Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 50, 1–15.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1109

Random walks on regular trees can not be slowed down*

Omer Angel† Jacob Richey‡ Yinon Spinka§ Amir Yehudayoff¶

Abstract

We study a permuted random walk process on a graph G. Given a fixed sequence of permutations on the vertices of G, the permuted random walker alternates between taking random walk steps, and applying the next permutation in the sequence to their current position. Existing work on permuted random walks includes results on hitting times, mixing times, and asymptotic speed. The usual random walk on a regular tree, or generally any non-amenable graph, has positive speed, i.e. the distance from the origin grows linearly. Our focus is understanding whether permuted walks can be slower than the corresponding non-permuted walk, by carefully choosing the permutation sequence. We show that on regular trees (including the line), the permuted random walk is always stochastically faster. The proof relies on a majorization inequality for probability measures, plus an isoperimetric inequality for the tree. We also quantify how much slower the permuted random walk can possibly be when it is coupled with the corresponding non-permuted walk.

Keywords: random walks; speed of random walk; permuted random walk.
MSC2020 subject classifications: 05C81.
Submitted to EJP on September 17, 2023, final version accepted on March 5, 2024.

1 Introduction

One of the classical results relating the geometry of a space to the behaviour of random walks on the space is that on any non-amenable graph the random walk has positive speed, in that $\liminf_{t\to\infty} t^{-1}|X_t|$ exists and is a.s. positive. On transitive graphs the limit is even an almost sure constant. In particular, on the d regular tree, denoted $\mathbb{T}_d$, the speed for the simple random walk is $\frac{d-2}{d}$, which is positive as long as $d > 2$. The motivation for this paper is the question: *Can we slow down the particle?*

*OA, JR and YS were supported in part by NSERC while working on this project. AY is partially supported by the BSF.

†University of British Columbia, Vancouver, Canada. E-mail: angel@math.ubc.ca
‡Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: jrichey@renyi.hu
§School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel. E-mail: yinonspi@tauex.tau.ac.il
¶Department of Mathematics, Technion-IIT, Haifa, Israel. E-mail: amir.yehudayoff@gmail.com

Suppose that after each step t of the random walk, we are allowed to apply some permutation π_t to the vertices of the tree, so that if the particle is at v it is transported to $\pi_t(v)$. If we observe the particle and can choose π_t accordingly, then we can constantly push it back to any vertex we wish, so that it never moves. Our main finding is that if the permutations do not depend on the location of the particle, then the particle can not be slowed down.

1.1 Permuted random walks

We start by considering lazy random walks, where the results are cleaner for mostly technical reasons (see the discussion below). We start by introducing some notation. Fix $d \geq 2$, and let $\mathbb{T} = \mathbb{T}_d$ denote the rooted infinite d-regular tree. The vertex set is denoted by $V = V(\mathbb{T}_d)$. The root of the tree is denoted by v_0. The depth $|v|$ of a vertex $v \in V$ is its distance from the root. The neighborhood $N(v)$ of v is the set of vertices u that are of distance at most one from v. Note that since we are considering lazy random walks, it is convenient to have $v \in N(v)$. Thus the size of $N(v)$ is $d+1$.

Let $(X_t)_{t=0}^{\infty}$ be a lazy random walk on $\mathbb{T}$ started at the root. The laziness parameter $\mathbb{P}(X_{t+1} = X_t)$ is chosen to be $\frac{1}{d+1}$. That is, $X_0 = v_0$ and X_{t+1} is a uniformly random element of $N(X_t)$. The (empirical) speed of (X_t) is defined to be the process $(t^{-1}|X_t|)$. The strong law of large numbers implies that the speed a.s. converges to $\frac{d-2}{d+1}$. Note that this also holds in the case $d = 2$ where $\mathbb{T}_2$ is the line and the speed is 0. For more on random walks on trees see e.g. [8, 10] and references therein.

The model we suggest for studying the slowing down of particles is as follows. Before the particle starts to move, we can choose a sequence $(\pi_t)_{t=1}^{\infty}$ of permutations of V. (These do not need to be finitary; any bijections of V will do.) The permutation π_t is applied on the random walk at time t. Thus the permuted random walk (Y_t) starts at the root, and its position at time $t+1$ is defined by $Y_{t+1} = \pi_{t+1}(Y'_{t+1})$, where Y'_{t+1} is a uniformly random vertex in $N(Y_t)$. The (empirical) speed of (Y_t) is the process $(t^{-1}|Y_t|)$. In contrast with (X_t), the permuted random walk may not have a limiting speed. The lower speed of the permuted random walk is defined by $\liminf_{t\to\infty} t^{-1}|Y_t|$.

Permuted random walk have been studied before, both on their own merit, and as a tool towards other ends. Pymar and Sousi [2] established uniform bounds on hitting times for permuted random walks on finite regular fgraphs. Ganguly and Peres [1] studied walks on an interval with a fixed uniform random permutation. Recently, Chatterjee and Diaconis [3, 5] demonstrated that mixing of certain Markov chains can be significantly sped up by adding a deterministic permutation after each move. In a different direction, Gouëzel [4] used permuted random walks to establish large deviation lower bounds on the speed of random walks on hyperbolic spaces without moment assumptions on the step distribution. One idea here is to condition on the long steps of the walk, and consider the process as a permuted version of a walk with bounded steps for which other methods can apply. The question at the heart of this paper arose following a presentation of that work.

Our main result is that no matter how we select the permutations (π_t), the permuted walk (Y_t) is not slower than (X_t).

Theorem 1.1. *For every $d \geq 2$, every sequence (π_t) of permutations of $V(\mathbb{T}_d)$, and every time $t \geq 0$, the depth of the permuted random walk $|Y_t|$ stochastically dominates the depth of the lazy random walk $|X_t|$. That is, for all $t, n \geq 0$,*

$$\mathbb{P}[|Y_t| \geq n] \geq \mathbb{P}[|X_t| \geq n].$$

In particular, $\mathbb{E}|Y_t| \geq \mathbb{E}|X_t|$ for all $t \geq 0$, and $\liminf_{t\to\infty} t^{-1}|Y_t| \geq \frac{d-2}{d+1}$ almost surely.

Remark 1.2. The processes (X_t) and (Y_t) in Theorem 1.1 correspond to a lazy random walk that stays put with probability $\frac{1}{d+1}$. Theorem 1.1 holds verbatim (with the obvious change to the constant $\frac{d-2}{d+1}$) as long as the probability to stay put is at least $\frac{1}{d+1}$. In particular, it holds when the chance to stay put is one half, which is a more common definition of the lazy random walk. For details, see the remark after the proof of Theorem 1.1.

