

MONOTONICITY FOR CONTINUOUS-TIME RANDOM WALKS

BY RUSSELL LYONS^a AND GRAHAM WHITE^b

Department of Mathematics, Indiana University, ^ardlyons@indiana.edu, ^bgrahamwhite@alumni.stanford.edu

Consider continuous-time random walks on Cayley graphs where the rate assigned to each edge depends only on the corresponding generator. We show that the limiting speed is monotone increasing in the rates for infinite Cayley graphs that arise from Coxeter systems but not for all Cayley graphs. On finite Cayley graphs, we show that the distance—in various senses—to stationarity is monotone decreasing in the rates for Coxeter systems and for abelian groups but not for all Cayley graphs. We also find several examples of surprising behaviour in the dependence of the distance to stationarity on the rates. This includes a counterexample to a conjecture on entropy of Benjamini, Lyons, and Schramm. We also show that the expected distance at any fixed time for random walks on $\mathbb{Z}^+$ is monotone increasing in the rates for arbitrary rate functions, which is not true on all of $\mathbb{Z}$. Various intermediate results are also of interest.

1. Introduction. We are interested in questions arising from the following scenario. Consider a connected, undirected graph G with each edge e labelled by a nonnegative real number $r(e)$. A random walk is defined on this graph by associating a Poisson process (“clock”) with rate $r(e)$ to each edge e . When this clock rings, if the walker is at either neighbouring vertex, then it moves along the edge e to the other vertex; otherwise it does not move. We choose the convention that the walk is left-continuous in time, which will be more convenient to describe our couplings. Note that if G is finite, then the stationary distribution of this walk is uniform.

Monotonicity in time or in the rate function $\mathbf{r} := r(\cdot)$ of the behaviour of such a random walk is complicated and not always intuitive, primarily because increasing $r(e)$ not only makes the walk more likely to cross e when at each of its endpoints, but also because it moves the walk sooner in both directions of traversing e . Let $p_t(x, y)$ be the transition probability from x to y at time t . Note that changing the rate function by a constant factor α is equivalent to changing the time t to $\alpha \cdot t$. When $x = y$ and t is fixed, there are examples where $p_t(x, x)$ is not monotone in the rate function (one such is given here in Example 4.7). Nevertheless, if G is finite, then the average of $p_t(x, x)$ over all vertices x of G is monotone, as shown by Benjamini and Schramm [17], Theorem 3.1. In particular, if the graph with rates is vertex-transitive, then $p_t(x, x)$ is monotone decreasing in the rates. This result was extended in various ways by [1, 14, 23, 24, 29], but all such extensions concerned only return probabilities.

In a different direction, Karlin and McGregor ([21], equation (45)) considered continuous-time random walks on $\mathbb{Z}^+ := \{0, 1, 2, \dots\}$ with arbitrary, symmetric rate functions. It follows from their result that, for each $n \geq 1$, the time to reach n from 0 stochastically decreases in $\mathbf{r}$. Indeed, they gave an explicit representation of that time as the sum of n independent, exponential random variables whose rates are the nonzero eigenvalues of the $(n+1) \times (n+1)$ tridiagonal Laplacian matrix Δ , where $\Delta(i, i) := r(i, i+1) + r(i, i-1)$, $\Delta(i, i+1) := -r(i, i+1)$, $\Delta(i, i-1) := -r(i, i-1)$, and all other entries are 0. Because Δ itself is

Received October 2021; revised September 2022.

MSC2020 subject classifications. Primary 05C81, 60G50; secondary 20F55, 60B15.

Key words and phrases. Cayley graphs, Coxeter groups, speed, mixing, majorization.

monotone increasing in $\mathbf{r}$ in the Loewner order, its eigenvalues are also increasing in the rate function. (This monotonicity is also what is behind all the results mentioned above on the return probabilities. We note that [21], equation (45), holds for more general birth and death chains; [22] gives a representation of transition probabilities via orthogonal polynomials.) See also [10–12, 27] for other proofs and extensions of the representation [21], equation (45).

We are motivated by the following observation [25], Exercise 13.24: If the 3-regular tree is regarded as the standard Cayley graph of the free product of three copies of $\mathbb{Z}/2\mathbb{Z}$ and the rates on all edges are 1 except for those corresponding to a fixed generator, where the rates are ρ , then the limiting rate of escape (i.e., the graph distance from the starting point divided by the time as time tends to infinity) is

$$(3\rho(\rho + 1) + (1 - \rho)\sqrt{16\rho + 9\rho^2})/(2(2 + \rho)) \quad \text{a.s.}$$

What is interesting about this formula is that it is monotone increasing in ρ but that this fact is not obvious. Is there a more intuitive explanation for such monotonicity? One might attempt to answer this via a coupling of two such random walks, where one is always at distance at least as great as the other, but no such coupling can be Markovian. In fact, even non-Markovian couplings cannot always have this property; see Proposition 2.37. Other questions of interest include these: For rates that depend only on the generators for the free product of more copies of $\mathbb{Z}/2\mathbb{Z}$, does monotonicity still hold? What about other Cayley graphs? What about other aspects of the random walk behaviour? For example, on a finite graph if some of the rates are increased, does this necessarily improve the convergence to the stationary (uniform) distribution? How does the total time spent at each vertex depend on the rates?

Consider a Cayley graph G of a group Γ generated by a finite set $S \subset \Gamma$. Write o for the identity element and $|x|$ for the graph distance between o and x . Given $r_s > 0$ for each $s \in S$, let $r(e) := r_s$ when e is an edge of G corresponding to $s \in S$. Throughout this paper we will assume that generating sets are symmetric—that is, if s is a generator, then so is s^{-1} , and that the corresponding rates r_s and $r_{s^{-1}}$ are equal.

Let $(Z_t)_{t \geq 0}$ be the corresponding random walk starting from o and $\sigma(\mathbf{r}) := \lim_{t \rightarrow \infty} \mathbf{E}[|Z_t|]/t$ be its limiting speed. This limit exists by a well-known subadditivity argument; the subadditive ergodic theorem also shows that $\lim_{t \rightarrow \infty} |Z_t|/t = \sigma(\mathbf{r})$ a.s. Our main result for infinite graphs is the following.

THEOREM 1.1. *If (Γ, S) is a Coxeter system, then $\sigma(\mathbf{r})$ is monotone increasing in $\mathbf{r}$. If (Γ, S) is irreducible and nonelementary hyperbolic, then $\sigma(\mathbf{r})$ is strictly increasing in $\mathbf{r}$.*

See Corollary 2.19 and Theorem 2.33 for the proofs. Cayley graphs that are trees are examples of Cayley graphs that arise from Coxeter systems. The theorem does not extend to all Cayley graphs; see Example 2.14.

We then consider finite Cayley graphs and examine the convergence to stationarity. It is plausible that increasing some of the rates might always improve the convergence to the stationary distribution (which is the uniform distribution), but this is not true in general. Our main positive result on this topic is the following.

THEOREM 1.2. *Let (Γ, S) be a finite Coxeter system. Let $\mathbf{r}$ and $\mathbf{r}'$ be two sets of rates on S with $r_s \leq r'_s$ for all $s \in S$. Let $t > 0$. Denote the corresponding transition probabilities by $p_t(x, y; \mathbf{r})$ and $p_t(x, y; \mathbf{r}')$. Then $p_t(o, \cdot; \mathbf{r})$ majorizes $p_t(o, \cdot; \mathbf{r}')$ with inequality if $\mathbf{r} \neq \mathbf{r}'$.*

This implies that increasing rates leads to distributions that are closer to the stationary distribution in many senses, including ℓ^p for $1 \leq p \leq \infty$ and relative entropy; see Definition 3.4 for the definition of majorization and Theorem 3.9 for the proof. The edge graph of the permutohedron is an example of a Coxeter Cayley graph, where Γ is a symmetric group. For certain special Markov chains in discrete time, a much stronger inequality for total variation distance was proved by [13], Theorem 8.3, which slightly extended a result of [28]. Similar strong inequalities, using what we call refresh rings, hold in our context as well; see Theorems 2.21 and 3.10.

Despite the impossibility of “perfect” couplings, Markovian couplings will be crucial to our proofs of Theorems 1.1 and 1.2.

In Section 3 we prove results of this kind for some other special cases, such as for every Cayley graph of an abelian group or for any group when distance from stationarity is measured with either the ℓ^2 - or ℓ^∞ -distance. We then discuss examples where increasing a rate worsens the distance to stationarity, which we find in the Cayley graphs of groups as small as the dihedral group D_5 and the symmetric group S_4 . How easy it is to find these examples depends on which ℓ^p -distance is used to measure the distance to stationarity. For instance, we find examples in dihedral groups for p very close to 4, both above and below 4, but are unable to find any for p equal to 4. In the symmetric group S_5 , we do find examples for p exactly equal to 4. Hermon and Kozma [18] find examples of generating sets for symmetric groups where increasing some rates by a tiny amount has the effect of increasing the mixing time by a large amount, which also means that the ℓ^1 -distance increases by a large amount. Their examples have unbounded degree as the size of the group tends to infinity.

A third result goes beyond the setting of Cayley graphs and concerns arbitrary rate functions $\mathbf{r}$ but is restricted to the ray graph, that is, the nearest-neighbor graph on $\mathbb{Z}^+$. Our result complements that of Karlin and McGregor given above.

THEOREM 1.3. *With $\mathbf{r}$ being an arbitrary positive rate function on the edges of the nearest-neighbor graph on $\mathbb{Z}^+$ and $t > 0$, $\mathbf{E}[|Z_t|]$ is monotone increasing in $\mathbf{r}$.*

See Corollary 4.4 for the proof. The theorem does not extend to walks on all of $\mathbb{Z}$; see Example 4.7.

Another aspect of the behaviour of random walk concerns the entropy: Let $h_t(x, \mathbf{r})$ be the (natural-log) entropy of Z_t when started at vertex x . The following conjecture [2], Conjecture 4.11, turns out to be false.

CONJECTURE 1.4. *Let G be a finite graph and $t > 0$. Then $|V(G)|^{-1} \sum_{x \in V(G)} h_t(x, \mathbf{r})$ is monotone increasing in $\mathbf{r}$.*

For a counterexample, let G be a star with six vertices, all edges but one having rate 1 and the other having rate either 10 or 20. Direct calculation shows that the mean entropies at time 1 are, respectively, approximately 1.626355024 and 1.626293845. Theorem 1.2 implies that the conjecture does hold for rates depending on generators in the setting of Coxeter systems, but changing the generators, even on dihedral groups, can make it fail; see Section 3.2. However, for the purposes of [2] it would suffice that there be a lower bound on the factor by which the mean entropy can decrease; we do not know whether this weakened statement is true.

Other monotonicity results for random walks include [30] and [26], which study monotonicity in time of $p_t(x, y)/p_t(x, x)$, and [8], which studies the expected range of symmetric random walks on $\mathbb{Z}^d$ when extra steps are included deterministically.

2. Coxeter systems.

2.1. Background on Coxeter systems. A Coxeter system (W, S) is a group W with generators S defined by a presentation of the form $\langle S \mid \forall s, s' \in S (ss')^{m(s,s')} \rangle$, where each $m(s, s')$ is either a positive integer or ∞ . Each $m(s, s)$ is equal to 1, and $m(s, s') \geq 2$ for $s \neq s'$. When $m(s, s') = \infty$, the interpretation is that no such relation is imposed. In particular, all generators are involutions. Furthermore, $m(s, s') \in \{1, 2\}$ iff s and s' commute. As always in this paper we assume that S is finite. It is customary to denote Coxeter groups with the letter W (for “Weyl”), and we will do that in order to make it easier to discern which of our results apply to Coxeter groups. Coxeter groups are abstractions of reflection groups; dihedral and symmetric groups are simple examples. Section 1.2 of [3] is devoted to these and other examples.

We always use *right Cayley graphs*: the vertex set is a group Γ , and the unoriented edges are $\{\{x, xs\}; x \in \Gamma, s \in S\}$. One could use multigraphs, but nothing would change for our questions of interest. The *Cayley diagram* of a group Γ with generators S is the corresponding Cayley graph with edges labelled by the corresponding generator; if the generator is not an involution, then also the edge is given an orientation so that multiplication (on the right) by the generator maps the tail of the edge to the head. For $\gamma \in \Gamma$, we write L_γ for left multiplication by γ , which is an automorphism of the Cayley diagram of Γ .

If (W, S) is a Coxeter system, then the Cayley graph of W with respect to S has many nice symmetry properties. One simple property is that the Cayley graph is bipartite, because all relations give even-length cycles. When G is a bipartite Cayley graph and L is a left multiplication that interchanges the endpoints of some edge, then we call L a *reflection*. We also call the set M of edges preserved by L a *wall*, a *hyperplane*, or a *mirror*. A key property for our purposes is the following, which will allow us to use arguments akin to the well-known reflection principle for one-dimensional random walks and Brownian motion.

