Submitted to the Annals of Probability

LINEAR COVER TIME IS EXPONENTIALLY UNLIKELY

BY QUENTIN DUBROFF[1,*] AND JEFF KAHN[1,†]

[1]*Department of Mathematics, Rutgers University,* [*]*qcd2@math.rutgers.edu;* [†]*jkahn@math.rutgers.edu*

Proving a 2009 conjecture of Itai Benjamini, we show:

Theorem. *For any C there is an $\varepsilon > 0$ such that for any simple graph G on V of size n, and $X_0,\ldots$ an ordinary random walk on G,*

$$\mathbb{P}(\{X_0,\ldots,X_{Cn}\}=V)<e^{-\varepsilon n}.$$

A first ingredient in the proof of this is a similar statement for Markov chains in which all transition probabilities are sufficiently small relative to C.

1. Introduction. We are motivated by a surprisingly basic question that we first heard from Ori Gurel–Gurevich in 2010: *is it true that for any fixed C and n-vertex simple graph G, the probability that a random walk on G covers $V(G)$ in Cn steps is exponentially small in n?* (Here and throughout we use Cn for what should really be $\lceil Cn\rceil$; some usage notes are included at the end of this section.)

A positive answer was conjectured by Itai Benjamini in 2009 ([3]; see also [4]), and given some support by a quite amazing argument of Benjamini, Gurel-Gurevich, and Morris [4], showing that the answer is yes if we assume any fixed bound Δ on the maximum degree of G (with the constant in the exponent then depending on Δ as well as C). That the answer is yes for trees was shown by Yehudayoff [13], who also observed that when G is an expander, a positive answer follows easily from the (less easy) large deviation bound of Gillman [8].

Here we answer the question:

THEOREM 1.1. *For any C there is an $\varepsilon>0$ such that for any simple graph G on V of size n, and $X_0,\ldots$ an ordinary random walk on G (with any rule for X_0),*

$$\mathbb{P}(\{X_0,\ldots,X_{Cn}\}=V)<e^{-\varepsilon n}.$$

The machine underlying the proof of Theorem 1.1 is the following statement for general Markov chains, which seems of independent interest. Here and in the corollary that follows, (X_t) is a Markov chain on V (of size n) with transition probabilities $\varphi(\cdot,\cdot)$, and X_I is the *set* $\{X_t : t\in I\}$ (for a set of indices I).

THEOREM 1.2. *For each C and $\beta>0$ there is a $\delta=\delta(C,\beta)>0$ such that if $W\subseteq V$, $|W|>\beta n$, $M\le C|W|$, and*

$$\max\{\varphi(v,w): v\in V, v\neq w\in W\}<\delta, \tag{1}$$

then

$$\mathbb{P}(X_{[M]}\supseteq W)=\exp[-\Omega_{C,\beta}|W|].$$

(For whatever it's worth, the δ we use is specified in (15).) In particular (roughly) the conclusion of Theorem 1.1 holds for *any* Markov chain in which the transition probabilities are small enough relative to C. This includes Theorem 1.1 itself when the minimum degree of G is sufficiently large:

MSC2020 subject classifications: Primary 05C81, 60J05.

Keywords and phrases: Markov chains, random walks on graphs, cover time.

COROLLARY 1.3. *For each C there are $\varepsilon > 0$ and d such that for any RW (X_t) on a graph of minimum degree at least d,*

$$\mathbb{P}(X_{[Cn]} = V) < e^{-\varepsilon n}.$$

This again seems interesting in its own right; e.g., we don't know another way to prove Theorem 1.1 even for the Hamming Cube ($\{0,1\}^m$ with the natural adjacencies), the scene of some of our early skirmishes with the present problem.

One might hope that Theorem 1.1 could now be handled by some combination of Corollary 1.3 and the ideas of [4], but this seems to be a dead end. (We did at least manage to "borrow" [4]'s title. The only antecedent we know of for what follows is [1, Lemma 2], whose sibling, the present Lemma 3.1, was our starting point. In particular, beautiful work of [7, 6, 14], showing (see [14, Theorem 1.1]) that cover time is "exponentially concentrated" in a different sense, seems unconnected to what we do here.)

The actual proof of Theorem 1.1 is based especially on the following easy consequence of Theorem 1.2, which again applies to general Markov chains (and in which φ_W refers to the "induced" chain on W; see "Usage" below).

COROLLARY 1.4. *With $|V| = n$, suppose the partition $V^0 \cup \cup_{i=1}^k V_i$ of V, and $U_i \subseteq V_i$ ($i \in [k]$), satisfy $|V_i| > \vartheta n$; $|V^0| < (1-\gamma_1)n$; $|U_i| > \gamma_2 |V_i|$; and, with $\delta = \delta(C/(\gamma_1\gamma_2), \gamma_2)$ (from Theorem 1.2),*

$$\max\{\varphi_{V_i}(v,w) : v \in V_i, v \neq w \in U_i\} < \delta. \tag{2}$$

Then

$$(\mathbb{P}(X_{[Cn]} = V) \leq)\ \ \mathbb{P}(X_{[Cn]} \supseteq \cup V_i) = \exp[-\Omega(n)], \tag{3}$$

where the implied constant depends on the constants C, γ_1, γ_2 and ϑ.

PROOF. Since $|\cup V_i| > \gamma_1 n$, if $X_{[Cn]} \supseteq \cup V_i$ then there is an i for which the first $(C/\gamma_1)|V_i|$ steps of the induced chain on V_i cover U_i, a set of size at least $\gamma_2|V_i|$. So Theorem 1.2 bounds the l.h.s. of (3) by $ke^{-\Omega(n)} = e^{-\Omega(n)}$. □

In what follows γ_1 and γ_2 will be "true" constants, meaning one not depending on C, and ϑ will be a function of C; thus the implied constant in (3) depends only on C and we have Theorem 1.1 whenever we can show the existence of the desired partition. Of course not every Markov chain admits such a partition (or we would have the nonsensical claim that Theorem 1.1 holds for general chains), but it seems possible that RW (again, on a simple G) does.

We will find it convenient to set (for the rest of the paper)

$$\gamma_1 = \gamma_2 = 0.1 =: \gamma,$$

but stress that any slightly small "true" constant would do as well.

QUESTION 1.5. Is it true that for each C there is a ϑ for which, for any G, RW on G admits a partition as in Corollary 1.4?

(Of course for Theorem 1.1 it would be enough to have a positive answer with γ also a function of C.)

In the event, we are only able to produce (more accurately, show existence of) such a partition under a pair of restrictions on G, but can also show that if either of these is violated then Theorem 1.1 holds for other reasons. Failure of the first restriction, which forbids too many large degrees, is handled by the next lemma, which may be thought of (not quite accurately because of the difference in the degree bounds) as a substantial extension of Corollary 1.3.

LEMMA 1.6. *For each D there is a Δ such that if*

$$|\{v : d_G(v) > \Delta\}| > \gamma n, \tag{4}$$

then $\mathbb{P}(X_{[Dn]} = V) = \exp[-\Omega(n)]$.

We postpone specifying the second restriction, which will be easier to do in the context of Section 5 (see Lemma 5.4 and (55)).

REMARK. Lemma 1.6 is the only place where we use simplicity of G, the rest of what we do being valid for general reversible chains. In that setting, Theorem 1.1 does not hold without some restriction, but e.g., the argument of Section 5 goes through essentially unchanged to show (with π denoting stationary distribution):

THEOREM 1.7. *Let (X_t) be a reversible Markov chain on V, and suppose there exists $W \subseteq V$ with $|W| \geq \alpha n$ and $\pi_v \leq K\pi_w$ $\forall v, w \in W$. Then*

$$\mathbb{P}(X_{[Dn]} = V) = \exp[-\Omega(n)],$$

where the implied constant depends on (the constants) α, K, and D.

Before closing this discussion we mention an obvious challenge:

QUESTION 1.8. Can one say anything reasonable about the value of ε in Theorem 1.1?

Whatever value can be extracted from our argument will be quite bad (We suspect it's not as bad as what could be gotten from [4], but are not volunteering to make this comparison.) As far as we know, it could be that, for slightly large C, complete graphs—for which $\mathbb{P}(\text{cover})$ is roughly $\exp[-e^{-C}n]$—are more or less the worst case; but note that for $C = 1$ (e.g.), the probability is larger for a path. At any rate, given how far we are from a decent value, there's clearly no point in trying to optimize anything, and we instead do what we can to keep things reasonably simple.

OUTLINE. Following brief preliminaries in Section 2, Theorem 1.2 and Lemma 1.6 are proved in Sections 3 and 4 (respectively), and the derivation of Theorem 1.1 via Corollary 1.4 is given in Section 5. To give some sense of Corollary 1.4, two "bonus" sections at the end of the paper return to known cases of Theorem 1.1 for which our machinery operates relatively simply: Section 6 treats expanders, and might be read as an interlude following Observation 5.1. Section 7, which reproves Yehudayoff's result for trees, can be read at any point (including this point). Finally, we have added an appendix treating a martingale concentration statement related to Section 3 (see following (10)), which is not needed for present purposes but might be of independent interest.