Note that the theorem is informative even for $d = 2$, where the limit speed is zero. However, the laziness is required for Theorem 1.1 to hold. Indeed, for the non-lazy walk on $\mathbb{T}_d$ we can have $\mathbb{E}|Y_1| < \mathbb{E}|X_1|$ (or for any other t).

Theorem 1.1 is a special case of a more general phenomenon, which we describe in the next two theorems. For a distribution p on V, define $p^* : \mathbb{N} \to [0,1]$ by letting $p^*(j)$ be the total mass of the j largest atoms in p, or equivalently,

$$p^*(s) = \max\{p(J) : J \subset V, |J| = s\}.$$

We say that a distribution p **majorizes** a distribution q if $p^*(j) \geq q^*(j)$ for all $j \in \mathbb{N}$.

Denote by p_t the distribution of X_t and by q_t the distribution of Y_t, depending implicitly on the fixed permutations (π_t). The stochastic domination asserted in Theorem 1.1 is a consequence of the following more technical statement. The main reasons are that the distribution p_t is spherically symmetric and monotone in depth; for more details, see section 3.

Theorem 1.3. *For every $t \geq 0$, the distribution p_t majorizes q_t.*

The fact that p_t majorizes q_t can be interpreted as saying that the amount of disorder in q_t is at least that of p_t. Concretely, the theorem implies that the Shannon entropy of Y_t is at least the Shannon entropy of X_t. There is no way to decrease the entropy of a lazy random walk on a regular tree by applying time dependent permutations.

A second interpretation of the theorem, which follows from the Birkhoff–von-Neumann decomposition of a bistochastic matrix into a convex combination of permutation matrices, is that for every t, there is a distribution r_t on permutations of V, so that if σ_t is sampled from r_t independently of X_t, then $(X_t, \sigma_t(X_t))$ has the same distribution as (X_t, Y_t). In other words, there is a distribution on a single permutation σ_t that allows to replace the iterative application of the t permutations $\pi_1, \dots, \pi_t$.

An even more general statement than Theorem 1.3 holds. Let $B_n = \{v \in V : |v| \leq n\}$ denote the ball of radius $n \geq -1$ in the tree[1] and $\partial B_n = B_n \setminus B_{n-1}$ the sphere of radius n. Fix an order $v_0, v_1, v_2, \dots$ of V with the following property: for every $i < j$, it holds that $|v_i| \leq |v_j|$ and if $|v_i| = |v_j|$ then the $d-1$ children of v_i appear in the order before the $d-1$ children of v_j. Initial segments of the form $\{v_0, v_1, \dots, v_i\}$ are called **quasi-balls**. Note that every ball is a quasi-ball. A distribution p on V is called **greedily arranged** if $p(v_i) \geq p(v_{i+1})$ for every i.

Theorem 1.4. *Let p and q be distributions on V and let p' and q' be the corresponding distributions after a single step of a lazy random walk started at p and q, respectively. If p is greedily arranged and majorizes q, then p' is greedily arranged and majorizes q'.*

Theorems 1.1 and 1.3 follow by a simple inductive argument from the last theorem using the following two observations. First, the initial distribution p_0 is greedily arranged, and majorizes q_0. Second, if a distribution p majorizes q, then it also majorizes any rearrangement of q (i.e., a distribution of the form $q \circ \pi$ for a permutation π of V). Thus Theorem 1.4 implies that for every $t \geq 0$ and every finite $J \subset V$,

$$p_t(B) \geq q_t(J), \tag{1.1}$$

where B is the quasi-ball of size $|B| = |J|$.

[1] The ball B_{-1} is empty.

1.2 Non-lazy random walks

The last result is particular to lazy random walks on regular trees. For non-lazy walks, it is too strong to be true. The distribution of a simple non-lazy random walk on a regular tree is not greedily arranged because the tree is bipartite; in particular, (1.1) may fail already for $t = 1$.

On the other hand, versions of the above theorems do hold for non-lazy walks, as we describe next. The limit speed of a simple (non-lazy) random walk on $\mathbb{T}_d$ is $\frac{d-2}{d}$ a.s. As noted, for such random walks, the same stochastic domination as in Theorem 1.1 does not hold. Nonetheless, we prove that it almost holds (at least for $d > 2$, when the tree is not the line).

Denote by $N'(v)$ the d neighbors of v not including v. Let (S_t) be a simple random walk so that $S_0 = v_0$ and S_{t+1} is uniform in $N'(S_t)$. Let (Z_t) be a permuted simple random walk so that $Z_0 = v_0$ and Z_{t+1} is $\pi_{t+1}(Z'_{t+1})$, where Z'_{t+1} is uniform in $N'(Z_t)$.

Theorem 1.5. *For every $d > 2$, every sequence (π_t) of permutations of $V(\mathbb{T}_d)$, and every time $t \geq 1$, we have that $|Z_t| + 2$ stochastically dominates $|S_t|$. In particular, $\mathbb{E}|Z_t| \geq \mathbb{E}|S_t| - 2$ for all $t \geq 0$, and $\liminf_{t\to\infty} t^{-1}|Z_t| \geq \frac{d-2}{d}$ almost surely.*

For $d = 2$, the bound $\frac{d-2}{d}$ on the lower speed of (Z_t) trivially holds, but the stronger claim in the theorem is false. One way to see this is to take π_t to be the identity up to some large time $2T$, and then map via π_{2T} all even integers in the range $[-2T, 2T]$ to all integers in $[-T, T]$ so that $\mathbb{E}|Z_{2T}| = \frac{1}{2}\mathbb{E}|S_{2T}|$.

We shall deduce Theorem 1.5 from the following modification of Theorem 1.4 which takes into account the periodicity of the non-lazy walk. The vertex set can be partitioned according to parity into $V_0 = \{v \in V : |v| = 0 \mod 2\}$ and $V_1 = V \setminus V_0$. A distribution p is called **half-greedily arranged** if it is supported on one of V_0 or V_1, and $p(v_i) \geq p(v_j)$ for every $i < j$ for which v_i and v_j have the same parity (using the same ordering of V as above).

Theorem 1.6. *Let p and q be distributions on V, and let p' and q' be the corresponding distributions after a single step of a non-lazy random walk started at p and q, respectively. If p is half-greedily arranged and majorizes q, then p' is half-greedily arranged and majorizes q'.*

Although the distribution of S_t is not greedily arranged, it is half-greedily arranged. The theorem thus implies that the distribution of S_t majorizes that of Z_t for every $t \geq 0$ (although the distribution of $|S_t|$ does not necessarily majorizes that of $|Z_t|$).

