THE SPEED OF A GENERAL RANDOM WALK REINFORCED BY ITS RECENT HISTORY

ROSS G. PINSKY

ABSTRACT. We consider several variants of a class of random walks whose increment distributions depend on the average value of the process over its most recent N steps. We investigate the speed of the process, and in particular, the limiting speed as the "history window" $N \to \infty$.

1. Introduction and Statement of Results

Over the past couple of decades, many papers have been devoted to the study of edge or vertex reinforced random walks and excited (also known as "cookie") random walks on $\mathbb{Z}$. These processes have a simple underlying transition mechanism—such as simple symmetric random walk—but this mechanism is "reinforced" or "excited" depending on the location of the random walk and its complete history at that location. For survey papers which include many references, see [4] and [3].

In this paper, we consider random walks on $\mathbb{R}$ with a simpler and very natural mechanism for reinforcement; namely, the reinforcement is catalyzed by the behavior of the random walk path over a bounded interval of its history, irrespective of its present location. In fact, we will define three versions of such a process, and analyze them to varying degrees. To define these processes, let $N, l \in \mathbb{N}$ with $l \le N$, let $\{P_i^{(\text{inc})}\}_{i=0}^l$ be probability measures on $\mathbb{R}$ with finite expectations $\mu_i = \int_{-\infty}^{\infty} x P_i^{(\text{inc})}(dx)$, and let $\{m_i\}_{i=1}^l$ be a sequence. We make the following assumption.

2010 *Mathematics Subject Classification.* 60F10, 60K35, 60J10.

Key words and phrases. random walk with reinforcement, Cramér's theorem, Legendre-Fenchel transform.

Assumption A. The sequence $\{\mu_i\}_{i=0}^l$ of the expectations corresponding to the measures $\{P_i^{(\text{inc})}\}_{i=0}^l$ is strictly increasing, and the sequence $\{m_i\}_{i=1}^l$ satisfies

$$m_i < \mu_i < m_{i+1},\ i = 1, \cdots, l-1;$$

$$\mu_0 < m_1 \text{ and } m_l < \mu_l.$$

In our notation for the processes, we suppress the dependence on all the above parameters with the exception of N. One version, the *instantaneous* version, will be denoted by $\{X_n^{N;I}\}_{n=0}^\infty$, a second version, the *delayed* version, will be denoted by $\{X_n^{N;D}\}_{n=0}^\infty$, and the third version, the *one-step delayed* version, will be denoted by $\{X_n^{N;D_1}\}_{n=0}^\infty$. Most of this paper will concern the delayed and the one-step delayed versions, but we will define the instantaneous version first, because this will make it easier to describe the other versions. For convenience, define $m_0 = -\infty$ and $m_{l+1} = +\infty$.

The instantaneous version $\{X_n^{N;I}\}_{n=0}^\infty$ is defined as follows. Let $X_0^{N;I} = 0$ and let $\{X_n^{N;I}\}_{n=1}^N$ be distributed like a random walk with increment distribution $P_{i_0}^{(\text{inc})}$, for some i_0. At each time $n \ge N+1$, the process looks back at its most recent N steps. If the average value of those steps falls in the range $[m_i, m_{i+1})$, then the process jumps with increment distribution $P_i^{(\text{inc})}$, $i = 0, \cdots, l$. That is, for $n \ge N$ and $i = 0, \cdots, l$,

$$P\big(X_{n+1}^{N;I} - X_n^{N;I} \in \cdot \mid \frac{1}{N}\big(X_n^{N;I} - X_{n-N}^{N;I}\big) \in [m_i, m_{i+1})\big) = P_i^{(\text{inc})}(\cdot). \tag{1.1}$$

The delayed version $\{X_n^{N;D}\}_{n=0}^\infty$ is defined similarly, the only difference being that this process is required to use any particular jump distribution at least N consecutive times, thereby insuring that the reinforcement that causes the process to switch from one increment distribution, say i, to another increment distribution is due to the behavior of the process while in the i regime. Thus, $\{X_n^{N;D}\}_{n=0}^N$ is defined identically to $\{X_n^{N;I}\}_{n=0}^N$, and for each time $n \ge N+1$, if the jump distribution used for $X_n^{N;D} - X_{n-1}^{N;D}$ was not used for $X_{n-N+1}^{N;D} - X_{n-N}^{N;D}$, then it is automatically used again for $X_{n+1}^{N;D} - X_n^{N;D}$, while otherwise the jump distribution for $X_{n+1}^{N;D} - X_n^{N;D}$ is determined by (1.1) (with $X^{N;I}$ replaced by $X^{N;D}$).

Finally the one-step delayed version $\{X_n^{N;D_1}\}_{n=0}^{\infty}$ can be thought of as the $\{X_n^{N;D}\}_{n=0}^{\infty}$ process with the restriction that the jump mechanism can only make nearest neighbor switches; that is, the jump mechanism can only make switches from i to $i\pm1$. More precisely, $\{X_n^{N;D_1}\}_{n=0}^{N}$ is defined identically to $\{X_n^{N;D}\}_{n=0}^{N}$, and for each time $n \geq N+1$, if the jump distribution used for $X_n^{N;D_1} - X_{n-1}^{N;D_1}$ was not used for $X_{n-N+1}^{N;D_1} - X_{n-N}^{N;D_1}$, then it is automatically used again for $X_{n+1}^{N;D_1} - X_n^{N;D_1}$. Otherwise, if the jump mechanism used for $X_n^{N;D_1} - X_{n-1}^{N-1;D_1}$ was j, and $\frac{1}{N}\big(X_n^{N;D_1} - X_{n-N}^{N;D_1}\big) \in [m_i, m_{i+1})$, then the jump mechanism used for $X_{n+1}^{N;D_1} - X_n^{N-1;D_1}$ is j, if $i=j$, is $j+1$, if $i>j$, and is $j-1$, if $i<j$.

Note that when $l=1$, that is, when there is only one threshold level, the delayed version and the one-step delayed version coincide.

We call each version of the process a random walk reinforced by its recent history. All three versions are natural models for the fortunes of various economic commodities, such as stocks, or for the popularity of various social trends, which respond positively to recent success and negatively to recent failure.

We call N the *history window* and $\{m_i\}_{i=1}^{l}$ the *threshold* levels. In Assumption C below, we specify a simple condition to ensure that the processes will almost surely jump an infinite number of times according to each of the $l+1$ increment distributions.

In this paper, we investigate the speeds of these processes. It's rather easy to show that the speeds exist almost surely and are almost surely constant,

Proposition 1. *Let Assumptions A and Assumption C (given below) hold. Then the speeds*

$$s^I(N, m_1, \cdots, m_l) := \lim_{n\to\infty} \frac{X_n^{N;I}}{n}, \quad s^D(N, m_1, \cdots, m_l) := \lim_{n\to\infty} \frac{X_n^{N;D}}{n}$$
$$\textit{and } s^{D_1}(N, m_1, \cdots, m_l) := \lim_{n\to\infty} \frac{X_n^{N;D_1}}{n}$$

exist almost surely and are almost surely constant.

Our main result concerns the limiting speed as the history window $N \to \infty$, with threshold levels increasing in an appropriate way, as we now specify. Relabeling the thresholds $\{m_i\}_{i=1}^l$ appearing in the definition of the processes by $\{m_i^{(N)}\}_{i=1}^l$, we make the following assumption.

Assumption B.

$$\begin{aligned}
&\lim_{N\to\infty} m_i^{(N)} = r_i,\ \ i = 1,\cdots,l;\\
&r_i < \mu_i < r_{i+1},\ \ i = 1,\cdots,l-1;\\
&\mu_0 < r_1, \quad r_l < \mu_l.
\end{aligned}$$

Here is the condition we impose to ensure that the processes will almost surely jump an infinite number of times according to each of the $l+1$ increment distributions.

Assumption C.

$$\begin{aligned}
&P_i^{(\text{inc})}\big((-\infty, r_i)\big) > 0 \text{ and } P_i^{(\text{inc})}\big([r_{i+1}, \infty)\big) > 0,\ \text{ for } i = 1,\cdots,l-1;\\
&P_0^{(\text{inc})}([r_1,\infty)) > 0, \quad P_l^{(\text{inc})}((-\infty, r_l)) > 0.
\end{aligned}$$

(Assumption C is a bit stronger than necessary to ensure that the process will almost surely jump an infinite number of times according to each of the $l+1$ increment distributions, but we use it so as to simplify the exposition.)

A key technical tool that will be used is Cramér's large deviations theorem for the empirical mean of an iid sequence. In order to have this at our disposal, we need to make a two-sided exponent moment assumption on the increment distributions $\{P_i^{(\text{inc})}\})_{i=0}^l$. Let

$$M_{P_i^{(\text{inc})}}(t) = \int_{-\infty}^{\infty} e^{tx} P_i^{(\text{inc})}(dx)$$

denote the moment generating function of the distribution $P_i^{(\text{inc})}$.

Assumption D. There exists a $t_0 > 0$ such that $M_{P_i^{(\text{inc})}}(\pm t_0) < \infty$, for $i = 0, 1, \cdots, l$.

Let $I_i(r)$ denote the Legendre-Fenchel transformation for the distribution $P_i^{(\mathrm{inc})}$, defined by

$$I_i(r)=\sup_{\lambda\in\mathbb{R}}\big(\lambda r-\log M_{P_i^{(\mathrm{inc})}}(\lambda)\big),\ r\in\mathbb{R}. \tag{1.2}$$

We recall several facts about I_i that we will need and that hold under Assumption D [1].

$$\begin{aligned}&I_i(r)<\infty \text{ if and only if either } r\le\mu_i \text{ and } P_i^{(\mathrm{inc})}(-\infty,r])>0,\ \text{ or}\\ &r>\mu_i \text{ and } P_i^{(\mathrm{inc})}([r,\infty))>0.\end{aligned} \tag{1.3}$$

Let $x_i^+=\sup\{x\in\mathbb{R}:I_i(x)<\infty\}$ and $x_i^-=\inf\{x\in\mathbb{R}:I_i(x)<\infty\}$. Then

$$\begin{aligned}&I_i(\mu_i)=0;\\ &I_i:[\mu_i,x_i^+)\to[0,\infty) \text{ is continuous and strictly increasing;}\\ &I_i:(x_i^-,\mu_i]\to[0,\infty) \text{ is continuous and strictly decreasing.}\end{aligned} \tag{1.4}$$

And we recall an elementary large deviations result, a version of Cramér's theorem, that holds under Assumption D [1]: if $S_n^{(i)}$ is the sum of n iid random variables distributed as $P_i^{(\mathrm{inc})}$, and $P_i^{(\mathrm{inc})}$ satisfies Assumption D, then

(1.5)

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\log P(\frac{S_n^{(i)}}{n}\ge r)=\lim_{n\to\infty}\frac{1}{n}\log P(\frac{S_n^{(i)}}{n}>r)=-I_i(r),\ \mu_i\le r<x_i^+;\\ &\lim_{n\to\infty}\frac{1}{n}\log P(\frac{S_n^{(i)}}{n}\le r)=\lim_{n\to\infty}\frac{1}{n}\log P(\frac{S_n^{(i)}}{n}<r)=-I_i(r),\ x_i^-<r\le\mu_i.\end{aligned}$$

Our main result gives the limiting speed as $N\to\infty$ for the delayed version and for the one-step delayed version. For the one-step delayed version, we need no more assumptions, but for the delayed version we need to make one additional assumption.