USAGE. We consider Markov chains $(X_t)_{t\geq 0}$ on default state space V of size n, as usual using π for stationary distribution (which will always be unique). We use $\mathbb{P}_v$ and $\mathbb{E}_v$ for probability and expectation given $X_0 = v$. For $B \subseteq V$, the *hitting time* of B is $T_B = \min\{t : X_t \in B\}$ (with, of course, $T_v = T_{\{v\}}$), and the *positive hitting time* is $T_B^+ := \min\{t > 0 : X_t \in B\}$ ($= T_B$ if $X_0 \notin B$).

We use φ for transition probabilities (and, as usual, φ^t for t-step transition probabilities) and φ_W for transition probabilities in the *induced chain* on $W \subseteq V$; that is, $\varphi_W(u,v) = \mathbb{P}_u(X_{T_W^+} = v)$. (This usage is not universal; e.g. [2] uses "chain watched on W" here and "induced chain" differently.)

Throughout $G = (V,E)$ is a (finite, connected) *simple* graph, with (again) $|V| = n$. Usage here is pretty standard: N_v for the neighborhood of (i.e. vertices adjacent to) v; d_v—or, if necessary, $d_G(v)$—for $|N_v|$ (the *degree* of v); and, for $X \subseteq V$, $N(X) = \cup_{x\in X} N_x$. We use *random walk (RW) on G* for a Markov chain on V with $\varphi(v,w) = \mathbb{1}_{\{v\sim w\}}/d_v$ (with any choice of X_0), recalling that then $\pi_v = d_v/(2|E|)$.

We use $[n]$ for $\{1,2,\dots,n\}$ and always assume n is large enough to support our arguments. To avoid needless clutter, we allow a few irrelevant abuses such as (usually) pretending large numbers are integers.

2. Preliminaries. We collect here only a few items that will be needed below (and that most readers might profitably skip). For general background on both Markov chains and martingales, see e.g. [10].

Recall that a Markov chain (with stationary distribution π) is *reversible* if, for any $u, v \in V$,

$$\pi_u \varphi(u,v) = \pi_v \varphi(v,u);$$

equivalently: for any $v_0, \dots, v_k \in V$,

$$\pi_{v_0}\varphi(v_0,v_1)\varphi(v_1,v_2)\cdots\varphi(v_{k-1},v_k) = \pi_{v_k}\varphi(v_k,v_{k-1})\varphi(v_{k-1},v_{k-2})\cdots\varphi(v_1,v_0). \tag{5}$$

(A reversible Markov chain is the same thing as RW on a weighted graph—-that is, with weights $\mathsf{w}(\cdot,\cdot)$ on edges and $\varphi(v,w) \propto \mathsf{w}(v,w)$—but we won't need this.)

The next two inequalities are for use in Section 5. The first bounds transition probabilities in terms of return probabilities. The second—monotonicity of return probabilities—will be used to deal with a *tiny* technical annoyance,

LEMMA 2.1. *[2, Lemma 3.20] For any two states v and w of a reversible Markov chain (and any s,t),*

$$\frac{\varphi^{t+s}(v,w)}{\pi_w} \leq \left[\frac{\varphi^{2t}(v,v)}{\pi_v}\frac{\varphi^{2s}(w,w)}{\pi_w}\right]^{1/2}.$$

LEMMA 2.2. *[10, Proposition 10.25] For any state v of a reversible Markov chain (and any t),*

$$\varphi^{2t+2}(v,v) \leq \varphi^{2t}(v,v).$$

The following basic martingale facts will be used in the proof of Theorem 1.2 (in Section 3). The first is a weak form of the Martingale Convergence Theorem; see e.g. [12, Theorem 5.1] and [10, Proposition A.11(i)] (and e.g. [12] for definitions).

THEOREM 2.3. *If X_s is a supermartingale with (for some L) $|X_s| \leq L$ for all $s \geq 0$, then there is a random variable X such that X_s converges to X with probability one, and*

$$\mathbb{E}X_s \leq \mathbb{E}X \ \ \forall s. \tag{6}$$

All limits in Section 3 are easily seen to exist *everywhere*, so for us the important part of Theorem 2.3 is (6).

Lastly, we recall (a special case of) the "Hoeffding-Azuma" Inequality:

THEOREM 2.4. *[10, Theorem A.10] If X_s is a martingale with $|X_{s+1} - X_s| \leq L$ for all $s \geq 0$, then*

$$\mathbb{P}(X_k - \mathbb{E}X_k > \eta) \leq e^{-\eta^2/(2kL^2)}.$$

3. Proof of Theorem 1.2. As mentioned above, our initial inspiration was Aldous' paper [1]. Our notation here is more or less his, and Lemma 3.1 was suggested by his Lemma 2.

The proof of Theorem 1.2, given at the end of this section, is a simple application of the material we are about to develop. Until then we keep the discussion slightly more general—if not as general as it might have been—to support a second application in the proof of Lemma 1.6 in Section 4.

We consider some $W \subseteq V$ and hope to show, under suitable assumptions, that

$$\mathbb{P}(X_{[M]} \supseteq W) < e^{-\Omega(n)}, \tag{7}$$

where the implied constant depends on C and $\beta < |W|/n$.

We will work with a parameter K, a (slightly large) function of C and β; but as the value of K plays no role in the first half (or so) of this discussion, we leave it unspecified until—in Lemma 3.2—it becomes relevant.

Given $W \subseteq V$, let $m = |W|$ and

$$\lambda = 1 - \max\{\varphi(v, w) : v \in V, v \neq w \in W\}. \tag{8}$$

(Though we've kept track of λ here, in our applications it will be at least $1/2$ and its precise value will be unimportant.)

Let $L = e^K$ and define random variables

$$H_v(t) = \prod_{i=0}^{t} [1 - \varphi(X_i, v)]$$

(so $H_v(-1) = 1$) and

$$r_v = \min\{t : H_v(t) < \lambda/L\}.$$

Write $a \wedge b$ for $\min(a, b)$ and parse $a \wedge b - 1 = (a \wedge b) - 1$. Define martingales

$$\xi_s^v = \mathbb{1}_{\{T_v > s \wedge r_v\}}/H_v(s \wedge r_v - 1)$$

and

$$\xi_s = \xi_s^W = \sum_{w \in W} \xi_s^w.$$

We omit the (standard, easy) proof that they *are* martingales (see e.g. the proof of [1, Lemma 1] for essentially the same argument, or the proof in Lemma 3.1 that S_k^I is a supermartingale for something more general).). We assume $X_0 \notin W$ (as we may since removing it changes essentially nothing), so

$$\mathbb{E}\xi_M = \xi_0 = m,$$

and observe that

$$|\xi_s - \xi_{s-1}| \leq L/\lambda^2. \tag{9}$$

[*Because*: with sums over v's with $T_v, r_v > s-1$ (i.e. those that can contribute here), we have

$$\begin{aligned} |\xi_s - \xi_{s-1}| &= |\sum(\xi_s^v - \xi_{s-1}^v)| \le \max\{L/\lambda, \sum \xi_{s-1}^v/(1-\varphi(X_{s-1},v)) - \xi_{s-1}^v\} \\ &= \max\{L/\lambda, \sum \tfrac{\varphi(X_{s-1},v)\xi_{s-1}^v}{1-\varphi(X_{s-1},v)}\} \le \max\{L/\lambda, (1/\lambda)\sum \varphi(X_{s-1},v)\xi_{s-1}^v\} \\ &\le \max\{L/\lambda, (L/\lambda^2)\sum \varphi(X_{s-1},v)\} \le L/\lambda^2. \end{aligned}$$

Here the second part of the max is an upper bound on $\sum(\xi_s^v - \xi_{s-1}^v)$ and the first bounds $\xi_{s-1}^v - \xi_s^v$ when there is some (necessarily unique) v for which $T_v = s$.]

The Hoeffding-Azuma Inequality, Theorem 2.4, thus gives

$$\mathbb{P}(\xi_M < m/2) = e^{-\Omega(m)}. \tag{10}$$

REMARK. Perhaps surprisingly, even ξ_∞ and the remaining ξ_s's are similarly concentrated. Since this seems interesting enough to record but isn't needed for the rest of what we do (and takes a little while to explain), we've added its proof as an appendix.

Let $\xi_\infty^v = \lim_{s\to\infty} \xi_s^v$. Define events

$$Q_v = \{\xi_\infty^v > 0\} \; (= \{T_v > r_v\}), \;\; R_v = \{r_v \le M\}, \text{ and } \; Q_v^* = Q_v \cap R_v. \tag{11}$$

Set $p = 1/L \; (= e^{-K})$.