1.3 The speed process

Theorems 1.1 and 1.5 establish stochastic domination of the distance of a standard (lazy/simple) random walk by the distance of a permuted random walk *at any particular time*. It is natural to wonder whether such stochastic domination holds for the corresponding processes, i.e., whether the two processes can be coupled so that the distance of the permuted walk is always at least the distance of the standard walk. Somewhat surprisingly, it turns out this is not always possible. We focus on lazy random walks for concreteness. As an example, consider a sequence of permutations π in which π_1 and π_2 are the identity permutation and π_3 is an automorphism of $\mathbb{T}$ which maps a neighbor of v_0 to v_0. A direct computation yields that

$$\mathbb{P}[|X_2| + |X_3| \leq 2] = \tfrac{1}{d+1} + \tfrac{4d}{(d+1)^3} < \tfrac{1}{d+1} + \tfrac{5d-1}{(d+1)^3} = \mathbb{P}[|Y_2| + |Y_3| \leq 2],$$

so that $(|Y_2|, |Y_3|)$ does not stochastically dominate $(|X_2|, |X_3|)$.

When $d = 2$, this effect can be repeated and magnified over time. The next result shows that for certain choices of permutations, even translations, there are infinitely many times at which the distance of the permuted random walk is much smaller (no matter how the two processes are coupled).

Theorem 1.7. *Fix $d = 2$. There exists a sequence of permutations (π_t) of $V(\mathbb{T}_2) \cong \mathbb{Z}$, all of which are translations, such that in any coupling of the lazy random walk process (X_t) and the permuted random walk process (Y_t), almost surely,*

$$\limsup_{t\to\infty} \frac{|X_t| - |Y_t|}{\sqrt{t \log\log t}} = \frac{\sqrt{3}}{2}. \tag{1.2}$$

When $d > 2$, on the other hand, we show that the above cannot occur (not even nearly) when the permutations are required to be automorphisms of $\mathbb{T}_d$. This is the content of the result below. We do not know how strong this effect can be for general permutations. For instance, we do not know whether it is always possible to couple the two processes so that, almost surely, $|X_t| \le |Y_t|$ for all large enough t.

Theorem 1.8. *For every $d > 2$ and every sequence of automorphisms (π_t) of $\mathbb{T}_d$, there exists a coupling of the lazy random walk process (X_t) and the permuted random walk process (Y_t) such that, almost surely,*

$$|Y_t| - |X_t| \ge t^{1/2-o(1)} \qquad \text{as } t \to \infty.$$

The theorem is interesting even when each π_t is the identity. It states that there is a way to couple two lazy random walks so that one is significantly more distant than the other. The result is tight is the sense that the $o(1)$ term cannot be dropped entirely. Our proof gives a quantitative estimate for this term and yields that $t^{1/2-o(1)}$ can be replaced with $\sqrt{t}/(\log^C t)$ for some constant $C > 0$. See Lemma 5.2 and the second remark following it.

1.4 A spectral argument

One natural approach towards proving the results above is using spectral methods (see [7] and references within). Specifically, the transition kernel on $\ell^2(V)$ is a contraction with norm $\rho < 1$, and application of a permutation is an isometry on $\ell^2(V)$. Thus $\|q_t\|_2 \le \rho^t$ decays exponentially. A positive lower bound on the lower speed of Y_t follows easily. Moreover, this argument holds for any non-amenable graph. However, the resulting bound on the speed is not sharp.

The proof of a spectral gap uses an isoperimetric inequality for the tree. Not surprisingly, our proofs also use isoperimetric inequalities; see Lemmas 2.2 and 2.3 below. Lemma 2.3 is a non-standard isoperimetric inequality, which takes into account the amount of “isolated” points in the set of interest. Proposition 2.1 is a significant generalization of the isoperimetric inequality using the language of majorization.

2 Isoperimetry

As noted, our arguments rely on isoperimetric properties of the tree. However, to get the strongest possible comparison between the permuted and regular random walks we need sharp isoperimetric inequalities, which we now proceed to prove.

Recall that $N(v)$ is the neighborhood of a vertex v, including v itself. For $J \subset V$, the neighborhood of J is defined by

$$N(J) = \bigcup_{v \in J} N(v).$$

To analyze the behavior of the random walk, we need to understand the boundary in more detail. For $J \subset V$ and $i \in [d+1]$, define

$$K_i(J) = \{v \in V : |N(v) \cap J| \ge i\}.$$

In particular, the set $K_1(J)$ is the neighborhood $N(J)$.

A **partition** is a sequence $\mu = (\mu_1, \mu_2, \ldots, \mu_\ell)$ with $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_\ell \geq 0$. Note that usually trailing 0's are omitted, but for us it is convenient to have the length of the partitions be fixed, so we may include 0's. The size of the partition is defined by $|\mu| = \sum_i \mu_i$. The dominance order on partitions is defined as follows. For partitions μ, λ, we write $\lambda \prec \mu$ if $|\mu| = |\lambda|$ and

$$\lambda_1 + \cdots + \lambda_r \leq \mu_1 + \cdots + \mu_r, \qquad \text{for all } r. \tag{2.1}$$

The following majorization statement is an extension of the standard isoperimetric inequality for the tree.

Proposition 2.1. *Let $J \subset V$ be finite and let B be the quasi-ball with $|B| = |J|$. Let $k_i = |K_i(J)|$ and $m_i = |K_i(B)|$. Then (k_i) dominates (m_i) as partitions: $(k_1, \ldots, k_{d+1}) \succ (m_1, \ldots, m_{d+1})$.*

To prove this result, we need a couple of lemmas on the isoperimetric behavior of the tree. Let $\kappa_1(J)$ denote the number of connected components induced by J. Let $\kappa_2(J)$ denote the number of connected components induced by J in the graph in which edges are added between all pairs of vertices that are at distance 2 from each other in the tree. The first lemma is a formula for $|N(J)|$ for general J:

Lemma 2.2. *For every finite $J \subset V$,*

$$|N(J)| = (d-1)|J| + \kappa_1(J) + \kappa_2(J).$$

Proof. We prove the claim by induction on $|J|$. The base case when $|J| = 0$ is trivial. Let J be non-empty. Let $v \in J$ be a vertex of maximum depth in J. Let $N_1 = N(v) \cap (J \setminus \{v\})$ and $N_2 = N(N(v)) \cap (J \setminus \{v\})$. The following two equalities hold:

$$\kappa_1(J \setminus \{v\}) = \kappa_1(J) - \mathbf{1}_{\{N_1 = \emptyset\}} \qquad \text{and} \qquad \kappa_2(J \setminus \{v\}) = \kappa_2(J) - \mathbf{1}_{\{N_2 = \emptyset\}}.$$