LEMMA 2.1. *Let w and ws be two adjacent vertices in the Cayley graph of a Coxeter system, (W, S) . Let M be the set of edges preserved by the automorphism $L := L_{ws}w^{-1}$. The map L interchanges the endpoints of each edge of M . The wall M separates the vertices closer to w from those closer to ws . Suppose that $|w| < |ws|$; then for all $v \in W$, v is closer to w than to ws iff $|v| < |Lv|$.*

These properties of walls are well known, but not all are easy to find explicitly stated in this form. Because they are crucial to our results, we provide a proof. Denote the shortest-path metric in a graph by dist .

PROOF. Let G be the Cayley graph. Refer to vertices closer to w than to ws as white and the others as black; also, call an edge grey if its endpoints have different colours. If y is white and z is black, then every path from y to z must include a grey edge.

Merely because G is a bipartite graph, no vertex is equidistant from w and ws . Because L preserves distances and interchanges w and ws , it, therefore, also interchanges white and black vertices. In particular, every preserved edge is grey, and the endpoints of grey edges are interchanged. We need to use more than bipartiteness to show that all grey edges belong to the wall.

Let (x, xs') be a grey edge with x white. First, merely by bipartiteness, $\text{dist}(w, x) = \text{dist}(ws, xs')$. We may write this last conclusion as $|w^{-1}x| = |sw^{-1}xs'|$. Suppose that $\mathbf{w}$ is a reduced word for $w^{-1}x$. Then $s\mathbf{w}s'$ is a longer word that, when reduced, has the same length as $\mathbf{w}$. By the deletion condition of Coxeter systems ([3], Proposition 1.4.7, [19], Corollary 5.8, or [9], Theorem 3.2.17), there are two letters from $s\mathbf{w}s'$ that can be deleted in order to obtain a

reduced word for the same element. Since x is white and ws is black, $|sw^{-1}x| = |w^{-1}x| + 1$, and so the word sw must be reduced; similarly, the word ws' is reduced. Checking cases then reveals that the two letters to delete must be the initial s and the final s' , which yields the word w . In other words, $sw^{-1}xs' = w^{-1}x$. Another way to say this is that $L_{ws^{-1}x} = xs'$. In particular, (x, xs') belongs to the wall, as claimed.

Finally, suppose that o is white. If v is white, then let $(x, xs') \in M$ be an edge belonging to a geodesic from o to Lv . We have $|v| \leq \text{dist}(o, x) + \text{dist}(x, v) = \text{dist}(o, x) + \text{dist}(xs', Lv) = |Lv| - 1 < |Lv|$. If v is, instead, black, then Lv is white, so $|Lv| < |L(Lv)| = |v|$. $\square$

COROLLARY 2.2. *Let (W, S) be a Coxeter system. The set of reflections equals the set of left multiplications $L_{ws^{-1}}$ for $w \in W$ and $s \in S$. If $a, b \in S$ and $w, x \in W$, then (w, wa) and (x, xb) belong to the same wall iff $waw^{-1} = xbx^{-1}$.*

PROOF. It is clear that $L_{ws^{-1}}$ interchanges the endpoints of the edge (w, ws) , and, conversely, if L_γ interchanges those endpoints, then $\gamma w = ws$, that is, $\gamma = wsw^{-1}$. For the last statement in the corollary, if the two edges belong to the same wall, defined, say, by a reflection L_γ , then by Lemma 2.1, $\gamma w = wa$ and $\gamma x = xb$, whence $waw^{-1} = \gamma = xbx^{-1}$. The converse is proved similarly. $\square$

Thus, there is a natural bijection between reflections and walls. We will identify a wall M with the (disconnected) graph formed by the edges $E(M)$ in the wall and their endpoints $V(M)$.

In the case of a free Coxeter group (i.e., the free product of copies of $\mathbb{Z}/2\mathbb{Z}$), whose Cayley graph is a tree, walls are the same as single edges. Figure 1 shows the Cayley diagram of the Coxeter group $\langle a, b, c \mid a^2, b^2, c^2, (ab)^7, (bc)^2, (ca)^3 \rangle$ corresponding to reflections in the sides of a hyperbolic triangle of angles $(\pi/2, \pi/3, \pi/7)$. Each white geodesic crosses through the edges of a wall. Note that, in this case, each wall contains edges corresponding to each of the three generators.

When we prove strict monotonicity of speed in the rates, we will need the following three lemmas, the first two of which are well known.

LEMMA 2.3. *Let M be a wall in a Coxeter Cayley graph. Then every geodesic uses at most one edge in M .*

PROOF. It suffices to show that if $(w, wa), (v, vb) \in M$ both belong to a path from w to v , then there is a shorter path from w to v . Indeed, if P is a subpath from wa to vb , then the path obtained from P by applying the reflection in M is a path from w to v that is shorter by at least two edges. $\square$

LEMMA 2.4. *Let (W, S) be a Coxeter system. There is some finite K such that for every wall M , every geodesic between two vertices of M stays within distance K of M .*

PROOF. By the parallel-wall theorem of Brink and Howlett ([4] or [7], Theorem C), there is some K so that for any pair (w, M) , where w is a vertex at distance at least K from a wall M , there is some other wall $M' \neq M$ that separates w from M . By Lemma 2.3 no geodesic between vertices of M can cross M' , whence it cannot go farther than K from M . $\square$

LEMMA 2.5. *Let M be a wall in a Coxeter system (W, S) . Let $W_M := \{w \in W; L_w V(M) = V(M)\}$. Then W_M acts quasi-transitively by graph automorphisms of M ; in other words, there are only finitely many orbits in $V(M)$ under the left action of W_M . Moreover, the orbit of $w \in V(M)$ is $\{v \in V(M); (v, vs) \in E(M)\}$, where s is the generator for which $(w, ws) \in E(M)$.*

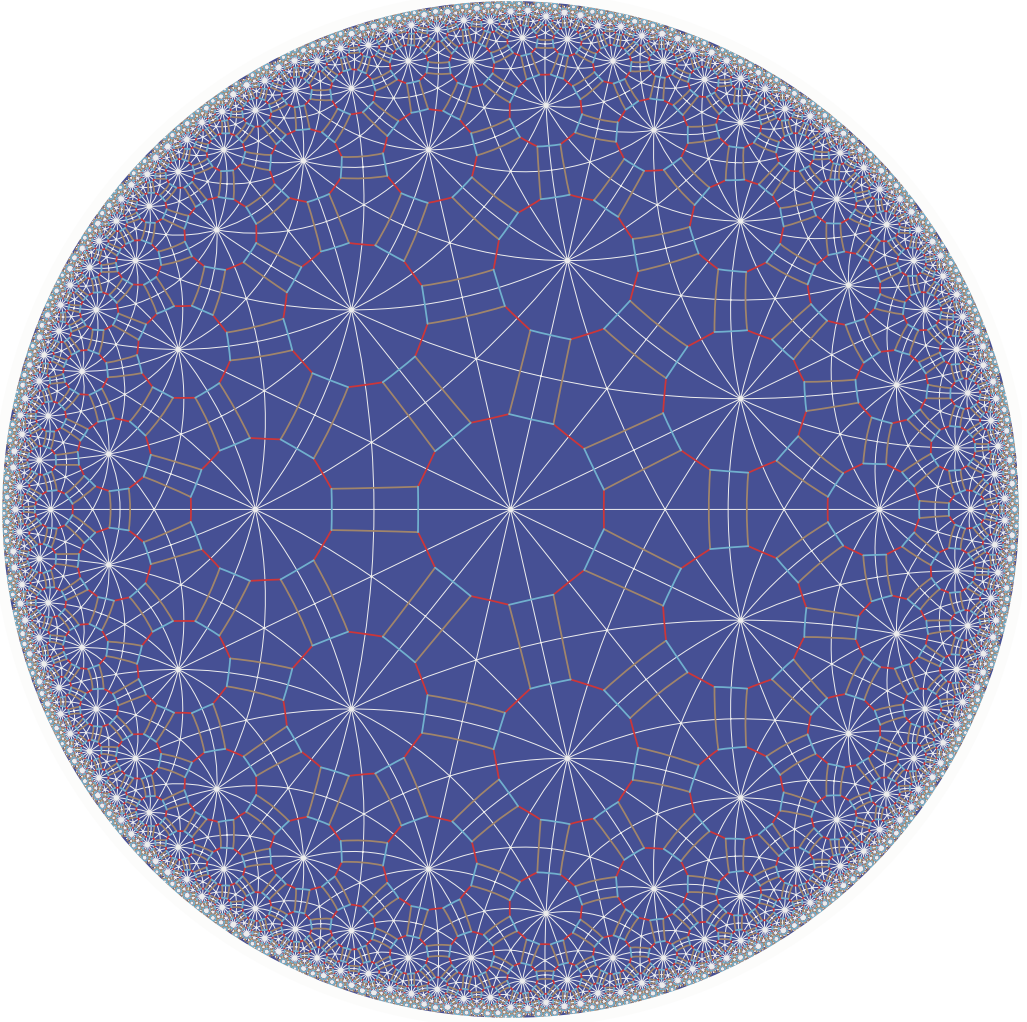


FIG. 1. A portion of the Cayley graph of the $(2, 3, 7)$ -triangle group, drawn by Matthias Weber. Edges are coloured red, light blue, or light brown according to the corresponding generator. Each dark blue face corresponds to a relation. Each white geodesic corresponds to a wall.

PROOF. Suppose that $(w, ws), (v, vs) \in E(M)$. Clearly, $L_{vw^{-1}}$ carries the first edge to the second. We claim that $vw^{-1} \in W_M$. To prove our claim, let $(x, xa) \in E(M)$. We have $ws w^{-1} = vs v^{-1} = xax^{-1}$ by Corollary 2.2, whence $vw^{-1}xa(vw^{-1}x)^{-1} = vs v^{-1}$ so that, indeed, $(vw^{-1}x, vw^{-1}xa) \in E(M)$. Therefore, for each pair of edges in M corresponding to the same generator, there is an element of W_M that takes one edge to the other. Since S is finite, it follows that this action is quasi-transitive. Finally, since left multiplications preserve the Cayley diagram, the orbit of w cannot be any larger than what is claimed. $\square$

2.2. Monotonicity of speed. A key technical device we will use to construct our couplings is the following alternative way to generate our random walks.

DEFINITION 2.6. If an edge e has an attached Poisson clock of rate r , we will generally consider this process as being controlled by a Poisson clock of rate $2r$, and when this new clock rings, there is a $\frac{1}{2}$ chance that the original clock rings, otherwise nothing happens. We will reserve the upper case R (for “refresh time”) for this Poisson process of twice the rate, with subscripts denoting the edge in question, for instance, R_e . We may use a fair coin flip to

decide whether the original clock rings; if so, we will say the coin flip is “move,” otherwise that the coin flip is “stay.” Alternatively, if the walk is at an endpoint of e at the time that R_e rings, then we may randomise the walk immediately after that ring to be at either endpoint of e with equal probability. In the case of Cayley graphs where the rates depend only on the generators, we observe that it suffices to use only one Poisson process per generator rather than one per edge; we then use a subscript corresponding to the generator.

Although our goal is Theorem 1.1, we begin for concreteness and clarity with a specific example, the 3-regular tree.

EXAMPLE 2.7. Consider the 3-regular tree as the Cayley graph of the group $\langle a, b, c \mid a^2, b^2, c^2 \rangle$, so that each edge is associated to one of the three generators, a , b , or c . Choose rates r_a , r_b , and r_c for each generator, and use those as the rates for each corresponding edge.

It is well known that the random walk of Example 2.7 is transient if and only if all three rates are positive.

Before we can discuss the dependence of the escape speed on the rates r_a , r_b , and r_c , we need some preliminary results.

PROPOSITION 2.8. *Consider Example 2.7. If w and w' are two adjacent vertices with w closer to the initial position than w' , then, for all $t > 0$, we have $\mathbf{P}[Z_t = w] > \mathbf{P}[Z_t = w']$.*

PROOF. Without loss of generality, assume that the edge between w and w' is labelled by the generator a , so that $w' = wa$.

We prove our result by giving a probability-preserving injection from paths (of the random walk up to time t) resulting in w' to paths resulting in w . For any path that ends at w' , let T be the first time $\leq t$ at which the path is at w and R_a rings. Modify the path by changing the outcome of the coin flip attached to this ring of R_a , and proceed with the rest of the path by multiplying by a , b , or c when appropriate, not changing any of the other rings of the Poisson processes or the coin flips. This has the effect of applying the reflection $L_{waw^{-1}}$ to each state after time T , so paths that ended at w' now end at w .

