MIXING AND RELAXATION TIME FOR RANDOM WALK ON WREATH PRODUCT GRAPHS

By Júlia Komjáthy [*]

Yuval Peres

Budapest University of Technology and Economics and Microsoft Research

Suppose that G and H are finite, connected graphs, G regular, X is a lazy random walk on G and Z is a reversible ergodic Markov chain on H. The generalized lamplighter chain $X^\diamond$ associated with X and Z is the random walk on the wreath product $H \wr G$, the graph whose vertices consist of pairs $(\underline{f}, x)$ where $\underline{f} = (f_v)_{v \in V(G)}$ is a labeling of the vertices of G by elements of H and x is a vertex in G. In each step, $X^\diamond$ moves from a configuration $(\underline{f}, x)$ by updating x to y using the transition rule of X and then independently updating both f_x and f_y according to the transition probabilities on H; f_z for $z \neq x, y$ remains unchanged. We estimate the mixing time of $X^\diamond$ in terms of the parameters of H and G. Further, we show that the relaxation time of $X^\diamond$ is the same order as the maximal expected hitting time of G plus $|G|$ times the relaxation time of the chain on H.

1. Introduction. Suppose that G and H are finite connected graphs with vertices $V(G), V(H)$ and edges $E(G), E(H)$, respectively. We refer to G as the base and H as the lamp graph, respectively. Let $\mathcal{X}(G) = \{\underline{f} \colon V(G) \to H\}$ be the set of markings of $V(G)$ by elements of H. The wreath product $H \wr G$ is the graph whose vertices are pairs $(\underline{f}, x)$ where $\underline{f} = (f_v)_{v \in V(G)} \in \mathcal{X}(G)$ and $x \in V(G)$. There is an edge between $(\underline{f}, x)$ and $(\underline{g}, y)$ if and only if $(x, y) \in E(G)$, $(f_x, g_x), (f_y, g_y) \in E(H)$ and $f_z = g_z$ for all $z \notin \{x, y\}$. Suppose that P and Q are transition matrices for Markov chains on G and on H, respectively. The generalized lamplighter walk $X^\diamond$ (with respect to the transition matrices P and Q) is the Markov chain on $H \wr G$ which moves from a configuration $(\underline{f}, x)$ by

1. picking y adjacent to x in G according to P, then
2. updating each of the values of f_x and f_y independently according to Q on H.

The state of lamps f_z at all other vertices $z \in G$ remain fixed. It is easy to see that if P and Q are irreducible, aperiodic and reversible with stationary distribution π_G and π_H, respectively, then the unique stationary distribution of $X^\diamond$ is the product measure

$$\pi^\diamond\big((\underline{f}, x)\big) = \pi_G(x) \cdot \prod_{v \in V(G)} \pi_H\left(f_v\right),$$

and $X^\diamond$ is itself reversible. In this article, we will be concerned with the special case that P is the transition matrix for the *lazy random walk* on G. In particular, P is given by

$$(1.1) \qquad P(x, y) := \begin{cases} \frac{1}{2} \text{ if } x = y, \\ \frac{1}{2d(x)} \text{ if } \{x, y\} \in E(G), \end{cases}$$

for $x, y \in V(G)$ and where $d(x)$ is the degree of x. We further assume that the transition matrix Q on H is irreducible and aperiodic. This and the assumption (1.1) guarantees that we avoid issues of periodicity.
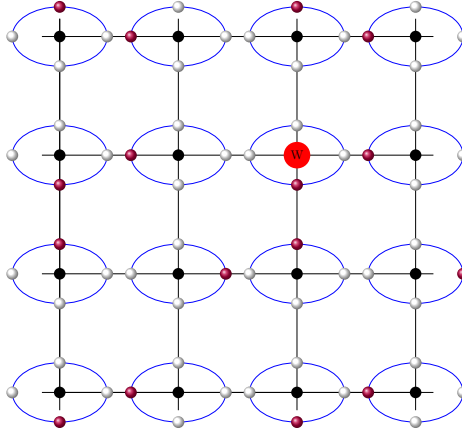

FIG 1. *A typical state of the generalized lamplighter walk. Here $H = \mathbf{Z}_4$ and $G = \mathbf{Z}_4^2$, the red bullets on each copy of H represents the state of the lamps over each vertex $v \in G$ and the walker is drawn as a red W bullet.*

1.1. *Main Results.*

In order to state our general result, we first need to review some basic terminology from the theory of Markov chains. Let P be the transition kernel for a lazy random walk on a finite, connected graph G with stationary distribution π.

The ε-*mixing time* of P on G in total variation distance is given by

$$(1.2) \qquad t_{\mathrm{mix}}(G, \varepsilon) := \min \left\{ t \geq 0 : \max_{x \in V(G)} \frac{1}{2} \sum_{y} \left| P^t(x, y) - \pi(y) \right| \leq \varepsilon \right\}.$$

Throughout, we set $t_{\mathrm{mix}}(G) := t_{\mathrm{mix}}(G, \frac{1}{4})$.

The *relaxation time* of a reversible Markov Chain with transition matrix P is

$$(1.3) \qquad t_{\mathrm{rel}}(G) := \frac{1}{1 - \lambda_2}$$

where λ_2 is the second largest eigenvalue of P.

The *maximal hitting time* of P is

$$(1.4) \qquad t_{\mathrm{hit}}(G) := \max_{x, y \in V(G)} \mathbf{E}_x[\tau_y],$$

where τ_y denotes the first time t that $X(t) = y$ and $\mathbf{E}_x$ stands for the expectation under the law in which $X(0) = x$. The random cover time τ_{cov} is the first time when all vertices have been visited by the walker X, and the cover time $t_{\mathrm{cov}}(G)$ is

$$(1.5) \qquad t_{\mathrm{cov}}(G) := \max_{x \in V(G)} \mathbf{E}_x[\tau_{\mathrm{cov}}].$$

The next needed concept is that of strong stationary times.

DEFINITION 1.1. *A randomized stopping time τ is called a strong stationary time for the Markov chain X_t on G if*

$$\mathbf{P}_x[X_\tau = y, \tau = t] = \pi(y)\mathbf{P}_x[\tau = t],$$

that is, the position of the walk when it stops at τ is independent of the value of τ.

The adjective randomized means that the stopping time can depend on some extra randomness, not just purely the trajectories of the Markov chain, for a precise definition see [13, Section 6.2.2].

DEFINITION 1.2. *A state $h(x) \in V(G)$ is called a halting state for a stopping time τ and initial state x if $\{X_t = h(x)\}$ implies $\{\tau \leq t\}$.*

Our main results are summarized in the following theorems:

THEOREM 1.3.   *Let us assume that G and H are connected graphs with G regular and the Markov chain on H is ergodic and reversible. Then there exist universal constants c_1, C_1 such that the relaxation time of the generalized lamplighter walk on $H \wr G$ satisfies*

$$(1.6) \qquad c_1 \leq \frac{t_{\mathrm{rel}}(H \wr G)}{t_{\mathrm{hit}}(G) + |G| t_{\mathrm{rel}}(H)} \leq C_1,$$

THEOREM 1.4.   *Assume that the conditions of Theorem 1.3 hold and further assume that the chain with transition matrix Q on H is lazy, i.e. $Q(x, x) \geq \frac{1}{2} \ \forall x \in H$. Then there exist universal constants c_2, C_2 such that the mixing time of the generalized lamplighter walk on $H \wr G$ satisfies*

$$(1.7) \qquad \begin{aligned} c_2 \left(t_{\mathrm{cov}}(G) + t_{\mathrm{rel}}(H)|G| \log |G| + |G| t_{\mathrm{mix}}(H) \right) &\leq t_{\mathrm{mix}}(H \wr G), \\ t_{\mathrm{mix}}(H \wr G) &\leq C_2 \left(t_{\mathrm{cov}}(G) + |G| t_{\mathrm{mix}}(H, \frac{1}{|G|}) \right). \end{aligned}$$

If further the Markov chain is such that

(A)   *There is a strong stationary time τ_H for the Markov chain on H which possesses a halting state $h(x)$ for every initial starting point $x \in H$,*

then the upper bound of 1.7 is sharp.

REMARK 1.5.   *The laziness assumption on the transition matrix Q on H is only used to get the term $c_2 |G| t_{\mathrm{mix}}(H)$ in (1.7). All the other bounds hold without the laziness assumption.*

REMARK 1.6.   *If the Markov Chain on H is such that*

$$t_{\mathrm{mix}}(H, \varepsilon) \leq t_{\mathrm{mix}}(H, 1/4) + t_{\mathrm{rel}}(H) \log \varepsilon,$$

then the upper bound matches the lower bound. This holds for many natural chains such as lazy random walk on hypercube $\mathbf{Z}_2^d$, tori $\mathbf{Z}_n^d$, some walks on the permutation group S_n (the random transpositions or random adjacent transpositions shuffle, and the top-to-random shuffle, for instance).

REMARK 1.7.   *Many examples where Assumption (**A**) holds are given in the thesis of Pak [17], including the cycle $\mathbf{Z}_n$, the hypercube $\mathbf{Z}_2^d$ and more generally tori $\mathbf{Z}_n^d, n, d \in \mathbf{N}$ and dihedral groups $\mathbf{Z}_2 \ltimes \mathbf{Z}_n, n \in \mathbf{N}$ are also obtained by the construction of strong stationary times with halting states on direct and semidirect product of groups. Further, Pak constructs strong stationary times possessing halting states for the random walk on k-sets of n-sets, i.e. on the group $S_n/(S_k \times S_{n-k})$, and on subsets of $n \times n$ matrices over the full linear group, i.e. on $GL(n, \mathbb{F}_q)/(GL(k, \mathbb{F}_q) \times GL(n-k, \mathbb{F}_q))$.*

1.2. *Previous Work.* The mixing time of $\mathbf{Z}_2 \wr G$ was first studied by Häggström and Jonasson in [11] in the case of G being the complete graph K_n and the one-dimensional cycle $\mathbf{Z}_n$. Generalizing their results, Peres and Revelle [18, Theorem 1.2, 1.3] proved that there exists constants c_i, C_i depending on ε such that for any transitive graph G,

$$c_1 t_{\mathrm{hit}}(G) \leq t_{\mathrm{rel}}(\mathbf{Z}_2 \wr G) \leq C_1 t_{\mathrm{hit}}(G),$$
$$c_2 t_{\mathrm{cov}}(G) \leq t_{\mathrm{mix}}(\mathbf{Z}_2 \wr G, \varepsilon) \leq C_2 t_{\mathrm{cov}}(G).$$

The vertex transitivity condition was dropped in [13, Theorem 19.1, 19.2]. These bounds match with Theorems 1.3 and 1.4 since $H_n = \mathbf{Z}_2$ implies that the terms not containing H_n in the denominator of (1.6) and in the bounds in (1.7) dominate.

