A Spectral Characterization for Concentration of the Cover Time

Jonathan Hermon[1] ⓘ

Abstract

We prove that for a sequence of finite vertex-transitive graphs of increasing sizes, the cover times are asymptotically concentrated if and only if the product of the spectral gap and the expected cover time diverges. In fact, we prove this for general reversible Markov chains under the much weaker assumption (than transitivity) that the maximal hitting time of a state is of the same order as the average hitting time.

Keywords Mixing times · Hitting times · Cover times · Vertex-transitive graphs · Spectral gap

Mathematics Subject Classification (2010) 60J10 · 60J27

1 Introduction

A big part of the modern theory of Markov chains is dedicated to the study of the hierarchy of different quantities associated with a Markov chain. It is a common theme that certain phenomena can be characterized by a simple criterion concerning whether or not two such quantities are strongly separated (i.e., are of strictly different orders). Often, one of these quantities is the inverse of the spectral gap. One instance is the cutoff phenomenon and the condition that the product of the mixing time and the spectral gap diverges, known as the *product condition* (a necessary condition for precutoff in total variation [32, Proposition 18.4] and a necessary and sufficient condition for cutoff in L_2 [14]). The condition that the product of the spectral gap and the maximal (expected) hitting time diverges is studied in [1] and [28, Theorem 1]).

Aldous' classic criterion for concentration of the cover time [4] is another such instance. Aldous' criterion asserts that for a sequence of Markov chains on finite state

✉ Jonathan Hermon
jonathan.hermon@statslab.cam.ac.uk

1   University of Cambridge, Cambridge, UK

spaces of diverging sizes $\frac{\tau_{cov}^{(n)}}{t_{cov}^{(n)}} \to 1$ in distribution if $t_{cov}^{(n)}/t_{hit}^{(n)} \to \infty$, where throughout the superscript '(n)' indicates that we are considering the nth Markov chain in the sequence, and where $\tau_{cov} = \inf\{t : \{X_s : s \leqslant t\} = V\}$ is the *cover time* of a Markov chain $(X_t)_{t \geqslant 0}$ on a finite state space V, defined to be the first time by which every state was visited at least once by the chain,

$$t_{cov} := \max_{x \in V} \mathbb{E}_x[\tau_{cov}]$$

its worst-case expectation and

$$t_{hit} := \max_{x,y} \mathbb{E}_x[T_y], \qquad \text{where} \qquad T_y := \inf\{t : X_t = y\},$$

is the maximal expected *hitting time* of a state. More precisely, Aldous [4, Proposition 1] showed that in the reversible case if $t_{hit}^{(n)} = \Omega(t_{cov}^{(n)})$, then there exists a sequence of initial states such that $\tau_{cov}^{(n)}/t_{cov}^{(n)}$ does not concentrate around any value.[1,2] Conversely (even without reversibility) $t_{cov}^{(n)} \leqslant \min_y \mathbb{E}_y[\tau_{cov}^{(n)}] + t_{hit}^{(n)}$ and when $t_{cov}^{(n)}/t_{hit}^{(n)} \to \infty$, we have that $\frac{t_{cov}^{(n)}}{\min_y \mathbb{E}_y[\tau_{cov}^{(n)}]} \to 1$ and that $\frac{\tau_{cov}^{(n)}}{t_{cov}^{(n)}} \to 1$ in distribution for every sequence of initial states.

Our Theorem 1 refines Aldous' criterion in the transitive setup by allowing one to replace the maximal hitting time in his result by the inverse of the spectral gap, which is positioned much lower in the aforementioned hierarchy of Markov chain parameters [see (3.10)]. Throughout, let gap $:= \lambda_2$ be the *spectral gap* of the considered chain, and $t_{rel} := \frac{1}{\lambda_2}$ its *relaxation time*, where $0 = \lambda_1 < \lambda_2 \leqslant \ldots \lambda_{|V|} \leqslant 2$ are the eigenvalues of the Laplacian $I - P$. When considering simple random walk (SRW) on a graph G, we often add parenthesis '(G)' to various quantities.

Theorem 1 *Let G_n be a sequence of finite connected vertex-transitive graphs of diverging sizes. Then $\frac{\tau_{cov}(G_n)}{t_{cov}(G_n)} \to 1$ in distribution iff $\mathrm{gap}(G_n)t_{cov}(G_n) \to \infty$.*

We note that in the setup of Theorem 1 if $\mathrm{gap}(G_n)t_{cov}(G_n) = O(1)$, then $\tau_{cov}^{(n)}/t_{cov}^{(n)}$ does not concentrate around any fixed value (by transitivity this holds for all initial states).[3]

Theorem 1 holds in the more general setup of reversible *transitive Markov chains*. That is, reversible Markov chains on a finite state space V whose transition matrix satisfies that for every $x, y \in V$ there is a bijection $f : V \to V$ such that $f(x) = y$ and $P(x, z) = P(y, f(z))$ for all $z \in V$. Theorem 2 extends Theorem 1 to a much larger class of Markov chains. Denote the *average hitting time* of an irreducible Markov chain on a finite state space V by

---

[1] The proposition is phrased for simple random walk, but the proof works for general reversible Markov chains. For a non-reversible counterexample, consider a walk on the cycle with a fixed clockwise bias.

[2] In fact, by [32, p. 274] starting from the stationary distribution π, one has that $\mathbb{P}_\pi[T_y > t_{hit}/4] \geqslant 1/4$ for some state y. By (3.4), $\mathbb{P}_\pi[T_y > it_{hit}] \geqslant (\mathbb{P}_\pi[T_y > t_{hit}/4])^{4i} \geqslant 2^{-8i}$.

[3] In fact, it is shown in [7] that for reversible chains there is always some state x and a set A of stationary probability at least $1/2$ such that $\mathrm{P}_x[T_A > t] \geqslant \exp(-\mathrm{gap}\,t)$ for all $t \geqslant 0$, where $T_A := \inf\{t : X_t \in A\}$.

$$\alpha := \sum_{x,y \in V} \pi(x)\pi(y)\mathbb{E}_x[T_y] = \sum_{y \in V} \pi(y)\alpha_y,$$

where throughout π denotes the stationary distribution and $\alpha_y := \mathbb{E}_\pi[T_y]$. Theorem 2 indeed generalizes Theorem 1, as (by Fact 3.1) for a transitive chain $\alpha \leqslant t_{\text{hit}} \leqslant 2\alpha$.

Theorem 2 *Consider a sequence of irreducible reversible Markov chains with finite state spaces $V^{(n)}$ and stationary distributions $\pi^{(n)}$. If[4] $\alpha^{(n)} \asymp t_{\text{hit}}^{(n)}$, we have that $\frac{\tau_{\text{cov}}^{(n)}}{t_{\text{cov}}^{(n)}} \to 1$ in distribution for every sequence of initial states if and only if $\text{gap}^{(n)} t_{\text{cov}}^{(n)} \to \infty$. Moreover, if $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\text{cov}}^{(n)}$, then*

$$\min_{v \in V^{(n)}} \max_{x \in V^{(n)}} \inf_{C \geqslant 1} -\frac{1}{C} \log P_v\left[T_x > C t_{\text{cov}}^{(n)}\right] \lesssim 1, \qquad (1.1)$$

and $\tau_{\text{cov}}^{(n)}/t_{\text{cov}}^{(n)}$ does not concentrate around any fixed value for any sequence of initial states.

As (3.2) $\frac{1}{\text{gap}} \leqslant \alpha \leqslant t_{\text{hit}} \leqslant t_{\text{cov}}$, by Theorem 2 $\frac{1}{\text{gap}^{(n)}} \asymp t_{\text{cov}}^{(n)}$ if and only if $\alpha^{(n)} \asymp t_{\text{cov}}^{(n)}$.

Remark 1.1 The condition $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\text{cov}}^{(n)}$ above (1.1) is implied by $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\text{hit}}^{(n)}$ in conjunction with $\text{gap}^{(n)} t_{\text{cov}}^{(n)} \asymp 1$ [see (3.10)]. Of course, $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\text{hit}}^{(n)}$ is stronger than the condition $\alpha^{(n)} \asymp t_{\text{hit}}^{(n)}$ used in the first part of the statement of Theorem 2. We believe that the condition $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\text{cov}}^{(n)}$ in (1.1) can be relaxed to $\alpha^{(n)} \asymp t_{\text{cov}}^{(n)}$.

Throughout, we work with the continuous-time rate 1 version of the chain. We remark that all our results are valid also in discrete time even if the chain is not lazy (i.e., if $P(x, x) = 0$ for some x). Moreover, in this case one *does not* need to replace gap by the absolute spectral gap (as is often the case when translating a result from the continuous-time or discrete-time lazy setups to the discrete-time non-lazy setup). This can be verified by an application of Wald's equation (used to argue that the expected cover time and hitting times are the same in both setups), together with the fact that Aldous' result [4] (which is used in the proofs of Theorems 1.1 and 1.2) applies to both setups.