Indeed, this map is a probability-preserving bijection from paths that ever cross the edge from w to w' and end at w' to paths that end at w and are at w at some ring of R_a . Crucially, it is impossible to get to w' without crossing the edge (w, w') , and hence without being at w when R_a rings. However, there is a positive probability that the walk ends at w without ever being at w when R_a rings, giving our strict inequality. $\square$

Proposition 2.8 actually has a much shorter proof: Let $\tau := \inf\{t > 0; Z_t = w\}$, and let $p_t(x, y)$ denote the transition probabilities of Z_t . Then

$$\begin{aligned} \mathbf{P}[Z_t = w'] &= \mathbf{E}[\mathbf{P}[Z_t = w' \mid \tau]] = \mathbf{E}[p_{t-\tau}(w, w'); \tau < t] \\ &< \mathbf{E}[p_{t-\tau}(w, w); \tau < t] = \mathbf{P}[Z_t = w] \end{aligned}$$

(compare (3.1)). The reason we did not give this proof is that Proposition 2.8 may be generalized to the Cayley graph of any Coxeter system but not with this short proof. This generalization is a continuous-time version of Theorem 1 of [34], and the proof is essentially the same. If M is a wall, let us call a time t an M -refresh time for our random walk if $Z_t \in M$ and the refresh clock R_a rings at time t , where a is the generator such that $(Z_t, Z_t a) \in M$. The main difference between the case of trees and the general case of Coxeter systems is that in the tree, every path that reaches w' must cross the edge between w and w' . In a general Coxeter system, this is no longer the case, so instead we need to consider the steps at which the walk might move from a state closer to w to one closer to w' , in other words, the M -refresh times for the wall determined by (w, w') .

THEOREM 2.9. *Let (W, S) be a Coxeter system. To each of the generators $s \in S$, attach a Poisson clock of rate r_s . Consider a random walk starting at the identity that moves from a current location w to ws when the s -clock rings. If w and w' are two adjacent vertices with $|w| < |w'|$, then, for all $t > 0$, we have $\mathbf{P}[Z_t = w] > \mathbf{P}[Z_t = w']$.*

PROOF. Let $w' = wa$. We follow the proof of Proposition 2.8 in this more general setting.

Fix $t > 0$. We construct a probability-preserving injection from paths of the random walk resulting in w' to paths resulting in w . Let M be the wall determined by (w, w') and L_M be the reflection in M . Let T be the first M -refresh time.

For a path ending at w' at time t , we have $T < t$. Modify such a path by changing the outcome of the coin flip attached to the refresh ring at time T , and proceed with the rest of the path by leaving the remaining sequence of refresh rings and coin flips unchanged. This has the effect of applying the reflection L_M to each state after time T , so paths that ended at w' now end at w .

This map is a probability-preserving bijection from the set of paths that end at w' at time t to the set of paths that end at w at time t and that have $T < t$. There is a positive probability that the walk ends at w without taking such a path, giving our strict inequality. $\square$

REMARK 2.10. Let (u_i) be any sequence of times and (a_i) be any sequence of generators. Note that Theorem 2.9 remains true if the random walk is conditioned to have R_{a_i} ring at time u_i for each i . The same holds if, in addition, we condition that R_s has no other rings for certain generators s up to time t . (Critically, R_{a_i} is the refresh clock of Definition 2.6, not the clock that always multiplies by the generator a_i .)

The proof is unchanged, because the constructed injection preserves the times at which each R_s rings.

We can express our argument in another way. The location Z_t is a function of the times and generators of the refresh rings that occur before time t and the results of the corresponding coin flips. Let a coin flip 0 represent “stay” and 1 represent “move.” For $s \in S$, let $\xi(s, 0)$ be the identity element and $\xi(s, 1) := s$. Given sequences $\mathbf{s} = (s_1, \dots, s_n) \in S^n$ and $\mathbf{b} = (b_1, \dots, b_n) \in \{0, 1\}^n$, let

$$\xi(\mathbf{s}, \mathbf{b}) := \xi(s_1, b_1)\xi(s_2, b_2) \cdots \xi(s_n, b_n) \in W.$$

Thus, if $\mathbf{s}$ gives the sequence of generators whose refresh rings occur before time t and $\mathbf{b}$ is the corresponding sequence of coin flips, we have $Z_t = \xi(\mathbf{s}, \mathbf{b})$. For any sequence $(a_1, \dots, a_n)$ and $0 \leq k \leq n$, write $(a_1, \dots, a_n)_k$ for the initial segment $(a_1, \dots, a_k)$. Our argument shows the following, which also easily implies Theorem 2.9.

THEOREM 2.11. *Let (W, S) be a Coxeter system. Let $\mathbf{s} = (s_1, \dots, s_n)$ be a finite sequence from S . Let $\mathbf{C} = (C_1, \dots, C_n)$ be independent, uniform $\{0, 1\}$ -valued random variables. If w and w' are two adjacent vertices with $|w| < |w'|$, then $\mathbf{P}[\xi(\mathbf{s}, \mathbf{C}) = w] \geq \mathbf{P}[\xi(\mathbf{s}, \mathbf{C}) = w']$.*

PROOF. Let M be the wall determined by (w, w') and L_M be the reflection in M . Given $\mathbf{b} = (b_1, \dots, b_n)$, define $\mathbf{b}' := (b_1, \dots, b_j, 1 - b_{j+1}, b_{j+2}, \dots, b_n)$ if there is some $k < n$ such that $\xi(\mathbf{s}_k, \mathbf{b}_k) \in M$ and j is the smallest such index, while $\mathbf{b}' := \mathbf{b}$ if there is no such k . Note that $(\mathbf{b}')' = \mathbf{b}$. Thus, $\mathbf{b} \mapsto \mathbf{b}'$ is a permutation of $\{0, 1\}^n$ with the property that $\xi(\mathbf{s}, \mathbf{b}') = L_M \xi(\mathbf{s}, \mathbf{b})$ iff $\mathbf{b}' \neq \mathbf{b}$. In particular, $\xi(\mathbf{s}, \mathbf{C}) = w'$ implies $\xi(\mathbf{s}, \mathbf{C}') = w$. Since $\mathbf{C}'$ has the same distribution as $\mathbf{C}$, the inequality follows. $\square$

We will now construct a Cayley graph where the conclusion of Theorem 2.9 fails.

EXAMPLE 2.12. Consider the cyclic group $\mathbb{Z}/8\mathbb{Z}$, with generators ± 1 , ± 2 , and ± 3 . Let the rates be chosen so that $r_1 = r_3$ are small and $r_2 = 1$. At time 2, there is a reasonable chance that the 2-clock has rung exactly twice and the -2 -clock has not rung, and the ± 1 -clocks and ± 3 -clocks are unlikely to have yet rung. Then the walker, started at 0, is more likely to be at 4 than at 3, even though 4 is distance 2 from 0, 3 is distance 1, and there is an edge between 3 and 4. Thus, Proposition 2.8 is not true in this setting.

The issue here is that while the vertices 3 and 4 are connected by an edge, and 4 is farther from 0 than 3 is, we have made this edge less likely than others, and there are paths from 0 to 4 that do not need such an unlikely edge. In the Coxeter-system setting, this cannot happen according to Theorem 2.9.

We may also modify this example so that the expected distance is not monotone in time; a discrete-time example of this phenomenon is due to Oded Schramm [25], Exercise 13.16(b).

EXAMPLE 2.13. Adjust Example 2.12 so that the group is $\mathbb{Z}/2n\mathbb{Z}$, with n at least 5. Take as generators ± 2 and every odd number, with ± 2 having high rates and the odd numbers having low rates.

After a time chosen so that the ± 2 -clocks have rung many times and the odd generator-clocks are unlikely to have rung, the walker is close to uniformly distributed on the even numbers with expected distance from 0 close to $2 - \frac{4}{n}$. After much more time, the walker will be close to uniformly distributed on all $2n$ states with expected distance close to $\frac{3}{2} - \frac{2}{n}$. Thus, the expected distance is not monotone.

This allows us to produce an example where the escape speed is not monotone in the rates.

EXAMPLE 2.14. Take the free product of Example 2.13 with $\mathbb{Z}/2\mathbb{Z}$, using the same generators as in that example, and one new generator a for the new factor of $\mathbb{Z}/2\mathbb{Z}$. Then the escape speed is not monotone in the rates $r_2 = r_{-2}$ and r_o associated, respectively, to ± 2 and the odd generators of the $\mathbb{Z}/2n\mathbb{Z}$ factor.

PROOF. Take the rate r_a to be 1, r_2 to be very large, r_o to be 0, and n to be large. Then the walker moves through copies of Example 2.13 at rate 1, backtracking only with rate proportional to $\frac{1}{n}$, and the average distance between entry and exit points in each copy of $\mathbb{Z}/2n\mathbb{Z}$ is $2 - \frac{4}{n} - o_{r_2}(1)$. Therefore, the escape speed is $3 - O(\frac{1}{n}) - o_{r_2}(1)$.

Now, increase the rate r_o to be as large as r_2 . The average distance between entry and exit points of each copy of $\mathbb{Z}/2n\mathbb{Z}$ is now only $\frac{3}{2} - \frac{2}{n} - o_{r_2}(1)$, so the escape speed has decreased to $\frac{5}{2} - O(\frac{1}{n}) - o_{r_2}(1)$. $\square$

REMARK 2.15. This example can be modified so that the escape speed is not monotone, either increasing or decreasing, in each rate separately, even for a rate on an involution.

The result of Example 2.14 should not necessarily be surprising; the graph distance is calculated using all edges, regardless of associated rates, so we shouldn't expect this necessarily to be a terribly meaningful quantity when it is affected by edges of very low rates. The following results start with trees, where this cannot occur, and we then show that Cayley graphs of Coxeter systems are symmetric enough that our results still apply. (While Examples 2.13 and 2.14 could equally well be implemented in Coxeter groups, replacing finite cyclic groups and $\mathbb{Z}$ with finite and infinite dihedral groups, the generators required would not be the Coxeter generators. Our results for Coxeter groups require that the generating set is composed of the Coxeter generators.)

We now return to Example 2.7.

PROPOSITION 2.16. *Let t_1 be an arbitrary time, and let $s \in S$ be one of the generators. Consider the following two random processes:*

1. *Run an instance Z^1 of the random walk of Example 2.7.*
2. *Run a separate instance Z^2 of the random walk of Example 2.7 until time t_1 , then trigger R_s at the (deterministic) time t_1 , and then continue to run the random walk.*

These two processes may be coupled so that at all times $t > t_1$, either they are in the same state or the second walker is farther from the origin than the first is.

PROOF. Given Z^1 , we will construct Z^2 via a special coupling to Z^1 . Let $A := \{w; |w| < |ws|\}$. Define γ to be the random vertex such that $Z_{t_1}^1 \in \{\gamma, \gamma s\}$ with $\gamma \in A$. Couple Z^1 and Z^2 up to time t_1 to be independent conditional that $Z_{t_1}^2 \in \{\gamma, \gamma s\}$, which is possible because Z^1 and Z^2 have the same law up to time t_1 . For all $w \in A$, Proposition 2.8 gives us that

$$(2.1) \quad \mathbf{P}[Z_{t_1}^1 = \gamma \mid \gamma = w] = \frac{\mathbf{P}[Z_{t_1}^1 = w]}{\mathbf{P}[Z_{t_1}^1 = w] + \mathbf{P}[Z_{t_1}^1 = ws]} > \frac{1}{2}.$$

We need $Z_{t_1^+}^2$ to be uniformly distributed in $\{\gamma, \gamma s\}$ independently of $(Z_t^2)_{0 \leq t \leq t_1}$ conditional on γ . We may ensure this by letting $Z_{t_1^+}^2$ equal γs whenever $Z_{t_1}^1 = \gamma s$, and sometimes even when $Z_{t_1}^1 = \gamma$; this is possible in light of (2.1) and maintains the required independence by the first step in our construction of Z^2 . We have now achieved that $|Z_{t_1}^1| \leq |Z_{t_1^+}^2|$.

In case $|Z_{t_1}^1| = |Z_{t_1^+}^2|$ (which has probability $2\mathbf{P}[Z_{t_1}^1 = \gamma s]$), we couple the two processes so that they stay together at all times after t_1 . Otherwise, we couple them after time t_1 so that the second is the first reflected by $L_{\gamma s \gamma^{-1}}$ (which is true at time t_1^+), until and unless they would move across the edge between γ and γs , in which case we couple the R_s -coin flips so that the walks agree from then on.

For any time $t > t_1$, this results in either (1) the two processes being at the same vertex at time t or (2) the first being at a vertex v that is closer to γ than to γs and the second being at $L_{\gamma s \gamma^{-1}}(v)$. In this case v is closer to the initial state than $L_{\gamma s \gamma^{-1}}(v)$ is by Lemma 2.1. $\square$

REMARK 2.17. As with Remark 2.10, let (u_i) be any sequence of times in $[0, \infty) \setminus \{t_1\}$ and (a_i) be any sequence of generators. Proposition 2.16 remains true if the processes Z^1 and Z^2 are each conditioned to have R_{a_i} ring at time u_i for each i . The same holds if, in addition, we condition that R_s has no other rings up to any given time t_2 for the process Z^1 .

