Vertex-reinforced random walk on Z has finite range

Robin Pemantle [1,2]
Stanislav Volkov [3]

ABSTRACT:

A stochastic process called Vertex-Reinforced Random Walk (VRRW) is defined in Pemantle (1988a). We consider this process in the case where the underlying graph is an infinite chain (i.e., the one-dimensional integer lattice). We show that the range is almost surely finite, that at least 5 points are visited infinitely often almost surely, and that with positive probability the range contains exactly 5 points. There are always points visited infinitely often but at a set of times of zero density, and we show that the number of visits to such a point to time n may be asymptotically n^α for a dense set of values $\alpha \in (0,1)$. The power law analysis relies on analysis of a related urn model.

Keywords: Vertex-reinforced random walk, Reinforced random walk, VRRW, Urn model, Bernard Friedman's urn

Subject classification: Primary: 60G17 ; Secondary 60J20

---

[1]Research supported in part by a Sloan Foundation Fellowship, and by a Presidential Faculty Fellowship

[2]Department of Mathematics, University of Wisconsin-Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706

[3]The Fields Institute for Research in Mathematical Sciences, 222 College Street, Toronto, Ontario, Canada, M5T3J1

1 Outline of results

For any process $X_0, X_1, X_2, \ldots$ taking values in the vertex set of a graph G (throughout this paper $G = \mathbf{Z}$), we define the augmented occupation numbers

$$Z(n,v) = 1 + \sum_{i=0}^{n} \mathbf{1}_{X_i = v}$$

to be the number of times plus one that the process visits site v up through time n. Let G be any locally finite graph, with the neighbor relation denoted by $\sim$, and define vertex-reinforced random walk (VRRW) on G with starting point $v \in V(G)$ to be the process $\{X_i : i \geq 0\}$ such that $X_0 = v$ and

$$\mathbf{P}(X_{n+1} = x \mid \mathcal{F}_n) = \mathbf{1}_{x \sim X_n} \frac{Z(n,x)}{\sum_{w \sim X_n} Z(n,w)} .$$

In other words, moves are restricted to the edges of G, with the probability of a move to a neighbor w being proportional to the augmented occupation of w at that time.

This is a special case of the weighted VRRW, defined by Pemantle (1988a, 1992), where each oriented edge $\vec{vw}$ carries a nonnegative weight $\lambda(v,w)$, and the transition probabilities are given by

$$\mathbf{P}(X_{n+1} = x \mid \mathcal{F}_n) = \mathbf{1}_{x \sim X_n} \frac{\lambda(X_n, x) Z(n,x)}{\sum_{w \sim X_n} \lambda(X_n, w) Z(n,w)} .$$

It is shown in Pemantle (1988a, 1992) that for generic symmetric values of λ and finite graphs G, the vector of normalized occupation measure, $(Z(n,v)/n)_{v \in V(G)}$, must converge to an element of a set of equilibrium points which is typically finite. Unfortunately, the case where λ is identically 1 is not generic, but rather degenerate from the point of view of the previous works, and so this one most natural case is left unanalyzed.

While the results of Pemantle (1992) do not extend to the case in this paper, it was conjectured there that in such cases the range of VRRW will be finite, and that in fact it will get "stuck" in a set of 3 points. In this paper we show that this behavior holds, in the sense of normalized occupation measure, at least with positive probability (Theorem 1.3

below). If one cares about the set of points visited infinitely often, rather than with positive density, the size of the set on which the walk gets stuck is 5 rather than 3. We obtain the following further results about the size of the range.

Theorem 1.1 *Let $R = \{k : X_n = k$ for some $n\}$ be the (random) range of the process $X_0, X_1, \ldots$. Then $\mathbf{P}(|R| = 5) > 0$ and $\mathbf{P}(|R| < \infty) = 1$.*

Theorem 1.2 *Let $R' = \{k : X_n = k$ infinitely often$\}$ be the essential range of the process $X_0, X_1, \ldots$. Then $\mathbf{P}(|R'| \leq 4) = 0$.*

Remark: Simulations appear to show that $\mathbf{P}(|R'| = 4)$ is nonzero, but these are evidently misleading.

We conjecture but cannot prove that $\mathbf{P}(|R'| = 5) = 1$. It is easy to see that if $R' = \{k, k+1, \ldots, k+j\}$, then $Z(n,k)/n$ and $Z(n,k+j)/n$ both converge to zero, or in other words, the occupation density goes to zero at the endpoints of the range. Quantitatively, we have

Theorem 1.3 *For any closed interval $I \subseteq (0,1)$ and any integer k, there is with positive probability an $\alpha \in I$ such that the following six events occur:*

(*i*) $R' = \{k-2, k-1, k, k+1, k+2\}$;
(*ii*) $\log Z(n, k+2)/\log n \to \alpha$;
(*iii*) $\log Z(n, k-2)/\log n \to 1-\alpha$;
(*iv*) $Z(n, k+1)/n \to \alpha/2$;
(*v*) $Z(n, k-1)/n \to (1-\alpha)/2$;
(*vi*) $Z(n, k)/n \to 1/2$.

We conjecture but cannot prove that this is the universal behavior, i.e., that there is always such an $\alpha \in (0,1)$.

The power law behavior in parts (ii) and (iii) of Theorem 1.3 rests on the analysis of a certain interacting urn process. This urn process is of a type studied in the doctoral dissertation of Athreya (1967), via embedding in a multiptype branching process. Since this is not generally available, and since our hypotheses and methods of proof are quite different, we include complete statements and proofs of the relevant results. The next section gives some background on urn processes and reinforced random walks. Proofs for the results on urns are given in Section 3, and proofs of Theorems 1.1, 1.2, and 1.3 are given in Section 4. The final section completes the proofs of some lemmas and poses a few open questions.

2 Background on urn processes and processes with reinforcement

This section begins with a brief survey of previously studied reinforced random processes. By popular demand, we have included more of a review than is strictly necessary for the analysis of VRRW, that being the generalization of Theorem 2.2 stated and proved in the next section.

The simplest (and one of the oldest) process with reinforcement is known as *Pólya's urn*, after the 1927 paper of Eggenberger and Pólya. In this model, there is an urn containing red and blue balls. At time 0 the urn contains r red balls and s blue balls. At each time $k \geq 1$, a ball is chosen uniformly from the contents of the urn, and is put back into the urn along with a extra balls of the same color. Thus if X_n denotes the number of red balls at time n and Y_n denotes the number of blue balls at time n, the dynamics are governed by

$$\begin{aligned}(X_{n+1}, Y_{n+1}) &= (X_n + a, Y_n) \qquad \text{with probability } \frac{X_n}{X_n + Y_n}; \\ (X_{n+1}, Y_{n+1}) &= (X_n, Y_n + a) \qquad \text{with probability } \frac{Y_n}{X_n + Y_n}.\end{aligned} \tag{2.1}$$

Eggenberger and Pólya showed that the proportion of red balls, $Z_n := X_n/(X_n + Y_n)$, converges almost surely, and that the limit is random. The distribution of the limit is a beta with parameters r/a and s/a (thus uniform over $[0,1]$ when $r = s = a = 1$). This random limit behavior is possible because Z_n is a martingale. In the sections to follow, we make use several times of this elementary principle:

Proposition 2.1 *If $Z_n = X_n/(X_n+Y_n)$, and if $(X_{n+1}, Y_{n+1}) = (X_n+1, Y_n)$ with probability $X_n/(X_n + Y_n)$, and $(X_n, Y_n + 1)$ otherwise, then $\mathbf{E}(Z_{n+1} \mid Z_n) = Z_n$. Also, $|Z_{n+1} - Z_n| < 1/(X_n + Y_n)$.* $\square$

In contrast to this is an variant suggested by B. Friedman (1949), where in addition to the a extra balls of the same color, one also adds b balls of the opposite color. This produces strikingly different behavior, even when $b << a$. Freedman (1965) showed that $Z_n \to 1/2$ almost surely, with $(Z_n - 1/2)/n^{-\gamma}$ converging to a nontrivial distribution for $\gamma > 0$ depending on a and b. To explain the differing behavior, note that $\{Z_n\}$ is not a martingale, but rather

$$\mathbf{E}(Z_{n+1} - Z_n \mid Z_n) = n^{-1}(f(Z_n) + o(1)) \tag{2.2}$$

where f is a function vanishing only at $1/2$. One could say that the drift, f, pushes $\{Z_n\}$ toward $1/2$, which is an attracting point for the one-dimensional vector field given by f.

