bound. It is well known that harmonic functions with at most linear growth on $\mathbb{R}^d$ are the affine maps (take the partial derivative along one direction, it is a bounded harmonic map on $\mathbb{R}^d$, and thus a constant map). $\square$

PROOF OF THEOREM 5. Let $d \geq 2$. The constant functions on ω are obviously harmonic. The projections of $x + \chi(x)$ where χ is the corrector (see Theorem 15) on each coordinate provide us with d linearly independent functions. These functions have linear growth. Therefore, the space of linear growth harmonic functions is at least $(d+1)$-dimensional. Thus we need to show that any harmonic function of linear growth is of the form $h(x + \chi(x))$.

The first step is to apply Theorem 8 and get that $\mathbb{E}(\Delta_n(\rho, X_1)^2) \leq 2(H_n - H_{n-1})$ and in particular is $\leq C/n$ on a subsequence. Hence, by Fatou's lemma, there is a random subsequence n_k such that $\mathbf{E}_\rho[\Delta_{n_k}(\rho, X_1)^2] \leq C/n_k$.

Now, let h be a harmonic function on ω with (at most) linear growth and with $h(0) = 0$. Proposition 19 allows us to extract a sequence m_k such that (h_{m_k}) converges uniformly on any compact subset of $\mathbb{R}^d$ to a continuous function $\tilde{h}$, and further one may take m_k to be a subsequence of any given sequence, so we may assume m_k is a subsequence of $\lfloor n_k^{1/2} \rfloor$. By Lemma 20 $\tilde{h}$ is linear. We get that, $f(x) := h(x) - \tilde{h}(x + \chi(x))$ is a harmonic function on ω with the following additional property: for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that for all $k > k_0$,

$$|f(x)| \leq \varepsilon m_k \qquad \text{for any } x \in \omega \text{ with } \mathbf{d}(x, \rho) < \frac{1}{\varepsilon} m_k.$$

that is, it is sublinear on a sequence of (space) scales. A simple calculation with the Gaussian upper bounds (22) and the fact that h has a linear bound shows that it is also sublinear on a sequence of time scales, that is,

$$\mathbf{E}_\rho\left[f(X_n)^2\right] \leq \varepsilon n \qquad \forall n \in \left[\tfrac{1}{2} m_k^2, 2m_k^2\right], k > k_0'. \tag{25}$$

Since the m_k were approximate square roots of a subsequence of the n_k, we may find a subsequence n_k' of n_k for which $\mathbf{E}[f(X_{n_k'})^2] \leq \varepsilon n_k'$.

We now repeat the argument of Theorem 3: Equation (11) still holds for every n:

$$\mathbf{E}_\rho|f(\rho) - f(X_1)| \leq \sqrt{2\mathbf{E}_\rho[\Delta_n(\rho, X_1)^2]\mathbf{E}_\rho[f^2(X_n)]}.$$

For our n_k' we have $\mathbf{E}[\Delta^2] \leq C/n_k'$, and with (25) we get $\mathbf{E}_\rho|f(\rho) - f(X_1)| \leq \sqrt{C\epsilon}$. Since ϵ was arbitrary, f must be constant. Since $f(0) = 0$ that constant is zero and $h = \tilde{h}(x + \chi(x))$. Therefore, any harmonic function with growth at most linear and equal to 0 at 0 belongs to a vector space of dimension d and the result follows. $\square$

A natural extension of the supercritical bond percolation setting is to look at random environments on $\mathbb{Z}^d$, such as the random conductance model. See [8, 17, 19, 74] for the existence of the corrector in different cases of this model. Similar results can probably be obtained in this setting.

5. Heat kernel derivative estimates. Our purpose in this section is to prove Theorem 6 which gives an upper bound for the (discrete) derivative of the heat kernel, $\mathbf{p}_n(x,y) - \mathbf{p}_{n-1}(x',y)$, for $x \sim x'$, where $\mathbf{p}_n(x,y) := \mathbf{P}_x(X_n = y)$.

We start with a lemma true on any graph. It relates the infinity norm of the gradient of the heat kernel to the infinity norm of the heat kernel and the entropy.