Assumption E. Let $l\ge2$ and let $0\le i,j\le l$. If $i\le j-2$, then

$$\sum_{k=i}^{j-1}I_k(r_{k+1})<I_i(r_j).$$

If $i \geq j+2$, then

$$\sum_{k=j+1}^{i} I_k(r_k) < I_i(r_j).$$

Remark. As we shall see, the import of Assumption E is that it guarantees that for $|j-i| \geq 2$, the probability that the jump mechanism goes directly from regime $P_i^{(\text{inc})}$ to regime $P_j^{(\text{inc})}$ is negligible compared to the probability that the jump mechanism goes from regime $P_i^{(\text{inc})}$ to regime $P_j^{(\text{inc})}$ by first passing through each of the regimes $P_k^{(\text{inc})}$ for k between i and j.

Assumption E is satisfied by many natural families of distributions.

Proposition 2. *Assumption E is satisfied if the $l+1$ jump distributions $\{P_i^{(inc)}\}_{i=0}^{l}$ are all of one of the following types:*
i. Gaussian distributions with a common variance;
ii. Exponential distributions;
iii. Poisson distributions;
iv. Geometric distributions;
v. Bernoulli distributions.

We can now state the main result.

Theorem 1. *Let Assumptions A,B,C, and D hold. Define*

$$\Lambda_i = I_i(r_{i+1}) + \sum_{k=1}^{i} \big(I_k(r_k) - I_k(r_{k+1})\big),\ 0 \leq i \leq l.$$

i. If $\max_{0\leq i\leq l} \Lambda_i$ occurs uniquely at $i=i_0$, then the speed $s^{D_1}(N, m_1^{(N)}, \cdots, m_l^{(N)})$ of the one-step delayed process $\{X_n^{N;D_1}\}_{n=0}^{\infty}$ satisfies

$$\lim_{N\to\infty} s^{D_1}(N, m_1^{(N)}, \cdots, m_l^{(N)}) = \mu_{i_0}.$$

ii. Let $l \geq 2$ and assume in addition that Assumption E holds. If $\max_{0\leq i\leq l} \Lambda_i$ occurs uniquely at $i=i_0$, then the speed $s^{D}(N, m_1^{(N)}, \cdots, m_l^{(N)})$ of the delayed process $\{X_n^{N;D}\}_{n=0}^{\infty}$ satisfies

$$\lim_{N\to\infty} s^{D}(N, m_1^{(N)}, \cdots, m_l^{(N)}) = \mu_{i_0}.$$

Example. The Legendre-Fenchel transformation of the Gaussian distribution $N(\mu,\sigma^2)$ is given by $I(r)=\frac{(r-\mu)^2}{2\sigma^2}$ (see section 6).
a. Let $P_i^{(\text{inc})}\sim N(\mu_i,\sigma^2)$. If

$$\arg\max_{i\in\{0,\cdots,l\}}\Big[(r_{i+1}-\mu_i)^2+\sum_{k=1}^{i}\Big((\mu_k-r_k)^2-(r_{k+1}-\mu_k)^2\Big)\Big]$$

occurs uniquely at i_0, then the limiting speed for the delayed or for the one-step delayed version is μ_{i_0}.
b. Let $P_i^{(\text{inc})}\sim N(\mu_i,\sigma_i^2)$. If

$$\arg\max_{i\in\{0,\cdots,l\}}\Big[\frac{(r_{i+1}-\mu_i)^2}{\sigma_i^2}+\sum_{k=1}^{i}\frac{(\mu_k-r_k)^2-(r_{k+1}-\mu_k)^2}{\sigma_k^2}\Big]$$

occurs uniquely at i_0, then the limiting speed for the one-step delayed version is μ_{i_0}.

In the instantaneous version, the passage from one regime, say i, to a neighboring regime, say $i+1$, will frequently be accompanied by a number of short time oscillations between the two regimes before the process securely ensconces itself in the new regime $i+1$. Because of technical difficulties related to these oscillations, we can only prove a theorem for the limiting speed of the instantaneous version in the case $l=1$.

Theorem 2. *Let $l=1$ and let Assumptions A,B,C, and D hold. The speed $s^I(N,m_1^{(N)})$ of the instantaneous process $\{X_n^{N;I}\}_{n=0}^{\infty}$ satisfies*

$$\lim_{N\to\infty}s^I(N,m_1^{(N)})=\begin{cases}\mu_0, & \textit{if } I_0(r_1)>I_1(r_1);\\ \mu_1, & \textit{if } I_1(r_1)>I_0(r_1).\end{cases}$$

Remark. Note that the result in Theorem 2, which treats the instantaneous version with $l=1$, looks the same as part (i) of Theorem 1 when $l=1$ (because $\Lambda_0=I_0(r_1)$ and $\Lambda_1=I_1(r_1)$). We expect that part (ii) of Theorem 1. which treats the delayed version with $l\geq 2$, also holds for the instantaneous version with $l\geq 2$.

In the instantaneous version, define the N-dimensional differences process $\{Z_n^{N;I}\}_{n=0}^{\infty}$ by

$$Z_n^{N;I} = (X_{n+1}^{N;I} - X_n^{N;I}, X_{n+2}^{N;I} - X_{n+1}^{N;I}, \cdots, X_{n+N}^{N;I} - X_{n+N-1}^{N;I}).$$

It is easy to see that this is a Markov process. In [5] we studied the speed of the instantaneous version $\{X_n^{N;I}\}_{n=0}^{\infty}$ under the assumption that the increment distributions $\{P_i^{(\text{inc})}\}_{i=0}^{l}$ are all Bernoulli distributions on $\{-1,1\}$; $P_i^{(\text{inc})} \sim \text{Ber}(p_i)$, so $\mu_i = 2p_i - 1$. Thus, those processes lived on $\mathbb{Z}$ and made only nearest-neighbor jumps. In that version, we were able to calculate explicitly the invariant measure π^N (defined on $\{-1,1\}^N$) of the differences process $\{Z_n^{N;I}\}_{n=0}^{\infty}$, and this allowed us to obtain an explicit formula for the speed $s^I(N, m_1, \cdots, m_l)$. What made the explicit calculation of the invariant distribution possible was the fact that π^N turned out to be constant on the level sets $\{z \in \{-1,1\}^N : \sum_{i=1}^N z_i = M\}$, for any M. Even in the case that the increment distributions $\{P_i^{(\text{inc})}\}_{i=0}^{l}$ are all supported on a fixed set of size three, the explicit calculation of the invariant measure π^N of the differences process does not seem possible in general. Exploiting this explicit formula for the speed $s^I(N, m_1^{(N)}, \cdots, m_l^{(N)})$ in the case of Bernoulli increment distributions, in [5] we proved the equivalent of Theorem 1 for the instantaneous version. (Recall from Proposition 2, that Assumption E is satisfied for Bernoulli distributions.) The expressions $\{\Lambda\}_{i=0}^{l}$ in the case of these Bernoulli distributions appear there in explicit form, but their connection to the Legendre-Fenchel transformation is not mentioned. The delicate borderline cases, when $\max_{0 \le i \le l} \Delta_i$ does not occur uniquely were also resolved, in each case of which the limiting speed was a certain linear combination of the speeds $\{\mu_i\}_{i=0}^{l}$. In this paper, we work on exponential scale, via (1.5), so we cannot handle the borderline cases.

We now turn to the organization of the rest of the paper. Theorem 1 is proved very quickly in section 3, but this is only after quite a number of technical propositions are proved in the long section 2. Here is a rough outline of the idea of the proof. Let $\{Y_m^{N;D}\}_{m=0}^{\infty}$ denote the Markov chain that follows the changes of the increment distribution utilized by the delayed version $\{X_n^{N;D}\}_{n=0}^{\infty}$ of the random walk reinforced by its recent history. Thus, $Y_0^{N;D} = i_0$, since the process $\{X_n^{N;D}\}_{n=0}^{\infty}$ starts out using the increment distribution $P_{i_0}^{(\mathrm{inc})}$. If the first time the process $\{X_n^{N;D}\}_{n=0}^{\infty}$ changes its increment distribution, it switches from distribution $P_{i_0}^{(\mathrm{inc})}$ to distribution $P_j^{(\mathrm{inc})}$, then $Y_1^{N;D} = j$. In general, $Y_m^{N;D} = k$, if after switching increment distribution m times, the process $\{X_n^{N;D}\}_{n=0}^{\infty}$ is using the increment distribution $P_k^{(\mathrm{inc})}$. Similarly, let $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$ denote the Markov chain that follows the changes of the increment distribution utilized by the one-step delayed version $\{X_n^{N;D_1}\}_{n=0}^{\infty}$ of the random walk reinforced by its recent history. (These Markov chains are introduced after the proof of Proposition 5 in section 2.) Propositions 3-5, the first three propositions of section 2, are the key technical results that are used to prove Proposition 6, which gives tight exponential estimates as $N \to \infty$ on the transition probabilities of the Markov chains $\{Y_m^{N;D}\}_{m=0}^{\infty}$ and $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$. Since $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$ is a birth and death chain, its invariant distribution can be written down explicitly in terms of its transition probabilities; thus we obtain tight exponential estimates on the behavior of this invariant measure as $N \to \infty$. Proposition 7 shows that under Assumption E, as $N \to \infty$ the invariant measure for $\{Y_m^{N;D}\}_{m=0}^{\infty}$ has the same exponential asymptotic as that of $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$. Its rather long proof is postponed to section 5. Propositions 8 and 9 calculate respectively the exponential order as $N \to \infty$ of the expected number of steps made by and the expected distance travelled by the delayed version or the one-step delayed version of the random walk reinforced by its recent history between the time it enters a particular increment distribution regime until it switches to a different increment distribution regime. The proof of Theorem 1 in section 3 follows easily from Propositions 8 and 9 along with the asymptotic behavior of the invariant measures for the Markov chains $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$ and $\{Y_m^{N;D}\}_{m=0}^{\infty}$. The proof of Proposition 1 for the delayed version and for the one-step delayed version is embedded in the proof of Theorem 1. The proof of Theorem 2 is given in section 4, and embedded in that proof is the proof of Proposition 1 for the instantaneous version. Finally, the proof of Proposition 2 is given in section 6.