When discussing processes with reinforcement, it is good to keep in mind the distinction *Pólya-like* ($f \equiv 0$) versus *Friedman-like* ($f \neq 0$ except at isolated points), which dictates the important aspects of the long-term behavior. A third category, *singular*, occurs when f has zeros on the boundary, in which case convergence happens at a slower rate.

The prototypical Friedman-like model is Robbins and Monro's (1951) stochastic approximation scheme, which obeys the law

$$\mathbf{E}(Z_{n+1} - Z_n \mid Z_n) = n^{-1} F(Z_n)$$

for a generic function F about which imprecise information can be obtained by sampling. The (unknown) zeros of F are then "found" by the $\{Z_n\}$ process. Since the 1950's, stochastic approximation has been an active research area; the overview by Kushner and Yin (1997) gives an idea of progress and techniques in stochastic approximation since then. The literature on formal models of learning contains many Friedman-like processes, in which the transition probabilities of a finite state (non-Markov) chain are updated based on the some kind of objective function. The chain then "learns" to spend most of its time at states with large values of the objective function. The first round of this literature appeared in the late sixties, e.g., Iosifescu and Theodorescu (1969), and a second round emerged with the study of neural nets. The common theme is self-organization by a system whose basic parameters are extremely simple.

Pólya-like models have appeared frequently in theoretical statistics, due to the fact that bounded martingales are mathematically equivalent to sequences of posteriors, given increasing σ-fields. For example, suppose an IID sequence of zeros and ones has an unknown mean p, with the prior on p being uniform on $[0,1]$ (or more generally, any beta distribution). Then the sample sequences $\{Z_n\}$ of a Pólya urn process can be interpreted as posterior means, where each red ball picked corresponds to observing a one and each blue ball picked corresponds to observing a zero. The so-called Bayes-Laplace estimate of the probability the sun will rise tomorrow and Greenwood and Yule's (1920) model for industrial accidents are both based on this interpretation. In modern times, Blackwell and McQueen (1973) contruct Ferguson's Dirichlet via an urn process, and Mauldin, Sudderth and Williams (1992) use a tree full of urns to construct a family of priors on distributions on $[0,1]$ with nice properties.

Pólya-like models have also arisen in modeling of self-organization and random limits. Arthur (1986) and Arthur et al (1987) use both Pólya- and Friedman-like urns to model the growth of industry and explain random clustering and market share patterns. Reinforced random walks were introduced by Coppersmith and Diaconis (1987) as another, somewhat simplified model of self-organized behavior. Although simplistic, urn models and reinforced random walks have been taken seriously in the modeling of physical phenomena; see for example Othmer and Stevens (1998), in which motion and aggregation of myxobacteria along slime trails are modeled by reinforced random walks and related stochastic cellular automata.

The VRRW studied by Pemantle (1988a , 1992) is a variant of their edge-reinforced random walk (ERRW). In ERRW, one keeps track of the number of times each edge has been crossed, the augmented occupation numbers being denoted $\{Z(n, \{v, w\}) : \{v, w\} \in E(G)\}$, and one chooses the next edge from among the edges adjacent to the present vertex, with probabilities proportional to the augmented occupation of each edge:

$$\mathbf{P}(X_{n+1} = x \mid \mathcal{F}_n) = \mathbf{1}_{x \sim X_n} \frac{Z(n, \{X_n, x\})}{\sum_{w \sim X_n} Z(n, \{X_n, w\})} .$$

Reinforcing edges rather than vertices makes a dramatic difference in the behavior of the process, because edge-reinforcement is Pólya-like and vertex-reinforcement is Friedman-like. A depiction of this difference via simulation may be found in Othmer and Stevens (1998). To see how to account for the difference theoretically, let v is a vertex in an acyclic graph, with incident edges $e_1, \ldots, e_k$. The successive edges chosen each time v is visited form a Pólya urn process and it is not hard to see that these are independent as v ranges over all vertices. (The analogue of this fact on a graph with cycles is much harder to formulate and prove.) Coppersmith and Diaconis (1987) proved that the normalized occupation measure of ERRW on a finite graph converges to a random vector having a nonzero density with respect to Lebesgue measure on the simplex. When $G = \mathbf{Z}$, Pemantle (1988b) shows that the process is a mixture of positive recurrent Markov chains, and in particular, that the normalized occupation measure converges to a limit that is everywhere positive.

The question of the behavior of either VRRW or ERRW on a lattice of dimension two or greater is still open. Some progress on ERRW has been made by generalizing the model so that the k^{th} crossing of each edge adds a_k to the occupation, where $\{a_k\}$ is a pre-specified sequence ($a_k \equiv 1$ in standard ERRW). A general recurrence/transience dichotomy for this model was obtained by Davis (1990) in one dimension, while Sellke (1994) has results on the coordinate processes for this model in two dimensions.

Our results for one-dimensional VRRW depend on an analysis of an urn model generalizing both the Pólya and the Friedman urn. Replace the dynamics (2.1) by the more general dynamics:

$$(X_{n+1}, Y_{n+1}) = (X_n + a, Y_n + b) \qquad \text{with probability } \frac{X_n}{X_n + Y_n}; \tag{2.3}$$
$$(X_{n+1}, Y_{n+1}) = (X_n + c, Y_n + d) \qquad \text{with probability } \frac{Y_n}{X_n + Y_n}.$$

There is no assumption that the number of balls be integral. When $\binom{a\ b}{c\ d}$ is a multiple of the identity matrix, we recover Pólya's urn, and when $a = d$ and $b = c$ are all nonzero, we recover Friedman's urn. In any case where $\binom{a\ b}{c\ d}$ has an eigenvector (v_1, v_2) with positive components, Freedman's analysis can be carried through to show that $X_n/(X_n + Y_n)$ converges to $v_1/(v_1 + v_2)$. Perhaps the cleanest way to do this is via embedding in a branching process, as described in Athreya and Ney (1972, chapter V, sec. 9). Thus in particular this holds when $bc > 0$. Two interesting cases are the singular cases, which can be reduced without loss of generality to the cases in the next two theorems. Theorem 2.2 was first proved by Athreya (1967) in a different form, while Theorem 2.3 is derivable from his results.

Theorem 2.2 *Suppose $a > d = 1$ and $b = c = 0$. Then X_n/Y_n^a converges almost surely to a random limit in $(0, \infty)$.*

Theorem 2.3 *Suppose $a = d = 1, b = 0$ and $c > 0$. Then $X_n/(cY_n) - \log Y_n$ converges to a random limit in $(-\infty, \infty)$.*

Remarks: (1) Theorem 2.3 is in a sense a finer result than Theorem 2.2, since it deals with the second order correction: Y_n is like $n/\log n$ multiplied by a specific constant, with a random correction of lower order: $X_n \approx cY_n(A + \log Y_n)$, where A is random. (2) The class of urns in Theorem 2.3 is not needed for analysis of VRRW on $\mathbf{Z}$, but is relevant to VRRW for a different reason. In the case $c = 1$, there is an isomorphism between the urn process and VRRW on the graph G with $V(G) = \{A, B\}$, having one edge between A and B and one loop connecting A to itself. Thus VRRW on G spends roughly time $n/\log n$ at B up to time n.

3 Urn model proofs

This section is devoted to proving Lemma 3.5, which generalizes Theorem 2.2 to allow random increments. Whereas Athreya (1967) proved version of this by embedding in a multitype branching process, we use martingale arguments (also considered by Athreya in some subcases). These turn out to be easier in the case of Theorem 2.3 than in the case of Lemma 3.5 below. Consequently, we first give a relatively short proof of Theorem 2.3 and then state and prove Lemma 3.5. Depending on your tastes, you may find the shorter proof or the more modular general proof easier to follow. Begin with the following easy lemma.

Lemma 3.1 *Let the nonnegative matrix* $\binom{a\ b}{c\ d}$ *satisfy* $(a+c)(b+d) > 0$ *and define an urn process as in (2.3). Then* $\min\{X_n, Y_n\} \to \infty$ *almost surely.*

Proof: The proportion of red balls at time n is always at least $X_0/(X_0+Y_0+n(a+b+c+d))$. Since the sum of these quantities is infinite, the Borel-Cantelli Lemma tells us that a red ball is chosen infinitely often. Similarly, a blue ball is chosen infinitely often. After each color has been chosen k times, $\min\{X_n, Y_n\}$ is at least $k \min\{a+c, b+d\}$. $\square$

The following general fact about convergence of random sequences is also useful.

