The simple harmonic urn

Edward Crane* Nicholas Georgiou* Stanislav Volkov†,‡ Andrew R. Wade§
Robert J. Waters*

9 August 2010

Abstract

We study a generalized Pólya urn model with two types of ball. If the drawn ball is red it is replaced together with a black ball, but if the drawn ball is black it is replaced and a red ball is thrown out of the urn. When only black balls remain, the rôles of the colours are swapped and the process restarts. We prove that the resulting Markov chain is transient but that if we throw out a ball every time the colours swap, the process is recurrent. We show that the embedded process obtained by observing the number of balls in the urn at the swapping times has a scaling limit that is essentially the square of a Bessel diffusion. We consider an oriented percolation model naturally associated with the urn process, and obtain detailed information about its structure, showing that the open subgraph is an infinite tree with a single end. We also study a natural continuous-time embedding of the urn process that demonstrates the relation to the simple harmonic oscillator; in this setting our transience result addresses an open problem in the recurrence theory of two-dimensional linear birth and death processes due to Kesten and Hutton. We obtain results on the area swept out by the process. We make use of connections between the urn process and birth-death processes, a uniform renewal process, the Eulerian numbers, and Lamperti's problem on processes with asymptotically small drifts; we prove some new results on some of these classical objects that may be of independent interest. For instance, we give sharp new asymptotics for the first two moments of the counting function of the uniform renewal process. Finally we discuss some related models of independent interest, including a 'Poisson earthquakes' Markov chain on the homeomorphisms of the plane.

Keywords: Urn model, recurrence classification, oriented percolation, uniform renewal process, two-dimensional linear birth and death process, Bessel process, coupling, Eulerian numbers.
AMS subject classification: Primary: 60J10; Secondary: 60J25, 60K05, 60K35.

1 Introduction

Urn models have a venerable history in probability theory, with classical contributions having been made by the Bernoullis and Laplace, among others. The modern view of many urn models is as prototypical reinforced stochastic processes. Classical urn schemes were often employed as 'thought experiments' in which to frame statistical questions; as stochastic processes, urn models have wide-ranging applications in economics, the physical sciences, and statistics. There is a large literature on urn models and their applications — see for example the monographs [20, 30]

*Heilbronn Institute for Mathematical Research, Department of Mathematics, University of Bristol, Bristol BS8 1TW, UK
†Department of Mathematics, University of Bristol, Bristol BS8 1TW, UK
‡Corresponding author
§Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

and the surveys [25, 34] — and some important contributions have been made in the last few years: see e.g. [18, 13].

A generalized Pólya urn with 2 types of ball, or 2 colours, is a discrete-time Markov chain $(X_n, Y_n)_{n\in\mathbb{Z}_+}$ on $\mathbb{Z}_+^2$, where $\mathbb{Z}_+ := \{0, 1, 2, \ldots\}$. The possible transitions of the chain are specified by a 2×2 *reinforcement matrix* $A = (a_{ij})_{i,j=1}^2$ and the transition probabilities depend on the current state:

$$\begin{aligned}\mathbb{P}\left((X_{n+1}, Y_{n+1}) = (X_n + a_{11}, Y_n + a_{12})\right) &= \frac{X_n}{X_n + Y_n}, \\ \mathbb{P}\left((X_{n+1}, Y_{n+1}) = (X_n + a_{21}, Y_n + a_{22})\right) &= \frac{Y_n}{X_n + Y_n}.\end{aligned} \tag{1}$$

This process can be viewed as an urn which at time n contains X_n red balls and Y_n black balls. At each stage, a ball is drawn from the urn at random, and then returned together with a_{i1} red balls and a_{i2} black balls, where $i = 1$ if the chosen ball is red and $i = 2$ if it is black.

A fundamental problem is to study the long-term behaviour of (X_n, Y_n), defined by (1), or some function thereof, such as the fraction of red balls $X_n/(X_n + Y_n)$. In many cases, coarse asymptotics for such quantities are governed by the eigenvalues of the reinforcement matrix A (see e.g. [4] or [5, §V.9]). However there are some interesting special cases (see e.g. [35]), and analysis of finer behaviour is in several cases still an open problem.

A large body of asymptotic theory is known under various conditions on A and its eigenvalues. Often it is assumed that all $a_{ij} \geq 0$; e.g., $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ specifies the *standard Pólya urn*, while $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ with $a, b > 0$ specifies a *Friedman urn*.

In general, the entries a_{ij} may be negative, meaning that balls can be thrown away as well as added, but nevertheless in the literature *tenability* is usually imposed. This is the condition that regardless of the stochastic path taken by the process, it is never required to remove a ball of a colour not currently present in the urn. For example the *Ehrenfest urn*, which models the diffusion of a gas between two chambers of a box, is tenable despite its reinforcement matrix $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ having some negative entries.

Departing from tenability, the *OK Corral model* is the 2-colour urn with reinforcement matrix $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$. This model for destructive competition was studied by Williams and McIlroy [43] and Kingman [23] (and earlier as a stochastic version of Lanchester's combat model; see e.g. [42] and references therein). Kingman and Volkov [24] showed that the OK Corral model can be viewed as a time-reversed Friedman urn with $a = 0$ and $b = 1$.

In this paper, we will study the 2-colour urn model with reinforcement matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{2}$$

To reiterate the urn model, at each time period we draw a ball at random from the urn; if it is red, we replace it and add an additional black ball, if it is black we replace it and throw out a red ball. The eigenvalues of A are $\pm i$, corresponding to the ordinary differential equation $\dot{\mathbf{v}} = A\mathbf{v}$, which governs the phase diagram of the simple harmonic oscillator. This explains the name *simple harmonic urn*. Naïvely, one might hope that the behaviour of the Markov chain is closely related to the paths in the phase diagram. We will see that it is, but that the exact behaviour is somewhat more subtle.

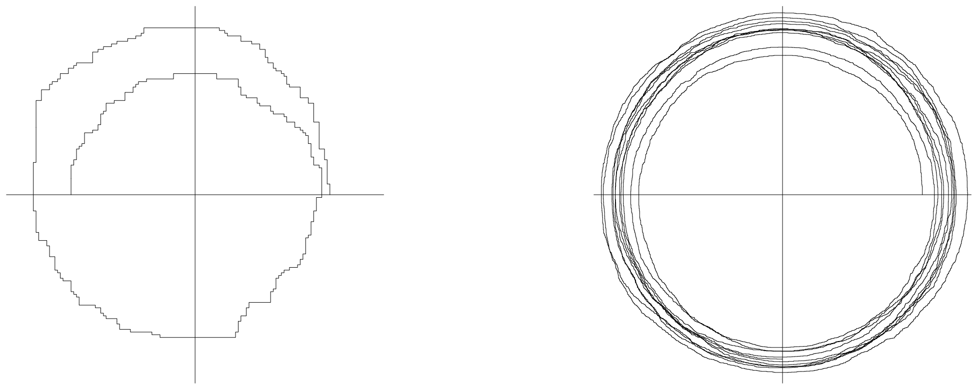

Figure 1: Two sample trajectories of the simple harmonic urn process, starting at $(50, 0)$ and running for about 600 steps (*left*) and starting at $(1000, 0)$ and running for $100,000$ steps (*right*).

2 Exact formulation of the model and main results

2.1 The simple harmonic urn process

The definition of the process given by the transition probabilities (1) and the matrix (2) only makes sense for $X_n, Y_n \geq 0$; however, it is easy to see that almost surely (a.s.) $X_n < 0$ eventually. Therefore, we reformulate the process (X_n, Y_n) rigorously as follows.

For $z_0 \in \mathbb{N} := \{1, 2, \ldots\}$ take $(X_0, Y_0) = (z_0, 0)$; we start on the positive x-axis for convenience but the choice of initial state does not affect any of our asymptotic results. For $n \in \mathbb{Z}_+$, given $(X_n, Y_n) = (x, y) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, we define the transition law of the process by

$$(X_{n+1},\, Y_{n+1}) \quad = \quad \begin{cases} (x, y + \operatorname{sgn}(x)), & \text{with probability } \frac{|x|}{|x|+|y|}, \\ \\ (x - \operatorname{sgn}(y), y), & \text{with probability } \frac{|y|}{|x|+|y|}, \end{cases} \tag{3}$$

where $\operatorname{sgn}(x) = -1, 0, 1$ if $x < 0$, $x = 0$, $x > 0$ respectively. The process $(X_n, Y_n)_{n \in \mathbb{Z}_+}$ is an irreducible Markov chain with state-space $\mathbb{Z}^2 \setminus \{(0, 0)\}$.

Let $\nu_0 := 0$, and recursively define stopping times

$$\nu_k := \min\{n > \nu_{k-1} : \ X_n Y_n = 0\}, \quad (k \in \mathbb{N}),$$

where throughout the paper we adopt the usual convention $\min \emptyset := \infty$. Thus $(\nu_k)_{k \in \mathbb{N}}$ is the sequence of times at which the process visits one of the axes.

It is easy to see that every ν_k is almost surely finite. Moreover, by construction, the process $(X_{\nu_k}, Y_{\nu_k})_{k \in \mathbb{N}}$ visits in cyclic (anticlockwise) order the half-lines $\{y > 0\}$, $\{x < 0\}$, $\{y < 0\}$, $\{x > 0\}$. It is natural (and fruitful) to consider the embedded process $(Z_k)_{k \in \mathbb{Z}_+}$ obtained by taking $Z_0 := z_0$ and $Z_k := |X_{\nu_k}| + |Y_{\nu_k}|$ $(k \in \mathbb{N})$.

If (X_n, Y_n) is viewed as a random walk on $\mathbb{Z}^2$, the process Z_k is the embedded process of the distances from 0 at the instances of hitting the axes. To interpret the process (X_n, Y_n) as the urn model described in Section 1 we need a slight modification to the description there. Starting with z_0 red balls, we run the process as described in Section 1, so the process traverses the first quadrant via an up/left path until the red balls run out (i.e. we first hit the half-line $\{y > 0\}$). Now we interchange the rôles of the red and black balls, and we still use y to count the black balls, but we switch to using $-x$ to count the number of red balls. Now the process traverses the second quadrant via a left/down path until the black balls run out, and so on. In the urn model, Z_k is the number of balls remaining in the urn when the urn becomes monochromatic for the k-th time ($k \in \mathbb{N}$).

The strong Markov property and the transition law of (X_n, Y_n) imply that Z_k is an irreducible Markov chain on $\mathbb{N}$. Since our two Markov chains just described are irreducible, there is the usual recurrence/transience dichotomy, in that either the process is recurrent, meaning that with probability 1 it returns infinitely often to any finite subset of the state space, or it is transient, meaning that with probability 1 it eventually escapes to infinity. Our main question is whether the process Z_k is recurrent or transient. It is easy to see that, by the nature of the embedding, this also determines whether the urn model (X_n, Y_n) is recurrent or transient.

Theorem 2.1. *The process Z_k is transient; hence so is the process (X_n, Y_n).*

Exploiting a connection between the increments of the process Z_k and a renewal process whose inter-arrival times are uniform on $(0, 1)$ will enable us to prove the following basic result.

Theorem 2.2. *Let $n \in \mathbb{N}$. Then*

$$\mathbb{E}\left[Z_{k+1} \mid Z_k = n\right] = n + \frac{2}{3} + O\left(\mathrm{e}^{\alpha_1 n}\right), \tag{4}$$

as $n \to \infty$, where $\alpha_1 + \beta_1 i = -(2.088843\ldots) + (7.461489\ldots)i$ is a root of $\lambda - 1 + \mathrm{e}^{-\lambda} = 0$.

The error term in (4) is sharp, and we obtain it from new (sharp) asymptotics for the uniform renewal process: see Lemma 4.5 and Corollary 6.5, which improve on known results. To prove Theorem 2.1 we need more than Theorem 2.2: we need to know about the second moments of the increments of Z_k, amongst other things; see Section 6. In fact we prove Theorem 2.1 using martingale arguments applied to $h(Z_k)$ for a well-chosen function h; the analysis of the function $h(Z_k)$ rests on a recurrence relation satisfied by the transition probabilities of Z_k, which are related to the Eulerian numbers (see Section 3).

2.2 The leaky simple harmonic urn

In fact the transience demonstrated in Theorem 2.1 is rather delicate, as one can see by simulating the process. To illustrate this, we consider a slight modification of the process, which we call the *leaky* simple harmonic urn. Suppose that each time the rôles of the colours are reversed, the addition of the next ball of the new colour causes one ball of the other colour to leak out of the urn; subsequently the usual simple harmonic urn transition law applies. If the total number of balls in the urn ever falls to one, then this modified rule causes the urn to become monochromatic at the next step, and again it contains only one ball. Thus there will only be one ball in total at all subsequent times, although it will alternate in colour. We will see that the system almost surely does reach this steady state, and we obtain almost sharp tail bounds on the time that it takes to do so. The leaky simple harmonic urn arises naturally in the context of a percolation model associated to the simple harmonic urn process, defined in Section 2.4 below.

As we did for the simple harmonic urn, we will represent the leaky urn by a Markov chain (X'_n, Y'_n). For this version of the model, it turns out to be more convenient to start just above the axis; we take $(X'_0, Y'_0) = (z_0, 1)$, where $z_0 \in \mathbb{N}$. The distribution of (X'_{n+1}, Y'_{n+1}) depends only on $(X'_n, Y'_n) = (x, y)$.

If $xy \neq 0$, the transition law is the same as that of the simple harmonic urn process. The difference is when $x = 0$ or $y = 0$; then the transition law is

$$(X'_{n+1}, Y'_{n+1}) = (-\operatorname{sgn}(y), y - \operatorname{sgn}(y)) \quad (x = 0)$$

$$(X'_{n+1}, Y'_{n+1}) = (x - \operatorname{sgn}(x), \operatorname{sgn}(x)) \quad (y = 0).$$

Now (X'_n, Y'_n) is a reducible Markov chain whose state-space has two communicating classes, the closed class $\mathcal{C} = \{(x, y) \in \mathbb{Z}^2 : |x| + |y| = 1\}$ and the class $\{(x, y) \in \mathbb{Z}^2 : |x| + |y| \geq 2\}$; if the process enters the closed class $\mathcal{C}$ it remains there for ever, cycling round the origin. Let τ be the hitting time of the set $\mathcal{C}$, that is

$$\tau := \inf\{n \in \mathbb{Z} : |X'_n| + |Y'_n| = 1\}.$$

Theorem 2.3. *For the leaky urn, $\mathbb{P}(\tau < \infty) = 1$. Moreover, for any $\varepsilon > 0$, $\mathbb{E}[\tau^{1-\varepsilon}] < \infty$ but $\mathbb{E}[\tau^{1+\varepsilon}] = \infty$.*

In contrast, Theorem 2.1 implies that the analogue of τ for the ordinary urn process has $\mathbb{P}(\tau = \infty) > 0$ if $z_0 \geq 2$.

2.3 The noisy simple harmonic urn

In view of Theorems 2.1 and 2.3, it is natural to ask about the properties of the hitting time τ if at the time when the balls of one colour run out we only discard a ball of the other colour with some probability $p \in (0, 1)$. For which p is τ a.s. finite? (Answer: for $p \geq 1/3$; see Corollary 2.7 below.)

We consider the following natural generalization of the model specified by (3) in order to probe more precisely the recurrence/transience transition. We call this generalization the *noisy* simple harmonic urn process. In a sense that we will describe, this model includes the leaky urn and also the intermittent leaky urn mentioned at the start of this section. The basic idea is to throw out (or add) a random number of balls at each time we are at an axis, generalizing the idea of the leaky urn. It is more convenient here to work with irreducible Markov chains, so we introduce a 'barrier' for our process. We now describe the model precisely.