In [16], it is shown that $t_{\mathrm{mix}}(\mathbf{Z}_2 \wr G_n) \sim \frac{1}{2} t_{\mathrm{cov}}(G_n)$ whenever (G_n) is a sequence of graphs satisfying some uniform local transience assumptions, including $G_n = \mathbf{Z}_n^d$ with $d \geq 3$ fixed.

Moving towards larger lamp spaces, if the base is the complete graph K_n and $|H_n| = o(n)$ one can determine the order of mixing time from [13, Theorem 20.7], since in this case the lamplighter chain is a product chain on $\prod_{i=1}^n H_n$. Levi [14] investigated random walks on wreath products when $H \neq \mathbf{Z}_2$. In particular, he determined the order of the mixing time of $K_{n^\lambda} \wr K_n$, $0 \leq \lambda \leq 1$, and he also had upper and lower bounds for the case $H_d \wr \mathbf{Z}_n$, i.e. H is the d-dimensional hypercube and the base is a cycle of length n, however, the bounds failed to match for general d and n. Further, Fill and Schoolfield [10] investigated the total variation and l_2 mixing time of $K_n \wr S_n$, where the base graph is the Cayley graph of the symmetric group S_n with transpositions chosen as the generator set, and the stationary distribution on K_n is not necessarily uniform.

The mixing time of $H_n = \mathbf{Z}_2$ is closely related to the cover time of the base graph, and thus it helps understanding the geometric structure of the last visited points by random walk [4–6, 16]. Further, larger lamp graphs give more information on the local time structure of the base graph G. This relates our work to the literature on blanket time (when all the local times of vertices are within a constant factor of each other) [3, 8, 20].

1.3. *Outline.* The remainder of this article is structured as follows. In Section 3 we state a few necessary theorems and lemmas about the Dirichlet form, strong stationary times, different notions of distances and their relations. In Lemmas 3.3 and 3.5 we construct a crucial stopping time $\tau^\diamond$ and a strong stationary time $\tau_2^\diamond$ on $H \wr G$ which we will use several times throughout the proofs later. Then we prove the main theorem about the relaxation time in Section 4, and the mixing time bounds in Section 5.

2. Notations. Throughout the paper, objects related to the base or the lamp graph will be indexed by G and H, respectively, and $\diamond$ always refers to an object related to the whole $H \wr G$. Unless misleading, G and H refers also to the vertex set of the graphs, i.e. $v \in G$ means $v \in V(G)$. $\mathbf{P}_\mu, \mathbf{E}_\mu$ denotes probability and expectation under the conditional law where the initial distribution of the Markov chain under investigation is μ. Similarly, $\mathbf{P}_x$ is the law under which the chain starts at x.

3. Preliminaries. In this section we collect the preliminary lemmas to be able to carry through the proofs quickly afterwards. The reader familiar with notions of strong stationary times, separation distance, and Dirichlet forms might want jump forward to Lemmas 3.3 and 3.5 immediately, and check the other lemmas here only when needed.

The first lemma is a common useful tool to prove lower bounds for relaxation times, by giving the variational characterization of the spectral gap. First we start with a definition.

Let P be a reversible transition matrix with stationary distribution π on the state space Ω and let $\mathbf{E}_\pi[\phi] := \sum_{y \in \Omega} \phi(y)\pi(y)$. The Dirichlet form associated to the pair (P, π) is defined for functions ϕ and η on Ω by

$$\mathcal{E}(\phi, \eta) := \langle (I - P)\phi, \eta \rangle_\pi = \sum_{y \in \Omega} (I - P)\phi(y)\eta(y)\pi(y).$$

It is not hard to see [13, Lemma 13.11] that

$$(3.1) \qquad \mathcal{E}(\phi) := \mathcal{E}(\phi, \phi) = \frac{1}{2}\mathbf{E}_\pi \left[(\phi(X_1) - \phi(X_0))^2 \right]$$

The next lemma relates the spectral gap of the chain to the Dirichlet form (for a short proof see [2] or [13, Lemma 13.12]):

LEMMA 3.1 (Variational characterization of the spectral gap). *The spectral gap $\gamma = 1 - \lambda_2$ of a reversible Markov Chain satisfies*

$$(3.2) \qquad \gamma = \min_{\phi:\mathbf{Var}_\pi\phi \neq 0} \frac{\mathcal{E}[\phi]}{\mathbf{Var}_\pi\phi},$$

where $\mathbf{Var}_\pi\phi = \mathbf{E}_\pi[\phi^2] - (\mathbf{E}_\pi[\phi])^2$.

A very useful object to prove the upper bound on t_{rel} and both bounds for t_{mix} is the concept of strong stationary times. Recall the definition from (1.1). It is not hard to see ([13, Lemma 6.9]) that this is equivalent to

$$(3.3) \qquad \mathbf{P}_x\left[X_t = y, \tau \leq t\right] = \pi(y)\mathbf{P}_x[\tau \leq t].$$

To be able to relate the tail of the strong stationary times to the mixing time of the graphs, we need another distance from stationary measure, called the separation distance:

$$(3.4) \qquad s_x(t) := \max_{y \in \Omega} \left[1 - \frac{P^t(x, y)}{\pi(y)} \right].$$

The relation between the separation distance and any strong stationary time τ is the following inequality from [2] or [7] or [13, Lemma 6.11]:

$$(3.5) \qquad \forall x \in \Omega : s_x(t) \le \mathbf{P}_x(\tau > t).$$

Throughout the paper, we will need a slightly stronger result than (3.5). Namely, by [7, Remark 3.39] or from the proof of (3.5) in [13, Lemma 6.11] it follows that in (3.5) equality holds *if τ has a halting state $h(x)$ for x*. Unfortunately, we just point out that the [13, Remark 6.12] is not true and the statement can not be reversed: the state $h(x, t)$ maximizing the separation distance at time t can also depend on t and thus the existence of a halting state is not necessarily needed to get equality in (3.5).

On the other hand, one can always construct τ such that (3.5) holds with equality for every $x \in \Omega$. This is a key ingredient to our proofs, so we cite it as a Theorem (with adjusted notation to the present paper).

THEOREM 3.2. *[Aldous, Diaconis][1, Proposition 3.2] Let $(X_t, t \ge 0)$ be an irreducible aperiodic Markov chain on a finite state space Ω with initial state x and stationary distribution π, and let $s_x(t)$ be the separation distance defined as in (3.4). Then*

1. *if τ is a strong stationary time for X_t, then $s_x(t) \le \mathbf{P}_x(\tau > t)$ for all $t \ge 0$.*
2. *Conversely, there exists a strong stationary time τ such that $s_x(t) = \mathbf{P}_x(\tau > t)$ holds with equality.*

Combining these, we will call a strong stationary time τ *separation optimal* if it achieves equality in (3.5). Mind that every stopping time possessing halting states is separation optimal, but the reversed statement is not necessarily true. The next two lemmas, which we will use several times, construct two stopping times for the graph $H \wr G$. The first one will be used to lower bound the separation distance and the second one upper bounds it.

We start with introducing the notation

$$(3.6) \qquad L_v(t) = 2 \sum_{i=0}^{t} \mathbf{1}(X_i = v) - \delta_{X_0, v} - \delta_{X_t, v}$$

for the number of moves on the lamp graph $H_v, v \in G$ by the walker up to time t. Slightly abusing terminology, we call it the local time at vertex $v \in G$.

Let us further denote the random walk with transition matrix Q on H by Z. Since the moves on the different lamp graphs $H_v, v \in G$ are taken independently given $L_v(t), v \in G$, we can define for each $v \in G$ an independent copy of the chain Z, denoted by Z_v, running on H_v. Thus, the position of the lamplighter chain at time t can be described as

$$(\underline{F}_t, X_t) = \left((Z_v(L_v(t)))_{v \in G}, X_t \right)$$

Below we will use copies of a strong stationary time τ_H for each $v \in G$, meaning that $\tau_H(v)$ is defined in terms of Z_v, and given the local times $L_v(t)$, $\tau_H(v)$-s are independent of each other.

LEMMA 3.3.   *Let τ_H be any strong stationary time for the Markov chain on H. Take the conditionally independent copies of $(\tau_H(v))_{v \in G}$ given the local times $L_v(t)$, realized on the lampgraphs H_v-s and define the stopping time $\tau^\diamond$ for $X^\diamond$ by*

$$(3.7) \qquad\qquad \tau^\diamond := \inf \left\{ t : \forall v \in G : \tau_H(v) \leq L_v(t) \right\}.$$

Then, for any starting state $(\underline{f}_0, x_0)$ we have

$$(3.8)\quad \mathbf{P}_{(\underline{f}_0, x_0)} \left[X_t^\diamond = (\underline{f}, x), \tau^\diamond = t \right] = \prod_{v \in G} \pi_H(f_v) \cdot \mathbf{P}_{(\underline{f}_0, x_0)} \left[X_t = x, \tau^\diamond = t \right].$$

If further τ_H has halting states then the vectors $(\underline{h}(f_v(0)), y)$ are halting state vectors for $\tau^\diamond$ and initial state $(\underline{f}_0, x_0)$ for every $y \in G$.

We postpone the proof and continue with a corollary of the lemma:

COROLLARY 3.4.   *Let τ_H be a strong stationary time for the Markov chain on H which has a halting state $h(z)$ for any $z \in H$. Then define $\tau^\diamond$ as in Lemma 3.3. Then for the separation distance on the lamplighter chain $H \wr G$ the following lower bound holds:*

$$s_{(\underline{f}_0, x_0)}(t) \geq \mathbf{P}_{(\underline{f}_0, x_0)} \left[\tau^\diamond > t \right].$$

PROOF. Observe that reaching the halting state vector $(\underline{h}(f_v(0)), x)$ implies the event $\tau^\diamond \leq t$ so we have
$$(3.9)$$
$$1 - \frac{\mathbf{P}_{(\underline{f}_0, x_0)} \left[X_t^\diamond = (\underline{h}(f_v(0)), x) \right]}{\pi_G(x) \prod_{v \in G} \pi_H\left(h(f_v(0)) \right)} = 1 - \frac{\mathbf{P}_{(\underline{f}_0, x_0)} \left[X_t^\diamond = (\underline{h}(f_v(0)), x), \tau^\diamond \leq t \right]}{\pi_G(x) \prod_{v \in G} \pi_H\left(h(f_v(0)) \right)}$$

Now pick a vertex $x_{x_0,t} \in G$ which minimizes $\mathbf{P}\left[X_t = x_{x_0,t} | \tau^{\diamond} \leq t\right]/\pi_G(x_{x_0,t})$. This quotient is less than 1 since both the numerator and the denominator are probability distributions on G. Then, using this and Lemma 3.3, the right hand side of (3.9) equals

$$1 - \frac{\mathbf{P}_{(\underline{f}_0, x_0)}\left[X_t = x_{x_0,t} | \tau^{\diamond} \leq t\right] \mathbf{P}_{(\underline{f}_0, x_0)}[\tau^{\diamond} \leq t]}{\pi_G(x_{x_0,t})} \geq 1 - \mathbf{P}_{(\underline{f}_0, x_0)}\left[\tau^{\diamond} \leq t\right].$$