We note that if G_n and $\widehat{G}_n$ are two sequences of finite connected graphs of uniformly bounded degree which are uniformly quasi-isometric (i.e., there exists some $K > 0$ such that G_n is K-quasi-isometric to $\widehat{G}_n$ for all n), then $\text{gap}(G_n) \asymp \text{gap}(\widehat{G}_n), \alpha(G_n) \asymp \alpha(\widehat{G}_n), t_{\text{hit}}(G_n) \asymp t_{\text{hit}}(\widehat{G}_n)$ and [22, Theorem 1.6] $t_{\text{cov}}(G_n) \asymp t_{\text{cov}}(\widehat{G}_n)$.[5] In particular,

---

[4] We write $o(1)$ for terms which vanish as $n \to \infty$. We write $f_n = o(g_n)$ or $f_n \ll g_n$ if $f_n/g_n = o(1)$. We write $f_n = O(g_n)$ and $f_n \lesssim g_n$ (and also $g_n = \Omega(f_n)$ and $g_n \gtrsim f_n$) if there exists a constant $C > 0$ such that $|f_n| \leqslant C|g_n|$ for all n. We write $f_n = \Theta(g_n)$ or $f_n \asymp g_n$ if $f_n = O(g_n)$ and $g_n = O(f_n)$.

[5] The fact that $\alpha(G_n) \asymp \alpha(\widehat{G}_n)$ follows from (3.2) via a standard comparison argument [19]. The claim that $t_{\text{hit}}(G_n) \asymp t_{\text{hit}}(\widehat{G}_n)$ can be seen from the commute-time identity (e.g., [32, Eq. (10.14)]) combined with the robustness of the effective-resistance under quasi-isometries (cf. the proof of Theorem 2.17 in [34]).

if $\widehat{G}_n$ is vertex-transitive, the sequence of SRWs on G_n satisfies the conditions of Theorem 2 [apart perhaps from $\min_{x \in V^{(n)}} \alpha_x \asymp t_{\mathrm{cov}}(G_n)$, used only in (1.1)].

The cover time of an $n \times n$ grid torus is concentrated [18], while that of the n-cycle is not. The following example shows that an $n \times \lceil n/\log^2 n \rceil$ grid torus is in some sense critical.

Example 1.1 Consider an $n \times m$ discrete (grid) torus (i.e., the Cayley graph of $\mathbb{Z}_n \times \mathbb{Z}_m$ w.r.t. the standard choice of generators). If $m = m(n) = O(n/\log^2 n)$, then its (expected) cover time is of order n^2, same as the inverse of its spectral gap. Conversely, if $n/\log^2 n \ll m \leqslant n$, the cover time is of order $mn(\log n)^2 \gg n^2$, while the spectral gap is $\Theta(\frac{1}{n^2})$.

Theorem 1 is a fairly immediate consequence of the following result (see Sect. 4 for the details).

Proposition 1.1 *For every transitive Markov chain on a finite state space V,*

$$t_{\mathrm{cov}} \geqslant \tfrac{1}{4} t_{\mathrm{hit}} \log\left(t_{\mathrm{hit}}\,\mathrm{gap}/12\right). \tag{1.2}$$

We note that whenever $\frac{1}{\mathrm{gap}} \lesssim |V|^a$ for some $a \in (0, 1)$ the bound offered by (1.2) is of the correct order, as by Matthews [35] $t_{\mathrm{cov}} \leqslant t_{\mathrm{hit}} \sum_{i=1}^{|V|-1} \frac{1}{i} \leqslant t_{\mathrm{hit}}(\log|V| + 1)$ (e.g., [32, Theorem 11.2]) and by (3.2) $t_{\mathrm{hit}} \geqslant \alpha \geqslant \frac{|V|}{2}$ (and so gap $t_{\mathrm{hit}} \gtrsim |V|^{1-a}$).

Theorem 2 (apart from (1.1)) is a fairly immediate consequence of the following extension of Proposition 1.1.

Proposition 1.2 *For every irreducible reversible Markov chain on a finite state space V with a stationary distribution π, we have that*

$$t_{\mathrm{cov}} \geqslant \frac{\alpha}{4} \log\left(M\left(t_{\mathrm{hit}}/\alpha, \alpha\,\mathrm{gap}\right)\right), \quad where \quad M(a, b) = \tfrac{8\log(8a)+b}{32a\log(8a)} \tag{1.3}$$

When $\alpha \asymp t_{\mathrm{hit}}$ and $t_{\mathrm{hit}} \gg 1/\mathrm{gap}$, we get that $M(\frac{t_{\mathrm{hit}}}{\alpha}, \alpha\,\mathrm{gap}) \gg 1$ and so $t_{\mathrm{cov}} \gg t_{\mathrm{hit}}$.

1.1 Related Work

Cover times have been studied extensively for over 35 years. This is a topic with rich ties to other objects such as the Gaussian free field [22]. There has been many works providing general bounds on the cover time, studying its evolution and its fluctuations in general [6,20,31,36,41], and in particular for the giant component of various random graphs [16], for trees [2,23], for the two-dimensional torus [9,15,18,21] and for higher-dimensional tori [8]. Feige [24,25] proved tight extremal upper and lower bounds on cover times of graphs (by SRW). For a more comprehensive review of the literature, see the related work section in [22] and the references therein. For further background on hitting times and cover times, see [1,5,32].

1.2 Organization of This Note

In Sect. 2, we present some open problems. In Sect. 3, we present some background on hitting times. In Sect. 4, we prove Theorems 1–2 and Propositions 1.1 and 1.2. Example 1.1 is analyzed in Sect. 5.

2 Open Problems

In the following seven questions, let $G_n = (V_n, E_n)$ be a sequence of finite connected vertex-transitive graphs of diverging sizes. Denote the degree of G_n by d_n and its diameter (i.e., the maximal graph distance between a pair of vertices) by $\mathrm{Diam}(G_n)$. Let $t_{\mathrm{rel}}(G_n) := \frac{1}{\mathrm{gap}(G_n)}$ be the *relaxation time*. The following two questions are the focus of an ongoing work with Nathanaël Berestycki (we believe both have an affirmative answer).

Question 2.1 *Assume that* $t_{\mathrm{rel}}(G_n) \ll t_{\mathrm{hit}}(G_n)$*. Is it the case that* $\frac{t_{\mathrm{cov}}(G_n)}{t_{\mathrm{hit}}(G_n) \log |V_n|} \to 1$*?*

Question 2.2 *Assume that* $\mathrm{Diam}(G_n)^2 \ll |V_n|$ *and* $d_n \asymp 1$*. Does* $\frac{t_{\mathrm{cov}}(G_n)}{t_{\mathrm{hit}}(G_n) \log |V_n|} \to 1$*?*

We now discuss some relaxations of the conditions from Questions 2.1 and 2.2. Let $\mathcal{R}(a, b)$ be the *effective resistance* between a and b and $\mathcal{R}_* := \max_{x,y} \mathcal{R}(x, y)$ (e.g., [32, Ch. 9] and [34, Ch. 2]). Let $o_n \in V_n$. Consider the conditions:

  (i) $\mathcal{R}_*(G_n) \asymp 1/d_n$ (where $\mathcal{R}_*(G)$ is $\mathcal{R}_*$ for SRW on G),
 (ii) $|\{x \in V_n : \mathcal{R}(o_n, x) \leqslant \delta \mathcal{R}_*(G_n)\}| = O(1)$ for all fixed $\delta \in (0, 1)$.

Question 2.3 *Does* $t_{\mathrm{rel}}(G_n) \log |V_n| \lesssim t_{\mathrm{hit}}(G_n)$ *imply (i)?*

Question 2.4 *Does* $t_{\mathrm{rel}}(G_n) \log |V_n| \ll t_{\mathrm{hit}}(G_n)$ *imply (ii) when* $d_n \asymp 1$*?*

Condition (i) arises in forthcoming work of Tessera and Tointon [39] as the analog of transience in a Varopoulos-type result for finite vertex-transitive graphs. In particular, they show (in the transitive setup) that it follows from the condition $\mathrm{Diam}(G_n)^2 \lesssim \frac{|V_n|}{\log |V_n|}$. Below we give some equivalent conditions to (i). One of them is $t_{\mathrm{hit}}(G_n) \asymp |V_n|$. Using the fact that $t_{\mathrm{rel}}(G) \leqslant 2d\mathrm{Diam}(G)^2$ for a vertex-transitive graph G of degree d [32, Theorem 13.26] (as well as $t_{\mathrm{hit}}(G_n) \gtrsim |V_n|$), it follows that when $d_n \asymp 1$,

$$\begin{aligned} \mathrm{Diam}(G_n)^2 \log |V_n| \lesssim |V_n| &\implies t_{\mathrm{rel}}(G_n) \log |V_n| \lesssim t_{\mathrm{hit}}(G_n), \text{ and likewise} \\ \mathrm{Diam}(G_n)^2 \log |V_n| \ll |V_n| &\implies t_{\mathrm{rel}}(G_n) \log |V_n| \ll t_{\mathrm{hit}}(G_n). \end{aligned} \quad (2.1)$$

Moderate growth is a certain technical growth condition introduced by Diaconis and Saloff-Coste in their seminal work [17]. For Cayley graphs, this condition is shown by Breuillard and Tointon [13] to be equivalent to the condition $c|V| \leqslant \mathrm{Diam}(G)^a$, in some precise quantitative sense, with these a and c being related to the parameters in the definition in [17]. This was recently extended to vertex-transitive graphs by Tessera and Tointon [40].