If we cover W in M steps, then $\xi_M \le (L/\lambda)|\{v \in W : Q_v^* \text{ holds}\}|$; so if also $\xi_M \ge m/2$, then

$$|\{v \in W : Q_v^* \text{ holds}\}| \ge \lambda mp/2. \tag{12}$$

So for (7) it is enough to show

$$\mathbb{P}((12)) = e^{-\Omega(m)}. \tag{13}$$

For the situations we have in mind, this will follow easily from the next two lemmas.

LEMMA 3.1. *For any $I \subseteq W$,*

$$(\mathbb{P}(\cap_{v\in I} Q_v^*) \le) \;\; \mathbb{P}(\cap_{v\in I} Q_v) \le p^{|I|}.$$

PROOF. Let $\mathcal{F}_k$ be the σ-field generated by $(X_0, \ldots, X_k)$ and consider the process

$$S_k^I = \mathbb{1}_{\cap_{v\in I}\{T_v > k\wedge r_v\}} \prod_{v\in I} H_v(k \wedge r_v - 1)^{-1}.$$

We will show that S_k^I is a supermartingale. The Martingale Convergence Theorem (Theorem 2.3) then says $S_k^I \to S_\infty^I$ (a.s.) and

$$\mathbb{E}S_\infty^I \le \mathbb{E}S_0^I = 1, \tag{14}$$

which, since

$$S_\infty^I = \mathbb{1}_{\cap_{v\in I} Q_v} \prod_{v\in I} H_v(r_v - 1)^{-1},$$

gives the desired

$$\mathbb{P}(\cap_{v\in I} Q_v) = \mathbb{E}\big[S_\infty^I \prod_{v\in I} H_v(r_v - 1)\big] \le \mathbb{E}[S_\infty^I]\Big(\max_{v\in I} H_v(r_v - 1)\Big)^{|I|} \le p^{|I|},$$

where the second inequality follows from (14) and the definitions of H_v, r_v, λ and p, according to which $H_v(r_v-1) \le \lambda^{-1}H_v(r_v) \le 1/L = p$.

To see that S_k^I is a supermartingale (here just extending the proof of [1, Lemma 2]), it is enough to show

$$\mathbb{E}[S_{k+1}^I|\mathcal{F}_k] \le S_k^I \text{ on } \{T_v > k \wedge r_v \ \forall v \in I\}$$

(since outside this conditioning set, $S_{k+1} = S_k = 0$). But here, with $J = \{v \in I : k < r_v\}$, we have

$$\begin{aligned}\mathbb{E}(S_{k+1}^I|\mathcal{F}_k) &= \mathbb{P}(\cap_{v\in J}\{T_v > k+1\}|\mathcal{F}_k) \prod_{v\in J} H_v(k)^{-1} \prod_{v\in I\setminus J} H_v(r_v-1)^{-1} \\ &= (1 - \sum_{v\in J}\varphi(X_k, v)) \prod_{v\in J} H_v(k)^{-1} \prod_{v\in I\setminus J} H_v(r_v-1)^{-1} \\ &\le \prod_{v\in J}(1-\varphi(X_k, v)) \prod_{v\in J} H_v(k)^{-1} \prod_{v\in I\setminus J} H_v(r_v-1)^{-1} \\ &= \prod_{v\in J} H_v(k-1)^{-1} \prod_{v\in I\setminus J} H_v(r_v-1)^{-1} = S_k^I. \qquad \square\end{aligned}$$

Recalling that $p = e^{-K}$, we now set

(15) $$K = \max\{[20e(64+C)]^2, \log(1/\beta)\}, \ \ \varepsilon = \lambda p/5, \text{ and } \ \delta = \varepsilon/K$$

The reasons for these choices will appear below (see (25)), and for now we just mention that (i) the more important constraint in the definition of K is the first, and (ii) the main thing to keep in mind here is that there is nothing preventing us from taking K as large, and δ as small, as needed to make things work (cf. "Perspective" following (25)); in particular, the only reason for the fussy specification of K is to make the role of this choice a little clearer below.

REMARK. While we won't do so here, we note that with a little more care we could take K linear in C, e.g. replacing the first part of the max in (15) by 10^5C (and thinking of C as possibly depending on n). We mention two reasons to be interested in this. First, our eventual bound for Theorem 1.2 (see (27)) is roughly $\exp[-e^{-K}n]$, which if $K = O(C)$ is something like the truth, in that the value for $G = K_n$ is about $\exp[-e^{-C}n]$. Second, taking $K = O(C)$—and now for simplicity sticking to Theorem 1.1—it is not hard to show that for small enough transition probabilities, even covering in time much less than $n\log n$ is (very) unlikely; precisely: for small enough $\varepsilon > 0$ there is $c > 0$ so that if all transition probabilities are less than $n^{-\varepsilon}$ then $\mathbb{P}(X_{[cn\log n]} = V) < \exp[-n^{-\Omega(1)}]$.

Define $\varphi_\delta(y,z) = \mathbb{1}_{\{\varphi(y,z)\le\delta\}}\varphi(y,z)$ and, for a multisubset Y of V,

$$\varphi_\delta(Y,z) = \sum_{y\in Y}\varphi_\delta(y,z).$$

(For Theorem 1.2 we could skip φ_δ and work with $\varphi(Y,z)$, defined in the natural way, but the present version will be needed in Section 4.)

For the next lemma we take $\mathcal{Z}$ to be the set of those $Z \subseteq W$ of size at least $2\varepsilon m$ for which

(16) there is a multisubset Y of V of size at most M with $\varphi_\delta(Y,z) > K/4 \ \ \forall z \in Z$.

LEMMA 3.2. *There is an* $\mathcal{I} \subseteq \binom{W}{\varepsilon m}$ *with*

$$|\mathcal{I}| < (15\varepsilon)^{-\varepsilon m} \tag{17}$$

such that

$$\text{each } Z \in \mathcal{Z} \text{ contains some } I \in \mathcal{I}. \tag{18}$$

REMARK. Our eventual bound on the probability in (13) will be (using Lemma 3.1 for the inequality)

$$\sum_{I \in \mathcal{I}} \mathbb{P}(\cap_{v \in I} Q_v^*) \leq |\mathcal{I}| p^{\varepsilon m},$$

so we want the r.h.s. of (17) to be small relative to $p^{-\varepsilon m}$, which will be true with the present bound since we will have $\lambda \geq 1/2$ (recall $\varepsilon = \lambda p/5$).

PROOF OF LEMMA 3.2. Fix $Z \in \mathcal{Z}$, let Y be as in (16), and set

$$W_0 = \{z \in W : \varphi_\delta(Y, z) > \sqrt{K}\},$$

noting that

$$|W_0| < M/\sqrt{K}. \tag{19}$$

Consider the random submultiset Y' of Y gotten by including members of Y independently, each with probability $32\delta/K$, and set

$$N_\delta(Y') = \{z \in W : \varphi_\delta(Y', z) \geq \delta\}.$$

We assert that with positive probability,

$$|Y'| < 33\delta M/K =: t, \tag{20}$$

$$|N_\delta(Y') \setminus W_0| < 64m/\sqrt{K}, \tag{21}$$

and

$$|N_\delta(Y') \cap Z| > \varepsilon m. \tag{22}$$

PROOF. Since $|Y'|$ is binomial with parameters $M' \leq M$ and $32\delta/K$, the probability of violating (20) is small.

For $z \in W \setminus W_0$, we have $\mathbb{E}\varphi_\delta(Y', z) \leq 32\delta/\sqrt{K}$, and (by Markov's Inequality) $\mathbb{P}(z \in N_\delta(Y')) \leq 32/\sqrt{K}$; so $\mathbb{E}|N_\delta(Y') \setminus W_0| \leq 32m/\sqrt{K}$, and a second application of Markov gives $\mathbb{P}((21) \text{ fails}) \leq 1/2$.