Clearly the separation distance is larger than the left hand side of (3.9), and the proof of the claim follows. Note that the proof only works if τ_H has a halting state and thus it is separation-optimal. $\square$

PROOF OF LEMMA 3.3. First we show that (3.8) holds using the conditional independence of $\tau_H(v)$-s given the number of moves $L_v(t)$ on the lamp graphs $H(v), v \in G$. Clearly, conditioning on the trajectory of the walker $\{X_1, \ldots, X_{t-1}, X_t = x\} := X[1,t]$ contains the knowledge of $L_v(t)$-s as well. We will omit to note the dependence of $\mathbf{P}$ on initial state $(\underline{f}_0, x_0)$ for notational convenience. The left hand side of condition (3.3) equals

$$\mathbf{P}\left[X_t^{\diamond} = (\underline{f}, x), \tau^{\diamond} \leq t\right] = \sum_{X_{[1,t]}} \mathbf{P}\left[X_t^{\diamond} = (\underline{f}, x), \tau^{\diamond} \leq t | X_{[1,t]}\right] \mathbf{P}\left[X_{[1,t]}\right].$$

Recall that Z_v stands for the Markov chain on the lamp graph H_v, and their conditional independence given $L_v(t)$-s. Due to (3.3) and τ_H being strong stationary for H we have for all $v \in G$ that

$$\mathbf{P}[Z_v(L_v(t)) = f_v, \tau_H(v) \leq L_v(t) | X_{[1,t]}] = \pi_H(f_v) \cdot \mathbf{P}[\tau_H(v) \leq L_v(t) | X_{[1,t]}].$$

Now we use that $\tau_H(v)$-s are conditionally independent given the local times to see that

$$\begin{aligned}
\mathbf{P}&\left[X_t^{\diamond} = (\underline{f}, x), \tau^{\diamond} \leq t | X_{[1,t]}\right] \\
&= \mathbf{P}\left[\forall v \in G : Z_v(L_v(t)) = f_v, \tau_H(v) \leq L_v(t), X_t = x, | X_{[1,t]}\right] \\
&= \prod_{v \in G} \pi_H(f_v) \prod_{v \in G} \mathbf{P}\left[\tau_H(v) \leq L_v(t) | X_{[1,t]}\right]
\end{aligned}$$

Note that the second product gives exactly $\mathbf{P}\left[\tau^{\diamond} \leq t | X_{[1,t]}\right]$, yielding

$$(3.10) \quad \mathbf{P}\left[X_t^{\diamond} = (\underline{f}, x), \tau^{\diamond} \leq t\right] = \prod_{v \in G} \pi_H(f_v) \sum_{X_{[1,t]}} \mathbf{P}\left[\tau^{\diamond} \leq t | X_{[1,t]}\right] \mathbf{P}[X_{[1,t]}]$$

As $X_t = x$ remains fixed over the summation, thus summing over all possible $X[1,t]$ trajectories yields

$$\mathbf{P}[X_t^{\diamond} = (\underline{f}, x), \tau^{\diamond} \leq t] = \prod_{v \in G} \pi_H(f_v) \mathbf{P}[X_t = x, \tau^{\diamond} \leq t].$$

To turn the inequality $\tau^\diamond \leq t$ inside the probability to equality can be done the same way as in (3.3) and left to the reader. To see that the vector of halting states $(\underline{h}(f_v(0)), y)$ is a halting state for $\tau^\diamond$ for any $y \in G$ is based on the simple fact that reaching the halting state vector $(h(f_v)_{v \in G}, y)$ means that all the halting states $h(f_v), v \in G$ have been reached on all the lamp graphs $H_v, v \in G$-s. Thus, by definition of the halting states, all the strong stationary times $\tau_H(v)$ have happened. Then, by its definition, $\tau^\diamond$ has happened as well. $\square$

Recall the definition (3.7) of $\tau^d iamond$. Then we can construct a strong stationary time for $H \wr G$, described in the next lemma.

LEMMA 3.5.   *Let $\tau^\diamond$ be the stopping time defined as in Lemma 3.3, and let $\tau_G(x)$ be a strong stationary time for G starting from $x \in G$ and define $\tau_2^\diamond$ by*

$$(3.11) \qquad\qquad \tau_2^\diamond := \tau^\diamond + \tau_G(X_{\tau^\diamond}),$$

where the chain is re-started at τ_G is started from $(\underline{F}_{\tau_d iamond}, X_{\tau_d iamond})$, run independently of the past and τ_G is measured in this walk. Then, $\tau_2^\diamond$ is a strong stationary time for $H \wr G$.

PROOF OF LEMMA 3.5.   The intuitive idea of the proof is based on the fact that τ_G is conditionally independent of τ_H-s and thus the lamp graphs stay stationary after reaching $\tau^\diamond$, and stationarity on G is reached by adding the term $\tau_G(X_{\tau^\diamond})$. The proof is not very difficult but it needs a delicate sequence of conditioning. To have shorter formulas, we write shortly $\mathbf{P}$ for $\mathbf{P}_{(\underline{f}_0, x_0)}$. First we condition on the events $\{\tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\}$ and make use of (3.8) from Lemma 3.3.

(3.12)
$$\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t\right] = \sum_{s \leq t; (\underline{g}, y)} \mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t | \tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\right]$$
$$\cdot \prod_{v \in G} \pi_H(g_v) \cdot \mathbf{P}\left[\tau^\diamond = s, X_s = y\right].$$

Now for the conditional probability inside the sum on the right hand side we have

$$\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t | \tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\right]$$
$$= \mathbf{P}\left[X_t^\diamond = (\underline{f}, x); \tau_G(y) \circ \theta_s = t - s | \tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\right]$$

where $\tau_G(y) \circ \theta_s$ means the time-shift of $\tau_G(y)$ by s, and we also used that τ_G is only depending on y. We claim that

$$\sum_{\underline{g}} \left(\mathbf{P}_{(\underline{g},y)} \left[X^{\diamond}_{t-s} = (\underline{f}, x), \tau_G(y) = t - s \right] \prod_{v \in G} \pi_H(g_v) \right)$$
$$= \mathbf{P}_y[X_{t-s} = x, \tau_G(y) = t - s] \prod_{v \in G} \pi_H(f_v)$$
$$= \pi_G(x) \mathbf{P}_y[\tau_G = t - s] \prod_{v \in G} \pi_H(f_v).$$

The first equality holds true since $\tau_G(y)$ is independent of the lampgraphs and the transition rules of $X^{\diamond}$ on $H \wr G$ tells us that the lamp-chains stay stationary. We omit the details of the proof. The second equality is just the strong stationarity property of τ_G. Thus, using this and rearranging the order of terms on the right hand side of (3.12) we end up with

$$\sum_{s \leq t, y \in G} \mathbf{P}_y[\tau_G = t - s] \mathbf{P}[\tau^{\diamond} = s, X_s = y] \cdot \pi_G(x) \prod_{v \in G} \pi_H(f_v).$$

Then, realizing that the sum is just $\mathbf{P}[\tau^{\diamond} + \tau_G(X_{\tau^{\diamond}}) = t]$ finishes the proof. $\square$

We continue with a lemma which relates the separation distance to the total variation distance: Let us define first

$$(3.13) \qquad d_x(t) := \|P^t(x, \cdot) - \pi(\cdot)\|_{\text{TV}} = \frac{1}{2} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)|.$$

The total variation distance of the chain from stationarity is defined as:

$$d(t) := \max_{x \in \Omega} d_x(t).$$

The next lemma relates the total and the separation distance:

LEMMA 3.6. *For any reversible Markov chain and any state $x \in \Omega$, the separation distance from initial vertex x satisfies:*

$$(3.14) \qquad\qquad\qquad d_x(t) \leq s_x(t)$$
$$(3.15) \qquad\qquad\qquad s_x(2t) \leq 4d(t)$$

PROOF. For a short proof of (3.14) see [2] or [13, Lemma 6.13], and combine [13, Lemma 19.3] with a triangle inequality to conclude (3.15). $\square$

We will also make use of the following lemma: ([13, Corollary 12.6])

LEMMA 3.7.   *For a reversible, irreducible and aperiodic Markov chain,*

$$\lim_{t\to\infty} d(t)^{1/t} = \lambda_*,$$

with $\lambda_* = \max\{|\lambda| : \lambda$ eigenvalue of P, $\lambda \neq 1\}$.

The two fundamental steps to prove Lemma 3.7 are the inequalities stating that for all $x \in \Omega$ we have

(3.16)
$$d_x(t) \leq s_x(t) \leq \frac{\lambda_*^t}{\pi_{\min}},$$
$$|\lambda_2|^t \leq 2d(t)$$

with $\pi_{\min} = \min_{y \in \Omega} \pi(y)$. This inequality follows from [13, Equation (12.11), (12.13)].

We note that Lemma 3.6 implies that the assertion of Lemma 3.7 stays valid if we replace $d(t)^{1/t}$ by the separation distance $s(t)^{1/t}$.

4. Relaxation time bounds.

4.1. *Proof of the lower bound of Theorem 1.3.*   We prove $c_1 = 1/(16 \log 2)$ in the lower bound of the statement of Theorem 1.3. First note that it is enough to prove that $t_{\mathrm{hit}}(G)$ and $|G|t_{\mathrm{rel}}(H)$ are both lower bounds, hence their average is a lower bound as well. First we start showing the latter.

Let us denote the second largest eigenvalue of Q by λ_H and the corresponding eigenfunction by ψ. It is clear that $\mathbf{E}_{\pi_H}(\psi) = 0$ and we can normalize it such that $\mathbf{Var}_{\pi_H}(\psi) = \mathbf{E}_{\pi_H}(\psi^2) = 1$ holds. Let us define

$$\phi : V(H \wr G) \to \mathbb{R}, \quad \phi((\underline{f}, x)) = \sum_{w \in G} \psi(f_w),$$

thus ϕ is actually not depending on the position of the walker, only on the configuration of the lamps. Let $X_t^\diamond = (\underline{F}_t, X_t)$ be the lamplighter chain with stationary initial distribution $\pi^\diamond$. In the sequel we will calculate the Dirichlet form (3.1) for ϕ at time t, first conditioning on the path $X[0, t]$ of the walker:

(4.1)
$$\mathcal{E}_t[\phi] = \frac{1}{2}\mathbf{E}_{\pi^\diamond}[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2]$$
$$= \frac{1}{2}\mathbf{E}_{\pi^\diamond}\left(\mathbf{E}_{\pi^\diamond}[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 | X[0, t]]\right)$$

We remind the reader that in each step of the lamplighter walk, the state of the lamp graph H_v is refreshed both at the departure and arrival site of the walker. Thus, knowing the trajectory of the walker implies that we also know $L_v(t)$, the number of steps made by the Markov chain Z_v on H_v. Moreover, the collection of random walks $(Z_v)_{v \in G}$ on the lamp graphs are independent given $L_v(t)$-s.

