Electronic Journal of Probability

Electron. J. Probab. **0** (2012), no. 0, 1–19.
ISSN: 1083-6489 DOI: 10.1214/EJP.vVOL-PID

Random walks colliding before getting trapped

Louigi Addario-Berry* Roberto I. Oliveira† Yuval Peres‡
Perla Sousi§

Abstract

Let P be the transition matrix of a finite, irreducible and reversible Markov chain. We say the continuous time Markov chain X has transition matrix P and speed λ if it jumps at rate λ according to the matrix P. Fix $\lambda_X, \lambda_Y, \lambda_Z \geq 0$, then let X, Y and Z be independent Markov chains with transition matrix P and speeds λ_X, λ_Y and λ_Z respectively, all started from the stationary distribution. What is the chance that X and Y meet before either of them collides with Z? For each choice of λ_X, λ_Y and λ_Z with $\max(\lambda_X, \lambda_Y) > 0$, we prove a lower bound for this probability which is uniform over all transitive, irreducible and reversible chains. In the case that $\lambda_X = \lambda_Y = 1$ and $\lambda_Z = 0$ we prove a strengthening of our main theorem using a martingale argument. We provide an example showing the transitivity assumption cannot be removed for general λ_X, λ_Y and λ_Z.

Keywords: Transitive Markov chains; mixing time; hitting times; meeting times; martingale..
AMS MSC 2010: 60J10.
Submitted to EJP on July 15, 2015, final version accepted on May 1, 2016.

1 Introduction

Consider three independent random walks X, Y, Z over the same finite connected graph. What is the probability that X, Y meet at the same vertex before either of them meets Z? If the initial distributions of the three walkers are the same, this probability is at least $1/3$ by symmetry, at least if we assume that ties (i.e. triple meetings) are broken symmetrically.

Now consider a similar problem where the initial states X_0, Y_0, Z_0 are all sampled independently from the same distribution, but Z stays put while X and Y move. What is the probability that X and Y meet before hitting Z?

There are several examples of bounds [1, 4, 5] relating the meeting time of two random walks to the hitting time of a fixed vertex by a single random walk. These typically provide upper bounds for meeting times in terms of worst-case or average

*McGill University, Montréal, Canada. E-mail: louigi.addario@mcgill.ca
†IMPA, Rio de Janeiro, Brazil. Supported by a *Bolsa de Produtividade em Pesquisa* from CNPq. E-mail: rimfo@impa.br
‡Microsoft Research, Redmond, Washington, USA. E-mail: peres@microsoft.com
§University of Cambridge, Cambridge, UK. E-mail: p.sousi@statslab.cam.ac.uk

hitting times, sometimes up to constant factors. In light of this, it seems natural to conjecture that the probability in question is at least $1/3$. However, the previous argument by symmetry fails. In fact, to the best of our knowledge, no known universal lower bound for this probability is known.

It will be convenient to consider the problem in continuous time. For the remainder of the paper let P be the transition matrix of an irreducible and reversible Markov chain on a finite state space with stationary distribution π. Let X and Y be two independent continuous time Markov chains that jump at rate 1 according to the transition matrix P and let $Z \sim \pi$ be independent of X and Y.

We define $M^{X,Y}$ to be the first time X and Y meet, i.e.

$$M^{X,Y} = \inf\{t \geq 0 : X_t = Y_t\}.$$

We also define:

$$M^{W,Z} = \inf\{t \geq 0 : W_t = Z\} \ (W \in \{X, Y\}).$$

We write $M^{\text{good}} = M^{X,Y}$ and $M^{\text{bad}} = M^{X,Z} \wedge M^{Y,Z}$.

1.1 Main results

Our first result proves a universal lower bound on the probability $\mathbb{P}\big(M^{\text{good}} < M^{\text{bad}}\big)$ for the class of transitive chains. First we recall the definition.

Definition 1.1. Fix a chain with transition matrix P and state space Ω. An *automorphism* of P is a bijection $\varphi : \Omega \to \Omega$ such that $P(z, w) = P(\varphi(z), \varphi(w))$ for all $z, w \in \Omega$. The chain P is *transitive* if for all $x, y \in \Omega$ there exists an automorphism φ of P with $\varphi(x) = y$.

Theorem 1.2. *Let P be the transition matrix of a finite, irreducible and reversible chain with two or more states. Assume X_0 and Y_0 are independent with law π. If P is transitive, then*

$$\mathbb{P}\big(M^{\text{good}} < M^{\text{bad}}\big) \geq \frac{1}{4}.$$

Next we consider a more general setup. We say that a random walk W has speed λ_W and transition matrix P, if it jumps at rate λ_W according to the matrix P.

Suppose again that P is the transition matrix of an irreducible and reversible Markov chain on a finite state space with stationary distribution π. Let $\lambda_X = 1$, $0 \leq \lambda_Y \leq 1$ and $0 \leq \lambda_Z < \infty$. Let X, Y and Z be three independent continuous time Markov chains with speeds λ_X, λ_Y and λ_Z respectively and transition matrix P.

For the remainder of the paper, we write $\mathbb{P}$ for the probability measure under which X_0, Y_0 and Z_0 are independent with law π. We also write $\mathbb{P}_{a,b,c}$ in the case when $(X_0, Y_0, Z_0) = (a, b, c)$. For computations that only involve two chains we drop one index writing only $\mathbb{P}_{a,b}$; which two chains are involved will always be clear from context. Likewise, we write $\mathbb{P}_a$ when only one chain is involved. We define $M^{X,Y}$ as above and redefine:

$$M^{W,Z} = \inf\{t \geq 0 : W_t = Z_t\} \ (W \in \{X, Y\}).$$

Note that when $\lambda_Z = 0$ this definition agrees with the previous one. We define $M^{\text{good}} = M^{X,Y}$ and $M^{\text{bad}} = M^{X,Z} \wedge M^{Y,Z}$ as before. Again we are interested in uniform lower bounds on the probability of the event $\{M^{\text{good}} < M^{\text{bad}}\}$ that have good dependence on the three speeds.

Theorem 1.3. *There exists $c > 0$ such that the following holds. Let P be the transition matrix of a transitive, irreducible and reversible chain with stationary distribution π and at least two states. Suppose that X, Y and Z are three independent continuous time Markov chains with speeds $\lambda_X = 1, \lambda_Y \leq 1$ and $0 \leq \lambda_Z < \infty$ and transition matrix P*

started from π. Then

$$\mathbb{P}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) \geq \frac{c}{(\sqrt{1+\lambda_Z}+\sqrt{\lambda_Y+\lambda_Z})^2}\,.$$

The proof shows that we may take $c = 1/352$, which implies a version of Theorem 1.2 with $1/4$ replaced by $1/1408$. The constant c most likely can be improved, but the dependence of the lower bound on λ_Z is sharp when $\lambda_Z \nearrow +\infty$. Indeed, if P is simple random walk over a large complete graph with n vertices, then

$$\mathbb{P}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) = \frac{1+\lambda_Y}{2(1+\lambda_Y+\lambda_Z)} - O\left(\frac{1}{n}\right),$$

where the term $O(1/n)$ corresponds to the possibility of meetings at time 0.

A key step in the proof of Theorem 1.3 is the following new occupation identity. We will use it to estimate the time that $(X_t,Y_t)_{t\geq 0}$ spends on the diagonal of Ω^2 up to time M^{bad}. It applies to all reversible chains and we believe it is of independent interest.

Lemma 1.4. *Suppose $(U_t)_{t\geq 0}$, $(V_t)_{t\geq 0}$ are independent, irreducible, continuous time reversible Markov chains with finite state spaces Ω_U and Ω_V respectively. Assume μ is a probability measure over $\Omega_U\times\Omega_V$ and that τ is a stopping time for the process $(U_t,V_t)_{t\geq 0}$ with the following properties.*

(1) $\mathbb{P}_\mu(\tau > 0) = 1$;
(2) $\mathbb{E}_\mu[\tau] < \infty$ *and*
(3) $\mathbb{P}_\mu(V_0 = \cdot) = \mathbb{P}_\mu(V_\tau = \cdot)$.

Then for all $v\in\Omega_V$

$$\mathbb{E}_\mu\left[\int_0^\tau \mathbf{1}(V_t = v)\,\mathrm{d}t\right] = \mathbb{E}_\mu[\tau]\;\pi_V(v)$$

where π_V is the stationary distribution of V.

It is natural to ask if our theorems can be extended to all (i.e. not necessarily transitive) chains. The next theorem shows that the answer is no for the more general Theorem 1.3. The theorem essentially asserts that there are graphs where typical meeting times are much smaller than typical hitting times.

Theorem 1.5. *For all $\varepsilon > 0$ there exists a finite connected graph G such that if P corresponds to simple random walk on G and $\lambda_X = 1, \lambda_Y = 0$ and $\lambda_Z = 1$, then $\mathbb{P}\big(M^{\mathrm{good}} \leq M^{\mathrm{bad}}\big) < \varepsilon$.*

On the other hand, we believe that for certain values of λ_X,λ_Y and λ_Z, universal lower bounds are possible without transitivity. Here is a concrete conjecture, which relates to the setting of Theorem 1.2.

Conjecture 1.6. *If $\lambda_Y = \lambda_X = 1$ and $\lambda_Z = 0$, the inequality*

$$\mathbb{P}\big(M^{\mathrm{good}} \leq M^{\mathrm{bad}}\big) \geq 1/3$$

holds for all finite irreducible and reversible chains P.