The proof is unchanged, except for applying Remark 2.10 when we invoke Proposition 2.8.

COROLLARY 2.18. *The escape speed in Example 2.7 is nondecreasing in the rates r_a , r_b , and r_c .*

PROOF. By symmetry it suffices to show that the escape speed is nondecreasing in r_a . It suffices for this to show that $\mathbf{E}[|Z_t|]$ is nondecreasing in r_a for every time t . But increasing the rate r_a just results in extra instances of R_a , so this is a consequence of Proposition 2.16 as follows.

Let (X_t) be our random walk and (Y_t) be the random walk where the rate r_a has been increased. We may couple X and Y so that they have the same R_a , R_b , and R_c rings at the same times, except that Y has some random number N_t of additional rings of R_a up to time t . For each i between 0 and N_t , let X^i be the random walk which agrees with X , except that it has the first i of these additional rings of R_a , so that $X^0 = X$ and $X^{N_t} = Y$.

For each i we may use Proposition 2.16 to couple X^i and X^{i+1} so that at time t the walk X^{i+1} is at least as far from the origin as X^i is. Because X^{i+1} has R_a -rings at the times that X^i does, this application of Proposition 2.16 requires the observation of Remark 2.17; we are also conditioning on the time of the extra R_a -ring. Combining these couplings gives a coupling between $X = X^0$ and $X^N = Y$ where at time t the walk Y is at least as far from the origin as X is. $\square$

As with Proposition 2.8, these results also apply to walks on the Cayley graphs of Coxeter systems.

COROLLARY 2.19. *The result of Proposition 2.16 is true for the Cayley graph of a Coxeter system, as is Corollary 2.18.*

PROOF. This result follows from Theorem 2.9 in the same way as Proposition 2.16 and Corollary 2.18 follow from Proposition 2.8, with a single change: in the proof of Proposition 2.16, the phrase

“until and unless they would move across the edge between γ and γs , in which case we couple the R_s -coin flips ...”

should be replaced by

“until and unless they would move across an edge (x, xb) in the wall determined by $(\gamma, \gamma s)$, in which case we couple the R_b -coin flips ...” $\square$

REMARK 2.20. A superficially more general result is that Corollary 2.19 holds for Cayley graphs of presentations

$$\Gamma := \langle S_1, S_2 \mid \forall s_1 \in S_1 \ \forall s \in S_1 \cup S_2 \ (s_1 s)^{n(s_1, s)}, \forall s_2 \in S_2 \ s_2^{2n(s_2)} \rangle,$$

where each $n(s_1, s)$ and $n(s_2)$ is either a positive integer or ∞ , each $n(s_1, s_1)$ is equal to 1, and $n(s_1, s) \geq 2$ for $s_1 \neq s$. To see why Corollary 2.19 holds for such Γ , let $\bar{S}_2$ be a copy of S_2 , and write $s_2 \mapsto \bar{s}_2$ for a bijection from $S_2 \rightarrow \bar{S}_2$. Consider the Coxeter system $W := \langle S_1, S_2, \bar{S}_2 \mid \forall s, s' \in S_1 \cup S_2 \cup \bar{S}_2 \ (ss')^{m(s, s')} \rangle$ with $m(s_1, s_2) := n(s_1, s_2) =: m(s_1, \bar{s}_2)$ when $s_1 \in S_1$ and $s_2 \in S_2$, $m(s_1, s'_1) := n(s_1, s'_1)$ when $s_1, s'_1 \in S_1$, and, finally, $m(s_2, s_2) = m(\bar{s}_2, \bar{s}_2) = 1$ and $m(s_2, \bar{s}_2) := n(s_2)$ when $s_2 \in S_2$. The Cayley graphs of Γ and W are the same, but to get the unoriented Cayley diagram of Γ from that of W , replace each label $\bar{s}_2$ with the label s_2 . As long as the generator rates for a random walk on W have the property that the rate for $\bar{s}_2$ is the same as the rate for s_2 , then the random walks on the two diagrams have the same law. The simplest case is when $S_1 = \emptyset$ and $n \equiv \infty$ on S_2 , in which case Γ is a free group with free generators.

Similarly to Theorem 2.11, we may deduce from our arguments the following, which also easily implies Corollary 2.19.

THEOREM 2.21. *Let (W, S) be a Coxeter system. Let $\mathbf{s}$ and $\mathbf{s}'$ be finite sequences from S of lengths n and n' , respectively, with $\mathbf{s}$ a proper subsequence of $\mathbf{s}'$. Let $\mathbf{C}$ be a Bernoulli(1/2) process. Then $\mathbf{E}[|\xi(\mathbf{s}', \mathbf{C}_{n'})|] \geq \mathbf{E}[|\xi(\mathbf{s}, \mathbf{C}_n)|]$.*

2.3. Strict monotonicity of speed. Corollaries 2.18 and 2.19 show that the escape speed is an increasing function of the rates r_s , for each generator s . We now consider the question of whether this function is *strictly* increasing. First, we give an example where this is not the case.

EXAMPLE 2.22. Consider the Coxeter system

$$\langle a, b, c \mid a^2, b^2, c^2 \rangle \times \langle d \mid d^2 \rangle.$$

This group is the direct product of a free Coxeter group, whose Cayley graph is the 3-regular tree, and the two-element group. Increasing the rate r_d does not change the escape speed of the random walk. As long as the three rates r_a , r_b , and r_c are all positive, increasing any of them does increase the escape speed, but the escape speed is zero if one of those three rates is zero.

More generally, we have the following behaviour in product groups.

PROPOSITION 2.23. Suppose that Γ_i is generated by S_i and has identity element o_i for $i = 1, 2$, and let $\mathbf{r}_i: S_i \rightarrow [0, \infty)$. Then for $\Gamma := \Gamma_1 \times \Gamma_2$ generated by $(S_1 \times \{o_2\}) \cup (\{o_1\} \times S_2)$ with rates $\mathbf{r}_1 \cup \mathbf{r}_2$, the escape speed in Γ is the sum of the escape speeds in Γ_i .

PROOF. We have $|(v_1, v_2)| = |v_1| + |v_2|$ for $v_i \in \Gamma_i$. In addition, if we write the random walk as $Z_t = (Z_t^{(1)}, Z_t^{(2)})$, then $Z_t^{(i)}$ are independent random walks in Γ_i with rates $\mathbf{r}_i$. $\square$

The *Coxeter diagram* of a Coxeter system (W, S) is the unoriented graph with vertices S , edges (s, s') for $m(s, s') \geq 3$, and edge labels $m(s, s')$ when $m(s, s') \geq 4$. A Coxeter system is called *irreducible* if its Coxeter diagram is connected. Every Coxeter system can be expressed as a direct product of irreducible Coxeter systems, namely, the subgroups generated by the vertices in the connected components of its Coxeter diagram. An infinite, irreducible Coxeter group is either *affine*, being a finite extension of an abelian group, or *nonelementary hyperbolic*, being nonelementary word hyperbolic; see [9], Chapter 12, especially Section 12.6, and [19], Sections 6.8 and 6.9, for more detailed information on hyperbolic Coxeter groups.

We quickly review the basic definitions for hyperbolic metric spaces. A metric space X in which every pair of points can be joined by a geodesic path is called δ -*hyperbolic* if, for each geodesic triangle in X , each side lies within the closed δ -neighbourhood of the union of the other two sides. If X is δ -hyperbolic for some $\delta \geq 0$, then X is called *hyperbolic*. A group is *word hyperbolic* if some (hence every) Cayley graph is hyperbolic in the shortest-path metric. A word hyperbolic group is called *elementary* if it is finite or virtually cyclic.

REMARK 2.24. In a finite or affine Coxeter group, the escape speed is zero regardless of the rates of each generator, so increasing the rate of a generator does not increase the escape speed. Indeed, the speed is zero of every nearest-neighbor, reversible, discrete-time random walk on a graph of subexponential volume growth and bounded edge weights, a consequence of the Varopoulos–Carne inequality [25], Theorem 13.4.

Thus, the same behaviour as in Example 2.22 appears for any direct product of a nonelementary hyperbolic Coxeter group with a finite or affine Coxeter group.

We will now prove that in an irreducible, nonelementary hyperbolic Coxeter system, increasing the rate of any generator strictly increases the escape speed. With the preceding remark, this will characterise those generators where increasing the rate must strictly increase the escape speed.

THEOREM 2.25. Let W be a Coxeter system. Write W as a direct product of irreducible Coxeter systems. If the rate corresponding to each generator is positive, then increasing a rate strictly increases the escape speed if and only if that generator belongs to a nonelementary hyperbolic irreducible factor, rather than a finite or affine factor.

To prove this, it suffices to examine random walks on irreducible, nonelementary hyperbolic Coxeter systems because of Proposition 2.23 and Remark 2.24.

We will need the following facts about the boundary of a Coxeter group. We begin by reviewing the definition of Gromov boundary and some of its basic properties. Recall that a graph is equipped with its shortest-path metric, dist . A *geodesic ray* in a graph G is a semiinfinite path all of whose finite subpaths are geodesic. Two geodesic rays $(x_n)_n$ and $(y_n)_n$ are *asymptotic* if $\sup_n \text{dist}(x_n, y_n) < \infty$. This is an equivalence relation, whose equivalence classes form the *boundary* ∂G of G . If $\xi \in \partial G$ and $x \in V(G)$, a geodesic ray starting from x that belongs to ξ will also be referred to as a geodesic from x to ξ . We define a topology on $V(G) \cup \partial G$ as follows. Fix a vertex o of G . Given $x_k, y \in V(G) \cup \partial G$ for $k \in \mathbb{N}$, we say that $\lim_{k \rightarrow \infty} x_k = y$ if for some $n_\infty, n_k \in \mathbb{N} \cup \{\infty\}$ ($k \in \mathbb{N}$), there are geodesics $(x_{k,n})_{n < n_k}$ from o to x_k and a geodesic $(y_n)_{n < n_\infty}$ from o to y such that, for every $n < n_\infty$, $x_{k,n} = y_n$ for all sufficiently large k . The closed sets are then those that are closed under sequential convergence. In case G is hyperbolic, ∂G is usually referred to as its *Gromov boundary*.

Consider now a Cayley graph with associated continuous-time Markov chain $(Z_t)_t$. The embedded discrete-time random walk records only the changes made by $(Z_t)_t$. This random walk has transition probability from w to ws equal to $r_s / \sum_{a \in S} r_a$, where S is the generating set. We can recover the law of $(Z_t)_t$ from its embedded discrete-time random walk by jumping at the times of a Poisson process of rate $\sum_{a \in S} r_a$. Many results for $(Z_t)_t$ follow from their corresponding results for the embedded random walk, such as this next one.

LEMMA 2.26 (Theorem 7.4 of [20]). *Random walk on a nonelementary word-hyperbolic group converges a.s. to a boundary point, provided all rates are positive.*

The law on ∂G of $\lim_{t \rightarrow \infty} Z_t$ is called *harmonic measure*, which depends on the starting point, Z_0 .

The following result is due to Gromov [16].

LEMMA 2.27 (Lemma 2.4.4 of [6]). *The action of a nonelementary word-hyperbolic group on its Gromov boundary is minimal, that is, every orbit is dense.*

LEMMA 2.28. *The harmonic measure of random walk on a nonelementary word-hyperbolic group has full support in its Gromov boundary.*

PROOF. Suppose that ξ is a point of the support when the walk starts at o . Then $\gamma\xi$ is a point of the support when the walk starts at γ . Since the random walk starting at o has positive chance to reach γ , it follows that $\gamma\xi$ is also in the support when the random walk starts at o . Thus, the result follows from Lemmas 2.26 and 2.27. $\square$

We will use the following result of [32].