Let $\kappa, \kappa_1, \kappa_2, \ldots$ be a sequence of i.i.d. $\mathbb{Z}$-valued random variables such that

$$\mathbb{E}[e^{\lambda|\kappa|}] < \infty, \tag{5}$$

for some $\lambda > 0$, so in particular $\mathbb{E}[\kappa]$ is finite. We now define the Markov chain $(\tilde{X}_n, \tilde{Y}_n)_{n \in \mathbb{Z}_+}$ for the noisy urn process. As for the leaky urn, we start one step above the axis: let $z_0 \in \mathbb{N}$, and take $(\tilde{X}_0, \tilde{Y}_0) = (z_0, 1)$. For $n \in \mathbb{Z}_+$, given $(\tilde{X}_n, \tilde{Y}_n) = (x, y) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, we define the transition law as follows. If $xy \neq 0$ then

$$(\tilde{X}_{n+1}, \tilde{Y}_{n+1}) = \begin{cases} (x, y + \operatorname{sgn}(x)), & \text{with probability } \frac{|x|}{|x|+|y|}, \\ \\ (x - \operatorname{sgn}(y), y), & \text{with probability } \frac{|y|}{|x|+|y|}, \end{cases}$$

while if $x = 0$ or $y = 0$ we have

$$\begin{aligned} (\tilde{X}_{n+1}, \tilde{Y}_{n+1}) &= (-\operatorname{sgn}(y), \operatorname{sgn}(y) \max(1, |y| - \kappa_n)) \quad (x = 0), \\ (\tilde{X}_{n+1}, \tilde{Y}_{n+1}) &= (\operatorname{sgn}(x) \max(1, |x| - \kappa_n), \operatorname{sgn}(x)) \quad (y = 0). \end{aligned}$$

In other words, the transition law is the same as (3) except when the process is on an axis at time n, in which case instead of just moving one step away in the anticlockwise perpendicular direction it also moves an additional distance κ_n parallel to the axis towards the origin (stopping distance 1 away if it would otherwise reach the next axis or overshoot). Then $(\tilde{X}_n, \tilde{Y}_n)_{n \in \mathbb{Z}_+}$ is an irreducible Markov chain on $\mathbb{Z}^2 \setminus \{(0, 0)\}$. The case where $\mathbb{P}(\kappa = 0) = 1$ corresponds to the original process (X_n, Y_n) starting one unit later in time.

A fundamental random variable is the first passage time to within distance 1 of the origin:

$$\tau := \min\{n \in \mathbb{Z}_+ : |\tilde{X}_n| + |\tilde{Y}_n| = 1\} = \min\{n \in \mathbb{Z}_+ : (\tilde{X}_n, \tilde{Y}_n) \in \mathcal{C}\}.$$

Define a sequence of stopping times $\tilde{\nu}_k$ by setting $\tilde{\nu}_0 := -1$ and for $k \in \mathbb{N}$,

$$\tilde{\nu}_k := \min\{n > \tilde{\nu}_{k-1} : \tilde{X}_n \tilde{Y}_n = 0\}.$$

As an analogue of Z_k, set $\tilde{Z}_0 := z_0$ and for $k \in \mathbb{N}$ define

$$\tilde{Z}_k := \max\{|\tilde{X}_{1+\tilde{\nu}_k}|, |\tilde{Y}_{1+\tilde{\nu}_k}|\} = |\tilde{X}_{1+\tilde{\nu}_k}| + |\tilde{Y}_{1+\tilde{\nu}_k}| - 1;$$

then $(\tilde{Z}_k)_{k\in\mathbb{Z}_+}$ is an irreducible Markov chain on $\mathbb{N}$. Define the return-time to the state 1 by

$$\tau_q := \min\{k \in \mathbb{N} : \tilde{Z}_k = 1\}, \tag{6}$$

where the subscript q signifies the fact that a time unit is one traversal of a quadrant here. By our embedding, $\tau = \tilde{\nu}_{\tau_q}$.

Note that in the case $\mathbb{P}(\kappa = 0) = 1$, $(\tilde{Z}_k)_{k\in\mathbb{Z}_+}$ has the same distribution as the original $(Z_k)_{k\in\mathbb{Z}_+}$. The noisy urn with $\mathbb{P}(\kappa = 1) = 1$ coincides with the leaky urn described in Section 2.2 up until the time τ (at which point the leaky urn becomes trapped in $\mathcal{C}$). Similarly the embedded process $\tilde{Z}_k$ with $\mathbb{P}(\kappa = 1) = 1$ coincides with the process of distances from the origin of the leaky urn at the times that it visits the axes, up until time τ_q (at which point the leaky urn remains at distance 1 forever). Thus in the $\mathbb{P}(\kappa = 1) = 1$ cases of all the results that follow in this section, τ and τ_q can be taken to be defined in terms of the leaky urn (X'_n, Y'_n).

The next result thus includes Theorem 2.1 and the first part of Theorem 2.3 as special cases.

Theorem 2.4. *Suppose that κ satisfies (5). Then the process $\tilde{Z}_k$ is*

(i) transient if $\mathbb{E}[\kappa] < 1/3$;

(ii) null-recurrent if $1/3 \leq \mathbb{E}[\kappa] \leq 2/3$;

(iii) positive-recurrent if $\mathbb{E}[\kappa] > 2/3$.

Of course, part (i) means that $\mathbb{P}(\tau_q < \infty) < 1$, part (ii) that $\mathbb{P}(\tau_q < \infty) = 1$ but $\mathbb{E}[\tau_q] = \infty$, and part (iii) that $\mathbb{E}[\tau_q] < \infty$. We can in fact obtain more information about the tails of τ_q:

Theorem 2.5. *Suppose that κ satisfies (5) and $\mathbb{E}[\kappa] \geq 1/3$. Then $\mathbb{E}[\tau_q^p] < \infty$ for $p < 3\mathbb{E}[\kappa] - 1$ and $\mathbb{E}[\tau_q^p] = \infty$ for $p > 3\mathbb{E}[\kappa] - 1$.*

It should be possible, with some extra work, to show that $\mathbb{E}[\tau_q^p] = \infty$ when $p = 3\mathbb{E}[\kappa] - 1$, using the sharper results of [2] in place of those from [3] that we use below in the proof of Theorem 2.5.

In the recurrent case, it is of interest to obtain more detailed results on the tail of τ (note that there is a change of time between τ and τ_q). We obtain the following upper and lower bounds, which are close to sharp.

Theorem 2.6. *Suppose that κ satisfies (5) and $\mathbb{E}[\kappa] \geq 1/3$. Then $\mathbb{E}[\tau^p] < \infty$ for $p < \frac{3\mathbb{E}[\kappa]-1}{2}$ and $\mathbb{E}[\tau^p] = \infty$ for $p > \frac{3\mathbb{E}[\kappa]-1}{2}$.*

Theorems 2.4 and 2.6 have an immediate corollary for the noisy urn process $(\tilde{X}_n, \tilde{Y}_n)$.

Corollary 2.7. *Suppose that κ satisfies (5). The noisy simple harmonic urn process $(\tilde{X}_n, \tilde{Y}_n)$ is recurrent if $\mathbb{E}[\kappa] \geq 1/3$ and transient if $\mathbb{E}[\kappa] < 1/3$. Moreover, the process is null-recurrent if $1/3 \leq \mathbb{E}[\kappa] < 1$ and positive-recurrent if $\mathbb{E}[\kappa] > 1$.*

This result is close to sharp but leaves open the question of whether the process is null- or positive-recurrent when $\mathbb{E}[\kappa] = 1$ (we suspect the former).

We also study the distributional limiting behaviour of $\tilde{Z}_k$ in the appropriate scaling regime when $\mathbb{E}[\kappa] < 2/3$. Again the case $\mathbb{P}(\kappa = 0) = 1$ reduces to the original Z_k.

Theorem 2.8. *Suppose that κ satisfies (5) and that $\mathbb{E}[\kappa] < 2/3$. Let $(D_t)_{t\in[0,1]}$ be a diffusion process taking values in $\mathbb{R}_+ := [0, \infty)$ with $D_0 = 0$ and infinitesimal mean $\mu(x)$ and variance $\sigma^2(x)$ given for $x \in \mathbb{R}_+$ by*

$$\mu(x) = \frac{2}{3} - \mathbb{E}[\kappa], \quad \sigma^2(x) = \frac{2}{3}x.$$

Then as $k \to \infty$,

$$(k^{-1}\tilde{Z}_{kt})_{t\in[0,1]} \to (D_t)_{t\in[0,1]},$$

where the convergence is in the sense of finite-dimensional distributions. Up to multiplication by a scalar, D_t is the square of a Bessel process with parameter $4 - 6\mathbb{E}[\kappa] > 0$.

Since a Bessel process with parameter $\gamma \in \mathbb{N}$ has the same law as the norm of a γ-dimensional Brownian motion, Theorem 2.8 says, for example, that if $\mathbb{E}[\kappa] = 0$ (e.g., for the original urn process) the scaling limit of $\tilde{Z}_t$ is a scalar multiple of the norm-square of 4-dimensional Brownian motion, while if $\mathbb{E}[\kappa] = 1/2$ the scaling limit is a scalar multiple of the square of one-dimensional Brownian motion.

To finish this section, consider the *area* swept out by the path of the noisy simple harmonic urn on its first excursion (i.e., up to time τ). Additional motivation for studying this random quantity is provided by the percolation model of Section 2.4. Formally, for $n \in \mathbb{N}$ let T_n be the area of the triangle with vertices $(0, 0)$, $(\tilde{X}_{n-1}, \tilde{Y}_{n-1})$, and $(\tilde{X}_n, \tilde{Y}_n)$, and define $A := \sum_{n=1}^{\tau} T_n$.

Theorem 2.9. *Suppose that κ satisfies (5).*

(i) Suppose that $\mathbb{E}[\kappa] < 1/3$. Then $\mathbb{P}(A = \infty) > 0$.

(ii) Suppose that $\mathbb{E}[\kappa] \geq 1/3$. Then $\mathbb{E}[A^p] < \infty$ for $p < \frac{3\mathbb{E}[\kappa]-1}{3}$.

In particular, part (ii) gives us information about the leaky urn model, which corresponds to the case where $\mathbb{P}(\kappa = 1) = 1$, at least up until the hitting time of the closed cycle; we can still make sense of the area swept out by the leaky urn up to this hitting time. We then have $\mathbb{E}[A^p] < \infty$ for $p < 2/3$, a result of significance for the percolation model of the next section. We suspect that the bounds in Theorem 2.9(ii) are tight. We do not prove this but have the following result in the case $\mathbb{P}(\kappa = 1) = 1$.

Theorem 2.10. *Suppose $\mathbb{P}(\kappa = 1) = 1$ (or equivalently take the leaky urn). Then $\mathbb{E}[A] = \infty$.*

2.4 A percolation model

Associated to the simple harmonic urn is a percolation model which we describe in this section. The percolation model, as well as being of interest in its own right, couples many different instances of the simple harmonic urn, and exhibits naturally an instance of the leaky version of the urn in terms of the planar dual percolation model. Our results on the simple harmonic urn will enable us to establish some interesting properties of the percolation model.

The simple harmonic urn can be viewed as a spatially inhomogeneous random walk on a directed graph whose vertices are $\mathbb{Z}^2 \setminus \{(0,0)\}$; we make this statement more precise shortly. In this section we will view the simple harmonic urn process not as a random path through a predetermined directed graph but as a deterministic path through a random directed graph. To do this it is helpful to consider a slightly larger state-space which keeps track of the number of times that the urn's path has wound around the origin. We construct this state-space as the vertex set of a graph G that is embedded in the Riemann surface $\mathcal{R}$ of the complex logarithm, which is the universal cover of $\mathbb{R}^2 \setminus \{(0,0)\}$. To construct G, we take the usual square-grid lattice and delete the vertex at the origin to obtain a graph on the vertex set $\mathbb{Z}^2 \setminus \{(0,0)\}$. Make this into a directed graph by orienting each edge in the direction of increasing argument; the paths of the simple harmonic urn only ever traverse edges in this direction. Leave undirected those edges along any of the co-ordinate axes; the paths of the simple harmonic urn never traverse these edges. Finally, we let G be the lift of this graph to the covering surface $\mathcal{R}$.

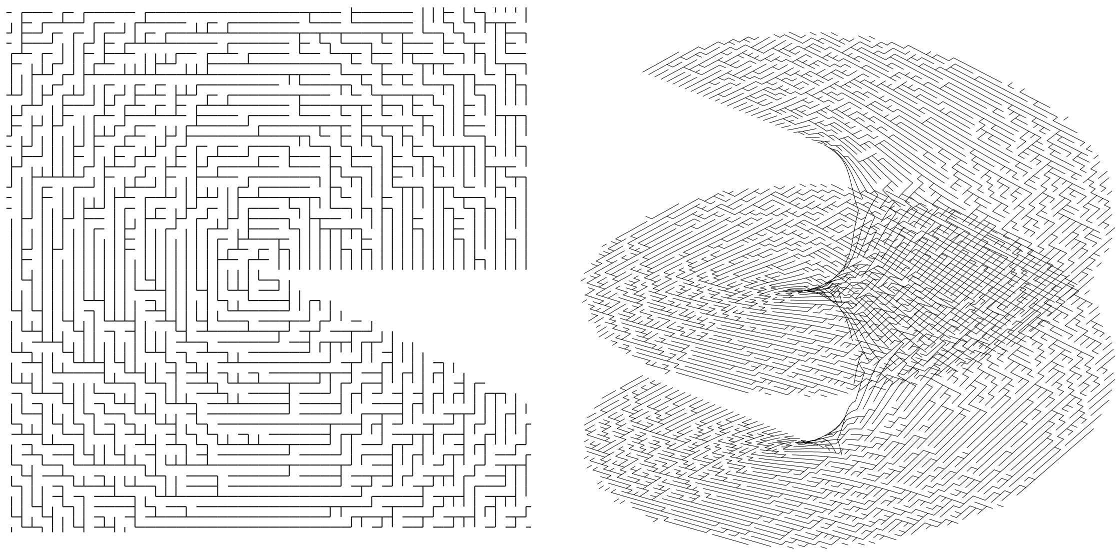

Figure 2: Simulated realizations of the simple harmonic urn percolation model: on a single sheet of $\mathcal{R}$ (*left*) and on a larger section (*right*).

We will interpret a path of the simple harmonic urn as the unique oriented path from some starting vertex through a random subgraph H of G. For each vertex v of G the graph H has precisely one of the out-edges from v that are in G. If the projection of v to $\mathbb{Z}^2$ is (x, y), then the graph H contains the edge from v that projects onto the edge from (x, y) to $(x - \operatorname{sgn}(y), y)$ with probability $|y|/(|x| + |y|)$, and otherwise it contains the edge from v that projects onto the edge from (x, y) to $(x, y + \operatorname{sgn}(x))$. These choices are to be made independently for all vertices v of G. In particular H does not have any edges that project onto either of the co-ordinate axes. The random directed graph H is an oriented percolation model that encodes a coupling of many different paths of the simple harmonic urn. To make this precise, let v_0 be any vertex of G. Then there is a unique oriented path $v_0, v_1, v_2, \ldots$ through H. That is, (v_i, v_{i+1}) is an edge of H for each $i \geq 0$. Let the projection of v_i from $\mathcal{R}$ to $\mathbb{R}^2$ be the point (X_i, Y_i). Then the sequence $(X_i, Y_i)_{i=0}^{\infty}$ is a sample of the simple harmonic urn process. If w_0 is another vertex of G, with unique oriented path $w_0, w_1, w_2, \ldots$, then its projection to $\mathbb{Z}^2$ is also a sample path of the simple harmonic urn process, but we will show (see Theorem 2.11 below) that with probability one the two paths eventually couple, which is to say that there exist random finite $m \geq 0$ and $n \geq 0$ such that for all $i \geq 0$ we have $v_{i+m} = w_{i+n}$. Thus the percolation model encodes many coalescing copies of the simple harmonic urn process. Next we show that it also encodes many copies of the leaky urn of Section 2.2.

We construct another random graph H' that is the dual percolation model to H. We begin with the planar dual of the square-grid lattice, which is another square-grid lattice with vertices at the points $(m+1/2, n+1/2)$, $m, n \in \mathbb{Z}$. We orient all the edges in the direction of decreasing argument, and lift to the covering surface $\mathcal{R}$ to obtain the dual graph G'. Now let H' be the directed subgraph of G' that consists of all those edges of G' that do not cross an edge of H. It turns out that H' can be viewed as an oriented percolation model that encodes a coupling of many different paths of the leaky simple harmonic urn.