Alexander Holroyd (personal communication) pointed out an example showing that for any $\delta > 0$ there exist transitive chains for which $\mathbb{P}\big(M^{\mathrm{good}} \leq M^{\mathrm{bad}}\big) \leq 1/3+\delta$. We describe this example in Section 6. This means that, if true, Conjecture 1.6 is best possible even for transitive chains. However, we note that any uniform lower bound

$$\mathbb{P}\big(M^{\mathrm{good}} \leq M^{\mathrm{bad}}\big) \geq c > 0$$

for all P, and for λ_X,λ_Y and λ_Z as in Conjecture 1.6, would be a new result.

Remark 1.7. Without reversibility, the conjecture fails badly. Consider a clockwise continuous time random walk on a cycle of length $2n$. More precisely, with $P = (p_{ij})_{1\le i,j\le n}$ we have $p_{ij} = 1$ if $j = (i+1) \bmod n$ and $p_{ij} = 0$ otherwise. The distance between independent random walkers behaves as continuous time simple symmetric random walk reflected at 0 and n. So started from stationarity, it typically takes such walkers time of order n^2 to meet. On the other hand, the hitting time of any point is at most of order n.

Before we continue, we say a few words about the main proof ideas and the structure of the rest of the document. The unifying theme of the proofs of Theorems 1.2 and 1.3 is the relationship between meeting times and hitting times of single vertices when P is transitive. Aldous and Fill [1, Chapter 14/Proposition 5 and Chapter 3] have related the expected values of these random variables via martingales. We use similar ideas to prove Theorem 1.2 in Section 2.

For the proof of Theorem 1.3, we need a stronger result establishing identities in distribution of meeting and hitting times, which (somewhat surprisingly) seems to be new: see Lemma 3.1 below. The proof of Theorem 1.3 requires several other tools, including small time estimates for hitting times given in Section 3.2, as well as the occupation time formula for product chains, Lemma 1.4. The start of Section 3 contains a succinct but rigorous birds-eye view of our approach, proving Theorem 1.3 *modulo* three lemmas, which are then proved in the remainder of the section, and the occupation identity Lemma 1.4, whose proof occupies Section 4.

The proof of Theorem 1.5 builds a graph with two parts: the “Up” part concentrates the bulk of the stationary measure, but the “Down” part is where meetings tend to happen, and they happen quickly. As a result, only a negligible fraction of the “Up” part is explored before X and Z meet, and the upshot is that $M^{X,Y} > M^{X,Z}$ with high probability. We prove Theorem 1.5 in Section 5.

2 The 1/4 lower bound

In this section we prove Theorem 1.2. The argument is fairly short, and much simpler than the one for the more general Theorem 1.3.

We need some preliminaries on hitting times. The hitting time of a state $z \in \Omega$ by X is the first time t at which $X_t = z$, i.e.

$$\tau_z^X := \inf\{t \ge 0 \,:\, X_t = z\} \tag{2.1}$$

We define τ_z^Y, τ_z^Z similarly and we also let

$$t^*_{\mathrm{hit}} := \max_{x\in\Omega} \mathbb{E}_\pi\left[\tau_x^X\right] \quad \text{and} \quad t_{\mathrm{hit}} := \max_{(x,z)\in\Omega} \mathbb{E}_x\left[\tau_z^X\right]. \tag{2.2}$$

Whenever there is no confusion, i.e. if there is a single chain in question, we will drop the dependence on X or Y from the notation of the hitting times.

Lemma 2.1. *For any irreducible and reversible chain with two or more states we have*

$$0 < t_{\mathrm{hit}} \le 2t^*_{\mathrm{hit}}.$$

Moreover, if X is transitive, then for all $x, z \in \Omega$ and all $t \ge 0$ we have

$$\mathbb{P}_x(\tau_z \le t) = \mathbb{P}_z(\tau_x \le t)\,.$$

Proof. For a proof of the first assertion see [3, Lemma 10.2]. (Note that mean hitting times are the same in discrete and continuous time.) A proof of the symmetry property specific to transitive chains can be found in [1, Lemma 1, Chapter 7]. □

We will also need the following lemma [1, Chapter 14/Proposition 5 and Chapter 3].

Lemma 2.2 (Aldous). *Let P be an irreducible and reversible transition matrix. Suppose that X and Y are independent continuous time Markov chains that jump at rate 1 according to the transition matrix P. For all $x,z \in \Omega$ define $f(x,z) := \mathbb{E}_x\left[\tau_z^X\right]$. Then $f(X_t,z)+t$, $f(Y_t,z)+t$ and $f(X_t,Y_t)+2t$ are martingales up to time*

$$S := M^{\mathrm{good}} \wedge M^{\mathrm{bad}} = M^{X,Y} \wedge \tau_z^X \wedge \tau_z^Y,$$

for any initial states $(x,y)\in\Omega^2$ and any $z\in\Omega$.

Proof of Theorem 1.2. Since X and Y are two independent copies of the same chain, we have $\mathbb{E}_a\left[\tau_b^X\right] = \mathbb{E}_a\left[\tau_b^Y\right]$ for all a,b. By Lemma 2.2 we now get that $(G_t)_{t\ge 0}$ is a martingale up to time S, where

$$G_t := \mathbb{E}_{X_t}\left[\tau_z^X\right] + \mathbb{E}_{Y_t}\left[\tau_z^X\right] - \mathbb{E}_{X_t}\left[\tau_{Y_t}^X\right] \quad (t \ge 0).$$

This martingale is bounded (because the state space is finite). The fact that the chain is finite and irreducible implies $S<\infty$ almost surely for all initial states. We deduce from optional stopping that

$$\mathbb{E}[G_0] = \mathbb{E}[G_S]\,. \tag{2.3}$$

The left hand side above is given by the quantity t^*_{hit} defined in (2.2). This is because

$$\mathbb{E}[G_0] = \mathbb{E}_\pi\left[\tau_z^X\right] + \mathbb{E}_\pi\left[\tau_z^X\right] - \sum_{y\in\Omega}\pi(y)\mathbb{E}_\pi\left[\tau_y^X\right] = \mathbb{E}_\pi\left[\tau_z^X\right], \tag{2.4}$$

where the second equality follows from the fact that for a transitive chain, $\mathbb{E}_\pi\left[\tau_y^X\right]$ is independent of y. Using this and (2.3) yields

$$t^*_{\mathrm{hit}} = \mathbb{E}[G_S]\,. \tag{2.5}$$

On the other hand, at time S we have two alternatives.

- If $\tau_z^X \wedge \tau_z^Y \le M^{\mathrm{good}}$, either $X_S = z$, and then $G_S = \mathbb{E}_{Y_S}\left[\tau_z^X\right] - \mathbb{E}_z\left[\tau_{Y_s}^X\right]$, or $Y_S = z$, in which case $G_S = \mathbb{E}_{X_S}\left[\tau_z^X\right] - \mathbb{E}_{X_S}\left[\tau_z^X\right]$. In both cases $G_S = 0$: this is obviously true in the second case, and follows from Lemma 2.1 in the first case.
- On the other hand, if $M^{\mathrm{good}} < \tau_z^X \wedge \tau_z^Y$, then $G_S = 2\mathbb{E}_{X_S}\left[\tau_z^X\right] \le 2t_{\mathrm{hit}}$.

We deduce that

$$G_S \le 2t_{\mathrm{hit}}\,\mathbf{1}(M^{\mathrm{good}} < \tau_z^X \wedge \tau_z^Y).$$

Plugging this into (2.5) gives

$$t^*_{\mathrm{hit}} \le 2\,t_{\mathrm{hit}}\,\mathbb{P}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big)\,.$$

Using that $t_{\mathrm{hit}} \le 2t^*_{\mathrm{hit}}$ from Lemma 2.1 finishes the proof. □

Before moving on, we first argue that the obvious “fix" to the proof of Theorem 1.2 does not work in general when X,Y and Z have differing speeds. Indeed, a straightforward extension of Lemma 2.2 establishes that

$$G_t := f(X_t,Z_t) + f(Y_t,Z_t) - \frac{1+\lambda_Y+2\lambda_Z}{1+\lambda_Y} f(X_t,Y_t)$$

is a martingale up to time S. One can see that in this case

$$\mathbb{E}[G_0] = \left(1 - \frac{2\lambda_Z}{1+\lambda_Y}\right) t^*_{\mathrm{hit}},$$

which easily yields

$$\mathbb{P}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) \geq \frac{1}{4} \cdot \left(1 - \frac{2\lambda_Z}{1+\lambda_Y}\right).$$

In particular, we obtain the same bound as in Theorem 1.2 provided that $\lambda_Z = 0$. However, this bound becomes useless when $\lambda_Z > (1+\lambda_Y)/2$. Other linear combinations of $f(X_t, Z_t)$, $f(Y_t, Z_t)$ and $f(X_t, Y_t)$ also fail to achieve our goal when λ_Z is large. So, in general a different strategy is needed.