LEMMA 2.29. *The conjugacy class of a reflection in an infinite, irreducible Coxeter system (W, S) is always infinite.*

PROOF. Let S be ordered as $(s_1, s_2, \dots, s_r)$, where $r := |S|$. Speyer [33] showed that $s_1 s_2 \cdots s_r s_1 s_2 \cdots s_r \cdots s_1 s_2 \cdots s_r$ is a reduced word for any number of repetitions of $s_1 s_2 \cdots s_r$. Write $w := s_1 s_2 \cdots s_r \in W$. Combining Speyer's result with Lemma 2.3, we deduce that the walls corresponding to the edges $(w^k, w^k s_1)$ are all distinct, whence the elements $w^k s_1 w^{-k}$ are all distinct. Since the ordering of S was arbitrary, every generator has infinitely many conjugates. Therefore, so does every conjugate of a generator, that is, every reflection. $\square$

LEMMA 2.30. *Let M be a wall in an irreducible Coxeter system with Cayley graph G . Then the closure $\overline{M}$ does not include all of ∂G .*

PROOF. Let $(v, vs) \in M$. By Lemma 2.29 there exist $w_k \in W$ such that $|w_k^{-1}vs v^{-1}w_k| \rightarrow \infty$ as $k \rightarrow \infty$. By Lemma 2.1, for every $w \in W$,

$$|w^{-1}vs v^{-1}w| = \text{dist}(w, L_{vs v^{-1}}w) = 1 + 2 \text{dist}(w, V(M)).$$

Therefore, $\text{dist}(w_k, V(M)) \rightarrow \infty$ as $k \rightarrow \infty$. Let $v_k \in V(M)$ satisfy $\text{dist}(w_k, v_k) = \text{dist}(w_k, V(M))$. By taking a subsequence if necessary, we may assume that there is some fixed $a \in S$ such that $(v_k, v_k a) \in E(M)$ for all k . If we translate the geodesic that goes from v_k to w_k so that it starts at v_1 , then as $k \rightarrow \infty$, it has a limit geodesic ray $(y_n)_n$, up to taking another subsequence if necessary. By Lemma 2.5 M is fixed under this translation, so $(y_n)_n$ is a geodesic ray from v_1 with $\text{dist}(y_n, V(M)) = \text{dist}(y_n, v_1) = n$. By Lemma 2.4 it follows that the equivalence class of $(y_n)_n$ does not belong to $\overline{M}$, as desired. $\square$

LEMMA 2.31. *Let M be a wall in an irreducible, nonelementary hyperbolic Coxeter system with Cayley graph G . Provided that all rates are positive, there is some $v \in V(M)$ such that the probability that random walk started at v never returns to M is positive.*

PROOF. If not, then by Lemma 2.5 and the strong Markov property, the number of visits to M would be infinite a.s., given that it starts on M . In other words, if the random walk starts on M , then it has a limit in $\overline{M} \cap \partial G$ a.s. by Lemma 2.26. But this contradicts Lemmas 2.28 and 2.30. $\square$

Given a Coxeter system (W, S) and rates $\mathbf{r}$, let $\eta(W, S; \mathbf{r})$ be the minimum over all walls M of the probability that there is no positive M -refresh time for the random walk with rates $\mathbf{r}$.

LEMMA 2.32. *Let (W, S) be an irreducible, nonelementary hyperbolic Coxeter system and $\mathbf{r}$ have all positive rates. Then $\eta(W, S; \mathbf{r}) > 0$.*

PROOF. Given any wall, M , there is an automorphism of the Cayley diagram that takes M to a wall containing o . By Lemmas 2.31 and 2.5, it follows that there are some $d < \infty$ and $c > 0$ such that, for all walls M and all vertices $w \in M$, there is a path of length at most d from w to some $v \notin M$ from which the probability that the random walk with rates $\mathbf{r}$ never visits M is at least c .

Given any wall, M , let w be the first vertex in M , if any, visited by our random walk. Consider a fixed path $(w = x_0, x_1, \dots, x_\ell = v)$ of length $\ell \leq d$ from w to some $v \notin M$ from which the probability that the random walk with rates $\mathbf{r}$ never visits M is at least c . By reflecting v in M if needed, we may assume that w and v are on the same side of M and hence that our fixed path does not cross M . The chance that the next ℓ steps of our random walk are exactly along this path and that there is no M -refresh time before visiting v is at least the chance that $R_{x_{i-1}x_i}^{-1}$ ($1 \leq i \leq \ell$) are the next ℓ refresh rings and that the first ℓ coin flips all tell the random walk to move, which equals

$$2^{-\ell} \prod_{i=1}^{\ell} \frac{r_{x_{i-1}x_i}^{-1}}{\sum_{s \in S} r_s}.$$

It follows that

$$\eta(W, S; \mathbf{r}) \geq \left(\frac{\min_{s \in S} r_s}{2 \sum_{s \in S} r_s} \right)^d c > 0. \quad \square$$

Gouëzel [15] showed that $\mathbf{r} \mapsto \sigma(\mathbf{r})$ is analytic for arbitrary generators of nonelementary word-hyperbolic groups, provided all rates in $\mathbf{r}$ are strictly positive.

THEOREM 2.33. *If (W, S) is an irreducible, nonelementary hyperbolic Coxeter system, then increasing the rate r_s of any generator s strictly increases the escape speed, provided all rates are positive. Moreover, $\partial\sigma(\mathbf{r})/\partial r_s \geq 2\eta(W, S; \mathbf{r})^2$.*

PROOF. In order to understand the consequences of increasing the rate of any generator, consider the construction in the proof of Corollary 2.19, taken from the proof of Corollary 2.18. In that construction we produce a sequence X^i of random walk paths, each successive path having one additional ring of the refresh clock, R_s . The key property of that construction is that, for each i , the paths X^i and X^{i+1} either eventually agree or have $|X^{i+1}| > |X^i|$ for all large times. To show that increasing a rate strictly increases the escape speed, it suffices to show that, asymptotically, a positive fraction of i fall into the latter case. In our proof of this, we will be very careful to avoid dealing with dependencies among these events for different i as well as with the dependencies on the number of such additional rings up to time t .

To be more precise, consider two sets of positive rates, $\mathbf{r}$ and $\mathbf{r}'$, where $\mathbf{r}'$ agrees with $\mathbf{r}$, except that $r'_s = r_s + \epsilon$ where $\epsilon > 0$. Let $(N_t)_t$ be a Poisson process of rate 2ϵ , which we use for the extra rings of R_s . The random walk corresponding to $\mathbf{r}$ is $X = X^0$, and the random walk corresponding to $\mathbf{r}$ plus i additional rings of R_s is X^i . We couple all X^i , as in the proof of Corollary 2.18. We have

$$\sigma(\mathbf{r}') - \sigma(\mathbf{r}) = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=0}^{N_t-1} (|X_t^{i+1}| - |X_t^i|) \quad \text{a.s.}$$

Let A_i be the event that $|X^{i+1}| > |X^i|$ for all large times. If $i < N_t$, then we have $|X_t^{i+1}| \geq |X_t^i|$, with equality only if A_i does not occur. Thus, $|X_t^{i+1}| - |X_t^i| \geq \mathbf{1}_{A_i}$, whence

$$\begin{aligned} \sigma(\mathbf{r}') - \sigma(\mathbf{r}) &\geq \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{i=0}^{N_t-1} \mathbf{1}_{A_i} \\ &= \lim_{t \rightarrow \infty} \frac{N_t}{t} \cdot \limsup_{t \rightarrow \infty} \frac{1}{N_t} \sum_{i=0}^{N_t-1} \mathbf{1}_{A_i} \\ &= 2\epsilon \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{A_i} \quad \text{a.s.} \end{aligned}$$

Taking expectation and using Fatou's lemma yields

$$\sigma(\mathbf{r}') - \sigma(\mathbf{r}) \geq 2\epsilon \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{P}(A_i).$$

Let T_i be the time of the i th ring of $(N_t)_t$, and let M_i be the wall containing the edge $(X_{T_i}^i, X_{T_i}^i s)$. Recall that a time t is an M_i -refresh time for X^i if $X_t^i \in M_i$ and the refresh clock R_a rings at time t , where a is the generator such that $(X_t^i, X_t^i a) \in M_i$. The event A_i occurs if (but not only if): (1) there is no M_i -refresh time $t < T_i$ for X^i , and (2) there is no M_i -refresh time $t > T_i$ for X^i . Here (1) guarantees that (3) $X_{T_i^+}^i \neq X_{T_i^+}^{i+1}$, while (2) guarantees that if (3) holds, then $X_t^i \neq X_t^{i+1}$ for all $t > T_i$. Note that the law of $(X_t^i)_{0 \leq t \leq T_i}$ is the same as the law of $((X_{T_i}^i)^{-1} X_{T_i-t}^i)_{0 \leq t \leq T_i}$ by reversibility. In addition, the reflection in M_i maps a path on one side of the wall to a path on the other side, preserving the times on the wall. Therefore, the probability of (1) is at least the probability of (2). The events (1) and (2) are also independent. Hence, $\mathbf{P}(A_i) \geq \eta(W, S; \mathbf{r})^2$, which yields the desired inequality $\sigma(\mathbf{r}') - \sigma(\mathbf{r}) \geq 2\epsilon \eta(W, S; \mathbf{r})^2 > 0$ in light of Lemma 2.32. $\square$

Combining Proposition 2.23 with Remark 2.24 and Theorem 2.33 gives Theorem 2.25.

REMARK 2.34. The assumption in Theorem 2.25 that each rate is positive is necessary; if enough of the rates are zero, then the escape speed may be zero, even when one of the rates is increased. In a scenario where some of the rates are zero, one should just remove those generators and apply Theorem 2.25 to the resulting smaller Coxeter system. Of course, if a rate is zero, increasing that rate will strictly increase the speed if the same holds when that rate is positive.

REMARK 2.35. The first part of Theorem 2.33 can be proved in the case of trees by using an implicit formula for the speed: Namely, [31] gives a formula for the speed of the embedded discrete-time random walk. Multiplying that speed by the sum of the rates gives the following formula for the continuous-time speed. For a free product of $p \geq 3$ copies of $\mathbb{Z}/2\mathbb{Z}$ and corresponding rates $r_1, \dots, r_p > 0$, write ζ for the unique positive solution to the equation $\sum_{i=1}^p (\sqrt{\zeta^2 + r_i^2} - r_i) = (p-2)\zeta$. (Existence and uniqueness follow from the fact that the left-hand side is a convex function of ζ that passes through $(0, 0)$ with derivative 0 and is asymptotic to $p\zeta$ as $\zeta \rightarrow \infty$.) The continuous-time speed $\sigma(\mathbf{r})$ is then $\zeta^{-1} \sum_{i=1}^p r_i (\sqrt{\zeta^2 + r_i^2} - r_i)$, which can also be written as $\sum_{i=1}^p \frac{r_i \zeta}{\sqrt{\zeta^2 + r_i^2}}$. Because $r_j \mapsto \sqrt{x^2 + r_j^2} - r_j$ is strictly decreasing for fixed $x > 0$, it follows that $\partial \zeta / \partial r_j > 0$. Now,

$$\frac{\partial \sigma(\mathbf{r})}{\partial r_j} = \frac{\partial \zeta}{\partial r_j} \frac{1}{\zeta} \left(-\sigma(\mathbf{r}) + \sum_{i=1}^p \frac{r_i \zeta}{\sqrt{\zeta^2 + r_i^2}} \right) + \frac{r_j}{\zeta} \left(\sqrt{(\zeta/r_j)^2 + 1} - 2 + \frac{1}{\sqrt{(\zeta/r_j)^2 + 1}} \right).$$

From the second formula for $\sigma(\mathbf{r})$, it is clear that the first term in parentheses is strictly positive; it is elementary that the second term in parentheses is also strictly positive. Thus, $\sigma(\mathbf{r})$ is indeed strictly increasing in each rate.

2.4. *Further questions.* The proof of Proposition 2.16 gives rise to the following question.

QUESTION 2.36. Given two random walks, with one having higher rates than the other, is it possible to couple them so that the more active walk is always at least as far from the initial position as the other?

It might be plausible that a slightly cleverer construction in the proof of Proposition 2.16 could produce such a coupling. Surprisingly, this is not possible in even the simplest of cases.

PROPOSITION 2.37. *Question 2.36 is impossible in the setting of Example 2.7, even in cases where $r_b = r_c = 0$. In such cases the walker is moving back and forth between just two states, at the rate r_a . Similarly, it is not always possible for $r_a, r_b, r_c > 0$.*

PROOF. Consider two such walks, one with $r_a = 1$ and the other with $r_a = 100$, and the behaviour of these over the first unit of time. The first process has a significant chance of, for instance, staying at the initial vertex until a time between 0.4 and 0.6, moving to the other vertex, and staying there until time 1. The more active process is very unlikely to take a path that remains farther from the starting vertex than such a path, because that would require staying at the other vertex between times 0.6 and 1, which is unlikely for such an active process.