Lemma 3.2 *Let* $\{Z_n : n \geq 0\}$ *be a random sequence measurable with respect to the filtration* $\{\mathcal{F}_n\}$. *Define*

$$\Delta_n = \mathbf{E}(Z_{n+1} - Z_n \,|\, \mathcal{F}_n) \quad ; \quad Q_n = \mathbf{E}((Z_{n+1} - Z_n)^2 \,|\, \mathcal{F}_n).$$

Then as n goes to infinity, Z_n *converges to a finite value almost surely on the event* $\sum_n \Delta_n < \infty$ *and* $\sum_n Q_n < \infty$.

Proof: Let τ_M be the first time n that $\sum_{j=0}^n Q_j > M$. Let

$$Z_n^{(M)} = Z_{n \wedge \tau_M} - \sum_{j=0}^{n \wedge \tau_M} \Delta_j \, .$$

Observe that $\{Z_n^{(M)}\}$ is a martingale with

$$\mathbf{E}((Z_{n+1}^{(M)} - Z_n^{(M)})^2 \,|\, \mathcal{F}_n) \leq \mathrm{Var}(Z_{n+1} - Z_n \,|\, \mathcal{F}_n)\mathbf{1}_{\tau_M > n} \leq Q_n \mathbf{1}_{\tau_M > n}$$

and so $Z_n^{(M)}$ converges almost surely and in L^2 to a finite limit, C_M. On the event $\{\sum_n Q_n < \infty\}$, τ_M will be infinite for sufficiently large M, and the sequence $\{Z_n\}$ will converge to $C_M + \sum_n \Delta_n$. □

Proof of Theorem 2.3: Let

$$Z_n = \frac{X_n}{cY_n} - \log Y_n \, .$$

We wish to apply Lemma 3.2 to $\{Z_n : n \geq 0\}$, so we must compute Δ_n and Q_n. It will turn out that $\Delta_n = O(1/Y_n^2)$ and $Q_n = O(n/Y_n^3)$, so we are going to need a preliminary lower bound on the growth rate of Y_n in order to see that these are almost surely summable.

Lemma 3.3 *For any $\epsilon > 0$, the function $X_n/Y_n^{1+\epsilon}$ is a supermartingale when X_n and Y_n are both at least $c+2$. It follows that Y_n is almost surely eventually greater than any power of n less than 1.*

Proof: To see that $X_n/Y_n^{1+\epsilon}$ is a supermartingale we compute the expected increment.

$$\begin{aligned}
&\mathbf{E}\left(\frac{X_{n+1}}{Y_{n+1}^{1+\epsilon}} - \frac{X_n}{Y_n^{1+\epsilon}} \,|\, \mathcal{F}_n\right) \\
= \;& \frac{X_n}{X_n + Y_n}\frac{1}{Y_n^{1+\epsilon}} + \frac{Y_n}{X_n + Y_n}\frac{c}{Y_n^{1+\epsilon}} - \frac{Y_n}{X_n + Y_n}\frac{X_n((Y_n+1)^{1+\epsilon} - Y_n^{1+\epsilon})}{Y_n^{1+\epsilon}(Y_n+1)^{1+\epsilon}} \\
= \;& \frac{1}{(X_n + Y_n)Y_n^{1+\epsilon}}\left(X_n + cY_n - X_n Y_n - (Y_n/(Y_n+1))^{1+\epsilon}\right) .
\end{aligned}$$

This is nonpositive when $\min\{X_n, Y_n\} \geq c+2$, proving that $X_n/Y_n^{1+\epsilon}$ is a supermartingale under this condition. By Lemma 3.1, both X_n and Y_n converge to infinity, so there is an almost surely finite $N = N(\epsilon, \omega)$ such that $\min\{X_n, Y_n\} \geq c+2$ for $n \geq N$, and consequently, $\{X_n/Y_n^{1+\epsilon} : n \geq N\}$ is a supermartingale. This implies that $\limsup_n X_n/Y_n^{1+\epsilon} < \infty$, and hence for any $0 < \epsilon < \delta$, that $\limsup_n X_n^{(1+\delta)^{-1}}/Y_n = 0$, proving the lemma. □

We continue with the proof of Theorem 2.3. We first compute the expected increment $\Delta_n := \mathbf{E}(Z_{n+1} - Z_n \,|\, \mathcal{F}_n)$ of Z_n.

$$\begin{aligned}
\Delta_n &= \frac{X_n}{X_n + Y_n}\frac{1}{cY_n} + \frac{Y_n}{X_n + Y_n}\frac{c}{cY_n} + \frac{Y_n}{X_n + Y_n}\left(-\frac{X_n}{cY_n(Y_n+1)} - \log\frac{Y_n + 1}{Y_n}\right) \\
&= \frac{1}{X_n + Y_n}\left(\frac{X_n}{cY_n} + 1 - \frac{X_n}{c(Y_n+1)} - Y_n \log(1 + 1/Y_n)\right) \\
&= \frac{1}{X_n + Y_n}\left(\frac{X_n}{cY_n(Y_n+1)} + O(\frac{1}{Y_n})\right) \\
&= O(\frac{1}{Y_n^2}). \qquad (3.1)
\end{aligned}$$

Now compute an upper bound for the quadratic variation $Q_n := \mathbf{E}((Z_{n+1} - Z_n)^2 \,|\, \mathcal{F}_n)$ as follows.

$$\begin{aligned}
Q_n &= \frac{X_n}{X_n + Y_n}\frac{1}{c^2Y_n^2} + \frac{Y_n}{X_n + Y_n}\left(\frac{1}{Y_n} - \frac{X_n}{cY_n(Y_n+1)} - \log(1 + \frac{1}{Y_n})\right)^2 \\
&\leq \frac{1}{c^2Y_n^2} + \frac{1}{Y^2} + \log^2(1 + \frac{1}{Y_n}) + \frac{Y_n}{X_n + Y_n}\frac{X_n^2}{c^2Y_n^4} \\
&\leq \frac{n}{CY_n^3} \qquad (3.2)
\end{aligned}$$

for an appropriate constant C, using the fact that $X_n \leq (1 + c)n + X_0$. We are now done: Lemma 3.3 together with (3.1) and (3.2) show that Δ_n and Q_n are almost surely summable, hence the conclusion of the theorem follows from Lemma 3.2. $\square$

Proof of Theorem 2.2: We now prove Lemmas 3.4 and 3.5, which together imply as a special case a result in the spirit of Theorem 2.2. We use supermartingales similar to those in Lemma 3.3.

Lemma 3.4 *Let (X_n, Y_n) be a positive process converging coordinatewise to infinity. Fix any $\beta > 1$ and suppose there is an $M = M(\beta) \leq \infty$ such that Y_n^β / X_n is a supermartingale once $X_n, Y_n \geq M$. Then*

$$\limsup_n \frac{\log Y_n}{\log X_n} \leq \frac{1}{\beta} \text{ on } \{M < \infty\}\,.$$

Similarly, if X_n/Y_n^β is a supermartingale once $X_n, Y_n \geq M'$, then

$$\liminf_n \frac{\log Y_n}{\log X_n} \geq \frac{1}{\beta} \text{ on } \{M < \infty\}.$$

PROOF: Let τ_m be the least $n \geq m$ for which $\min\{X_n, Y_n\} < M$. Then

$$\{Y_{n\wedge\tau_m}^\beta / X_{n\wedge\tau_m} : n \geq m\}$$

is a nonnegative supermartingale, so converges almost surely to a limit $L(m)$. When $\tau_m = \infty$, it follows that $L(m)$ is the almost sure limit of Y_n^β/X_n. Thus when $\tau_m = \infty$, $Y_n = [((L(m) + o(1))X_n]^{1/\beta}$. If $L(m) > 0$ this implies $\limsup \log Y_n / \log X_n = 1/\beta$, while if $L(m) = 0$, the lim sup may be strictly less than $1/\beta$. On the event $\{M < \infty\}$ an m exists with $\tau_m = \infty$, which finishes the proof of the first assertion. The proof of the second assertion is similar. $\square$

Lemma 3.5 *Let (X_n, Y_n) be a process generalizing the urn process in Theorem 2.2 as follows. For each n, with probability $X_n/(X_n+Y_n)$, there is a $W > 0$ such that $X_{n+1} = W + X_n$ and $Y_{n+1} = Y_n$; with probability $Y_n/(X_n + Y_n)$, we have $X_{n+1} = X_n$ and $Y_{n+1} = Y_n + 1$. Suppose further that with probability 1,*

$$\mathbf{E}(W \,|\, \mathcal{F}_n, X_{n+1} > X_n) \in [a, b] \tag{3.3}$$

and

$$\mathbf{E}(W^2 \,|\, \mathcal{F}_n, X_{n+1} > X_n) \leq K \tag{3.4}$$

for some positive constant K and some $0 < a \leq b$. Then for any $0 < \beta < a$, there is an M such that whenever $Y_n \geq M$, the function Y_n^β/X_n is a supermartingale. Likewise, for any $\beta > b$ there is an M' such that X_n/Y_n^β is a supermartingale whenever $Y_n \geq M'$.