3 Towards the theorem for general speeds

In this section we prove Theorem 1.3, assuming a handful of results whose proofs we briefly postpone. Our proof of Theorem 1.3 is based on the analysis of the time that (X, Y) spends on the diagonal $\Delta = \{(x, x) : x \in \Omega\}$ prior to time M^{bad}, i.e.

$$T := \int_0^{M^{\mathrm{bad}}} \mathbf{1}(X_t = Y_t)\,\mathrm{d}t = \int_0^{\infty} \mathbf{1}(X_t = Y_t, t < M^{\mathrm{bad}})\,\mathrm{d}t. \tag{3.1}$$

In order to gain some intuition for this quantity, note that $T > 0$ if and only if $M^{\mathrm{good}} < M^{\mathrm{bad}}$, i.e. (X, Y) visits the diagonal before X or Y meet Z. The proof relies on obtaining lower and upper bounds for $\mathbb{E}[T]$.

We first derive an expression for $\mathbb{E}[T]$ that requires only reversibility. For any $t > 0$, using reversibility and the definitions of M^{good} and M^{bad}, we have

$$\begin{aligned}
\mathbb{P}\big(X_t = Y_t,\, t < M^{\mathrm{bad}}\big) &= \mathbb{P}\left(\{X_t = Y_t\} \cap \{\forall s \leq t\, :\, X_s \neq Z_s \text{ and } Y_s \neq Z_s\}\right)\\
&= \mathbb{P}\left(\{X_0 = Y_0\} \cap \{\forall s \leq t\, :\, X_s \neq Z_s \text{ and } Y_s \neq Z_s\}\right)\\
&= \mathbb{P}\left(M^{\mathrm{good}} = 0, t < M^{\mathrm{bad}}\right).
\end{aligned}$$

By Fubini's theorem, it follows that

$$\begin{aligned}
\mathbb{E}[T] &= \int_0^{\infty} \mathbb{P}\left(X_t = Y_t,\, t < M^{\mathrm{bad}}\right)\,\mathrm{d}t\\
&= \int_0^{\infty} \mathbb{P}\left(M^{\mathrm{good}} = 0, t < M^{\mathrm{bad}}\right)\,\mathrm{d}t\\
&= \mathbb{E}\left[M^{\mathrm{bad}}\mathbf{1}(M^{\mathrm{good}} = 0)\right]\\
&= \sum_{x,z} \pi(x)^2 \pi(z) \mathbb{E}_{(x,x,z)}\left[M^{\mathrm{bad}}\right].
\end{aligned}$$

The lower bound on $\mathbb{E}[T]$ now exploits the following distributional identity, which (surprisingly) appears to be new; its proof appears in Section 3.1. Recall from Section 2 that $\tau_z^X := \inf\{t \geq 0 : X_t = z\}$ is the hitting time of $z \in \Omega$ by X.

Lemma 3.1. *Let P be a reversible and transitive transition matrix. Let X, Y and Z be three independent continuous time Markov chains with speeds $\lambda_X = 1, \lambda_Y \geq 0$ and $\lambda_Z \geq 0$ and transition matrix P. Assume also $\lambda_Y + \lambda_Z > 0$. Then for all $(x, z) \in \Omega^2$, the distribution of $\frac{\tau_z^X}{\lambda_Y + \lambda_Z}$ under $\mathbb{P}_x$ is the same as the distribution of $M^{Y,Z}$ under $\mathbb{P}_{(x,z)}$.*

Since $M^{\mathrm{bad}} = M^{X,Y} \wedge M^{X,Z}$, it follows that for a transitive chain with n states, for any fixed $t > 0$,

$$\begin{aligned}
\mathbb{E}[T] &\geq t \sum_{x,z} \pi(x)^2 \pi(z) \mathbb{P}_{(x,x,z)}\left(M^{X,Z} \wedge M^{Y,Z} > t\right)\\
&\geq t \sum_{x,z} \pi(x)^2 \pi(z) (1 - \mathbb{P}_{(x,z)}\left(M^{X,Z} \leq t\right) - \mathbb{P}_{(x,z)}\left(M^{Y,Z} \leq t\right))_+
\end{aligned}$$

$$
\begin{aligned}
&= t \sum_{x,z} \pi(x)^2 \pi(z) (1 - \mathbb{P}_x \left(\tau_z^X \leq t(1+\lambda_Z) \right) - \mathbb{P}_x \left(\tau_z^X \leq t(\lambda_Y + \lambda_Z) \right))_+ \\
&= \frac{t}{n} \sum_{x,z} \pi(x) \pi(z) (1 - \mathbb{P}_x \left(\tau_z^X \leq t(1+\lambda_Z) \right) - \mathbb{P}_x \left(\tau_z^X \leq t(\lambda_Y + \lambda_Z) \right))_+ . \qquad (3.2)
\end{aligned}
$$

Transitivity is used in the last step, to ensure that $\pi(x) = 1/n$ for all x.

We can not expect useful bounds on the lower tail of τ_z^X when the starting point x and the point z are arbitrary (think of adjacent vertices on a graph). The next lemma shows that, for transitive chains, we may nevertheless find a large set of states z for which $\tau_z^X \geq \theta\, t_{\text{hit}}^*$ with high probability when θ is small. We hereafter drop the superscript X from τ_z^X to simplify notation.

Lemma 3.2. *Suppose that P is irreducible, reversible and transitive. Then for any $x \in \Omega$, there exists a subset $A_x \subset \Omega$ with $\pi(A_x) \geq 1/2$ such that, for any $\theta > 0$,*

$$
\frac{1}{\pi(A_x)} \sum_{z \in A_x} \pi(z) \mathbb{P}_x \left(\tau_z \leq \theta\, t_{\text{hit}}^* \right) \leq \sqrt{\theta}
$$

The proof of Lemma 3.2 appears in Section 3.2. We conclude our lower bound on $\mathbb{E}[T]$ by applying the lemma with

$$
\theta := \frac{1}{4\, (\sqrt{1+\lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2} ,
$$

and taking $t = \theta t_{\text{hit}}^*$ in (3.2). We obtain:

$$
\begin{aligned}
\mathbb{E}[T] &\geq \frac{\theta t_{\text{hit}}^*}{n} \sum_x \pi(x) \sum_{z \in A_x} \pi(z) (1 - \mathbb{P}_x \left(\tau_z \leq t_{\text{hit}}^* \theta (1+\lambda_Z) \right) - \mathbb{P}_x \left(\tau_z \leq t_{\text{hit}}^* \theta (\lambda_Y + \lambda_Z) \right))_+ \\
&\geq \frac{\theta t_{\text{hit}}^*}{n} \sum_x \pi(x) \pi(A_x) (1 - \sqrt{\theta(1+\lambda_Z)} - \sqrt{\theta(\lambda_Y + \lambda_Z)}) \\
&\geq \frac{t_{\text{hit}}^*}{16\, n\, (\sqrt{1+\lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2} . \qquad (3.3)
\end{aligned}
$$

The required upper bound for $\mathbb{E}\left[T\right]$ is given by the next lemma.

Lemma 3.3. *Let P be the transition matrix of a transitive, irreducible and reversible chain on $n \geq 2$ states with stationary distribution π. Suppose that X, Y and Z are three independent continuous time Markov chains with speeds $\lambda_X = 1$, $\lambda_Y \leq 1$ and $0 \leq \lambda_Z < \infty$ and transition matrix P started from π. Then*

$$
\mathbb{E}[T] \leq \frac{22\, t_{\text{hit}}^*}{n} \mathbb{P} \left(M^{\text{good}} < M^{\text{bad}} \right) .
$$

We prove Lemma 3.3, in Section 3.3, by applying the occupation identity Lemma 1.4 to a carefully chosen stopping time. The role of transitivity in this step is to control the the law of the stopping state. With this lemma in hand, the proof of Theorem 1.3 is easily completed.

Proof of Theorem 1.3. Combining Lemma 3.3 with (3.3) gives

$$
\frac{t_{\text{hit}}^*}{16\, n\, (\sqrt{1+\lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2} \leq \mathbb{E}\left[T\right] \leq \frac{22\, t_{\text{hit}}^*}{n} \mathbb{P} \left(M^{\text{good}} < M^{\text{bad}} \right) ,
$$

from which the desired lower bound on $\mathbb{P} \left(M^{\text{good}} < M^{\text{bad}} \right)$ is immediate. □

The remainder of Section 3 is devoted to the proofs of Lemmas 3.1, 3.2 and 3.3.

3.1 Proof of Lemma 3.1

Define the functions

$$g_{(x,z)}(t) := \mathbb{P}_x\left(\tau_z \le (\lambda_Y + \lambda_Z)\, t\right) \text{ and } f_{(x,z)}(t) := \mathbb{P}_{(x,z)}\left(M^{Y,Z} \le t\right)\ \ ((x,z) \in \Omega^2,\, t \ge 0).$$

We will be done once we show that $g_{(x,z)}(t) = f_{(x,z)}(t)$ for all $(x,z) \in \Omega^2$ and $t \ge 0$. These equalities are true (by inspection) when $t = 0$. We are going to show that the functions $(f_{(x,z)}(\cdot))_{(x,z)\in\Omega^2}$ and $(g_{(x,z)}(\cdot))_{(x,z)\in\Omega^2}$ satisfy the same linear system of ordinary differential equations (with the derivatives at $t = 0$ interpreted as right derivatives). Then the equality for all $t \ge 0$ will follow from the general uniqueness theory of linear ODE's.