The induction hypothesis implies

$$|N(J \setminus \{v\})| = (d-1)|J| + \kappa_1(J) + \kappa_2(J) - (d - 1 + \mathbf{1}_{\{N_1 = \emptyset\}} + \mathbf{1}_{\{N_2 = \emptyset\}}).$$

It remains to show that

$$|N(J)| - |N(J \setminus \{v\})| = d - 1 + \mathbf{1}_{\{N_1 = \emptyset\}} + \mathbf{1}_{\{N_2 = \emptyset\}}.$$

The left-hand side equals

$$|N(J) \setminus N(J \setminus \{v\})| = |N(v) \setminus N(J \setminus \{v\})| = d + 1 - |N(v) \cap N(J \setminus \{v\})|.$$

So we need to show that

$$|N(v) \cap N(J \setminus \{v\})| = 2 - \mathbf{1}_{\{N_1 = \emptyset\}} + \mathbf{1}_{\{N_2 = \emptyset\}} = \mathbf{1}_{\{N_1 \neq \emptyset\}} + \mathbf{1}_{\{N_2 \neq \emptyset\}}.$$

By the choice of v, there are at most two vertices in $N(v) \cap N(J \setminus \{v\})$; the vertex v and its parent. The vertex v is in $N(J \setminus \{v\})$ iff $N_1 \neq \emptyset$. Its parent is in $N(J \setminus \{v\})$ iff $N_2 \neq \emptyset$. $\square$

For the next lemma, we also need the following definitions. The sets of isolated points in J and connected points in J are defined by

$$\mathsf{iso}(J) = \{v \in J : N(v) \cap J = \{v\}\} \quad \text{and} \quad \mathsf{con}(J) = J \setminus \mathsf{iso}(J).$$

Lemma 2.3. *For every non-empty $J \subset V$,*

$$|N(J)| \geq \begin{cases} 1 + d|J| & \mathsf{con}(J) = \emptyset \\ 2 + d|\mathsf{iso}(J)| + (d-1)|\mathsf{con}(J)| & \mathsf{con}(J) \neq \emptyset. \end{cases}$$

Proof. Using Lemma 2.2,

$$\begin{aligned} |N(J)| &= (d-1)(|\mathsf{iso}(J)| + |\mathsf{con}(J)|) + \kappa_1(J) + \kappa_2(J) \\ &\geq (d-1)(|\mathsf{iso}(J)| + |\mathsf{con}(J)|) + (|\mathsf{iso}(J)| + \mathbf{1}_{\{\mathsf{con}(J)\neq\emptyset\}}) + 1. \end{aligned}$$ □

Proof of Proposition 2.1. The fact that (k_i) and (m_i) are decreasing is obvious. These are partitions of the same size $s = (d+1)|J|$. If $|J| = 1$ then the statement trivially holds, so we can assume $|J| > 1$. The choice of order on V implies there is $n \geq 0$ so that $B_n \subseteq B \subsetneq B_{n+1}$, where B_n is the ball of radius n. We can write

$$|J| = |B| = |B_n| + a(d-1) + c,$$

where a, c are non-negative integers so that $c < d-1$.

The tree is simple enough so that we can compute all the m_i's in terms of these:

$$\begin{aligned} & m_1 = (d-1)|J| + 2, \\ & m_2 = |J|, \\ \forall\, 3 \leq i \leq c+2 \qquad & m_i = |B_{n-1}| + a + 1, \\ \forall\, c+3 \leq i \leq d+1 \qquad & m_i = |B_{n-1}| + a. \end{aligned}$$

The case $r = 1$ of (2.1) now holds by Lemma 2.2:

$$k_1 = |N(J)| \geq (d-1)|J| + 2 = m_1.$$

The case $r = 2$ is proved as follows. If $k_2 = 0$ then $k_i = 0$ for all $i \geq 2$ and the proof is complete. On the other hand, if $k_2 \geq 1$ then by Lemma 2.3, and because $\mathsf{con}(J) \subseteq K_2(J)$,

$$k_1 + k_2 \geq (d-1)|J| + |\mathsf{iso}(J)| + \mathbf{1}_{\{\mathsf{con}(J)\neq\emptyset\}} + 1 + k_2 \geq d|J| + 2 = m_1 + m_2.$$

For $r \in \{3, 4, \ldots, d\}$, proceed by induction. Because $k_r \geq k_{r+1} \geq \ldots \geq k_{d+1}$, we have

$$k_r \geq a_r := \frac{s - \sum_{i\in[r-1]} k_i}{d-r+2}.$$

By induction,

$$\begin{aligned} \sum_{i\in[r]} k_i &\geq a_r + \sum_{i\in[r-1]} k_i \\ &= \frac{s}{d-r+2} + \frac{d-r+1}{d-r+2} \sum_{i\in[r-1]} k_i \\ &\geq \frac{s}{d-r+2} + \frac{d-r+1}{d-r+2} \sum_{i\in[r-1]} m_i \\ &= b_r + \sum_{i\in[r-1]} m_i, \end{aligned}$$

where

$$b_r := \frac{s - \sum_{i\in[r-1]} m_i}{d-r+2}.$$

If $r \le c+2$, then

$$b_r = \frac{(c+2-r+1)(|B_{n-1}|+a+1)+(d+1-c-2)(|B_{n-1}|+a)}{d-r+2}$$
$$= \frac{(d-r+2)(|B_{n-1}|+a)+c+2-r}{d-r+2},$$

and if $r > c+2$, then

$$b_r = \frac{(d-r+2)(|B_{n-1}|+a)}{d-r+2}.$$

It follows that $m_r = \lceil b_r \rceil$. All k_i's and m_i's are integers, so the desired inequality follows. □

3 Lazy random walks

The following proposition presents the key link between the isoperimetric inequality and the behavior of random walks.

Proposition 3.1. *Let $J \subset V$ be finite and let B be the quasi-ball with $|B| = |J|$. Let $k_i = |K_i(J)|$ and $m_i = |K_i(B)|$. For every distribution q,*

$$\sum_{i\in[d+1]} q^*(k_i) \le \sum_{i\in[d+1]} q^*(m_i).$$

This may seem surprising until one realizes that q^* can be any function on $\mathbb{N}$ that is increasing from 0 to 1 and is concave. The proof of Proposition 3.1 is based on the following majorization inequality, known as the Hardy–Littlewood–Pólya inequality and Karamata's inequality, a version of which was first proved by Schur; see e.g. [6, Theorem 3.C.1]. Note that the definition of the dominance order $\mu \succ \lambda$ extends verbatim to partitions of a real number with real instead of integer parts, and so this applies also for non-integer dominated sequences. In our setting, k_i and m_i are integers.