2. A series of propositions

We will use the following notation throughout the paper.

$$a_N \approx b_N \text{ means } \lim_{N\to\infty}\frac{1}{N}\log a_N = \lim_{N\to\infty}\frac{1}{N}\log b_N;$$
$$a_N \lessapprox b_N \text{ means } \limsup_{N\to\infty}\frac{1}{N}\big(\log a_N - \log b_N\big) \le 0;$$
$$a_N \lessapprox\!\lessapprox b_N \text{ means } \limsup_{N\to\infty}\frac{1}{N}\big(\log a_N - \log b_N\big) < 0.$$

The random walk with increment distribution $P_i^{(\text{inc})}$ will be denoted by $\{S_n^{(i)}\}_{n=0}^{\infty}$. Also, we will use the notation

$$S_{j,k}^{(i)} = S_k^{(i)} - S_j^{(i)}, \text{ for } 0 \le j < k.$$

In order to reduce the cumbersome notation, we define as follows $Z_n^{N,i;j,k}$, for $n \geq 1$ and $1 \leq i \leq l-1$, with $1 \leq j \leq i$ and $i+1 \leq k \leq l$:

(2.1)

$$Z_n^{N,i;j,k} = 1, \text{ if}$$

$$\max\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) \geq r_k \text{ and}$$

$$\min\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) \geq r_j;$$

$$Z_n^{N,i;j,k} = -1, \text{ if}$$

$$\max\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) < r_k \text{ and}$$

$$\min\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) < r_j;$$

$$Z_n^{N,i;j,k} = -11, \text{ if}$$

$$\max\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) \geq r_k \text{ and}$$

$$\min\Big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\Big) < r_j;$$

$$Z_n^{N,i;j,k} = 0, \text{ otherwise.}$$

Note that $\{Z_n^{N,i;j,k}\}_{n=1}^\infty$ are identically distributed, and that each of $\{Z_{2n}^{N,i;j,k}\}_{n=1}^\infty$ and $\{Z_{2n-1}^{N,i;j,k}\}_{n=1}^\infty$ is an independent sequence.

We begin with three key propositions with rather involved proofs. These propositions serve as a basis for the rest of the results in this section.

Proposition 3. *Let* $1 \leq i \leq l-1,\ 0 \leq j \leq i,\ i+1 \leq k \leq l$. *Then*

$$\begin{aligned} &P(Z_1^{N,i;j,k} = 1) \approx e^{-NI_i(r_k)}; \\ &P(Z_1^{N,i;j,k} = -1) \approx e^{-NI_i(r_j)}; \qquad (2.2) \\ &P(Z_1^{N,i;j,k} = -11) \lessapprox e^{-N\big(I_i(r_j)+I_i(r_k)\big)}. \end{aligned}$$

Proof. We will prove the first and third formulas in (2.2); the second one is proved analogous to the first. For the first formula, we may assume that

$I_i(r_k)<\infty$, since otherwise, by (1.3), the formula clearly holds. By (1.5), we have

$$P(Z_1^{N,i;j,k}=1)\le P\big(\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\le$$
$$\sum_{j=0}^{N-1}P(\frac{S^{(i)}_{j,N+j}}{N}\ge r_k)\approx Ne^{-NI_i(r_k)}\approx e^{-NI_i(r_k)}. \tag{2.3}$$

Also

(2.4)

$$P(Z_1^{N,i;j,k}=1)=$$
$$P\big(\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\times$$
$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_j|\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big).$$

By (1.5),

(2.5)

$$P\big(\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\ge P(\frac{S^{(i)}_{0,N}}{N}\ge r_k)\approx e^{-NI_i(r_k)}.$$

Using the clearly positive correlations for the first inequality below, and using (1.5) for the approximation below, we have

(2.6)

$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_j|\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\ge$$
$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_j\big)=$$
$$1-P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)<r_j\big)\ge 1-NP(\frac{S^{(i)}_{0,N}}{N}<r_j)\approx 1,\ \text{as } N\to\infty.$$

The first formula in (2.2) now follows from (2.3)-(2.6).

We now turn to the third formula in (2.2). We have

(2.7)

$$P(Z^{N,i;j,k}=-11)=P\big(\max\big(\frac{S^{(i)}_{0,N}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\times$$
$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)<r_j|\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big).$$

By (1.5),

(2.8)

$$P\big(\max\big(\frac{S^{(i)}_{0,N}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\le NP\big(\frac{S^{(i)}_{0,N}}{N}\ge r_k\big)\approx e^{-NI_i(r_k)}.$$

By the clearly negative correlations, and by (1.5),

(2.9)

$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)<r_j|\max\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)\ge r_k\big)\le$$
$$P\big(\min\big(\frac{S^{(i)}_{0,N}}{N},\frac{S^{(i)}_{1,N+1}}{N},\cdots,\frac{S^{(i)}_{N-1,2N-1}}{N}\big)<r_j\big)\le NP\big(\frac{S^{(i)}_{0,N}}{N}\le r_j\big)\approx e^{-NI_i(r_j)}.$$

The third formula in (2.2) follows from (2.7)-(2.9). □

Proposition 4. *Let $1\le i\le l-1,\ 0\le j\le i,\ i+1\le k\le l$, and assume that either $j=i$ or $k=i+1$. Define*

$$\tau^{N,i;j,k}=\inf\big\{n\ge 0:\frac{S^{(i)}_{n,N+n}}{N}\notin[r_j,r_k)\big\}. \tag{2.10}$$

Then

$$\begin{aligned}&P\big(\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}<r_j\big)\approx e^{-N\big(I_i(r_j)-I_i(r_k)\big)^+};\\&P\big(\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}\ge r_k\big)\approx e^{-N\big(I_i(r_k)-I_i(r_j)\big)^+}.\end{aligned} \tag{2.11}$$

Proof. By Assumption C and the assumption that either $j=i$ or $k=i+1$, it follows that $\tau^{N,i;j,k}<\infty$ a.s. Also, by Assumption C and (1.3), at least one of $I_i(r_k)$ and $I_i(r_j)$ is finite, so $(I_i(r_j)-I_i(r_k))^+\in[0,\infty]$ is well-defined. Assume without loss of generality that $I_i(r_j)\ge I_i(r_k)$. If $I_i(r_j)>I_i(r_k)$, then it suffices to prove the first formula in (2.11) since the two terms on the left hand side of (2.11) add up to one. If $I_i(r_j)=I_i(r_k)$, then the proofs of the two formulas in (2.11) are almost identical. Thus, in this case too we will prove only the first formula. Suppressing the dependence on N, let

(2.12)

$$\sigma^{(e)}_{i;j,k}=\inf\big\{2n\ge 2:Z^{N,i;j,k}_{2n}\ne 0\big\},\ \ \sigma^{(o)}_{i;j,k}=\inf\big\{2n-1\ge 1,Z^{N,i;j,k}_{2n-1}\ne 0\big\}.$$

Using Proposition 3 and the fact that each of $\{Z_{2n}^{N,i;j,k}\}_{n=1}^{\infty}$ and $\{Z_{2n-1}^{N,i;j,k}\}_{n=1}^{\infty}$ is an iid sequence, it follows that

$$\begin{aligned} &P(Z_{\sigma_{i;j,k}^{(*)}}^{N,i;j,k}=-1)\approx e^{-N\big(I_i(r_j)-I_i(r_k)\big)},\\ &P(Z_{\sigma_{i;j,k}^{(*)}}^{N,i;j,k}=-11)\lessapprox e^{-NI_i(r_j)},\\ &\text{both when } \sigma_{i;j,k}^{(*)}=\sigma_{i;j,k}^{(e)} \text{ and when } \sigma_{i;j,k}^{(*)}=\sigma_{i;j,k}^{(o)}. \end{aligned} \tag{2.13}$$

Now

$$\Big\{\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}<r_j\Big\}\subset\big\{Z^{N,i,j,k}_{\sigma^{(e)}_{i;j,k}}\in\{-1,-11\}\big\}\cup\big\{Z^{N,i,j,k}_{\sigma^{(0)}_{i;j,k}}\in\{-1,-11\}\big\};$$

thus, it follows from (2.13) that

$$P\big(\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}<r_j\big)\lessapprox e^{-N\big(I_i(r_j)-I_i(r_k)\big)}. \tag{2.14}$$

To prove an inequality in the other direction, let $a_N=P(Z_1^{(N,i;j,k)}=-1)$ and $b_N=P(Z_1^{(N,i;j,k)}\in\{1,-11\})$, where we have suppressed the dependence on i,j,k. From Proposition 3,

$$a_N\approx e^{-NI_i(r_j)},\quad b_N\approx e^{-NI_i(r_k)}. \tag{2.15}$$

We have for any positive integer M,

$$\big\{\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}<r_j\Big\}\supset\big(\cap_{n=1}^{2M}\{Z_n^{N,i;j,k}\in\{0,-1\}\}\big)\cap\big(\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k}=-1\}\big). \tag{2.16}$$

Thus,

$$\begin{aligned} P(\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N}<r_j)\ge P(\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k}=-1\})\times\\ P\big(\cap_{n=1}^{2M}\{Z_n^{N,i;j,k}\in\{0,-1\}\}\big|\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k}=-1\}\big). \end{aligned} \tag{2.17}$$

Since $\{Z_{2n}^{N,i;j,k}\}_{n=1}^{M}$ are iid, it follows that

$$P(\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k}=-1\})=1-(1-a_N)^M. \tag{2.18}$$

From the definitions, it follows that

$$\{Z_n^{N,i;j,k}\in\{0,-1\}\}=\cap_{m=(n-1)N}^{nN-1}\{\frac{S^{(i)}_{m,N+m}}{N}<r_k\} \tag{2.19}$$

and

(2.20)

$$\{Z_{2n}^{N,i;j,k} = -1\} = \big(\cap_{m=(2n-1)N}^{2nN-1}\{\frac{S_{m,N+m}^{(i)}}{N} < r_k\}\big)\cap\big(\cup_{m=(2n-1)N}^{2nN-1}\{\frac{S_{m,N+m}^{(i)}}{N} < r_j\}\big).$$