We can calculate the conditional expectation on the right hand side of (4.1) by using the argument above and the fact that $\mathbf{E}_{\pi_H}(\psi) = 0$ as follows:

$$(4.2)$$
$$\mathbf{E}_{\pi^\diamond}\left[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 | X[0,t]\right] = \sum_{v \in G} \mathbf{E}_{\pi^\diamond}\left[\left(\psi(Z_v(L_v(t))) - \psi(Z_v(0))\right)^2 \big| L_v(t)\right]$$

Next, the product form of the stationary measure $\pi^\diamond$ ensures that we can move to π_H inside the sum and condition on the starting state $Z_v(0)$:

$$\mathbf{E}_{\pi^\diamond}\left[\left(\psi\left(Z_v(L_v(t))\right) - \psi\left(Z_v(0)\right)\right)^2 \big| L_v(t)\right]$$
$$= 2\mathbf{E}_{\pi_H}\psi^2 - 2\mathbf{E}_{\pi_H}\left[\psi\left(Z_v(0)\right)\mathbf{E}_{Z_v(0)}[\psi\left(Z_v(L_v(t))\right)|Z_v(0)]\right],$$

Since ψ was chosen to be the second eigenfunction for Q, clearly $\mathbf{E}_{Z_v(0)}[\psi\left(Z_v(L_v(t))\right)] = \lambda_H^{L_v(t)}\psi(Z_v(0))$. Using the normalization $\mathbf{E}_{\pi_H}[\psi^2] = 1$, we arrive at

$$(4.3) \qquad \mathbf{E}_{\pi^\diamond}\left[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 | X[0,t]\right] = 2|G| - 2\sum_{v \in G} \lambda_H^{L_v(t)}$$

Since $\sum_{v \in G} L_v(t) = 2t$ and the function λ_H^x is convex, Jensen's inequality implies that

$$\sum_{v \in G} \lambda_H^{L_v(t)} \geq |G| \cdot \lambda_H^{2t/|G|}.$$

Combining this with (4.3) and (4.1) and setting $t := t^* = |G|t_{\mathrm{rel}}(H) = |G|/(1 - \lambda_H)$ we arrive at

$$\mathcal{E}_t(\phi) \leq |G|\left(1 - e^{2\frac{\log \lambda_H}{1 - \lambda_H}}\right) \leq |G|\left(1 - 2^{-4-1}\right),$$

where in the last step we assumed $\lambda_H > 1/2$, since in this case we have $(1 - \lambda_H)^{-1}\log\lambda_H > -2\log 2$. On the other hand, if $\lambda_H < 1/2$, than $t_{\mathrm{rel}}(H) < 2$ and we will use the other lower bound $t_{\mathrm{hit}}(G)$ which is at least of order $|G|$. Dividing by $\mathbf{Var}_{\pi^\diamond}\phi = |G|$, and using the variational characterization of the spectral gap in Lemma 3.1, we get that the spectral gap γ_{t^*} at time t^* satisfies

$$\gamma_{t^*} \leq 1 - 2^{-5}.$$

Since γ_t is by definition the spectral gap of the chain at time t, we have

$$(4.4) \qquad\qquad 1 - \lambda_2(H \wr G)^{t^*} \leq 1 - 2^{-5}.$$

Thus

$$5\log 2 \geq t^* \left(1 - \lambda_2(H \wr G)\right),$$

so we get a lower bound $t_{\mathrm{rel}}(H \wr G) \geq \frac{1}{5\log 2}|G|t_{\mathrm{rel}}(H)$.

To get the lower bound $t_{\mathrm{hit}}(G)/4$ we adjust the proof for $0 - 1$ lamps ($H = \mathbf{Z}_2$) [13, Theorem 19.1] to our setting. First pick a vertex $w \in G$ which maximizes the expected hitting time $\mathbf{E}_{\pi_G}(\tau_w)$. As before, we will use the second eigenfunction ψ with eigenvalue λ_H with $\mathbf{E}_{\pi_H}(\psi) = 0, \mathbf{E}_{\pi_H}(\psi^2) = 1$ and define

$$\phi\left((\underline{f}, x)\right) := \psi(f_w).$$

Easy to see with the same conditioning argument we used in (4.2) and (4.3) that the Dirichlet form at time t equals

$$\mathcal{E}_t(\phi) = 1 - \mathbf{E}_{\pi^\diamond}\left[\lambda_H^{L_w(t)}\right]$$

Now we will show that $\mathbf{E}_{\pi^\diamond}\left[\lambda_H^{L_w(t)}\right] \geq 1/4$. To see this we first note that for any t we have for the hitting time τ_w of $w \in G$

$$\mathbf{E}_v(\tau_w) \leq t + t_{\mathrm{hit}}\mathbf{P}_v[\tau_w > t]$$
$$\mathbf{E}_{\pi_G}(\tau_w) \leq t + t_{\mathrm{hit}}\mathbf{P}_{\pi_G}[\tau_w > t]$$

To see the first line: either the walk hits w before time t, or the expected additional time it takes to arrive at w is bounded by t_{hit} regardless of where it is at time t. The second line follows by averaging over π_G.

Next, [13, Lemma 10.2] states that $t_{\mathrm{hit}} \leq 2\max_v \mathbf{E}_\pi[\tau_v]$ holds for every irreducible Markov chain. We exactly picked w such that it maximizes $\mathbf{E}_{\pi_G}(\tau_v)$, so we have $t_{\mathrm{hit}} \leq 2\mathbf{E}_{\pi_G}[\tau_w]$, so multiplying the previous displayed inequality by 2 gives

$$t_{\mathrm{hit}} \leq 2t + 2t_{\mathrm{hit}}\mathbf{P}_{\pi_G}[\tau_w > t]$$

Now substituting $t = t_{\mathrm{hit}}/4$ and rearranging terms results in

$$\mathbf{P}_{\pi_G}\left[\tau_w > \frac{t_{\mathrm{hit}}}{4}\right] \geq \frac{1}{4}.$$

Since $\{L_w(t_{\mathrm{hit}}/4) = 0\} = \{\tau_w > t_{\mathrm{hit}}/4\}$, we can use this inequality to obtain the upper bound

$$\mathcal{E}_{t_{\mathrm{hit}}/4}(\phi) = 1 - \mathbf{E}_{\pi^\diamond}\left[\lambda_H^{L_w(t_{\mathrm{hit}}/4)}\right] \leq 1 - \mathbf{P}_{\pi_G}[\tau_w > t_{\mathrm{hit}}/4] \leq 1 - \frac{1}{4} = \frac{3}{4}.$$

Analogous to the last lines of the proof of the lower bound above, (see (4.4)) we obtain the other desired lower bound:

$$t_{\text{rel}}(H) \geq \frac{1}{2\log 2}\frac{1}{4}t_{\text{hit}}(G).$$

Putting together the two bounds we get

$$t_{\text{rel}}(H \wr G) \geq \max\left\{\frac{1}{8\log 2}t_{\text{hit}}(G), \frac{1}{5\log 2}|G|t_{\text{rel}}(H)\right\}$$
$$\geq \frac{1}{16\log 2}\left(t_{\text{hit}}(G) + |G|t_{\text{rel}}(H)\right).$$

4.2. *Proof of the upper bound of Theorem* (1.3). To prove the upper bound, we will estimate the tail behavior of the strong stationary time $\tau_2^\diamond$ in Lemma 3.5, relate it to $s^\diamond(t)$, the separation distance on $H \wr G$, and then use Lemmas 3.7 and 3.6 to see that $s^\diamond(t)^{1/t} \to \lambda^\diamond$. We will use separation-optimal τ_H and τ_G in the construction of $\tau_2^\diamond$. The existence is guaranteed by Theorem 3.2. We will use $\mathbf{P}$ for $\mathbf{P}_{(\underline{f},x)}$ for notational convenience. Combining (3.5) and the fact that $\tau^\diamond$ happens when all the stopping times $\tau_H(v), v \in G$ have happened on the lamp graphs, by union bound we have for any choice of $0 < \alpha < 1$

$$s^\diamond_{(\underline{f},x)}(t) \leq \mathbf{P}_{(\underline{f},x)}\left[\tau_2^\diamond > t\right] \leq \mathbf{P}_{(\underline{f},x)}\left[\tau^\diamond > \alpha t\right] + \mathbf{P}_{(\underline{f},x)}\left[\tau_2^\diamond > t|\tau^\diamond < \alpha t\right]$$

$$(4.5) \qquad \leq \quad \mathbf{P}[\tau_{\text{cov}} > \alpha t/3]$$

$$(4.6) \qquad + \mathbf{P}\left[\exists w \in G : L_w(\alpha t) < \tfrac{\alpha t}{2|G|}\big|\tau_{\text{cov}} \leq \alpha t/3\right]$$

$$(4.7) \qquad + \mathbf{P}\left[\exists w \in G : \tau_H(w) > L_w(\alpha t) \,\big|\forall v \in G : L_v(\alpha t) \geq \frac{\alpha t}{2|G|}\right]$$

$$(4.8) \qquad + \max_{(\underline{g},y)}\mathbf{P}_{(\underline{g},y)}\left[\tau_G > (1-\alpha)t\right]$$

Namely, there are four possibilities: The first option is that there is a state $w \in G$ which is not hit yet, i.e. the cover time of the chain is greater than $\alpha t/3$: giving the term (4.5). The constant $1/3$ could have been chosen differently, we picked $\alpha t/3$ such that the remaining $2\alpha t/3$ time still should be enough to gain large enough local time on the vertices $v \in G$. Secondly, even though any state w on the graph G is reached before time $\alpha t/3$, the remaining time was not enough to have at least $\alpha t/2|G|$ many moves on some lamp graph $H(w)$, term (4.6). The third option is that even though there have been many moves on all the lamp graphs, there is a vertex $w \in G$ where $\tau_H(w)$ has not happened yet, yielding the term (4.7). We will handle

the three terms separately. The fourth term handles the case where the strong stationary time τ_G is too large. (For convenience, we will write t instead of αt in estimating the first three formulas.)

We can estimate the first term (4.5) by a union bound:

$$(4.9) \qquad \mathbf{P}[\tau_{\mathrm{cov}} > t/3] \leq \mathbf{P}[\exists w : \tau_w > t/3] \leq |G| 2 e^{-\frac{\log 2}{6} \frac{t}{t_{\mathrm{hit}}}},$$

where t_{hit} is the maximal hitting time of the graph G, see (1.4). To see this, use Markov's inequality on the hitting time of $w \in G$ to obtain that for all starting states $v \in G$ we have $\mathbf{P}_v[\tau_w > 2t_{\mathrm{hit}}] \leq 1/2$, and then run the chain in blocks of $2t_{\mathrm{hit}}$. In each block we hit w with probability at least $1/2$, so we have

$$\mathbf{P}_v[\tau_w > K(2t_{\mathrm{hit}})] \leq \frac{1}{2^K}.$$

To get it for general t, we can move from $\lfloor t/t_{\mathrm{hit}} \rfloor$ to t/t_{hit} by adding an extra factor of 2, and (4.9) immediately follows by a union bound.