Theorem 3.2. *Let $I \subset \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be concave. For any $t \in \mathbb{N}$, if $\mu, \lambda \in I^t$ are two partitions such that $\mu \succ \lambda$, then*

$$\sum_{i\in[t]} f(\mu_i) \le \sum_{i\in[t]} f(\lambda_i).$$

Proof of Proposition 3.1. To apply Theorem 3.2 and Proposition 2.1 we need to extend q^* to a concave function. By construction, the function $q^* : \mathbb{N} \to [0,1]$ is increasing and can be written as $q^*(j) = \sum_{i\in[j]} D(i)$ where $D : \mathbb{N} \to [0,1]$ is a decreasing function. Thus extending q^* to $\mathbb{R}_+$ by a piecewise linear interpolation is increasing and concave. □

The following fact helps to establish when a distribution is greedily arranged.

Fact 3.3. *Let B be a quasi-ball and let $i \in [d+1]$. Then $K_i(B)$ is a quasi-ball.*

Proof. Write B as $B = \{v_0, v_1, \dots, v_i\}$. The choice of order on V implies there is $n \ge 0$ so that $B_n \subseteq B \subsetneq B_{n+1}$, and we can write $|B| = |B_n| + a(d-1) + c$, where a, c are non-negative integers so that $c < d-1$. Analyze the different $K_i(B)$'s as follows. The set $K_1(B) = N(B)$ contains B_{n+1} and some of the smallest elements in ∂B_{n+2}. The set $K_2(B)$ is equal to B. For $i \in \{3, \dots, c+2\}$, the set $K_i(B)$ contains B_{n-1} and the $a+1$ smallest elements in ∂B_n. For $i \in \{c+3, \dots, d+1\}$, the set $K_i(B)$ contains B_{n-1} and the a smallest elements in ∂B_n. □

We are now ready to complete the proof of our main results.

Proof of Theorem 1.4. Let $J \subset V$ and let B be the quasi-ball of the same size. For $i \in [d+1]$, let $k_i = |K_i(J)|$ and $m_i = |K_i(B)|$. We have

$$\begin{aligned} q'(J) &= \frac{1}{d+1} \sum_{i \in [d+1]} q(K_i(J)) \\ &\leq \frac{1}{d+1} \sum_{i \in [d+1]} q^*(k_i) && (3.1) \\ &\leq \frac{1}{d+1} \sum_{i \in [d+1]} q^*(m_i) && (3.2) \\ &\leq \frac{1}{d+1} \sum_{i \in [d+1]} p^*(m_i) && (3.3) \\ &= \frac{1}{d+1} \sum_{i \in [d+1]} p(K_i(B)) && (3.4) \\ &= p'(B). \end{aligned}$$

Here, the first and last equalities follow from the definition of the lazy random walk; (3.1) follows from the definition of q^*; (3.2) follows from Proposition 3.1; (3.3) holds because p majorizes q; and (3.4) follows from Fact 3.3 and the assumption that p is greedily arranged.

Let us now explain how the above implies the theorem, i.e., that p' is greedily arranged and majorizes q'. Let B be a quasi-ball. Since $q'(J) \leq p'(B)$ for any $J \subset V$ having $|J| = |B|$, we have that $(q')^*(|B|) \leq p'(B) \leq (p')^*(|B|)$. In particular, p' majorizes q'. Moreover, using this with $q = p$ yields that $p'(B) = (p')^*(|B|)$, which implies that p' is greedily arranged. □

Proof of Theorem 1.1. Theorem 1.4 implies (1.1) and in particular $p_t(B_n) \geq q_t(B_n)$ for all $n, t \geq 0$. In other words, $|Y_t|$ stochastically dominates $|X_t|$ for every $t \geq 0$. This implies that $\mathbb{E}|Y_t| \geq \mathbb{E}|X_t|$. It remains to show that $\liminf_{t\to\infty} t^{-1}|Y_t| \geq \frac{d-2}{d+1}$ almost surely. For every $\varepsilon > 0$, standard concentration bounds show that for some constants $c, C > 0$,

$$\mathbb{P}\left[t^{-1}|X_t| < \tfrac{d-2}{d+1} - \varepsilon\right] \leq Ce^{-ct}.$$

Since $|Y_t|$ stochastically dominates $|X_t|$ for every $t \geq 0$, the same holds with Y_t instead of X_t. The Borel–Cantelli lemma completes the proof. □

Remark 3.4. Theorem 1.4, and thus also Theorems 1.1 and 1.3, extend to the lazy random walk in which the probability to stay put is any $\gamma \geq \frac{1}{d+1}$. The idea is that if q'_γ is the result of a lazy random walk step applied to a distribution q with lazyness γ, then for any $\gamma > \delta$,

$$q'_\gamma = \big(\gamma - \delta\big)q + \big(1 - \gamma + \delta\big)q'_\delta.$$

We apply this with $\gamma > \delta = \frac{1}{d+1}$ to get

$$\begin{aligned} q'_\gamma(J) &= \Big(\gamma - \frac{1}{d+1}\Big)q(J) + \Big(1 - \gamma + \frac{1}{d+1}\Big)\frac{1}{d+1}\sum_{i\in[d+1]} q(K_i(J)) \\ &\leq \Big(\gamma - \frac{1}{d+1}\Big)p(B) + \Big(1 - \gamma + \frac{1}{d+1}\Big)\frac{1}{d+1}\sum_{i\in[d+1]} p(K_i(B)) = p'(B). \end{aligned}$$

4 Simple random walks

In this section, we consider simple (non-lazy) walks. The argument is similar to the lazy case, and we omit some of the details that are unchanged. For $J \subset V$, let

$$N'(J) = \bigcup_{v \in J} N'(v).$$

The main difficulty stems from the fact that the tree is bipartite. The **half-ball** B'_n is the set of the form

$$B'_n = \{v \in B_n : |v| \equiv n \mod 2\}.$$

A **half-quasi-ball** is the intersection of a quasi-ball with either V_0 or with V_1. Half-qausi-balls have parities. A **half-greedily arranged** distribution is a distribution supported on a quasi-ball.

Lemma 4.1. *For every non-empty $J \subset V$,*

$$|N'(J)| \geq 1 + (d-1)|J|.$$

Proof. First assume that J is contained in either V_0 or $V_1 = V \setminus V_0$. In this case, $\kappa_1(J) = |J|$ so that Proposition 2.2 implies that

$$|N'(J)| = |N(J)| - |J| = (d-1)|J| + \kappa_2(J) \geq (d-1)|J| + 1.$$

Second, for arbitrary J, we have $N'(v) \cap N'(w) = \emptyset$ if $|v| \neq |w| \mod 2$. The result follows by applying the above to $J \cap V_0$ and $J \cap V_1$ separately. □

For $J \subset V$ and $i \in [d]$, define

$$K'_i(J) = \{v \in V : |N'(v) \cap J| \geq i\}.$$

Fix J and let B be a half-quasi-ball of the same size. Let $k_i = |K'_i(J)|$ and $m_i = |K'_i(B)|$.