From (2.19) and (2.20), the events $\cap_{n=1}^{2M}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}$ and $\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k} = -1\}$ are positively correlated, so

$$P\big(\cap_{n=1}^{2M}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}\big|\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k} = -1\}\big) \geq P\big(\cap_{n=1}^{2M}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}\big).$$

From (2.19), the events $\big\{\{Z_n^{N,i;j,k} \in \{0,-1\}\}\big\}_{n=1}^{2M}$ are positively correlated, so

$$P\big(\cap_{n=1}^{2M}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}\big) \geq \big(P(Z_n^{N,i;j,k} \in \{0,-1\})\big)^{2M} = (1-b_N)^{2M}.$$

Thus,

(2.21)

$$P\big(\cap_{n=1}^{2M}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}\big|\cup_{n=1}^{M}\{Z_{2n}^{N,i;j,k} = -1\}\big) \geq (1-b_N)^{2M}.$$

If $I_i(r_j) > I_i(r_k)$, choose $M = [\frac{1}{b_N}]$ and note that $\lim_{N\to\infty}\frac{a_N}{b_N} = 0$. Since $1 - a_N \leq e^{-a_N}$, from (2.18),

$$P(\cup_{n=1}^{[\frac{1}{b_N}]}\{Z_{2n}^{N,i;j,k} = -1\}) \geq 1 - e^{-a_n[\frac{1}{b_N}]} \geq \frac{a_N}{2b_N}, \text{ for large } N. \tag{2.22}$$

From (2.21),

$$\liminf_{N\to\infty} P\big(\cap_{n=1}^{2[\frac{1}{b_N}]}\big\{Z_n^{N,i;j,k} \in \{0,-1\}\big\}\big|\cup_{n=1}^{[\frac{1}{b_N}]}\{Z_{2n}^{N,i;j,k} = -1\}\big) \geq e^{-2}. \tag{2.23}$$

From (2.15), (2.17), (2.22) and (2.23), we conclude that

$$P\big(\frac{S^{(i)}_{\tau^{N,i;j,k},N+\tau^{N,i;j,k}}}{N} < r_j\big) \gtrapprox e^{-N\big(I_i(r_j)-I_i(r_k)\big)}. \tag{2.24}$$

If $I_i(r_j) = I_i(r_k)$ and $\lim_{N\to\infty}\frac{a_N}{b_N} = 0$, then the above proof still works and gives (2.24). Otherwise, without loss of generality, assume that $c := \lim_{N\to\infty}\frac{b_N}{a_N} \in [0,\infty)$. Choose $M = [\frac{1}{a_N}]$. Then from (2.18),

$$\lim_{N\to\infty} P(\cup_{n=1}^{[\frac{1}{a_N}]}\{Z_{2n}^{N,i;j,k} = -1\}) = 1 - e^{-1}, \tag{2.25}$$

and from (2.21),

(2.26)
$$\liminf_{N\to\infty} P\big(\cap_{n=1}^{2[\frac{1}{a_N}]}\ \{Z_n^{N,i;j,k}\in\{0,-1\}\}\big|\cup_{n=1}^{2[\frac{1}{a_N}]}\ \{Z_{2n}^{N,i;j,k}=-1\}\big)\ge e^{-2c}.$$

From (2.17), (2.25) and (2.26), we obtain (2.24). The first formula in (2.11) follows from (2.14) and (2.24). □

Proposition 5. *Using the notation of Proposition 4,*

(2.27)
$$P(\frac{S^{(i)}_{\tau^{N,i;i,i+1},N+\tau^{N,i;i,i+1}}}{N}<r_j)\lessapprox e^{-N\Big(\max\big(I_i(r_j)-I_i(r_i),\,I_i(r_j)-I_i(r_{i+1})\big)\Big)},$$
$$i=2,\cdots,l-1,\ 1\le j\le i-1;$$
$$P(\frac{S^{(i)}_{\tau^{N,i;i,i+1},N+\tau^{N,i;i,i+1}}}{N}\ge r_k)\lessapprox e^{-N\Big(\max\big(I_i(r_k)-I_i(r_{i+1}),\,I_i(r_k)-I_i(r_i)\big)\Big)},$$
$$i=1,\cdots,l-2,\ i+2\le k\le l.$$

Define

(2.28)
$$\tau^{N,l;l,l+1}=\inf\{n\ge 0:\frac{S^{(l)}_{n,N+n}}{N}<r_l)\};$$
$$\tau^{N,0;0,1}=\inf\{n\ge 0:\frac{S^{(0)}_{n,N+n}}{N}\ge r_1\}.$$

Then

(2.29)
$$P(\frac{S^{(l)}_{\tau^{N,l;l,l+1},N+\tau^{N,l;l,l+1}}}{N}<r_j)\lessapprox e^{-N\big(I_l(r_j)-I_l(r_l)\big)},\ 1\le j\le l-1;$$
$$P(\frac{S^{(0)}_{\tau^{N,0;0,1},N+\tau^{N,0;0,1}}}{N}\ge r_k)\lessapprox e^{-N\big(I_l(r_k)-I_l(r_1)\big)},\ 2\le k\le l.$$

Proof. The statements of Proposition 3 and Proposition 4 involve certain two-sided hitting times related to a random walk with increment distribution $P_i^{(\text{inc})}$, with $1\le i\le l-1$. Similar one-sided results could have been written down for $i=0$ and $i=l$. We refrained from including them in order not to incur the necessity of additional notation and an additional analogous proof. The proof of (2.27) uses Proposition 3 and Proposition 4. Similarly, the proof of (2.29) uses the corresponding one-sided results. Thus, we give the proof of (2.27), the proof of (2.29) following similarly using the corresponding one-sided results.

Without loss of generality, assume that $I_i(r_i) \geq I_i(r_{i+1})$. Applying Proposition 4 with $k = i+1$ and $1 \leq j \leq i-1$, we have

$$P(\frac{S^{(i)}_{\tau^{N,i;j,i+1},N+\tau^{N,i;j,i+1}}}{N} < r_j) \approx e^{-N\big(I_i(r_j)-I_i(r_{i+1})\big)}.$$

From this, the first inequality in (2.27) holds.

We turn to the second inequality in (2.27). Fix $k \in \{i+2, \cdots, l\}$. Define $\hat{Z}_n^{N,i;i,i+1}$ similar to $Z_n^{N,i;i,i+1}$ in (2.1), but with certain changes: If the condition for $Z_n^{N,i;i,i+1} = 1$ holds and, in addition,

$$\max\big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\big) < r_k. \tag{2.30}$$

then define $\hat{Z}_n^{N,i;i,i+1} = 1$; otherwise, define $\hat{Z}_n^{N,i;i,i+1} = 2$. Define $\hat{Z}_n^{N,i;i,i+1} = -1$ if and only if $Z_n^{N,i;i,i+1} = -1$. If the condition for $Z_n^{N,i;i,i+1} = -11$ holds and, in addition, (2.30) holds, define $\hat{Z}_n^{N,i;i,i+1} = -11$; otherwise, define $\hat{Z}_n^{N,i;i,i+1} = -12$. Define $\hat{Z}_n^{N,i;i,i+1} = 0$ if and only if $Z_n^{N,i;i,i+1} = 0$. Note that $\{\hat{Z}_n^{N,i;i,i+1}\}_{n=1}^{\infty}$ are identically distributed, and that each of $\{\hat{Z}_{2n}^{N,i;i,i+1}\}_{n=1}^{\infty}$ and $\{\hat{Z}_{2n-1}^{N,i;i,i+1}\}_{n=1}^{\infty}$ is an independent sequence. By (1.5) we have

$$\begin{aligned} &P(\hat{Z}_1^{N,i;i,i+1} \in \{2,-12\}) \leq \\ &P(\max\big(\frac{S^{(i)}_{(n-1)N,nN}}{N}, \frac{S^{(i)}_{(n-1)N+1,nN+1}}{N}, \cdots, \frac{S^{(i)}_{(n-1)N+N-1,nN+N-1}}{N}\big) \geq r_k) \leq \\ &\sum_{j=0}^{N-1} P(\frac{S^{(i)}_{(n-1)N+j,nN+j}}{N} \geq r_k) \approx Ne^{-NI_i(r_k)} \approx e^{-NI_i(r_k)}. \end{aligned} \tag{2.31}$$

Using this, it follows that (2.2) holds with $Z_1^{N,i;i,i+1}$ replaced by $\hat{Z}_1^{N,i;i,i+1}$:

$$\begin{aligned} &P(\hat{Z}_1^{N,i;i,i+1} = 1) \approx e^{-NI_i(r_{i+1})}; \\ &P(\hat{Z}_1^{N,i;i,i+1} = -1) \approx e^{-NI_i(r_i)}; \\ &P(\hat{Z}_1^{N,i;i,i+1} = -11) \lessapprox e^{-N\big(I_i(r_i)+I_i(r_{i+1})\big)}. \end{aligned} \tag{2.32}$$

Define $\hat{\sigma}^{(e)}_{i;i,i+1}$ and $\hat{\sigma}^{(o)}_{i;i,i+1}$ for $\hat{Z}^{N,i;i,i+1}_n$ exactly as $\sigma^{(e)}_{i;i,i+1}$ and $\sigma^{(o)}_{i;i,i+1}$ were defined for $Z^{N,i;i,i+1}_n$ in (2.12). Using (2.31), (2.32) and the fact that each of $\{\hat{Z}^{N,i;i,i+1}_{2n}\}_{n=1}^\infty$ and $\{\hat{Z}^{N,i;i,i+1}_{2n-1}\}_{n=1}^\infty$ is an iid sequence, it follows that

$$P(\hat{Z}^{N,i;i,i+1}_{\hat{\sigma}^{(*)}_{i;i,i+1}} \in \{2,-12\}) \lessapprox e^{-N\left(I_i(r_k)-I_i(r_{i+1})\right)}, \tag{2.33}$$

both when $\hat{\sigma}^{(*)}_{i;i,i+1} = \hat{\sigma}^{(e)}_{i;i,i+1}$ and when $\hat{\sigma}^{(*)}_{i;i,i+1} = \hat{\sigma}^{(o)}_{i;i,i+1}$. (Recall that we are assuming that $I_i(r_i) \ge I_i(r_{i+1})$.) Now

$$\Big\{\frac{S^{(i)}_{\tau^{N,i;i,i+1},N+\tau^{N,i;i,i+1}}}{N} \ge r_k\Big\} \subset \big\{\hat{Z}^{N,i;i,i+1}_{\hat{\sigma}^{(e)}_{i;i,i+1}} \in \{2,-12\}\big\} \cup \big\{\hat{Z}^{N,i,i,i+1}_{\hat{\sigma}^{(0)}_{i;i,i+1}} \in \{2,-12\}\big\};$$