To prove that f and g satisfy the same system of ODE's, we will use a standard formula for the cummulative distribution function of a hitting time. If $(V_t)_{t\ge0}$ is an irreducible continuous time Markov chain over a set Ω_V with transition rates $q(v,w)$, and $A \subset \Omega_V$ is a nonempty subset of the state space, the hitting time τ_A^V of A by V satisfies

$$\frac{d}{dt}\mathbb{P}_v\left(\tau_A^V \le t\right) = \begin{cases} 0, & v \in A; \\ \sum_{w\in V} q(v,w)\left(\mathbb{P}_w\left(\tau_A^V \le t\right) - \mathbb{P}_v\left(\tau_A^V \le t\right)\right), & v \in \Omega_V\backslash A. \end{cases} \tag{3.4}$$

(The derivative is understood as a right derivative at time $t = 0$.)

We first apply (3.4) to the product chain $(V_t)_{t\ge0} = (Y_t, Z_t)_{t\ge0}$, with $\Omega_V = \Omega^2$, and $A = \Delta := \{(x,x)\,:\, x \in \Omega\}$ the diagonal set. In this case $\tau_A^V = M^{Y,Z}$, and a straightforward computation with the transition rates gives:

$$\frac{d}{dt}\, f_{(x,z)}(t) = \begin{cases} 0, & x = z \\ \lambda_Y \sum_{x'\in\Omega} P(x,x')\left(f_{(x',z)}(t) - f_{(x,z)}(t)\right) & \\ +\lambda_Z \sum_{z'\in\Omega} P(z,z')\left(f_{(x,z')}(t) - f_{(x,z)}(t)\right), & x \ne z. \end{cases} \tag{3.5}$$

We now apply the same formula (3.4) with $(V_t)_{t\ge0} = (X_t)_{t\ge0}$. Note that $g_{(x,z)}(t) := \mathbb{P}_x\left(\tau_z \le s(t)\right)$ where $s(t) = (\lambda_Y + \lambda_Z)\, t$, so the chain rule implies

$$\frac{d}{dt}\, g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ (\lambda_Y + \lambda_Z)\, \sum_{x'\in\Omega} P(x,x')\left(g_{(x',z)}(t) - g_{(x,z)}(t)\right), & x \ne z. \end{cases} \tag{3.6}$$

We will now make crucial use of transitivity, which allows us to use Lemma 2.1 to deduce that $\mathbb{P}_x\left(\tau_z^X \le (\lambda_Y + \lambda_Z)\, t\right)$ is symmetric in x and z, i.e.

$$\mathbb{P}_x\left(\tau_z^X \le (\lambda_Y + \lambda_Z)\, t\right) = \mathbb{P}_z\left(\tau_x^X \le (\lambda_Y + \lambda_Z)\, t\right),$$

that is $g_{(x,z)}(\cdot) = g_{(z,x)}(\cdot)$ for all x, z. This allows us to reverse the roles of x and z in (3.6) to obtain:

$$\frac{d}{dt}\, g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ (\lambda_Y + \lambda_Z)\, \sum_{z'\in\Omega} P(z,z')\left(g_{(x,z')}(t) - g_{(x,z)}(t)\right), & x \ne z. \end{cases} \tag{3.7}$$

We add the two formulas (3.6) and (3.7) with weights $\lambda_Y/(\lambda_Y + \lambda_Z)$ and $\lambda_Z/(\lambda_Y + \lambda_Z)$ respectively. The upshot is:

$$\frac{d}{dt}\, g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ \lambda_Y \sum_{x'\in\Omega} P(x,x')\left(g_{(x',z)}(t) - g_{(x,z)}(t)\right) & \\ +\lambda_Z \sum_{z'\in\Omega} P(z,z')\left(g_{(x,z')}(t) - g_{(x,z)}(t)\right), & x \ne z. \end{cases}$$

This is precisely the system of ODEs we obtained for the f's in (3.5), and it concludes the proof. □

3.2 Small time estimates for hitting times

In this section we prove Lemma 3.2. Throughout the section we drop the superscript X from τ_z^X.

Recall from, e.g., [1, Section 3.4] that a reversible transition matrix P is always diagonalizable with real eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \lambda_3, \ldots$. When P is also irreducible, $\lambda_2 < 1$, and we may define the relaxation time via $t_{\rm rel} = (1-\lambda_*)^{-1}$. We will use the next lemma to prove Lemma 3.2 in the case that $t_{\rm rel}$ and $t_{\rm hit}$ have similar magnitude.

Lemma 3.4. *Let X be an irreducible and reversible chain. There exist $x \in \Omega$ and a subset $A \subset \Omega$ with stationary measure $\pi(A) \geq 1/2$ such that, if $\tau_A := \min_{z\in A} \tau_z$, then for any $t > 0$,*

$$\mathbb{P}_x\left(\tau_A > t\right) \geq e^{-t/t_{\rm rel}}.$$

Proof. The first step is to note that P has a non-zero eigenfunction $\varphi : \Omega \to \mathbb{R}$ such that

$$P\varphi = \lambda_2\,\varphi = \left(1 - \frac{1}{t_{\rm rel}}\right)\varphi.$$

By the general theory of reversible chains [1, Section 3.4], this eigenfunction is orthogonal to the constant eigenfunction in the inner product induced by π. In particular, it must take both positive and negative values. We may assume without loss of generality that the set

$$A := \{z \in \Omega \,:\, \varphi(z) \leq 0\}$$

has measure $\pi(A) \geq 1/2$ (if that is not the case, replace φ with $-\varphi$). Choose $x \in \Omega$ with $\varphi(x) > 0$ as large as possible. Next, note that

$$\forall t > 0,\, u \in \Omega \,:\, \mathbb{E}_u\left[\varphi(X_t)\right] = [e^{t(P-I)}\,\varphi](u) = e^{-t/t_{\rm rel}}\,\varphi(u). \tag{3.8}$$

In particular, for all $u \in A$ and $s \geq 0$ we have $\mathbb{E}_u[\varphi(X_s)] \leq 0$. Since $X_{\tau_A} \in A$, the strong Markov property then gives

$$\begin{aligned}\mathbb{E}_x\left[\varphi(X_t)\,\mathbf{1}(\tau_A \leq t)\right] &= \mathbb{E}_x\left[\mathbb{E}_{X_{\tau_A}}\left[\varphi(X_{t-\tau_A})\right]\mathbf{1}(\tau_A \leq t)\right] \\ &\leq \mathbb{E}_x\left[0 \cdot \mathbf{1}(\tau_A \leq t)\right] = 0.\end{aligned}$$

Plugging this into (3.8) with the choice $u = x$, and recalling $\varphi(X_t) \leq \varphi(x)$ always, we obtain

$$e^{-t/t_{\rm rel}}\varphi(x) = \mathbb{E}_x\left[\varphi(X_t)\right] \leq \mathbb{E}_x\left[\varphi(X_t)\,\mathbf{1}(\tau_A > t)\right] \leq \varphi(x)\,\mathbb{P}_x\left(\tau_A > t\right).$$

Dividing both sides by $\varphi(x)$ (which is > 0) finishes the proof. □

Our next result is a different small-time estimate, which will be useful when $t_{\rm rel} \ll t_{\rm hit}$.

Lemma 3.5. *Let P be an irreducible and reversible chain. Then, for all $s > 0$,*

$$\mathbb{P}_\pi\left(\tau_z > s\right) \geq 1 - \frac{s + t_{\rm rel}}{\mathbb{E}_\pi\left[\tau_z\right] + t_{\rm rel}}$$

To prove this lemma we will use the claim below which follows from estimates in Aldous and Brown [2].

Claim 3.6 ([2]). *Define*

$$f(s) := \mathbb{E}_\pi[\tau_z - s \mid \tau_z > s]\,,\, s \geq 0.$$

Then f is an increasing function and $\sup_s f(s) \leq \mathbb{E}_\pi[\tau_z] + t_{\rm rel}$.

Proof of Lemma 3.5. Note that this claim implies the lemma once we note that

$$\tau_z \le s + (\tau_z - s)\,\mathbf{1}(\tau_z > s) \Rightarrow \mathbb{E}_\pi\left[\tau_z\right] \le s + f(s)\,\mathbb{P}_\pi\left(\tau_z > s\right),$$

use $f(s) \le \mathbb{E}_\pi[\tau_z] + t_{\mathrm{rel}}$ and then rearrange terms. □

Proof of Claim 3.6. The proof is based on some estimates in Aldous and Brown [2] which we recall and reprove here for convenience.

To prove the claim, let $Q = I - P$ and Q_z be the restriction of Q to $\Omega \setminus \{z\}$. We recall the complete positivity of the law of τ_z starting from π (see for instance [2, eqn. (18)]): there exist non-negative constants $(p_i, 1 \le i \le m)$ such that for all t

$$\mathbb{P}_\pi(\tau_z > t) = \sum_{i=1}^m p_i e^{-\gamma_i t},$$

where $0 < \gamma_1 < \ldots < \gamma_m$ are the distinct eigenvalues of $-Q_z$. Note that $\gamma_1^{-1} = \mathbb{E}_\alpha[\tau_z]$, where α is any quasistationary distribution on $\Omega \setminus \{z\}$ corresponding to the eigenvalue γ_1.