To explain this, we define a mapping Φ from the vertices of G' to $\mathbb{Z}^2$. Let (x, y) be the co-ordinates of the projection of $v \in G'$ to the shifted square lattice $\mathbb{Z}^2 + (1/2, 1/2)$. Then

$$\Phi(v) = \left(x + \frac{1}{2} \operatorname{sgn} y, -(y - \frac{1}{2} \operatorname{sgn} x) \right).$$

Thus we project from $\mathcal{R}$ to $\mathbb{R}^2$, move to the nearest lattice point in the clockwise direction, and then reflect in the x-axis. If v_0 is any vertex of H', there is a unique oriented path $v_0, v_1, v_2, \ldots$ through H', this time winding clockwise. Take $v_0 = (z_0 - 1/2, 1/2)$. A little thought shows that the sequence $(X'_i, Y'_i) = \Phi(v_i)$ has the distribution of the leaky simple harmonic urn process. This is because the choice of edge in H' from v is determined by the choice of edge in H from the nearest point of G in the clockwise direction. The map Φ is not quite a graph homomorphism onto the square lattice because of its behaviour at the axes; e.g., it sends $(3\frac{1}{2}, \frac{1}{2})$ and $(3\frac{1}{2}, -\frac{1}{2})$ to $(4, 0)$ and $(3, 1)$ respectively. The decrease of 1 in the x-co-ordinate corresponds to the leaked ball in the leaky urn model. If some v_i has projection (x_i, y_i) with $|x_i| + |y_i| = 1$, then the same is true of all subsequent vertices in the path. This corresponds to the closed class $\mathcal{C}$.

From results on our urn processes, we will deduce the following quite subtle properties of the percolation model H. Let $I(v)$ denote the number of vertices in the *in-graph* of the vertex v in H, which is the subgraph of H induced by all vertices from which it is possible to reach v by following an oriented path.

Theorem 2.11. *Almost surely, the random oriented graph H is, ignoring orientations, an infinite tree with a single semi-infinite end in the out direction. In particular, for any v, $I(v) < \infty$ a.s. and moreover $\mathbb{E}[I(v)^p] < \infty$ for any $p < 2/3$; however, $\mathbb{E}[I(v)] = \infty$.*

The dual graph H' is also an infinite tree a.s., with a single semi-infinite end in the out direction. It has a doubly-infinite oriented path and the in-graph of any vertex not on this path is finite a.s.

2.5 A continuous-time fast embedding of the simple harmonic urn

There is a natural continuous-time embedding of the simple harmonic urn process. Let $(A(t), B(t))_{t \in \mathbb{R}_+}$ be a $\mathbb{Z}^2$-valued continuous-time Markov chain with $A(0) = a_0$, $B(0) = b_0$, and transition rates

$$\begin{aligned} \mathbb{P}(A(t + \mathrm{d}t) = A(t) - \operatorname{sgn}(B(t))) &= |B(t)| \mathrm{d}t, \\ \mathbb{P}(B(t + \mathrm{d}t) = B(t) + \operatorname{sgn}(A(t))) &= |A(t)| \mathrm{d}t. \end{aligned}$$

Given that $(A(t), B(t)) = (a, b)$, the wait until the next jump after time t is an exponential random variable with mean $1/(|a| + |b|)$. The next jump is a change in the first co-ordinate with probability $|b|/(|a| + |b|)$, so the process considered at its sequence of jump times does indeed follow the law of the simple harmonic urn. Note that the process does not explode in finite time since the jump rate at (a, b) is $|a| + |b|$, and $|X_n| + |Y_n| = O(n)$ (as jumps are of size 1), so $\sum_n (|X_n| + |Y_n|)^{-1} = \infty$ a.s..

The process $(A(t), B(t))$ is an example of a *two-dimensional linear birth and death process.* The recurrence classification of such processes defined on $\mathbb{Z}_+^2$ was studied by Kesten [22] and Hutton [17]. Our case (which has $B_{1,1} + B_{2,2} = 0$ in their notation) was not covered by the results in [22, 17]; Hutton remarks [17, p. 638] that "we do not yet know whether this case is recurrent or transient." In the $\mathbb{Z}_+^2$ setting of [22, 17], the boundaries of the quadrant would become absorbing in our case. The model on $\mathbb{Z}^2$ considered here thus seems a natural setting in which to pose the recurrence/transience question left open by [22, 17]. Our Theorem 2.1 implies that $(A(t), B(t))$ is in fact transient.

We call $(A(t), B(t))$ the *fast* embedding of the urn since typically many jumps occur in unit time (the process jumps faster the farther away from the origin it is). There is another continuous-time embedding of the urn model that is also very useful in its analysis, the *slow* embedding described in Section 3 below.

The mean of the process $(A(t), B(t))$ precisely follows the simple harmonic oscillation suggested by the name of the model. This fact is most neatly expressed in the complex plane $\mathbb{C}$. Recall that a complex martingale is a complex-valued stochastic process whose real and imaginary parts are both martingales.

Lemma 2.12. *The process $(M_t)_{t\in\mathbb{R}_+}$ defined by*

$$M_t := \mathrm{e}^{-it}\,(A(t) + iB(t))$$

is a complex martingale. In particular, for $t > t_0$ and $z \in \mathbb{C}$,

$$\mathbb{E}[A(t) + iB(t) \mid A(t_0) + iB(t_0) = z] \;=\; z\mathrm{e}^{i(t-t_0)}.$$

As can be seen directly from the definition, the continuous-time Markov chain $(A(t), B(t))$ admits a constant invariant measure; this fact is closely related to the 'simple harmonic flea circus' that we describe in Section 10.1.

Returning to the dynamics of the process, what is the expected time taken to traverse a quadrant in the fast continuous-time embedding? Define $\tau_f := \inf\{t \in \mathbb{R}_+ : A(t) = 0\}$. We use the notation $\mathbb{P}_n(\,\cdot\,)$ for $\mathbb{P}(\,\cdot \mid A(0) = n, B(0) = 0)$, and similarly for $\mathbb{E}_n$. Numerical calculations strongly suggest the following:

Conjecture 2.13. *Let $n \in \mathbb{N}$. With $\alpha_1 \approx -2.0888$ as in Theorem 2.2 above, as $n \to \infty$,*

$$\mathbb{E}_n[\tau_f] = \pi/2 + O(\mathrm{e}^{\alpha_1 n}/\sqrt{n}).$$

We present a possible approach to the resolution of Conjecture 2.13 in Section 9.3; it turns out that $\mathbb{E}_n[\tau_f]$ can be expressed as a rational polynomial of degree n evaluated at e. The best result that we have been able to prove along the lines of Conjecture 2.13 is the following, which shows not only that $\mathbb{E}_n[\tau_f]$ is close to $\pi/2$ but also that τ_f itself is concentrated about $\pi/2$.

Theorem 2.14. *Let $n \in \mathbb{N}$. For any $\delta > 0$, as $n \to \infty$,*

$$\mathbb{E}_n[\tau_f] = \pi/2 + O(n^{\delta-(1/2)}), \tag{7}$$

$$\mathbb{E}_n[|\tau_f - (\pi/2)|^2] = O(n^{\delta-(1/2)}). \tag{8}$$

In the continuous-time fast embedding the paths of the simple harmonic urn are a discrete stochastic approximation to continuous circular motion at angular velocity 1, with the radius of the motion growing approximately linearly in line with the transience of the process. Therefore a natural quantity to examine is the area enclosed by a path of the urn across the first quadrant, together with the two co-ordinate axes. For a typical path starting at $(n, 0)$ we would expect this to be roughly $\pi n^2/4$, this being the area enclosed by a quarter-circle of radius n about the origin. We use the percolation model to obtain an exact relation between the expected area enclosed and the expected time taken for the urn to traverse the first quadrant.

Theorem 2.15. *For $n \in \mathbb{N}$, for any $\delta > 0$,*

$$\mathbb{E}_n[\text{Area enclosed by a single traversal}] = \sum_{m=1}^{n} m\mathbb{E}_m[\tau_f] \; = \; \frac{\pi n^2}{4} + O(n^{(3/2)+\delta}).$$

In view of the first equality in Theorem 2.15 and Conjecture 2.13, we suspect a sharp version of the asymptotic expression for the expected area enclosed to be

$$\mathbb{E}_n[\text{Area enclosed by a single traversal}] = \frac{\pi n(n+1)}{4} + c + O\left(\sqrt{n}\mathrm{e}^{\alpha_1 n}\right),$$

for some constant $c \in \mathbb{R}$.

2.6 Outline of the paper and related literature

The outline of the remainder the the paper is as follows. We begin with a study of the discrete-time embedded process Z_k in the original urn model. In Section 3 we use a decoupling argument to obtain an explicit formula, involving the Eulerian numbers, for the transition probabilities of Z_k. In Section 4 we study the drift of the process Z_k and prove Theorem 2.2. We make use of an attractive coupling with the renewal process based on the uniform distribution. Then in Section 5 we give a short, stand-alone proof of our basic result, Theorem 2.1. In Section 6 we study the increments of the process Z_k, obtaining tail bounds and moment estimates. As a by-product of our results we obtain (in Lemma 4.5 and Corollary 6.5) sharp expressions for the first two moments of the counting function of the uniform renewal process, improving on existing results in the literature. In Section 7 we study the asymptotic behaviour of the noisy urn embedded process $\tilde{Z}_k$, building on our results on Z_k. Here we make use of powerful results of Lamperti and others on processes with asymptotically zero drift, which we can apply to the process $\tilde{Z}_k^{1/2}$. Then in Section 8 we complete the proofs of Theorems 2.3–2.6, 2.8, 2.9 and 2.11. In Section 9 we study the continuous-time fast embedding described in Section 2.5, and in Sections 9.1 and 9.2 present proofs of Theorems 2.10, 2.14, and 2.15. In Section 9.3 we give some curious exact formulae for the expected area and time described in Section 2.5. Finally, in Section 10 we collect some results on several models that are not directly relevant to our main theorems but that demonstrate further some of the surprising richness of the phenomena associated with the simple harmonic urn and its generalizations.

We finish this section with some brief remarks on modelling applications related to the simple harmonic urn. The simple harmonic urn model has some similarities to R.F. Green's urn model for cannibalism (see e.g. [36]). The cyclic nature of the model is similar to that of various stochastic or deterministic models of certain planar systems with feedback: see for instance [12] and references therein. Finally, one may view the simple harmonic urn as a gated polling model with two queues and a single server. The server serves one queue, while new arrivals are directed to the other queue. The service rate is proportional to the ratio of the numbers of customers in the two queues. Customers arrive at the unserved queue at times of a Poisson process of constant rate. Once the served queue becomes empty, the server switches to the other queue, and a new secondary queue is started. This model gives a third continuous-time embedding of the simple harmonic urn, which we do not study any further in this paper. This polling model differs from typical polling models studied in the literature (see e.g. [29]) in that the service rate depends upon the current state of the system. One possible interpretation of this unusual service rate could be that the customers in the primary queue are in fact served by the waiting customers in the secondary queue.

3 Transition probabilities for Z_k

In this section we derive an exact formula for the transition probabilities of the Markov chain $(Z_k)_{k\in\mathbb{Z}_+}$ (see Lemma 3.3 below). We use a coupling (or rather 'decoupling') idea that is sometimes attributed to Samuel Karlin and Herman Rubin. This construction was used in [24] to study the OK Corral gunfight model, and is closely related to the embedding of a generic generalized Pólya urn in a multi-type branching process [4, 5]. The construction yields another continuous-time embedding of the urn process, which, by way of contrast to the embedding described in Section 2.5, we refer to as the *slow* embedding of the urn.

We couple the segment of the urn process (X_n, Y_n) between times $\nu_k + 1$ and ν_{k+1} with certain birth and death processes, as follows. Let $\lambda_k := 1/k$. Consider two *independent* $\mathbb{Z}_+$-valued continuous-time Markov chains, $U(t)$ and $V(t)$, $t \in \mathbb{R}_+$, where $U(t)$ is a pure death process with the transition rate

$$\mathbb{P}(U(t + \mathrm{d}t) = U(t) - 1 \mid U(t) = a) = \lambda_a \mathrm{d}t,$$

and $V(t)$ is a pure birth process with

$$\mathbb{P}(V(t + \mathrm{d}t) = V(t) + 1 \mid V(t) = b) = \lambda_b \mathrm{d}t.$$

Set $U(0) = z$ and $V(0) = 1$.

From the standard exponential holding-time characterization for continuous-time Markov chains and the properties of independent exponential random variables, it follows that the embedded process $(U(t), V(t))$ considered at the times when either of its coordinates changes has the same distribution as the simple harmonic urn (X_n, Y_n) described above when (X_n, Y_n) is traversing the first quadrant. More precisely, let $\theta_0 := 0$ and define the jump times of the process $V(t) - U(t)$ for $n \in \mathbb{N}$:

$$\theta_n := \inf\{t > \theta_{n-1} : U(t) < U(\theta_{n-1}) \text{ or } V(t) > V(\theta_{n-1})\}.$$

Since $\lambda_b \leq 1$ for all b, the processes $U(t)$, $V(t)$ a.s. do not explode in finite time, so $\theta_n \to \infty$ a.s. as $n \to \infty$. Define $\eta := \min\{n \in \mathbb{N} : U(\theta_n) = 0\}$ and set

$$T := \theta_\eta = \inf\{t > 0 : U(t) = 0\},$$

the extinction time of $U(t)$. The coupling yields the following result (cf [5, §V.9.2]).

Lemma 3.1. *Let $k \in \mathbb{Z}_+$ and $z \in \mathbb{N}$. The sequence $(U(\theta_n), V(\theta_n))$, $n = 0, 1, \ldots, \eta$, with $(U(0), V(0)) = (z, 1)$, has the same distribution as each of the two sequences*

(i) $(|X_n|, |Y_n|)$, $n = \nu_k + 1, \ldots, \nu_{k+1}$, *conditioned on* $Z_k = z$ *and* $Y_{\nu_k} = 0$;

(ii) $(|Y_n|, |X_n|)$, $n = \nu_k + 1, \ldots, \nu_{k+1}$, *conditioned on* $Z_k = z$ *and* $X_{\nu_k} = 0$.

Note that we set $V(0) = 1$ since $(X_{\nu_k+1}, Y_{\nu_k+1})$ is always one step in the 'anticlockwise' lattice direction away from (X_{ν_k}, Y_{ν_k}). Let

$$T'_w := \inf\{t > 0 : V(t) = w\}.$$

We can represent the times T and T'_w as sums of exponential random variables. Write

$$T_z = \sum_{k=1}^{z} k\xi_k, \quad \text{and} \quad T'_w = \sum_{k=1}^{w-1} k\zeta_k, \tag{9}$$

where $\xi_1, \zeta_1, \xi_2, \zeta_2, \ldots$ are independent exponential random variables with mean 1. Then setting $T = T_{U(0)}$, (9) gives useful representations of T and T'_w.

As an immediate illustration of the power of this embedding, observe that $Z_{k+1} \leq Z_k$ if and only if V has not reached $U(0)+1$ by the time of the extinction of U, i.e. $T'_{U(0)+1} > T$. But (9) shows that T'_{z+1} and T_z are identically distributed continuous random variables, so:

Lemma 3.2. *For $z \in \mathbb{N}$, $\mathbb{P}\left(Z_{k+1} \leq Z_k \mid Z_k = z\right) = \mathbb{P}\left(T'_{U(0)+1} > T_{U(0)}\right) = \frac{1}{2}$.*

We now proceed to derive from the coupling described in Lemma 3.1 an exact formula for the transition probabilities of the Markov chain $(Z_k)_{k \in \mathbb{Z}_+}$. Define $p(n,m) = \mathbb{P}\left(Z_{k+1} = m \mid Z_k = n\right)$. It turns out that $p(n,m)$ may be expressed in terms of the *Eulerian numbers* $A(n,k)$, which are the positive integers defined for $n \in \mathbb{N}$ by

$$A(n,k) = \sum_{i=0}^{k} (-1)^i \binom{n+1}{i} (k-i)^n, \quad k \in \{1, \ldots, n\}.$$

The Eulerian numbers have several combinatorial interpretations and have many interesting properties; see for example Bóna [6, Chapter 1].