Proposition 4.2. *For any distribution q on V,*

$$\sum_{i \in [d]} q^*(k_i) \leq \sum_{i \in [d]} q^*(m_i).$$

Proof. As before, the proposition follows from Theorem 3.2 once we show that $(k_1, \ldots, k_d) \succ (m_1, \ldots, m_d)$. The fact that these are partitions is obvious, and they have the same size since $d|J| = \sum_{i \in [d]} k_i = \sum_{i \in [d]} m_i$. It remains to establish (2.1) for these partitions. Write

$$|J| = |B| = |B'_n| + a(d-1) + c,$$

where $B'_n \subset B \subsetneq B'_{n+2}$, and $a, c \geq 0$ are integers so that $c \leq d-2$. The values of the m_i's are now as follows: $m_1 = (d-1)|J| + 1$, for $2 \leq i \leq c+1$, we have $m_i = |B'_{n-2}| + a + 1$, and for $c+2 \leq i \leq d$, we have $m_i = |B'_{n-2}| + a$. The inequality $\sum_{i \in [r]} k_i \geq \sum_{i \in [r]} m_i$ for $r = 1$ follows from Lemma 4.1. For $r \geq 2$, one proceeds by induction in a similar manner as in the proof of Proposition 2.1. □

Fact 4.3. *Let B be a half-quasi-ball, and $i \in [d]$. Then, $K'_i(B)$ is a half-quasi-ball of opposite parity than B.*

Proof of Theorem 1.6. Let $J \subset V$ and let B be the half-quasi-ball of the same size as J and with opposite parity than p. Let $k_i = |K'_i(J)|$ and $m_i = |K'_i(B)|$. We have

$$\begin{aligned} q'(J) &= \frac{1}{d} \sum_{i \in [d]} q(K'_i(J)) \\ &\le \frac{1}{d} \sum_{i \in [d]} q^*(k_i) && (4.1) \\ &\le \frac{1}{d} \sum_{i \in [d]} q^*(m_i) && (4.2) \\ &\le \frac{1}{d} \sum_{i \in [d]} p^*(m_i) && (4.3) \\ &= \frac{1}{d} \sum_{i \in [d]} p(K'_i(B)) && (4.4) \\ &= p'(B), \end{aligned}$$

where the first and last equalities follow from the definition of the non-lazy random walk; (4.1) follows from the definition of q^*; (4.2) follows from Proposition 4.2; (4.3) holds because p majorizes q; and (4.4) follows from Fact 4.3 and the assumption that p is half-greedily arranged. The result follows in the same way as in the proof of Theorem 1.4. □

Proof of Theorem 1.5. Denote by p_t the distribution of S_t, and denote by q_t the distribution of Z_t. Since $d > 2$, we have $|B_n| \le 1 + d(d-1)^n \le |B'_{n+1}|$. We then have

$$\begin{aligned} q_t(B_n) &\le q_t^*(|B_n|) && (4.5) \\ &\le p_t^*(|B_n|) && (4.6) \\ &\le p_t^*(|B'_{n+1}|) && (4.7) \\ &\le p_t(B_{n+2}), && (4.8) \end{aligned}$$

where (4.5) holds by definition of q_t^*; (4.6) holds by Theorem 1.6 and induction on t; (4.7) holds because $|B_n| \le |B'_{n+1}|$; and (4.8) holds because p_t is half-greedily arranged, and because $B'_{n+1} \cup B'_{n+2} \subseteq B_{n+2}$. The rest of the proof proceeds in a similar manner as in the proof of Theorem 1.1. □

5 Exceptional times

In this section we consider the possible slow-down of a random walk on $\mathbb{Z} \cong \mathbb{T}_2$ and on $\mathbb{T}_d$ for $d > 2$. While the domination of Theorem 1.1 still applies, we ask here whether (π_t) may be chosen so that there are exceptional times where $|Y_t|$ is much smaller than $|X_t|$. We prove Theorem 1.7 on the *existence* of exceptional times of slowing down on $\mathbb{Z}$. In contrast, we prove Theorem 1.8 on the *non-existence* of such times on T_d when $d > 2$ and the permutations are restricted to automorphisms. This section is mostly independent of the previous parts of the paper.

5.1 Exceptional times for $\mathbb{Z}$

Proof of Theorem 1.7. The permutations (π_t) are all translations of $\mathbb{Z}$. Consequently, the permutations commute not just with each other but with the steps of the random walk. We shall define an integer sequence ℓ_t, and define the permutations π_t by $\pi_t \pi_{t-1} \cdots \pi_1(v) = v - \ell_t$. Thus the process $(Y_t + \ell_t)$ has the same law as the random walk (X_t). However, the coupling between the processes may not be such that $Y_t = X_t - \ell_t$, even though that is one possible coupling.

To define (ℓ_t), let $\phi(t)$ denote the integer part of $(\frac{4}{3}t\log\log t)^{1/2}$. Let $f(t)$ be a positive integer-valued non-decreasing function growing to infinity slower than $\phi(t)$. Let $(b_j)_{j=0}^{\infty}$ be defined by $b_0 = 1$ and $b_{j+1} = b_j + f(b_j)$ for all $j \geq 0$. Let (ℓ_t) be defined by ℓ_{b_j+i} is the integer part of $\phi(b_j)\cdot(\frac{4i}{f(b_j)} - 2)$ for all $j \geq 0$ and $0 \leq i < f(b_j)$. Intuitively, for each j, the numbers of the form ℓ_{b_j+i} are uniformly and densely placed in the interval between $-2\phi(b_j)$ and $2\phi(b_j)$.

Fix $\varepsilon > 0$ and consider the set T_ε of times t at which $X_t \geq (1-\varepsilon)\phi(t)$. By the law of the iterated logarithm for the lazy random walk (X_t), we have that the cardinality of T_ε is almost surely infinite. By the same law, almost surely, the set T' of times t at which $|Y_t + \ell_t| \leq 1.5\phi(t)$ contains all but finitely many positive integers.