Using the fact that for vertex-transitive graphs of moderate growth $t_{\mathrm{rel}}(G_n) \gtrsim \mathrm{Diam}(G_n)^2$ [17],[6] it appears that the main ingredient for establishing the converse implications to the ones in (2.1) when $d_n \asymp 1$ is providing an affirmative answer to Question 2.3. Indeed for graphs of sufficiently large growth, the condition $\mathrm{Diam}(G_n)^2 \ll \frac{|V_n|}{\log |V_n|}$ holds for free.

Question 2.5 *Do the reverse implications in* (2.1) *hold? What can be said without the assumption that* $d_n \asymp 1$?

Question 2.6 *Assume that* $\mathrm{Diam}(G_n)^2 \ll \frac{|V_n|}{\log |V_n|}$. *Is it the case that for all fixed* $\delta \in (0, 1)$, $|\{x \in V_n : \mathcal{R}(o_n, x) \leqslant \delta \mathcal{R}_*(G_n)\}| = O(d_n)$?

For vertex-transitive graphs, condition (i) is equivalent to $t_{\mathrm{hit}}(G_n) \asymp |V_n|$ and condition (ii) is equivalent to the condition that $|\{x \in V_n : \mathbb{E}_{o_n}[T_x] \leqslant \delta t_{\mathrm{hit}}(G_n)\}| = O(1)$ for all fixed $\delta \in (0, 1)$. Indeed, by the commute-time identity (e.g., [32, Proposition 10.7]), for SRW on a graph $G = (V, E)$, we have that

$$\tfrac{1}{2}(\mathbb{E}_o[T_x] + \mathbb{E}_x[T_o]) = |E|\mathcal{R}(o, x). \tag{2.2}$$

If $G = (V, E)$ is also vertex-transitive, then $\mathbb{E}_o[T_x] = \mathbb{E}_x[T_o]$ [1, Proposition 2] (see also [32, Proposition 10.10], and this also follows from (3.5)), and so $\mathbb{E}_o[T_x] = |E|\mathcal{R}(o, x)$. Hence,

$$\mathcal{B}_{\mathrm{eff-res}}(o, \delta\mathcal{R}_*(G)) := \{x \in V : \mathcal{R}(o, x) \leqslant \delta\mathcal{R}_*(G)\} = \{x : \mathbb{E}_o[T_x] \leqslant \delta t_{\mathrm{hit}}(G)\},$$

and (even without transitivity) $t_{\mathrm{hit}}(G_n) \asymp |V_n|$ iff $\mathcal{R}_*(G_n) \asymp 1/d_n$.

Question 2.7 *In the above setup, is it the case that* $t_{\mathrm{hit}}(G_n) \asymp |V_n|$ *if and only if* $t_{\mathrm{cov}}(G_n) \asymp |V_n| \log |V_n|$?

The implication $t_{\mathrm{hit}}(G_n) \asymp |V_n| \implies t_{\mathrm{cov}}(G) \asymp |V_n| \log |V_n|$ follows (even without transitivity) from (3.10) and the bound $t_{\mathrm{cov}}(G) \geqslant (1 - o(1))|V| \log |V|$ [24] holding for every $G = (V, E)$.

It is not clear what is the correct analog of transience for a sequence of finite graphs. One technical definition is given in [36]. Here we discuss two different conditions. Our goal is to motivate Questions 2.3–2.5, relate them back to τ_{cov} and stimulate future research.

A natural informal definition of being *uniformly locally transient* for a sequence $G_n = (V_n, E_n)$ of d_n-regular graphs is that the expected number of returns to the origin of the walk by the mixing time (or by time $|V_n|$) is $O(1)$ (uniformly in the choice of the origin). If we think of "mixing" as "reaching infinity," then this is a natural analog of transience in the infinite setup. It is hence natural that such a condition can be phrased in terms of effective resistance. To make this precise, one can consider the following equivalent conditions:

- $\mathcal{R}_*(G_n) \asymp 1/d_n$ (as discussed above, this is equivalent to $t_{\mathrm{hit}} \asymp |V_n|$).

---

[6] See also [? §8.1], where it is noted that argument from [17] is valid for vertex-transitive graphs of moderate growth, not just for Cayley graphs.

- $\min_{x,y\in V_n: x\neq y} P_x[T_x^+ > T_y] \gtrsim 1$, where $T_x^+ := \inf\{t : X_t = x, X_s \neq x \text{ for some } s < t\}$.
- $\max_{x,y\in V_n} \int_0^{t_{\mathrm{rel}}(G_n)} H_t^{G_n}(x,x)\mathrm{d}t = O(1)$ (where $H_t^{G_n}(\cdot, \cdot\cdot)$ are the time t transition probabilities for SRW on G_n; equivalently, $\max_{x\in V_n} \int_0^{t_{\mathrm{rel}}(G_n)} H_t^{G_n}(x,y)\mathrm{d}t = O(1)$).
- $\max_{x,y\in V_n} \int_0^{|V_n|\vee t_{\mathrm{mix}}^{(\infty)}(G_n)} H_t^{G_n}(x,y)\mathrm{d}t = O(1)$, where $a \vee b := \max\{a,b\}$ and $t_{\mathrm{mix}}^{(\infty)} := \inf\{t : \max_{a,b}|\frac{H_t(a,b)}{\pi(b)} - 1| \leqslant 1/2\}$ is the L_∞ mixing time.

Consider a sequence of finite connected graphs $G_n := (V_n, E_n)$. Let $d_{\max}^{(n)}$ and $d_{\min}^{(n)}$ be the maximal and minimal (respectively) degree of G_n. We assume that $d_{\max}^{(n)} \asymp d_{\min}^{(n)}$ so that π_{G_n} the stationary distribution of SRW on G_n satisfies $\max_{v\in V_n} \pi_{G_n}(v) \asymp \frac{1}{|V_n|} \asymp \min_{v\in V_n} \pi_{G_n}(v)$. A natural informal definition for saying that G_n is *uniformly globally transient* is that the walk either returns to the origin rapidly (i.e., in $O(1)$ time units) or else it is unlikely to return before getting mixed. To make this precise, we consider the condition

$$\text{condition (a):} \qquad \lim_{s\to\infty} \limsup_{n\to\infty} \max_{x,y\in V_n} \int_{s\wedge t_{\mathrm{rel}}(G_n)}^{t_{\mathrm{rel}}(G_n)} H_t^{G_n}(x,y)\mathrm{d}t = 0.$$

In fact, condition (a) is equivalent to (b): $\limsup_{n\to\infty} \max_{x\in V_n} \int_{s\wedge t_{\mathrm{rel}}(G_n)}^{t_{\mathrm{rel}}(G_n)} H_t^{G_n}(x,x)\mathrm{d}t \to 0$, as $s \to \infty$, as well as to (c): $\limsup_{n\to\infty} \max_{x\in V_n} \int_{s\wedge t_{\mathrm{mix}}^{(\infty)}(G_n)}^{t_{\mathrm{mix}}^{(\infty)}(G_n)} H_t^{G_n}(x,x)\mathrm{d}t \to 0$, as $s \to \infty$ (provided $d_{\max}^{(n)} \asymp d_{\min}^{(n)}$). These conditions imply that

(1) $\mathcal{R}_*(G_n) \asymp 1/d_{\max}^{(n)}$ (recall that we assume $d_{\max}^{(n)} \asymp d_{\min}^{(n)}$),
(2) $\max_{o\in V_n} |\{x \in V_n : \mathbb{E}_o[T_x] \leqslant \mathbb{E}_{\pi_{G_n}}[T_y] - \delta|V_n|\}| = O(1)$ for all fixed $\delta > 0$,
(3) $t_{\mathrm{rel}}(G_n) = o(t_{\mathrm{hit}}(G_n))$.