Lemma 3.3. *For $n, m \in \mathbb{N}$, the transition probability $p(n,m)$ is given by*

$$\begin{aligned} p(n,m) &= m \sum_{r=0}^{m} (-1)^r \frac{(m-r)^{n+m-1}}{r!\,(n+m-r)!} \\ &= \frac{m}{(m+n)!} A(n+m-1, n). \end{aligned}$$

We give two proofs of Lemma 3.3, both using the coupling of Lemma 3.1 but in quite different ways. The first uses moment generating functions and is similar to calculations in [24], while the second involves a time-reversal of the death process and makes use of the recurrence relation satisfied by the Eulerian numbers. Each proof uses ideas that will be useful later on.

First proof of Lemma 3.3. By Lemma 3.1, the conditional distribution of Z_{k+1} on $Z_k = n$ coincides with the distribution conditional of $V(T)$ on $U(0) = n$. So

$$\mathbb{P}(Z_{k+1} > m \mid Z_k = n) = \mathbb{P}(V(T) > m \mid U(0) = n) = \mathbb{P}(T_n > T'_{m+1}), \tag{10}$$

using the representations in (9). Thus from (9) and (10), writing

$$R_{n,m} = \sum_{i=1}^{n} i\xi_i - \sum_{j=1}^{m} j\zeta_j,$$

we have that $\mathbb{P}(Z_{k+1} > m \mid Z_k = n) = \mathbb{P}(R_{n,m} > 0)$. The density of $R_{n,m}$ can be calculated using the moment generating function and partial fractions; for $t \geq 0$,

$$\begin{aligned} \mathbb{E}[\mathrm{e}^{tR_{n,m}}] &= \prod_{i=1}^{n} \frac{1/i}{1/i - t} \times \prod_{j=1}^{m} \frac{1/j}{1/j + t} = \prod_{i=1}^{n} \frac{1}{1 - it} \times \prod_{j=1}^{m} \frac{1}{1 + jt} \\ &= \sum_{i=1}^{n} \frac{a_i}{1 - it} + \sum_{j=1}^{m} \frac{b_j}{1 + jt}, \end{aligned}$$

for some coefficients $a_i = a_{i;n,m}$ and $b_j = b_{j;n,m}$. Multiplying both sides of the last displayed equality by $\prod_{i=1}^n(1-it)\prod_{j=1}^m(1+jt)$ and setting $t=1/i$ we obtain

$$a_i = \prod_{\substack{j=1\\ j\neq i}}^n \frac{1}{1-(j/i)}\prod_{k=1}^m \frac{1}{1+(k/i)} = (-1)^{n-i} i^{n+m-1}\prod_{j=1}^{i-1}\frac{1}{i-j}\prod_{j=i+1}^{n}\frac{1}{j-i}\prod_{k=1}^{m}\frac{1}{k+i}.$$

Simplifying, and then proceeding similarly but taking $t=-1/j$ to identify b_j, we obtain

$$a_i = \frac{(-1)^{n-i} i^{n+m}}{(n-i)!(m+i)!}, \quad \text{and} \quad b_j = \frac{(-1)^{m-j} j^{n+m}}{(m-j)!(n+j)!}.$$

Consequently the density of $R_{n,m}$ is

$$r(x) = \begin{cases} \sum_{i=1}^n a_i i^{-1} \mathrm{e}^{-x/i}, & \text{if } x \geq 0; \\ \sum_{j=1}^m b_j j^{-1} \mathrm{e}^{x/j}, & \text{if } x < 0. \end{cases}$$

Thus we obtain

$$\begin{aligned} \mathbb{P}(Z_{k+1} > m \mid Z_k = n) &= \mathbb{P}(Z_{k+1} \geq m+1 \mid Z_k = n) = \mathbb{P}(R_{n,m} \geq 0) = \sum_{k=1}^n a_{k;n,m} \\ &= \sum_{k=0}^n \frac{(-1)^{n-k} k^{n+m}}{(n-k)!(m+k)!} = \sum_{i=0}^n \frac{(-1)^i (n-i)^{n+m}}{i!(m+n-i)!} \\ &= \frac{1}{(m+n)!}\sum_{i=0}^n (-1)^i \binom{m+n}{i} (n-i)^{n+m}. \end{aligned} \tag{11}$$

It follows that

$$\begin{aligned} p(n,m) &= \mathbb{P}(Z_{k+1} \geq m \mid Z_k = n) - \mathbb{P}(Z_{k+1} \geq m+1 \mid Z_k = n) \\ &= \sum_{i=0}^n \frac{(-1)^i (n-i)^{n+m-1}}{i!(m-1+n-i)!} - \sum_{i=0}^n \frac{(-1)^i (n-i)^{n+m}}{i!(m+n-i)!} \\ &= \sum_{i=0}^n \frac{(-1)^i (n-i)^{n+m-1}}{i!(m+n-i)!}\left[(m+n-i)-(n-i)\right] \\ &= \frac{m}{(m+n)!}\sum_{i=0}^n (-1)^i \binom{m+n}{i}(n-i)^{n+m-1} = \frac{m}{(n+m)!} A(m+n-1,n), \end{aligned}$$

as required. □

Second proof of Lemma 3.3. Consider the birth process $W(t)$ defined by $W(0)=1$ and

$$W(t) = \min\{z \in \mathbb{Z}_+ : \ T_z > t\}, \quad (t>0),$$

where T_z is defined as in (9). The inter-arrival times of $W(t)$ are $(i\xi_i)_{i=1}^z$ and, given $U(0)=z$, the death process $U(t)$ has the same inter-arrival times but taken in the reverse order. The processes $V(t)$ and $W(t)$ are independent and identically distributed. Define for $n,m \in \mathbb{N}$,

$$r(n,m) = \mathbb{P}\left(\exists\, t>0 : W(t)=n,\ V(t)=m \mid V(0)=W(0)=1\right).$$

If $Z_k=n$, then Z_{k+1} is the value of V when W first reaches the value $n+1$; $Z_{k+1}=m$ if and only if the process (W,V) reaches (n,m) *and* then makes the transition to $(n+1,m)$. Since (W,V) is Markov, this occurs with probability $r(n,m)\frac{m}{n+m}$. So for $n,m\in\mathbb{N}$,

$$p(n,m) = \frac{m}{n+m} r(n,m). \tag{12}$$

Conditioning on the site from which (W,V) jumps to (n,m), we get, for $n,m \in \mathbb{N}$, $n+m \geq 3$,

$$r(n,m) \;=\; \frac{m}{n+m-1}\, r(n-1,m) \;+\; \frac{n}{n+m-1}\, r(n,m-1), \tag{13}$$

where $r(0,m) = r(n,0) = 0$. It is easy to check that $r(k,1) = r(1,k) = 1/k!$ for all $k \in \mathbb{N}$. It will be helpful to define

$$s(n,m) \;=\; (n+m-1)!\, r(n,m)\,.$$

Then we have for $n,m \in \mathbb{N}$, $n+m \geq 3$,

$$s(n,m) \;=\; m\, s(n-1,m) \;+\; n\, s(n,m-1)\,,$$

$$s(k,1) = s(1,k) = 1 \quad \text{for all } k \in \mathbb{N}.$$

These constraints completely determine the *positive integers* $s(n,m)$ for all $m,n \in \mathbb{N}$. Since the Eulerian numbers $A(n+m-1,m)$ satisfy the same initial conditions and recurrence relation [6, Thm. 1.7], we have $s(n,m) = A(m+n-1,m)$, which together with (12) gives the desired formula for $p(n,m)$. □

It is evident from (13) and its initial conditions that $r(n,m) = r(m,n)$ for all $n,m \in \mathbb{N}$. So

$$n\, p(n,m) \;=\; m\, p(m,n)\,, \tag{14}$$

Therefore the σ-finite measure π on $\mathbb{N}$ defined by $\pi(n) = n$ satisfies the detailed balance equations and hence is invariant for $p(\cdot,\cdot)$. In fact there is a pathwise relation of the same type, which we now describe. We call a sequence $\omega = (x_j,y_j)_{j=0}^k$ $(k \geq 2)$ of points in $\mathbb{Z}_+^2$ an *admissible traversal* if $y_0 = x_k = 0$, $x_0 \geq 1$, $y_k \geq 1$, each point (x_j,y_j), $2 \leq j \leq k-1$ is one of $(x_{j-1}-1, y_{j-1})$, $(x_{j-1}, y_{j-1}+1)$, and $(x_1,y_1) = (x_0, y_0+1)$, $(x_k,y_k) = (x_{k-1}-1, y_{k-1})$. If ω is an admissible traversal, then so is the time-reversed and reflected path $\omega' = (y_{k-j}, x_{k-j})_{j=0}^k$. In fact, conditioning on the endpoints, ω and ω' have the same probability of being realized by the simple harmonic urn:

Lemma 3.4. *For any admissible traversal $(x_j,y_j)_{j=0}^k$ with $x_0 = n \in \mathbb{N}$, $y_k = m \in \mathbb{N}$,*

$$\mathbb{P}((X_j,Y_j)_{j=0}^{\nu_1} = (x_j,y_j)_{j=0}^k \mid Z_0 = n, Z_1 = m) = \mathbb{P}((X_j,Y_j)_{j=0}^{\nu_1} = (y_{k-j},x_{k-j})_{j=0}^k \mid Z_0 = m, Z_1 = n).$$

Proof. Let $\omega = (x_j,y_j)_{j=0}^k$ be an admissible traversal, and define

$$p = p(\omega) = \mathbb{P}((X_j,Y_j)_{j=0}^{\nu_1} = (x_j,y_j)_{j=0}^k, Z_1 = m \mid Z_0 = n), \text{ and}$$

$$p' = p'(\omega) = \mathbb{P}((X_j,Y_j)_{j=0}^{\nu_1} = (y_{k-j},x_{k-j})_{j=0}^k, Z_1 = n \mid Z_0 = m),$$

so that p' is the probability of the reflected and time-reversed path. To prove the lemma it suffices to show that for any ω with $(x_0,y_0) = (n,0)$ and $(x_k,y_k) = (0,m)$, $p(\omega)/p(n,m) = p'(\omega)/p(m,n)$. In light of (14), it therefore suffices to show that $np = mp'$. To see this, we use the Markov property along the path ω to obtain

$$p = \prod_{j=0}^{k-1} (x_j+y_j)^{-1} (x_j \mathbf{1}_{\{x_{j+1}=x_j\}} + y_j \mathbf{1}_{\{y_{j+1}=y_j\}}),$$

while, using the Markov property along the reflection and reversal of ω,

$$p' = \prod_{j=0}^{k-1} (x_{k-j}+y_{k-j})^{-1} (x_{k-j} \mathbf{1}_{\{x_{k-j-1}=x_{k-j}\}} + y_{k-j} \mathbf{1}_{\{y_{k-j-1}=y_{k-j}\}})$$

$$= \prod_{i=0}^{k-1} (x_{i+1} + y_{i+1})^{-1} (x_i \mathbf{1}_{\{x_{i+1}=x_i\}} + y_i \mathbf{1}_{\{y_{i+1}=y_i\}}),$$

making the change of variable $i = k - j - 1$. Dividing the two products for p and p' yields, after cancellation, $p/p' = (x_k + y_k)/(x_0 + y_0) = m/n$, as required. □

Remarks. Of course by summing over paths in the equality $np(\omega) = mp'(\omega)$ we could use the argument in the last proof to prove (14). The reversibility and the invariant measure exhibited in Lemma 3.4 and (14) will appear naturally in terms of a stationary model in Section 10.1.

4 Proof of Theorem 2.2 via the uniform renewal process

In this section we study the asymptotic behaviour of $\mathbb{E}[Z_{k+1} \mid Z_k = n]$ as $n \to \infty$. The explicit expression for the distribution of Z_{k+1} given $Z_k = n$ obtained in Lemma 3.3 turns out not to be very convenient to use directly. Thus we proceed somewhat indirectly and exploit a connection with a renewal process whose inter-arrival times are uniform on $(0, 1)$. Here and subsequently we use $U(0, 1)$ to denote the uniform distribution on $(0, 1)$.

Let $\chi_1, \chi_2, \chi_3, \ldots$ be an i.i.d. sequence of $U(0, 1)$ random variables. Consider the renewal sequence S_i, $i \in \mathbb{Z}_+$ defined by $S_0 := 0$ and, for $i \geq 1$, $S_i := \sum_{j=1}^{i} \chi_j$. For $t \geq 0$ define the counting process

$$N(t) := \min\{i \in \mathbb{Z}_+ : S_i > t\} = 1 + \max\{i \in \mathbb{Z}_+ : S_i \leq t\}\,, \tag{15}$$

so a.s., $N(t) \geq t + 1$. In the language of classical renewal theory, $\mathbb{E}[N(t)]$ is a renewal function (note that we are counting the renewal at time 0). The next result establishes the connection between the uniform renewal process and the simple harmonic urn.

Lemma 4.1. *For each $n \in \mathbb{N}$, the conditional distribution of Z_{k+1} on $Z_k = n$ equals the distribution of $N(n) - n$. In particular, for $n \in \mathbb{N}$, $\mathbb{E}[Z_{k+1} \mid Z_k = n] = \mathbb{E}[N(n)] - n$.*

The proof of Lemma 4.1 amounts to showing that $\mathbb{P}(N(n) = n + m) = p(n, m)$ as given by Lemma 3.3. This equality is Theorem 3 in [41], and it may be verified combinatorially using the interpretation of $A(n, k)$ as the number of permutations of $\{1, \ldots, n\}$ with exactly $k - 1$ falls, together with the observation that for $n \in \mathbb{N}$, $N(n)$ is the position of the n^{th} fall in the sequence $\psi_1, \psi_2, \ldots$, where $\psi_k = S_k \bmod 1$, another sequence of i.i.d. $U(0, 1)$ random variables. Here we will give a neat proof of Lemma 4.1 using the coupling exhibited above in Section 3.

Proof of Lemma 4.1. Consider a doubly-infinite sequence $(\xi_i)_{i \in \mathbb{Z}}$ of independent exponential random variables with mean 1. Taking $\zeta_k = \xi_{-k}$, we can write $R_{n,m}$ (as defined in the first proof of Lemma 3.3) as $\sum_{i=-m}^{n} i\xi_i$. Define $S_{n,m} = \sum_{i=-m}^{n} \xi_i$. For fixed $n \in \mathbb{N}$, $m \in \mathbb{Z}_+$ we consider normalized partial sums

$$\chi'_j = \left(\sum_{i=-m}^{j-1-m} \xi_i \right) / S_{n,m}, \quad j \in \{1, \ldots, n + m\}.$$

Since $(S_{j-1-m,m})_{j=1}^{n+m}$ are the first $n+m$ points of a unit-rate Poisson process on $\mathbb{R}_+$, the vector $(\chi'_1, \chi'_2, \ldots, \chi'_{n+m})$ is distributed as the vector of increasing order statistics of the $n + m$ i.i.d. $U(0, 1)$ random variables $\chi_1, \ldots, \chi_{n+m}$. In particular,

$$\mathbb{P}\left(N(n) > n + m\right) = \mathbb{P}\left(\sum_{i=1}^{n+m} \chi_i \leq n \right) = \mathbb{P}\left(\sum_{i=1}^{n+m} \chi'_i \leq n \right),$$

using the fact that, by (15), $\{N(n) > r\} = \{S_r \le n\}$ for $r \in \mathbb{Z}_+$ and $n > 0$. But

$$n - \sum_{i=1}^{n+m} \chi'_i = \sum_{i=m+1}^{m+n} (1 - \chi'_i) - \sum_{i=1}^{m} \chi'_i = \left(\sum_{i=-m}^{n} i\xi_i \right) / S_{n,m} = R_{n,m}/S_{n,m}.$$

So, using the equation two lines above (11),

$$\mathbb{P}\left(N(n) - n > m\right) = \mathbb{P}\left(R_{n,m} \ge 0\right) = \mathbb{P}\left(Z_{k+1} > m \mid Z_k = n\right).$$

Thus $N(n) - n$ has the same distribution as Z_{k+1} conditional on $Z_k = n$. □

In view of Lemma 4.1, to study $\mathbb{E}[Z_{k+1} \mid Z_k = n]$ we need to study $\mathbb{E}[N(n)]$.