For the third term (4.7) we claim the following upper bound holds:
(4.10)
$$\mathbf{P}\left[\exists w : \tau_H(w) \geq L_w(t) \big| \forall v : L_v(t) > \tfrac{t}{2|G|}\right] \leq |G| \frac{1}{\pi_{\min}(H)} e^{-\frac{t}{2|G|t_{\mathrm{rel}}(H)}}.$$

To see this we estimate the probability of the event $\{\tau_H(w) \geq L_w(t) \big| L_w(t) \geq \frac{t}{2|G|}\}$ on a single lamp graph and then use a union bound to lose a factor $|G|$ and arrive at the right hand side. First note that according to Lemma 3.7, the tail of the strong stationary time τ_H is driven by λ_H^t. More precisely, using the inequality (3.16) we have that for any initial state $h \in H$:

$$\mathbf{P}_h\left[\tau_H(w) \geq \frac{t}{2|G|}\right] \leq s_H\left(\frac{t}{2|G|}\right) \leq \frac{1}{\pi_{\min}(H)} \lambda_H^{t/2|G|}$$
$$\leq \frac{1}{\pi_{\min}(H)} \exp\left\{-\frac{(1-\lambda_H)t}{2|G|}\right\}.$$

Since we have made at least $L_w(t) \geq \frac{t}{2|G|}$ steps on each coordinate, the claim (4.10) follows. The fourth term (4.8) can be handled analogously and yields an error probability $\exp\{-ct/t_{\mathrm{rel}}(G)\}$ which then, taking the power of $1/t$ and limit as in Lemma 3.7, will lead to a term of order $t_{\mathrm{rel}}(G)$. Then, taking into account that $t_{\mathrm{rel}}(G) \leq ct_{\mathrm{mix}}(G) \leq Ct_{\mathrm{hit}}(G)$ holds for any lazy reversible chain (see e.g. [13, Chapter 11.6,12.4]), we can ignore this term.

The intuition behind the estimates below for the second term (4.6) is that since the total time was at least $2t/3$ after hitting, regularity of G implies that the average number of moves on a lamp graph equals $4t/(3|G|)$ by the

double refreshment at any visit to the vertex. Thus, the probability of having less than $t/(2|G|)$ moves must be small.

More precisely, we introduce the excursion-lengths to a vertex $w \in G$: Let us define for all $w \in G$ the first return time to state w as

$$R(w) = \inf\{t > 0 : X_t = w | X_0 = w\}.$$

The strong Markov property implies that the length of the i-th excursion $R_i(w)$, defined as the time spent between the $(i-1)$th and ith visit to w, are i.i.d random variables distributed as the first return time $R(w)$.

Thus, having not enough local time on some site $w \in G$ can be expressed in terms of the excursion lengths $R_i(w)$-s as follows:

$$(4.11) \quad \mathbf{P}\left[\exists w : L_w(t) \le \tfrac{t}{2|G|} \big| \tau_{\text{cov}} \le \frac{t}{3}\right] \le |G| \max_{w \in G} \mathbf{P}_w\left[\sum_{i=1}^{\lceil t/4|G| \rceil} R_i(w) \ge \frac{2t}{3}\right],$$

since conditioning on hitting before $t/3$ ensures that we had at least $2t/3$ steps to gain the $t/4|G|$ visits to w, and by the definition (3.6) of $L_v(t)$, this guarantees that $L_w(t) < t/2|G|$.

We aim to estimate the right hand side of (4.11) using the moment generating function of the first return time $R(w)$. To be able to carry out the estimates we need a bound on the tail behavior of the return times. A very similar argument can be used to the one we used for the tail of the cover time (4.9), namely the following holds:

$$\mathbf{P}_w\left[R(w) > 2t_{\text{hit}} + 1\right] = \mathbf{P}_w\left[X_1 \ne w\right] \mathbf{E}\left[\mathbf{P}_{X_1}(\tau_w > 2t_{\text{hit}} | X_1)\right]$$
$$\le \frac{\mathbf{E}[\mathbf{E}_{X_1}(\tau_w)]}{2t_{\text{hit}}} \le \frac{1}{2}.$$

Running the chains in blocks of $2t_{\text{hit}} + 1$, one can see that in each block the chain has a chance at least $1/2$ to return to w, so we have for each $t > 2t_{\text{hit}} + 1$

$$(4.12) \qquad \mathbf{P}\left[R_w > t\right] \le 2\left(\frac{1}{2}\right)^{\frac{t}{2t_{\text{hit}}+1}} = 3\exp\left\{-\frac{\log 2}{2}\frac{t}{t_{\text{hit}}}\right\},$$

where the factor 3 comes from ignoring to take the integer part of t/t_{hit} and neglecting the $+1$ term in the denominator.

We can use this tail behavior to estimate the moment generating function

$$\mathbf{E}\left[e^{\beta R_w}\right] \le e^{2\beta|G|} + \mathbf{E}[e^{\beta R_w}\mathbf{1}\{R_w > 2|G|\}]$$
$$= e^{2\beta|G|} + \int_{e^{2\beta|G|}}^{\infty} \mathbf{P}[e^{\beta R_w} > z]\mathrm{d}z$$

where we cut the expectation at $2|G|$. Using the bounds in (4.12) yields:

$$\mathbf{E}\left[e^{\beta R_w}\right] \le e^{2\beta|G|} + \int_{e^{2\beta|G|}}^{\infty} \mathbf{P}[R_w > \frac{1}{\beta}\log z]\mathrm{d}z$$

$$\le e^{2\beta|G|} + 2\int_{e^{2\beta|G|}}^{\infty} z^{-\frac{\log 2}{2\beta t_{\mathrm{hit}}}}\,\mathrm{d}z.$$

Setting arbitrary $\beta < \log 2/(2t_{\mathrm{hit}})$ makes the second term integrable, and with the special choice of $\beta = \frac{\log 2}{4t_{\mathrm{hit}}}$ we obtain the following estimate:

$$(4.13) \qquad \mathbf{E}\left[e^{\beta R_w}\right] \le e^{2\beta|G|} + 2e^{-2\beta|G|} \le e^{(2+\delta)\beta|G|}$$

with an appropriately chosen $0 < \delta < 1/3$. Now we apply Markov's inequality to the function $e^{\beta\sum_{i=1}^{t/4|G|} R_i(w)}$ to estimate the right hand side of (4.11):

$$(4.14) \qquad \mathbf{P}_w\left[\sum_{i=1}^{\lceil t/4|G|\rceil} R_i(w) \ge 2t/3\right] \le e^{-\frac{2}{3}\beta t} \cdot \mathbf{E}\left[e^{\beta R_w}\right]^{\frac{t}{4|G|}},$$

where we also used the independence of the excursions $R_i(w)$-s. Using the estimate in (4.13) to bound the right hand side we gain that

$$\mathbf{P}_w\left[\sum_{i=1}^{\lceil t/4|G|\rceil} R_i(w) \ge 2t/3\right] \le e^{-\frac{2}{3}\beta t} \cdot e^{(2+\delta)\beta|G|\cdot\frac{t}{4|G|}}$$

$$(4.15)$$

$$\le e^{-\frac{1-\tilde{\delta}}{6}\beta t} = \exp\left\{-\frac{(1-\tilde{\delta})\log 2}{24}\frac{t}{t_{\mathrm{hit}}}\right\},$$

where we used $\beta = \log 2/(4t_{\mathrm{hit}})$, and modified $\tilde{\delta} := 3\delta/2 \le 1/2$. Using the relation of the local time to the excursion lengths in (4.11) we finally get that the second term (4.6) is bounded from above by

$$(4.16) \qquad \mathbf{P}\left[\exists w : L_w(t) \le \frac{t}{2|G|}\Big|t_{\mathrm{cov}} \le t/3\right] \le |G|\exp\left\{-\frac{\log 2}{48}\frac{t}{t_{\mathrm{hit}}}\right\}.$$

Mind that all the estimates (4.9), (4.10) and (4.16) were independent of the initial state $(\underline{f}, x) \in H \wr G$, so using the second inequality in (3.16) and maximizing over all possible initial states yields us

$$(4.17)$$

$$|\lambda_2|^t \le 2d^\diamond(t) \le 2s^\diamond(t) \le \frac{4|G|}{\pi_{\min}(H)}\exp\left\{-\frac{t}{2|G|t_{\mathrm{rel}}(H)}\right\}$$

$$+ 4\exp\left\{-\frac{(1-\tilde{\delta})\log 2}{24}\frac{t}{t_{\mathrm{hit}}}\right\} + 4|G|\exp\left\{-\frac{\log 2}{6}\frac{t}{t_{\mathrm{hit}}}\right\}$$

In the final step we apply Lemma 3.7: we take the power $1/t$ and limit as t tends to infinity with fixed graph sizes $|G|$ and $|H|$ on the right hand side of (4.2) to get an upper bound on λ_2. Then we use that $(1 - e^{-x}) \leq x + o(x)$ for small x and obtain the bound on t_{rel} finally:

$$t_{\mathrm{rel}}(H \wr G) \leq \max \left\{ 2|G|t_{\mathrm{rel}}(H), \frac{48}{\log 2} t_{\mathrm{hit}} \right\}.$$

This finishes the proof of the upper bound on the relaxation time.

5. Mixing time bounds. Based on the fact that H has a separation-optimal strong stationary time τ_H, the idea of the proofs is to relate the separation distance to the tail behavior of the stopping times $\tau^\diamond$ and $\tau_2^\diamond$ constructed in Lemmas 3.3 and 3.5, respectively. Then these estimates are turned into bounds of the total variation distance using the relations in Lemma 3.6. This method gives us the upper bound in (1.7) and the corresponding lower bound under the assumption **(A)**. For the lower bound without the assumption, we will need slightly different methods.

5.1. *Proof of the upper bound of Theorem 1.4.* The idea of the proof is to use appropriate top quantiles of the strong stationary time τ_H on H, and give an upper bound on the tail of the strong stationary time $\tau_2^\diamond$ defined in Lemma 3.5. Throughout, we (only) need that τ_H and τ_G in the construction of $\tau_2^\diamond$ are separation-optimal. The existence is guaranteed by Theorem 3.2. (Thus, τ_H does not necessarily possess halting states.)

Let us denote the worst-case initial state top ε-quantile of a stopping time τ as

$$(5.1) \qquad t_\varepsilon^{\mathrm{quant}}(\tau) := \max_{y \in \Omega} \inf\{t : \mathbf{P}_y[\tau > t] \leq \varepsilon\}$$

We continue with the definition of the blanket time:

$$(5.2) \qquad \mathcal{B}_2 := \inf_t \left\{ \forall v, w \in G : \frac{L_w(t)}{L_v(t)} \leq 2 \right\}.$$

Let us further denote

$$(5.3) \qquad \mathbf{B}_2 := \max_{v \in G} \mathbf{E}_v(\mathcal{B}_2)$$

It is known from [8] that there exist universal constants C and C' such that $C' t_{\mathrm{cov}} \leq \mathbf{B}_2 \leq C t_{\mathrm{cov}}$.