LEMMA 21. *Let G be a graph of maximal degree d. Then for any $x, x', y \in G$ with $x \sim x'$,*

$$\begin{aligned}
&\big(\mathbf{p}_{2n}(x,y) - \mathbf{p}_{2n-1}(x',y)\big)^2 \\
&\quad \le 4d(d+1) \cdot \Delta_n(x,x')^2 \cdot \max_{\substack{a,b\in\mathbf{B}_x(2n):\\ \mathbf{d}(a,b)\ge\mathbf{d}(x,y)/2}} \mathbf{p}_n(a,b) \cdot \max_{a,b\in\mathbf{B}_x(2n)} \mathbf{p}_n(a,b),
\end{aligned} \tag{26}$$

where Δ_n is defined in (8).

PROOF. Markov's property gives that

$$\mathbf{p}_{2n}(x,y) - \mathbf{p}_{2n-1}(x',y) = \sum_{a\in G} \big(\mathbf{p}_n(x,a) - \mathbf{p}_{n-1}(x',a)\big)\mathbf{p}_n(a,y).$$

Let us split the sum on $a \in G$ into two sums $I + II$, where I is the sum over $a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)$, and II on the remaining a. Using Cauchy–Schwarz we can write

$$I^2 \le \Bigg(\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \big(\mathbf{p}_n(x,a) - \mathbf{p}_{n-1}(x',a)\big)^2\Bigg)\Bigg(\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)^2\Bigg).$$

For the first term, bound the denominator in the definition of Δ_n by its maximum and get

$$\begin{aligned}
&\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \big(\mathbf{p}_n(x,a) - \mathbf{p}_{n-1}(x',a)\big)^2 \\
&\quad \le \Delta_n(x,x')^2 \cdot \max_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \big\{\mathbf{p}_n(x,a) + \mathbf{p}_{n-1}(x',a)\big\}.
\end{aligned}$$

For the second term write

$$\begin{aligned}
\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)^2 &\le \Big(\max_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)\Big) \cdot \Bigg(\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)\Bigg) \\
&\le \Big(\max_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)\Big) \cdot \Bigg(\sum_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} d \cdot \mathbf{p}_n(y,a)\Bigg) \\
&\le d \cdot \Big(\max_{a\in\mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a,y)\Big).
\end{aligned}$$

Together we get

$$I^2 \le d \cdot \Delta_n(x, x')^2 \cdot \max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \{\mathbf{p}_n(x, a) + \mathbf{p}_{n-1}(x', a)\} \\ \times \max_{a \in \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y). \tag{27}$$

Now, the second maximum in the right-hand side of (27) is a maximum on a smaller set than the first maximum in (26) [note that points in $\mathbf{B}_x(\mathbf{d}(x, y)/2)$ are at distance larger than $\mathbf{d}(x, y)/2$ from y]. Similarly, the first maximum is smaller than $(1 + d)$ times the second maximum of (26). Therefore, the product of maxima is smaller than

$$(d + 1) \cdot \max_{\substack{a,b \in \mathbf{B}_x(2n): \\ \mathbf{d}(a,b) \ge \mathbf{d}(x,y)/2}} \mathbf{p}_n(a, b) \cdot \max_{a,b \in \mathbf{B}_x(2n)} \mathbf{p}_n(a, b).$$

The estimate for II is similar:

$$II^2 \le d \cdot \Delta_n(x, x')^2 \cdot \max_{a \notin \mathbf{B}_x(\mathbf{d}(x,y)/2)} \{\mathbf{p}_n(x, a) + \mathbf{p}_{n-1}(x', a)\} \cdot \max_{a \notin \mathbf{B}_x(\mathbf{d}(x,y)/2)} \mathbf{p}_n(a, y).$$

It is easy to obtain the same bound again, except the estimates are reversed (i.e., what was bounded by the first term before is now bounded by the second term). We sum up:

$$\begin{aligned}(\mathbf{p}_{2n}(x, y) - \mathbf{p}_{2n-1}(x', y))^2 &= (I + II)^2 \le 2(I^2 + II^2) \\ &\le 4d(d + 1) \cdot \Delta_n(x, x')^2 \cdot \max_{\substack{a,b \in \mathbf{B}_x(2n): \\ \mathbf{d}(a,b) \ge \mathbf{d}(x,y)/2}} \mathbf{p}_n(a, b) \\ &\quad \times \max_{a,b \in \mathbf{B}_x(2n)} \mathbf{p}_n(a, b). \qquad \square\end{aligned}$$