Using the above representation we can rewrite f as follows

$$f(s) = \int_s^\infty \frac{\mathbb{P}_\pi(\tau_z > t)}{\mathbb{P}_\pi(\tau_z > s)}\,\mathrm{d}t = \frac{\sum_{i=1}^m p_i e^{-\gamma_i s}/\gamma_i}{\sum_{i=1}^m p_i e^{-\gamma_i s}}.$$

A straightforward differentiation now gives that f is increasing. From the above expression we also deduce

$$\lim_{s\to\infty} f(s) = \frac{1}{\gamma_1} = \mathbb{E}_\alpha[\tau_z]\,. \tag{3.9}$$

From [2, Corollary 4] we have

$$\mathbb{E}_\alpha[\tau_z] \le \mathbb{E}_\pi[\tau_z] + t_{\mathrm{rel}}.$$

Therefore, using this, the fact that f is increasing and (3.9) we conclude that for all s

$$f(s) \le \mathbb{E}_\alpha[\tau_z] \le \mathbb{E}_\pi[\tau_z] + t_{\mathrm{rel}}$$

which completes the proof of the claim. □

Proof of Lemma 3.2. Fix $\theta > 0$. We will consider two cases separately: $t_{\mathrm{rel}} < \sqrt{\theta}\, t^*_{\mathrm{hit}}$ and $t_{\mathrm{rel}} \ge \sqrt{\theta}\, t^*_{\mathrm{hit}}$.

Suppose first that $t_{\mathrm{rel}} \ge \sqrt{\theta}\, t^*_{\mathrm{hit}}$. By Lemma 3.4 there exist $x \in \Omega$ and a set $A = A_x$ with $\pi(A) \ge 1/2$ such that

$$\mathbb{P}_x\left(\tau_A > t\right) \ge e^{-t/t_{\mathrm{rel}}}.$$

Since the chain is transitive, this in fact holds for all x, with corresponding sets A_x. Since $\mathbb{P}_x\left(\tau_z > t\right) \ge \mathbb{P}_x\left(\tau_A > t\right)$ for all $z \in A$, we obtain

$$\frac{1}{\pi(A_x)} \sum_{z\in A_x} \pi(z)\,\mathbb{P}_x\left(\tau_z > \theta\, t^*_{\mathrm{hit}}\right) \ge e^{-\theta\, t^*_{\mathrm{hit}}/t_{\mathrm{rel}}} \ge 1 - \frac{\theta\, t^*_{\mathrm{hit}}}{t_{\mathrm{rel}}} \ge 1 - \sqrt{\theta}$$

which concludes the proof in this case.

Suppose next that $t_{\mathrm{rel}} < \sqrt{\theta}\, t^*_{\mathrm{hit}}$. In this case it suffices to prove

$$\forall z \in \Omega\, :\, \mathbb{P}_\pi\left(\tau_z > \theta\, t^*_{\mathrm{hit}}\right) \ge 1 - \sqrt{\theta}. \tag{3.10}$$

To see that this suffices, we use the fact that P is transitive and apply Lemma 2.1 to obtain that $\mathbb{P}_x(\tau_z > t)$ is symmetric in x and z. As a result, (3.10) implies

$$1 - \sqrt{\theta} \le \mathbb{P}_\pi\left(\tau_x > \theta\, t^*_{\rm hit}\right) = \sum_{z\in\Omega} \pi(z)\mathbb{P}_z\left(\tau_x > \theta\, t^*_{\rm hit}\right) = \sum_{z\in\Omega} \pi(z)\mathbb{P}_x\left(\tau_z > \theta\, t^*_{\rm hit}\right),$$

and this implies the lemma with the choice of $A_x = \Omega$.

It remains to prove (3.10). Since P is transitive, $\mathbb{E}_\pi[\tau_z] = t^*_{\rm hit}$ is independent of z. Moreover, we are assuming that $t_{\rm rel} \le \sqrt{\theta}\, t^*_{\rm hit}$, so

$$\mathbb{E}_\pi[\tau_z] + t_{\rm rel} \le (1+\sqrt{\theta})\, t^*_{\rm hit}.$$

Lemma 3.5 gives

$$\mathbb{P}\left(\tau_z > \theta\, t^*_{\rm hit}\right) \ge 1 - \frac{\theta + \sqrt{\theta}}{1+\sqrt{\theta}} = 1 - \sqrt{\theta}.$$

This finishes the proof of (3.10) and of the lemma. □

3.3 The upper bound on $\mathbb{E}[T]$

In this section we prove Lemma 3.3. We start by recalling a result relating meeting times with deterministic trajectories to the quantity $t^*_{\rm hit}$. Notice that transitivity is *not* required.

Lemma 3.7. *Let X be an irreducible and reversible Markov chain taking values in Ω and $\mathbf{h} = (h_t)_{t\ge 0}$ a deterministic, càdlàg, Ω-valued trajectory. If*

$$\tau_{\mathbf{h}} := \inf\{t \ge 0 \,:\, X_t = h_t\},$$

then for any $x \in \Omega$,

$$\mathbb{E}_x[\tau_{\mathbf{h}}] \le 11\, t^*_{\rm hit}.$$

Proof. In [5], using [4, Lemma 1.7], it is proved that

$$\mathbb{E}_x[\tau_{\mathbf{h}}] \le c\, t_{\rm hit}.$$

for a universal constant $c > 0$, where $t_{\rm hit}$ is as in (2.2). Inspection of the proof [5] shows that $c \le 4 + 5/4$, therefore $2c \le 11$. Lemma 2.1 finishes the proof. □

Proof of Lemma 3.3. The first point is to note that, for any t, $X_t = Y_t$ implies $t \ge M^{\rm good}$. In particular, $M^{\rm good} \ge M^{\rm bad}$ implies that $T = 0$, and any $t \le M^{\rm good}$ does not contribute to the integral defining T (cf. (3.1)). We deduce:

$$T = T \circ \Theta_{M^{\rm good}}\, \mathbf{1}(M^{\rm good} < M^{\rm bad}), \tag{3.11}$$

where Θ denotes the time shift operator.

Now consider the distribution μ of $(X_{M^{\rm good}}, Y_{M^{\rm good}}, Z_{M^{\rm good}})$ conditionally on $M^{\rm good} < M^{\rm bad}$. Note that $\{M^{\rm good} < M^{\rm bad}\}$ is measurable with respect to the σ-field $\mathcal{F}_{M^{\rm good}}$ generated by the process up to time $M^{\rm good}$. Equation (3.11) and the strong Markov property imply:

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}\left[\mathbf{1}(M^{\rm good} < M^{\rm bad})\, \mathbb{E}\left[T \circ \Theta_{M^{\rm good}} \mid \mathcal{F}_{M^{\rm good}}\right]\right] \\
&= \mathbb{E}\left[\mathbf{1}(M^{\rm good} < M^{\rm bad})\, \mathbb{E}_{(X_{M^{\rm good}}, Y_{M^{\rm good}}, Z_{M^{\rm good}})}[T]\right] \qquad (3.12)\\
&= \mathbb{P}\left(M^{\rm good} < M^{\rm bad}\right)\, \mathbb{E}_\mu[T] \qquad (3.13)
\end{aligned}$$

$$= \mathbb{P}\left(M^{\mathrm{good}} < M^{\mathrm{bad}}\right) \sum_{x\in\Omega} \mathbb{E}_\mu\left[\int_0^{M^{\mathrm{bad}}} \mathbf{1}((X_t,Y_t)=(x,x))\,\mathrm{d}t\right]. \tag{3.14}$$

We will now use the *occupation time identity* of Lemma 1.4 to compute the RHS of the display. More specifically, we will apply this lemma to (3.14), with $V_t := (X_t, Y_t)$, $U_t = Z_t$, μ as above and a time τ to be specified. The following claim gives us the distribution of $V_0 = (X_0, Y_0)$ under μ.

Claim 3.8. *The measure μ is invariant under automorphisms of P. Thus, if (X_0, Y_0, Z_0) has distribution μ, then (X_0, Y_0) is uniform over $\Delta = \{(x,x) : x \in \Omega\}$.*

Proof. The values of $M^{\mathrm{good}}, M^{\mathrm{bad}}$ and the law of X, Y, Z under $\mathbb{P}(\cdot)$ are all invariant by automorphisms, so μ must be invariant as well. Moreover, going back to the definition of μ we see at once that $X_0 = Y_0$ μ-almost surely. Since the automorphism group of P is transitive over Ω, the events $\{(X_{M^{\mathrm{good}}}, Y_{M^{\mathrm{good}}}) = (x,x)\}$ ($x \in \Omega$) must all be equally likely under μ. □

We still need to define τ in order to apply Lemma 1.4 to (3.14). A seemingly natural choice would be $\tau = M^{\mathrm{bad}}$, but this would violate the third condition of the lemma: $X_{M^{\mathrm{bad}}}, Y_{M^{\mathrm{bad}}}$ in general are not uniform over Δ. We take instead

$$\tau = \inf\{t \geq M^{\mathrm{bad}} : X_t = Y_t\} = M^{\mathrm{bad}} + M^{\mathrm{good}} \circ \Theta_{M^{\mathrm{bad}}},$$

noting that

$$\int_0^{M^{\mathrm{bad}}} \mathbf{1}((X_t,Y_t)=(x,x))\,\mathrm{d}t = \int_0^{\tau} \mathbf{1}((X_t,Y_t)=(x,x))\,\mathrm{d}t$$

because there are no visits to the diagonal between times M^{bad} and τ. Analogously to the previous claim, we observe that

Claim 3.9. *The law of (X_τ, Y_τ, Z_τ) under $\mathbb{P}_\mu$ is invariant under automorphisms of P. Therefore, (X_τ, Y_τ) is uniform over Δ.*