When G_n is vertex-transitive, conditions (1)–(3) are equivalent to condition (a) and also to conditions (1), (2') and (3), where condition (2') is: $|\mathcal{B}_{\mathrm{eff-res}}(o, \delta\mathcal{R}_*(G_n))| = O(1)$, for all fixed $\delta \in (0, 1)$ (this is condition (ii) above). While we omit the proof of these equivalence, we stress that some of them are not at all obvious.

We strongly believe that (in the transitive setup) these conditions imply that $\frac{t_{\mathrm{cov}}(G_n)}{t_{\mathrm{hit}}(G_n)\log|V_n|} \to 1$ and that the distribution of the cover time exhibits Gumbel fluctuations.

Considering an $n \times n$ discrete torus shows that even under transitivity the condition $t_{\mathrm{rel}}(G_n)\log|V_n| \asymp t_{\mathrm{hit}}(G_n)$ does not imply that $\mathcal{R}_*(G_n) \asymp 1$. Another illustrative example is an $n \times n \times f(n)$ discrete torus. It is not hard to verify that when $f(n) = \lceil\log n\rceil$, we have that $t_{\mathrm{rel}}(G_n)\log|V_n| \asymp t_{\mathrm{hit}}(G_n)$ and $\mathcal{R}_*(G_n) \asymp 1$. However for every $\delta \in (0, 1)$, it holds that $|\{x \in V_n : \mathcal{R}(o_n, x) \leqslant \delta\mathcal{R}_*(G_n)\}| = \Omega_\delta(n^{2c\delta}\log n)$ for some absolute constant $c \in (0, 1)$, and so condition (ii) fails. Conversely, if $\log n \ll f(n) \lesssim n$, then $t_{\mathrm{rel}}(G_n)\log|V_n| \ll t_{\mathrm{hit}}(G_n)$ and conditions (i) and (ii) hold.

Considering a Cartesian product of the n-cycle with a vertex-transitive expander of size $f(n) \gg n$ shows that the condition $t_{\mathrm{rel}}(G_n)\log|V_n| \ll t_{\mathrm{hit}}(G_n)$ above is not a necessary condition for conditions (i) and (ii) to hold.

In light of Example 1.1, the following question naturally arises.

Question 2.8 *Let G_n be a sequence of finite connected vertex-transitive graphs of diverging sizes and uniformly bounded degrees. Assume that along every subsequence $\frac{\tau_{\mathrm{cov}}(G_n)}{t_{\mathrm{cov}}(G_n)}$ does not converge to 1 in distribution. Is it the case that when viewing G_n as a metric space with the graph distance as its metric, after rescaling distances by a $\frac{\mathrm{Diam}(G_n)}{f(\mathrm{Diam}(G_n))}$ factor, for every $f : \mathbb{N} \to \mathbb{R}_+$ satisfying $1 \ll f(k) = o((\log k)^2)$, the pointed Gromov–Hausdorff scaling limit exists and is $\mathbb{R}$?*

This question is the cover time analog of a question from [12][7], where it is shown that for a sequence of finite vertex-transitive graphs G_n of fixed degree and increasing sizes satisfying that their mixing times are proportional to their maximal hitting times, G_n rescaled by their diameters converges in the Gromov–Hausdorff topology to the unit circle S^1.

3 Hitting-Times Preliminaries

Let $(X_t)_{t=0}^{\infty}$ be an irreducible reversible Markov chain on a finite state space V with transition matrix P and stationary distribution π. Denote the law of the continuous-time rate 1 version of the chain starting from vertex x (resp. initial distribution μ) by P_x (respectively, P_μ). Denote the corresponding expectation by $\mathbb{E}_x$ (respectively, $\mathbb{E}_\mu$). Let $H_t := e^{-t(I-P)}$ be its heat kernel (so that $H_t(\cdot, \cdot)$ are the time t transition probabilities).

We now present some background on hitting times. The random target identity (e.g., [32, Lemma 10.1]) asserts that $\sum_y \pi(y)\mathbb{E}_x[T_y]$ is independent of x and hence equals α, while for all $x \in V$ we have that (e.g., [32, Proposition 10.26])

$$\alpha_x := \mathbb{E}_\pi[T_x] = \frac{1}{\pi(x)} \sum_{i=0}^{\infty} \left(P^i(x, x) - \pi(x) \right) = \frac{1}{\pi(x)} \int_0^{\infty} (H_t(x, x) - \pi(x))\, \mathrm{d}t,$$

(3.1)

Averaging over x yields the eigentime identity ([5, Proposition 3.13])

$$\alpha = \sum_{x,y} \pi(x)\pi(y)\mathbb{E}_x[T_y] = \sum_y \sum_{i=0}^{\infty} (P^i(y, y) - \pi(y))$$

$$= \sum_{i=0}^{\infty} [\mathrm{Trace}(P^i) - 1] = \sum_{i \geqslant 2} \frac{1}{\lambda_i}.$$

(3.2)

For a transitive Markov chain $\mathbb{E}_x[T_y] = \mathbb{E}_y[T_x]$ for all x, y [1, Proposition 2] [see also (3.6)] and so

$$\forall x, \quad \alpha = \mathbb{E}_\pi[T_x].$$

---

[7] In the corresponding question from [12] the condition $t_{\mathrm{mix}}(G_n) \asymp t_{\mathrm{hit}}(G_n)$ replaces the condition on $\frac{\tau_{\mathrm{cov}}(G_n)}{t_{\mathrm{cov}}(G_n)}$ from Question 2.8. Moreover, in [12] we take to satisfy $1 \ll f(k) = o(\log k)$, instead of $1 \ll f(k) = o((\log k)^2)$ as in Question 2.8.

Let $U \sim \pi$ be independent of the chain. As $T_x \leqslant T_U + \inf\{t : X_{t+T_U} = x\}$, using the random target identity, as well as the strong Markov property yields:

Fact 3.1 [32, Lemma 10.2] $\max_x \alpha_x \leqslant t_{\text{hit}} \leqslant \alpha + \max_x \alpha_x \leqslant 2\max_x \alpha_x$.

The following material can be found at [5, §3.5]. Under reversibility, for any set A the law of its hitting time $T_A := \inf\{t : X_t \in A\}$ under initial distribution π conditioned on $A^{\complement}$ is a mixture of exponential distributions, whose minimal parameter $\lambda(A)$ is the Dirichlet eigenvalue of the set $A^{\complement}$. There exists a distribution μ_A, known as the *quasi-stationary distribution* of $A^{\complement}$, under which T_A has an Exponential distribution of parameter $\lambda(A)$. It follows that $\lambda(A) \geqslant \frac{1}{\max_a \mathbb{E}_a[T_A]}$. We see that for all $t \geqslant 0$,

$$\mathrm{P}_\pi[T_y > t] \leqslant \exp(-t/t_{\text{hit}}), \quad \text{and so} \quad \mathbb{E}_\pi[T_y^2] \leqslant 2t_{\text{hit}}^2. \tag{3.3}$$

Using the above description of the law of T_A, it is not hard to show [5, p. 86] that

$$\forall s, t \geqslant 0, \qquad \mathrm{P}_\pi[T_y > t + s \mid T_y \geqslant s] \geqslant \mathrm{P}_\pi[T_y > t]. \tag{3.4}$$

Let

$$Z_{x,y} := \int_0^\infty (H_t(x, y) - \pi(y))\, \mathrm{d}t.$$

By (3.1) $\pi(y)\alpha_y := Z_{y,y}$ for all y. We also have that for all $x, y \in V$ (e.g., [5, §2.2])

$$\alpha_y - \mathbb{E}_x[T_y] = Z_{x,y}/\pi(y) \stackrel{\text{(under reversibility)}}{=} Z_{y,x}/\pi(x) = \alpha_x - \mathbb{E}_y[T_x]. \tag{3.5}$$

Hence by (3.1), (this is justified in more detail below)

$$\mathbb{E}_x[T_y] = \frac{1}{\pi(y)} \int_0^\infty (H_t(y, y) - H_t(x, y))\, \mathrm{d}t. \tag{3.6}$$

Indeed, it follows from the spectral decomposition (e.g., [32, §12.1]) that for all x and all $s, t \geqslant 0$ we have that

$$0 < H_{t+s}(x, x) - \pi(x) \leqslant e^{-s/t_{\text{rel}}}(H_t(x, x) - \pi(x)). \tag{3.7}$$

By a standard application of reversibility and the Cauchy–Schwartz inequality (e.g., [5, Eq. (3.59)]), $|\frac{H_t(y,x)}{\pi(x)} - 1|^2 = |\sum_z \pi(z)\frac{H_{t/2}(y,z)-\pi(z)}{\pi(z)}\frac{H_{t/2}(x,z)-\pi(z)}{\pi(z)}|^2 \leqslant \left(\frac{H_t(x,x)}{\pi(x)} - 1\right)$ $\left(\frac{H_t(y,y)}{\pi(y)} - 1\right)$ (where we have used reversibility to get $\sum_z \pi(z)\left(\frac{H_s(a,z)-\pi(z)}{\pi(z)}\right)^2 = \frac{H_{2s}(a,a)}{\pi(a)} - 1$). Hence,

$$\left|\frac{H_t(y,x)}{\pi(x)} - 1\right| \leqslant \frac{1}{2}\left(\frac{H_t(x,x)}{\pi(x)} - 1\right) + \frac{1}{2}\left(\frac{H_t(y,y)}{\pi(y)} - 1\right). \tag{3.8}$$