In Section 2 it was always enough to discuss behavior (say of $H_n - H_{n-1}$) on a sequence n_k. Here it is no longer enough and we need an estimate that holds for all n. Hence we prove:

LEMMA 22. *For supercritical percolation, $H_n - H_{n-1} \le C/n$ for every n, where C is a constant depending only on d and p.*

PROOF. The heat kernel estimates (1) show, after a little calculation, that

$$\mathbf{H}_n = \frac{d}{2} \log n + O(1) \qquad \forall n > n_0(\omega). \tag{28}$$

For $n \le n_0(\omega)$ we can use a much rougher bound, say $\mathbf{H}_n \le d \log(2n)$ which follows from the fact that for any cluster ω the distribution of R_n is supported on the cube $\{-n, \dots, n\}^d$ and any measure has entropy smaller than the entropy of the uniform measure on its support. Since $n_0(\omega)$ has a stretched exponential tail, we can integrate over the environment and get that $H_n = \frac{d}{2}\log n + O(1)$. This means that $H_n - H_{n/2} \le C$ for some C. Using the fact that $H_n - H_{n-1}$ is decreasing (Corollary 10 on page 2344) proves the lemma. $\square$

PROOF OF THEOREM 6. As before, percolation can be seen as a stationary random graph, and it is sufficient to prove

$$\mathbb{E}\big((\mathbf{p}_{2n}(\rho,x) - \mathbf{p}_{2n-1}(X_1,x))^2 \cdot \mathbf{1}_{\{x\in\omega\}}\big) \le \frac{C_3'}{n^{d+1}} \exp\big(-C_4'\mathbf{d}(x,\rho)^2/n\big),$$

where C_3' and C_4' depend only on d and the percolation probability p.

Again we use the variables $n_y(\omega)$ from (1) and (2). Take ε to be given by the stretched exponential bound (2) for $n_y(\omega)$. Note that we can restrict ourselves to $|x| \le n^{1/2+\varepsilon/3}$, since in the regime $|x| \ge n^{1/2+\varepsilon/3}$, the heat kernel decreases fast enough so that one can tune the constant C_4' in order to obtain the result for free. Fix therefore $|x| \le n^{1/2+\varepsilon/3}$. Let $N(\omega) = \max\{n_y(\omega) : y \in \mathbf{B}_\rho(n)\}$. The Gaussian estimates (1) imply that for a.e. environment ω such that $x \in \omega$, whenever $n \ge N(\omega)$, we have

$$\begin{aligned} &\max_{\substack{a,b\in\mathbf{B}_\rho(2n):\\ \mathbf{d}(a,b)>\mathbf{d}(\rho,x)/2}} \mathbf{p}_n(a,b) \le \frac{C_3}{n^{d/2}} \exp\big[-C_4\mathbf{d}(x,\rho)^2/n\big] \quad \text{and} \\ &\max_{a,b\in\mathbf{B}_\rho(2n)} \mathbf{p}_n(a,b) \le \frac{C_3}{n^{d/2}}. \end{aligned} \tag{29}$$

Averaging (26) on the environments satisfying $N(\omega) \le n$ (for which we have (29)), we find

$$\begin{aligned} &\mathbb{E}\big[(\mathbf{p}_{2n}(\rho,x) - \mathbf{p}_{2n-1}(X_1,x))^2 \cdot \mathbf{1}_{\{x\in\omega\}}\mathbf{1}_{\{N(\omega)\le n\}}\big] \\ &\quad \le 4d(d+1) \cdot \mathbb{E}\big[\Delta_n(\rho,X_1)^2\big] \cdot \frac{C_3^2}{n^d} \exp\big[-C_4\mathbf{d}(x,\rho)^2/n\big]. \end{aligned}$$