Proof. $\mathbb{P}_\mu$ and τ are invariant by automorphisms, so the law of (X_τ, Y_τ, Z_τ) is also invariant. Moreover, since $\tau = M^{\mathrm{bad}} + M^{\mathrm{good}} \circ \Theta_{M^{\mathrm{bad}}}$, we have $X_\tau = Y_\tau$, and uniformity over Δ follows as in the previous claim. □

We now see that all conditions of Lemma 1.4 are satisfied, so for all $x \in \Omega$

$$\begin{aligned}\mathbb{E}_\mu\left[\int_0^{M^{\mathrm{bad}}} \mathbf{1}((X_t,Y_t)=(x,x))\,\mathrm{d}t\right] &= \mathbb{E}\left[\int_0^{\tau} \mathbf{1}((X_t,Y_t)=(x,x))\,\mathrm{d}t\right]\\ &= \pi(x)^2\,\mathbb{E}_\mu[\tau] = \frac{\mathbb{E}_\mu[\tau]}{n^2}.\end{aligned}$$

Combining this with (3.14), and recalling $\tau = M^{\mathrm{bad}} + M^{\mathrm{good}} \circ \Theta_{M^{\mathrm{bad}}}$, we obtain

$$\mathbb{E}[T] = \mathbb{P}\left(M^{\mathrm{good}} < M^{\mathrm{bad}}\right)\frac{\mathbb{E}_\mu[\tau]}{n} = \mathbb{P}\left(M^{\mathrm{good}} < M^{\mathrm{bad}}\right)\left(\frac{\mathbb{E}_\mu\left[M^{\mathrm{bad}}\right] + \mathbb{E}_\nu\left[M^{\mathrm{good}}\right]}{n}\right) \tag{3.15}$$

for some distribution ν over Ω^2.

Equation (3.15) gives an exact expression for $\mathbb{E}[T]$. Our last step is a simple upper bound for the RHS of this identity. Note that $M^{\mathrm{bad}} \leq M^{X,Z}$ and $M^{\mathrm{good}} = M^{X,Y}$, so that M^{good} and M^{bad} are upper bounded by meeting times between $(X_t)_{t\geq 0}$ and *independent* trajectories. We may apply Lemma 3.7 conditionally on these trajectories to obtain $\mathbb{E}_\mu\left[M^{\mathrm{bad}}\right], \mathbb{E}_\nu\left[M^{\mathrm{good}}\right] \leq 11\, t^*_{\mathrm{hit}}$. Plugging this back into (3.15) finishes the proof. □

4 The occupation time identity

It is well known that a finite irreducible chain $(V_t)_{t\geq 0}$ with state space Ω_V, started from a point x and stopped at a stopping time $\tau > 0$ with $V_\tau = x$ almost surely, satisfies

$$\forall\, v \in \Omega_V \,:\, \mathbb{E}_x\left[\int_0^\tau \mathbf{1}(X_t = v)\,\mathrm{d}t\right] = \pi_V(v)\,\mathbb{E}_x\left[\tau\right],$$

where π_V is the unique stationary measure of $(V_t)_{t\geq 0}$ (some simple conditions on τ are necessary for this). There are also extensions of this lemma to the case where V_0 and V_τ are not necessarily equal, but have the same distribution [1, Proposition 2.4, Chapter 2]. In this section we prove Lemma 1.4 from the Introduction, which extends this idea even further, and shows that τ may be a stopping time for a “larger" Markov chain.

Proof of Lemma 1.4. In this proof we will interchange integrals, expectations and summations several times. Instead of justifying this at each step, we note right away that all of these interchanges are valid, because the integrands are non-negative.

Consider the row vector h with nonnegative coordinates

$$h(v) := \mathbb{E}_\mu\left[\int_0^\tau \mathbf{1}(V_t = v)\,\mathrm{d}t\right] = \int_0^\infty \mathbb{P}_\mu\left(V_t = v, \tau > t\right)\,\mathrm{d}t \quad (v \in \Omega_V).$$

Note that $\sum_{v\in V} h(v) = \mathbb{E}_\mu\left[\tau\right] > 0$ because $\tau > 0$ a.s.. Letting Q be the generator of $(V_t)_{t\geq 0}$, we will show below that

$$hQ = 0. \tag{4.1}$$

This identity implies that $h/\mathbb{E}_\mu\left[\tau\right]$ is one invariant probability distribution for V. Since π_V is the unique invariant distribution, we deduce that for all $v \in \Omega_V$

$$\frac{h(v)}{\mathbb{E}_\mu\left[\tau\right]} = \pi_V(v),$$

which is precisely what we need to prove.

We will derive $hQ = 0$ from the limit

$$\forall v \in \Omega_V \,:\, h\,Q(v) = \lim_{\varepsilon\searrow 0} \frac{[h\,e^{\varepsilon\,Q}](v) - h(v)}{\varepsilon}. \tag{4.2}$$

In order to compute the limit we recall $e^{\varepsilon\,Q}(w,v) = \mathbb{P}_w\left(V_\varepsilon = v\right)$ for all $w, v \in \Omega_V$. Therefore

$$[h\,e^{\varepsilon\,Q}](v) = \sum_{w\in\Omega_V} h(w)\,\mathbb{P}_w\left(V_\varepsilon = v\right) = \int_0^\infty \sum_{w\in\Omega_V} \mathbb{P}_\mu\left(V_t = w, \tau > t\right)\,\mathbb{P}_w\left(V_\varepsilon = v\right)\,\mathrm{d}t.$$

Crucially, the fact that τ is a stopping time for U, V implies that the event $\{V_t = w, \tau > t\}$ is measurable with respect to $(U_s, V_s)_{s\leq t}$. Using that V and U evolve independently and the Markov property for V implies

$$\begin{aligned}\mathbb{P}_\mu\left(V_t = w, \tau > t\right)\,\mathbb{P}_w\left(V_\varepsilon = v\right) &= \mathbb{P}_\mu\left(V_t = w, \tau > t\right)\,\mathbb{P}_\mu\left(V_{t+\varepsilon} = v \mid V_t = w, \tau > t\right)\\ &= \mathbb{P}_\mu\left(V_t = w, V_{t+\varepsilon} = v, \tau > t\right).\end{aligned}$$

Plugging this back in the previous display gives

$$\begin{aligned}[h\,e^{\varepsilon\,Q}](v) &= \int_0^\infty \mathbb{P}_\mu\left(V_{t+\varepsilon} = v, \tau > t\right)dt\\ &= \int_0^\infty \mathbb{P}_\mu\left(V_{t+\varepsilon} = v, \tau > t+\varepsilon\right)\,\mathrm{d}t + \int_0^\infty \mathbb{P}_\mu\left(V_{t+\varepsilon} = v, t \leq \tau \leq t+\varepsilon\right)\,\mathrm{d}t\end{aligned}$$

$$=: \quad (I) + (II). \tag{4.3}$$

The first term is

$$(I) = \int_{\varepsilon}^{\infty} \mathbb{P}_{\mu}\left(V_t = v, \tau > t\right)\, \mathrm{d}t = h(v) - \int_0^{\varepsilon} \mathbb{P}_{\mu}\left(V_t = v, \tau > t\right)\, \mathrm{d}t,$$

so

$$\frac{(I) - h(v)}{\varepsilon} \to -\mathbb{P}_{\mu}\left(V_0 = v, \tau > 0\right) = -\mathbb{P}_{\mu}\left(V_0 = v\right) \tag{4.4}$$

because $\tau > 0$ always. Regarding the second term, we have

$$(II) = \int_0^{\infty} \mathbb{P}_{\mu}(V_t = v, V_\tau = v, \tau \le t \le \tau + \varepsilon)\, \mathrm{d}t + \int_0^{\infty} \mathbb{P}_{\mu}(V_t = v, V_\tau \ne v, \tau \le t \le \tau + \varepsilon)\, \mathrm{d}t. \tag{4.5}$$

For the first term on the right hand side above we obtain

$$\lim_{\varepsilon \to 0} \frac{\int_0^{\infty} \mathbb{P}_{\mu}(V_t = v, V_\tau = v, \tau \le t \le \tau + \varepsilon)\, \mathrm{d}t}{\varepsilon} = \mathbb{P}_{\mu}(V_\tau = v)\,. \tag{4.6}$$

As for the second term in the sum in (4.5) we get

$$\begin{aligned} &\int_0^{\infty} \mathbb{P}_{\mu}(V_t = v, V_\tau \ne v, \tau \le t \le \tau + \varepsilon)\, \mathrm{d}t \\ &= \int_0^{\infty} \mathbb{E}_{\mu}[\mathbb{P}_{\mu}(V_t = v, V_\tau \ne v, \tau \le t \le \tau + \varepsilon \mid \tau)]\, \mathrm{d}t \\ &= \int_0^{\infty} \mathbb{E}_{\mu}[\mathbf{1}(\tau \le t \le \tau + \varepsilon)\mathbb{P}_{\mu}(V_t = v, V_\tau \ne v \mid \tau)]\, \mathrm{d}t. \end{aligned}$$