thus, it follows from (2.33) that

$$P\big(\frac{S^{(i)}_{\tau^{N,i;i,i+1},N+\tau^{N,i;i,i+1}}}{N} \ge r_k\big) \lessapprox e^{-N\left(I_i(r_k)-I_i(r_{i+1})\right)}, \tag{2.34}$$

giving the second inequality in (2.27). □

Recall the processes $\{Y^{N;D}_m\}_{m=0}^\infty$ and $\{Y^{N;D_1}_m\}_{m=0}^\infty$ mentioned at the end of section 1. They denote the Markov processes that follow the changes of the increment distribution utilized by the delayed version $\{X^{N;D}_n\}_{n=0}^\infty$ and by the one-step delayed version $\{X^{N;D_1}_n\}_{n=0}^\infty$ of the random walk reinforced by its recent history. Thus, $Y^{N;D}_0 = i_0$, since the process $\{X^{N;D}_n\}_{n=0}^\infty$ starts out using the increment distribution $P^{(\text{inc})}_{i_0}$. If the first time the process $\{X^{N;D}_n\}_{n=0}^\infty$ changes its increment distribution, it switches from distribution $P^{(\text{inc})}_{i_0}$ to distribution $P^{(\text{inc})}_j$, then $Y^{N;D}_1 = j$. In general, $Y^{N;D}_m = k$, if after switching increment distribution m times, the process $\{X^{N;D}_n\}_{n=0}^\infty$ is using the increment distribution $P^{(\text{inc})}_k$. The process $\{Y^{N;D_1}_m\}_{m=0}^\infty$ is defined analogously. We denote the transitions for $\{Y^{N;D}_m\}_{m=0}^\infty$ by

$$p^{N;D}_{i,j} = P(Y^{N;D}_{m+1} = j | Y^{N;D}_m = i),\ i,j \in \{0,\cdots,l\},\ i \ne j$$

and for $\{Y^{N;D_1}_m\}_{m=0}^\infty$ by

$$p^{N;D_1}_{i,j} = P(Y^{N;D_1}_{m+1} = j | Y^{N;D_1}_m = i),\ i,j \in \{0,\cdots,l\},\ i \ne j.$$

We call $\{Y_m^{N;D}\}_{m=0}^\infty$ and $\{Y_m^{N;D_1}\}_{m=0}^\infty$ the increments processes. Using Propositions 3-5, the following estimates on their transition probabilities are almost immediate.

Proposition 6.

(2.35)

$$p_{i,i+1}^{N;D},\ p_{i,i+1}^{N;D_1} \approx e^{-N\left(I_i(r_{i+1})-I_i(r_i)\right)^+};\ i\in\{1,\cdots,l-1\};$$

$$p_{i,i-1}^{N;D},\ p_{i,i-1}^{N;D_1} \approx e^{-N\left(I_i(r_i)-I_i(r_{i+1})\right)^+};\ i\in\{1,\cdots,l-1\};$$

$$p_{0,1}^{N;D},\ p_{0,1}^{N;D_1}\approx 1,\quad p_{l,l-1}^{N;D},\ p_{l,l-1}^{N;D_1}\approx 1;$$

$$p_{i,j}^{N;D} \lessapprox e^{-N\left(\max\left(I_i(r_j)-I_i(r_{i+1}),\,I_i(r_j)-I_i(r_i)\right)\right)},\ 1\le i\le l-2,\ i+2\le j\le l;$$

$$p_{0,j}^{N;D}\lessapprox e^{-N(I_0(r_j)-I_0(r_1))},\ 2\le j\le l;$$

$$p_{i,j}^{N;D} \lessapprox e^{-N\left(\max\left(I_i(r_{j+1})-I_i(r_i),\,I_i(r_{j+1})-I_i(r_{i+1})\right)\right)},\ 2\le i\le l-1,\ 0\le j\le i-2;$$

$$p_{l,j}^{N;D}\lessapprox e^{-N(I_l(r_{j+1})-I_l(r_l))},\ 0\le j\le l-2.$$

Proof. Noting that

$$p_{i,j}^{N;D}=P(r_j\le \frac{S^{(i)}_{\tau^{N,i;i,i+1},N+\tau^{N,i;i,i+1}}}{N}<r_{j+1}),$$

the proposition for the delayed version follows immediately from Proposition 4 (with j and k there set to i and $i+1$ respectively) and Proposition 5. Since $p_{i,i+1}^{N;D_1}=\sum_{j=i+1}^l p_{i,j}^{N;D}$ and $p_{i,i-1}^{N;D_1}=\sum_{j=0}^{i-1}p_{i,j}^{N;D}$, the one-step delayed version follows from the delayed version. □

Denote respectively the invariant distributions of the Markov chains $\{Y_m^{N;D}\}_{m=0}^\infty$ and $\{Y_m^{N;D_1}\}_{m=0}^\infty$ on $\{0,\cdots,l\}$ by $\nu^{N;D}$ and $\nu^{N;D_1}$. The Markov chain $\{Y_m^{N;D_1}\}_{m=0}^\infty$ is a birth and death process, thus reversible, so its invariant distribution can be calculated explicitly, via the detailed balance equations: $\nu^{N;D_1}(i)p_{i,i+1}^{N;D_1}=\nu^{N;D_1}(i+1)p_{i+1,i}^{N;D_1},\ \ i=0,\cdots,l-1$. As is well-known, we obtain

$$\begin{aligned}
&\Pi_N\nu^{N,D_1}(0)=1;\\
&\Pi_N\nu^{N,D_1}(k)=\prod_{i=1}^k\frac{p^{N;D_1}_{i-1,i}}{p^{N;D_1}_{i,i-1}},\ k=1,\cdots,l,\\
&\text{where } \Pi_N=1+\sum_{k=1}^l\prod_{i=1}^k\frac{p^{N;D_1}_{i-1,i}}{p^{N;D_1}_{i,i-1}}.
\end{aligned}\tag{2.36}$$

Under Assumption E, we prove that $\nu^{N,D}$, which we cannot calculate explicitly, and ν^{N,D_1} are equivalent up to exponential order.

Proposition 7. *Assume that Assumption E holds. Then*

$$\frac{\nu^{N,D_1}(k)}{\nu^{N,D}(k)}\approx 1,\quad k=0,\cdots,l.\tag{2.37}$$

The rather long proof of this proposition, which is needed for the proof of Theorem 1 for the delayed version but not for the one-step delayed version, is postponed until section 5 so as not to interfere with the flow of the proof of that theorem.

Recalling the definition of $\tau^{N,i;i,i+1}$ from (2.10), note that anytime the delayed or the one-step delayed version of the random walk reinforced by its recent history switches to regime i, the number of steps during which it will operate in this regime before moving to a different regime is distributed as $\tau^{N,i;i,i+1}$, and the distance it travelled between its entrance into regime i and its exit to another regime is distributed as $S^{(i)}_{\tau^{N,i;i,i+1}+N}$. The next two propositions calculate the expected values of these two distributions.

Proposition 8.

$$\begin{aligned}
&E\tau^{N,i;i,i+1}\approx e^{N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)},\ 1\le i\le l-1;\\
&E\tau^{N,l,l,l+1}\approx e^{NI_l(r_l)};\\
&E\tau^{N,0;0,1}\approx e^{NI_0(r_1)}.
\end{aligned}\tag{2.38}$$

Proof. Using the notation from the proof of Proposition 4, for any positive integer L, we have

$$\{\tau^{N,i;i,i+1}\ge 2LN\}=\{Z_n^{N,i;i,i+1}=0,\ \text{for all } n=1,\cdots,2L\}=\{\sigma^{(e)}_{i;i,i+1}>2L,\sigma^{(o)}_{i;i,i+1}>2L-1\}. \tag{2.39}$$

Since $\sigma^{(e)}_{i;i,i+1}$ and $\sigma^{(o)}_{i;i,i+1}+1$ have the same distribution, it follows that

$$P(\tau^{N,i;i,i+1}\ge 2LN)\le 2P(\sigma^{(e)}_{i;i,i+1}>2L),\ L\ge 1. \tag{2.40}$$

We have

$$\sum_{L=0}^{\infty}P(\tau^{N,i;i,i+1}\ge 2LN)\ge\sum_{m=0}^{\infty}(\frac{m}{2N}+1)P(\tau^{N,i;i,i+1}=m)=1+\frac{1}{2N}E\tau^{N,i;i,i+1}. \tag{2.41}$$

From the definition of $\sigma^{(e)}_{i;i,i+1}$ along with Proposition 3, $\sigma^{(e)}_{i;i,i+1}$ is distributed according to a geometric distribution with parameter $p\approx e^{-N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}$; thus, $E\sigma^{(e)}_{i;i,i+1}\approx e^{N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}$. Consequently,

$$\sum_{L=0}^{\infty}P(\sigma^{(e)}_{i;i,i+1}>2L)\le\sum_{L=1}^{\infty}P(\sigma^{(e)}_{i;i,i+1}\ge L)=E\sigma^{(e)}_{i;i,i+1}\approx e^{N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}. \tag{2.42}$$

From (2.40)-(2.42), we obtain

$$E\tau^{N,i;i,i+1}\lessapprox e^{N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}. \tag{2.43}$$

From Proposition 3 and the definition of $\sigma^{(e)}_{i;i,i+1}$ and $\sigma^{(o)}_{i;i,i+1}$, we have for any $\epsilon>0$ and sufficiently large N,

$$P(\sigma^{(e)}_{i;i,i+1}>2L,\sigma^{(o)}_{i;i,i+1}>2L-1)=P(Z_n^{N,i;i,i+1}=0,\ \text{for } n=1,\cdots,2L)\ge 1-2LP(Z_1^{N,i;i,i+1}\neq 0)\ge 1-2Le^{\epsilon N-N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}. \tag{2.44}$$

Letting $L_{N,\epsilon}=[e^{-2\epsilon N+N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}]$, it follows from (2.39) and (2.44) that $\lim_{N\to\infty}P(\tau^{N,i;i,i+1}\ge 2L_{N,\epsilon}N)=1$. Since $\epsilon>0$ is arbitrary, it follows that

$$E\tau^{N,i;i,i+1}\gtrapprox e^{N\min\big(I_i(r_i),\,I_i(r_{i+1})\big)}. \tag{2.45}$$

The first formula in (2.38) follows from (2.43) and (2.45). The second and third formulas are proved similarly using one-sided hitting times—see the first paragraph of the proof of Proposition 5. □

Proposition 9.