PROOF: By a Taylor expansion, there exist constants c_1 and c_2 such that for any w, and any sufficiently large x,

$$\frac{1}{x+w} \leq \frac{1}{x} - \frac{w}{x^2} + c_1 \frac{w^2}{x^3}.$$

Also,

$$(y+1)^\beta \le y^\beta + \beta y^{\beta-1} + c_2 \beta y^{\beta-2}.$$

The expected increment $\Delta_n := Y_{n+1}^\beta / X_{n+1} - Y_n^\beta / X_n$, conditional on $\mathcal{F}_n$, is given by

$$\frac{X_n}{X_n+Y_n}\mathbf{E}\left(\frac{Y_n^\beta}{X_n+W} - \frac{Y_n^\beta}{X_n}\right) + \frac{Y_n}{X_n+Y_n}\mathbf{E}\left(\frac{(Y_n+1)^\beta}{X_n} - \frac{Y_n^\beta}{X_n}\right) .$$

Plugging in the Taylor estimates above yields

$$\mathbf{E}\Delta_n \le \frac{X_n}{X_n+Y_n}\frac{Y_n^\beta}{X_n^2}\left(-\mathbf{E}W + c_1\frac{\mathbf{E}W^2}{X_n}\right) + \frac{Y_n}{X_n+Y_n}\frac{\beta Y_n^{\beta-1}}{X_n}\left(1+\frac{c_2}{Y_n}\right) .$$

The assumptions on W imply that

$$\mathbf{E}\Delta_n \le \frac{Y_n^\beta}{(X_n+Y_n)X_n}\left(\beta - a + \frac{c_2}{Y_n} + \frac{c_1 K}{X_n}\right) .$$

When $\beta < a$ and X_n and Y_n are sufficiently large, then this is nonpositive. Choosing M large enough so that the constants c_1 and c_2 in the Taylor expansion are valid whenever $X_n, Y_n \ge M$, we have proved the first assertion of the lemma.

The proof of the second assertion is similar. Choose c so that

$$(y+1)^{-\beta} \le y^{-\beta} - \beta y^{-\beta-1}(1-\frac{c}{y})$$

whenever $y \ge 1$. The expected increment $\Delta_n := X_{n+1}/Y_{n+1}^\beta - X_n/Y_n^\beta$, conditional on $\mathcal{F}_n$, is given by

$$\frac{X_n}{X_n+Y_n}\mathbf{E}\frac{W}{Y_n^\beta} + \frac{Y_n}{X_n+Y_n}\mathbf{E}\left(\frac{X_n}{(Y_n+1)^\beta} - \frac{X_n}{Y_n^\beta}\right) .$$

Thus

$$\mathbf{E}\Delta_n \le \frac{X_n}{(X_n+Y_n)Y_n^\beta}\left(b - \beta + \frac{\beta c}{Y_n}\right) ,$$

proving the lemma for $M' = \beta c/(b-\beta)$. □

Finally, we show Lemma 3.4 and Lemma 3.5 together imply the first order of approximation in Theorem 2.2, namely that $\log X_n / \log Y_n \to a$ almost surely. The urn process in

Theorem 2.2 satisfies the conditions of Lemma 3.5 with $K = a^2$ and $[a,b] = \{a\}$. Thus for any $0 < \beta < a$, we may plug the conclusion of Lemma 3.5 into Lemma 3.4 to see that

$$\limsup_n \frac{\log Y_n}{\log X_n} \le \frac{1}{\beta}\,.$$

Similarly, for any $\beta > a$, we plug the conclusion of Lemma 3.5 into Lemma 3.4 to see that

$$\liminf_n \frac{\log Y_n}{\log X_n} \ge \frac{1}{\beta}\,.$$

Since β may be chosen arbitrarily close to a, we see that $\log Y_n / \log X_n \to 1/a$ as $n \to \infty$. $\square$

4 Proof of Theorem 1.3

In the next section we will prove the following lemma:

Lemma 4.1 *There is an $\epsilon > 0$ such that for all integers $m > 0$,*

$$\mathbf{P}(m + 3 \in R \,|\, m \in R) \le 1 - \epsilon.$$

The first statement of Theorem 1.1 follows from Theorem 1.3. The second statement follows directly from Lemma 4.1: by induction, $\mathbf{P}(3n \in R) \le (1-\epsilon)^n$, which goes to zero as $n \to \infty$. Hence we concentrate on the proof of Theorem 1.3. Begin with a lemma.

Lemma 4.2 *Let $J = [a,b]$ be an interval of integers containing zero. Let $\mathbf{P}$ denote the law of VRRW on $\mathbf{Z}$ as before, and let $\mathbf{P}_J$ denote the law of a VRRW on the interval J, both started from 0. Then the following four conditions are equivalent.*

(i) $\mathbf{P}(R \subseteq J) > 0$;

(ii) $\mathbf{P}(R' \subseteq J) > 0$;

(iii) $\mathbf{P}\left(\sum_n \mathbf{1}_{X_n = a} Z(n, a+1)^{-1} + \sum_n \mathbf{1}_{X_n = b} Z(n, b-1)^{-1} < \infty\right) > 0$.

(iv) $\mathbf{P}_J\left(\sum_n \mathbf{1}_{X_n = a} Z(n, a+1)^{-1} + \sum_n \mathbf{1}_{X_n = b} Z(n, b-1)^{-1} < \infty\right) > 0$.

PROOF: We define a coupling, i.e., a measure Q on pairs of paths $(\{X_n : n \geq 0\}, \{X'_n : n \geq 0\})$ such that the first coordinate of Q has law $\mathbf{P}$ and the second has law $\mathbf{P}_J$. To do so, choose $\{X_n\}$ according to $\mathbf{P}$ and let τ be the first time n that $X_n \in \{a-1, b+1\}$. Let $X'_n = X_n$ for $n < \tau$, let $X'_\tau = a+1$ if $X_\tau = a-1$, let $X'_\tau = b-1$ if $X_\tau = b+1$, and let X'_n be chosen from the transition probabilities for $\mathbf{P}_J$ independently of $\{X_n : n \geq 0\}$ when $n > \tau$.