On the event $\{\tau \le t \le \tau + \varepsilon\}$ in order to have $V_t = v$ and $V_\tau \ne v$, there must exist at least one jump of the Markov chain in the time interval $[\tau, t]$, which on this event has length less than ε. Therefore, we obtain that on the event $\{\tau \le t \le \tau + \varepsilon\}$

$$\mathbb{P}_{\mu}(V_t = v, V_\tau \ne v \mid \tau) = O(\varepsilon).$$

Therefore we deduce

$$\int_0^{\infty} \mathbb{E}_{\mu}[\mathbf{1}(\tau \le t \le \tau + \varepsilon)\mathbb{P}_{\mu}(V_t = v, V_\tau \ne v \mid \tau)]\, \mathrm{d}t = O(\varepsilon^2).$$

Hence this together with (4.6) gives that

$$\frac{(II)}{\varepsilon} \to \mathbb{P}_{\mu}\left(V_\tau = v\right) \text{ as } \varepsilon \searrow 0.$$

Combining this with (4.4) and (4.3) gives:

$$\frac{[h\, e^{\varepsilon Q}](v) - h(v)}{\varepsilon} \to \mathbb{P}_{\mu}\left(V_\tau = v\right) - \mathbb{P}_{\mu}\left(V_0 = v\right)$$

Our assumption that $\mathbb{P}_{\mu}\left(V_0 = \cdot\right) = \mathbb{P}_{\mu}\left(V_\tau = \cdot\right)$ implies that the right hand side above is zero. Plugging this back into (4.2) gives $hQ = 0$ and finishes the proof. □

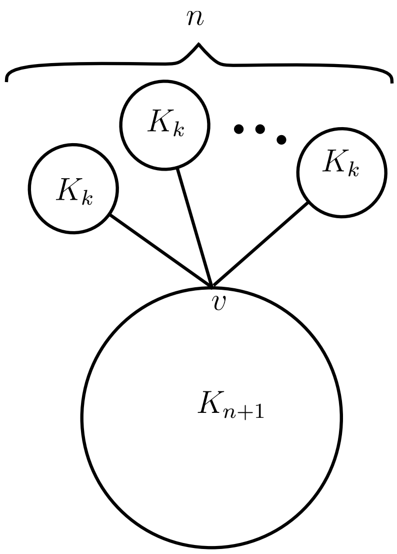

Figure 1: The graph G

5 Non transitive chains

The goal of this section is to prove Theorem 1.5. Throughout the section we fix $\varepsilon > 0$ and let $C \in \mathbb{N}$ be a perfect square satisfying $C \geq 6/\varepsilon^2$. In what follows K_r is the complete graph on $r \in \mathbb{N} \setminus \{0\}$ vertices.

For $n \in \mathbb{N}$ a perfect square, construct a graph G_n as follows: begin from a clique K_{n+1} and n disjoint copies of K_k with $k = \sqrt{Cn}$. Fix a vertex $v \in K_{n+1}$ and add exactly one edge from v to each copy of K_k. See Figure 1 for a depiction of the graph.

Let Ω be the vertex set of G_n. We call Down the set of vertices belonging to K_{n+1} and Up $= \Omega \setminus$ Down the rest.

Let P be the transition matrix of a simple random walk over G and π its stationary distribution. Let X, Y and Z be independent random walks starting from π with transition matrix P and speeds $\lambda_X = 1, \lambda_Y = 0$ and $\lambda_Z = 1$.

The idea is that by choosing ε sufficiently small, the stationary measure of Down becomes arbitrarily small. So if we start X, Y and Z according to π, then it is very likely they will all start from different cliques in Up. Let T be the $\sqrt{n}$-th time that X visits the vertex v. We will show that as $n \to \infty$ the probability that X and Z collide after time T is arbitrarily small. Moreover, we will show that the probability that X and Y collide before T is arbitrarily small as $n \to \infty$. Combining these two assertions will complete the proof.

For all $r \geq 0$ we define $\tau_v^{(r)}$ to be the time of the r-th visit to v. Formally,

$$\tau_v^{(0)} = \inf\{t \geq 0 : \, X_t = v\}$$

and for $i \geq 1$ we define

$$\tau_v^{(i)} = \inf\{t > \tau_v^{(i-1)} : \, X_t = v, \, X_{t-} \neq v\}.$$

Lemma 5.1. *There exists $\alpha = \alpha(C) > 0$ independent of n such that for all $x, z \in \Omega$ and all $r \geq 1$ we have*

$$\mathbb{P}_{x,z}\Big(M^{X,Z} > \tau_v^{(r)}\Big) \leq (1-\alpha)^{r-1}.$$

Proof. First note that by the strong Markov property we have for all $r \geq 1$

$$\sup_{x,z} \mathbb{P}_{x,z}\Big(M^{X,Z} > \tau_v^{(r)}\Big) \leq \sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(r-1)}\Big).$$

Using the strong Markov property again, for $r \geq 1$ we obtain

$$\begin{aligned}\sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(r)}\Big) &= \sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(r)} \;\Big|\; M^{X,Z} > \tau_v^{(1)}\Big)\, \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(1)}\Big) \\ &\leq \sup_w \mathbb{P}_{v,w}\Big(M^{X,Z} > \tau_v^{(r-1)}\Big) \sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(1)}\Big).\end{aligned}$$

By induction for all $r \geq 1$ this yields

$$\sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(r)}\Big) \leq \Big(\sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(1)}\Big)\Big)^r.$$

So we complete the proof by showing that

$$\sup_z \mathbb{P}_{v,z}\Big(M^{X,Z} > \tau_v^{(1)}\Big) \leq 1 - \alpha \tag{5.1}$$

for a positive constant α depending only on C.

Let $\tau = \inf\{t \geq 0 : Z_t \in \mathrm{Down} \setminus \{v\}\}$ and fix $w \in \mathrm{Down} \setminus \{v\}$. By symmetry, for all z we then have

$$\mathbb{P}_{v,z}\Big(M^{X,Z} \leq \tau_v^{(1)}\Big) \geq \frac{1}{2} \cdot \min_z \mathbb{P}_{w,z}\Big(\tau \leq \tau_v^{(0)}\Big) \min_{a,b \,\in\, \mathrm{Down}\setminus\{v\}} \mathbb{P}_{a,b}\Big(M^{X,Z} \leq \tau_v^{(0)}\Big),$$

where the factor $1/2$ corresponds to the probability that the first time X jumps it goes to $\mathrm{Down} \setminus \{v\}$.

If $X_0 = a \in \mathrm{Down} \setminus \{v\}$, then $\tau_v^{(0)} = \tau_v^X$, and hence if also $b \in \mathrm{Down} \setminus \{v\}$, then

$$\mathbb{P}_{a,b}\Big(M^{X,Z} \leq \tau_v^{(0)}\Big) \geq \mathbb{P}_{a,b}\big(M^{X,Z} \leq \tau_v^X \wedge \tau_v^Z\big) = \frac{1}{2}.$$

It remains to show that for a positive constant c_1 we have

$$\min_z \mathbb{P}_{w,z}\Big(\tau \leq \tau_v^{(0)}\Big) \geq c_1 > 0. \tag{5.2}$$

If $z \in \mathrm{Down} \setminus \{v\}$, then this probability is 1 and if $z = v$ it is easily seen to be at least $1/4$. So we assume that $z \in \mathrm{Up}$. Let x be the unique neighbour of v lying in the same clique as z. Then the time τ can be expressed as $\tau = T_{z,x} + T_{x,v} + T_{v,\mathrm{Down}\setminus\{v\}}$, where the time $T_{r,S}$ stands for the first hitting time of S starting from r. Using this, it is then not hard to see that there exists a positive constant c such that uniformly over all $z \in \mathrm{Up}$ we have $\mathbb{E}[\tau] \leq ck^2$. Moreover, if $X_0 \in \mathrm{Down} \setminus \{v\}$, then $\tau_v^{(0)}$ is an exponential random variable with mean n. By Markov's inequality we obtain

$$\mathbb{P}_{w,z}\Big(\tau \leq \tau_v^{(0)}\Big) \geq \mathbb{P}_z(\tau \leq 2\mathbb{E}[\tau]) \cdot \mathbb{P}_w\Big(\tau_v^{(0)} \geq 2\mathbb{E}[\tau]\Big) \geq \frac{1}{2} \cdot \int_{2\mathbb{E}[\tau]}^{\infty} n e^{-ns}\, ds = \frac{1}{2} e^{-2n\mathbb{E}[\tau]}.$$

Note that this bound does not depend on z. Since $k = \sqrt{Cn}$ and $\mathbb{E}[\tau] \leq ck^2$ the bound in (5.2) follows. □

Proof of Theorem 1.5. We show that for n sufficiently large, the graph $G = G_n$ satisfies the claim of the theorem.