Lemma 4.2. *As $n \to \infty$,*

$$\mathbb{E}[N(n)] - \left(2n + \frac{2}{3}\right) \to 0.$$

Proof. This is a consequence of the renewal theorem. For a general non-arithmetic renewal process whose inter-arrival times have mean μ and variance σ^2, let $U(t)$ be the expectation of the number of arrivals up to time t, including the initial arrival at time 0. Then

$$U(t) - \frac{t}{\mu} \to \frac{\sigma^2 + \mu^2}{2\mu^2}, \quad \text{as } t \to \infty. \tag{16}$$

We believe this is due to Smith [38]. See e.g. Feller [11, §XI.3, Thm. 1], Cox [8, §4] or Asmussen [1, §V, Prop. 6.1]. When the inter-arrival distribution is $U(0,1)$, we have $U(t) = \mathbb{E}[N(t)]$ with the notation of (15), and in this case $\mu = 1/2$ and $\sigma^2 = 1/12$. □

Together with Lemma 4.1, Lemma 4.2 gives the following result.

Corollary 4.3. $\mathbb{E}[Z_{k+1} \mid Z_k = n] - n \to \frac{2}{3}$ *as $n \to \infty$.* □

To obtain the exponential error bound in (4) above, we need to know more about the rate of convergence in Corollary 4.3 and hence in Lemma 4.2. The existence of a bound like (4) for *some* $\alpha_1 < 0$ follows from known results: Stone [40] gave an exponentially small error bound in the renewal theorem (16) for inter-arrival distributions with exponentially decaying tails, and an exponential bound also follows from the coupling proof of the renewal theorem (see e.g. Asmussen [1, §VII, Thm. 2.10 and Problem 2.2]). However, in this particular case we can solve the renewal equation exactly and deduce the asymptotics more precisely, identifying a (sharp) value for α_1 in (4). The first step is the following result.

Lemma 4.4. *Let $\chi_1, \chi_2, \ldots$ be an i.i.d. sequence of $U(0,1)$ random variables. For $t \in \mathbb{R}_+$,*

$$\mathbb{P}\left(\sum_{i=1}^{k} \chi_i \le t\right) = \sum_{i=0}^{k \wedge \lfloor t \rfloor} \frac{(t-i)^k (-1)^i}{i!(k-i)!},$$

and

$$\mathbb{E}[N(t)] = U(t) = \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} \chi_i \le t\right) = \sum_{i=0}^{\lfloor t \rfloor} \frac{(i-t)^i \, \mathrm{e}^{t-i}}{i!}. \tag{17}$$

Proof. The first formula is classical (see e.g. [11, p. 27]); according to Feller [10, p. 285], it is due to Lagrange. The second formula follows from observing (with an empty sum being 0)

$$U(t) = \mathbb{E} \sum_{k=0}^{\infty} \mathbf{1}\left\{\sum_{i=1}^{k} \chi_i \le t\right\} = \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} \chi_i \le t\right),$$

and exchanging the order in the consequent double sum (which is absolutely convergent). $\square$

We next obtain a more tractable explicit formula for the expression in (17). Define for $t \geq 0$

$$f(t) := \sum_{i=0}^{\lfloor t \rfloor} \frac{(i-t)^i \, \mathrm{e}^{t-i}}{i!} .$$

It is easy to verify (see also [1, p. 148]) that f is continuous on $[0,\infty)$ and satisfies

$$\begin{aligned} f(t) &= \mathrm{e}^t, & (0 \leq t \leq 1), \\ f'(t) &= f(t) - f(t-1), & (t \geq 1). \end{aligned} \tag{18}$$

Lemma 4.5. *For all $t > 0$,*

$$f(t) = 2t + \frac{2}{3} + \sum_{\substack{\gamma \in \mathbb{C}:\ \gamma \neq 0, \\ \gamma = 1 - \exp(-\gamma)}} \frac{1}{\gamma} \mathrm{e}^{\gamma t} . \tag{19}$$

The sum is absolutely convergent, uniformly for t in (ε, ∞) for any $\varepsilon > 0$.

Proof. The Laplace transform $\mathcal{L}f(\lambda)$ of f exists for $\mathrm{Re}(\lambda) > 0$ since $f(t) = 2t + 2/3 + o(1)$ as $t \to \infty$, by (17) and Lemma 4.2. Using the differential-delay equation (18) we obtain

$$\mathcal{L}f(\lambda) = \frac{1}{\lambda - 1 + \mathrm{e}^{-\lambda}} .$$

The principal part of $\mathcal{L}f$ at 0 is $\frac{2}{\lambda^2} + \frac{2}{3\lambda}$. There are simple poles at the non-zero roots of $\lambda - 1 + \mathrm{e}^{-\lambda}$, which occur in complex conjugate pairs $\alpha_n \pm i\beta_n$, where $\alpha = \alpha_1 > \alpha_2 > \cdots$ and $0 < \beta_1 < \beta_2 < \cdots$. In fact, $\alpha_n = -\log(2\pi n) + o(1)$ and $\beta_n = (2n + \frac{1}{2})\pi + o(1)$ as $n \to \infty$. For $\gamma = \alpha_n + i\beta_n$ the absolute value of the term $\mathrm{e}^{\gamma t}/\gamma$ in the right-hand side of (19) is $1/(|\gamma||1-\gamma|^t)$, hence the sum converges absolutely, uniformly on any interval (ε, ∞), $\varepsilon > 0$.

To establish (19) we will compute the Bromwich integral (inverting the Laplace transform), using a carefully chosen sequence of rectangular contours:

$$f(t) = \lim_{R \to \infty} \int_{\varepsilon - iR}^{\varepsilon + iR} \frac{\mathrm{e}^{\lambda t}}{\lambda - 1 + \exp(-\lambda)} \, \mathrm{d}\lambda .$$

To evaluate this limit for a particular value of $t > 0$, we will take $\varepsilon = 1/t$ and integrate around a sequence C_n of rectangular contours, with vertices at $(1/t) \pm (2n - \frac{1}{2})\pi i$ and $-2\log n \pm (2n - \frac{1}{2})\pi i$. The integrand along the vertical segment at real part $-2\log n$ is bounded by $(1 + o(1))/n^2$ and the integrand along the horizontal segments is bounded by $\mathrm{e}/(2n - \frac{1}{2})\pi$ because the imaginary parts of λ and $\mathrm{e}^{-\lambda}$ have the same sign there, so $|\lambda - 1 + \mathrm{e}^{-\lambda}| \geq \mathrm{Im}(\lambda)$. It follows that the integrals along these three arcs all tend to zero as $n \to \infty$. Each pole lies inside all but finitely many of the contours C_n, so the principal value of the Bromwich integral is the sum of the residues of $\mathrm{e}^{\lambda t}/(\lambda - 1 + \exp(-\lambda))$. The residue at 0 is $2t + 2/3$, and the residue at $\gamma = \alpha_n + i\beta_n$ is $\mathrm{e}^{\gamma t}/\gamma$. Thus we obtain (19). $\square$

Proof of Theorem 2.2. The statement of the theorem follows from Lemma 4.5, since by Lemma 4.1 and (17) we have $\mathbb{E}[Z_{k+1} \mid Z_k = n] = f(n) - n$ for $n \in \mathbb{N}$. $\square$

Remarks. According to Feller [11, Problem 2, p. 385], equation (17) "is frequently rediscovered in queuing theory, but it reveals little about the nature of U." We have not found the formula (19) in the literature. The dominant term in $f(t) - 2t - 2/3$ as $t \to \infty$ is $\mathrm{e}^{\gamma_1 t}/\gamma_1 + \mathrm{e}^{\overline{\gamma_1} t}/\overline{\gamma_1}$, i.e.

$$\frac{1}{\alpha_1^2 + \beta_1^2} \mathrm{e}^{\alpha_1 t} (\beta_1 \sin(\beta_1 t) + \alpha_1 \cos(\beta_1 t)) ,$$

which changes sign infinitely often. After subtracting this term, the remainder is $O\left(e^{\alpha_2 t}\right)$. The method that we have used for analysing the asymptotic behaviour of solutions to the renewal equation was proposed by A.J. Lotka and was put on a firm basis by Feller [9]; Laplace transform inversions of this kind were dealt with by Churchill [7].

5 Proof of Theorem 2.1

The recurrence relation (13) for $r(n,m)$ permits a direct proof of Theorem 2.1 (transience), without appealing to the more general Theorem 2.4, via standard martingale arguments applied to $h(Z_k)$ for a judicious choice of function h. This is the subject of this section.

Rewriting (13) in terms of p yields the following recurrence relation, which does not seem simple to prove by conditioning on a step in the urn model; for $n,m \in \mathbb{N}$, $n+m \geq 3$,

$$\binom{n+m}{m} p(n,m) \; = \; p(n-1,m) + \left(\frac{n}{m-1}\right) p(n,m-1), \tag{20}$$

where if $m=1$ we interpret the right-hand side of (20) as just $p(n-1,1)$, and where $p(0,m) = p(n,0) = 0$. Note $p(1,1) = 1/2$. For ease of notation, for any function F we will write $\mathbb{E}_n[F(Z)]$ for $\mathbb{E}[F(Z_{k+1}) \mid Z_k = n]$, which, by the Markov property, does not depend on k.

Lemma 5.1. *Let $\alpha_1 \approx -2.0888$ be as in Theorem 2.2. Then for $n \geq 2$,*

$$\mathbb{E}_n\left[\frac{1}{Z}\right] = \frac{\mathbb{E}_n[Z] - \mathbb{E}_{n-1}[Z]}{n} = \frac{1}{n} + O\left(e^{\alpha_1 n}\right),$$

$$\mathbb{E}_n\left[\frac{1}{Z^2(Z+1)}\right] = \frac{\mathbb{E}_{n-1}[1/Z] - \mathbb{E}_n[1/Z]}{n} = \frac{1}{n^2(n-1)} + O\left(e^{\alpha_1 n}\right),$$

where the asymptotics refer to the limits as $n \to \infty$.

Proof. We use the recurrence relation (20) satisfied by the transition probabilities of Z_k. First, multiply both sides of (20) by m, to get for $n,m \in \mathbb{N}$, $m+n \geq 3$,

$$(n+m)p(n,m) = mp(n-1,m) + np(n,m-1) + \frac{n}{m-1}p(n,m-1),$$

where $p(n,0) = p(0,m) = 0$. Summing over $m \in \mathbb{N}$ we obtain for $n \geq 2$,

$$n + \mathbb{E}_n[Z] = \mathbb{E}_{n-1}[Z] + n + n\mathbb{E}_n[1/Z],$$

which yields the first equation of the lemma after an application of (4).

For the second equation, divide (20) through by m to get for $n,m \in \mathbb{N}$, $m+n \geq 3$,

$$\frac{(n+m)}{m^2}p(n,m) \; = \; \frac{1}{m}p(n-1,m) + \frac{n}{m(m-1)}p(n,m-1).$$

On summing over $m \in \mathbb{N}$ this gives, for $n \geq 2$,

$$n\mathbb{E}_n[1/Z^2] + \mathbb{E}_n[1/Z] = \mathbb{E}_{n-1}[1/Z] + n\mathbb{E}_n\left[\frac{1}{(Z+1)Z}\right],$$

which gives the second equation when we apply the asymptotic part of the first equation. □

Proof of Theorem 2.1. Note that $h(x) = \frac{1}{x} - \frac{1}{x^2(x+1)}$ satisfies $h(n) > 0$ for all $n \in \mathbb{N}$ while $h(n) \to 0$ as $n \to \infty$. By Lemma 5.1 we have

$$\mathbb{E}_n[h(Z)] = \mathbb{E}_n[1/Z] - \mathbb{E}_n\left[\frac{1}{Z^2(Z+1)}\right] = \frac{1}{n} - \frac{1}{n^2(n-1)} + O\left(e^{\alpha_1 n}\right),$$

which is less than $h(n)$ for n sufficiently large. In particular, $h(Z_k)$ is a positive supermartingale for Z_k outside a finite set. Hence a standard result such as [1, Prop. 5.4, p. 22] implies that the Markov chain (Z_k) is transient. □

6 Moment and tail estimates for $Z_{k+1} - Z_k$

In order to study the asymptotic behaviour of $(Z_k)_{k \in \mathbb{Z}_+}$, we build on the analysis of Section 4 to obtain more information about the increments $Z_{k+1} - Z_k$. We write $\Delta_k := Z_{k+1} - Z_k$ ($k \in \mathbb{Z}_+$). From the relation to the uniform renewal process, by Lemma 4.1, we have that

$$\mathbb{P}(\Delta_k > x \mid Z_k = n) = \mathbb{P}(N(n) > 2n + x) = \mathbb{P}\left(\sum_{i=1}^{2n+x} \chi_i \le n \right), \tag{21}$$

where $\chi_1, \chi_2, \ldots$ are i.i.d. $U(0,1)$ random variables, using the notation at (15).

Lemma 6.2 below gives a tail bound for $|\Delta_k|$ based on (21) and a sharp bound for the moment generating function of a $U(0,1)$ random variable, for which we have not been able to find a reference and which we state first since it may be of interest in its own right.

Lemma 6.1. *For χ a $U(0,1)$ variable with moment generating function given for $\lambda \in \mathbb{R}$ by*

$$\phi(\lambda) = \mathbb{E}[\mathrm{e}^{\lambda \chi}] = \frac{\mathrm{e}^{\lambda} - 1}{\lambda}, \tag{22}$$

we have

$$\log \phi(-\lambda) \le -\frac{\lambda}{2} + \frac{\lambda^2}{24} \quad (\lambda \ge 0); \qquad \log \phi(\lambda) \le \frac{\lambda}{2} + \frac{\lambda^2}{24} \quad (\lambda \ge 0).$$

Proof. Consider the first of the two stated inequalities. Exponentiating and multiplying both sides by $\lambda \mathrm{e}^{\lambda/2}$, this is equivalent to

$$2 \sinh(\lambda/2) \le \lambda \exp(\lambda^2/24) \tag{23}$$

for all $\lambda \ge 0$. The inequality (23) is easily verified since both sides are entire functions with non-negative Taylor coefficients and the right-hand series dominates the left-hand series term by term, because $6^n n! \le (2n+1)!$ for all $n \in \mathbb{N}$. The second stated inequality reduces to (23) also on exponentiating and multiplying through by $\lambda \mathrm{e}^{-\lambda/2}$. □

Now we can state our tail bound for $|\Delta_k|$. The bound in Lemma 6.2 is a slight improvement on that provided by Bernstein's inequality in this particular case; the latter yields a weaker bound with $4x$ instead of $2x$ in the denominator of the exponential.

Lemma 6.2. *For $n \in \mathbb{N}$ and any integer $x \ge 0$ we have*

$$\mathbb{P}(|\Delta_k| > x \mid Z_k = n) \le 2 \exp\left\{ -\frac{3x^2}{4n + 2x} \right\}.$$

Proof. From (21) and Markov's inequality, we obtain for $x \ge 0$ and any $\lambda \ge 0$,

$$\mathbb{P}(\Delta_k > x \mid Z_k = n) = \mathbb{P}\left(\exp\left\{ -\lambda \sum_{i=1}^{2n+x} \chi_i \right\} \ge \mathrm{e}^{-\lambda n} \right) \le \exp\left\{ \lambda n + (2n+x) \log \phi(-\lambda) \right\},$$

where ϕ is given by (22). With $\lambda = 6x/(2n+x)$ the first inequality of Lemma 6.1 yields

$$\mathbb{P}(\Delta_k > x \mid Z_k = n) \le \exp\left\{ -\frac{x\lambda}{4} \right\} = \exp\left\{ -\frac{3x^2}{4n+2x} \right\}.$$

On the other hand, for $x \in [0, n-1]$, from (21) and Markov's inequality once more,

$$\mathbb{P}(\Delta_k \le -x \mid Z_k = n) = \mathbb{P}\left(\sum_{i=1}^{2n-x} \chi_i > n \right) = \mathbb{P}\left(\exp\left\{ \lambda \sum_{i=1}^{2n-x} \chi_i \right\} > \mathrm{e}^{\lambda n} \right)$$

$$\leq \quad \exp\{-\lambda n + (2n - x)\log\phi(\lambda)\}.$$

On setting $\lambda = 6x/(2n-x)$, the second inequality in Lemma 6.1 yields, for any $x \in [0, n-1]$,

$$\mathbb{P}(\Delta_k < -x \mid Z_k = n) \leq \exp\left\{-\frac{3x^2}{4n-2x}\right\} \leq \exp\left\{-\frac{3x^2}{4n+2x}\right\},$$

while $\mathbb{P}(\Delta_k < -n \mid Z_k = n) = 0$. Combining the left and right tail bounds completes the proof. □

Next from Lemma 6.2 we obtain the following large deviation and moment bounds for Δ_k.