Fix $t \in T_\varepsilon \cap T'$ sufficiently large. Let $j \geq 0$ be such that $b_j \leq t < b_{j+1}$. Since $|Y_t + \ell_t| \leq 1.5\phi(t) < 2\phi(b_j)$, there exists $0 \leq i < f(b_j)$ such that

$$|(Y_t + \ell_t) - \ell_{b_j+i}| \leq \varepsilon\phi(t).$$

At time $t' = b_j + i$, we have

$$X_{t'} \geq X_t - |t - t'| \geq (1-\varepsilon)\phi(t) - f(b_j) \geq (1-\varepsilon)\phi(t) - f(t) > 0,$$

and

$$|Y_{t'}| = |(Y_{t'} + \ell_{t'}) - \ell_{t'}| \leq |(Y_t + \ell_t) - \ell_{t'}| + |t - t'| \leq \varepsilon\phi(t) + f(t).$$

Thus,

$$|X_{t'}| - |Y_{t'}| \geq (1-2\varepsilon)\phi(t) - 2f(t) \geq (1-3\varepsilon)\phi(t').$$

We conclude that almost surely,

$$\limsup_{t\to\infty} \frac{|X_t| - |Y_t|}{\phi(t)} \geq 1.$$

Since $\limsup_{t\to\infty} \frac{|X_t|}{\phi(t)} = 1$ almost surely, we have equality above. □

5.2 No exceptional times for $d > 2$

We split the proof of Theorem 1.8 into two parts for readability, and in order to emphasize the missing piece for lifting the automorphism restriction.

Lemma 5.1. *For every $d > 2$ and every sequence of automorphisms (π_t) of $\mathbb{T}_d$, there exists a coupling of the lazy random walk process (X_t) and the permuted random walk process (Y_t) such that, almost surely,*

$$|Y_t| \geq |X_t| - 2\log t \qquad \text{for all } t \text{ large enough.}$$

Proof. Using that (π_t) consists only of automorphisms, it is not hard to check that $(\pi_t\pi_{t-1}\cdots\pi_1 X_t)$ has the same distribution as the permuted random walk process (Y_t). Thus, setting $Y_t = \pi_t\pi_{t-1}\cdots\pi_1 X_t$ describes a coupling between (X_t) and (Y_t).

To see that this coupling satisfies the claimed property, note that $|X_t| - |Y_t| > k$ implies that either $|X_t| < k$ or $(\pi_t\pi_{t-1}\cdots\pi_1)^{-1}v_0 \in T_k(X_t)$, where $T_k(x)$, defined when $|x| \geq k$, is the connected component (subtree) of $\{v \in V(\mathbb{T}) : |v| \geq k\}$ containing x. Indeed, on the complementary event $\{|X_t| \geq k\} \cap \{(\pi_t\pi_{t-1}\cdots\pi_1)^{-1}v_0 \notin T_k(X_t)\}$, we have

$$\begin{aligned} |Y_t| &= \mathrm{dist}_{\mathbb{T}}(X_t, (\pi_t\pi_{t-1}\cdots\pi_1)^{-1}v_0) && (5.1)\\ &\geq \mathrm{dist}_{\mathbb{T}}(X_t, u) && (5.2)\\ &= |X_t| - k, && (5.3)\end{aligned}$$

where u is the unique vertex in $T_k(X_t)$ with $|u| = k$. Here (5.1) follows from the fact that automorphisms preserve the graph distance, and (5.2) uses that $\mathbb{T}$ is a tree, so any path from $(\pi_t \pi_{t-1} \cdots \pi_1)^{-1} v_0$ to X_t must contain u.

Since X_t is uniform given its depth $|X_t|$, we see that

$$\mathbb{P}(|X_t| - |Y_t| > k) \le \mathbb{P}(|X_t| < k) + \frac{1}{|\partial B_k|},$$

where $\partial B_k = \{v \in V(\mathbb{T}) : |v| = k\}$ and $|\partial B_k| = d(d-1)^{k-1}$. Standard concentration bounds on the speed of (X_t) now imply that $\sum_{t=1}^{\infty} \mathbb{P}(|X_t| - |Y_t| > 2\log t) < \infty$, and the Borel–Cantelli lemma completes the proof. □

Lemma 5.2. *Let (X_t) be a non-trivial nearest-neighbor random walk on $\mathbb{Z}$ (possibly biased and with any laziness). There is a coupling of (X_t) with another copy of itself (X'_t) such that for some constant $C > 0$, almost surely,*

$$X'_t - X_t \ge \frac{\sqrt{t}}{(\log t)^C} \qquad \text{for all } t \text{ large enough.}$$

Remark 5.3. For positively biased random walks, X_t and X'_t are eventually positive so that the conclusion is equivalent to $|X'_t| - |X_t| \ge \sqrt{t}/(\log t)^C$. By interchanging the roles of (X_t) and (X'_t), the same statement is seen to hold also for negatively biased random walks. For unbiased random walks, on the other hand, it holds that $X'_t = 0$ infinitely often.

Remark 5.4. The term $\sqrt{t}/(\log t)^C$ is not optimal, but it cannot be improved to $c\sqrt{t}$. Indeed, in any coupling, the probability of the event $\{X_t \ge 0, X'_t < c\sqrt{t}\}$ is bounded from below, so that Fatou's lemma gives that $X'_t - X_t < c\sqrt{t}$ infinitely often with positive probability.

Proof. We may always couple (X_t) and (X'_t) so that they stay put at the same times (and this set of times has density less than 1). It therefore suffices to handle the non-lazy case. We thus assume that $\mathbb{P}(X_1 - X_0 = 1) = p$ and $\mathbb{P}(X_1 - X_0 = -1) = 1 - p$ for some $p \in (0, 1)$.

The main step is to construct a coupling between two Binomial (n, p) random variables B_n and B'_n such that

$$\mathbb{P}\big(B'_n - B_n \ge \tfrac{\sqrt{n}}{\log^2 n}\big) \ge c' \qquad \text{and} \qquad \mathbb{P}(B'_n < B_n) \le \tfrac{C'}{\log^2 n},$$

where $c', C' > 0$ are constants that depend on p but not on n. Let m be the integer part of $\sqrt{n}/\log^2 n$ and consider the two intervals

$$I_n = [pn - \sqrt{n}, pn] \cap \mathbb{N} \qquad \text{and} \qquad J_n = [pn - \sqrt{n}, pn - m] \cap \mathbb{N}.$$

Denote $f(i) = \mathbb{P}(B_n = i)$ and observe that $\frac{f(i)}{f(i-1)} = \frac{p}{1-p} \cdot \frac{n-i+1}{i} \ge 1$ whenever $i \le p(n+1)$. Thus, $f(i)$ is increasing for $i \in I_n$, and $f(i) \le f(i+m)$ for $i \in J_n$. It follows that there is a coupling such that

$$\begin{aligned} B_n \notin I_n &\implies B'_n = B_n, \\ B_n \in J_n &\implies B'_n = B_n + m, \\ B_n \in I_n \setminus J_n &\implies B'_n \in I_n. \end{aligned}$$

The central limit theorem implies that $\mathbb{P}(B_n \in J_n)$ converges as $n \to \infty$ to some positive constant $c = c(p)$. Since f is bounded from above by $C/\sqrt{n}$ for some constant $C = C(p)$, we have that $\mathbb{P}(B_n \in I_n \setminus J_n) \leq Cm/\sqrt{n} \leq C/\log^2 n$. This completes the construction of a coupling between B_n and B'_n with the claimed properties.