Thus, our first goal is to show that at time

$$8\mathbf{B}_2 + |G|t_{1/16|G|}^{\mathrm{quant}}(\tau_H) + t_{1/16}^{\mathrm{quant}}(\tau_G) =: 8\mathbf{B}_2 + |G|t_H^u + t_G =: t^\diamond$$

we have for any starting state $(\underline{f}, x)$ that

$$(5.4) \qquad\qquad \mathbf{P}_{(\underline{f},x)}[\tau_2^\diamond > t^\diamond] \leq \frac{1}{4}.$$

We remind the reader that $\tau_2^\diamond = \tau^\diamond + \tau_G(X_{\tau^\diamond})$ and thus the following union bound holds:

$$(5.5) \qquad \begin{aligned} \mathbf{P}\left[\tau_2^\diamond > t^\diamond\right] &\leq \mathbf{P}\left[\mathcal{B}_2 > 8\mathbf{B}_2\right] + \mathbf{P}\left[\tau^\diamond > |G|t_H^u + 8\mathbf{B}_2|\mathcal{B}_2 \leq 8\mathbf{B}_2\right] \\ &+ \max_{v \in G} \mathbf{P}_v\left[\tau_G > t_G|\mathcal{B}_2 \leq 8\mathbf{B}_2, \tau^\diamond < 8\mathbf{B}_2 + |G|t_H^u\right], \end{aligned}$$

where in the third term we mean that we restart the chain after time $8\mathbf{B}_2 + |G|t_H^u$, and measure τ_G starting from there. The first term on the right hand side is less than $1/8$ by Markov's inequality, the third is less than $1/16$ by the definition of the worst case quantile. The second term can be handled by conditioning on the local time sequence of vertices and on the blanket time: (for shorter notation we introduce $t_1 := |G|t_H^u + 8\mathbf{B}_2$)
(5.6)
$$\mathbf{P}\left[\tau^\diamond > |G|t_H^u + 8\mathbf{B}_2|\mathcal{B}_2 \leq 8\mathbf{B}_2\right]$$
$$= \sum_{s \leq 8\mathbf{B}_2, (L_v(t_1))_v} \mathbf{P}\left[\exists w \colon \{\tau_H(w) > L_w(t_1)\}\big| (L_v(t_1))_v, \mathcal{B}_2 = s\right] \cdot \mathbf{P}\left[(L_v(t_1))_v, \mathcal{B}_2 = s\right]$$

The fact that $\mathcal{B}_2 \leq 8\mathbf{B}_2$ means that the number of visits to every vertex $v \in G$ must be greater than half of the average, which is at least $\frac{1}{2}t_H^u$. Since $L_w(t)$ is twice the number of visits by (3.6), $\{\tau_H(w) > L_w(t_1)\} \subseteq \{\tau_H(w) > t_H^u\}$. By the definition of the quantiles,

$$\mathbf{P}_h\left[\tau_H(w) > t_H^u\right] \leq \frac{1}{16|G|}$$

holds for every $h \in H$ and $w \in G$. Applying a simple union bound on the conditional probability on the right hand side of (5.6) yields

$$\begin{aligned} \mathbf{P}_{(\underline{f},x)}\left[\tau^\diamond > t_1|\mathcal{B}_2 \leq 8\mathbf{B}_2\right] &\leq \sum_{s \leq 8\mathbf{B}, (L_v(t_1))_v} \left(|G|\frac{1}{16|G|}\right) \mathbf{P}\left[(L_v(t_1))_v, \mathcal{B}_2 = s\right] \\ &\leq \frac{1}{16}, \end{aligned}$$

where we used that the sum of the probabilities on the right hand side is at most 1. Combining these estimates with (5.5) yields (5.4). It remains to relate the worst-case quantiles to the total variation mixing times. Here we will make use of the separation-optimal property of τ_H and τ_G. Now

just consider the walk on G. Let us start the walker on G from an initial state $x_0 \in G$ for which the maximum is attained in the definition (5.1) of the quantile $t^{\text{quant}}_{1/16}(\tau_G)$. Then, by (3.15) we have that one step before the quantile we have

$$\frac{1}{16} \le \mathbf{P}_{x_0}\left[\tau_G > t^{\text{quant}}_{1/16}(\tau_G) - 1\right] = s_{x_0}\left(t^{\text{quant}}_{1/16}(\tau_G) - 1\right)$$
$$\le 4d\left(\frac{1}{2}(t^{\text{quant}}_{1/16}(\tau_G) - 1)\right).$$

This immediately implies that $\frac{1}{2}(t^{\text{quant}}_{1/16}(\tau_G) - 1) \le t_{\text{mix}}\left(G, \frac{1}{64}\right)$. By the sub-multiplicative property of the total variation distance $d(kt) \le 2^k d(t)^k$ we have that $t_{\text{mix}}(G, \frac{1}{64}) \le 6t_{\text{mix}}\left(G, \frac{1}{4}\right)$. So we arrive at

$$(5.7) \qquad\qquad t^{\text{quant}}_{1/16}(\tau_G) - 1 \le 12t_{\text{mix}}(G)$$

Similarly, starting all the lamps from the position h_0 where the maximum is attained in the definition of $t^u_H = t^{\text{quant}}_{1/16|G|}(\tau_H)$, one step before the quantile we have

$$\frac{1}{16|G|} \le \mathbf{P}_{h_0}\left[\tau_H > t^u_H - 1\right] = s_{h_0}\left(t^u_H - 1\right) \le 4d\left((t^u_H - 1)/2\right)$$

So we have

$$(5.8) \qquad\qquad \frac{1}{2}(t^{\text{quant}}_{1/16|G|}(\tau_H) - 1) \le t_{\text{mix}}\left(H, \frac{1}{64|G|}\right).$$

On the other hand, on the whole lamplighter chain $H \wr G$ we need the other direction: For every starting state $(\underline{f}, x)$ (3.14) and (5.4) implies that

$$d_{(\underline{f},x)}(t^\diamond) \le s_{(\underline{f},x)}(t^\diamond) \le \mathbf{P}_{(\underline{f},x)}\left[\tau^\diamond_2 > t^\diamond\right] \le 1/4$$

Maximizing over all states $(\underline{f}, x)$ yields

$$(5.9) \qquad\qquad t_{\text{mix}}(H \wr G) \le t^\diamond.$$

Putting the estimates in (5.7) and (5.8) to (5.9), we get that

$$t_{\text{mix}}(H \wr G) \le t^\diamond \le 8\mathbf{B}_2(G) + 12t_{\text{mix}}(G) + 1 + 2|G|\left(t_{\text{mix}}\left(H, \frac{1}{64|G|}\right) + \frac{1}{2}\right).$$

Since $\mathbf{B}_2(G) \le Ct_{\text{cov}}(G)$, and $t_{\text{mix}}(G) \le 2t_{\text{hit}}(\mathbf{G}) \le 2t_{\text{cov}}(G)$ for any G (see for instance [13]), the assertion of Theorem 1.4 follows with $C_2 = 8(C + 3)$,

where C is the universal constant relating the blanket time $\mathbf{B}_2$ to the cover time t_{cov} in [8].

We remark why we did not make the constant C_2 explicit: If the blanket time $\mathbf{B}_2$ were not used in our estimates, the error probability that some vertex $w \in G$ does not have enough local time would need to be added. This, similarly to the term (4.6) behaves like $|G|e^{-c\left(t_{\text{cov}}+|G|t_{\text{mix}}(H,\frac{1}{G})\right)/t_{\text{hit}}}$. If we do not assume anything about the relation of $t_{\text{hit}}(G)$ and $t_{\text{cov}}(G)$ and on $t_{\text{mix}}(H,\frac{1}{G})$, then this error term will not necessarily be small. For example, if G_n is a cycle of length n, H_n is a sequence of expander graphs, then $t_{\text{cov}}(G_n) = t_{\text{hit}}(G_n) = \Theta(n^2)$, and $t_{\text{mix}}(H,\frac{1}{G}) = \log|H| \cdot \log|G| = \log|H|\log n$, and we see that the term is not small if $\log|H| = o(n/\log n)$.

5.2. *Proof of the lower bounds of Theorem 1.4.* As we did with the relaxation time, it is enough to prove that all the bounds are lower bounds separately, then take an average. First we start showing that the upper bound is sharp in 1.7 under the assumption that there is a strong stationary time τ_H with halting states.

5.2.1. *Lower bound under Assumption (A).* We first aim to show that $c \, |G|t_{\text{mix}}(H,\frac{1}{|G|}) \leq t_{\text{mix}}(H \wr G)$.

Consider the stopping time $\tau^\diamond$ constructed in Lemma 3.3. Corollary 3.4 tells us that the tail of $\tau^\diamond$ lower bounds the separation distance at time t. We again emphasize that this bound holds only if τ_H in the construction of $\tau^\diamond$ is not only separation optimal but it also *has a halting state*. Our first goal is to lower bound the tail of $\tau^\diamond$, then relate it to the total variation distance.

First set

$$(5.10) \qquad t_H^l := t_{|G|^{-1/2}/2}^{\text{quant}}(\tau_H) - 1, \quad t^\diamond := \frac{1}{4}|G|t_H^\ell,$$

clearly this time is nontrivial if $t_{|G|^{-1/2}/2}^{\text{quant}}(\tau_H) \neq 1$. We handle the case if it equals 1 later. We can estimate the upper tail of $\tau^\diamond$ by conditioning on the number of moves on the lamp graphs $H_v, v \in G$:

(5.11)
$$\mathbf{P}\left[\tau^\diamond > t^\diamond\right] \geq \mathbf{P}\left[\exists w \in G : \tau_H(w) > L_w(t^\diamond)\right]$$
$$\geq \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in G : \tau_H(w) > L_w(t^\diamond)\big|(L_v(t^\diamond))_v\right]\mathbf{P}\left[(L_v(t^\diamond))_v\right]$$

For each sequence $(L_v(t^\diamond))_{v \in G}$ we define the random set

$$S_{(L_v)_v} := \left\{w \in G : L_w(t^\diamond) \leq t_H^\ell\right\}$$

Since $\sum_v L_v(t^\diamond) = 2t^\diamond = \frac{1}{2}|G|t_H^\ell$, we have that for arbitrary local time configuration $(L_v(t^\diamond))_v$,

$$(5.12) \qquad\qquad |S_{(L_v)_v}| \geq |G|/2.$$

Thus we can lower bound (5.11) by restricting the event only to those $w \in G$ coordinates which belong to this set, i.e. whose local time is small:
(5.13)
$$\mathbf{P}\left[\tau^\diamond > t^\diamond\right] \geq \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > L_w(t^\diamond)\big|(L_v(t^\diamond))_v\right] \mathbf{P}\left[(L_v(t^\diamond))_v\right]$$
$$\geq \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > t_H^\ell\big|(L_v(t^\diamond))_v\right] \mathbf{P}\left[(L_v(t^\diamond))_v\right],$$

where in the second line we used that for $w \in S_{(L_v)_v}$ we have $\{\tau_H(w) > L_w(t^\diamond)\} \supseteq \{\tau_H(w) > t_H^\ell\}$. Conditioned on the sequence $(L_v(t^\diamond))_v$, the times $\tau_H(w)$ for $w \in S_{(L_v)_v}$ are independent. On each lamp graph $H(v)$ let us pick the starting state to be $h_0 \in H$ where the maximum is attained in the definition of $t_{|G|^{-1/2}/2}^{\text{quant}}(\tau_H)$. Since t_H is one step *before* the quantile, we have

$$(5.14) \qquad \mathbf{P}_{h_0}\left[\tau_H(w) > t_{|G|^{-1/2}}^{\text{quant}}(\tau_H) - 1\right] \geq |G|^{-1/2}/2.$$

We need to start the lamp-chains from the worst-case scenario $h_0 \in H$ for two reasons: First, we needed to define the quantile as in (5.1) to be able to relate it to the total variation mixing time on H, see below. Then, the fact that $t_\varepsilon^{\text{quant}}$ was defined as the worst-case starting state quantile means that for other starting states the quantile may be smaller, and the lower bound can possibly fail.