Combining (3.7) and (3.8), we see that

$$\max_{a,b} \int_0^\infty \frac{|H_t(a,b) - \pi(b)|}{\pi(b)} dt = \max_a \int_0^\infty \frac{H_t(a,a) - \pi(a)}{\pi(a)} dt$$

$$\leqslant \max_a \int_0^\infty \frac{e^{-t/t_{\mathrm{rel}}}}{\pi(a)} dt < \infty.$$

Hence, $\int_0^\infty (H_t(y,y) - \pi(y)) dt - \int_0^\infty (H_t(x,y) - \pi(y)) dt = \int_0^\infty (H_t(y,y) - H_t(x,y)) dt$, and so (3.6) indeed follows from (3.1) and (3.5). Combining (3.7) and (3.8) also yields:

Lemma 3.1 *For every irreducible, reversible Markov chain on a finite state space V with a stationary distribution π, we have that*

$$\forall x, y \in V, s \geqslant 0, \qquad \int_s^\infty \left(\frac{H_t(y,x)}{\pi(x)} - 1 \right) dt \leqslant \tfrac{1}{2} e^{-s/t_{\mathrm{rel}}} (\alpha_x + \alpha_y). \quad (3.9)$$

Proof Using (3.8), (3.7) and (3.1) (in this order)

$$\int_s^\infty \left(\frac{H_t(y,x)}{\pi(x)} - 1 \right) dt \leqslant \tfrac{1}{2} \int_s^\infty \left(\frac{H_t(x,x)}{\pi(x)} - 1 \right) dt + \tfrac{1}{2} \int_s^\infty \left(\frac{H_t(y,y)}{\pi(y)} - 1 \right) dt$$

$$\leqslant \tfrac{1}{2} e^{-s/t_{\mathrm{rel}}} \int_0^\infty \left(\frac{H_t(x,x)}{\pi(x)} - 1 \right) dt + \tfrac{1}{2} e^{-s/t_{\mathrm{rel}}} \int_0^\infty \left(\frac{H_t(y,y)}{\pi(y)} - 1 \right) dt$$

$$= \tfrac{1}{2} e^{-s/t_{\mathrm{rel}}} (\alpha_x + \alpha_y).$$

$\square$

Recall that $t_{\mathrm{mix}}^{\mathrm{TV}} := \inf\{t : \max_a \sum_b |H_t(a,b) - \pi(b)| \leqslant 1/2\}$ and $t_{\mathrm{mix}}^{(\infty)} := \inf\{t : \max_{a,b} |\frac{H_t(a,b)}{\pi(b)} - 1| \leqslant 1/2\}$ are the *total variation* and the L_∞ *mixing times*. Under reversibility, the quantities considered above satisfy the following hierarchy (e.g., [32, Theorems 10.22 and 12.5]):

$$\tfrac{1}{9\lambda_2} \log 4 \leqslant \tfrac{1}{9} t_{\mathrm{mix}}^{\mathrm{TV}} \leqslant \tfrac{1}{9} t_{\mathrm{mix}}^{(\infty)} \leqslant t_{\mathrm{hit}} \leqslant t_{\mathrm{cov}} \leqslant t_{\mathrm{hit}}(\log |V| + 1), \quad (3.10)$$

where the last inequality is due to Matthews' [35] (see [32, Ch. 11] for a neat presentation). It is interesting to note that for reversible chains $t_{\mathrm{mix}}^{\mathrm{TV}} \leqslant C \min_x \max_y \mathbb{E}_y[T_x]$, for some absolute constant C. This follows from the results of Lovász and Winkler [33] concerning what they call the "forget time".[8] See equation (3.7) in [29] for a more elementary derivation.

---

[8] Under reversibility, it follows from their result that $t_{\mathrm{stop}} \leqslant C_1 t_{\mathrm{forgettime}} \leqslant C_1 \min_x \max_y \mathbb{E}_y[T_x]$, while Aldous [3] showed that $t_{\mathrm{mix}}^{\mathrm{TV}} \leqslant C_2 t_{\mathrm{stop}}$ (see also Peres and Sousi [38]), for some absolute constants C_1, C_2, where t_{stop} is the expectation of a mean optimal stopping rule (4.3) starting from the worst initial state.

4 Proof of Theorems 1–2 and Propositions 1.1 and 1.2

We will show that for every irreducible reversible Markov chain on a finite state space V,

$$\forall x \in V, \varepsilon \in (0, 1), \qquad \pi(B(x, \varepsilon)) \leqslant \frac{2 \log\left(\frac{2t_{\text{hit}}}{\varepsilon \alpha}\right)}{2 \log\left(\frac{2t_{\text{hit}}}{\varepsilon \alpha}\right) + \varepsilon \alpha \, \text{gap}}, \qquad (4.1)$$

where $B(x, \varepsilon) := \{z \in V : \mathbb{E}_z[T_x] \leqslant \alpha_x - \varepsilon \alpha \text{ or } \mathbb{E}_x[T_z] \leqslant \alpha_z - \varepsilon \alpha\}$, whereas if it is transitive,

$$\forall x \in V, \varepsilon \in (0, 1), \qquad |\{z : \mathbb{E}_z[T_x] \leqslant (1 - \varepsilon)\alpha\}| \leqslant \frac{2 \log(2/\varepsilon)}{2 \log(2/\varepsilon) + \varepsilon \alpha \, \text{gap}} |V|. \qquad (4.2)$$

We first prove (1.2) and (1.3) assuming (4.2) and (4.1), whose proofs are deferred to the end of the section.

Proof of (1.2) and (1.3) We first prove (1.2). By (3.2) $|V| \geqslant \alpha$ gap (this is used in the first inequality below). By Fact 3.1 $\alpha \geqslant \frac{1}{2} t_{\text{hit}}$. Using the fact that for transitive chains $\mathbb{E}_a[T_b] = \mathbb{E}_b[T_a]$ [1], by (4.2) with $\varepsilon = \frac{1}{2}$ there exists a set $B \subseteq V$ of size at least

$$\left\lceil |V| / \left(|V| \frac{4 \log 4}{4 \log 4 + \alpha \, \text{gap}}\right) \right\rceil \geqslant \frac{\alpha \, \text{gap}}{4 \log 4} \geqslant \frac{t_{\text{hit}} \, \text{gap}}{8 \log 4} \geqslant \frac{t_{\text{hit}} \, \text{gap}}{12},$$

such that for all $a, b \in B$ we have that $\mathbb{E}_a[T_b] \geqslant \alpha/2 \geqslant t_{\text{hit}}/4$. The claim now follows from Matthews' method [35] (see [32, Proposition 11.4]), which asserts that $t_{\text{cov}} \geqslant \min_{a,b \in B : a \neq b} \mathbb{E}_a[T_b] \sum_{i=1}^{|B|-1} \frac{1}{i}$ (using $\sum_{i=1}^{k-1} \frac{1}{i} \geqslant \log k$).

We now prove (1.3). We first use the Paley–Zygmund inequality to argue that

$$\pi(D) \geqslant \frac{\alpha^2}{4 \sum_z \pi(z) \alpha_z^2} \geqslant \frac{\alpha}{4 t_{\text{hit}}}, \qquad \text{where} \qquad D := \{z : \alpha_z \geqslant \alpha/2\}.$$

Hence by (4.1) with $\varepsilon = \frac{1}{4}$ there exists a subset B of D of stationary probability at least

$$\left\lceil \pi(D) / \left(\frac{8 \log\left(\frac{8t_{\text{hit}}}{\alpha}\right)}{8 \log\left(\frac{8t_{\text{hit}}}{\alpha}\right) + \text{gap} \cdot \alpha}\right) \right\rceil,$$

such that for all $a, b \in B$ we have that $\mathbb{E}_a[T_b] \geqslant \alpha_b - \alpha/4 \geqslant \alpha/4$ (using the fact that $\alpha_x \geqslant \alpha/2$ for all $x \in D$). The proof is concluded as above using Matthews' method. □

We now prove Theorems 1 and 2.

Proof of Theorems 1 and 2 without (1.1) Recall that Aldous [4] showed that in the reversible setup $\frac{\tau_{\text{cov}}^{(n)}}{t_{\text{cov}}^{(n)}} \to 1$ in distribution for all sequences of initial states iff $t_{\text{hit}}^{(n)} \ll t_{\text{cov}}^{(n)}$.