Observe that

$$Q(\tau = n+1 \,|\, \tau > n) = \mathbf{1}_{X'_n = a} Z(n, a+1)^{-1} + \mathbf{1}_{X'_n = a} Z(n, a+1)^{-1}.$$

Thus by Borel-Cantelli, $Q(\tau = \infty) > 0$ if and only if condition (iv) is satisfied. The event $\{\tau = \infty\}$ is the same as the event $\{R \subseteq J\}$, proving the equivalence of (i) and (iv). Similarly, from the equation

$$Q(\tau = n+1 \,|\, \tau > n) = \mathbf{1}_{X_n = a} Z(n, a+1)^{-1} + \mathbf{1}_{X_n = a} Z(n, a+1)^{-1}$$

one sees that (i) and (iii) are equivalent. The implication $(i) \Rightarrow (ii)$ is clear. Finally, to see that (ii) implies (iii), assume (ii). Thus with positive probability, $Z(n, a-1) + Z(n, b+1)$ is bounded as $n \to \infty$. By Borel-Cantelli, this means that

$$\mathbf{P}\left[\sum_n \mathbf{P}(X_n \in \{a-1, b+1\} \,|\, \mathcal{F}_n) < \infty\right] > 0\,.$$

This sum is an upper bound for the sum in (iii), hence the sum in (iii) is finite with positive probability. $\square$

Corollary 4.3 *Suppose that* $Z(n, a+1)$ *and* $Z(n, b-1)$ *are* $\Theta(n)$*, i.e.,* $\liminf Z(n, a+1)/n > 0$ *and* $\liminf Z(n, b-1)/n > 0$*. Then* $\mathbf{P}(R \subseteq J) > 0$ *if and only if*

$$\mathbf{P}_J\left(\sum_n \frac{Z(n,a) + Z(n,b)}{n^2} < \infty\right) > 0.$$

PROOF: Let σ_m be the first n for which $Z(n, a+1) = m$ and let ρ_m be the first n for which $Z(n, b-1) = m$. Summing by parts gives

$$\begin{aligned}
&\sum_k \mathbf{1}_{X_k=a} Z(k, a+1)^{-1} + \sum_k \mathbf{1}_{X_k=b} Z(k, b-1)^{-1} \\
= \quad & \sum_n \frac{Z(\sigma_{n+1}, a) - Z(\sigma_n, a)}{n} + \frac{Z(\rho_{n+1}, b) - Z(\rho_n, b)}{n} \\
= \quad & \sum_n \frac{Z(\sigma_n, a) - 1}{n^2 - n} + \frac{Z(\rho_n, b) - 1}{n^2 - n} .
\end{aligned}$$

Since $Z(n, r)$ is increasing in n for all r and we have assumed $\sigma_n, \rho_n = O(n)$, this proves the corollary. □

PROOF OF THEOREM 1.3: There are four steps to the proof. The first is to reduce to a VRRW on the five points $-2, -1, 0, 1$ and 2. The second is to show that this VRRW can, with positive probability, have $2Z(n,1)/n$ remain in the interval I, while simultaneously $Z(n,2)$ and $Z(n,-2)$ remain less than $n^{1-\epsilon}$ for a prescribed $\epsilon = \epsilon(I) > 0$. The third step is to show that when these two things happen, then actually $2Z(n,1)/n$ converges to some $\alpha \in I$. The fourth step is to see that whenever $2Z(n,1)/n$ converges, then $Z(n,2)$ almost surely obeys the power law

$$\lim_{n\to\infty} \frac{\log Z(n,2)}{\log n} = \lim_{n\to\infty} \frac{2Z(n,1)}{n} .$$

Step 1. This step is essentially done. If we show that for $J = \{k-2, k-1, k, k+1, k+2\}$, the $\mathbf{P}_J$ probability of properties (ii) - (vi) holding simultaneously is positive, then the conclusion of the theorem follows from (ii) - (v) and Corollary 4.3. The argument is the same for every k, so from now on we assume without loss of generality that $k = 0$, and set about proving Theorem 1.3 for $\mathbf{P}_J$ in place of $\mathbf{P}$, where $J = \{-2, -1, 0, 1, 2\}$.

Step 2. For the remainder of the argument, fix an interval $I = [c, d] \subseteq (0, 1)$ and a positive $\epsilon \leq \min\{c, 1-d, d-c\}/10$. Let $\beta = (1-\epsilon)^{-1}$. Also fix an integer N_0 and define stopping times depending on N_0 as follows. Let τ_1 be the least $n \geq N_0$ such that

$2Z(n,1)/n \notin I$. Let τ_2 be the least $n \geq N_0$ such that $Z(n,2) \geq n^{1-\epsilon}$ and let τ_3 be the least $n \geq N_0$ such that $Z(n,-2) \geq n^{1-\epsilon}$. Let $\tau = \tau_1 \wedge \tau_2 \wedge \tau_3$. Let $\{z_i : -2 \leq i \leq 2\}$ be a quintuple of integers. Our goal in this step is to identify an N_0 and a quintuple z_i such that

$$\mathbf{P}_J(\tau = \infty \mid Z(N_0, i) = z_i : -2 \leq i \leq 2) > 0\,; \tag{4.1}$$

in fact we will show it is near 1. We assume (4.1) for the moment, and continue with steps 3 and 4.

Step 3. Let κ_n be the time of the n^{th} return to the state 0 and define

$$V_n = \frac{Z(\kappa_n, 1)}{Z(\kappa_n, 1) + Z(\kappa_n, -1)}\,.$$

We will see below that when $\kappa_n < \tau_2 \wedge \tau_3$,

$$|\mathbf{E}_J(V_{n+1} - V_n \mid \mathcal{F}_{\kappa_n})| \;\leq\; C\kappa_n^{-1-\epsilon}; \tag{4.2}$$

$$\mathbf{E}_J((V_{n+1} - V_n)^2 \mid \mathcal{F}_{\kappa_n}) \;\leq\; C'\kappa_n^{-2}. \tag{4.3}$$

Plugging these two bounds into Lemma 3.2 shows that whenever $\tau = \infty$, the sequence V_n must converge, to a value necessarily in I. This gives us parts (iv) and (v) of the theorem, with part (vi) already following from step 2.

Step 4. We claim that for fixed r and s, whenever

$$r \leq \liminf V_n \leq \limsup V_n \leq s \tag{4.4}$$

and $\tau = \infty$, then

$$\frac{1}{s} \leq \liminf \frac{\log Z(n,2)}{\log n} \leq \limsup \frac{\log Z(n,2)}{\log n} \leq \frac{1}{r}\,.$$

To prove the claim, define the return times $\{\alpha_n : n \geq 0\}$ to state 1 by letting $\alpha_n = \min\{n > \alpha_{n-1} : X_n = 1\}$, and α_{-1} is set equal to $N_0 - 1$, for some N_0 Let

$$U_n = Z(\alpha_n, 0);$$

$$U_n' = Z(\alpha_n, 2);$$

For any $\delta > 0$, we show that the conditions of Lemma 3.5 are satisfied with $(X_n, Y_n) = (U_n, U_n')$ and $[a, b] = [r - \delta, s + \delta]$. Indeed, between times α_n and α_{n+1}, VRRW will either visit state 2 once or will visit state 0 some number of times $W \geq 1$. The probabilities of these disjoint cases are respectively $U_n/(U_n + U_n')$ and $U_n'/(U_n + U_n')$. Let N_1 be the least $N \geq N_0$ such that $r - \delta \leq \inf_{n \geq N} V_n \leq \sup_{n \geq N} V_n \leq s + \delta$; when (4.4) holds, N_1 will be finite. We need to show that when $n \geq N_1$, then (3.3) and (3.4) hold, with $[a, b] = [r - \delta, s + \delta]$. Since $\mathbf{P}_J(W \geq k)$ is equal to the probability that on the first $k - 1$ visits to state 0 after time α_n the VRRW moves to the left, the assumption that $n \geq N_1$ implies that

$$(1 - s - \delta)^k \leq \mathbf{P}_J(W \geq k) \leq (1 - r + \delta)^k,$$

which gives $1/(s + \delta) \leq \mathbf{E}_J W \leq 1/(r - \delta)$ and $\mathbf{E}_J W^2 \leq K$ for some constant $K = K(r, \delta)$. The conclusion of Lemma 3.5 is that $Y_n^{1/(s+2\delta)}/X_n$ and $X_n/Y_n^{1/(r-2\delta)}$ are supermartingales for $n \geq M$, where M will be finite when (4.4) holds. We then apply Lemma 3.4 together with the fact that $\alpha_n = O(n)$ on $\{\tau = \infty\}$ to see that

$$\frac{1}{s + 2\delta} \leq \liminf \frac{\log Z(n, 2)}{\log n} \leq \limsup \frac{\log Z(n, 2)}{\log n} \leq \frac{1}{r - 2\delta}$$

on (4.4) when $\tau = \infty$. Sending δ to 0 proves the claim.

Applying the claim simultaneously to all intervals (r, s) with rational endpoints, we see that conclusion (ii) of Theorem 1.3 holds with probability 1 whenever $\tau = \infty$. An identical argument establishes conclusion (iii). Since we have shown that $\mathbf{P}_J(\tau = \infty)$ may be made arbitrarily close to 1 by suitable choice of $\{z_i : -2 \leq i \leq 2\}$, we are done with all four steps, modulo the verification of (4.1), and of (4.2) and (4.3).