Finally, set, for $z \in Z$ and $y \in Y$, $\psi_z = \varphi_\delta(Y', z)$ and $\zeta_y = \mathbb{1}_{\{y \in Y'\}}$. Then $\psi_z = \sum\{\zeta_y \varphi(y, z) : \varphi(y, z) \leq \delta\}$, $\mathbb{E}\psi_z > 8\delta$, and $\mathrm{Var}(\psi_z) < \delta \mathbb{E}\psi_z$, implying (e.g. by the second moment method; this is reason for the 32) $\vartheta := \max_{z \in Z} \mathbb{P}(\psi_z < \delta) < 1/6$. On the other hand, Markov gives $\mathbb{P}(|Z \setminus N_\delta(Y')| > 3\vartheta|Z|) < 1/3$, so $|N_\delta(Y') \cap Z| > (1 - 3\vartheta)|Z| > \varepsilon m$ with probability at least 2/3, and the assertion follows. □

It follows (by the "probabilistic method") that there *is* some multiset Y' satisfying (20), (22), and (from (19) and (21); recall $M \leq Cm$)

$$|N_\delta(Y')| < (64 + C)m/\sqrt{K}.$$

Thus with

$$\mathcal{Y} = \{Y' \text{ a multisubset of } V : |Y'| \leq t,\ |N_\delta(Y')| \leq (64 + C)m/\sqrt{K}\}, \tag{23}$$

we find that

$$\mathcal{I} := \bigcup_{Y' \in \mathcal{Y}} \binom{N_\delta(Y')}{\varepsilon m}$$

satisfies (18). But it also satisfies (17):

Noting that

$$|\mathcal{Y}| \le \binom{n+t}{t} < \binom{2m/\beta}{33C\delta m/K}$$

(using the value of t from (20) and the hypotheses $m := |W| > \beta n$ and $M \le Cm$), and recalling (15) and the bound on $|N_\delta(Y')|$ in (23), we have

$$|\mathcal{I}| < \binom{2m/\beta}{33C\delta m/K}\binom{(64+C)m/\sqrt{K}}{\varepsilon m} \tag{24}$$

$$< \left(\frac{2eK}{33C\beta\delta}\right)^{(33CK^{-2})\varepsilon m}\left(\frac{e(64+C)}{\varepsilon\sqrt{K}}\right)^{\varepsilon m} < (15\varepsilon)^{-\varepsilon m} \tag{25}$$

where we use $\binom{n}{k} \le (en/k)^k$, and for the final inequality: $\beta > e^{-K}$ bounds the first term in (25) by (say) $e^{(70C/K)\varepsilon m}$, while the definition of K (see (15)) bounds the second by $(20\varepsilon)^{-\varepsilon m}$. □

PERSPECTIVE. There is less here than meets the eye: the main point is the $1/\sqrt{K}$ of (21), which, since *we* choose K, can be used to make the second factor on the r.h.s. of (24) much smaller than $\varepsilon^{-\varepsilon m}$; though we've taken δ only (roughly) as small as necessary to make the first factor irrelevant, there was nothing to stop us from making it smaller, so this factor was not really an issue; the remaining terms (including the canceling m's) may safely be ignored.

PROOF OF THEOREM 1.2. We prove this with $\delta(C,\beta)$ the δ of (15), noting that (1) then gives $\varphi_\delta(v,w) = \varphi(v,w)$ for relevant v, w, whence $\lambda \approx 1$. As observed above, we just need (13); namely, with $Z = \{w \in W : Q_w^* \text{ holds}\}$,

$$\mathbb{P}(|Z| \ge \lambda mp/2) < e^{-\Omega(m)}. \tag{26}$$

If Q_w^* holds (in particular $M \ge r_w$), then $\lambda/L \ge \prod_{t \le M}(1 - \varphi(X_t, w))$, implying (with some room since the $\varphi(v,w)$'s are small) $\sum_{t \le M} \varphi(X_t, w) > K/2$; thus Z satisfies (16) (with $Y = \{X_1, \dots, X_M\}$). So if the event in (26) holds, then $Z \in \mathcal{Z}$ and we have $\cap_{v \in I} Q_w^*$ for some $I \in \mathcal{I}$, which according to Lemma 3.1 (and (17)) occurs with probability at most

$$|\mathcal{I}|p^{\varepsilon m} = e^{-\Omega(m)}. \tag{27}$$

□

(So here the $K/4$ in (16) could have been $K/2$, but we will need slightly more room in Section 4.)

4. Proof of Lemma 1.6. Our main new point here is Claim 4.1, given which Lemma 1.6 will be another simple application of the material of Section 3. We begin by setting parameters, in particular the Δ of the lemma, noting again that these fairly careful specifications are meant to make the arithmetic below easier to track (if one *cares* to track it), but that there is nothing delicate in these choices, since there are no constraints on Δ (beyond its being a function of D and $\gamma = 0.1$). With this advisory, we take $\beta = \gamma/2$, $C = D/\beta$ and $M = Dn$ $(= \beta Cn)$; K, ε, δ as in (15) (again, with $p = e^{-K}$);

$$d = 1/\delta \quad \text{and} \quad \varrho = \gamma Kp/(160);$$

and, finally,

$$\Delta = 16Dd^2/(\gamma\varrho) \tag{28}$$

(so Δ is roughly e^{3K}).

Set

$$B = \{v : d_v > \Delta\} \text{ and } S = \{v : d_v \le d\}.$$

CLAIM 4.1. There is a $W \subseteq B$ of size at least $|B|/2$ such that, with $S^* = N(W) \cap S$,

$$\mathbb{P}(|\{t \in [M] : X_t \in S^*\}| > \varrho n) < e^{-\Omega(n)}. \tag{29}$$

PROOF. We first observe that for all t,

$$\sum_{v \in B} \varphi^t(v, S) \le (d/\Delta)|S|. \tag{30}$$

PROOF. Using reversibility (which implies $\pi(v)\varphi^t(v,w) = \pi(w)\varphi^t(w,v)$; see (5)), we have

$$\sum_{v \in B} \varphi^t(v,S) = \sum_{v \in B}\sum_{w \in S} \varphi^t(v,w) = \sum_{w \in S}\sum_{v \in B} \varphi^t(w,v) d_w/d_v$$
$$\le (d/\Delta) \sum_{w \in S} \varphi^t(w,B) \le (d/\Delta)|S|. \qquad \square$$

It follows that for all T (now just using $|S| \le n$),

$$|B|^{-1} \sum_{v \in B} \mathbb{P}_v(X_{[T]} \cap S \ne \emptyset) \le Td/(\gamma\Delta)$$

and

$$|\{v \in B : \mathbb{P}_v(X_{[T]} \cap S \ne \emptyset) \ge 2Td/(\gamma\Delta)\}| \le |B|/2;$$

so if we set $T = \gamma\Delta/(4d)$ and take

$$W = \{v \in B : \mathbb{P}_v(X_{[T]} \cap S \ne \emptyset) < 2Td/(\gamma\Delta) \ (= 1/2)\},$$

then

$$|W| \ge |B|/2$$

and we just have to show

$$W \text{ satisfies (29).} \tag{31}$$

To see this, let ξ_i be the time between the $(i-1)$st and ith visits to S^*. Then (independent of history up to the $(i-1)$st visit),

$$\mathbb{P}(\xi_i > T) > 1/(2d). \tag{32}$$

(Starting from $v \in S^*$, we're in W at the first step with probability at least $1/d$ and then with probability at least 1/2 the time to return to S^* is at least T.) With $\psi_i = \mathbb{1}_{\{\xi_i > T\}}$ and $\psi = \sum_{i=1}^{\varrho n} \psi_i$, visiting S^* more than ϱn times (the event in (29)) requires

$$\psi < M/T = \varrho n/(4d).$$

But ψ stochastically dominates $\psi' \sim \text{Bin}(\varrho n, 1/(2d))$ (by (32)), and $\mathbb{P}(\psi' < \varrho n/(4d)) < e^{-\Omega(n)}$.

This completes the proofs of (31) and Claim 4.1. $\square$

PROOF OF LEMMA 1.6. We use the machinery of Section 3 with W as in Claim 4.1 (and $|W| = m$) and other parameters as in the first paragraph of this section.

[One picky adjustment: If v is pendant (i.e. of degree one) with unique neighbor w, then $\varphi(v,w) = 1$ and the λ of (8) can be zero. But, except when $X_0 = v$, transitions from pendant vertices have no effect on anything in Section 3, since the mandatory next vertex has already been seen and is no longer contributing to the martingale. So for the present application we may without penalty modify (8) to require $d_G(v) \geq 2$ (and—getting sillier—exclude the unique neighbor of X_0 from W if X_0 happens to be pendant); thus we assume for this little discussion that $\lambda \geq 1/2$.]

Define events

$$E = \{|\{v \in W : Q_v^* \text{ holds}\}| > \lambda pm/2\}$$

and

$$F = \{|\{t \in [M] : X_t \in S^*\}| \leq \varrho n\}$$

(S^* as in Claim 4.1). As earlier, we just need to show (13) (namely, $\mathbb{P}(E) < e^{-\Omega(n)}$), which in view of (29) will follow from

$$\mathbb{P}(E \cap F) < e^{-\Omega(n)}.$$

To see this, note first that, with

$$Z_0 = \left\{v : \sum\{\varphi(X_t,v) : t \in [M], X_t \in S^*\} > K/4\right\},$$

F implies $|Z_0| < 4\varrho n/K$.