Combining (5.14) with (5.12) and the conditional independence gives us the following stochastic domination from below to the event in (5.13)

$$\mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > t_H^\ell\big|(L_w(t^\diamond))_w\right] \geq \mathbf{P}[V > 0],$$

where V is a Binomial random variable with parameters $\left(|G|/2, |G|^{-1/2}/2\right)$. Clearly, for $|G| > 8 > 16(\log 2)^2$ we have

$$\mathbf{P}[V > 0] = 1 - \left(1 - \frac{1}{2|G|^{1/2}}\right)^{|G|/2} \geq 1 - e^{-\frac{|G|^{1/2}}{4}}.$$

Combining this with (5.13) and summing over all possible $(L_v(t^\diamond))_{v \in G}$ sequences we easily get that

$$\mathbf{P}\left[\tau^\diamond > t^\diamond\right] \geq 1 - e^{-\frac{|G|^{1/2}}{4}}.$$

Then, by Corollary 3.4 we have

$$s^{\diamond}_{(\underline{h}_0,x)}(t^{\diamond}) \geq 1 - e^{-\frac{|G|^{1/2}}{4}}.$$

In the next few steps we relate the tail of $\tau^{\diamond}$ and τ_H to the mixing time of the graphs. First, combining the previous inequality with (3.15) implies that for the starting state $(\underline{h}_0, x)$ the following inequalities hold:

$$1 - e^{-|G|^{1/2}/4} \leq s^{\diamond}_{(\underline{h}_0,x)}(t^{\diamond}) \leq 4d^{\diamond}(t^{\diamond}/2).$$

These immediately imply

$$(5.15) \qquad t_{\mathrm{mix}}(H \wr G, \tfrac{1}{8}) \geq \frac{1}{2}t^{\diamond} = \frac{1}{8}|G|t^{\ell}_H$$

Now we will relate $t^{\ell}_H = t^{\mathrm{quant}}_{|G|^{-1/2}/2}(\tau_H) - 1$ to the mixing time on H. Since t_H investigates the worst case initial-state scenario, by inequality (3.5) for any starting state $h \in H$ we have

$$s_h(t_H + 1) \leq \mathbf{P}_h\left[\tau_H \geq t_H + 1\right] \leq |G|^{-1/2}/2$$

Using $d_h(t) \leq s_h(t)$ (see Lemma 3.6) and maximizing over all $h \in H$ we get that

$$(5.16) \qquad d_H(t_H + 1) \leq |G|^{-1/2}/2.$$

On the other hand, the total variation distance for any Markov chain has the following sub-multiplicative property for any integer k, see [13, Section 4.5]:

$$(5.17) \qquad d(kt) \leq 2^k d(t)^k.$$

Taking $t = t_H + 1$ and combining with (5.16) we have that

$$d_H(2(t_H + 1)) \leq 4d_H(t_H + 1)^2 \leq 4\frac{1}{4|G|},$$

which immediately implies

$$t_{\mathrm{mix}}(H, 1/|G|) \leq 2(t_H + 1).$$

Combining this with (5.15) yields the desired lower bound:

$$\frac{1}{16}|G|\left(t_{\mathrm{mix}}\left(H, \tfrac{1}{|G|}\right) - 2\right) \leq t_{\mathrm{mix}}(H \wr G, \tfrac{1}{8}).$$

Mind that the term -2 in the brackets can be dropped when picking a possibly smaller constant and take the graph large enough. The case when $t^{\text{quant}}_{|G|^{-1/2}/2}(\tau_H) = 1$ can be handled the following way: first mind that we can exchange the quantile for arbitrary $0 < \alpha < 1$, and look at the proof with $t^{\text{quant}}_{|G|^{-\alpha}/2}(\tau_H)$. If this is still $= 1$ for all α, that means that $\tau_H = 1$ a.s. In this case, it is enough to hit the vertices to mix immediately and thus the mixing time $|G| t_{\text{mix}}(H)$ is of smaller order than the cover time $t_{\text{cov}}(G)$. The case when $|G| \leq 8$ but $|H| \to \infty$ is easy to see since in this case $t_{\text{mix}}(H, \frac{1}{|G|}) \leq 2 t_{\text{mix}}(H)$ and one can argue that mixing on $H \wr G$ requires mixing on a single lamp graph H_w for a fixed $w \in G$. Thus the lower bound remains valid.

The cover time of G is already a lower bound for the $0 - 1$ lamps case by [18], hence also for general lamps, but, for completeness, we adjust the proof in [13, Theorem 19.2] to our setting. By Lemma 3.3 we can estimate the separation distance on $H \wr G$ as

$$
\begin{aligned}
s^{\diamond}_{(\underline{f},x)}(t) &\geq \mathbf{P}_{(\underline{f},x)}\left[\tau^{\diamond} > t\right] \\
\text{(5.18)} \qquad &\geq \mathbf{P}_{(\underline{f},x)}\left[\exists w \in G : \tau_H(w) > L_w(t)\right] \\
&\geq \mathbf{P}_{(\underline{f},x)}\left[\exists w \in G : L_w(t) = 0\right] = \mathbf{P}_{(\underline{f},x)}\left[\tau_{\text{cov}} > t\right].
\end{aligned}
$$

Now, using the submultiplicativity of $d(t)$ in (5.17) and the relation of the separation distance and the total variation distance in (3.15), we have that at time $8 t_{\text{mix}}(H \wr G, 1/4)$:

$$
s^{\diamond}_{(\underline{f},x)}\left(8 t_{\text{mix}}(H \wr G, \tfrac{1}{4})\right) \leq 4 d^{\diamond}\left(4 t_{\text{mix}}(H \wr G, \tfrac{1}{4})\right) \leq 4 \frac{2^4}{4^4} \leq \frac{1}{4}
$$

Combining with (5.18) yields that for every starting state we have

$$
\mathbf{P}_{(\underline{f},x)}\left[\tau_{\text{cov}} > 8 t_{\text{mix}}(H \wr G, 1/4)\right] \leq 1/4.
$$

Thus, run the chain in blocks of $8 t_{\text{mix}}(H \wr G, 1/4)$ and conclude that in each block it covers with probability at least $3/4$. Thus, the cover time is dominated by $8 t_{\text{mix}}(H \wr G, 1/4)$ times a geometric random variable with success probability $3/4$, so we have

$$
\mathbf{E}_{(\underline{f},x)}\left[\tau_{\text{cov}}\right] \leq 11 t_{\text{mix}}(H \wr G, 1/4).
$$

Maximizing the left hand side over all possible starting states yields $t_{\text{cov}}(G) \leq 11 t_{\text{mix}}(H \wr G, 1/4)$, finishing the proof.

5.2.2. *Proof of the lower bound of Theorem 1.4, without assumption* **(A)**. Now we turn to the general case and first show that $c\ t_{\mathrm{rel}}(H)|G|\log|G|$ is a lower bound. No laziness assumption on the chain on H is needed to get this bound. We will use a distinguishing function method. Namely, take an eigenfunction ϕ_2 of the transition matrix Q on H corresponding to the second eigenvalue λ_H. Then let us define $\psi : H \wr G \to \mathbb{C}$:

$$(5.19) \qquad\qquad \psi((\underline{f}, x)) := \sum_{v \in G} \phi_2(f_v).$$

One can always normalize such that

$$\mathbf{E}_{\pi^\diamond}(\psi) = \sum_{v \in G} \mathbf{E}_\pi[\phi_2] = 0 \quad \mathbf{Var}_{\pi^\diamond}(\psi) = \sum_{v \in G} \mathbf{Var}_\pi(\phi_2) = |G| \cdot 1$$

This normalization has two useful consequences: First, by Chebyshev's inequality, the set $A = \{\psi < 2|G|^{1/2}\}$ has measure at least $3/4$ under stationarity. Second, $\phi_2(g_0) := \max_{g \in H} \phi_2(g) > 1$, otherwise the variance would be less than 1. We aim to show that the set A has measure less then $1/2$ at time $ct_{\mathrm{rel}}(H)|G|\log|G|$ and then we are done by using the following characterization of the total variation distance, see [2, 13]:

$$\|\nu - \mu\|_{\mathrm{TV}} = \sup_{A \subset \Omega} \{\nu(A) - \mu(A)\}.$$

Let us start all the lamp graphs from $g_0 \in H$ where the maximum is attained for ϕ_2. Then we can condition on the local time sequence and use the eigenvalue property of ϕ_2 to obtain

$$(5.20) \qquad
\begin{aligned}
\mathbf{E}_{(\underline{g}_0, x)}\left[\psi((\underline{F}_t, X_t))\right] &= \mathbf{E}\left[\mathbf{E}\left[\sum_{w \in G} \phi_2(F_w(t))\,|\,(L_v(t))_v\right]\right] \\
&= \phi_2(g_0)\mathbf{E}_x\left[\sum_{w \in G} \lambda_H^{L_w(t)}\right].
\end{aligned}$$

Since $\sum_v L_v(t) = 2t$, we can apply Jensen's inequality on the function $y \mapsto \lambda_H^y$ to get a lower bound on the expectation:

$$(5.21) \qquad \mathbf{E}_x\left[\sum_{w \in G} \lambda_H^{L_w(t)}\right] \geq |G|\lambda_H^{\frac{2t}{|G|}} = |G|\left(1 - \frac{1}{t_{\mathrm{rel}}(H)}\right)^{\frac{2t}{|G|}}.$$

By giving a lower bound on the right hand side we must assume here that $\lambda_H > 0$, or equivalently $t_{\mathrm{rel}}(H) > C > 1$. Thus, first we handle the other

case, i.e. when $t_{\mathrm{rel}}(H) < 2$. Then the lower bound we are about to show is of order $|G| \log |G|$ which is always at most the order of $t_{\mathrm{cov}}(G)$, due to a result by Feige [9] stating that for simple random walk on any connected graph G, $t_{\mathrm{cov}}(G) \geq (1 + o(1))|G| \log |G|$.