By (1.2) in the transitive setup, and by (1.3) in the setup of Theorem 2, this occurs iff $\mathrm{gap}^{(n)} t_{\mathrm{cov}}^{(n)} \to \infty$. Indeed, the condition $\mathrm{gap}^{(n)} t_{\mathrm{cov}}^{(n)} \to \infty$ is necessary for $t_{\mathrm{hit}}^{(n)} \ll t_{\mathrm{cov}}^{(n)}$ by (3.10). Conversely, if $\mathrm{gap}^{(n)} t_{\mathrm{cov}}^{(n)} \to \infty$ and along a subsequence $t_{\mathrm{hit}}^{(n)} \asymp t_{\mathrm{cov}}^{(n)}$, then along this subsequence $\mathrm{gap}^{(n)} t_{\mathrm{hit}}^{(n)} \to \infty$, which by (1.2) in the transitive setup and by (1.3) in the setup of Theorem 2 implies that $t_{\mathrm{hit}}^{(n)} \ll t_{\mathrm{cov}}^{(n)}$ along this subsequence. A contradiction! □

4.1 Proof of (1.1) and Stopping Rules

Before proving (1.1), we first recall a notion of mixing, first introduced by Aldous [3] in the continuous-time setup and later studied in discrete time by Lovász and Winkler [33] who developed a rich theory and also by Peres and Sousi [38] and by Oliveira [37]:

$$t_{\mathrm{stop}} := \max_x t_{\mathrm{stop},x}, \qquad \text{where}$$
$$t_{\mathrm{stop},x} := \inf\{\mathbb{E}_x[T] : T \text{ is a stopping rule such that } \mathrm{P}_x[X_T \in \cdot] = \pi(\cdot)\}, \quad (4.3)$$

and where a *stopping rule* is a stopping time, possibly with respect to a filtration larger than the natural filtration. A stopping rule attaining the infimum in (4.3) is called *mean optimal*. Lovász and Winkler [33] showed that for every initial state x there exists a mean optimal stopping rule T and [33, Theorem 2.2] that T is mean optimal iff there exists a state y such that a.s.

$$T \leqslant T_y.$$

Such a state is called a *halting state*. While they work in discrete time, a standard application of Wald's equation can be used to translate their results to the continuous-time setup (cf. [38]; alternatively, one can simply check that the arguments in [33] can be carried out directly in continuous-time).

Aldous [3] showed that under reversibility $\frac{1}{C} \leqslant t_{\mathrm{stop}}/t_{\mathrm{mix}}^{\mathrm{TV}} \leqslant C$ for some universal constant C. This was refined by Peres and Sousi [38] and independently by Oliveira [37] (see also [32, Ch. 24]) who in particular showed that for reversible Markov chains $t_{\mathrm{mix}} \asymp \max_{x,A:\pi(A) \geqslant \alpha} \mathbb{E}_x[T_A]$ for all fixed $\alpha > 1/2$ (this was extended also to $\alpha = 1/2$ in [26]). For more connections between hitting times and mixing times, we refer the reader to [7,27,30].

Proof of (1.1) We suppress the dependence on n. Fix some x and some mean optimal stopping rule T (such that $\mathrm{P}_x[X_T \in \cdot] = \pi(\cdot)$). Let y be an halting state. Then by the strong Markov property, and the fact that $T \leqslant T_y$ a.s., we have that for all t,

$$\mathrm{P}_x[T_y \geqslant t] \geqslant \mathrm{P}_x[T_y - T \geqslant t] = \mathrm{P}_\pi[T_y \geqslant t].$$

By (3.4)

$$P_\pi[T_y \geqslant i\alpha_y] \geqslant (P_\pi[T_y \geqslant \alpha_y/2])^{2i} \geqslant \left(\frac{\alpha_y^2}{4\mathbb{E}_\pi[T_y^2]}\right)^{2i}$$

where we have used the Paley–Zygmund inequality in the second inequality. Using the assumption $\min_z \alpha_z \asymp t_{\text{cov}}$, we conclude the proof by arguing that $\alpha_y^2 \asymp \mathbb{E}_\pi[T_y^2]$. Indeed, by the aforementioned assumption $\alpha_y \asymp t_{\text{hit}}$, whereas $\mathbb{E}_\pi[T_y^2] \leqslant 2t_{\text{hit}}^2$ by (3.3). □

4.2 Proof of (4.2) and (4.1)

To conclude the proof of Propositions 1.1 and 1.2, it remains to prove (4.2) and (4.1).

Proof of (4.2) Let $\varepsilon > 0$. Let $\delta := \varepsilon/2$. Consider $s(\delta) := \frac{|\log \delta|}{\lambda_2}$. Let $x \in V$ and

$$A = A(x, \delta) := \left\{ z : \int_0^{s(\delta)} H_t(z, x)\mathrm{d}t \geqslant \frac{\delta\alpha + s(\delta)}{|V|} \right\}.$$

By Markov's inequality, we have that

$$\pi(A) \leqslant \frac{\sum_{z \in V} \pi(z) \int_0^{s(\delta)} H_t(z,x)\mathrm{d}t}{(\delta\alpha + s(\delta))/|V|} = \frac{s(\delta)}{s(\delta) + \delta\alpha} = \frac{|\log \delta|}{|\log \delta| + \delta\lambda_2\alpha}.$$

By (3.5) and (3.9) as well as by the definition of A, for all $z \notin A$, we have that

$$\alpha - \mathbb{E}_z[T_x] = \int_0^{s(\delta)} (H_t(z, x)|V| - 1)\, dt + \int_{s(\delta)}^\infty (H_t(z, x)|V| - 1)\, dt \leqslant \delta\alpha + \delta\alpha = \varepsilon\alpha.$$

□

Proof of (4.1) Let $x \in V$. Let $\varepsilon > 0$. Let $\delta = \delta(\varepsilon) := \frac{\varepsilon\alpha}{2t_{\text{hit}}}$. Consider $s(\delta) := \frac{|\log \delta|}{\lambda_2}$ and

$$A = A(x, \delta) := \left\{ z : \int_0^{s(\delta)} H_t(z, x)\mathrm{d}t \geqslant \pi(x)\left(\tfrac{\varepsilon}{2}\alpha + s(\delta)\right) \right\}.$$

By Markov's inequality, we have that

$$\pi(A) \leqslant \frac{\sum_{z \in V} \pi(z) \int_0^{s(\delta)} H_t(z,x)\mathrm{d}t}{\left(\tfrac{\varepsilon}{2}\alpha + s(\delta)\right)\pi(x)} = \frac{s(\delta)}{s(\delta) + \tfrac{\varepsilon}{2}\alpha} = \frac{2|\log \delta|}{2|\log \delta| + \varepsilon\lambda_2\alpha}.$$

By (3.5) and (3.9) as well as by the definitions of A and δ, for all $z \notin A$ we have that

$$\alpha_x - \mathbb{E}_z[T_x] = \int_0^{s(\delta)} \left(\frac{H_t(z,x)}{\pi(x)} - 1\right) dt + \int_{s(\delta)}^\infty \left(\frac{H_t(z,x)}{\pi(x)} - 1\right) dt \leqslant \tfrac{\varepsilon}{2}\alpha + \tfrac{1}{2}\delta(\alpha_x + \alpha_z) \leqslant \varepsilon\alpha.$$

By (3.5), we have that $\alpha_z - \mathbb{E}_x[T_z] = \alpha_x - \mathbb{E}_z[T_x] \leqslant \varepsilon\alpha$. □

5 Analysis of Example 1.1

Let G_n be an $n \times m$ grid torus with $m = m(n) \leqslant n$. It is well known that for SRW on the n-cycle the spectral gap is $\asymp n^{-2}$ (cf. [32, Example 12.10]). Hence by general results about product chains (e.g., [32, Corollary 12.13]), $\text{gap}(G_n) \asymp n^{-2}$ (uniformly for all $m(n) \leqslant n$).