As in the proof of Theorem 1.2, if Q_v^* holds, then $\lambda/L \geq \prod_{t \leq M}(1 - \varphi(X_t,v))$, which in the present situation (i.e. where $\lambda \geq 1/2$) implies

$$\sum_{t \leq M} \varphi(X_t,v) \geq (2\log 2)^{-1} \log(L/\lambda) > K/2;$$

so if also $v \in W \setminus Z_0$, then (since $d = 1/\delta$),

$$\sum_{t \leq M} \varphi_\delta(X_t,v) > K/4.$$

So if $E \wedge F$ holds, then

$$Z := \{v \in W : Q_v^* \text{ holds}\} \setminus Z_0$$

satisfies (16) with

$$Y = \{X_t : t \in [M], X_t \notin S^*\},$$

and (with minor arithmetic, again using $\lambda \geq 1/2$)

$$|Z| > \lambda pm/2 - 4\varrho n/K > 2\varepsilon m;$$

that is, $Z \in \mathcal{Z}$. We thus have $\cap_{v \in I} Q_v^*$ for some $I \in \mathcal{I}$, and, by Lemma 3.1,

$$\mathbb{P}(E \wedge F) \leq |\mathcal{I}| p^{\varepsilon m} = e^{-\Omega(m)}.$$

□

5. Partitions. Here we prove Theorem 1.1. As mentioned earlier, this will be based on Corollary 1.4 provided we exclude two possibilities—(4) and (52)—that imply the conclusion of the theorem for other reasons (as shown earlier in Lemma 1.6 and soon in Lemma 5.4).

We fix C and consider a walk of length Cn on the (n-vertex) graph G. Let $\delta = (2/3)\delta(C/\gamma^2,\gamma)$ (see Theorem 1.2 for $\delta(\cdot,\cdot)$, Corollary 1.4 for our intended use, and (41) for the silly reason for the 2/3), and let Δ be as in Lemma 1.6 with C in place of D, and

$$\theta = \delta^2.$$

(This extra parameter could be skipped, but is included as it will appear pretty often.)

Set (for any v and R)

$$B_v(R) = \{w \neq v : \mathbb{P}_v(w \in X_{[R]}) > \delta/2\}$$

and

$$B'_v(R) = \{w \neq v : \mathbb{P}_v(T_w \leq \min\{R, T_v^+\}) > \delta/2\}.$$

(We don't actually need the superset $B_v(R)$ of $B'_v(R)$, but keep it to point out that the upper bound shown in Lemma 5.3 doesn't use the extra constraint in $B'_v(R)$.)

PREVIEW. For any specification of V_i's we will take

$$U_i = \{w \in V_i : \max_{w \neq v \in V_i} \varphi_i(v,w) < \delta\}, \tag{33}$$

Thinking of v's that cause exclusions from these U_i's, we say $v \in W \subseteq V$ is *good for W* (or just *good* if the identity of W is clear) if

$$\max_{v \neq w \in W} \varphi_W(v,w) < \delta. \tag{34}$$

We are hoping for V_i's in which few vertices are bad (not good), in which case we can use the trivial

$$|V_i \setminus U_i| \leq |\{v \in V_i : v \text{ is bad for } V_i\}|/\delta. \tag{35}$$

Perhaps surprisingly—and luckily, since other options seem difficult—much of our production of such V_i's (all but what's covered by Lemma 5.2) can be based on the following easy point.

OBSERVATION 5.1. *For any R, sufficient conditions for v to be good for W are*

$$W \cap B'_v(R) = \emptyset \tag{36}$$

and

$$\mathbb{P}_v(X_{[R]} \cap W = \emptyset) < \delta/2. \tag{37}$$

(These are enough since then for $w \in W \setminus \{v\}$,

$$\varphi_W(v,w) \leq \mathbb{P}_v(T_w \leq \min\{R, T_v^+\}) + \mathbb{P}_v(X_{[R]} \cap W = \emptyset) < \delta.)$$

Note. As mentioned earlier, a reader interested in a warm-up for what we're about to do might find this a good time to take a look at Section 6.

Before turning to our main line of argument we dispose of an easy case. Say v is (δ,R)-*recurrent* if

$$\mathbb{P}_v(T_v^+ \leq R) > 1 - \delta \tag{38}$$

and (δ,R)-*transient* otherwise.

LEMMA 5.2. *If, for some R,*

$$|\{v : v \text{ is } (\delta, R)\text{-recurrent}\}| > 2\gamma n, \tag{39}$$

then G admits a partition as in Corollary 1.4 with $\vartheta = \delta\gamma/(2R)$.

PROOF. Let S be the set in (39). Notice that, for any v, w,

$$\text{if } d_G(w) \ge d_G(v) \text{ and } v \in B'_w(R), \text{ then } w \in B'_v(R) \tag{40}$$

(since $\mathbb{P}_v(T_w \le \min\{R, T_v^+\}) = (d_G(w)/d_G(v))\mathbb{P}_w(T_v \le \min\{R, T_w^+\})$; see (5)).

Let Γ be the graph on S with $v \sim w$ if $w \in B'_v(R)$ or vice versa. Order V by some "$\prec$" with $v \prec w \Rightarrow d_G(v) \le d_G(w)$ and notice that (40) implies (the first inequality in)

$$d^+(v) \le |B'_v(R)| < 2R/\delta \quad \forall v \in S$$

(where $d^+(v) = |\{w : v \prec w \sim v\}|$), whence the chromatic number of G is at most $2R/\delta$ (just color the vertices greedily, from last to first w.r.t. "$\prec$"). We now take $\{W_j\}$ to be a (proper) $(2R/\delta)$-coloring of Γ, and notice that, for any j and distinct $v, w \in W_j$,

$$\varphi_{W_j}(v, w) \le \mathbb{P}_v(T_w \le \min\{R, T_v^+\}) + \mathbb{P}_v(T_v^+ > R) < 3\delta/2. \tag{41}$$

(as in (2)). We also have (with ϑ as in Lemma 5.2)

$$\sum\{|W_j| : |W_j| \le \vartheta n\} \le (2R/\delta)\vartheta n = \gamma n;$$

so we satisfy the demands of Corollary 1.4 by taking $\{V_i\} = \{W_j : |W_j| > \vartheta n\}$ and $U_i = V_i$ $\forall i$ (and $V^0 = V \setminus \cup V_i$).

(For clarity we just note that the bound we actually need in (41) is $\delta(C/\gamma, 1) > 3\delta/2$.) $\square$

We now turn to the main argument. Fix k with

$$\delta^{k-3} < (16\sqrt{\Delta})^{-1}$$

($k = 5$ will do since Δ is roughly δ^{-3}), and let N be minimum with

$$(1-\delta)^N < \delta^k. \tag{42}$$

Let $R_0 = 1$ and, for $i \ge 1$,

$$R_i = 4CN\delta^{-10}R_{i-1}.$$

Choose $i < 10\delta^{-k}$ for which at least $.9n$ vertices v satisfy

$$\mathbb{P}_v(T_v^+ \in (R_{i-1}, R_i]) < \delta^k. \tag{43}$$

Parameters we will use (collected here to have them in one place, though it will take us a little while to get to the Q_i's) are then:

$$R' = NR_{i-1}, \;\; Q = (4/\delta)R', \;\; R = R_i,$$

$$Q_1 = Q\delta^{-3}, \;\; \text{and} \;\; Q_2 = Q_1\delta^{-2}\theta^{-1} = Q_1\delta^{-4}$$

(so $R = CQ_2/\theta$). We also abbreviate

$$(R', R] = I,$$

since this interval will appear frequently. (The ratios between parameters are generous but convenient, in particular supporting occasional use of inequalities of the form $e^{-1/\delta} < \delta^{O(1)}$, which hold since δ is small.) For minor reasons at (50) we want—and, to avoid very silly distractions, will just assume—

$$R'+1 \text{ is even.} \tag{44}$$

In view of Lemmas 1.6 and 5.2, we may assume

$$\text{at least } .6n \text{ vertices } v \text{ are } (\delta,R)\text{-transient, have } d_G(v)<\Delta, \text{ and satisfy (43).} \tag{45}$$

Let $\mathcal{T}$ be the set of such v's.

LEMMA 5.3. *For any* $v\in\mathcal{T}$, $|B_v(R)\cap\mathcal{T}|\le Q$.

PROOF. We first observe that

$$\mathbb{P}_v(v\in X_I)<2\delta^{k-1}.$$

[Because: If $\{t\ge 0: X_t=v\}=\{t_0<t_1<t_2<\cdots\}$, then $v\in X_I$ implies that either

$$t_u-t_{u-1}\le R'/N \quad \forall u\in[N] \tag{46}$$

or, for some $j\le N$,

$$t_u-t_{u-1}\le R'/N \;\text{ for }\; u\in[j-1] \;\text{ and }\; t_{j-1}+R'/N<t_j\le R. \tag{47}$$

But by (42) and (43) the probabilities of (46) and (47) are less than (respectively) δ^k and

$$\delta^k\sum_{j\in[N]}(1-\delta)^{j-1}<\delta^{k-1}.]$$

Set $\ell_v(I)=|\{t\in I: X_t=v\}|$ and notice that, for any $v\in\mathcal{T}$,

$$\mathbb{E}_v\ell_v(I)<2\delta^{k-2} \tag{48}$$

(since $\mathbb{E}_v\ell_v(I)=\sum_{u\ge1}\mathbb{P}_v(\ell_v(I)\ge u)\le 2\delta^{k-1}\sum_{u\ge1}(1-\delta)^{u-1}=2\delta^{k-2}$).