When $t_{\mathrm{rel}}(H) > 2$, we can use that $1 - x > e^{-1.5x}$ when $0 < x < 0.5$ to get a lower bound on the right hand side of (5.21). Then set $t = ct_{\mathrm{rel}}(H)|G| \log |G|$ turning the estimate in (5.20) into

$$\mathbf{E}_{(\underline{g}_0, x)}[\psi((\underline{F}_t, X_t))] \geq |G|^{1-3c} \phi_2(g_0).$$

We can easily upper bound the conditional variance as follows:

$$\mathbf{Var}\left[\psi_t | (L_v(t))_{v \in G}\right] \leq \sum_{w \in G} \mathbf{E}_{g_0}\left[\phi_2^2(F_w(t)) | L_w(t)\right] \leq |G| \phi_2^2(g_0).$$

Now, let us estimate the measure of set A at time t by using the lower bound on the expectation:

$$\mathbf{P}_{(\underline{g}_0, x)}\left[\psi_t \leq 2|G|^{1/2}\right] \leq \mathbf{P}_{(\underline{g}_0, x)}\left[|\psi_t - \mathbf{E}(\psi_t)| \geq \phi_2(g_0)|G|^{1-3c} - 2|G|^{1/2}\right]$$

Now we use that $\phi_2(g_0) > 1$ and if $c < 1/6$ then on the right hand side, the term $\phi_2(g_0)|G|^{1-3c}$ dominates, so for $|G|$ large enough we can drop the negative term and compensate it with a multiplicative factor of $1/2$, say. Thus, condition on the local time sequence first and see that for any sequence $(L_v(t))_{v \in G}$ Chebyshev's inequality yields:

$$\mathbf{P}_{(\underline{g}_0, x)}\left[\psi_t \in A \,\middle|\, (L_v(t))_{v \in G}\right] \leq \frac{\mathbf{Var}\left[\psi_t | (L_v(t))_{v \in G}\right]}{1/4 \phi_2^2(g_0)|G|^{2-6c}}$$

Combining this with the estimate on the conditional variance above yields that

$$\mathbf{P}_{(\underline{g}_0, x)}\left[\psi_t \in A \,\middle|\, (L_v(t))_v\right] \leq \frac{4}{|G|^{1-6c}}.$$

This bound is independent of the local time sequence, so the law of total probability says we have the same upper bound without conditioning on the local times. Now setting $c < 1/6$ an $|G|$ large enough we see that the right hand side can be made smaller than $1/2$, finishing the proof.

To see that the cover time is a lower bound in the general case, couple the chain on $H \wr G$ to $\mathbf{Z}_2 \wr G$, i.e. jump to stationary distribution on H_v once the walker on the base hits vertex v and use [18] or [13] to see that $t_{\mathrm{cov}}(G) \leq t_{\mathrm{mix}}(Z_2 \wr G) \leq t_{\mathrm{mix}}(H \wr G)$.

Next we show that $c|G|t_{\mathrm{mix}}(H)$ is a lower bound if the chain on H is lazy.

Let us start with a definition for general Markov chain X on Ω

$$t_{\text{stop}}(G) := \max_{x \in \Omega} \min \left\{ \mathbf{E}[\tau]; \tau \text{ stopping time s.t. } \mathbf{P}_x[X_\tau = y] = \pi(y) \ \forall y \in \Omega \right\}.$$

We call a stopping time mean-optimal if $\mathbf{E}[\tau] = t_{\text{stop}}(G)$. Lovász and Winkler [15] show that optimal stopping rules always exist for irreducible Markov chains. We aim to show that

$$\frac{1}{2}|G| \cdot t_{\text{stop}}(H) \leq t_{\text{stop}}(H \wr G).$$

Take a mean optimal stopping time τ^* on $H \wr G$ reaching minimal expectation, i.e. $\mathbf{E}_{(\underline{f}^*, x^*)}[\tau^*] = t_{\text{stop}}(H \wr G)$ for some $(\underline{f}^*, x^*) \in H \wr G$ and $\mathbf{E}_{(\underline{f}, x)}[\tau^*] \leq t_{\text{stop}}(H \wr G)$ for $(\underline{f}, x) \neq (\underline{f}^*, x^*)$.

We use this τ^* to define a stopping rule $\tau_H(v)$ on H_v, for every $v \in G$. Namely, do the following: look at a coordinate $v \in G$ and at the chain restricted to the lamp graph H_v, i.e. only the moves which are done on the coordinate H_v. Then, stop the chain on H_v when τ^* stops on the whole $H \wr G$.

Start the chain from any $(\underline{f}_0, x_0)$. Since $\sum_{v \in G} L_v(t) = 2t$, we have

$$\sum_{v \in G} \mathbf{E}_{f_v(0)}[\tau_H(v)] = \mathbf{E}_{(\underline{f}_0, x_0)} \left[\sum_{v \in G} L_v(\tau^*) \right] = 2\mathbf{E}_{(\underline{f}_0, x_0)}[\tau^*].$$

Take the vertex $w \in G$ (which can depend on x_0), which minimizes the expectation $\mathbf{E}_{f_v(0)}[\tau_H(w)]$. Clearly for this vertex the expected value must be less than the average:

$$\mathbf{E}_{f_0}[\tau_H] \leq \frac{2}{|G|} \mathbf{E}_{(\underline{f}_0, x_0)}[\tau^*]$$

The left hand side is at least as large as what a mean-optimal stopping rule on H can achieve, and the right hand side is at most $\frac{2}{|G|} t_{\text{stop}}(H \wr G)$. Thus we arrive at

$$\frac{1}{2}|G| t_{\text{stop}}(H) \leq t_{\text{stop}}(H \wr G).$$

In the last step we use the equivalence from the paper [19, Corollary 2.5] stating that t_{stop} and t_{mix} are equivalent up to universal constants for lazy reversible chains and get that

$$c_1 |G| t_{\text{mix}}(H) \leq t_{\text{mix}}(H \wr G).$$

6. Further directions. The next step of understanding generalized lamplighters walks might be to investigate which properties on G and H are needed to exhibit cutoff (for a definition see [2, 13]), or to determine the mixing time in the uniform metric.

For $\mathbf{Z}_2 \wr G$, already [11] implies a total variation cutoff with threshold $\frac{1}{2} t_{\mathrm{cov}}(K_n)$ for G being the complete graph and that there is no cutoff if G is a cycle of length n. The results of [18] include a proof of total variation cutoff for $\mathbf{Z}_2 \wr \mathbf{Z}_n^2$ with threshold $t_{\mathrm{cov}}(\mathbf{Z}_n^2)$. The results in [16] also includes cutoff at $1/2 t_{\mathrm{cov}}(G_n)$, with some uniform local transience assumptions on G_n. Further, Levi [14] proved that the wreath product of two complete graphs $K_{n^\lambda} \wr K_n$, $0 \le \lambda \le 1$ exhibits a cutoff at $(1+\lambda)/2n \log n$.

For the mixing time in the uniform metric, we know [18, Theorem 1.4] that if G is a regular graph such that $t_{\mathrm{hit}}(G) \le K|G|$, then there exists constants c, C depending only on K such that

$$(6.1) \qquad c|G|(t_{\mathrm{rel}}(G) + \log|G|) \le t_{\mathrm{u}}(\mathbf{Z}_2 \wr G) \le C|G|(t_{\mathrm{mix}}(G) + \log|G|).$$

These bounds fail to match in general. For example, for the hypercube $\mathbf{Z}_2^d$, $t_{\mathrm{rel}}(\mathbf{Z}_2^d) = \Theta(d)$ [13, Example 12.15] while $t_{\mathrm{mix}}(\mathbf{Z}_2^d) = \Theta(d \log d)$ [13, Theorem 18.3]. Then [12] showed that the lower bound is sharp in (6.1) under conditions which are satisfied by the $d(n)$ dimension tori $G_n = \mathbf{Z}_n^{d(n)}$ for arbitrary chosen n and $d(n)$.

References.

[1] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93:243–297, 1986.

[2] David Aldous and James Allen Fill. *Reversible Markov Chains and Random Walks on Graphs*. University of California, Berkeley, 2002.

[3] Martin T. Barlow, Jian Ding, Asaf Nachmias, and Yuval Peres. The evolution of the cover time. *Combin. Probab. Comput.*, 20(3):331–345, 2011.

[4] M.J.A. M. Brummelhuis and H. J. Hilhorst. Covering of a finite lattice by a random walk. *Physica A.*, 176:387–408, 1991.

[5] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. Math.*, 160(2):433–464, 2004.

[6] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Late points for random walk in two dimensions. *Ann. Probab.*, 34:219–263, 2006.

[7] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[8] Jian Ding, Yuval Peres, and James R. Lee. Cover times, blanket times, and majorizing measures. *Annals of Mathematics*, 175(3):1409–1471, 2012.

[9] Uriel Feige. A tight lower bound on the cover time for random walks on graphs. *Random Structures and Algorithms*, 6(4):433–438, 1995.

[10] James Fill and Jr. Clyde Schoolfield. Mixing times for markov chains on wreath products and related homogeneous spaces. *Electron. J. Probab.*, 6:no. 11, 1–22, 2001.

[11] O. Häggström and J. Jonasson. Rates of convergence of lamplighter processes. *Stochastic Processes and their Applications*, 67:227–249, 1997.

[12] J. Komjathy, J. P. Miller, and Y. Peres. Uniform mixing time for random walk on lamplighter graphs. `http://arxiv.org/pdf/1109.4281.pdf`.

[13] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.

[14] Nathan Levy. Mixing time for lamplighter graphs. Master's thesis, University of California, Berkeley, 2006.

[15] László Lovász and Peter Winkler. Efficient stopping rules for markov chains. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, STOC '95, pages 76–82, New York, NY, USA, 1995. ACM.

[16] Jason P. Miller and Yuval Peres. Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *Annals of Probability*, 2011.

[17] Igor Pak. *Random Walks on Groups: Strong Uniform Time Approach*. PhD thesis, Harvard University, 1997.

[18] Yuval Peres and David Revelle. Mixing times for random walks on finite lamplighter groups. *Electronic Journal of Probability*, 9:825–845, 2004.

[19] Yuval Peres and Perla Sousi. Mixing times are hitting times of large sets. `http://arxiv.org/abs/1108.0133`.

[20] Peter Winkler and David Zuckerman. Multiple cover time. *Random Structures Algorithms*, 9(4):403–411, 1996.

BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS, `komyju@math.bme.hu` MICROSOFT RESEARCH `peres@microsoft.com`