Cleanup step. First we prove (4.2) and (4.3). Let $\Delta_n = V_{n+1} - V_n$. We estimate Δ_n in three pieces. Let A be twice the least integer greater than $2/\epsilon$. Write $\Delta_n = R_n + S_n + T_n$ where

$$\begin{aligned}
R_n &= \frac{Z(\kappa_n + 2, 1)}{Z(\kappa_n + 2, 1) + Z(\kappa_n + 2, -1)} - V_n\,, \\
S_n &= \frac{Z(\kappa_{n+1} \wedge (\kappa_n + A), 1)}{Z(\kappa_{n+1} \wedge (\kappa_n + A), 1) + Z(\kappa_{n+1} \wedge (\kappa_n + A), -1)} - V_n - R_n\,, \\
T_n &= \Delta_n - R_n - S_n.
\end{aligned}$$

By Proposition 2.1, $\mathbf{E}_J(R_n \,|\, \mathcal{F}_{\kappa_n}) = 0$ and $R_n^2 \le \kappa_n^{-2}$. By the same token, $S_n^2 \le A^2 \kappa_n^{-2}$ and we easily see that

$$\mathbf{E}_J(|S_n| \,|\, \mathcal{F}_{\kappa_n}) \le \frac{A}{\kappa_n} \mathbf{P}_J(\kappa_{n+1} > \kappa_n + 2 \,|\, \mathcal{F}_{\kappa_n}) \le \frac{2N_0^{1-\epsilon}}{N_0 - 2N_0^{1-\epsilon}} \frac{A}{\kappa_n^{1+\epsilon}}$$

when $\kappa_n < \tau_2 \wedge \tau_3$ (as $n/2 - n^{1-\epsilon} \le \kappa_n \le n$ and $Z(\kappa_n, 0) = \kappa_n$). Finally, since $T_n \le 1$, we have

$$\mathbf{E}_J(|T_n|^i \,|\, \mathcal{F}_{\kappa_n}) \le \mathbf{P}_J(\kappa_{n+1} > \kappa_n + A \,|\, \mathcal{F}_{\kappa_n})$$

for $i = 1, 2$. The RHS is just the probability of at least $A/2$ successive moves from state 1 to state 2 or state -1 to state -2. This probability is at most the maximum of

$$\prod_{i \le A/2} \frac{Z(\kappa_n, 2) + i}{Z(\kappa_n, 2) + Z(\kappa_n, 0) + i}$$

and the same expression with 2 replaced by -2. Since $(\kappa_n^{-\epsilon})^{A/2} < \kappa_n^{-2}$ by choice of A, and since the terms in the product are at most a constant multiple of $\kappa_n^{-\epsilon}$ by the assumption that $\kappa_n < \tau_2 \wedge \tau_3$, the RHS is bounded by a constant multiple of κ_n^{-2}. Having bounded the conditional expectations of R_n^2, S_n^2 and T_n^2 by multiples of κ_n^{-2} and the magnitudes of the conditional expectations of R_n, S_n and T_n by constant multiples of $\kappa_n^{-1-\epsilon}$, we have established (4.2) and (4.3).

To establish (4.1), we will show that all of the three probabilities $\mathbf{P}_J(\tau_2 \le \tau_1 < \infty \,|\, \mathcal{F}_{N_0}))$, $\mathbf{P}_J(\tau_3 \le \tau_1 < \infty \,|\, \mathcal{F}_{N_0})$ and $\mathbf{P}_J(\tau_1 < \tau_2 \wedge \tau_3 \,|\, \mathcal{F}_{N_0})$ are simultaneously small when the values $Z(N_0, i) = z_i$ are chosen appropriately. As in step 4, we define the return times α_n to state 1 by $\alpha_{-1} = N_0 - 1$ and $\alpha_{n+1} = \min\{k > \alpha_n : X_k = 1\}$. Again set $U_n = Z(\alpha_n, 0)$ and $U_n' = Z(\alpha_n, 2)$. As in step 4, the process (U_n, U_n') evolves as the urns in Theorem 2.2, where again we let $W = U_{n+1} - U_n$. Assume that $\alpha_n < \tau_1 \wedge \tau_2$. A lower bound for the probability that $W > K$ is the probability that from state 1 the VRRW visits 0 and then visits states -1 and 0 K times in alternation. Thus

$$\begin{aligned} \mathbf{E}_J(W \,|\, \mathcal{F}_{\alpha_n}, W > 0) &\ge 1 + \sum_{i=2}^{K} \mathbf{P}_J(W > i - 1 \,|\, \mathcal{F}_{\alpha_n}, W > 0) \\ &\ge 1 + \sum_{i=2}^{K} \prod_{j=1}^{i-1} \frac{Z(\alpha_n, -1) + j - 1}{Z(\alpha_n, -1) + Z(\alpha_n, 1) + j - 1} \,. \end{aligned}$$

There is a $\phi(K)$ such that when $N_0 \geq \phi(K)$ and $\alpha_n < \tau_1 \wedge \tau_2$ then each factor in the product is at least $1 - d - 2\epsilon$. Thus for sufficiently large K and $\alpha_n < \tau_1 \wedge \tau_2$, the RHS is at least $1/(d+3\epsilon)$. It is trivial to see that $\mathbf{E}_J(W^2 \,|\, \mathcal{F}_n)$ is bounded. Thus setting $a = 1/(d+3\epsilon)$ and $\beta = 1/(d+4\epsilon)$, we apply Lemma 3.5 to see that $(U'_n)^\beta / U_n$ is a supermartingale when $\alpha_n < \tau_1 \wedge \tau_2$. More formally, let ρ be the least n for which $\alpha_n \geq \tau_1 \wedge \tau_2$. Setting $Y_n = U'_{n\wedge\rho}$ and $X_n = U_{n\wedge\rho}$, the process Y_n^β / X_n is a supermartingale. Since $d + 4\epsilon < 1 - \epsilon$, we see that by definition of τ_1 that $Y_n^\beta / X_n > 1$ if $\tau_2 \leq \tau_1 < \infty$, where n is the least j for which $\alpha_j \geq \tau_1$. Therefore, by the supermartingale optional stopping theorem we arrive at

$$\mathbf{P}_J(\tau_2 \leq \tau_1 < \infty \,|\, \mathcal{F}_{N_0}) \leq \frac{z_2^{1/(d+4\epsilon)}}{z_0} \,.$$

An entirely analogous argument with the states -2 and -1 in place of 2 and 1 and c in place of $1 - d$ yields the analogous bound

$$\mathbf{P}_J(\tau_3 \leq \tau_1 < \infty \,|\, \mathcal{F}_{N_0}) \leq \frac{z_{-2}^{1/(1-c-4\epsilon)}}{z_0} \,.$$

Finally, we need to see how to make $\mathbf{P}_J(\tau_1 < \tau_2 \wedge \tau_3 \,|\, \mathcal{F}_{N_0})$ small. Let ρ be the least n for which $\kappa_n \geq \tau_2 \wedge \tau_3$. Then by (4.2),

$$|\mathbf{E} W_\rho - W_0| \leq \sum_{m \geq \kappa_n} C \kappa_n^{-1-\epsilon} \leq C_1(\epsilon) N_0^{-\epsilon} \,.$$

Similarly, (4.3) gives

$$\mathrm{Var}(W_\rho - W_0) \leq \sum_{m \geq \kappa_n} C' \kappa_n^{-1} \leq C'_1 N_0^{-1} \,.$$

On the other hand, if $\tau_1 < \tau_2 \wedge \tau_3$ then

$$|W_\rho - W_0| \geq \min\{W_0 - c, d - W_0\} \,.$$

Chebyshev's inequality applied to $W_\rho - W_0$ then shows that

$$\mathbf{P}_j(\tau_1 < \tau_2 \wedge \tau_3) \leq \frac{C'_1 N_0^{-1}}{(\min\{W_0 - c, d - W_0\} - C_1 N_0^{-\epsilon})^2} \,.$$

When N_0 is sufficiently large, and $2z_1/N_0$ is sufficiently close to $(c+d)/2$, this is at most $C_2 N_0^{-1}$. Thus we have shown how to pick $z_{-2}, \ldots, z_2$ so that $\mathbf{P}(\tau < \infty \,|\, \mathcal{F}_{N_0})$ can be made arbitrarily small, which finishes the proof of (4.1) and of Theorem 1.3. $\square$

5 Remaining proofs and open questions

The proof of Lemma 4.1 is quite similar to the proof of Theorem 1.3. We give an outline for the argument, leaving out details that are the same as in the proof of Theorem 1.3.