We now apply Theorem 8 to bound $\mathbb{E}[\Delta_n^2]$ by $2(H_n - H_{n-1})$. Recall also that Lemma 22 says that for supercritical percolation $H_n - H_{n-1} \le C/n$ for all n. Together these give

$$\mathbb{E}\big[(\mathbf{p}_{2n}(\rho,x) - \mathbf{p}_{2n-1}(\tilde{X}_1,x))^2 \cdot \mathbf{1}_{\{x\in\omega\}}\mathbf{1}_{\{N(\omega)\le n\}}\big] \le \frac{C}{n^{d+1}} \exp\big[-C_4\mathbf{d}(x,\rho)^2/n\big].$$

We do not need to control the behavior of the gradient on $\{N(\omega) > n\}$ since this event has probability at most $Cn^d e^{-n^\varepsilon}$. Hence in the regime $|x| \le n^{1/2+\varepsilon/3}$ we find

$$\begin{aligned} &\mathbb{E}\big[(\mathbf{p}_{2n}(\rho,x) - \mathbf{p}_{2n-1}(\tilde{X}_1,x))^2 \cdot \mathbf{1}_{\{x\in\omega\}}\mathbf{1}_{\{N(\omega)> n\}}\big] \\ &\quad \le \mathbb{P}_p\big(N(\omega) > n\big) \le \frac{C}{n^{d+1}} \exp\big[-C_4\mathbf{d}(x,\rho)^2/n\big]. \end{aligned}$$

Putting all the pieces together, we obtain the result. □

The proof involved only Gaussian estimates at mesoscopic scale and the entropy argument. It extends to other contexts such as random conductances satisfying the uniform elliptic condition (see Example 2.1). One may then get, using convolution, annealed second space-derivative and first time-derivative estimates for the heat kernel using the first space-derivative estimates. We refer to Section 5 of [32] for more details.

6. Open questions. This article must be understood as an introduction and some initial steps in the subject. There are many natural questions on harmonic functions which remain open. We present few of them in this section.

Minimal growth harmonic functions. The question of minimal growth harmonic functions was implicitly studied in the literature: the failure of the Liouville property corresponds to a special case of minimal growth. When the Liouville property is true, it becomes interesting to determine the minimal growth. Even the deterministic case (i.e., transitive or Cayley graphs) has interesting phenomenology, and we plan to analyze some examples in a future paper. Note that groups always admit linear growth harmonic functions [55, 73, 75]. This is no longer the case for stationary random graphs. When the random walk is subdiffusive (note that the random walk on Cayley graphs is at least diffusive, a result due to Erschler; see Lee and Peres [60]), Theorem $3'$ (page 2347) implies a phenomenon which is specific to random environments.

COROLLARY 23. *Let (G, ν, ρ) be a stationary random graph with polynomial growth such that the random walk is (strictly) subdiffusive. Then, almost surely there do not exist linear growth harmonic functions.*

Therefore graphical fractals, UIPQ, critical Galton–Watson trees conditioned to survive and the incipient infinite cluster (IIC) do not admit linear growth harmonic functions. We mention that it was already proved [11] that the uniform infinite planar triangulation is almost surely Liouville. There are no nonconstant harmonic functions on the critical Galton–Watson tree or on the IIC, as both have infinitely many cut vertices. Indeed, the Galton–Watson tree is well known to be one-ended and hence, as a tree, must have infinitely many cut vertices. The existence of cut points for the IIC is essentially known, but we did not find a reference and including a full proof would take us too far off-topic.

QUESTION 1. Do there exist nonconstant harmonic functions with polynomial growth on the UIPQ?

If such functions exist, we may ask the following question:

QUESTION 2. What is the minimal growth of a nonconstant harmonic function on the UIPQ?

Space of harmonic functions with polynomial growth. Cayley graphs with polynomial growth automatically satisfy the volume doubling property and the Poincaré inequality, thus implying that spaces of harmonic functions with prescribed polynomial growth are finite dimensional. The possibility of such behavior in the case of stationary random graphs of polynomial volume growth is a legitimate question. For example:

QUESTION 3. Is the space of harmonic functions with some prescribed polynomial growth on the UIPQ finite dimensional?

Dimension of spaces of harmonic functions. The computation of the dimension of spaces of harmonic functions does not restrict to the case of linear growth harmonic functions. For a graph G and $k > 0$, let $d_k[G]$ be the dimension of the space of harmonic functions with growth bounded by a polynomial of degree k.