Lemma 6.3. *Suppose that $\varepsilon > 0$. Then for some $C < \infty$ and all $n \in \mathbb{N}$,*

$$\mathbb{P}(|\Delta_k| > n^{(1/2)+\varepsilon} \mid Z_k = n) \leq C\exp\{-n^{\varepsilon}\}. \tag{24}$$

Also for each $r \in \mathbb{N}$, there exists $C(r) < \infty$ such that for any $n \in \mathbb{N}$,

$$\mathbb{E}[|\Delta_k|^r \mid Z_k = n] \leq C(r)n^{r/2}. \tag{25}$$

Proof. The bound (24) is straightforward from Lemma 6.2. For $r \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}[|\Delta_k|^r \mid Z_k = n] &\leq \int_0^\infty \mathbb{P}(|\Delta_k| \geq \lfloor x^{1/r} \rfloor \mid Z_k = n)\,\mathrm{d}x \\ &\leq C\int_0^{n^r} \exp\left\{-\frac{x^{2/r}}{2n}\right\}\mathrm{d}x + C\int_{n^r}^\infty \exp\left\{-\frac{x^{1/r}}{2}\right\}\mathrm{d}x, \end{aligned} \tag{26}$$

for some $C < \infty$, by Lemma 6.2. With the substitution $y = x^{1/r}$, the second integral on the last line of (26) is seen to be $O(n^{r-1}\mathrm{e}^{-n})$ by asymptotics for the incomplete Gamma function. The first integral on the last line of (26), with the substitution $y = (2n)^{-1}x^{2/r}$, is equal to

$$\frac{(2n)^{r/2}r}{2}\int_0^{n/2} \mathrm{e}^{-y}y^{(r/2)-1}\mathrm{d}y \leq \Gamma(r/2)(2n)^{r/2}r/2.$$

Combining the last two upper bounds we verify (25). □

The next result gives sharp asymptotics for the first two moments of $\Delta_k = Z_{k+1} - Z_k$.

Lemma 6.4. *Let $\alpha_1 \approx -2.0888$ be as in Theorem 2.2. Then as $n \to \infty$,*

$$\mathbb{E}[\Delta_k \mid Z_k = n] \;=\; \frac{2}{3} + O(\mathrm{e}^{\alpha_1 n}), \tag{27}$$

$$\mathbb{E}[\Delta_k^2 \mid Z_k = n] \;=\; \frac{2}{3}n + \frac{2}{3} + O(n\mathrm{e}^{\alpha_1 n}). \tag{28}$$

Proof. The equation (27) is immediate from (4). Now we observe that $J_n := X_n^2 + Y_n^2 - n$ is a martingale. Indeed, for any $(x, y) \in \mathbb{Z}^2$,

$$\mathbb{E}[J_{n+1} - J_n \mid (X_n, Y_n) = (x, y)] = \frac{|x|}{|x|+|y|}(2y\,\mathrm{sgn}(x) + 1) + \frac{|y|}{|x|+|y|}(-2x\,\mathrm{sgn}(y) + 1) - 1 = 0.$$

Between times ν_k and ν_{k+1}, the urn takes $Z_k + Z_{k+1}$ steps, so $\nu_{k+1} - \nu_k = Z_k + Z_{k+1}$. Moreover, $J_{\nu_k} = Z_k^2 - \nu_k$. Applying the optional stopping theorem at ν_k and ν_{k+1} we have that

$$J_{\nu_k} = Z_k^2 - \nu_k = \mathbb{E}[J_{\nu_{k+1}} \mid Z_k] = \mathbb{E}[Z_{k+1}^2 - \nu_{k+1} \mid Z_k] = \mathbb{E}[Z_{k+1}^2 - Z_{k+1} \mid Z_k] - \nu_k - Z_k.$$

The optional stopping theorem is applicable here since a.s. $J_n \leq Cn^2$ for some $C < \infty$ and all n, while there is an exponential tail-bound for $\nu_{k+1} - \nu_k$ (see Lemma 8.1 below). Rearranging the equation in the last display, it follows that for $n \in \mathbb{N}$,

$$\mathbb{E}[Z_{k+1}^2 \mid Z_k = n] = n^2 + n + \mathbb{E}[Z_{k+1} \mid Z_k = n]. \tag{29}$$

Writing $\Delta_k = Z_{k+1} - Z_k$, we have that

$$\mathbb{E}[\Delta_k^2 \mid Z_k = n] = \mathbb{E}[Z_{k+1}^2 \mid Z_k = n] - 2n\mathbb{E}[Z_{k+1} \mid Z_k = n] + n^2,$$

which with (29) and (4) gives (28). □

Remark. In view of Lemma 4.1, we could have used renewal theory (e.g., [39]) to estimate $\mathbb{E}[\Delta_k^2 \mid Z_k = n]$. However, no result we could find in the literature would yield a bound as sharp as that in (28).

Lemma 4.1 with (27) and (28) implies an ancillary result on the $U(0,1)$ renewal process:

Corollary 6.5. *Let $N(t)$ be the counting function of the uniform renewal process, as defined by (15). Then with $\alpha_1 \approx -2.0888$ as in Theorem 2.2, as $t \to \infty$,*

$$\mathbb{E}[N(t)^2] = 4t^2 + \frac{10}{3}t + \frac{2}{3} + O(t\mathrm{e}^{\alpha_1 t}); \quad \mathbb{V}\mathrm{ar}[N(t)] = \frac{2}{3}t + \frac{2}{9} + O(t\mathrm{e}^{\alpha_1 t}).$$

These asymptotic results are both sharper than any we have seen in the literature; see e.g. [19, 41] in the particular case of the uniform renewal process or [39] for the general case. We remark that the formula given in [41, p. 231] for $\mathbb{E}[N(t)^2]$ contains an error (in [41] the renewal at 0 is not counted, so the notation $m_k(\cdot)$ there is equivalent to our $\mathbb{E}[(N(\cdot) - 1)^k]$).

7 Asymptotic analysis of the noisy urn

7.1 Connection to Lamperti's problem

In this section we study the noisy urn model described in Section 2.3. To study the asymptotic behaviour of $(\tilde{Z}_k)_{k \in \mathbb{Z}_+}$, it turns out to be more convenient to work with the process $(W_k)_{k \in \mathbb{Z}_+}$ defined by $W_k = \tilde{Z}_k^{1/2}$, since the latter process has asymptotically-zero drift, in a sense to be made precise shortly, and such processes have been well-studied in the literature.

Let $(W_k)_{k \in \mathbb{Z}_+}$ be an irreducible time-homogeneous Markov chain whose state-space is an unbounded countable subset of $\mathbb{R}_+$. Define the increment moment functions

$$\mu_r(x) := \mathbb{E}\left[(W_{k+1} - W_k)^r \mid W_k = x\right]; \tag{30}$$

by the Markov property, when the corresponding moments exist the $\mu_r(x)$ are genuine functions of x. Given a reasonable choice of scale for the process W_k, it is common that $\mu_2(x)$ be uniformly bounded away from 0 and ∞. In this case, under some mild additional regularity conditions, the regime where $x|\mu_1(x)| = O(1)$ is critical from the point of view of the recurrence classification of W_k. For a nearest-neighbour random walk on $\mathbb{Z}_+$ this fact had been known for a long time (see [15]), but a study of this and many other aspects of the problem, in much greater generality (with absence of the Markovian and countable state-space assumptions), was carried out by Lamperti [26, 27, 28] using martingale techniques. Thus the analysis of processes with asymptotically zero drift (i.e. $\mu_1(x) \to 0$) is sometimes known as *Lamperti's problem.*

We will next state some consequences of Lamperti's results that we will use. For convenience, we impose conditions that are stronger than Lamperti's. We suppose that for each $r \in \mathbb{N}$,

$$\sup_x |\mu_r(x)| < \infty. \tag{31}$$

The recurrence and transience properties of W_k were studied by Lamperti [26, 28] and his results were refined by Menshikov, Asymont and Iasnogorodskii [31]. Parts (i) and (ii) of the following result are consequences of Theorems 3.1 and 3.2 of [26] with Theorem 2.1 of [28], while part (iii) is a consequence of Theorem 3 of [31] (which is in fact a much sharper result).

Proposition 7.1. *[26, 28, 31] Let (W_k) be an irreducible Markov chain on a countable unbounded subset of $\mathbb{R}_+$. Suppose that (31) holds, and that there exists $v > 0$ such that $\mu_2(x) > v$ for all x sufficiently large. Then the following recurrence criteria are valid.*

(i) W_k is transient if there exist $\delta, x_0 \in (0, \infty)$ such that for all $x > x_0$,

$$2x\mu_1(x) - \mu_2(x) > \delta.$$

(ii) W_k is positive-recurrent if there exist $\delta, x_0 \in (0, \infty)$ such that for all $x > x_0$,

$$2x\mu_1(x) + \mu_2(x) < -\delta.$$

(iii) W_k is null-recurrent if there exists $x_0 \in (0, \infty)$ such that for all $x > x_0$,

$$2x|\mu_1(x)| \leq \left(1 + \frac{1}{\log x}\right) \mu_2(x).$$

In [27], Lamperti proved the existence of weak-sense limiting diffusions for certain processes satisfying parts (i) or (iii) of Proposition 7.1. To state Lamperti's result we need some more notation. To describe the time-homogeneous diffusions on $\mathbb{R}_+$ that arise here, it will suffice to describe the infinitesimal mean $\mu(x)$ and infinitesimal variance $\sigma^2(x)$; see e.g. [21, Ch 15]. The transition functions p of our diffusions will then satisfy the Kolmogorov backwards equation

$$\frac{\partial p}{\partial t} = \mu(x)\frac{\partial p}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2 p}{\partial x^2}.$$

Let $(H_t^{\alpha,\beta})_{t\in[0,1]}$ denote a diffusion process on $\mathbb{R}_+$ with infinitesimal mean and variance

$$\mu(x) = \frac{\alpha}{x}, \quad \sigma^2(x) = \beta. \tag{32}$$

The particular case of a diffusion satisfying (32) with $\beta = 1$ and $\alpha = (\gamma - 1)/2$ for some $\gamma \in \mathbb{R}$ is a Bessel process with parameter γ; in this case we use the notation $V_t^\gamma = H_t^{(\gamma-1)/2,1}$. Recall that for $\gamma \in \mathbb{N}$, the law of $(V_t^\gamma)_{t\in[0,1]}$ is the same as that of $\|B_t\|_{t\in[0,1]}$ where $(B_t)_{t\in[0,1]}$ is standard γ-dimensional Brownian motion. In fact, any $H_t^{\alpha,\beta}$ is related to a Bessel process via simple scaling, as the next result shows.

Lemma 7.2. *Let $\alpha \in \mathbb{R}$ and $\beta > 0$. The diffusion process $H_t^{\alpha,\beta}$ is a scaled Bessel process:*

$$(H_t^{\alpha,\beta})_{t\in[0,1]} \text{ has the same law as } (\beta^{1/2}V_t^\gamma)_{t\in[0,1]}, \quad \text{with } \gamma = 1 + \frac{2\alpha}{\beta}.$$

Proof. By the Itô transformation formula (cf p. 173 of [21]), for any $\beta > 0$ the process $(\beta^{1/2}V_t^\gamma)_{t\in[0,1]}$ is a diffusion process on $[0, 1]$ with infinitesimal mean $\mu(x) = \beta(\gamma - 1)/(2x)$ and infinitesimal variance $\sigma(x) = \beta$, from which we obtain the result. □

We need the following form of Lamperti's invariance principle [27, Thms. 2.1, 5.1, and A.2].

Proposition 7.3. *[27] Let (W_k) be an irreducible Markov chain on a countable unbounded subset of $\mathbb{R}_+$. Suppose that (31) holds, and that*

$$\lim_{x\to\infty} \mu_2(x) = \beta > 0, \qquad \lim_{x\to\infty} x\mu_1(x) = \alpha > -(\beta/2).$$

Let $(H_t^{\alpha,\beta})_{t\geq 0}$ be a diffusion process as defined at (32). Then as $k \to \infty$,

$$(k^{-1/2} W_{kt})_{t\in[0,1]} \to (H_t^{\alpha,\beta})_{t\in[0,1]}$$

in the sense of convergence of finite-dimensional distributions. Marginally,

$$\lim_{k\to\infty} \mathbb{P}(k^{-1/2} W_k \leq y) = \frac{2}{(2\beta)^{(\alpha/\beta)+(1/2)}\Gamma((\alpha/\beta)+(1/2))} \int_0^y r^{2\alpha/\beta} \exp(-r^2/(2\beta)) \mathrm{d}r.$$

7.2 Increment moment estimates for W_k

Now consider the process $(W_k)_{k\in\mathbb{Z}_+}$ where $W_k = \tilde{Z}_k^{1/2}$; this is a Markov chain with a countable state space (since $\tilde{Z}_k$ is), so fits into the framework described in Section 7.1 above. Lemma 7.7 below shows that indeed W_k is an instance of Lamperti's problem in the critical regime. First we need some simple properties of the random variable κ.

Lemma 7.4. *If κ satisfies (5) for $\lambda > 0$, then*

$$\mathbb{P}(|\kappa| \geq x) \leq \exp\{-\lambda x\}, \quad (x \geq 0), \quad \text{and} \tag{33}$$

$$\mathbb{E}[|\kappa|^r] < \infty, \quad (r \geq 0). \tag{34}$$

Proof. (33) is immediate from Markov's inequality and (5), and (34) is also straightforward. □

Now we can start our analysis of the noisy urn and the associated process $\tilde{Z}_k$. Recall that $\tilde{Z}_k$ is defined as $\max\{|\tilde{X}_{\tilde{\nu}_k+1}|, |\tilde{Y}_{\tilde{\nu}_k+1}|\}$. By definition of the noisy urn process, if we start at unit distance away from an axis (in the anticlockwise sense), the path of the noisy urn until it hits the next axis has the same distribution as the corresponding path in the original simple harmonic urn. Since we refer to this fact often, we state it as a lemma.

Lemma 7.5. *Given $\tilde{Z}_k = z$, the path $(\tilde{X}_n, \tilde{Y}_n)$ for $n = \tilde{\nu}_k+1, \ldots, \tilde{\nu}_{k+1}$ has the same distribution as the path (X_n, Y_n) for $n = \nu_k + 1, \ldots, \nu_{k+1}$ given $Z_k = z$. In particular, $\tilde{Z}_{k+1}$ conditioned on $\tilde{Z}_k = z$ has the same distribution as $Z_{k+1} - \min\{\kappa, Z_{k+1} - 1\} = Z_{k+1} - \kappa + (\kappa + 1 - Z_{k+1})\mathbf{1}\{\kappa \geq Z_{k+1}\}$ conditioned on $Z_k = z$.*

Recall that $\Delta_k = Z_{k+1} - Z_k$, and write $\tilde{\Delta}_k = \tilde{Z}_{k+1} - \tilde{Z}_k$. The next result is an analogue of Lemmas 6.3 and 6.4 for $\tilde{\Delta}_k$.