If $r_b, r_c > 0$, then one can choose a sufficiently small t and make two choices of r_a , the smaller one being equal to $1/t$, so that a similar analysis holds. $\square$

PROPOSITION 2.38. *The coupling of Question 2.36 is not possible for two walks on a recurrent graph, one of which has edges all with rate ρ_1 and the other all with rate $\rho_2 > \rho_1 > 0$.*

PROOF. Let both random walks start at o . If such a coupling were possible, then we could similarly couple random walks with rates $(c\rho_1, c\rho_2)$ for any $c > 0$. In particular, we could couple with rates $(\rho_1(\rho_2/\rho_1)^n, \rho_2(\rho_2/\rho_1)^n)$ for every nonnegative integer n . Combining such couplings, we could couple with rates $(\rho_1, c\rho_2)$ for arbitrarily large c . By recurrence the times between visits of the random walk with rate ρ_2 to the vertex o have a distribution that is dominated by a geometric sum of exponential random variables with rate ρ_2 , that is, by an exponential random variable, which means that given $\epsilon > 0$, for all sufficiently large c the probability is at least $1 - \epsilon$ that o is visited in every interval of length ϵ in $[0, 1]$ by the random walk with rate $c\rho_2$. By the coupling the same is true for the random walk with rate ρ_1 , which forces that random walk to be at o a.s. during the entire interval $[0, 1]$, a contradiction. $\square$

Question 2.36 appears to be difficult to decide even in other simple settings.

QUESTION 2.39. On which transient Cayley graphs is the coupling of Question 2.36 possible for two walks, one of which has edges all with rate ρ_1 and the other all with rate $\rho_2 > \rho_1 > 0$?

It is even possible to ask a version of this question for a chain with only three states.

QUESTION 2.40. Is the coupling of Question 2.36 possible for two walks on the set $\{0, 1, 2\}$ that both start at 0, where walk i moves from 0 to 1, 1 to 0, and 1 to 2 with rate ρ_i ? Note that once the walker reaches the state 2, it stays there.

Neither Question 2.39 nor 2.40 can be answered by a Markovian coupling, because it would be possible for both walks to be at a neighbor of the starting vertex and then the more active walk has a greater chance to move back to the starting vertex than the less active walk does.

Finally, here are two more questions whose answers are unknown to us.

QUESTION 2.41. Is the escape speed concave in the rates for Coxeter systems?

QUESTION 2.42. Suppose that we have two random fields of rates on a Coxeter system with the first being at most the second a.s. Suppose that the law of each field is invariant under left group multiplication. Is the expected speed of the first at most that of the second?

3. Approach to stationarity. We will now consider random walks on finite groups and how the distance of the law of Z_t to the stationary distribution changes when the random walk rates are increased.

QUESTION 3.1. Given a continuous-time random walk $(Z_t)_t$ on a finite group with positive rates on the generators and a time t , consider the distance between the distribution of Z_t and the stationary distribution. Does the distance necessarily decrease when a random walk rate is increased?

It seems natural to think that the distance would decrease, as increasing a rate might be thought of as injecting additional randomness into the walk. In this section we will show

that the answer is “yes” in several families of nice examples and then give several examples where the answer is “no” in general, including some demonstrating rather unusual behaviour. Furthermore, there appear to be interesting patterns in the sorts of examples that we are able to produce.

There are several different measures of the distance from stationarity. Of course, the stationary distribution is uniform.

DEFINITION 3.2. For $p \in [1, \infty)$, the ℓ^p -distance between distributions f and g is

$$\ell^p(f, g) := \left(\sum_v |f(v) - g(v)|^p \right)^{\frac{1}{p}},$$

where the sum is taken over the states v of the chain. The ℓ^∞ -distance between distributions f and g is $\ell^\infty(f, g) := \max_v |f(v) - g(v)|$. The *Hellinger distance* between distributions f and g is

$$\left(\frac{1}{2} \sum_v |\sqrt{f(v)} - \sqrt{g(v)}|^2 \right)^{\frac{1}{2}}.$$

DEFINITION 3.3. The *entropy* of a distribution f is the sum over states v of $-f(v) \times \log f(v)$.

Observe that because the function $x \log x$ is convex on $(0, 1)$, the distribution of maximum entropy is the uniform distribution. Thus in this context, Question 3.1 asks whether increasing one of the rates causes the entropy of Z_t to increase.

One particularly strong way in which Question 3.1 may be answered in the affirmative would be if increasing one of the rates results in decreasing the vector of transition probabilities in the majorization order.

DEFINITION 3.4. Let f and g be vectors of length n . Denote by $f_{[1]} \geq f_{[2]} \geq \cdots \geq f_{[n]}$ the decreasing rearrangement of f , and similarly for $g_{[i]}$. We say that f *majorizes* g if, for each i ,

$$f_{[1]} + \cdots + f_{[i]} \geq g_{[1]} + \cdots + g_{[i]}$$

with equality when $i = n$ (which is automatic if f and g are probability vectors). Majorization defines a partial order on probability vectors, with the largest elements being the extreme points such as $(1, 0, 0, \dots, 0)$ and the smallest element being $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$.

Recall the inequality of Hardy–Littlewood–Pólya–Karamata that if f majorizes g and ϕ is convex, then $\sum_{i=1}^n \phi(f_i) \geq \sum_{i=1}^n \phi(g_i)$; if ϕ is strictly convex, then equality holds only if $f = g$. For example, if f and g are probability vectors and $\phi(x) = |x - 1/n|^p$ for some $p \geq 1$, then we may conclude that the ℓ^p -distance between f and the uniform distribution is at least the ℓ^p -distance between g and the uniform distribution. By letting $p \rightarrow \infty$, we get the same for the ℓ^∞ -distance. A similar inequality holds for Hellinger distance and, with opposite sign, for entropy. Recall also that f majorizes g iff g is gotten by applying a doubly stochastic matrix to f , such as by replacing (f_i, f_j) by $(\alpha f_i + (1 - \alpha)f_j, (1 - \alpha)f_i + \alpha f_j)$, where $\alpha \in [0, 1]$. This is equivalent to moving some part of the larger of f_i and f_j to the smaller of the two. More generally, doubly stochastic matrices are convex combinations of permutation matrices.

In particular, for continuous-time random walk on any finite graph, the distribution of Z_t decreases in the majorization partial order as t increases. In some cases increasing one of the rates causes the distribution Z_t to decrease in the majorization partial order, which implies that the ℓ^p -distance to uniform has decreased for each $p \geq 1$ and that the entropy has increased.

3.1. *Positive examples.* There are several special cases where the answer to Question 3.1 is positive.

3.1.1. *Coxeter systems.* Throughout this subsection we work in a Coxeter system, (W, S) . We will show that increasing the generator rates has the effect of decreasing in the majorization order the probability distribution of Z_t at any fixed time t , which, therefore, implies decreasing distance to the uniform distribution in all senses mentioned. In order to prove this, we will use the *Bruhat order*, the partial order $\leq$ on W defined by taking the transitive and reflexive closure of the relation

$$\{(v, Lv); L \text{ is a reflection and } |v| < |Lv|\}.$$

We first extend Theorem 2.9 to the Bruhat order.

THEOREM 3.5. *Let $x < y$ in the Bruhat order and $t > 0$. Then $p_t(o, x) > p_t(o, y)$.*

PROOF. It suffices to prove this when y is a reflection L of x in a wall M . Let T be the first M -refresh time of the random walk. For every path with $Z_t = y$, we have $T < t$ in light of Lemma 2.1. If we change the coin flip at time T , then the walk will end instead at $Ly = x$. This defines a probability-preserving bijection of the set of paths and coin flips (on $[0, t]$) from o to y to the set of paths and coin flips from o to x that have $T < t$. Since there is a positive-probability set of paths from o to x during which there is no M -refresh time, we get the desired strict inequality. $\square$

We now look at how partial orders relate to majorization, based on [5].

DEFINITION 3.6. Let $\preceq$ denote a partial order on a set, A . A function $f: A \rightarrow \mathbb{R}$ is called *decreasing* if $f(x) \geq f(y)$ whenever $x \preceq y$. A subset $B \subseteq A$ is called *decreasing* if its indicator $\mathbf{1}_B$ is decreasing. If f and g are two functions on A , then $f \preceq$ -majorizes g if f and g are decreasing and $\sum_{x \in B} f(x) \geq \sum_{x \in B} g(x)$ for all decreasing sets B , with equality when $B = A$.

LEMMA 3.7. *Let $\preceq$ denote a partial order on a finite set, A . If $f \preceq$ -majorizes g , then f majorizes g .*

PROOF. For $i < |A|$, there is a decreasing set B of cardinality i on which g attains its i largest values. The sum of the i largest values of f is at least the sum of f over B , which, by definition, is at least the sum of g over B . $\square$

COROLLARY 3.8. *Let $\mathbf{r}$ and $\mathbf{r}'$ be two sets of rates on S and t and t' be two positive times. Let Z and Z' be the corresponding random walks on W . Suppose there is a coupling of Z_t and $Z'_{t'}$ such that $Z_t \leq Z'_{t'}$ a.s. in the Bruhat order. Then the law of Z_t majorizes the law of $Z'_{t'}$.*

PROOF. Let f be the law of Z_t and g be the law of $Z'_{t'}$. By Theorem 3.5 both f and g are decreasing. If B is a decreasing subset of W , then $\mathbf{P}[Z_t \in B] \geq \mathbf{P}[Z'_{t'} \in B]$ because of the coupling, whence $f \preceq$ -majorizes g . Thus, the result follows from Lemma 3.7. $\square$

THEOREM 3.9. *Let $\mathbf{r}$ and $\mathbf{r}'$ be two sets of rates on S with $r_s \leq r'_s$ for all $s \in S$. Let $t > 0$. Denote the corresponding transition probabilities by $p_t(x, y; \mathbf{r})$ and $p_t(x, y; \mathbf{r}')$. If W is finite, then $p_t(o, \cdot; \mathbf{r})$ majorizes $p_t(o, \cdot; \mathbf{r}')$ with inequality if $\mathbf{r} \neq \mathbf{r}'$.*

PROOF. Suppose that $\mathbf{r} = \mathbf{r}'$, except for $r_s < r'_s$ for one specific s . Consider how many extra rings in $[0, t)$ there are of R_s with the rates $\mathbf{r}'$. If there is only one, then the argument of Corollary 2.19 showed, with the same notation, that we may couple Z_t^1 and Z_t^2 so that $Z_t^1 \leq Z_t^2$ in the Bruhat order, and inequality holds with positive probability. By induction this same conclusion extends to any finite set of additional rings of R_s . Therefore, we may couple the two random walks so that $Z_t^1 \leq Z_t^2$ conditional on the extra rings of R_s . Hence, we may couple them so that this inequality holds without such conditioning. By Corollary 3.8 we deduce that the distribution of Z_t^1 majorizes that of Z_t^2 . For the general case of rates, we may change the rates one by one until $\mathbf{r}$ becomes $\mathbf{r}'$, still yielding majorization. $\square$

Of course, there is a superficial generalization of this result, as in Remark 2.20. With the proper definition, this result also holds for infinite W . Similarly to Theorem 2.21, we may deduce from our arguments the following, which also easily implies Theorem 3.9.

THEOREM 3.10. *Let (W, S) be a Coxeter system. Let $\mathbf{s}$ and $\mathbf{s}'$ be finite sequences from S of lengths n and n' , respectively, with $\mathbf{s}$ a proper subsequence of $\mathbf{s}'$. Let $\mathbf{C}$ be a Bernoulli(1/2) process. Then the probability distribution of $\xi(\mathbf{s}, \mathbf{C}_n)$ strictly majorizes that of $\xi(\mathbf{s}', \mathbf{C}_{n'})$.*

REMARK 3.11. There are additional consequences of the argument used in Theorem 3.5. Suppose that M is a wall. Write M^+ for the set of vertices v for which there is a path joining v to o without using an edge of M , and write M^- for the remainder. Since the Cayley graph of (W, S) is connected, Lemmas 2.3 and 2.1 imply that the reflection L_M in M interchanges M^+ and M^- . Let $\tau_{M^-} := \inf\{t > 0; Z_t \in M^-\}$. A direct analogue of the reflection principle (for one-dimensional random walks and Brownian motion) is that, for all $t > 0$ and all $A \subseteq M^+$,

$$\mathbf{P}[\tau_{M^-} < t, Z_t \in A] = \mathbf{P}[Z_t \in L_M A].$$

Therefore,

$$\mathbf{P}[\tau_{M^-} < t] = 2\mathbf{P}[Z_t \in M^-].$$

Write $\tau_v := \inf\{t > 0; Z_t = v\}$. If $x < y$, then τ_y strictly stochastically dominates τ_x , that is, for all $t > 0$, $\mathbf{P}[\tau_x < t] > \mathbf{P}[\tau_y < t]$. In particular, $\mathbf{P}[\tau_x < \infty] \geq \mathbf{P}[\tau_y < \infty]$. Let $\lambda_v(t)$ denote the Lebesgue measure of the set of times in $[0, t]$ when the random walk is at v . As a consequence of the strict stochastic domination inequality, we obtain that $\lambda_x(t)$ strictly stochastically dominates $\lambda_y(t)$ when $x < y$ and $t > 0$. We leave the arguments to the reader.