The similarity between $\mathbb{Z}^d$ and the infinite cluster of percolation might extend to the dimension of the space of harmonic functions with arbitrary polynomial growth. More precisely, we ask the following question:

QUESTION 4. Are the families $(d_k[\omega])_{k>0}$ and $(d_k[\mathbb{Z}^d])_{k>0}$ equal almost surely?

In particular, an interesting intermediate step toward this question would be to show that there is no harmonic function with noninteger growth.

It is natural to ask if an invariance principle for the random walk in the random environment ω implies that the sequence $(d_k[\omega])$ coincides with $(d_k[\mathbb{Z}^d])$. On $\mathbb{Z}^d$, diffusivity and the invariance principle are robust under rough isometry. Therefore, one can ask if $(d_k[G])_{k\geq 0}$ is invariant under rough isometry for these kind of graphs. This is not true in general. For instance, the Liouville property is not invariant under rough isometry; see [63] for the first example or [14] for a simpler one.

More generally, one can ask whether a small perturbation of a Cayley graph modifies drastically the harmonic functions on it. For instance, consider percolation on a Cayley graph G such that $p_u(G)$ (the infimum of the values for which there exists a *unique* infinite cluster) is strictly smaller than 1. Fix $p > p_u(G)$, and set $\omega(G)$ to be the unique infinite cluster of the percolation with parameter p.

QUESTION 5. Are the dimensions of spaces of harmonic functions with a given growth equal for G and $\omega(G)$?

Note that the question, in the case of bounded harmonic functions on the infinite percolation cluster for nonamenable Cayley graphs, was addressed in [13].

In the context of Cayley graphs, the space of harmonic functions with a certain growth rate is crucial in the study of the underlying group. Indeed, the latter acts on harmonic functions naturally. In the random setting, we do not have this interpretation. Nevertheless, an interesting question is to understand what information on the random graph is encoded in the sequence $(d_k[G])_{k\geq 0}$. In particular, the following question would be a first step in this direction:

QUESTION 6. Consider a random subgraph G of $\mathbb{Z}^d$. What are the requirements to ensure that $(d_k[G])_{k\geq 0}$ equals $(d_k[\mathbb{Z}^d])_{k\geq 0}$?

Acknowledgments. We wish to thank Ofer Zeitouni for pointing out that the entropy argument works also in nonreversible setting; Russell Lyons for directing us to [39] and Jean-Dominique Deuschel for a discussion on connectivity leading to the last part of Section 2. This paper was written during the visit of the second author to the Weizmann Institute in Israel. The first author is the incumbent of the Renee and Jay Weiss Professorial Chair.

REFERENCES

[1] ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. MR2354165

[2] ANTAL, P. and PISZTORA, A. (1996). On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** 1036–1048. MR1404543

[3] AVEZ, A. (1976). Harmonic functions on groups. In *Differential Geometry and Relativity. Mathematical Phys. and Appl. Math.* **3** 27–32. Reidel, Dordrecht. MR0507229

[4] BARLOW, M. T. (1998). *Diffusions on fractals*. *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1995). *Lecture Notes in Math.* **1690** 1–121. Springer, Berlin. MR1668115

[5] BARLOW, M. T. (2004). Which values of the volume growth and escape time exponent are possible for a graph? *Rev. Mat. Iberoam.* **20** 1–31. MR2076770

[6] BARLOW, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32** 3024–3084. MR2094438

[7] BARLOW, M. T. and BASS, R. F. (1999). Brownian motion and harmonic analysis on Sierpinski carpets. *Canad. J. Math.* **51** 673–744. MR1701339

[8] BARLOW, M. T. and DEUSCHEL, J.-D. (2010). Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.* **38** 234–276. MR2599199

[9] BARLOW, M. T. and HAMBLY, B. M. (2009). Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab.* **14** 1–27. MR2471657

[10] BARLOW, M. T. and PERKINS, E. A. (1989). Symmetric Markov chains in $\mathbf{Z}^d$: How fast can they move? *Probab. Theory Related Fields* **82** 95–108. MR0997432

[11] BENJAMINI, I. and CURIEN, N. (2012). Ergodic theory on stationary random graphs. *Electron. J. Probab.* **17** 20. MR2994841