The above coupling between B_n and B'_n is relevant because X_t has the same law as $2B_t - t$. Consider the times $t_n = 2^n$ for $n \geq 1$. We construct the coupling between (X_t) and (X'_t) so that it is Markovian at these times. Fix $n \geq 1$ and suppose we have already coupled $(X_t)_{t \leq t_n}$ and $(X'_t)_{t \leq t_n}$ in some manner (the coupling for $n = 1$ can be done arbitrarily). We now describe the (conditional) coupling between the processes in the time range $(t_n, t_{n+1}]$. This coupling only depends on X_{t_n} and X'_{t_n}. The law of $(X_t - X_{t_n})_{t_n \leq t \leq t_{n+1}}$ and $(X'_t - X'_{t_n})_{t_n \leq t \leq t_{n+1}}$ is entirely independent of the past (conditioned on time t_n). These are two random walks of length $t_{n+1} - t_n = 2^n$, which we denote by $(S_i)_{i=0}^{2^n}$ and $(S'_i)_{i=0}^{2^n}$. To couple these walks, we first couple the endpoints S_{2^n} and S'_{2^n} using the above coupling between B_{2^n} and B'_{2^n} (pushed forward by the map $x \mapsto 2x - 2^n$). Given the endpoints, we couple the walks so that $S'_i \geq S_i$ for all i when $S'_{2^n} \geq S_{2^n}$, and arbitrarily otherwise. The former can be done by first sampling (S_i) and then uniformly choosing $\frac{1}{2}(S'_{2^n} - S_{2^n})$ coordinates i among those where the increment $S_i - S_{i+1}$ is -1 and setting the corresponding increments $S'_i - S'_{i-1}$ to $+1$ there (with all other increments remaining the same for both). This completes the description of the coupling between (X_t) and (X'_t).

It remains to check that the constructed coupling has the claimed property. Let $\Delta_n = X_{t_{n+1}} - X_{t_n}$ and $\Delta'_n = X'_{t_{n+1}} - X'_{t_n}$. Define events

$$E_n = \{\Delta'_n - \Delta_n \geq 2^{n/2}/n^2\} \qquad \text{and} \qquad F_n = \{\Delta'_n < \Delta_n\}.$$

Since F_n has probability at most C'/n^2, only finitely many of the F_n occur almost surely. Let N_1 be the smallest positive integer such that F_n does not occur for any $n \geq N_1$. Observe that $X'_t - X_t$ is non-decreasing for $t \geq t_{N_1}$. Since $\{E_n\}_{n=1}^{\infty}$ are independent events, each of probability at least c', infinitely many of them occur almost surely. Moreover, almost surely, for any n large enough, at least one of $E_{n-1}, E_{n-2}, \ldots, E_{n-C''\log n}$ occurs, where $C'' > 0$ is some large constant. Let N_2 be the smallest positive integer so that this holds for $n \geq N_2$. Observe that if $n - C''\log n \geq \max\{N_1, N_2\}$ and $t_n \leq t \leq t_{n+1}$, then letting $n - C''\log n \leq m < n$ be such that E_m occurs, we obtain that

$$\begin{aligned} X'_t - X_t &\geq X'_{t_{m+1}} - X_{t_{m+1}} \\ &= \Delta'_m - \Delta_m + X'_{t_m} - X_{t_m} \\ &\geq 2^{m/2}/m^2 + X'_{t_N} - X_{t_N} \\ &\geq \sqrt{t} \cdot e^{-5C\log\log t}, \end{aligned}$$

where the last inequality holds for t large enough. □

Proof of Theorem 1.8. Let (X'_t) denote a copy of the lazy random walk (X_t). By the first lemma, (X'_t) and (Y_t) can be coupled so that $|X'_t| - |Y_t| \leq 2\log t$ eventually. By the second lemma (and the first remark following it), the walks $(|X_t|)$ and $(|X'_t|)$ can be coupled so that $|X'_t| - |X_t| \geq \sqrt{t}/\log^C t$ eventually. Extend this coupling to a coupling of (X_t) and (X'_t). The processes (X_t) and (Y_t) are now coupled so that $|Y_t| - |X_t| \geq \sqrt{t}/\log^C t - 2\log t \geq \sqrt{t}/\log^{2C} t$ eventually. □

Removing the automorphism assumption in Theorem 5.1, even at the expense of increasing the upper bound on $|X_t| - |Y_t|$ from $2\log t$ to $\sqrt{t}/(\log t)^C$ for a sufficiently large constant C, would suffice in order to lift the automorphism assumption in Theorem 1.8.

References

[1] Ganguly, S. & Peres, Y. Permuted random walk exits typically in linear time. *2014 Proceedings Of The Eleventh Workshop On Analytic Algorithmics And Combinatorics (ANALCO)*, pp. 74-81 (2014) MR3248355

[2] Pymar, R. & Sousi, P. A permuted random walk exits faster. *ArXiv Preprint arXiv:1304.6704*. (2013) MR3225972

[3] Chatterjee, S. & Diaconis, P. Correction to: Speeding up Markov chains with deterministic jumps. *Probability Theory And Related Fields*. **181**, 377-400 (2021) MR4341077

[4] Gouëzel, S. Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunisian Journal Of Mathematics*. **4**, 635-671 (2023) MR4533553

[5] Chatterjee, S. & Diaconis, P. Speeding up Markov chains with deterministic jumps. *Probability Theory And Related Fields*. **178** (2020) MR4168397

[6] Marshall, A., Olkin, I. & Arnold, B. Inequalities: theory of majorization and its applications. (Springer, 1979) MR0552278

[7] Morris, B. & Peres, Y. Evolving sets, mixing and heat kernel bounds. *Probability Theory And Related Fields*. **133**, 245-266 (2005) MR2198701

[8] Lyons, R. & Peres, Y. Probability on trees and networks. (Cambridge University Press, 2017) MR3616205

[9] Virág, B. On the speed of random walks on graphs. *The Annals Of Probability*. **28**, 379-394 (2000) MR1756009

[10] Virág, B. Anchored expansion and random walk. *GAFA*. **10**, 1588-1605 (2000) MR1810755

Acknowledgments. OA would like to thank the American Institute of Math, where this project was initiated, and the Technion, where the collaboration began.