It is not hard to verify that for n large enough, in G_n we have

$$\pi(\mathrm{Down}) \leq \frac{2}{C}.$$

Let A be the set of pairs (x,y) such that $y \in \mathrm{Up}$ and x is not in the same clique as y. Then let $E = \{(X_0, Y_0) \in A\}$. By the preceding bound $\mathbb{P}(E^c) \le 3/C = \varepsilon/2$ for n large enough. We then have

$$\begin{aligned}\mathbb{P}\big(M^{\mathrm{good}} \le M^{\mathrm{bad}}\big) &\le \mathbb{P}\big(M^{X,Y} \le M^{X,Z}\big)\\ &= \mathbb{P}\big(M^{X,Y} \le M^{X,Z}, E\big) + \mathbb{P}\big(M^{X,Y} \le M^{X,Z}, E^c\big) &(5.3)\\ &\le \sup_{x,y,z\colon (x,y)\in A} \mathbb{P}_{x,y,z}\big(M^{X,Y} \le M^{X,Z}\big) + \frac{\varepsilon}{2}. &(5.4)\end{aligned}$$

Therefore, it suffices to upper bound the last probability appearing above. Fix any $z \in \Omega$ and $(x,y) \in A$. For r to be determined later we have

$$\mathbb{P}_{x,y,z}\big(M^{X,Y} \le M^{X,Z}\big) \le \mathbb{P}_{x,y}\Big(M^{X,Y} \le \tau_v^{(r)}\Big) + \mathbb{P}_{x,z}\Big(M^{X,Z} > \tau_v^{(r)}\Big). \qquad (5.5)$$

Because Y is not moving, we have $M^{X,Y} = \tau_y$, where $y = Y_0$. Since x, y are not in the same clique of Up, if $\tau_y \le \tau_v^{(r)}$, then there exists $1 \le i \le r$ such that $\tau_v^{(i-1)} < \tau_y \le \tau_v^{(i)}$. By the strong Markov property and union bound we obtain

$$\mathbb{P}_{x,y}\Big(M^{X,Y} \le \tau_v^{(r)}\Big) \le r\mathbb{P}_v\Big(\tau_y \le \tau_v^{(1)}\Big) \le \frac{r}{2n},$$

since, when $X_0 = v$, in order to hit $y \in \mathrm{Up}$ before returning to v, the first time X moves it must jump into the clique that contains y.

Using the above bound and Lemma 5.1 in (5.5) we deduce

$$\mathbb{P}_{x,y,z}\big(M^{X,Y} \le M^{X,Z}\big) \le \frac{r}{2n} + (1-\alpha)^{r-1}.$$

Taking $r = \sqrt{n}$ or any other function of n that goes to infinity slower than n gives that

$$\mathbb{P}_{x,y,z}\big(M^{X,Y} \le M^{X,Z}\big) \to 0 \quad \text{as } n \to \infty.$$

We conclude from (5.3) that

$$\mathbb{P}\big(M^{\mathrm{good}} \le M^{\mathrm{bad}}\big) < \varepsilon$$

and this finishes the proof. □

6 Sharpness of Conjecture 1.6

In this section we describe the example pointed out by Alexander Holroyd, mentioned in the Introduction, of a family of transitive graphs for which $\mathbb{P}\big(M^{\mathrm{good}} \le M^{\mathrm{bad}}\big) \le 1/3 + \delta$. In what follows we take $\lambda_X = \lambda_Y = 1$ and $\lambda_Z = 0$.

To construct the example, fix $\varepsilon \in (0,1)$ and consider the chain with state space $\{0,1\}^n$ in which the j'th coordinate changes value (from 0 to 1 or vice-versa) at rate $q_j = \varepsilon^{j-1}(1-\varepsilon)/(1-\varepsilon^n)$; note that $\sum_{i=1}^n q_i = 1$. The idea is that for small ε, earlier coordinates change state much more quickly than later coordinates, so the primary obstacle to both meeting and hitting is simply the largest coordinate in which the value differs. For $u, v \in \{0,1\}^n$, let $k(u,v) = \max\{i : u_i \ne v_i\}$, or $k(u,v) = 0$ if $u = v$.

We claim that for $x, y, z \in \{0,1\}^n$, if $k(x,y) > \min(k(x,z), k(y,z))$ then we have $\mathbb{P}_{x,y,z}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) < 2\varepsilon n$. Assuming this, and taking $\varepsilon = \delta/(2n)$, it follows by symmetry that, starting from stationarity,

$$\mathbb{P}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) \le \delta + \mathbb{P}(k(X_0,Y_0) < \min(k(X_0,Z_0), k(Y_0,Z_0)) < \delta + \frac{1}{3}.$$

It thus remains to prove the preceding claim.

Fix $x, y, z \in \{0,1\}^n$ with $k(x,y) > \min(k(x,z), k(y,z))$, and assume by symmetry that $k(x,z) < k(x,y)$. For $1 \le k \le n$, let $\tau_k = \min\{t : X_t^{(i)} = z_i, 1 \le i \le k\}$ be the first time that X_t and z agree in the first k coordinates. It is convenient to set $\tau_0 = 0$. Also, let $\sigma_k^X = \min\{t : \exists i \ge k,\ X_t^{(i)} \ne X_0^{(i)}\}$ be the first time one of the last $n-k+1$ coordinates of X changes, and define σ_k^Y accordingly.

We will show that for all $1 \le k < n$,

$$\mathbb{P}\big(\tau_k < \sigma_{k+1}^X\big) \ge (1-\varepsilon)^k \ge 1 - k\varepsilon. \tag{6.1}$$

Note that τ_k, σ_{k+1}^X and σ_{k+1}^Y are all independent. Furthermore, σ_{k+1}^X and σ_{k+1}^Y are identically distributed, so if the preceding inequality holds as written then it also holds with σ_{k+1}^Y in place of σ_{k+1}^X. We finish proving the claim assuming that (6.1) holds, then conclude by proving (6.1).

At time $\tau_{k(x,z)}$, the first $k(x,z)$ coordinates of X agree with those of z. If $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^X$ then the remaining coordinates of X and z also agree (because they did at time 0 and they have not changed), so $M^{X,Z} = \tau_{k(x,z)}$. Similarly, if $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^X$ and $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^Y$ then $M^{\mathrm{bad}} < M^{\mathrm{good}}$. It then follows, using (6.1) and the subsequent observation, that

$$\mathbb{P}_{x,y,z}\big(M^{\mathrm{good}} < M^{\mathrm{bad}}\big) \le \mathbb{P}\Big(\sigma_{k(x,z)+1}^X \le \tau_{k(x,z)}\Big) + \mathbb{P}\Big(\sigma_{k(x,z)+1}^Y \le \tau_{k(x,z)}\Big) \le 2k(x,z)\varepsilon < 2n\varepsilon\,,$$

as claimed. It thus remains to prove (6.1). In what follows we write $\sigma_k = \sigma_k^X$.

Fix $1 \le k < n$, and note that σ_k is exponential with rate $\sum_{j=k}^n q_j = \varepsilon^{k-1}(1-\varepsilon^{n+1-k})/(1-\varepsilon^n)$. Furthermore, $\sigma_k < \sigma_{k+1}$ precisely if the k'th coordinate of X changes before any larger coordinate. It follows that $\mathbb{P}(\sigma_k < \sigma_{k+1}) = q_k / \sum_{j=k}^n q_j$.

Suppose that $X_0^{(k)} = z_k$. In this case to have $\tau_k < \sigma_{k+1}$ it suffices that $\tau_{k-1} < \sigma_k$, so

$$\mathbb{P}\Big(\tau_k < \sigma_{k+1} \ \Big|\ X_0^{(k)} = z_k\Big) \ge \mathbb{P}(\tau_{k-1} < \sigma_k)\,.$$

If $X_0^{(k)} \ne z_k$ then the k-th coordinate must change before time τ_k, so to have $\tau_k < \sigma_{k+1}$ it is necessary that $\sigma_k < \sigma_{k+1}$.

By the strong Markov property, we then have

$$\begin{aligned}
\mathbb{P}\Big(\tau_k < \sigma_{k+1} \ \Big|\ X_0^{(k)} \ne z_k\Big) &= \mathbb{P}(\sigma_k < \sigma_{k+1})\, \mathbb{P}\Big(\tau_k < \sigma_{k+1} \ \Big|\ X_0^{(k)} = z_k\Big) \\
&= \frac{q_k}{\sum_{j=k}^n q_j} \cdot \mathbb{P}\Big(\tau_k < \sigma_{k+1} \ \Big|\ X_0^{(k)} = z_k\Big) \\
&\ge \frac{1-\varepsilon}{1-\varepsilon^{n+1-k}} \mathbb{P}(\tau_{k-1} < \sigma_k)\ .
\end{aligned}$$

We thus have the unconditional bound

$$\mathbb{P}(\tau_k < \sigma_{k+1}) \ge \frac{1-\varepsilon}{1-\varepsilon^{n+1-k}} \mathbb{P}(\tau_{k-1} < \sigma_k) > (1-\varepsilon)\mathbb{P}(\tau_{k-1} < \sigma_k)\,,$$

This bound holds for all $1 \le k < n$; since $\tau_0 = 0$ we also have $\mathbb{P}(\tau_0 < \sigma_1) = 1$, and (6.1) follows.

References

[1] David Aldous and James Fill, *Reversible Markov chains and random walks on graphs*, In preparation, http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] Aldous, D. J. and Brown, M. (1992). Inequalities for rare events in time-reversible Markov chains. I. In *Stochastic inequalities (Seattle, WA, 1991)*. IMS Lecture Notes Monogr. Ser., Vol. **22**. Inst. Math. Statist., Hayward, CA, 1–16. MR-1228050

[3] Levin, D. A., Peres, Y., and Wilmer, E. L. (2009). *Markov chains and mixing times*. American Mathematical Society, Providence, RI. With a chapter by James G. Propp and David B. Wilson. MR-2466937

[4] Oliveira, R. I. (2012). On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.* **364**, 4, 2109–2128. MR-2869200

[5] Sousi, P. and Winkler, P. (2014). Mixing times and moving targets. *Combin. Probab. Comput.* **23**, 3, 460–476. MR-3189422

Acknowledgments. The authors thank the Bellairs Institute, Microsoft Research and the Isaac Newton Institute, where parts of this work were completed.