[12] BENJAMINI, I. and CURIEN, N. (2013). Simple random walk on the uniform infinite planar quadrangulation: Subdiffusivity via pioneer points. *Geom. Funct. Anal.* **23** 501–531. MR3053754

[13] BENJAMINI, I., LYONS, R. and SCHRAMM, O. (1999). Percolation perturbations in potential theory and random walks. In *Random Walks and Discrete Potential Theory* (*Cortona*, 1997). *Sympos*. *Math*. **XXXIX** 56–84. Cambridge Univ. Press, Cambridge. MR1802426

[14] BENJAMINI, I. and SCHRAMM, O. (1996). Harmonic functions on planar and almost planar graphs and manifolds, via circle packings. *Invent*. *Math*. **126** 565–587. MR1419007

[15] BENJAMINI, I. and SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron*. *J*. *Probab*. **6** 13 pp. (electronic). MR1873300

[16] BERGER, N. and BISKUP, M. (2007). Quenched invariance principle for simple random walk on percolation clusters. *Probab*. *Theory Related Fields* **137** 83–120. MR2278453

[17] BERGER, N., BISKUP, M., HOFFMAN, C. E. and KOZMA, G. (2008). Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **44** 374–392. MR2446329

[18] BERGER, N. and DEUSCHEL, J.-D. (2014). A quenched invariance principle for non-elliptic random walk in i.i.d. balanced random environment. *Probab*. *Theory Related Fields* **158** 91–126. MR3152781

[19] BISKUP, M. and PRESCOTT, T. M. (2007). Functional CLT for random walk among bounded random conductances. *Electron*. *J*. *Probab*. **12** 1323–1348. MR2354160

[20] BOLTHAUSEN, E., SZNITMAN, A.-S. and ZEITOUNI, O. (2003). Cut points and diffusive random walks in random environment. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **39** 527–555. MR1978990

[21] BUSER, P. (1982). A note on the isoperimetric constant. *Ann*. *Sci*. *École Norm*. *Sup*. (4) **15** 213–230. MR0683635

[22] CAPUTO, P., FAGGIONATO, A. and PRESCOTT, T. (2013). Invariance principle for Mott variable range hopping and other walks on point processes. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **49** 654–697. MR3112430

[23] CARNE, T. K. (1985). A transmutation formula for Markov chains. *Bull*. *Sci*. *Math*. (2) **109** 399–405. MR0837740

[24] CHASSAING, P. and DURHUUS, B. (2006). Local limit of labeled trees and expected volume growth in a random quadrangulation. *Ann*. *Probab*. **34** 879–917. MR2243873

[25] CHOQUET, G. and DENY, J. (1960). Sur l'équation de convolution $\mu = \mu * \sigma$. *C*. *R*. *Math*. *Acad*. *Sci*. *Paris* **250** 799–801. MR0119041

[26] COLDING, T. H. and MINICOZZI, W. P. II (1997). Harmonic functions on manifolds. *Ann*. *of Math*. (2) **146** 725–747. MR1491451

[27] CONLON, J. G. and NADDAF, A. (2000). Green's functions for elliptic and parabolic equations with random coefficients. *New York J*. *Math*. **6** 153–225 (electronic). MR1781430

[28] CSISZÁR, I. (1963). Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Magyar Tud*. *Akad*. *Mat*. *Kutató Int*. *Közl*. **8** 85–108. MR0164374

[29] CSISZÁR, I. (1966). A note on Jensen's inequality. *Studia Sci*. *Math*. *Hungar*. **1** 185–188. MR0214714

[30] DELMOTTE, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev*. *Mat*. *Iberoam*. **15** 181–232. MR1681641

[31] DELMOTTE, T. (1998). Harnack inequalities on graphs. In *Séminaire de Théorie Spectrale et Géométrie*, *Vol*. 16, *Année* 1997–1998. *Sémin*. *Théor*. *Spectr*. *Géom*. **16** 217–228. Univ. Grenoble I, Saint-Martin-d'Hères. MR1666463

[32] DELMOTTE, T. and DEUSCHEL, J.-D. (2005). On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab*. *Theory Related Fields* **133** 358–390. MR2198017