Lemma 7.6. *Suppose that (5) holds. Let $\varepsilon > 0$. Then for some $C < \infty$ and all $n \in \mathbb{N}$,*

$$\mathbb{P}(|\tilde{\Delta}_k| > n^{(1/2)+\varepsilon} \mid \tilde{Z}_k = n) \leq C \exp\{-n^{\varepsilon/3}\}. \tag{35}$$

Also, for any $r \in \mathbb{N}$, there exists $C < \infty$ such that for any $n \in \mathbb{N}$,

$$\mathbb{E}[|\tilde{\Delta}_k|^r \mid \tilde{Z}_k = n] \leq C n^{r/2}. \tag{36}$$

Moreover, there exists $\gamma > 0$ for which, as $n \to \infty$,

$$\mathbb{E}[\tilde{\Delta}_k \mid \tilde{Z}_k = n] = \frac{2}{3} - \mathbb{E}[\kappa] + O(\mathrm{e}^{-\gamma n}), \tag{37}$$

$$\mathbb{E}[\tilde{\Delta}_k^2 \mid \tilde{Z}_k = n] = \frac{2}{3} n + O(1). \tag{38}$$

Proof. By the final statement in Lemma 7.5, for any $r \geq 0$,

$$\mathbb{P}(|\tilde{\Delta}_k| > r \mid \tilde{Z}_k = n) \leq \mathbb{P}(|\Delta_k - \kappa| > r \mid Z_k = n).$$

We have for any $\varepsilon > 0$,

$$\mathbb{P}(|\Delta_k - \kappa| > n^{(1/2)+\varepsilon} \mid Z_k = n) \leq \mathbb{P}(|\Delta_k| > n^{(1+\varepsilon)/2} \mid Z_k = n) + \mathbb{P}(|\kappa| > n^{(1+\varepsilon)/2}),$$

for all n large enough. Using the bounds in (24) and (33) we obtain (35). For $r \in \mathbb{N}$,

$$\mathbb{E}[|\tilde{\Delta}_k|^r \mid \tilde{Z}_k = n] \leq \mathbb{E}[(|\Delta_k| + |\kappa|)^r \mid Z_k = n].$$

Then with Minkowski's inequality, (25) and (34) we obtain (36).

Next we have from Lemma 7.5 and (27) that

$$\begin{aligned}
\mathbb{E}[\tilde{\Delta}_k \mid \tilde{Z}_k = n] &= \mathbb{E}[\Delta_k - \kappa + (\kappa + 1 - Z_{k+1})\mathbf{1}\{\kappa \geq Z_{k+1}\}\} \mid Z_k = n] \\
&= \frac{2}{3} + O(\mathrm{e}^{\alpha_1 n}) - \mathbb{E}[\kappa] + \mathbb{E}[(\kappa + 1 - Z_{k+1})\mathbf{1}\{\kappa \geq Z_{k+1}\} \mid Z_k = n].
\end{aligned}$$

By the Cauchy–Schwarz inequality, (34) and the bound $0 \leq \kappa + 1 - Z_{k+1} \leq \kappa$, the last term here is bounded by a constant times the square-root of

$$\mathbb{P}(\kappa \geq Z_{k+1} \mid Z_k = n) \leq \mathbb{P}(|\Delta_k| \geq n/2 \mid Z_k = n) + \mathbb{P}(|\kappa| > n/2) = O(\exp\{-\lambda n/2\}),$$

using the bounds (24) and (33). Hence we obtain (37). Similarly, from (28), we obtain (38). □

Now we can give the main result of this section on the increments of the process $(W_k)_{k \in \mathbb{Z}_+}$.

Lemma 7.7. *Suppose that κ satisfies (5). With $\mu_r(x)$ as defined by (30), we have*

$$\sup_x |\mu_r(x)| < \infty, \tag{39}$$

for each $r \in \mathbb{N}$. Moreover as $x \to \infty$,

$$\mu_1(x) = \frac{1 - 2\mathbb{E}[\kappa]}{4x} + O(x^{-2}); \quad \mu_2(x) = \frac{1}{6} + O(x^{-1}). \tag{40}$$

Proof. For the duration of this proof we write $\mathbb{E}_{x^2}[\,\cdot\,]$ for $\mathbb{E}[\,\cdot \mid \tilde{Z}_k = x^2] = \mathbb{E}[\,\cdot \mid W_k = x]$. For $r \in \mathbb{N}$ and $x \geq 0$, from (30),

$$|\mu_r(x)| \leq \mathbb{E}_{x^2}[|\tilde{Z}_{k+1}^{1/2} - \tilde{Z}_k^{1/2}|^r] = x^r \mathbb{E}_{x^2}\left[|(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1|^r\right]. \tag{41}$$

Fix $\varepsilon > 0$ and write $A(n) := \{|\tilde{\Delta}_k| > n^{(1/2)+\varepsilon}\}$ and $A^c(n)$ for the complementary event. Now for some $C < \infty$ and all $x \geq 1$, by Taylor's theorem,

$$|(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1|^r \mathbf{1}_{A^c(x^2)} \leq C x^{-2r} |\tilde{\Delta}_k|^r \mathbf{1}_{A^c(x^2)}.$$

Hence

$$\mathbb{E}_{x^2}\left[|(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1|^r \mathbf{1}_{A^c(x^2)}\right] \leq C x^{-2r} \mathbb{E}_{x^2}[|\tilde{\Delta}_k|^r] = O(x^{-r}), \tag{42}$$

by (36). On the other hand, using the fact that for $y \geq -1$, $0 \leq (1+y)^{1/2} \leq 1 + (y/2)$, we have

$$\begin{aligned}
\mathbb{E}_{x^2}\left[|(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1|^r \mathbf{1}_{A(x^2)}\right] &\leq \mathbb{E}_{x^2}\left[\left(1 + (1/2)x^{-2}|\tilde{\Delta}_k|\right)^r \mathbf{1}_{A(x^2)}\right] \\
&\leq \left(\mathbb{E}_{x^2}[(1 + |\tilde{\Delta}_k|)^{2r}]\right)^{1/2} \left(\mathbb{P}(A(x^2) \mid \tilde{Z}_k = x^2)\right)^{1/2},
\end{aligned}$$

for $x \geq 1$, by Cauchy–Schwarz. Using (36) to bound the expectation here and (35) to bound the probability, we obtain, for any $r \in \mathbb{N}$,

$$\mathbb{E}_{x^2}\left[|(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1|^r \mathbf{1}_{A(x^2)}\right] = O(\exp\{-x^{\varepsilon/2}\}). \tag{43}$$

Combining (42) and (43) with (41) we obtain (39).

Now we prove (40). We have that for $x \geq 0$,

$$\begin{aligned} \mu_1(x) &= \mathbb{E}_{x^2}[W_{k+1} - W_k] = x\mathbb{E}_{x^2}[(1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1] \\ &= x\mathbb{E}_{x^2}[((1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1)\mathbf{1}_{A^c(x^2)}] + O(\exp\{-x^{\varepsilon/3}\}), \end{aligned} \tag{44}$$

using (43). By Taylor's theorem with Lagrange form for the remainder we have

$$\begin{aligned} & x\mathbb{E}_{x^2}[((1 + x^{-2}\tilde{\Delta}_k)^{1/2} - 1)\mathbf{1}_{A^c(x^2)}] \\ &= \frac{1}{2x}\mathbb{E}_{x^2}[\tilde{\Delta}_k \mathbf{1}_{A^c(x^2)}] - \frac{1}{8x^3}\mathbb{E}_{x^2}[\tilde{\Delta}_k^2 \mathbf{1}_{A^c(x^2)}] + O(x^{-5}\mathbb{E}_{x^2}[|\tilde{\Delta}_k|^3]). \end{aligned} \tag{45}$$

Here we have that $x^{-5}\mathbb{E}_{x^2}[|\tilde{\Delta}_k|^3] = O(x^{-2})$, by (36), while for $r \in \mathbb{N}$, we obtain

$$\mathbb{E}_{x^2}[\tilde{\Delta}_k^r \mathbf{1}_{A^c(x^2)}] = \mathbb{E}_{x^2}[\tilde{\Delta}_k^r] + O((\mathbb{E}_{x^2}[|\tilde{\Delta}_k|^{2r}]\mathbb{P}(A(x^2) \mid \tilde{Z}_k = x^2))^{1/2}),$$

by Cauchy–Schwarz. Using (35) again and combining (44) with (45) we obtain

$$\mu_1(x) = \frac{1}{2x}\mathbb{E}_{x^2}[\tilde{\Delta}_k] - \frac{1}{8x^3}\mathbb{E}_{x^2}[\tilde{\Delta}_k^2] + O(x^{-2}).$$

Thus from (37) and (38) we obtain the expression for μ_1 in (40). Now we use the fact that

$$(W_{k+1} - W_k)^2 = W_{k+1}^2 - W_k^2 - 2W_k(W_{k+1} - W_k) = \tilde{Z}_{k+1} - \tilde{Z}_k - 2W_k(W_{k+1} - W_k)$$

to obtain $\mu_2(x) = \mathbb{E}_{x^2}[\tilde{\Delta}_k] - 2x\mu_1(x)$, which with (37) yields the expression for μ_2 in (40). □

8 Proofs of theorems

8.1 Proofs of Theorems 2.3, 2.4, 2.5, and 2.8

First we work with the noisy urn model of Section 2.3. Given the moment estimates of Lemma 7.7, we can now apply the general results described in Section 7.1 and [3].

Proof of Theorem 2.4. First observe that $(\tilde{Z}_k)_{k\in\mathbb{Z}_+}$ is transient, null-, or positive-recurrent exactly when $(W_k)_{k\in\mathbb{Z}_+}$ is. From Lemma 7.7, we have that

$$2x\mu_1(x) - \mu_2(x) = \frac{1}{3} - \mathbb{E}[\kappa] + O(x^{-1}); \quad 2x\mu_1(x) + \mu_2(x) = \frac{2}{3} - \mathbb{E}[\kappa] + O(x^{-1}).$$

Now apply Proposition 7.1. □

Proof of Theorem 2.5. By the definition of τ_q at (6), τ_q is also the first hitting time of 1 by $(W_k)_{k\in\mathbb{N}}$. Then with Lemma 7.7 we can apply results of Aspandiiarov, Iasnogorodski and Menshikov [3, Props. 1 and 2], which generalize those of Lamperti [28] and give conditions on μ_1 and μ_2 for existence and non-existence of passage-time moments, to obtain the stated result. □

Proof of Theorem 2.8. First Proposition 7.3 and Lemma 7.7 imply that, as $n \to \infty$,

$$(n^{-1/2}W_{nt})_{t\in[0,1]} \to (H_t^{\alpha,\beta})_{t\in[0,1]},$$

in the sense of finite-dimensional distributions, where $\alpha = (1 - 2\mathbb{E}[\kappa])/4$ and $\beta = 1/6$, provided $\mathbb{E}[\kappa] < 2/3$. By the Itô transformation formula (cf p. 173 of [21]), with $H_t^{\alpha,\beta}$ as defined at (32), $(H_t^{\alpha,\beta})^2$ is a diffusion process with infinitesimal mean $\mu(x) = \beta + 2\alpha$ and infinitesimal variance $\sigma^2(x) = 4\beta x$. In particular $(H_t^{\alpha,\beta})^2$ has the same law as the process denoted D_t in the statement of Theorem 2.8. Convergence of finite-dimensional distributions for $(n^{-1}W_{nt}^2)_{t\in[0,1]} = (n^{-1}\tilde{Z}_{nt})_{t\in[0,1]}$ follows. The final statement in the theorem follows from Lemma 7.2. □

Next consider the leaky urn model of Section 2.2.

Proof of Theorem 2.3. This is an immediate consequence of the $\mathbb{P}(\kappa = 1) = 1$ cases of Theorems 2.4 and 2.6. □

Remark. There is a short proof of the first part of Theorem 2.3 due to the existence of a particular martingale. Consider the process Q'_n defined by $Q'_n = Q(X'_n, Y'_n)$ where

$$Q(x,y) := \left(x + \frac{1}{2}\operatorname{sgn}(y) - \frac{1}{2}\mathbf{1}_{\{y=0\}}\operatorname{sgn}(x)\right)^2 + \left(y - \frac{1}{2}\operatorname{sgn}(x) - \frac{1}{2}\mathbf{1}_{\{x=0\}}\operatorname{sgn}(y)\right)^2 .$$

It turns out that Q'_n is a (non-negative) martingale. Thus it converges a.s. as $n \to \infty$. But since $Q(x,y) \to \infty$ as $\|(x,y)\| \to \infty$, we must have that eventually (X'_n, Y'_n) gets trapped in the closed class $\mathcal{C}$. So $\mathbb{P}(\tau < \infty) = 1$.

8.2 Proofs of Theorems 2.6 and 2.9

The proofs of Theorems 2.6 and 2.9 that we give in this section both rely on the good estimates we have for the embedded process $\tilde{Z}_k$ to analyse the noisy urn $(\tilde{X}_n, \tilde{Y}_n)$. The main additional ingredient is to relate the two different time-scales. The first result concerns the time to traverse a quadrant.

Lemma 8.1. *Let $k \in \mathbb{Z}_+$. The distribution of $\tilde{\nu}_{k+1} - \tilde{\nu}_k$ given $\tilde{Z}_k = n$ coincides with that of $Z_{k+1} + Z_k$ given $Z_k = n$. In addition*

$$\tilde{\nu}_{k+1} - \tilde{\nu}_k = |\tilde{X}_{\tilde{\nu}_{k+1}}| + |\tilde{Y}_{\tilde{\nu}_{k+1}}| + \tilde{Z}_k. \tag{46}$$

Moreover

$$\mathbb{P}(\tilde{\nu}_{k+1} - \tilde{\nu}_k > 3n \mid \tilde{Z}_k = n) = O(\exp\{-n^{1/2}\}). \tag{47}$$

Proof. Without loss of generality, suppose that we are traversing the first quadrant. Starting at time $\tilde{\nu}_k + 1$, Lemma 7.5 implies that the time until hitting the next axis, $\tilde{\nu}_{k+1} - \tilde{\nu}_k - 1$, has the same distribution as the time taken for the original simple harmonic urn to hit the next axis, starting from $(Z_k, 1)$. In this time, the simple harmonic urn must make Z_k horizontal jumps and $Z_{k+1} - 1$ vertical jumps. Thus $\tilde{\nu}_{k+1} - \tilde{\nu}_k - 1$ has the same distribution as $Z_{k+1} + Z_k - 1$, conditional on $Z_k = \tilde{Z}_k$. Thus we obtain the first statement in the lemma. For equation (46), note that between times $\tilde{\nu}_k + 1$ and $\tilde{\nu}_{k+1}$ the noisy urn must make $\tilde{Z}_k$ horizontal steps and (in this case) $|\tilde{Y}_{\tilde{\nu}_{k+1}}| - 1$ vertical steps. Finally we have from the first statement of the lemma that

$$\mathbb{P}(\tilde{\nu}_{k+1} - \tilde{\nu}_k > 3n \mid \tilde{Z}_k = n) = \mathbb{P}(Z_{k+1} > 2n \mid Z_k = n),$$

and then (47) follows from (24). □

Very roughly speaking, the key to our Theorems 2.6 and 2.9 is the fact that $\tau \approx \sum_{k=0}^{\tau_q} W_k^2$ and $A \approx \sum_{k=0}^{\tau_q} W_k^4$. Thus to study τ and A we need to look at sums of powers of W_k over a *single excursion*. First we will give results for $S_\alpha := \sum_{k=0}^{\tau_q} W_k^\alpha$, $\alpha \geq 0$. Then we quantify the approximations '≈' for τ and A by a series of bounds.

Let $M := \max_{0\leq k\leq \tau_q} W_k$ denote the maximum of the first excursion of W_k. For ease of notation, for the rest of this section we set $r := 6\mathbb{E}[\kappa] - 3$.

Lemma 8.2. *Suppose that $r > -1$. Then for any $\varepsilon > 0$, for all x sufficiently large*

$$x^{-1-r-\varepsilon} \le \mathbb{P}(M \ge x) \le x^{-1-r+\varepsilon}.$$

In particular, for any $\varepsilon > 0$, $\mathbb{E}[M^{1+r+\varepsilon}] = \infty$ but $\mathbb{E}[M^{1+r-\varepsilon}] < \infty$.