We first consider the case that $m \in [n/\log n, n]$. We now prove that $t_{\text{cov}} \lesssim mn(\log n)^2$ for such m. The same bound for $m \in [\frac{n}{(\log n)^2}, \frac{n}{\log n}]$ will be given at the end of the section. Let $H_t := e^{-t(I-P)}$ be the heat kernel of the continuous-time SRW on G_n. Let $H_t^{(1,k)}$ be the heat kernel of the continuous-time SRW on the k-cycle. Then $\max_{a,b} H_t^{(1,k)}(a,b) \leqslant \frac{C_0}{\sqrt{t+1}}$ for $t \leqslant k^2$ (this follows by the local CLT, e.g., [11, §4.4], or from some more general considerations, e.g., [32, Theorem 17.17]), while by the Poincaré inequality

$$\max_{a,b} |H_{t+k^2}^{(1,k)}(a,b)/k - 1| \leqslant \max_a |H_{k^2}^{(1,k)}(a,a)/k - 1|e^{-ct/k^2} \leqslant C_1 e^{-ct/k^2}$$

(for some absolute constants c, C_0, C_1) for all $t \geqslant 0$. Because the continuous-time chain evolves independently in the two coordinates, at rate $\frac{1}{2}$ along each coordinate (e.g., [32, p. 288 (20.35)]) for all $a = (a_1, a_2)$ and $b = (b_1, b_2)$

$$H_t(a,b) = H_{t/2}^{(1,n)}(a_1, b_1)H_{t/2}^{(1,m)}(a_2, b_2). \tag{5.1}$$

Denote the vertex set of G_n by V_n and the uniform distribution on V_n by π. It follows that for all $y \in V_n$ we have that $\int_{2n^2}^{\infty} (H_t(y,y) - \pi(y))\,dt \leqslant C_2 \int_0^{2n^2} (H_t(y,y) - \pi(y))\,dt$ and so

$$\int_0^{\infty} (H_t(y,y) - \pi(y))\,dt \leqslant C_2 \int_0^{2m^2} H_t(y,y)dt + C_2 \int_{2m^2}^{2n^2} H_t(y,y)dt$$
$$\leqslant C_3 \int_0^{2m^2} \frac{dt}{1+t} + C_4 \int_{2m^2}^{2n^2} \frac{dt}{m\sqrt{t/2}} \leqslant C_5\left(\frac{n}{m} + \log m\right). \tag{5.2}$$

Hence by Fact 3.1 and (3.1)

$$t_{\text{hit}}(G_n) \leqslant 2\alpha(G_n) \leqslant 2C_5(n^2 + nm\log m),$$

and so by (3.10)

$$t_{\text{cov}}(G_n) \leqslant C_6(n^2 + nm\log m)\log n.$$

For $m \in [n/\log n, n]$, we get that $t_{\text{cov}}(G_n) \lesssim nm(\log n)^2$.

We now prove a matching lower bound for $m \in [n/(\log n)^2, n]$. If $x, y \in V_n$ are of graph distance at least $\sqrt{m}$, then by the local CLT (e.g., [11, §4.4])

$$\forall t \geqslant 0, \qquad H_t(y,y) - H_t(x,y) \geqslant c_1(t+1)^{-1}\mathbf{1}_{\{t \leqslant 2m\}},$$

where we have used the fact that for transitive chains $H_t(y, y) \geqslant H_t(x, y)$ for all x, y and all t (e.g., [5, Eq. (3.60)]). Hence by (3.6) for such x, y we have that

$$\mathbb{E}_y[T_x] = \mathbb{E}_x[T_y] = mn \int_0^\infty (H_t(y, y) - H_t(x, y)) \, \mathrm{d}t \geqslant c_2 mn \log m, \quad (5.3)$$

By considering a collection of vertices $A \subset V_n$ of size $\Omega(n)$ such that for any distinct $a, b \in A$ we have that the distance of a from b is at least $\sqrt{m}$ we get by (5.3) and Matthews' argument that $t_{\mathrm{cov}}(G_n) \gtrsim nm(\log n)^2$ (for $m \in [n/(\log n)^2, n]$).

We now treat $m \in [1, n/(\log n)^2]$. Using our upper bound on the spectral gap, it remains only need to show that $t_{\mathrm{cov}}(G_n) \lesssim n^2$ in this regime. This requires a more careful analysis than the one above.

It is not hard to show that there exists an absolute constant $p \in (0, 1)$ such that for every n and every $C \geqslant 1$, SRW on the n-cycle satisfies that all vertices are visited at least $Cn/4$ times by time Cn^2 with probability at least p. For an argument which is specific for the cycle cf. [10, Lemma 6.6]. This also follows from a general result from [22] which says that the blanket-time (see [22] for a definition) is proportional to the cover time. Thus with probability at least p, by time $8Cn^2$ for all $i \in [n]$ the walk spends at least Cn steps at each strip $S_i := \{i\} \times [m]$, where $[k] := \{1, \ldots, k\}$. We now exploit this fact to obtain the bound $t_{\mathrm{cov}}(G_n) \lesssim n^2$.

For a set A, we define the *induced chain* on A, denoted by $(Y_k^A)_{k=0}^\infty$, to be the chain $(X_t)_{t \geqslant 0}$ viewed only at times at which it visits A. That is $t_0(A) := \inf\{t \geqslant 0 : X_t \in A\}$, $Y_0^A := X_{t_0(A)}$ and inductively,

$$t_{k+1}(A) := \inf\{t > t_k(A) : X_t \in A, X_s \neq X_{t_k(A)} \text{ for some } s \in (t_k(A), t]\}$$

and $Y_{k+1}^A := X_{t_{k+1}(A)}$. Observe that if $\tau_{\mathrm{cov}}^A := \inf\{t : \{X_s : s \leqslant t\} \supseteq A\}$ and $\tau_{\mathrm{cov}}^{\mathrm{induced}, A} := \inf\{k : \{Y_i^A : i \leqslant k\} = A\}$ are the cover times of A w.r.t. the original chain and the induced chain on A, respectively, and $\mathrm{P}^{\mathrm{induced}, A}$ and $\mathbb{E}^{\mathrm{induced}, A}$ are the law and expectation of the induced chain on A, then for any partition $A_1, \ldots, A_\ell$ of V we have that for all $v \in V$

$$\forall t \geqslant 0, k \in \mathbb{N}, \quad \mathrm{P}_v[\tau_{\mathrm{cov}} > t] \leqslant \mathrm{P}_v[\max_{i \in [\ell]} t_k(A_i) > t]$$
$$+ \sum_{j \in [\ell]} \max_{a \in A_j} \mathrm{P}_a^{\mathrm{induced}, A_j}[\tau_{\mathrm{cov}}^{\mathrm{induced}, A_j} > k],$$

where we have used the fact that for all $j \in [\ell]$ and $k \in \mathbb{N}$ we have that

$$\mathrm{P}_v[\tau_{\mathrm{cov}}^{A_j} > t, \max_{i \in [\ell]} t_k(A_i) \leqslant t] \leqslant \mathrm{P}_v[\{X_{t_i(A_j)}^{A_j} : i \leqslant k\} \neq A_j]$$
$$= \max_{a \in A_j} \mathrm{P}_a^{\mathrm{induced}, A_j}[\tau_{\mathrm{cov}}^{\mathrm{induced}, A_j} > k].$$

We now argue that it suffices to show that the induced chain on each strip S_i satisfies that the probability that it is not covered in Cn steps is $\ll 1/n$, provided that C is

sufficiently large (note that here steps are counted w.r.t. the induced chain, i.e., these are the number of visits to each strip). That is, by symmetry, it suffices to show that $\max_{a \in S_1} P_a^{\text{induced}, S_1}[\tau_{\text{cov}}^{\text{induced}, S_1} > \lceil Cn \rceil] = o(1/n)$. Indeed, once this is established for some $C \geq 1$, then using the above with the partition $S_1, \ldots, S_n$ we get that

$$P_v[\tau_{\text{cov}} > 8Cn^2] \leq P_v[\max_{i \in [n]} t_{\lceil Cn \rceil}(S_i) > 8Cn^2] + n \max_{a \in S_1} P_a^{\text{induced}, S_1}[\tau_{\text{cov}}^{\text{induced}, S_1} > \lceil Cn \rceil].$$

The r.h.s. is at most $1 - p + o(1)$ where $p \in (0, 1)$ is as above. Using the obvious submultiplicativity property

$$\forall s \geq 0, r \in \mathbb{N}, \qquad \max_v P_v[\tau_{\text{cov}} > sr] \leq (\max_v P_v[\tau_{\text{cov}} > s])^r,$$

this implies that $t_{\text{cov}}(G_n) \lesssim n^2$ as desired.

Denote $M := \max_{x, y \in S_1} \mathbb{E}_x^{\text{induced}, S_1}[T_y]$. We will show that $M \leq C_6 m \log m$. By Markov's inequality $P_a^{\text{induced}}[T_b > 2M] \leq \frac{1}{2}$ for all a, b in the same strip. By the Markov property, for all a, b in the same strip, we have that

$$P_a^{\text{induced}}[T_b > (2M)\lceil \log_2(n^3) \rceil] \leq 2^{-\log_2(n^3)} = n^{-3}.$$

By a union bound over all m vertices in that strip, we obtain the desired tail estimate on the cover time of a single strip w.r.t. the induced chain, as $(2M)\lceil \log_2(n^3) \rceil \lesssim m \log m \log n \lesssim n$.