$$ES^{(i)}_{\tau^{N,i;i,i+1}+N} = \mu_i(E\tau^{N,i;i,i+1} + N), \ 0 \le i \le l. \tag{2.46}$$

Proof. Let $\{W_n\}_{n=1}^\infty$ be iid random variables distributed according to $P_i^{(\mathrm{inc})}$ and consider the filtration $\mathcal{F}_n = \sigma\Big(W_1, \cdots, W_n\Big), \ n \ge 1$. We can write $S_n^{(i)} = \sum_{j=1}^n W_j$. Now $M_{n+N} := S^{(i)}_{n+N} - (n+N)\mu_i$, $n \ge 0$, is a martingale with respect to $\{\mathcal{F}_{n+N}\}_{n=0}^\infty$. Note that $N + \tau^{N,i,i,i+1}$ is a stopping time with respect to $\{\mathcal{F}_{n+N}\}_{n=0}^\infty$. So by Doob's optional sampling theorem,

$$ES_{(\tau^{N,i,i,i+1}+N)\wedge L} - \mu_i E((\tau^{N,i,i,i+1} + N) \wedge L) = 0, \ L \ge 0.$$

Letting $L \to \infty$ and using (2.38), we obtain (2.46). □

3. Proof of Theorem 1

We prove parts (i) and (ii) together. Recall that $\nu^{N,D}$ and ν^{N,D_1} denote the invariant distributions respectively of the increments processes $\{Y_m^{N;D}\}_{m=0}^\infty$ and $\{Y_m^{N;D_1}\}_{m=0}^\infty$. Thus, by the ergodic theorem, as $m \to \infty$ the asymptotic proportion of switches of the increment process $\{Y_m^{N;D}\}_{m=0}^\infty$ for the delayed process $\big(\{Y_m^{N;D_1}\}_{m=0}^\infty$ for the one-step delayed process$\big)$ to the regime i is $\nu^{N,D}(i)$ $\big(\nu^{N,D_1}(i)\big)$. Recalling the definition of $\tau^{N,i;i,i+1}$ from (2.10), note that anytime the delayed or the one-step delayed version of the random walk reinforced by its recent history switches to regime i, the number of steps during which it will operate in this regime before moving to a different regime is distributed as $\tau^{N,i;i,i+1}$, and the distance travelled by the process between its entrance into regime i and its exit to another regime is distributed as $S^{(i)}_{\tau^{N,i;i,i+1}+N}$. Also, this random number of steps and this random distance travelled are independent of the random number of steps the process spent and the random distance it travelled in any regime in the past before the present entrance into regime i. From these observations, it is standard to deduce that the speeds $s^D(N, m_1, \cdots, m_l)$ and $s^{D_1}(N, m_1, \cdots, m_l)$, defined in Proposition 1, exist almost surely and are almost surely given by the constants

$$s^D(N, m_1, \cdots, m_l) = \frac{\sum_{i=0}^{l} \nu^{N,D}(i) ES^{(i)}_{\tau^{N,i;i,i+1}+N}}{\sum_{i=0}^{l} \nu^{N,D}(i) E\tau^{N,i;i,i+1}};$$
$$(3.1)\qquad s^{D_1}(N, m_1, \cdots, m_l) = \frac{\sum_{i=0}^{l} \nu^{N,D_1}(i) ES^{(i)}_{\tau^{N,i;i,i+1}+N}}{\sum_{i=0}^{l} \nu^{N,D_1}(i) E\tau^{N,i;i,i+1}}.$$

This proves Proposition 1 for the delay and the one-step delayed versions.

By Propositions 8 and 9,

$$ES^{(i)}_{\tau^{N,i;i,i+1}+N} \approx \mu_i e^{N \min\left(I_i(r_i), I_i(r_{i+1})\right)}, \text{ for } 1 \le i \le l-1;$$
$$(3.2)\qquad ES^{(0)}_{\tau^{N,0;0,1}+N} \approx \mu_0 e^{N I_0(r_1)},\ ES^{(l)}_{\tau^{N,l;l,l+1}+N} \approx \mu_l e^{N I_l(r_l)}.$$

From (2.36) and Proposition 6, we have

(3.3)

$$\nu^{N;D_1}(0) \approx \frac{1}{\Pi_N};$$
$$\nu^{N;D_1}(1) \approx \frac{1}{\Pi_N} e^{N\left(I_1(r_1)-I_1(r_2)\right)^+};$$
$$\nu^{N;D_1}(i) \approx \frac{1}{\Pi_N} e^{N\left(I_1(r_1)-I_1(r_2)\right)^+} \prod_{k=2}^{i} e^{N\left(\left(I_k(r_k)-I_k(r_{k+1})\right)^+ - \left(I_{k-1}(r_k)-I_{k-1}(r_{k-1})\right)^+\right)},$$
$1 \le i \le l-1;$
$$\nu^{N;D_1}(l) \approx \frac{1}{\Pi_N} e^{N\left(I_1(r_1)-I_1(r_2)\right)^+} \prod_{k=2}^{l-1} e^{N\left(\left(I_k(r_k)-I_k(r_{k+1})\right)^+ - \left(I_{k-1}(r_k)-I_{k-1}(r_{k-1})\right)^+\right)} \times$$
$$e^{-N\left(I_{l-1}(r_l)-I_{l-1}(r_{l-1})\right)^+}.$$

Noting that $\left(I_k(r_k) - I_k(r_{k+1})\right)^+ - \left(I_k(r_{k+1}) - I_k(r_k)\right)^+ = I_k(r_k) - I_k(r_{k+1})$ and recalling the definition of $\{\Lambda_i\}_{i=0}^{l}$ in the statement of Theorem 1, it follows from (3.2) and (3.3) that

(3.4)

$$\nu^{N,D_1}(i) ES^{(i)}_{\tau^{N,i;i,i+1}+N} \approx \frac{1}{\Pi_N} \mu_i e^{I_i(r_i)+\sum_{k=1}^{i-1}\left(I_k(r_k)-I_k(r_{k+1})\right)} = \frac{1}{\Pi_N} \mu_i e^{N\Lambda_i},\ 0 \le i \le l.$$

Substituting (3.4) into the second equation in (3.1), recalling from Propositions 8 and 9 that $\frac{ES^{(i)}_{\tau^{N,i;i,i+1}+N}}{E\tau^{N,i;i,i+1}} \approx \mu_i$, and letting $N \to \infty$ gives part (i) of

Theorem 1 for the one-step delayed version. Under Assumption E, it follows from Proposition 7 that (3.3) also holds with D_1 replaced by D. Thus, part (ii) of Theorem 1 holds for the delayed version. □

4. Proof of Theorem 2

For the proof of Theorem 2, we need the following lemma.

Lemma 1. *Let $\{Z_n\}_{n=1}^\infty$ be iid random variables satisfying $EZ_1 = \mu$ and let $S_n = \sum_{i=1}^n Z_i$. Then for every $r < \mu$,*

$$P(\frac{S_n}{n} \ge r,\ n = 1, 2, \cdots) > 0. \tag{4.1}$$

Proof. By the strong law of large numbers, $\lim_{n\to\infty} \frac{S_n}{n} = \mu$ a.s. Thus, for every $r < \mu$, there exists an N_r such that $P(\frac{S_n}{n} \ge r,\ n > N_r) > 0$. Clearly, $P(\frac{S_n}{n} \ge r,\ n = 1, \cdots, N_r) > 0$. Since the events $\{\frac{S_n}{n} \ge r\ n = 1, \cdots, N_r\}$ and $\{\frac{S_n}{n} \ge r,\ n > N_r\}$ are positively correlated, we have

$$P(\frac{S_n}{n} \ge r,\ n = 1, 2, \cdots) =$$
$$P(\frac{S_n}{n} \ge r,\ n = 1, \cdots, N_r) P(\frac{S_n}{n} \ge r,\ n > N_r | \frac{S_n}{n} \ge r,\ n = 1, \cdots, N_r) > 0.$$

□

We now turn to the proof of the theorem.

Proof of Theorem 2. Without loss of generality, assume that $I_0(r_1) < I_1(r_1)$. We need to prove that the limiting speed of $\{X_n^{N;I}\}_{n=0}^\infty$ is μ_1. Define

$$c := P(\frac{S_n^{(1)}}{n} \ge r_1,\ n = 1, 2, \cdots) > 0, \tag{4.2}$$

where the positivity of c follows from Lemma 1. We will compare our process $\{X_n^{N;I}\}_{n=0}^\infty$ to another process $\{W_n^N\}_{n=0}^\infty$. Define $W_0^N = 0$ and let $\{W_n^N\}_{n=1}^N$ be distributed like a random walk with increment distribution $P_0^{(\text{inc})}$. After time N, the process continues to use the increment distribution $P_0^{(\text{inc})}$ up until and including the first $n > N$ such that $\frac{W_n^N - W_{n-N}^N}{N} \ge r_1$. At this point the increment distribution is reassessed. With probability c it switches to the increment distribution $P_1^{(\text{inc})}$, while with probability $1-c$ it remains with the increment distribution $P_0^{(\text{inc})}$. The continuation of the process is defined inductively as follows. After every reassessment, the process automatically uses its newly chosen increment distribution, call it i, for the next N steps. After those N steps, if $i=0$, then the process continues to use increment distribution $P_0^{(\text{inc})}$ until the first time that the average of the most recent N steps of the process is larger or equal to r_1, at which point the process makes a reassessment just like the one noted above, while, on the other hand, if $i=1$, then the process continues to use increment distribution $P_1^{(\text{inc})}$ until the first time that the average of the most recent N steps of the process is smaller than r_1, at which point the process makes a "reassessment" and automatically reverts to the increment distribution $P_0^{(\text{inc})}$. In light of (4.2), the process $\{W_n^N\}_{n=0}^{\infty}$ is stochastically dominated by our original process $\{X_n^{N;I}\}_{n=0}^{\infty}$. Thus, to complete the proof of the theorem, it suffices to show that the limiting speed as $N\to\infty$ of the $\{W_n^N\}_{n=0}^{\infty}$ process is μ_1.