It follows that for distinct $v,w\in\mathcal{T}$,

$$(\mathbb{P}_v(w\in X_I)\le)\quad \mathbb{E}_v\ell_w(I)=\sum_{t\in I}\varphi^t(v,w)$$

$$\le\sqrt{\Delta}\left[\sum_{t\in I}\varphi^{2\lfloor t/2\rfloor}(v,v)\sum_{t\in I}\varphi^{2\lceil t/2\rceil}(w,w)\right]^{1/2} \tag{49}$$

$$\le\sqrt{\Delta}\left[2\sum_{t\in I}\varphi^t(v,v)2\sum_{t\in I}\varphi^t(w,w)\right]^{1/2} \tag{50}$$

$$<4\sqrt{\Delta}\cdot\delta^{k-2}<\delta/4, \tag{51}$$

where (49) is Lemma 2.1 and Cauchy-Schwarz, (50) uses Lemma 2.2 and (44), and (51) is given by (48).

Thus, finally,

$$B_v(R)\cap\mathcal{T}\subseteq\{w\in\mathcal{T}\setminus\{v\}:\mathbb{P}_v(w\in X_{[R']})>\delta/4\},$$

a set of size at most $(4/\delta)R'=Q$. □

MORE PREVIEW. In what follows, aiming for Corollary 1.4, we will discard $V \setminus \mathcal{T}$ (that is, include it in V^0) and consider a random partition of $\mathcal{T}$, hoping to use Observation 5.1 (and the discussion preceding it) to say that (with good probability) much of $\mathcal{T}$ lies in blocks that behave as the corollary requires. Roughly speaking, what we get from Lemma 5.3 is likelihood of (36): if the number of blocks in our random partition is much larger than Q, then the block containing v is unlikely to meet $B'_v(R)$.

For (37) a natural intuition is that "transience" (failure of (38)) implies that, for the walk started from v, $X_{[R]}$ is likely to be large, which, suitably quantified, does imply that (37) is likely (for v and its random block W). This intuition turns out to be not quite correct, but, as shown in Lemma 5.4, if it is wrong too often then the conclusion of Theorem 1.1 holds for other (simpler) reasons.

Set

$$\mathcal{D} = \{v \in \mathcal{T} : \mathbb{P}_v(|X_{[R]} \cap \mathcal{T}| < Q_1) > \theta\delta\}.$$

LEMMA 5.4. *If*

$$|\mathcal{D}| \geq 2\theta n, \tag{52}$$

then $\mathbb{P}(X_{[Cn]} \supseteq V) = e^{-\Omega(n)}$.

PROOF. We first claim that

$$\text{for any } v \in V, \quad \mathbb{P}_v(|X_{[R]} \cap \mathcal{D}| > Q_2) < \exp[-1/\delta]. \tag{53}$$

PROOF. With (X_t) started from v, let $t_0 = \min\{t : X_t \in \mathcal{D}\}$ and, for $i \geq 1$,

$$t_i = \min\{t : X_t \in \mathcal{D},\ |X_{(t_{i-1},t]} \cap \mathcal{T}| \geq Q_1\}.$$

(That is, t_i is the first time that the walk is in $\mathcal{D}$, having seen at least Q_1 distinct vertices of $\mathcal{T}$ since t_{i-1}.)

For the event in (53) we must have (*very* generously)

$$t_i - t_{i-1} \leq R \quad \forall i \in [Q_2/Q_1],$$

which, since each $X_{t_{i-1}}$ is in $\mathcal{D}$, occurs with probably less than $(1-\theta\delta)^{Q_2/Q_1} < e^{-1/\delta}$. □

We can now show

$$\mathbb{P}(|X_{[Cn]} \cap \mathcal{D}| \geq 2CQ_2 n/R) < e^{-\Omega(n)}, \tag{54}$$

which gives the lemma since $CQ_2/R = \theta$.

PROOF OF (54). For $i \in [Cn/R]$ let ξ_i be the indicator of

$$\{|X_{((i-1)R,iR]} \cap \mathcal{D}| > Q_2\}.$$

Then

$$|X_{[Cn]} \cap \mathcal{D}| \leq R \sum \xi_i + CQ_2 n/R,$$

so the event in (54) requires $\xi := \sum \xi_i > CQ_2 n/R^2$. But ξ is stochastically dominated by $\xi' \sim \text{Bin}(Cn/R, e^{-1/\delta})$ (by (53)), and $\mathbb{P}(\xi' > CQ_2 n/R^2) < e^{-\Omega(n)}$. □

So we may assume

$$|\mathcal{D}| < 2\theta n. \tag{55}$$

For the partition of Corollary 1.4, we include $V \setminus \mathcal{T}$ in V^0 and will mainly be interested in $\mathcal{T} \setminus \mathcal{D}$. Setting

$$\zeta = \theta/Q,$$

we randomly (uniformly) partition $\mathcal{T}$ into ζ^{-1} blocks, usually called W, and want to say that each $v \in \mathcal{T} \setminus \mathcal{D}$ is likely to be good (meaning, of course, good in its block).

LEMMA 5.5. *If $v \in \mathcal{T} \setminus \mathcal{D}$ then $\mathbb{P}(v$ bad$) < 4\theta$.*

PROOF. We want to say that, at least for $v \in \mathcal{T} \setminus \mathcal{D}$, (36) and (37) are likely for v and the block W containing it. For (36) this is just

$$\mathbb{P}(W \cap B'_v(R) \neq \emptyset) < \zeta|B'_v(R) \cap \mathcal{T}| < \zeta Q = \theta \tag{56}$$

(this just requires $v \in \mathcal{T}$; see Lemma 5.3).

For (37) (now using $v \notin \mathcal{D}$), with unsubscripted $\mathbb{P}$ referring to the choice of the block W containing v *and* the walk from v, we have

$$\begin{aligned} \mathbb{E}_W[\mathbb{P}_v(X_{[R]} \cap W = \emptyset)] &= \mathbb{P}(X_{[R]} \cap W = \emptyset) \\ &< \mathbb{P}_v(|X_{[R]} \cap \mathcal{T}| < Q_1) + e^{-\zeta Q_1} < \theta\delta + e^{-1/\delta} =: q; \end{aligned} \tag{57}$$

so by Markov's inequality,

$$\mathbb{P}_W[\mathbb{P}_v(X_{[R]} \cap W = \emptyset) \geq \delta/2] < 2q/\delta < 3\theta. \tag{58}$$

Combining (58) and (56) now completes the proof of Lemma 5.5. □

Again considering our random partition, and using (55) and Lemma 5.5, we find that there *exists* a partition $\{W_i : i \in [\zeta^{-1}]\}$ of $\mathcal{T}$ with (say)

$$|W_i| > \zeta n/2 \ \forall i \tag{59}$$

and

$$|\{v : v \text{ bad}\}| < 5\theta|\mathcal{T} \setminus \mathcal{D}| + |\mathcal{D}| < 7\theta n$$

(where, again, "v bad" means bad in its W_i).

Say W_i is *nice* if

$$|\{v \in W_i : v \text{ bad}\}| < \delta|W_i|/2,$$

noting that this implies

$$|U_i| > |W_i|/2$$

(recalling that U_i was defined in (33) and using (35)).

On the other hand,

$$\sum\{|W_i| : W_i \text{ not nice}\} \leq (2/\delta)|\{v : v \text{ bad}\}| < 14\theta n/\delta,$$

whence $\sum\{|W_i| : W_i \text{ nice}\} > |\mathcal{T}| - 14\theta n/\delta > .5n$ (see (45)); so, with $\vartheta = \zeta/2$ (see (59)), the collection $\{V_j\}$ of nice W_i's, with $V^0 = V \setminus \cup V_j$, is the desired partition. □

6. Expanders. As promised near the end of Section 1, this and the next section give separate treatment to two previously known cases of Theorem 1.1, as relatively simple illustrations of the use of Corollary 1.4. Here we provide (a little sketchily) a simpler substitute for much of Section 5 in the case of expanders (for which, as said earlier, Theorem 1.1 was observed in [13] to follow easily from [8]). Note we are still using the defaults $G=(V,E)$ and $|V|=n$.