3.1.2. Abelian groups and conjugacy classes. In an abelian group, the answer to Question 3.1 is always positive, because increasing a rate results in extra multiplications by random group elements, and in an abelian group, we may consider these extra multiplications to take place at the end of the random walk.

As a generalization of this observation, consider a random walk on a group G with a generator g , every conjugate of which is also a generator. If all of the rates on this conjugacy class are increased the same amount, then this results in multiplying by additional group elements partway through the walk. Those extra elements are uniformly distributed in the conjugacy class of g , independently of all other steps. This is equivalent to multiplying by the same number of independent, uniformly chosen conjugates of g at the end. Thus, this also moves the resulting distribution down in the majorization order.

3.1.3. *All groups with special distances.* As long as we measure distance to stationarity in special ways, increasing rates will always decrease distance on any group.

PROPOSITION 3.12. *For any finite group, increasing any of the rates always decreases the ℓ^∞ -distance and the ℓ^2 -distance to the uniform distribution.*

PROOF. This relies on some well-known calculations. We have

$$\begin{aligned}
 |p_t(x, y) - 1/n| &= \left| \sum_z (p_{t/2}(x, z) - 1/n)(p_{t/2}(z, y) - 1/n) \right| \\
 &= \left| \sum_z (p_{t/2}(x, z) - 1/n)(p_{t/2}(y, z) - 1/n) \right| \\
 (3.1) \quad &\leq \sqrt{\sum_z (p_{t/2}(x, z) - 1/n)^2 \sum_z (p_{t/2}(y, z) - 1/n)^2} \\
 &= \sqrt{(p_t(x, x) - 1/n)(p_t(y, y) - 1/n)} \\
 &= p_t(o, o) - 1/n.
 \end{aligned}$$

Note that $p_t(o, o) \geq 1/n$ because the first two lines above, without absolute values, show that

$$(3.2) \quad p_t(o, o) - 1/n = \sum_z (p_{t/2}(o, z) - 1/n)^2 \geq 0.$$

In particular, (3.1) gives

$$\max_{x,y} |p_t(x, y) - 1/n| = p_t(o, o) - 1/n.$$

When any rate is increased, $p_t(o, o)$ decreases, as we noted in the [Introduction](#). Thus, the ℓ^∞ -distance decreases. In light of (3.2), the ℓ^2 -distance at time t equals the square root of the ℓ^∞ -distance at time $2t$, whence it is also decreasing in the rates. $\square$

3.2. *Negative examples.* Symmetric groups give Coxeter systems when generated by adjacent transpositions. Changing generators, however, may yield entirely different behaviour. Indeed, every finite group is a subgroup of a symmetric group. Therefore, any counterexample can be exhibited on a symmetric group by adding generators to those for the subgroup and making their rates 0 or very close to 0. However, symmetric groups being very large means that it can be hard to find examples by searching in symmetric groups.

In this section many of our examples were found by random numerical searching and numerical calculation—choosing sets of random generators and rates and exploring the consequences of increasing one of those rates. As we move through the different types of examples, we will attempt to give some idea of how difficult it was to find them, as an indication of the frequency of similar examples. We will usually be considering groups with up to about 120 elements and numbers of random samples in the hundreds of thousands. In cases where we were unable to find examples, we will indicate when we spent enough effort that we were surprised not to find them.

In most cases we used sets of four generators, with rates between about 1 and 10. Indeed, when choosing five generators at random, the most extreme examples often either included the identity or two copies of the same generator, suggesting that it is easier to find these examples with four distinct generators than with five.

For many of the groups we examined, it was not difficult to find examples where increasing a rate increased the ℓ^1 -distance from stationarity. This includes the symmetric groups S_4

and S_5 , dihedral groups with as few as 10 elements, and dicyclic groups with as few as 20 elements. We did not find these examples for split metacyclic groups. We will restrict our discussion now to dihedral and symmetric groups, as two quite different families of groups.

For the dihedral group D_{11} with three involutions as generators, we found an example where the entropy was not monotone in the rates. Note that dihedral groups are Coxeter groups, with Coxeter generators being two involutions.

As we increase p from 1 toward 2, it becomes more difficult to find examples where the ℓ^p -distance increases, which is consistent with Proposition 3.12. For instance, in the dihedral group D_5 we can find examples for values of p between 1 and 1.4, and in the group D_7 , for values of p between 1 and 1.8. It also seems easier to find these kinds of examples in dihedral groups D_n when n is prime, compared to a composite number of similar size, such as D_{31} or D_{41} compared to D_{30} or D_{40} .

While Proposition 3.12 says that there are no examples of this kind for $p = 2$, it is possible to get very close to this value. For instance, it was not uncommon for us to find examples exhibiting this behaviour for p up to 1.99 or from 2.01, and individually optimised examples worked for p as large as 1.9999.

The next natural question is what happens for values of p larger than two. In the case of dihedral groups, we found examples of sets of generators and rates that increase the ℓ^p -distance from uniform for a range of values of p between $2 + \epsilon$ and $4 - \epsilon$, for ϵ as small as 0.001, but we were unable to find any examples where the ℓ^p -distance increases for $p = 4$, searching in dihedral groups as large as D_{131} . This seems rather surprising. A generic example of this sort has the ℓ^p -distance increasing for p in some interval contained in $(2, 4)$. When the rates are adjusted so that the right endpoint of this interval moves closer to $p = 4$, the left endpoint seems to move closer to $p = 2$ and vice versa.

Likewise, we found an example in the dihedral group D_{41} for which the ℓ^p -distance increases for values of p between 4.001 and 5.995. Interestingly, this example also increases the ℓ^p -distance for p up to 1.997.

We were unable to find dihedral examples either at $p = 6$ or for larger values of p . Proposition 3.12 explains why distance cannot increase for $p = 2$, but we have no similar explanation for $p = 4$ or $p = 6$. While the ℓ^p -distance may increase for ranges of p less than 2 or between 2 and 4, we were unable to find any dihedral examples exhibiting this behaviour in both ranges simultaneously.

The proof of Proposition 3.12 relates the behaviour of the ℓ^2 -distance at time t to the ℓ^∞ -distance at time $2t$. One might then hope that if there are examples where the ℓ^p -distance increases for p near 2, then there should be similar examples where the ℓ^p -distance increases for p near ∞ , that is, for p very large, at twice the time. Yet we were unable to find such dihedral examples.

We looked for similar examples to these in the symmetric group S_5 . As with the dihedral groups, we found examples where the ℓ^p -distance increased for all p up to a value quite close to 2 and examples where it increased for most p between 2 and 4, but also new types of behaviour, including one where the ℓ^p -distance increased for p between 2.004 and 4.02 and several where it increased for p from about 3 up to as large as 300.

We did not find examples where the distance increased at $p = 4$ in dicyclic groups.

4. A ray. We now turn our attention from Cayley graphs to an infinite one-ended path and allow arbitrary rates on the edges. We label vertices by the nonnegative integers, and each pair of consecutive numbers is connected by an edge. For each positive integer i , let r_i be the (nonnegative) rate of the Poisson clock on the edge between $i - 1$ and i . Let $i_0 := \inf\{j; r_j = 0\} \in [1, \infty]$. For simplicity, we assume that the rates are such that explosions do not occur. We will see several interesting phenomena in this case. The random walk will begin from 0.

PROPOSITION 4.1. *For each time $t > 0$, the probability $p_t(0, i)$ that the walker is at location i is a strictly decreasing function of i for those $i < i_0$.*

PROOF. We first prove that $p_t(0, i)$ is weakly decreasing in i .

Let $f_t(i, j)$ be the probability that a random walk starting from i visits j at some time before t . Because there are no explosions, $\lim_{N \rightarrow \infty} f_t(0, N) = 0$. Consider the sequence of chains whose rates are r_i for $i \leq N$ and 0 for $i > N$; write $p_t^N(i, j)$ for their transition probabilities. Since $|p_t(0, i) - p_t^N(0, i)| \leq f_t(0, N)$, it follows that $\lim_{N \rightarrow \infty} p_t^N(0, i) = p_t(0, i)$ for all i . Therefore, it suffices to prove the claim for a chain where only finitely many rates are nonzero. We now assume that condition.

Consider the refresh times R_i of Definition 2.6. There are only finitely many such times before t a.s. If we condition on the sequence of all refresh rings that occur between time 0 and time t , ordered by time, then the probability distribution function of the walker's location at time t is a nonincreasing function of i . This is because the coin flip, or randomization, at each time of R_i acts on the walker's distribution function at that time by averaging the probabilities at $i - 1$ and i , and this operation preserves monotonicity. The walker is initially at position 0 with probability 1, and this initial distribution function is monotone.

Integrating these conditional distribution functions with respect to the distribution of all refresh rings before time t completes the proof of the claim.

Next, we show that the transition probabilities are strictly decreasing. We no longer assume that only finitely many rates are nonzero. Let $\tilde{p}_t(0, \cdot)$ denote the probability distribution conditional on the set of refresh rings before time t . There is a collection of positive probability of sets of refresh rings for each of which $\tilde{p}_t(0, i - 1) > \tilde{p}_t(0, i)$ for $i < i_0$. Because $\tilde{p}_t(0, i - 1) \geq \tilde{p}_t(0, i)$ for each possible set of refresh rings and $p_t(0, i)$ is the expectation of $\tilde{p}_t(0, i)$, it follows that $p_t(0, i)$ is strictly decreasing for $0 \leq i < i_0$. $\square$

This analysis also lets us discuss the speed at which the walker moves away from 0.

LEMMA 4.2. *In the language of the proof of Proposition 4.1, consider the following two distribution functions on $\mathbb{Z}^+$. The function f_1 is defined by starting with a unit mass at 0 and applying any fixed finite sequence of refresh rings. The function f_2 is defined similarly, using the same sequence of refresh rings, except with one additional refresh ring R_j occurring partway through the sequence. Then f_2 dominates f_1 , that is, $\sum_{i=0}^k f_1(i) \geq \sum_{i=0}^k f_2(i)$ for every $k \in \mathbb{Z}^+$.*

PROOF. Consider the evolution of the difference $f_2 - f_1$. This is zero until the extra R_j is applied to f_2 . At this point, $f_2 - f_1$ is zero at all points, except $j - 1$ and j . At $j - 1$ it is negative (or zero), at j it is positive (or zero), and these two differences have the same size (but opposite sign).

At this point the difference $f_2 - f_1$ has the property that when summed over the states 0 to any k , it is nonpositive. This property is preserved by the application of any of the R_i , which completes the proof. $\square$

This implies that the location of the random walk at each time is stochastically strictly increasing in each of the rates r_j .

COROLLARY 4.3. *At every time t and nonnegative integer $i < i_0$, the probability that the walker is at a position between 0 and i is a strictly decreasing function of each r_j for $j < i_0$.*

PROOF. As in the proof of Proposition 4.1, we first prove the statement without the strictness under the assumption that $i_0 < \infty$. When some R_j rings one additional time, then the result follows from Lemma 4.2. Applying this fact repeatedly gives the result. It then follows even if $i_0 = \infty$, but without the strictness part.

To show strictness, it again suffices, for each i , to exhibit a collection of positive probability of sequences of refresh rings with a marked additional ring of R_j that each give a strict inequality. Indeed, such a collection is formed by the sequences that when restricted to the refresh rings for edges $k \leq i \vee j$ are sequences of the form $(R_1, R_2, \dots, R_{i \wedge j}, \dots, R_{i \vee j})$, where (one of) the R_j is marked additional. $\square$

COROLLARY 4.4. *At every time t , the expected distance of the walker from 0 is a strictly increasing function of each r_j for $j < i_0$.*

PROOF. The expected distance from 0 is the sum over all i of the probabilities that the walker is farther away than i . Hence, the result follows from Corollary 4.3. $\square$

We are also interested in how much time the walker spends at each vertex.

PROPOSITION 4.5. *For our random walk on the infinite one-ended path, let t be an arbitrary time. Then the time spent at 0 between time 0 and time t stochastically dominates the time spent at 1.*

PROOF. As in previous proofs, we may assume that $i_0 < \infty$. Without loss of generality, rescale time so that the rate r_1 is equal to 1. Fix $t > 0$. We define recursively the following sequences of times. Let $A_1 := 0$. Then let B_k be the infimum of the times in $[0, t]$ after time A_k at which the walker is at the vertex 1 if there is such a time and $B_k := t$ if not, and A_{k+1} be the infimum of the times in $[0, t]$ after B_k at which the walker is at 0 if there is such a time and t if not, for each k . For each k , let X_k be the amount of time the walker spends at vertex 0 between time A_k and A_{k+1} , and Y_k be the amount of time spent at 1 between B_k and B_{k+1} .