Let $\{Y_m^N\}_{m=0}^{\infty}$ denote the Markov chain that follows the changes in the increment distribution at the reassessment times (similar to $\{Y_m^{N;D}\}_{m=0}^{\infty}$ and $\{Y_m^{N;D_1}\}_{m=0}^{\infty}$). This is a Markov process on the two states $\{0,1\}$ with transition probability from 0 to 0 equal to $1-c$, from 0 to 1 equal to c, and from 1 to 0 equal to 1. Its invariant distribution ν^N satisfies $\nu^N(0)=\frac{1}{1+c}$ and $\nu^N(1)=\frac{c}{1+c}$. Anytime the process $\{W_n^N\}_{n=0}^{\infty}$ reassesses and finds itself in regime i, the number of steps during which it will operate in this regime before its next reassessment is distributed as $\tau^{N,i;i,i+1}$, and the distance travelled by the process during this time is distributed as $S^{(i)}_{\tau^{N,i;i,i+1}+N}$. Thus, just as in the derivation of (3.1), we conclude that the speed s^N of $\{W_n^N\}_{n=0}^{\infty}$ is given by

$$s^N=\frac{\nu^N(0)ES^{(0)}_{\tau^{N,0;0,1}+N}+\nu^N(1)ES^{(1)}_{\tau^{N,1;1,2}+N}}{\nu^N(0)E\tau^{N,0;0,1}+\nu^N(1)E\tau^{N,1;1,2}}. \tag{4.3}$$

This proves Proposition 1 for the instantaneous version.

From Propositions 8 and 9 (with $l=1$), we have $E\tau^{N,0;0,1}\approx e^{NI_0(r_1)}$, $E\tau^{N,1;1,2}\approx e^{NI_1(r_1)}$, $ES^{(0)}_{\tau^{N,0;0,1}+N}\approx\mu_0e^{NI_0(r_1)}$ and $ES^{(1)}_{\tau^{N,1;1,2}+N}\approx\mu_1e^{NI_1(r_1)}$. Substituting this in (4.3) and recalling the assumption that $I_1(r_1)>I_0(r_1)$, it follows that the limiting speed of $\{W^N_n\}_{n=0}^\infty$ as $N \to \infty$ is given by $\lim_{N\to\infty} s^N = \mu_1$. □

5. Proof of Proposition 7

For the proof of Proposition 7, we need the following lemma.

Lemma 2. *If Assumption E holds, then for* $i \le j-2$,

$$p^{N;D}_{i,j} \lessapprox \prod_{k=i}^{j-1} p^{N;D}_{k,k+1};$$
$$p^{N;D}_{j,i} \lessapprox \prod_{k=i+1}^{j} p^{N;D}_{k,k-1}. \tag{5.1}$$

Proof. We'll prove the first inequality in the statement of the proposition with the help of the first inequality in Assumption E; the second inequality in the statement of the proposition is proved in the same way, using the second inequality in Assumption E. By Proposition 6, if $i \ge 1$, it suffices to show that

$$\sum_{k=i}^{j-1} \big(I_k(r_{k+1}) - I_k(r_k)\big)^+ < \max\Big(I_i(r_j) - I_i(r_{i+1}), I_i(r_j) - I_i(r_i)\Big), \tag{5.2}$$

while if $i = 0$, it suffices to show that

$$\sum_{k=1}^{j-1} \big(I_k(r_{k+1}) - I_k(r_k)\big)^+ < I_0(r_j) - I_0(r_1). \tag{5.3}$$

Consider the case $i \ge 1$. Assume first that $I_i(r_{i+1}) \ge I_i(r_i)$. Then in order to show (5.2), it suffices to show that

$$I_i(r_{i+1}) - I_i(r_i) + \sum_{k=i+1}^{j-1} I_k(r_{k+1}) < I_i(r_j) - I_i(r_i),$$

but this is exactly the first inequality in Assumption E. Now assume that $I_i(r_{i+1}) < I_i(r_i)$. Then in order to show (5.2), it suffices to show that

$$\sum_{k=i+1}^{j-1} I_k(r_{k+1}) < I_i(r_j) - I_i(r_{i+1}),$$

which is again the first inequality in Assumption E.

Now consider the case $i = 0$. In order to show (5.3), it suffices to show that

$$\sum_{k=1}^{j-1} I_k(r_{k+1}) < I_0(r_j) - I_0(r_1).$$

But this is a particular case of the first inequality in Assumption E. □

We now turn to the proof of the proposition.

Proof of Proposition 7. We will show that

$$\nu^{N,D}(k) \approx \frac{1}{p_{k,k-1}} \sum_{j=0}^{k-1} \nu^{N,D}(j) p_{j,k}^{N;D}, \; k = 1, \cdots, l. \tag{5.4}$$

Using this we prove the proposition as follows. In light of Proposition 6 and (2.36), it suffices to show that

$$\nu^{N,D}(j) \approx (\prod_{i=1}^{j} \frac{p_{i-1,i}^{N;D}}{p_{i,i-1}^{N;D}}) \nu^{N,D}(0), \text{ for } j \in \{1, \cdots, l\}. \tag{5.5}$$

From (5.4) we have

$$\nu^{N,D}(1) \approx \frac{p_{0,1}^{N;D}}{p_{1,0}^{N;D}} \nu^{N,D}(0). \tag{5.6}$$

Now make the inductive assumption that

$$\nu^{N,D}(j) \approx (\prod_{i=1}^{j} \frac{p_{i-1,i}^{N;D}}{p_{i,i-1}^{N;D}}) \nu^{N,D}(0), \text{ for } j < k. \tag{5.7}$$

Then using (5.4) for the first approximation and the inductive assumption (5.7) for the second approximation, we have

(5.8)
$$\nu^{N,D}(k) \approx \frac{1}{p_{k,k-1}^{N;D}} \sum_{j=0}^{k-1} \nu^{N,D}(j) p_{j,k}^{N;D} \approx (\prod_{i=1}^{k} \frac{p_{i-1,i}^{N;D}}{p_{i,i-1}^{N;D}}) \nu^{N,D}(0) + \frac{1}{p_{k,k-1}^{N;D}} \sum_{j=0}^{k-2} \nu^{N,D}(j) p_{j,k}^{N;D}.$$

For $j \in \{0, \cdots, k-2\}$, we have

$$\frac{p_{j,k}^{N;D}}{p_{k,k-1}^{N;D}} \nu^{N,D}(j) \approx \frac{p_{j,k}^{N;D}}{p_{k,k-1}^{N;D}} (\prod_{i=1}^{j} \frac{p_{i-1,i}^{N;D}}{p_{i,i-1}^{N;D}}) \nu^{N,D}(0) \lessapprox\lessapprox (\prod_{i=1}^{k} \frac{p_{i-1,i}^{N;D}}{p_{i,i-1}^{N;D}}) \nu^{N,D}(0), \tag{5.9}$$

where the approximation follows from the inductive assumption (5.7) and the inequality follows from Lemma 2. From (5.8) and (5.9), it follows that

(5.7) also holds for k. This completes the inductive proof of (5.5). We now return to prove (5.4).

Since $\nu^{N,D}(l)=\sum_{j=0}^{l-1}\nu^{N,D}(j)p^{N;D}_{j,l}$ and $p^{N;D}_{l,l-1}\approx 1$, (5.4) holds for $k=l$. We now show directly that it also holds for $k=l-1$. Then we will use induction to obtain (5.4) for all k. We have

$$\begin{aligned}
&\nu^{N,D}(l-1)=\sum_{j=0}^{l-2}\nu^{N,D}(j)p^{N;D}_{j,l-1}+\nu^{N,D}(l)p^{N;D}_{l,l-1}=\\
&\sum_{j=0}^{l-2}\nu^{N,D}(j)p^{N;D}_{j,l-1}+p^{N;D}_{l,l-1}\sum_{j=0}^{l-1}\nu^{N,D}(j)p^{N;D}_{j,l}=\\
&\sum_{j=0}^{l-2}\nu^{N,D}(j)\big(p^{N;D}_{j,l-1}+p^{N;D}_{l,l-1}p^{N;D}_{j,l}\big)+p^{N;D}_{l-1,l}p^{N;D}_{l,l-1}\nu^{N,D}(l-1).
\end{aligned}\tag{5.10}$$

By Lemma 2, $p^{N;D}_{j,l}\lessapprox\lessapprox p^{N;D}_{j,l-1}$, for $j\le l-2$, thus from (5.10) we have

$$\nu^{N,D}(l-1)\approx\frac{1}{1-p^{N;D}_{l-1,l}p^{N;D}_{l,l-1}}\sum_{j=0}^{l-2}\nu^{N,D}(j)p^{N;D}_{j,l-1}.\tag{5.11}$$

Using the notation $p^{N;D}_{k,\le j}:=\sum_{i=0}^{j}p^{N;D}_{k,i}$, for $j\le k-2$, and using Lemma 2 for the approximation, we have

$$\begin{aligned}
&1-p^{N;D}_{l-1,l}p^{N;D}_{l,l-1}=1-\big(1-p_{l-1,l-2}-p^{N;D}_{l-1,\le l-3}\big)\big(1-p^{N;D}_{l,\le l-2}\big)=\\
&p_{l-1,l-2}+p^{N;D}_{l-1,\le l-3}+p^{N;D}_{l,\le l-2}-p_{l-1,l-2}p^{N;D}_{l,\le l-2}-p^{N;D}_{l-1,\le l-3}p^{N;D}_{l,\le l-2}\approx\\
&p_{l-1,l-2}.
\end{aligned}\tag{5.12}$$

Now (5.4) for $k=l-1$ follows from (5.11) and (5.12).

We now make the inductive assumption that (5.4) holds for $k>k_0$, where $k_0\in\{0,\cdots,l-2\}$, and prove that it holds for $k=k_0$. We write

$$\nu^{N,D}(k_0)=\sum_{j=0}^{k_0-1}\nu^{N,D}(j)p^{N;D}_{j,k_0}+\nu^{N,D}(k_0+1)p^{N;D}_{k_0+1,k_0}+\sum_{j=k_0+2}^{l}\nu^{N,D}(j)p^{N;D}_{j,k_0}.\tag{5.13}$$

Using the inductive assumption, we have

$$\nu^{N,D}(k_0+1)p^{N;D}_{k_0+1,k_0}\approx\sum_{j=0}^{k_0-1}\nu^{N,D}(j)p^{N;D}_{j,k_0+1}+\nu^{N,D}(k_0)p^{N;D}_{k_0,k_0+1},\tag{5.14}$$

and by Lemma 2,

$$\sum_{j=0}^{k_0-1} \nu^{N,D}(j) p^{N;D}_{j,k_0+1} \lessapprox \sum_{j=0}^{k_0-1} \nu^{N,D}(j) p^{N;D}_{j,k_0}. \tag{5.15}$$

We now consider the terms in the final expression on the right hand side of (5.13), beginning with the first one, when $j = k_0 + 2$. By the inductive assumption,

$$\nu^{N,D}(k_0+2) p^{N;D}_{k_0+2,k_0} \approx \frac{p^{N;D}_{k_0+2,k_0}}{p^{N;D}_{k_0+2,k_0+1}} \sum_{j=0}^{k_0+1} \nu^{N,D}(j) p_{j,k_0+2}. \tag{5.16}$$