[33] DERRIENNIC, Y. (1980). Quelques applications du théorème ergodique sous-additif. In *Conference on Random Walks* (*Kleebach*, 1979) (*French*). *Astérisque* **74** 183–201, 4. Soc. Math., France, Paris. MR0588163

[34] DEUSCHEL, J.-D. and KÖSTERS, H. (2008). The quenched invariance principle for random walks in random environments admitting a bounded cycle representation. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 574–591. MR2451058

[35] DE GIORGI, E. (1957). Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* (3) **3** 25–43. MR0093649

[36] DE MASI, A., FERRARI, P. A., GOLDSTEIN, S. and WICK, W. D. (1989). An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55** 787–855. MR1003538

[37] DISERTORI, M., SPENCER, T. and ZIRNBAUER, M. R. (2010). Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model. *Comm. Math. Phys.* **300** 435–486. MR2728731

[38] DUDLEY, R. M. (2002). *Real Analysis and Probability. Cambridge Studies in Advanced Mathematics* **74**. Cambridge Univ. Press, Cambridge. MR1932358

[39] ERSCHLER, A. and KARLSSON, A. (2010). Homomorphisms to $\mathbb{R}$ constructed from random walks. *Ann. Inst. Fourier* (*Grenoble*) **60** 2095–2113. MR2791651

[40] FERRARI, P. A., GRISI, R. M. and GROISMAN, P. (2012). Harmonic deformation of Delaunay triangulations. *Stochastic Process. Appl.* **122** 2185–2210. MR2921977

[41] FURSTENBERG, H. (1963). A Poisson formula for semi-simple Lie groups. *Ann. of Math.* (2) **77** 335–386. MR0146298

[42] FURSTENBERG, H. (1971). Random walks and discrete subgroups of Lie groups. In *Advances in Probability and Related Topics*, *Vol.* 1 1–63. Dekker, New York. MR0284569

[43] GLORIA, A. and OTTO, F. (2011). An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.* **39** 779–856. MR2789576

[44] GRIMMETT, G. (1999). *Percolation*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* **321**. Springer, Berlin. MR1707339

[45] GUO, X. and ZEITOUNI, O. (2012). Quenched invariance principle for random walks in balanced random environment. *Probab. Theory Related Fields* **152** 207–230. MR2875757

[46] HEBISCH, W. and SALOFF-COSTE, L. (1993). Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.* **21** 673–709. MR1217561

[47] HEYDENREICH, M., VAN DER HOFSTAD, R. and HULSHOF, T. (2011). High-dimensional incipient infinite clusters revisited. Preprint. Available at http://arxiv.org/abs/1108.4325.

[48] JÁRAI, A. A. (2003). Incipient infinite percolation clusters in 2D. *Ann. Probab.* **31** 444–485. MR1959799

[49] JERISON, D. (1986). The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J.* **53** 503–523. MR0850547

[50] KAIMANOVICH, V. A. and SOBIECZKY, F. (2010). Stochastic homogenization of horospheric tree products. In *Probabilistic Approach to Geometry. Adv. Stud. Pure Math.* **57** 199–229. Math. Soc. Japan, Tokyo. MR2648261

[51] KAĬMANOVICH, V. A. and VERSHIK, A. M. (1983). Random walks on discrete groups: Boundary and entropy. *Ann. Probab.* **11** 457–490. MR0704539

[52] KESTEN, H. (1986). The incipient infinite cluster in two-dimensional percolation. *Probab. Theory Related Fields* **73** 369–394. MR0859839

[53] KESTEN, H. (1986). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** 425–487. MR0871905

[54] KIPNIS, C. and VARADHAN, S. R. S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** 1–19. MR0834478

[55] KLEINER, B. (2010). A new proof of Gromov's theorem on groups of polynomial growth. *J. Amer. Math. Soc.* **23** 815–829. MR2629989

[56] KOZLOV, S. M. (1985). The averaging method and walks in inhomogeneous environments. *Uspekhi Mat. Nauk* **40** 61–120, 238. Russian version available at mathnet.ru. MR0786087

[57] KOZMA, G. and NACHMIAS, A. (2009). The Alexander–Orbach conjecture holds in high dimensions. *Invent. Math.* **178** 635–654. MR2551766

[58] KRIKUN, M. (2008). On one property of distances in the infinite random quadrangulation. Preprint. Available at http://arxiv.org/abs/0805.1907.