SKETCH OF PROOF OF LEMMA 4.1: The first reduction is to analyze VRRW on $(-\infty, m+2]$. Let τ_m be the first time m is reached. The hardest part, because it requires a simultaneous induction on two different stopping times, similar to (4.1), is the following:

> Claim 1: There is a constant $\delta > 0$ such that for all m, the probability is at least δ that inequalities (i) and (ii) hold for every $n \geq \tau_m$:
>
> $$\begin{array}{ll} (i) & Z(n,m) \geq Z(n,m+2)^2\,; \\ (ii) & Z(n,m-1) \geq 2Z(n,m+1)\,. \end{array}$$

The other essential ingredient is

> Claim 2: $Z(n,m+1) \leq (1/4)Z(n,m+2)^2 \mathbf{1}_{A_n}$ finitely often almost surely, where A_M is the event that (i) and (ii) of the previous claim are true for all $n \in [\tau_m, M]$.

Assume these two claims and let σ_k be the time of the k^{th} visit to site $m+2$. From the first claim, the decreasing limit A_∞ has probability at least δ. Since $Z(\sigma_k, m+2) = k+1$, it follows from the second claim that on A_∞,

$$\sum_{k=1}^{\infty} Z(\sigma_k, m+1)^{-1} < \infty.$$

Fix $M, \epsilon > 0$ such that with probability at least ϵ, $\sum_{k=1}^{\infty} Z(\sigma_k, m+1)^{-1} < M$. As in Corollary 4.3 it then follows for VRRW on $\mathbf{Z}$ that $\mathbf{P}(m+3 \notin R \,|\, m \in R) > 0$ and in fact that a lower bound is $\delta := \epsilon \exp(-2M)$.

To prove the first claim, stop the walk the first time either condition (i) or (ii) is violated. Consider first the process $\{(U_n, V_n)\} := \{(Z(\rho_n, m-1), Z(\rho_n, m+1))\}$, where ρ_n are the successive hitting times of site m. At each step, precisely one of the coordinates is updated, with Polya-like probabilities, so by Proposition 2.1, the expected increment of $U_n/(U_n+V_n)$ is given by the contributions from increments of magnitude greater than 1:

$$\mathbf{E}\frac{U_{n+1}}{U_{n+1}+V_{n+1}} - \frac{U_n}{U_n+V_n} = \mathbf{E}\frac{(U_{n+1}-U_n-1)^+ - (V_{n+1}-V_n-1)^+}{U_{n+1}+V_{n+1}} .$$

The term $(U_{n+1}-U_n-1)^+$ is nonnegative and the term

$$\mathbf{E}\frac{-(V_{n+1}-V_n-1)^+}{U_{n+1}+V_{n+1}}$$

is of order

$$\frac{1}{U_n+V_n}\frac{Z(\rho_n, m+2)}{Z(\rho_n, m)} = O(n^{-1}n^{-1/2})$$

by condition (ii). This expresses $U_n/(U_n+V_n)$ as a martingale plus a drift term whose negative part is summable.

Consider next the process $\{(U_n', V_n')\} := \{(Z(\rho_n, m), Z(\rho_n, m+2))\}$, where ρ_n are now the successive hitting times of $m+1$. Again the updates are in a single coordinate chosen with Pólya-like probabilities, with the increment in V_n' being 1 and the increment in U_n' having conditional mean at least 3. Using Lemma 3.5, just as in Step 4 of the proof of Theorem 1.3, we see that $(V_n')^2/U_n'$ is a supermartingale.

The optional stopping theorem now shows that from an appropriate initial position, the probability of stopping due to a violation of (i) or (ii) is arbitrarily low. The initial position (or one at least as good) can be attained with a probability bounded away from zero (unless m is visited only finitely often, which is even better!) so the claim is proved.

Finally, to prove the second claim, let $X_n = Z(\rho_{n+1}, m+1) - Z(\rho_n, m+1)$ where now ρ_n are the successive hitting times of site $m+2$. On the event $A_{\rho_{n+1}}$, the probability of a transition to $m+2$ from $m+1$ between times ρ_n and ρ_{n+1} is bounded above by $1/(n+1)$ (use condition (i) and $Z(\cdot, m+2) = n$). Thus X_n stochastically dominates a geometric of mean n, and it is easy to verify that

$$\sum_{k=1}^{n} X_k < \frac{n^2}{4} \text{ finitely often, a.s.,}$$

which proves the second claim and hence the lemma. $\square$

Proof of Theorem 1.2: It is straightforward that $\mathbf{P}(|R'| = 2) = \mathbf{P}(|R'| = 3) = 0$, so we suppose that $R' = J := \{-2, -1, 0, 1\}$ and show that this leads to a contradiction. To simplify notation, let A_n, B_n, C_n and D_n denote $Z(n, x)$ for $x = -2, -1, 0$ and 1 respectively. The assumption $R' = J$ yields that there exist N such that $X_n \in J$ as soon as $n \geq N$. Throughout the rest of the proof, we assume that $n \geq N$.

Let κ_m be the time of the m^{th} return to -1 (clearly, κ_m's are $< \infty$ on the event $R' = J$). Between the times κ_m and κ_{m+1}, the random walk goes from -1 either to -2 and returns to -1, or to 0 and (possibly) bounces between 0 and 1 before going back to -1. Therefore, $W_m := A_{\kappa_m}/(A_{\kappa_m} + C_{\kappa_m})$ is a non-negative supermartingale which converges a.s. to some random variable W. Consider two cases:

- $W(\omega) = w > 0$;
- $W(\omega) = 0$.

In the first case, there exists $N_1 > N_0$ such that $W_m > w/2$ for $n > N_1$. As in Lemma 4.2, the probability never to jump from -2 to -3 is positive only whenever

$$\sum_n \frac{A_{n+1} - A_n}{B_n} < \infty$$

which is equivalent to the following sum being finite

$$\sum_{n \geq N_1} \frac{A_n}{B_n} \frac{B_n - B_{n-1}}{B_{n-1}} \geq \frac{w}{2} \sum \frac{B_n - B_{n-1}}{B_{n-1}} \tag{5.5}$$

(here we used the obvious inequality $B_n \leq A_n + C_n$). However, the sum in the RHS of (5.5) is a tail of the harmonic series and therefore diverges.

Before we proceed to the second case we observe that by the same arguments we can restrict the problem to the case when both $A_n/(A_n + C_n)$ and $D_n/(D_n + B_n)$ go to zero.

As a result, $A_n/C_n \to 0$ and $D_n/B_n \to 0$ as well. Taking into account that $B_n \le C_n + A_n$ and $C_n \le B_n + D_n$ we conclude that

$$\frac{B_n}{C_n} \to 1, \ \frac{2B_n}{n} \to 1, \ \frac{2C_n}{n} \to 1. \tag{5.6}$$

Let τ_m be the time of the m^{th} visit to -1 or 0 skipping at least one step,

$$\tau_m := \inf\{n > \tau_{m-1} + 1 : \ X_n \in \{-1, 0\}\}, \quad \tau_0 := N_0.$$

Define

$$U_m = \frac{B_{\tau_m} + C_{\tau_m}}{2}, \quad V_m = A_{\tau_m} + D_{\tau_m}.$$

If $X_{\tau_m} = -1$ ($X_{\tau_m} = 0$ resp.), then between the times τ_m and τ_{m+1} VRRW will either (1) go to the left (right resp.) and back, or (2) go to the right (left resp.) and back, or (3) go twice to the right (left resp.) and make one step back. Consequently, either $U_{m+1} = U_m + 1$ and $V_{m+1} = V_m$ (when (2) takes place) or $U_{m+1} \le U_m + 1$ and $V_{m+1} = V_m + 1$ (when (1) or (3) takes place). We claim that the probability of the latter event, denoted by F, is greater than $V_m/(U_m + V_m)$ when m is large enough. This, in turn, will imply that the process (U, V) can be coupled with some general urn model process described by (2.3) such that $U_m \le X'_m$ and $V_m \ge Y'_m$.