It remains to show that $M = O(m \log m)$. The induced chain is itself a transitive chain. In particular, its stationary distribution is the uniform distribution. Let $x, y \in S_1$. By Wald's equation, and the fact that the expected return time to S_1 from any $a \in S_1$ is n, we have that $\mathbb{E}_x[T_y] = n\mathbb{E}_x^{\text{induced}, S_1}[T_y]$, and so by (3.6) we have that

$$\mathbb{E}_x^{\text{induced}, S_1}[T_y] = m \int_0^\infty [H_t(y, y) - H_t(x, y)]dt. \tag{5.4}$$

Finally, for $x = (1, x_2), y = (1, y_2) \in S_1$ by the local CLT (e.g., [11, §4.4]) and (5.1)

$$H_t(y, y) - H_t(x, y) \leq H_{t/2}^{(1,n)}(1, 1) \max_{a,b} \left(H_{t/2}^{(1,m)}(1, a) - H_{t/2}^{(1,m)}(1, b) \right)$$

$$\lesssim (t+1)^{-1}\mathbf{1}_{\{t \leq 2m^2\}} + \frac{1}{m\sqrt{t}}e^{-ct/m^2}\mathbf{1}_{\{2m^2 < t \leq 2n^2\}} + \frac{1}{mn}e^{-ct/m^2}\mathbf{1}_{\{t > 2n^2\}}, \tag{5.5}$$

where we have used the Poincaré inequality for SRW on the m-cycle. Integrating we see that for $m \in [n/\log n, n]$ indeed

$$\max_{x,y \in S_1} \int_0^\infty [H_t(y, y) - H_t(x, y)]dt \lesssim \log m. \tag{5.6}$$

By (5.4), we get that indeed $M \lesssim m \log m$, as desired.

We now show that also for $m \in [n/(\log n)^2, n/\log n]$ we have $t_{\mathrm{cov}} \lesssim nm(\log n)^2$. Indeed by (5.6) $M \lesssim m \log m$ and so by the above analysis, we can bound t_{cov} from above, up to a constant factor, by the expected time until all strips are visited for at least $CM \log n$ time units. As $M \log n \gtrsim n$, it is not hard to show that this time is of order $n \times (CM \log n) \asymp nm(\log n)^2$ by using the fact that there exists $C > 0$ such that with probability at least $p > 0$ all strips are visited for at least n time units by time Cn^2. To see this, consider a sequence of consecutive time intervals of length Cn^2 and use the Markov property to argue that during each interval, with probability at least $p > 0$, independently of the previous time intervals, all strips are visited for at least n time units). □

References

1. Aldous, D.: Hitting times for random walks on vertex-transitive graphs. Math. Proc. Camb. Philos. Soc. **106**(1), 179–191 (1989)
2. Aldous, D.: Random walk covering of some special trees. J. Math. Anal. Appl. **157**(1), 271–283 (1991)
3. Aldous, D.: Some inequalities for reversible Markov chains. J. Lond. Math. Soc. **2**(3), 564–576 (1982)
4. Aldous, D.: Threshold limits for cover times. J. Theor. Probab. **4**(1), 197–211 (1991)
5. Aldous, D., Fill, J.: Reversible Markov chains and random walks on graphs. Unfinished manuscript. Available at the first author's
6. Barlow, M.T., Ding, J., Nachmias, A., Peres, Y.: The evolution of the cover time. Comb. Probab. Comput. **20**(3), 331–345 (2011)
7. Basu, R., Hermon, J., Peres, Y.: Characterization of cutoff for reversible Markov chains. Ann. Probab. **45**(3), 1448–1487 (2017)
8. Belius, D.: Gumbel fluctuations for cover times in the discrete torus. Probab. Theory Relat. Fields **157**(3–4), 635–689 (2013)
9. Belius, D., Kistler, N.: The subleading order of two dimensional cover times. Probab. Theory Relat. Fields **167**(1–2), 461–552 (2017)
10. Benjamini, I., Hermon, J.: Recurrence of Markov chain traces. To appear in Ann. Henri Poincaré Probab. Stat. Arxiv preprint (2019)
11. Benjamini, I., Fontes, L. R., Hermon, J., Machado, F. P.: On an epidemic model on finite graphs. To appear in Ann. Appl. Probab. Arxiv preprint (2019)
12. Benjamini, I., Hermon, J., Tessera, R., Tointon, M.: Transitive graphs with proportional mixing and hitting times scale to the unit-circle. In preparation (2019+)
13. Breuillard, E., Tointon, M.C.H.: Nilprogressions and groups with moderate growth. Adv. Math. **289**, 1008–1055 (2016)
14. Chen, G.Y., Saloff-Coste, L.: The cutoff phenomenon for ergodic Markov processes. Electron. J. Probab. **13**(3), 26–78 (2008)
15. Comets, F., Gallesco, C., Popov, S., Vachkovskaia, M.: On large deviations for the cover time of two-dimensional torus. Electron. J. Probab. **18**(96), 18 (2013)
16. Cooper, C., Frieze, A.: The cover time of the giant component of a random graph. Random Struct. Algorithms **32**(4), 401–439 (2008)

17. Diaconis, P., Saloff-Coste, L.: Moderate growth and random walk on finite groups. Geom. Funct. Anal. **4**(1), 1–36 (1994)
18. Dembo, A., Peres, Y., Rosen, J., Zeitouni, O.: Cover times for Brownian motion and random walks in two dimensions. Ann. Math. (2) **160**(2), 433–464 (2004)
19. Diaconis, P., Saloff-Coste, L.: Walks on generating sets of groups. Invent. Math. **134**(2), 251–299 (1998)
20. Ding, J.: Asymptotics of cover times via Gaussian free fields: bounded-degree graphs and general trees. Ann. Probab. **42**(2), 464–496 (2014)
21. Ding, J.: On cover times for 2D lattices. Electron. J. Probab. **17**(45), 18 (2012)
22. Ding, J., Lee, J.R., Peres, Y.: Cover times, blanket times, and majorizing measures. Ann. Math. (2) **175**(3), 1409–1471 (2012)
23. Ding, J., Zeitouni, O.: A sharp estimate for cover times on binary trees. Stoch. Process. Appl. **122**(5), 2117–2133 (2012)
24. Feige, U.: A tight lower bound on the cover time for random walks on graphs. Random Struct. Algorithms **6**(4), 433–438 (1995)
25. Feige, U.: A tight upper bound on the cover time for random walks on graphs. Random Struct. Algorithms **6**(1), 51–54 (1995)
26. Griffiths, S., Kang, R., Oliveira, R., Patel, V.: Tight inequalities among set hitting times in Markov chains. Proc. Am. Math. Soc. **142**(9), 3285–3298 (2014)
27. Hermon, J.: A technical report on hitting times, mixing and cutoff. ALEA Lat. Am. J. Probab. Math. Stat. **15**(1), 101–120 (2018)
28. Hermon, J.: Intersection times for critical branching random walk. Arxiv preprint (2019)
29. Hermon, J., Lacoin, H., Peres, Y.: Total variation and separation cutoffs are not equivalent and neither one implies the other. Electron. J. Probab. **21**(44), 36 (2016)
30. Hermon, J., Peres, Y.: A characterization of L_2 mixing and hypercontractivity via hitting times and maximal inequalities. Probab. Theory Relat. Fields **170**(3–4), 769–800 (2018)
31. Kahn, J., Kim, J.H., Lovász, L., Vu, V.H.: The cover time, the blanket time, and the Matthews bound. In: 41st Annual Symposium on Foundations of Computer Science, pp. 467–475. IEEE Computer Society Press, Los Alamitos (2000)
32. Levin, D., Peres, Y.: Markov Chains and Mixing Times. American Mathematical Society, Providence, RI. With contributions by Elizabeth L. Wilmer and a chapter by James G. Propp and David B. Wilson (2017)
33. Lovász, L., Winkler, P.: Mixing times. Microsurv. Discrete Probab. **41**, 85–134 (1998)
34. Lyons, R., Peres, Y.: Probability on Trees and Networks. Cambridge Series in Statistical and Probabilistic Mathematics, vol. 42. Cambridge University Press, New York (2016)
35. Matthews, P.: Covering problems for Markov chains. Ann. Probab. **16**(3), 1215–1228 (1988)
36. Miller, J., Peres, Y.: Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. Ann. Probab. **40**(2), 535–577 (2012)
37. Oliveira, R.: Mixing and hitting times for finite Markov chains. Electron. J. Probab. **17**(70), 12 (2012)
38. Peres, Y., Sousi, P.: Mixing times are hitting times of large sets. J. Theor. Probab. **28**(2), 488–519 (2015)
39. Tessera, R., Tointon, M.: A finitary structure theorem for vertex-transitive graphs of polynomial growth. Arxiv preprint (2019)
40. Tessera, R., Tointon, M.: Sharp relations between volume growth, isoperimetry and resistance in vertex-transitive graphs. In preparation (2019+)
41. Zhai, A.: Exponential concentration of cover times. Electron. J. Probab. **23**(32), 22 (2018)