[59] LAWLER, G. F. (1982/83). Weak convergence of a random walk in a random environment. *Comm. Math. Phys.* **87** 81–87. MR0680649

[60] LEE, J. R. and PERES, Y. (2013). Harmonic maps on amenable groups and a diffusive lower bound for random walks. *Ann. Probab.* **41** 3392–3419. MR3127886

[61] LI, P. (1993). The theory of harmonic functions and its relation to geometry. In *Differential Geometry: Partial Differential Equations on Manifolds (Los Angeles, CA,* 1990). *Proc. Sympos. Pure Math.* **54** 307–315. Amer. Math. Soc., Providence, RI. MR1216591

[62] LYONS, R., PEMANTLE, R. and PERES, Y. (1995). Ergodic theory on Galton–Watson trees: Speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems* **15** 593–619. MR1336708

[63] LYONS, T. (1987). Instability of the Liouville property for quasi-isometric Riemannian manifolds and reversible Markov chains. *J. Differential Geom.* **26** 33–66. MR0892030

[64] LYONS, R. and PERES, Y. Probability on trees and networks. Book draft. Available at http://mypage.iu.edu/~rdlyons/.

[65] MARGULIS, G. A. (1966). Positive harmonic functions on nilpotent groups. *Soviet Math. Dokl.* **7** 241–244. MR0222217

[66] MATHIEU, P. and PIATNITSKI, A. (2007). Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463** 2287–2307. MR2345229

[67] MOSER, J. (1961). On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.* **14** 577–591. MR0159138

[68] MOSER, J. (1964). A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.* **17** 101–134. MR0159139

[69] NASH, J. (1958). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* **80** 931–954. MR0100158

[70] PAPANICOLAOU, G. C. and VARADHAN, S. R. S. (1982). Diffusions with random coefficients. In *Statistics and Probability: Essays in Honor of C. R. Rao* 547–552. North-Holland, Amsterdam. MR0659505

[71] PITTET, C. and SALOFF-COSTE, L. A survey on the relationships between volume growth, isoperimetry, and the behavior of simple random walk on Cayley graphs, with examples. Preprint. Available at http://www.math.cornell.edu/~lsc/surv.ps.gz.

[72] SARIG, O. Lecture notes on ergodic theory. Available at http://www.wisdom.weizmann.ac.il/~sarigo/506/ErgodicNotes.pdf.

[73] SHALOM, Y. and TAO, T. (2010). A finitary version of Gromov's polynomial growth theorem. *Geom. Funct. Anal.* **20** 1502–1547. MR2739001

[74] SIDORAVICIUS, V. and SZNITMAN, A.-S. (2004). Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields* **129** 219–244. MR2063376

[75] TAO, T. (2010). A proof of Gromov's theorem. Blog. Available at http://terrytao.wordpress.com/2010/02/18/a-proof-of-gromovs-theorem/.

[76] VAN DER HOFSTAD, R. and JÁRAI, A. A. (2004). The incipient infinite cluster for high-dimensional unoriented percolation. *J. Stat. Phys.* **114** 625–663. MR2035627

[77] VAROPOULOS, N. T. (1985). Long range estimates for Markov chains. *Bull. Sci. Math.* (2) **109** 225–252. MR0822826

[78] WOESS, W. (1994). Random walks on infinite graphs and groups—a survey on selected topics. *Bull. Lond. Math. Soc.* **26** 1–60. MR1246471

[79] YAU, S. T. (1975). Harmonic functions on complete Riemannian manifolds. *Comm. Pure Appl. Math.* **28** 201–228. MR0431040

I. BENJAMINI
G. KOZMA
DEPARTMENT OF MATHEMATICS
WEIZMANN INSTITUTE OF SCIENCE
REHOVOT
ISRAEL
E-MAIL: itai.benjamini@weizmann.ac.il
gady.kozma@weizmann.ac.il

H. DUMINIL-COPIN
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ DE GENÈVE
GENÈVE
SWITZERLAND
E-MAIL: hugo.duminil@unige.ch

A. YADIN
BEN GURION UNIVERSITY OF THE NEGEV
BEER SHEVA
ISRAEL
E-MAIL: yadina@bgu.ac.il