To prove that $\mathbf{P}(F) \ge V_m/(U_m + V_m)$ we consider the quantity $A_n - B_n + C_n - D_n$ which is “almost” invariant for $n \ge N$. Namely, there exists a (possibly negative) constant K, depending on the history of VRRW before time N only, such that $A_n - B_n + C_n - D_n = K$ whenever $X_n = -1$ and $A_n - B_n + C_n - D_n = K + 1$ whenever $X_n = 0$. If we denote $t_m := B_{\tau_m} + D_{\tau_m}$, then $A_{\tau_m} + C_{\tau_m}$ equals $t_m + K$ or $t_m + K + 1$ when VRRW is at -1 or at 0 respectively. In the second case

$$\mathbf{P}(F) = \frac{D}{t} + \left(1 - \frac{D}{t}\right) \frac{A}{t + K + 1} = \frac{V - AD/(t + K + 1)}{t} - \frac{A(K + 1)}{t(t + K + 1)}$$

(we omit the indices for simplicty). Taking into account that $2U + V = 2t + K + 1$, $V > A$ and $AD \le V^2/4$, we obtain

$$\mathbf{P}(F) - \frac{V}{U + V} \quad \ge \quad \frac{V - V^2/(4t + 4K + 4)}{t} - \frac{A(K + 1)}{t(t + K + 1)} - \frac{V}{t + (V + K + 1)/2}$$

$$\geq \quad \frac{V(V+K-1)}{t(2t+V+K+1)} - \frac{V^2}{t(4t+4K+4)} - \frac{V(|K|+1)}{t(t+K+1)}$$

As $m \to \infty$ (and, therefore, $n \to \infty$) we have $V_m \to \infty$ and $V_m = o(t_m)$, whence

$$\mathbf{P}(F) - \frac{V}{U+V} = \frac{V^2}{4t^2}\left(1 - \Theta\left(\frac{V}{t}\right) - \Theta\left(\frac{1}{V}\right)\right) = \frac{V^2}{4t^2}(1 - o(1))$$

is non-negative for $m \geq M$ where M is some constant. The case $X_{\tau_m} = -1$ can be analyzed in the similar way.

We have shown that for large m the process (U_m, V_m) can be coupled with the process (X'_m, Y'_m) obeying the law (2.3) with $a = d = 1$, $b = 0$ and $c = 1$ such that $U_m \leq X'_m$ and $V_m \geq Y'_m$. By Theorem 2.3, there exists

$$\lim_{m\to\infty} \frac{X'_m}{Y'_m} - \log(Y'_m) \in (-\infty, \infty)$$

which, in turn, implies the existence of a random variable $\zeta \in (0, \infty)$ such that

$$X'_m \leq Y'_m \log(2\zeta Y'_m) \quad \text{for all } m \geq M.$$

Clearly, $X'_m / Y'_m \to \infty$, so

$$Y'_m \geq \frac{X'_m}{\log(2\zeta X'_m)}$$

for all m larger than some $M_1 \geq M$. Since $U_m \leq X'_m$, $V_m \geq Y'_m$ and the function $f(x) = x/\log(2\zeta x)$ is increasing for large x, $V_m \geq U_m/\log(2\zeta U'_m)$. Furthermore, $\tau_m \to \infty$ and $\tau_{m+1} \leq \tau_m + 3$, so we asymptotically have

$$A_n + D_n \geq \frac{(B_n + C_n)/2}{\log(\zeta(B_n + C_n))} \simeq \frac{n}{2\log(n)}$$

by (5.6).

The event that VRRW does not jump off J can occur only when the sum

$$\sum \frac{A_{n+1} - A_n}{B_n} + \frac{D_{n+1} - D_n}{C_n}$$

is finite. Summing by parts as in Corollary 4.3, we obtain that this is equivalent to the finiteness of the sum

$$\sum_n \frac{A_n + D_n}{n^2} \geq const + \sum_n \frac{(B_n + C_n)/2}{\log(\zeta(B_n + C_n))} \frac{1}{n^2} \simeq const + \sum_n \frac{1}{2n \log(\zeta n)}$$

which diverges. Therefore, $\mathbf{P}(R' = J) = 0$, completing the proof. $\square$

We end with some questions. The strongest conjecture about VRRW on $\mathbf{Z}$ is the one stated after Theorem 1.3, to the effect that the behavior described in Theorem 1.3 happens with probability 1. Some smaller steps toward this would be to prove that the set of sites visited with positive density must be connected and to prove that α can never be 0 or 1. This would, for example, rule out that sites -1 and 0 are visited with density $1/2$, and by time n the numbers of visits to the sites $1, 2, \ldots$ are asymptotically $n/\log n, n/(\log n \log\log n), \ldots$. Another graph on which VRRW may have interesting behavior is $\mathbf{Z}^2$. Ferrari and Meilijson (personal communication, 1996) also have some results about VRRW on a tree.

A further question is that of stochastically comparing VRRW with different histories. For example, we originally thought we could prove a version of Lemma 4.1 in which it was shown that $\mathbf{P}(m+3 \in R \,|\, m \in R) \leq \mathbf{P}(3 \in R)$, by showing that the extra weight to the left of m the first time m is reached can only help the range stay bounded above by $m+2$. We were unable to do this, by coupling or martingale arguments, but believe that some such comparison must hold. The easiest to state are false.

References

[1] Arthur, B. (1986). Industry location patterns and the importance of history. Center for Economic Policy Research Paper No. 84: Stanford.

[2] Arthur, B., Ermoliev, Y. and Kaniovski, Y. (1987). Path-dependent processes and the emergence of macro-structure. *Eur. J. Oper. Res.* **30** 294 - 303.

[3] Athreya, K. (1967). Doctoral dissertation, Stanford University.

[4] Athreya, K., and Ney, P. (1972). Branching Processes. Springer: New York.

[5] Blackwell, D. and McQueen, J. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Math. Stat.* **1** 353 - 355.

[6] Coppersmith, D. and Diaconis, P. (1987). Random walks with reinforcement. Unpublished manuscript.

[7] Davis, B. (1990). Reinforced random walk. *Prob. Th. Rel. Fields* **84** 203 - 229.

[8] Eggenberger, F. and Pólya, G. (1923). Über die Statistik Verketter Vorgänge. *Zeit. Angew. Math. Mech.* **3** 279 - 289.

[9] Feller, W. (1971). Introduction to probability theory and its applications, 2nd edition. John Wiley and Sons: New York.

[10] Freedman, D. (1965). Bernard Friedman's urn. *Ann. Math. Stat.* **36** 956 - 970.

[11] Friedman, B. (1949). A simple urn model. *Comm. Pure Appl. Math.* **2** 59 - 70.

[12] Greenwood, M. and Yule, G. U. (1920). Inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J. Royal Stat. Soc.* **83** 255 - 279.

[13] Iosifescu, M. and Theodorescu, R. (1969). Random prcesses and learning. Springer: New York.

[14] Kushner, H. and Yin, G. (1997). Stochastic approximation algorithms and applications. Applications of Mathematics, 35. Springer-Verlag: New York.

[15] Mauldin, D., Sudderth, B. and Williams, S. (1992). Pólya trees and random distributions. *Ann. Math. Stat.* **20** 1203 - 1221.

[16] Othmer, H. and Stevens, A. (1998). Aggregation, blowup, and collapse: the ABC's of taxis in reinforced random walk. *SIAM J. Appl. Math.* **57** 1044 - 1081.

[17] Pemantle, R. (1988a). Random processes with reinforcement. Massachsetts Institute of Technology doctoral dissertation.

[18] Pemantle, R. (1988b). Phase transition in reinforced random walk and RWRE on trees. *Ann. Probab.* **16** 1229 - 1241.

[19] Pemantle, R. (1992). Vertex Reinforced random walk. *Prob. Th. and Rel. Fields* **92**, 117 - 136.

[20] Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400 - 407.

[21] Sellke, T. (1994). Reinforced random walk on the d-dimensional integer lattice. *Preprint.*

Robin Pemantle
Department of Mathematics, University of Wisconsin-Madison
Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706

Stanislav Volkov
The Fields Institute for Research in Mathematical Sciences
222 College Street, Toronto, Ontario, Canada, M5T3J1