Using Lemma 2,

$$\begin{aligned} &\frac{p^{N;D}_{k_0+2,k_0}}{p^{N;D}_{k_0+2,k_0+1}} \sum_{j=0}^{k_0-1} \nu^{N,D}(j) p_{j,k_0+2} \lessapprox \sum_{j=0}^{k_0-1} \nu^{N,D}(j) p^{N;D}_{j,k_0}; \\ &\frac{p^{N;D}_{k_0+2,k_0}}{p^{N;D}_{k_0+2,k_0+1}} p^{N;D}_{k_0,k_0+2} \nu^{N,D}(k_0) \lessapprox p^{N;D}_{k_0,k_0+2} \nu^{N,D}(k_0). \end{aligned} \tag{5.17}$$

Also, using the inductive assumption for the approximation and Lemma 2 for the inequality, we have

$$\begin{aligned} &\frac{p^{N;D}_{k_0+2,k_0}}{p^{N;D}_{k_0+2,k_0+1}} p^{N;D}_{k_0+1,k_0+2}\, \nu^{N,D}(k_0+1) \approx \frac{p^{N;D}_{k_0+2,k_0}}{p^{N;D}_{k_0+2,k_0+1}} p^{N;D}_{k_0+1,k_0+2} \frac{1}{p^{N;D}_{k_0+1,k_0}} \sum_{j=0}^{k_0} \nu^{N,D}(j) p^{N;D}_{j,k_0+1} \\ &\lessapprox \sum_{j=0}^{k_0-1} \nu^{N,D}(j) p^{N;D}_{j,k_0+1} + p^{N;D}_{k_0,k_0+1} \nu^{N,D}(k_0). \end{aligned} \tag{5.18}$$

From (5.16)-(5.18) and Lemma 2, we conclude that

$$\nu^{N,D}(k_0+2) p^{N;D}_{k_0+2,k_0} \lessapprox \sum_{j=0}^{k_0-1} \nu^{N,D}(j) p^{N;D}_{j,k_0} + p^{N;D}_{k_0,k_0+1} \nu^{N,D}(k_0). \tag{5.19}$$

We now consider the term $\nu^{N,D}(j) p^{N;D}_{j,k_0}$ in the final expression on the right hand side of (5.13) for $k_0 + 2 < j \le l$. Using the inductive assumption twice and using Lemma 2, we have

(5.20)

$$\nu^{N,D}(j)p^{N;D}_{j,k_0} \approx \frac{p^{N;D}_{j,k_0}}{p^{N;D}_{j,j-1}}\sum_{t=0}^{j-1}\nu^{N,D}(t)p^{N;D}_{t,j} \lessapprox\lessapprox p^{N;D}_{j-1,k_0}\sum_{t=0}^{j-1}\nu^{N,D}(t)p^{N;D}_{t,j} \lessapprox$$

$$p^{N;D}_{j-1,k_0}\big(\sum_{t=0}^{j-2}\nu^{N;D}(t)p^{N;D}_{t,j-1}+\nu^{N;D}(j-1)p^{N;D}_{j-1,j}\big) \approx$$

$$p^{N;D}_{j-1,k_0}\Big(p^{N;D}_{j-1,j-2}\,\nu^{N,D}(j-1)+p^{N;D}_{j-1,j}\nu^{N,D}(j-1)\Big) \le \nu^{N,D}(j-1)p^{N;D}_{j-1,k_0}.$$

From (5.19) and (5.20) we conclude that

(5.21)

$$\nu^{N,D}(j)p^{N;D}_{j,k_0} \lessapprox\lessapprox \sum_{j=0}^{k_0-1}\nu^{N,D}(j)p^{N;D}_{j,k_0}+p^{N;D}_{k_0,k_0+1}\nu^{N,D}(k_0),\ \text{for } k_0+2\le j\le l.$$

From (5.13)-(5.15) and (5.21), it follows that $\nu^{N,D}(k_0)\approx\sum_{j=0}^{k_0-1}\nu^{N,D}(j)p^{N;D}_{j,k_0}+\nu^{N,D}(k_0)p^{N;D}_{k_0,k_0+1}$. Now (5.4) for $k=k_0$ follows from this and Proposition 6. □

6. Proof of Proposition 2

Except for the family of Gaussian distributions, all the families of distributions appearing in the proposition are one-parameter families for which for any μ there is at most one member of the family with mean μ. For each family of distributions, let I_μ denote the Legendre-Fenchel transformation for the distribution of that family with mean μ. (In the case of the Gaussian distribution, we fix the variance and suppress that dependence.) By induction, to prove the proposition, it suffices to prove that for each family of distributions,

$$\begin{aligned} &I_\mu(r)+I_\nu(s)<I_\mu(s),\ \mu<r<\nu<s;\\ &I_\mu(r)+I_\nu(s)<I_\mu(s),\ s<\nu<r<\mu. \end{aligned} \tag{6.1}$$

For each of the distributions, we will prove the first formula in (6.1), the second one following similarly. For the normal distribution, the proof of this formula will be immediate. For the rest of the distributions, we will employ the following strategy. Suppressing the dependence on r and ν, which we consider to be fixed, we define the function

$$H(\mu, s) = I_\mu(s) - I_\mu(r) - I_\nu(s), \text{ for } \mu \leq r < \nu \leq s.$$

It is easy to check that $H(r, s) = I_r(s) - I_\nu(s) > 0$ and $H(\mu, \nu) = I_\mu(\nu) - I_\mu(r) > 0$. Thus, to prove the first formula in (6.1) it suffices to show that

$$\frac{\partial H}{\partial \mu}(\mu, s) \leq 0, \ \mu < r; \quad \frac{\partial H}{\partial s}(\mu, s) \geq 0, \ s > \nu.$$

We now turn to each of the distributions in the proposition. The formulas for the Legendre-Fenchel transformation can be derived directly from their definition in (1.2). They can be found in various places in the literature; for example, some of them appear in [2]. We specify each one on the subset of $\mathbb{R}$ on which it is finite.

Gaussian distributions with fixed variance: The Legendre-Fenchel transformation of the $\mathrm{N}(\mu, \sigma^2)$-distribution is given by $I_\mu(r) = \frac{(r-\mu)^2}{2\sigma^2}$, for $r \in \mathbb{R}$. Thus the first formula in (6.1) reduces to $(r-\mu)^2 + (s-\nu)^2 < (s-\mu)^2$, which holds due to the inequality $A^2 + B^2 < (A+B)^2$, for $A, B > 0$.

Exponential distributions: The exponential distribution $\mathrm{Exp}(\frac{1}{\mu})$ with parameter $\frac{1}{\mu} > 0$ has mean μ. Its Legendre-Fenchel transformation is given by $I_\mu(r) = \frac{r}{\mu} - \log\frac{r}{\mu} - 1$, for $r > 0$. Thus,

$$H(\mu, s) = \frac{s-r}{\mu} - \frac{s}{\nu} + \log\frac{r}{\nu} + 1.$$

We have $\frac{\partial H}{\partial \mu}(\mu, s) = \frac{r-s}{\mu^2} < 0$ and $\frac{\partial H}{\partial s}(\mu, s) = \frac{1}{\mu} - \frac{1}{\nu} > 0$..

Poisson distributions: The Poisson distribution $\mathrm{Poiss}(\mu)$ with parameter $\mu > 0$ has mean μ. Its Legendre-Fenchel transformation is given by $I_\mu(r) = r\log\frac{r}{\mu} - (r-\mu)$, for $r \geq 0$. Thus,

$$H(\mu, s) = s\log\frac{\nu}{\mu} - r\log\frac{r}{\mu} + r - \nu.$$

We have $\frac{\partial H}{\partial \mu}(\mu, s) = \frac{r-s}{\mu} < 0$ and $\frac{\partial H}{\partial s}(\mu, s) = \log\frac{\nu}{\mu} > 0$.

Geometric distributions: The geometric distribution $\text{Geom}(\frac{1}{\mu})$ supported on $\mathbb{N}$ with parameter $\frac{1}{\mu}\in(0,1)$ has mean μ. Its Legendre-Fenchel transformation is given by $I_\mu(r)=(r-1)\log\frac{r-1}{r}\frac{\mu}{\mu-1}+\log\frac{\mu}{r}$, for $r\ge 1$. Thus,

$$H(\mu,s)=(s-1)\log\frac{\mu(\nu-1)}{\nu(\mu-1)}-(r-1)\log\frac{\mu(r-1)}{r(\mu-1)}+\log\frac{r}{\nu}.$$

We have $\frac{\partial H}{\partial\mu}(\mu,s)=(s-r)(\frac{1}{\mu}-\frac{1}{\mu-1})<0$ and $\frac{\partial H}{\partial s}(\mu,s)=\log\frac{\mu(\nu-1)}{\nu(\mu-1)}>0$.

Bernoulli distributions The Bernoulli distribution $\text{Ber}(\mu)$ with parameter $\mu\in(0,1)$ has mean μ. Its Legendre-Fenchel transformation is given by $I_\mu(r)=r\log\frac{r}{\mu}+(1-r)\log\frac{1-r}{1-\mu}$, for $r\in[0,1]$. Thus,

$$H(\mu,s)=s\log\frac{\nu}{\mu}+(1-s)\log\frac{1-\nu}{1-\mu}-r\log\frac{r}{\mu}-(1-r)\log\frac{1-r}{1-\mu}.$$

We have $\frac{\partial H}{\partial\mu}(\mu,s)=(r-s)(\frac{1}{1-\mu}+\frac{1}{\mu})<0$ and $\frac{\partial H}{\partial s}(\mu,s)=\log\frac{\nu(1-\mu)}{\mu(1-\nu)}>0$. □

References

[1] Dembo, A. and Zeitouni, O., *Large Deviations Techniques and Applications* 2nd edition, Springer-Verlag, New York, (1998).

[2] Durrett, R. *Probability: theory and examples.* Third edition, Brooks/Cole (2005).

[3] Kosygina, E. and Zerner M. *Excited random walks: results, methods, open problems*, Bull. Inst. Math. Acad. Sin. (N.S.) **8** (2013), 105-157.

[4] Pemantle, R. *A survey of random processes with reinforcement*, Probab. Surv. **4** (2007), 1-79.

[5] Pinsky, R. *The speed of a random walk excited by its recent history*, Adv. in Appl. Probab. 48 (2016), 215-234.

Department of Mathematics, Technion—Israel Institute of Technology, Haifa, 32000, Israel

E-mail address: `pinsky@math.technion.ac.il`

URL: `http://www.math.technion.ac.il/~pinsky/`