If we didn't stop counting at time t , then both X_k and Y_k would be distributed as exponential random variables with rate 1, because they count the time spent at vertex 0 or 1, respectively, until the Poisson clock corresponding to the edge $(0, 1)$ rings while the walk is at one of its endpoints, and this clock has rate 1. Taking this into account, we see that the conditional distribution of X_k , given A_k and (X_i, Y_i) for all $i < k$, is the same as the distribution of $\min(\text{Exp}(1), t - A_k)$. We can't give such a concrete expression for the conditional distribution of Y_k , because the walker may move to 2 and beyond before moving back to 1 and then to 0. We can say, though, that Y_k given A_k , X_k , and (X_i, Y_i) for all $i < k$ is stochastically dominated by $\min(\text{Exp}(1), t - A_k - X_k)$. In particular, this means that X_k dominates Y_k , given (X_i, Y_i) for all $i < k$. It follows that the sum of the X_k stochastically dominates the sum of the Y_k . Since the total time spent at 0 between 0 and t is the sum of the X_k and the time spent at 1 is the sum of the Y_k , we obtain the desired result. $\square$

In the preceding proof, note that we gave an explicit expression for X_k and an upper bound for Y_k . If the walker was allowed to move away from the vertex 0 without moving to 1, then we would only have obtained an upper bound for X_k , which would not allow us to compare X_k and Y_k . Example 4.6 illustrates what may go wrong in such a scenario.

Surprisingly, even the simplest generalizations of Proposition 4.5 are not true. The following example shows that in the same setting, the walker need not spend more time at vertex 1 than at 2.

EXAMPLE 4.6. Fix $t > 0$. Set the rate r_1 to be extremely large, so that the total time spent at vertex 0 or 1 before time t is very likely to be almost the same. Take r_2 to be 1 and r_3 to be zero (thus, the walker will never reach any state beyond 2, and the other rates are irrelevant). Then at time t , the probability that the walker has spent more than time $\frac{2t}{3}$ at vertex 1 is almost zero, because it spends almost equal amounts of time at 0 and 1. However, there is probability bounded away from 0 that the walker spends more time than $\frac{2t}{3}$ at state 2; for instance, it could move to state 2 by time $\frac{t}{4}$, and then the edge between states 1 and 2 might not ring again before time t .

Therefore, the time spent at 1 does not stochastically dominate the time spent at 2.

The next example illustrates how Corollary 4.4 may fail in cases that are only slightly more complicated. The state space will still be a path, but the walker will not start at an end.

EXAMPLE 4.7. Expand the setting so that there is a vertex for each integer i , with r_i still the rate for moving between $i - 1$ and i . The walker still begins at 0, but there are now possible states on either side of 0.

If $r_i = 0$ for $i \neq 0, 1$ and $r_1 = 1$, then one can check by explicit calculation that $p_t(0, 0)$ is larger when $r_0 = 3$ than it is when $r_0 = 2$, whence the expected distance from 0, being $1 - p_t(0, 0)$, is smaller when $r_0 = 3$ than it is when $r_0 = 2$.

For a more extreme example, choose arbitrary positive integers k and n , and set the rates on edges between $-k$ and 0 to each be 1 and rates on edges between 0 and $n - 1$ to be very large and the same. Rates outside these ranges are zero. At times that are much less than 1 but large compared to the reciprocal of the large rate, the walker is approximately equally likely to be at any state between 0 and $n - 1$, for an expected distance of $\frac{n-1}{2}$. At large times the walker is approximately equally likely to be at any of the $k + n$ states, for an expected distance of $\frac{n(n-1)+k(k+1)}{2(n+k)}$. For some values of k and n , the expected distance from the starting state has decreased between these two regimes.

If we take $k = \alpha n$ and consider the limit as k and n increase, the most extreme ratio occurs when $\alpha = \sqrt{2} - 1$, which gives a ratio between the small-time and large-time expected distances of $\frac{1+\sqrt{2}}{2}$.

One permissible generalization of Proposition 4.5 is to settings where the excursions from the two vertices in question have the same distribution. This includes the setting of Cayley graphs with rates depending only on the generators.

PROPOSITION 4.8. *Consider random walk on a graph whose edge rates are preserved by some group of graph automorphisms that acts transitively on the vertices. Then the time spent between 0 and t at the initial vertex u stochastically dominates the time spent at each other vertex v .*

PROOF. Let $\lambda(x, t)$ be the amount of time between 0 and t spent at the vertex x by a random walk started at x at time 0. Because the graph is vertex-transitive, the law of this quantity does not depend on x . The duration $\lambda(x, t)$ is also increasing in t . Let τ be a random variable independent of the random walk and whose distribution is the same as that of the hitting time of the vertex v when the random walk is started from u . Then we are comparing $\lambda(u, t)$ with $\lambda(v, (t - \tau)^+)$, whose law is the same as that of $\lambda(u, (t - \tau)^+)$. The random variable τ is positive, which completes the proof. $\square$

Acknowledgments. We thank Kavita Ramanan for early discussions of some of these questions. We also thank Yiping Hu for asking us about convergence to stationarity. Thanks to Yuval Peres, who alerted us to [28] and [13] after seeing an earlier draft of our results. We also thank one of the referees for informing us of [18]. We are grateful to Matthias Weber for permission to include Figure 1.

Funding. The first author was supported by the National Science Foundation under grants DMS-1612363 and DMS-1954086 and also by the Simons Foundation.

REFERENCES

- [1] ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. [MR2354165](#) <https://doi.org/10.1214/EJP.v12-463>
- [2] BENJAMINI, I., LYONS, R. and SCHRAMM, O. (1999). Percolation perturbations in potential theory and random walks. In *Random Walks and Discrete Potential Theory (Cortona, 1997)*. *Sympos. Math.*, XXXIX 56–84. Cambridge Univ. Press, Cambridge. [MR1802426](#)
- [3] BJÖRNER, A. and BRENTI, F. (2005). *Combinatorics of Coxeter Groups*. *Graduate Texts in Mathematics* **231**. Springer, New York. [MR2133266](#)
- [4] BRINK, B. and HOWLETT, R. B. (1993). A finiteness property and an automatic structure for Coxeter groups. *Math. Ann.* **296** 179–190. [MR1213378](#) <https://doi.org/10.1007/BF01445101>
- [5] BRUALDI, R. A. and DAHL, G. (2013). Majorization for partially ordered sets. *Discrete Math.* **313** 2592–2601. [MR3095434](#) <https://doi.org/10.1016/j.disc.2013.08.003>
- [6] CALEGARI, D. (2013). The ergodic theory of hyperbolic groups. In *Geometry and Topology down Under*. *Contemp. Math.* **597** 15–52. Amer. Math. Soc., Providence, RI. [MR3186668](#) <https://doi.org/10.1090/conm/597/11762>
- [7] CAPRACE, P.-E. (2006). Conjugacy of 2-spherical subgroups of Coxeter groups and parallel walls. *Algebr. Geom. Topol.* **6** 1987–2029. [MR2263057](#) <https://doi.org/10.2140/agt.2006.6.1987>
- [8] CHEN, L.-C. and SUN, R. (2014). A monotonicity result for the range of a perturbed random walk. *J. Theoret. Probab.* **27** 997–1010. [MR3245995](#) <https://doi.org/10.1007/s10959-012-0472-x>
- [9] DAVIS, M. W. (2008). *The Geometry and Topology of Coxeter Groups*. *London Mathematical Society Monographs Series* **32**. Princeton Univ. Press, Princeton, NJ. [MR2360474](#)
- [10] DIACONIS, P. and MICLO, L. (2009). On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.* **22** 558–586. [MR2530103](#) <https://doi.org/10.1007/s10959-009-0234-6>
- [11] FILL, J. A. (2009). The passage time distribution for a birth-and-death chain: Strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.* **22** 543–557. [MR2530102](#) <https://doi.org/10.1007/s10959-009-0235-5>
- [12] FILL, J. A. (2009). On hitting times and fastest strong stationary times for skip-free and more general chains. *J. Theoret. Probab.* **22** 587–600. [MR2530104](#) <https://doi.org/10.1007/s10959-009-0233-7>
- [13] FILL, J. A. and KAHN, J. (2013). Comparison inequalities and fastest-mixing Markov chains. *Ann. Appl. Probab.* **23** 1778–1816. [MR3114917](#) <https://doi.org/10.1214/12-AAP886>
- [14] FONTES, L. R. G. and MATHIEU, P. (2006). On symmetric random walks with random conductances on $\mathbb{Z}^d$. *Probab. Theory Related Fields* **134** 565–602. [MR2214905](#) <https://doi.org/10.1007/s00440-005-0448-1>
- [15] GOUËZEL, S. (2017). Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Discrete Anal.* **7**. [MR3651925](#) <https://doi.org/10.19086/da.1639>
- [16] GROMOV, M. (1987). Hyperbolic groups. In *Essays in Group Theory*. *Math. Sci. Res. Inst. Publ.* **8** 75–263. Springer, New York. [MR0919829](#) https://doi.org/10.1007/978-1-4613-9586-7_3
- [17] HEICKLEN, D. and HOFFMAN, C. (2005). Return probabilities of a simple random walk on percolation clusters. *Electron. J. Probab.* **10** 250–302. [MR2120245](#) <https://doi.org/10.1214/EJP.v10-240>
- [18] HERMON, J. and KOZMA, G. (2021). Sensitivity of mixing times of Cayley graphs. Preprint. Available at [arXiv:2008.07517](#).
- [19] HUMPHREYS, J. E. (1990). *Reflection Groups and Coxeter Groups*. *Cambridge Studies in Advanced Mathematics* **29**. Cambridge Univ. Press, Cambridge. [MR1066460](#) <https://doi.org/10.1017/CBO9780511623646>
- [20] KAIMANOVICH, V. A. (2000). The Poisson formula for groups with hyperbolic properties. *Ann. of Math.* (2) **152** 659–692. [MR1815698](#) <https://doi.org/10.2307/2661351>
- [21] KARLIN, S. and MCGREGOR, J. (1959). Coincidence properties of birth and death processes. *Pacific J. Math.* **9** 1109–1140. [MR0114247](#)

- [22] KARLIN, S. and MCGREGOR, J. L. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.* **85** 489–546. [MR0091566](#) <https://doi.org/10.2307/1992942>
- [23] LYONS, R. (2017). Comparing graphs of different sizes. *Combin. Probab. Comput.* **26** 681–696. [MR3681977](#) <https://doi.org/10.1017/S096354831700013X>
- [24] LYONS, R. (2018). Monotonicity of average return probabilities for random walks in random environments. In *Unimodularity in Randomly Generated Graphs. Contemp. Math.* **719** 1–9. Amer. Math. Soc., Providence, RI. [MR3880007](#) <https://doi.org/10.1090/conm/719/14464>
- [25] LYONS, R. and PERES, Y. (2016). *Probability on Trees and Networks. Cambridge Series in Statistical and Probabilistic Mathematics* **42**. Cambridge Univ. Press, New York. [MR3616205](#) <https://doi.org/10.1017/9781316672815>
- [26] MCMURRAY PRICE, T. (2017). An inequality for the heat kernel on an Abelian Cayley graph. *Electron. Commun. Probab.* **22** 57. [MR3718707](#) <https://doi.org/10.1214/17-ECP84>
- [27] MICLO, L. (2010). On absorption times and Dirichlet eigenvalues. *ESAIM Probab. Stat.* **14** 117–150. [MR2654550](#) <https://doi.org/10.1051/ps:2008037>
- [28] PERES, Y. and WINKLER, P. (2013). Can extra updates delay mixing? *Comm. Math. Phys.* **323** 1007–1016. [MR3106501](#) <https://doi.org/10.1007/s00220-013-1776-0>
- [29] PITTET, C. and SALOFF-COSTE, L. (2000). On the stability of the behavior of random walks on groups. *J. Geom. Anal.* **10** 713–737. [MR1817783](#) <https://doi.org/10.1007/BF02921994>
- [30] REGEV, O. and SHINKAR, I. (2016). A counterexample to monotonicity of relative mass in random walks. *Electron. Commun. Probab.* **21** 8. [MR3485377](#) <https://doi.org/10.1214/16-ECP4392>
- [31] SAWYER, S. and STEGER, T. (1987). The rate of escape for anisotropic random walks in a tree. *Probab. Theory Related Fields* **76** 207–230. [MR0906775](#) <https://doi.org/10.1007/BF00319984>
- [32] SPEYER, D. E. Infinite Coxeter groups with a non-trivial finite conjugacy class? MathOverflow. Available at <https://mathoverflow.net/q/82921> (version: 2011-12-08).
- [33] SPEYER, D. E. (2009). Powers of Coxeter elements in infinite groups are reduced. *Proc. Amer. Math. Soc.* **137** 1295–1302. [MR2465651](#) <https://doi.org/10.1090/S0002-9939-08-09638-X>
- [34] WHITE, G. (2016). The weak Bruhat order for random walks on Coxeter groups. Available at [arXiv:1611.04098](https://arxiv.org/abs/1611.04098).