Suppose the transition matrix, P, of RW on G has eigenvalues $1=\lambda_1\geq\cdots\geq\lambda_n\geq-1$ (as guaranteed by Perron-Frobenius). We call G an *ε-expander* if $\max\{|\lambda_2|,|\lambda_n|\}<1-\varepsilon$. We should show:

THEOREM 6.1. *For RW (X_t) on an ε-expander G,*

$$\mathbb{P}(X_{[Cn]}=V)=\exp[-\Omega_{\varepsilon,C}|V|].$$

(Note δ, Δ are still as in the second paragraph of Section 5.) In view of Lemma 1.6, we may assume at least $(1-\gamma)n$ vertices of G have degree at most Δ. Let $\mathcal{T}$ be the set of such vertices. Application of Observation 5.1 here will be based on the next two assertions.

PROPOSITION 6.2. *[11, Theorem 5.1] For an ε-expander G and $S\subseteq V$,*

$$|\varphi^t(v,S)-\pi_S|\leq\sqrt{\pi_S/\pi_v}\,(1-\varepsilon)^t.$$

PROPOSITION 6.3. *For RW on an ε-expander G and $S\subseteq V$,*

$$\mathbb{P}(T_S>t)<(1-\pi_S/2)^{\varepsilon t/(2\log n)}.$$

[We include the trivial proof: Set $s=2\log n/\varepsilon$. Proposition 6.2 gives (say) $\mathbb{P}(X_{r+s}\in S|X_r=v)>\pi_S/2$ for any r and v, so

$$\mathbb{P}(T_S>t)\leq\mathbb{P}(X_{ks}\notin S\;\forall k\in[t/s])<(1-\pi_S/2)^{\varepsilon t/(2\log n)}.]$$

Now thinking of (36), we observe that there is a fixed Q such that for any $R=o(n)$ and v,

$$(|B'_v(R)\cap\mathcal{T}|\leq)\;\;|B_v(R)\cap\mathcal{T}|<Q. \tag{60}$$

[Because: By Proposition 6.2, there is a fixed T (depending on $\varepsilon,\delta,\Delta$) so that, for any $w\in\mathcal{T}$,

$$\mathbb{P}_v(w\in X_{(T,R]})<\sqrt{\Delta}\varepsilon^{-1}(1-\varepsilon)^T+R\pi_w<\delta/4;$$

so $B_v(R)\cap\mathcal{T}\subseteq\{w:\mathbb{P}_v(w\in X_{[T]})>\delta/4\}$, a set of size less than $4T/\delta=:Q$.]

On the other hand, Proposition 6.3 *guarantees* (37) whenever $R=\omega(\log n)$ and $|W|=\Omega(n)$.

Now set $R=\sqrt{n}$ (we need $\log n\ll R\ll n$) and $\zeta=\theta/Q$ (recall $\theta=\delta^2$), and consider a random (uniform) partition, $\{W_i\}$, of $\mathcal{T}$ into ζ^{-1} blocks. By Observation 5.1 and the discussion above, the probability that $v\in\mathcal{T}$ is bad in its block W is less than

$$\mathbb{P}(W\cap B_v(R))=\emptyset)+\mathbb{P}(|W|<\zeta n/2)<\zeta Q+o(1)=\theta+o(1).$$

The rest of this is essentially the same as the end of Section 5 (following the proof of Lemma 5.5 and omitting $\mathcal{D}$); so we won't duplicate, but briefly: The preceding discussion shows existence of a partition $\{W_i\}$ of $\mathcal{T}$ with (say) $|W_i|>\zeta n/2=:\vartheta n\;\forall i$, and only $2\theta n$ bad v's. We then discard (add to $V\setminus\mathcal{T}$ to form V^0) any W_i's that are "not nice," meaning $|\{v\in W_i: v \text{ bad}\}|>\delta|W_i|/2$, and take $\{V_j\}=\{\text{nice } W_i\text{'s}\}$.

(The definition of "nice" is chosen so that W_i nice implies $|U_i|>|W_i|/2$ (see (33) for U_i), and the bound on the number of bad v's, with $\theta\ll\delta$, implies that the number of discarded vertices is small.) □

REMARK. This could also have been handled deterministically, as in the proof of Lemma 5.2, but the intention here was to parallel the main argument of Section 5.

7. Trees. Here we give the promised alternate proof of Theorem 1.1 for trees. This is again based on Corollary 1.4, but now without Observation 5.1. The proof is constructive (unlike that of Section 5) and gives more than the corollary requires:

THEOREM 7.1. *For RW on a tree T, $\delta > 0$ and $t = 1/\delta$, there is a partition $V = W_1 \cup \cdots \cup W_k$ with $k \le (t+1)t^{t+1}$ and (for all i)*

$$\max\{\varphi_{W_i}(v,w) : v, w \in W_i, v \ne w\} \le \delta.$$

(To get a partition as in Corollary 1.4 from this, set $\vartheta = (2k)^{-1}$, and take $\{V_j\} = \{W_i : |W_i| \ge \vartheta n\}$, $U_i = V_i$, and $V^0 = V \setminus \cup V_i$, noting that $|V^0| \le n/2$.)

Our construction is based especially on the following easy property of trees (see e.g. [11, Prop. 2.3]), in which $d(\cdot,\cdot)$ is distance.

PROPOSITION 7.2. *For distinct vertices v, w of T, $\mathbb{P}_v(T_w < T_v^+) \le 1/d(v,w)$.*

USAGE. We regard trees as rooted. As usual, v is an *ancestor* of w (and w a *descendant* of v) if v lies on the path joining w to the root. We use D_v for the set of descendants of v, $v \wedge w$ for the most recent common ancestor of v, w (the one furthest from the root), and L_i for the set of vertices at distance i from the root.

We will find it convenient to treat partitions as colorings (of V). We say $W \subseteq V$ is *safe* if

$$\max\{\varphi_W(v,w) : v, w \in W, v \ne w\} \le \delta,$$

and a coloring σ is safe if $\sigma^{-1}(c)$ is safe for every c. Since (trivially) $\varphi_W(v,w) \le \varphi_{W'}(v,w)$ whenever $v, w \in W \subseteq W'$, Proposition 7.2 implies

(61) if $W_1, \dots$ are safe and $d(W_i, W_j) \ge 1/\delta$ $\forall i \ne j$, then $\cup W_i$ is safe.

For the partition of Theorem 7.1 the main thing we have to show is:

CLAIM 7.3. For any T, there is a safe coloring of L_t with at most $(t+1)t^t$ colors.

PROOF OF THEOREM 7.1 GIVEN CLAIM 7.3. Let $\mathcal{D}_q$, $q \in [t]$, be disjoint sets of colors, each of size $(t+1)t^t$. By (61) it is enough to find, for each q and $i \equiv q \pmod t$, a safe coloring of L_i using colors from $\mathcal{D}_q$. For $i \ge t$ this is accomplished by applying Claim 7.3 to $D_v \cap L_i$ for each $v \in L_{i-t}$ (and again using (61)); for smaller i, we can apply the claim to the tree gotten from T by adding a new root and a path of length $t - i$ joining it to the root of T. (Or check that the proof of the claim also applies here.) □

PROOF OF CLAIM 7.3. Let $\mathcal{C}_B, \mathcal{C}_1, \dots, \mathcal{C}_t$ be disjoint sets of colors of size t^t. We color L_t in stages. For a given stage, we use U for the set of uncolored vertices at the beginning of the stage, and, for $v \in L_0 \cup \ldots \cup L_{t-1}$, $U_v = D_v \cap U$. The process continues until $|U| \le t^t$, at which point we complete the coloring by assigning distinct colors from $\mathcal{C}_B$ to the vertices of U.

If $|U| > t^t$, we choose $v \in L_i$ with $|U_v| > t^{t-i}$ and i as large as possible (so $|U_w| \le t^{t-j}$ for each j and $w \in D_v \cap L_j$). Call $S \subseteq U_v$ *primitive* (w.r.t. v) if $w \wedge z = v$ for all distinct $w, z \in S$. For $j = 1, \dots,$ let S_j be a maximal primitive subset of $U_v \setminus (S_1 \cup \ldots \cup S_{j-1})$, ending, say at S_ℓ, as soon as the largest surviving primitive set has size less than t. Thus each of $|S_1|, \dots, |S_\ell|$ is at least t and, by our choice of i,

$$\ell \ (\le \max\{|U_w| : w \text{ a child of } v\}) \ \le t^{t-i-1};$$

so we may assign $S_1,\dots,S_\ell$ distinct colors from $\mathcal{C}_i$ (and could have taken $|\mathcal{C}_i|=t^{t-i-1}$). This completes the stage and leaves v with fewer than t^{t-i} uncolored descendants (since fewer than t of its children now have *any* uncolored descendants. Since each v is "processed" at most once, we eventually have $|U|\le t^t$ and (as above) finish the coloring using $\mathcal{C}_B$.

It remains to show that the coloring, σ, is safe. Suppose instead that $\sigma_w=\sigma_z=c$ (for some $w\neq z$ and c). Since $|\sigma^{-1}(c)|\le 1$ for $c\in\mathcal{C}_B$, we have $c\in\mathcal{C}_i$ for some i. But then (e.g.) w was colored as part of a primitive set $S=\{w_1,\dots,w_s\}$, with $s\ge t$ and common ancestor $v\in L_i$; so, since the path from z to w includes v, we have $\varphi_c(z,w_j)\ge\varphi_c(z,w)$ $\forall j$ (with equality if $z\neq w_j$), where $\varphi_c=\varphi_W$ with $W=\sigma^{-1}(c)$. Thus $\varphi_c(z,w)\le 1/s\le\delta$. $\square$

APPENDIX: CONCENTRATION

Usage here is as in Section 3, and v will always be a vertex of W. As promised following (10), we show that each ξ_s $(=\xi_s^W)$ is exponentially concentrated about its mean.

THEOREM A.4. *For any* $\vartheta>0$,

$$\mathbb{P}(|\xi_s-m|>\vartheta m)\le 2e^{-\vartheta^2\lambda^4 m/(8L)^2}.$$

Since $Q_v=\{\xi_\infty^v>0\}$, this gives exponential tail bounds for $|\{v:Q_v\}|$. (This isn't quite concentration about the mean since we only know $|\{v:Q_v\}|L\le\xi_\infty\le(L/\lambda)|\{v:Q_v\}|$.)

Theorem A.4 is proved using a better martingale analysis, based on an idea from [9]. We set $Z_i=\xi_i-\xi_{i-1}$ and $Z=\sum_{i=1}^s Z_i$ $(=\xi_s-\xi_0)$, and as usual want to bound $\mathbb{E}[e^{\zeta Z}]$ (with $\zeta>0$ to be specified). The main point here, an instance of [9, Lemma 3.4], is that we can replace the usual product of worst case bounds in

$$\mathbb{E}[e^{\zeta Z}]\le\prod_{i=1}^s\max_{H_{i-1}}\mathbb{E}[e^{\zeta Z_i}\mid H_{i-1}]$$

by a worst case product:

LEMMA A.5. *With each H_i ranging over events $\{X_0=x_0,X_1=x_1,\dots,X_i=x_i\}$,*

$$\mathbb{E}[e^{\zeta Z}]\le\max\Big\{\prod_{i=1}^s\mathbb{E}[e^{\zeta Z_i}\mid H_{i-1}]:H_0\supseteq H_1\supseteq\cdots\supseteq H_{s-1}\Big\}. \tag{62}$$

The next observation will be used to bound the factors in (62).

PROPOSITION A.6. *[9, Proposition 3.8] Suppose the $\Re$-valued random variable Y with $\mathbb{E}[Y]=0$ satisfies*

$$|Y|\le c$$

and

$$\mathbb{E}[|Y|]\le M.$$

Then for $|\zeta|c\le 1$,

$$\mathbb{E}[e^{\zeta Y}]\le e^{8\zeta^2 Mc}.$$

Let $H_i=\{X_0=x_0,X_1=x_1,\dots,X_i=x_i\}$ (as in Lemma A.5). Applying Proposition A.6 to each $Z_i\mid H_{i-1}$, with $c=L/\lambda^2$ (see (9)) and $M=M_i:=\mathbb{E}[|Z_i\mid H_{i-1}|]$, gives

$$\prod_{i=1}^s\mathbb{E}[e^{\zeta Z_i}\mid H_{i-1}]\le e^{8\zeta^2(L/\lambda^2)\sum M_i}\quad\text{for }|\zeta|c\le 1. \tag{63}$$

CLAIM A.7. $\sum_{i=1}^{s} M_i \le 2Lm/\lambda^2$.

We need the following easy observation. For $\underline{p} = (p_i)_{i=1}^{s}$ with $p_i \in [0, 1)$, let

$$f(\underline{p}) = \sum_{i=1}^{s} p_i \prod_{j<i} (1-p_i)^{-1}, \qquad g(\underline{p}) = \prod_{j=1}^{s} (1-p_i)^{-1}.$$

PROPOSITION A.8. $f(\underline{p}) \le g(\underline{p}) - 1$.

PROOF. We prove the equivalent

$$\prod_{i=1}^{s} (1-p_i) + \sum_{i=1}^{s} p_i \prod_{j \ge i} (1-p_i) \le 1$$

by induction on $s \ge 1$. The base case is obvious, and for the induction step we just observe that the left hand side is

$$(1-p_1)^2 \prod_{i=2}^{s} (1-p_i) + \sum_{i=2}^{s} p_i \prod_{j \ge i} (1-p_i) \le \prod_{i=2}^{s} (1-p_i) + \sum_{i=2}^{s} p_i \prod_{j \ge i} (1-p_i) \le 1. \qquad \square$$

PROOF OF CLAIM A.7. With sums over v's (in W) with $T_v, r_v > i-1$ (cf. the discussion following (9)), we have

$$M_i = \mathbb{E}|\sum (\xi_i^v - \xi_{i-1}^v)| \le \sum \mathbb{E}|\xi_i^v - \xi_{i-1}^v| = \sum 2\varphi(x_{i-1}, v)\xi_{i-1}^v. \tag{64}$$

Thus, using Proposition A.8 for (65), we have

$$\begin{aligned} \sum_{i=1}^{s} M_i &\le 2 \sum_{i=1}^{s} \sum_{v} \varphi(x_{i-1}, v)\xi_{i-1}^v \\ &= 2 \sum_{v} \sum \{\varphi(x_{i-1}, v)\xi_{i-1}^v : i-1 < T_v \wedge r_v\} \\ &\le 2 \sum_{v} (H_v(T_v \wedge r_v)^{-1} - 1) \qquad (65) \\ &\le 2mL/\lambda^2. \end{aligned} \qquad \square$$

PROOF OF THEOREM A.4. Lemma A.5, with (63) and Claim A.7, gives

$$\mathbb{E}[e^{\zeta Z}] \le e^{16L^2 m\zeta^2/\lambda^4}$$

whenever $|\zeta| \le \lambda^2/L$. So for any $\zeta \in (0, \lambda^2/L]$,

$$\mathbb{P}(Z > \vartheta m) = \mathbb{P}(e^{\zeta Z} > e^{\zeta\vartheta m}) \le \exp[16L^2 m\zeta^2/\lambda^4 - \zeta\vartheta m].$$

Since $\xi_s < (L/\lambda)m$, we may assume $\vartheta \le L/\lambda$ (or the theorem is trivial). Setting $\zeta = \vartheta\lambda^4/(32L^2)$ to minimize the exponent, we have

$$\mathbb{P}(Z > \vartheta m) \le \exp[-\vartheta^2\lambda^4 m/(8L)^2].$$

Similarly

$$\mathbb{P}(Z < -\vartheta m) \le \exp[-\vartheta^2\lambda^4 m/(8L)^2],$$

completing the proof. $\square$

Acknowledgments. We thank Bhargav Narayanan for helpful conversations and Ori Gurel-Gurevich for telling us the problem, long ago.

Funding. J.K. was supported by NSF Grant DMS1954035.

REFERENCES

[1] D. Aldous, Lower bounds for covering times for reversible Markov chains and random walks on graphs, *J Theor. Probab.* **2** (1989), 91–100.

[2] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, Unfinished monograph, available at http://www.stat.berkeley.edu/ aldous/RWG/book.html 2002.

[3] I. Benjamini, personal communication.

[4] I. Benjamini, O. Gurel-Gurevich, and B. Morris, Linear cover time is exponentially unlikely, *Probab. Theory Relat. Fields* **155** (2013), 451–461.

[5] B. Bollobás, *Modern Graph Theory*, Springer-Verlag, New York, 1998.

[6] J. Ding, Asymptotics of cover times via Gaussian free fields: Bounded-degree graphs and general trees, *Ann. Probab.* **42** (2014), 464–496.

[7] J. Ding, J. Lee and Y. Peres, Cover times, blanket times, and majorizing measures, *Ann. Math.* **175** (2012), 1409-1471.

[8] D. Gillman, A Chernoff bound for random walks on expander graphs, *SIAM J. Comput.* **27** (1998), 1203–1220.

[9] J. Kahn, Asymptotically good list-colorings, *J. Combin. Theory Ser. A* **73** (1996), 1-59.

[10] D. Levin, Y. Peres, and E. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, Providence, 2017. With a chapter by James G. Propp and David B. Wilson.

[11] L. Lovász, Random walks on graphs: a survey, *Combinatorics, Paul Erdős is eighty* **2** (1993), 1-46.

[12] S. Karlin and H. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1975.

[13] A. Yehudayoff, Linear cover time for trees is exponentially unlikely, *Chic. J. Theor. Comput. Sci.* **2012** (2012).

[14] A. Zhai, Exponential concentration of cover times, *Electron. J. Probab.* **23** (2018), Paper No. 32, 22 pp.