Proof. It follows from Lemma 7.7 and some routine Taylor's theorem computations that for any $\varepsilon > 0$ there exists $w_0 \in [1, \infty)$ such that for any $x \ge w_0$,

$$\mathbb{E}[W_{k+1}^{1+r+\varepsilon} - W_k^{1+r+\varepsilon} \mid W_k = x] \ge 0,$$
$$\mathbb{E}[W_{k+1}^{1+r-\varepsilon} - W_k^{1+r-\varepsilon} \mid W_k = x] \le 0.$$

Let $\eta := \min\{k \in \mathbb{Z}_+ : W_k \le w_0\}$ and $\sigma_x := \min\{k \in \mathbb{Z}_+ : W_k \ge x\}$. Recall that $(W_k)_{k\in\mathbb{Z}_+}$ is an irreducible time-homogeneous Markov chain on a countable subset of $[1, \infty)$. It follows that to prove the lemma it suffices to show that, for some $w \ge 2w_0$, for any $\varepsilon > 0$,

$$x^{-1-r-\varepsilon} \le \mathbb{P}(\sigma_x < \eta \mid W_0 = w) \le x^{-1-r+\varepsilon}, \tag{48}$$

for all x large enough.

We first prove the lower bound in (48). Fix $x > w$. We have that $W_{k\wedge\eta\wedge\sigma_x}^{1+r+\varepsilon}$ is a submartingale, and, since W_k is an irreducible Markov chain, $\eta < \infty$ and $\sigma_x < \infty$ a.s.. Hence

$$\mathbb{P}(\sigma_x < \eta)\mathbb{E}[W_{\sigma_x}^{1+r+\varepsilon}] + (1 - \mathbb{P}(\sigma_x < \eta))\mathbb{E}[W_\eta^{1+r+\varepsilon}] \ge w^{1+r+\varepsilon}.$$

Here $W_\eta \le w_0$ a.s., and for some $C \in (0, \infty)$ and all $x > w$,

$$\mathbb{E}[W_{\sigma_x}^{1+r+\varepsilon}] \le \mathbb{E}\left[(x + (W_{\sigma_x} - W_{\sigma_x-1}))^{1+r+\varepsilon}\right] \le Cx^{1+r+\varepsilon},$$

since $\mathbb{E}[(W_{\sigma_x} - W_{\sigma_x-1})^{1+r+\varepsilon}]$ is uniformly bounded in x, by equation (39). It follows that

$$\mathbb{P}(\sigma_x < \eta)\left[Cx^{1+r+\varepsilon} - w_0^{1+r+\varepsilon}\right] \ge w^{1+r+\varepsilon} - w_0^{1+r+\varepsilon} > 0,$$

which yields the lower bound in (48). The upper bound follows by a similar argument based on the supermartingale property of $W_{k\wedge\eta\wedge\sigma_x}^{1+r-\varepsilon}$. □

The next result gives the desired moment bounds for S_α.

Lemma 8.3. *Let $\alpha \ge 0$ and $r > -1$. Then $\mathbb{E}[S_\alpha^p] < \infty$ if $p < \frac{1+r}{\alpha+2}$ and $\mathbb{E}[S_\alpha^p] = \infty$ if $p > \frac{1+r}{\alpha+2}$.*

Proof. First we prove the upper bound. Clearly $S_\alpha \le (1+\tau_q)M^\alpha$. Then, by Hölder's inequality,

$$\mathbb{E}[S_\alpha^p] \le (\mathbb{E}[(1+\tau_q)^{(2+\alpha)p/2}])^{\frac{2}{2+\alpha}}(\mathbb{E}[M^{(2+\alpha)p}])^{\frac{\alpha}{2+\alpha}}.$$

For $p < \frac{1+r}{\alpha+2}$ we have $(2+\alpha)p/2 < (1+r)/2 = 3\mathbb{E}[\kappa] - 1$ and $(2+\alpha)p < 1+r$ so that Lemma 8.2 and Theorem 2.5 give the upper bound.

For the lower bound we claim that there exists $C \in (0, \infty)$ such that

$$\mathbb{P}(S_\alpha \ge x) \ge \frac{1}{2}\mathbb{P}(M \ge Cx^{\frac{1}{\alpha+2}}), \tag{49}$$

for all x large enough. Given the claim (49), we have, for any $\varepsilon > 0$,

$$\mathbb{E}[S_\alpha^p] \ge \frac{p}{2}\int_1^\infty x^{p-1}\mathbb{P}(M \ge Cx^{\frac{1}{\alpha+2}})\mathrm{d}x \ge \frac{p}{2}\int_1^\infty x^{p-1}x^{-\frac{1+r}{\alpha+2}-\varepsilon}\mathrm{d}x,$$

by Lemma 8.2. Thus $\mathbb{E}[S_\alpha^p] = \infty$ for $p > \frac{1+r}{\alpha+2}$. It remains to verify (49). Fix $y > 2$. Let $\mathcal{F}_k = \sigma(W_1, \dots, W_k)$, and define stopping times

$$\sigma_1 = \min\{k \in \mathbb{N} : W_k \geq y\}; \quad \sigma_2 = \min\{k \geq \sigma_1 : W_k \leq y/2\}.$$

Then $\{\sigma_1 < \tau_q\}$, i.e., the event that W_k reaches y before 1, is $\mathcal{F}_{\sigma_1}$-measurable. Now

$$\mathbb{P}\left(\{\sigma_1 < \tau_q\} \cap \{\sigma_2 \geq \sigma_1 + \delta y^2\}\right) = \mathbb{E}\left[\mathbf{1}\{\sigma_1 < \tau_q\}\mathbb{P}(\sigma_2 \geq \sigma_1 + \delta y^2 \mid \mathcal{F}_{\sigma_1})\right]. \tag{50}$$

We claim that there exists $\delta > 0$ so that

$$\mathbb{P}(\sigma_2 \geq \sigma_1 + \delta y^2 \mid \mathcal{F}_{\sigma_1}) \geq \frac{1}{2}, \text{ a.s.} \tag{51}$$

Let $D_k = (y - W_k)^2 \mathbf{1}\{W_k < y\}$. Then, with $\Delta_k = W_{k+1} - W_k$,

$$\mathbb{E}[D_{k+1} - D_k \mid \mathcal{F}_k] \leq 2(W_k - y)\mathbb{E}[\Delta_k \mid \mathcal{F}_k] + \mathbb{E}[\Delta_k^2 \mid \mathcal{F}_t].$$

Lemma 7.7 implies that on $\{W_k > y/2\}$ this last display is bounded above by some $C < \infty$ not depending on y. Hence an appropriate maximal inequality [32, Lemma 3.1] implies (since $D_{\sigma_1} = 0$) that $\mathbb{P}(\max_{0 \leq s \leq k} D_{(\sigma_1+s)\wedge\sigma_2} \geq w) \leq Ck/w$. Then, since $D_{\sigma_2} \geq y^2/4$, we have

$$\mathbb{P}(\sigma_2 \leq \sigma_1 + \delta y^2 \mid \mathcal{F}_{\sigma_1}) \leq \mathbb{P}\left(\max_{1 \leq s \leq \delta y^2} D_{(\sigma_1+s)\wedge\sigma_2} \geq (y^2/4) \mid \mathcal{F}_{\sigma_1}\right) \leq \frac{C\delta y^2}{(y^2/4)} \leq \frac{1}{2}, \text{ a.s.,}$$

for $\delta > 0$ small enough. Hence (51) follows. Combining (50) and (51) we get

$$\mathbb{P}\left(\{\sigma_1 < \tau_q\} \cap \{\sigma_2 \geq \sigma_1 + \delta y^2\}\right) \geq \frac{1}{2}\mathbb{P}(\sigma_1 < \tau_q) = \frac{1}{2}\mathbb{P}(M \geq y).$$

Moreover, on $\{\sigma_1 < \tau_q\} \cap \{\sigma_2 \geq \sigma_1 + \delta y^2\}$ we have that $W_s \geq y/2$ for all $\sigma_1 \leq s < \sigma_2$, of which there are at least δy^2 values; hence $S_\alpha \geq \delta y^2 \times (y/2)^\alpha$. Now taking $x = 2^{-\alpha}\delta y^{2+\alpha}$ we obtain (49), and so complete the proof. □

Next we need a technical lemma.

Lemma 8.4. *Let $p \geq 0$. Then for any $\varepsilon > 0$ there exists $C < \infty$ such that*

$$\mathbb{E}\left[\left(\sum_{k=1}^{\tau_q} |\kappa_{\tilde{\nu}_k}|\right)^p\right] \leq C\mathbb{E}[\tau_q^{p+\varepsilon}]. \tag{52}$$

Proof. For any $s \in (0, 1)$,

$$\mathbb{P}\left(\sum_{k=1}^{\tau_q} |\kappa_{\tilde{\nu}_k}| > x\right) \leq \mathbb{P}(\tau_q > x^s) + \mathbb{P}\left(\sum_{k=1}^{x^s} |\kappa_{\tilde{\nu}_k}| > x\right).$$

For any random variable X, $\mathbb{E}[X^p] = p\int_0^\infty x^{p-1}\mathbb{P}(X > x)\mathrm{d}x \leq 1 + p\int_1^\infty x^{p-1}\mathbb{P}(X > x)\mathrm{d}x$; so

$$\mathbb{E}\left[\left(\sum_{k=1}^{\tau_q} |\kappa_{\tilde{\nu}_k}|\right)^p\right] \leq 1 + p\int_1^\infty x^{p-1}\mathbb{P}(\tau_q > x^s)\mathrm{d}x + p\int_1^\infty x^{p-1}\mathbb{P}\left(\sum_{k=1}^{x^s} |\kappa_{\tilde{\nu}_k}| > x\right)\mathrm{d}x. \tag{53}$$

Here we have that

$$\mathbb{P}\left(\sum_{k=1}^{x^s} |\kappa_{\tilde{\nu}_k}| > x\right) \leq \mathbb{P}\left(\bigcup_{k=1}^{x^s} \{|\kappa_{\tilde{\nu}_k}| > x^{1-s}\}\right) \leq \sum_{k=1}^{x^s} \mathbb{P}(|\kappa| > x^{1-s}),$$

by Boole's inequality. Then Markov's inequality and the moment bound (5) yield

$$\mathbb{P}\left(\sum_{k=1}^{x^s}|\kappa_{\tilde{\nu}_k}|>x\right)\leq x^s\mathbb{P}(\mathrm{e}^{\lambda|\kappa|}>\mathrm{e}^{x^{1-s}})\leq x^s\mathbb{E}[\mathrm{e}^{\lambda|\kappa|}]\mathrm{e}^{-x^{1-s}}. \tag{54}$$

It follows that, since $s<1$, the final integral in (53) is finite for any p. Also, from Markov's inequality, for any $\varepsilon>0$,

$$\int_1^\infty x^{p-1}\mathbb{P}(\tau_q>x^s)\mathrm{d}x\leq\mathbb{E}[\tau_q^{p+\varepsilon}]\int_1^\infty x^{p-1-s(p+\varepsilon)}\mathrm{d}x;$$

taking s close to 1 this last integral is finite, and (52) follows (noting $\tau_q\geq 1$ by definition). □

Proof of Theorem 2.6. By the definitions of τ and τ_q we have that $\tau=\tilde{\nu}_{\tau_q}=-1+\sum_{k=1}^{\tau_q}(\tilde{\nu}_k-\tilde{\nu}_{k-1})$, recalling $\tilde{\nu}_0=-1$. Hence by Lemma 8.1,

$$\tau=-1+\sum_{k=1}^{\tau_q}(W_{k-1}^2+W_k^2)+R,$$

for R a random variable such that $|R|\leq\sum_{k=1}^{\tau_q}|\kappa_{\tilde{\nu}_k}|$. It follows that

$$-1+\sum_{k=0}^{\tau_q}W_k^2-|R|\leq\tau\leq 2\sum_{k=0}^{\tau_q}W_k^2+|R|. \tag{55}$$

Lemma 8.4 implies that for any $\varepsilon>0$ there exists $C<\infty$ such that $\mathbb{E}[|R|^p]\leq C\mathbb{E}[\tau_q^{p+\varepsilon}]$. The $\mathbb{E}[\kappa]>1/3$ case of the theorem now follows from (55) with Theorem 2.5, Lemma 8.3 and Minkowski's inequality. In the $\mathbb{E}[\kappa]=1/3$ case, it is required to prove that $\mathbb{E}[\tau^p]=\infty$ for any $p>0$; this follows from the $\mathbb{E}[\kappa]=1/3$ case of Theorem 2.5 and the fact that $\tau\geq\tau_q$ a.s.. □

Proof of Theorem 2.9. First note that we can write

$$A=\sum_{n=1}^{\tau}T_n=\sum_{n=1}^{\tilde{\nu}_{\tau_q}}T_n=\sum_{k=1}^{\tau_q}A_k,$$

where $A_1=\sum_{n=1}^{\tilde{\nu}_1}T_n$ and $A_k=\sum_{n=\tilde{\nu}_{k-1}+1}^{\tilde{\nu}_k}T_n$ ($k\geq 2$) is the area swept out in traversing a quadrant for the kth time. Since $A_k\geq 1/2$, part (i) of the theorem is immediate from part (i) of Theorem 2.4. For part (ii), we have that

$$\begin{aligned}A_k&\leq(\tilde{Z}_k+|\kappa_{\tilde{\nu}_k}|)(\tilde{Z}_{k-1}+|\kappa_{\tilde{\nu}_{k-1}}|)\\&\leq W_k^4+W_{k-1}^4+W_{k-1}^2|\kappa_{\tilde{\nu}_k}|+W_k^2|\kappa_{\tilde{\nu}_{k-1}}|+|\kappa_{\tilde{\nu}_{k-1}}||\kappa_{\tilde{\nu}_k}|.\end{aligned}$$

Thus

$$A\leq 2\sum_{k=0}^{\tau_q}W_k^4+R_1+R_2+R_3,$$

where $R_1=\sum_{k=1}^{\tau_q}W_{k-1}^2|\kappa_{\tilde{\nu}_k}|$, $R_2=\sum_{k=1}^{\tau_q}W_k^2|\kappa_{\tilde{\nu}_{k-1}}|$ and $R_3=\sum_{k=1}^{\tau_q}|\kappa_{\tilde{\nu}_{k-1}}||\kappa_{\tilde{\nu}_k}|$. Here $\sum_{k=0}^{\tau_q}W_k^4$ has finite pth moment for $p<\frac{3\mathbb{E}[\kappa]-1}{3}$, by Lemma 8.3. Next we deal with the terms R_1,R_2 and R_3. Consider R_1. We have that, by Hölder's inequality, $\mathbb{E}[|R_1|^p]$ is at most

$$\mathbb{E}\left[\left(\sum_{k=1}^{\tau_q}W_{k-1}^2\right)^{3p/2}\right]^{2/3}\mathbb{E}\left[\left(\sum_{k=0}^{\tau_q}|\kappa_{\tilde{\nu}_k}|\right)^{3p}\right]^{1/3}\leq C'\mathbb{E}\left[\left(\sum_{k=0}^{\tau_q}W_k^2\right)^{3p/2}\right]^{2/3}\mathbb{E}[\tau_q^{3p+\varepsilon}]^{1/3},$$

for any $\varepsilon > 0$, by (52). Lemma 8.3 and Theorem 2.5 show that this is finite provided $p < \frac{3\mathbb{E}[\kappa]-1}{3}$ (taking ε small enough). A similar argument holds for R_2. Finally,

$$\mathbb{E}[|R_3|^p] \leq \mathbb{E}\left[\left(\sum_{k=0}^{\tau_q} |\kappa_{\tilde{\nu}_k}|\right)^{2p}\right] \leq C''\mathbb{E}[\tau_q^{2p+\varepsilon}],$$

for any $\varepsilon > 0$, by (52). For ε small enough, this is also finite when $p < \frac{3\mathbb{E}[\kappa]-1}{3}$ by Theorem 2.5. These estimates and Minkowski's inequality then complete the proof. □

8.3 Proof of Theorem 2.11

We now turn to the percolation model described in Section 2.4.

Lemma 8.5. *Let v and v' be any two vertices of G. Then with probability 1 there exists a vertex $w \in G$ such that the unique semi-infinite oriented paths in H from v and v' both pass through w.*

Proof. Without loss of generality, suppose v, v' are distinct vertices in G on the positive x-axis on the same sheet of $\mathcal{R}$. Let $Z_0 = |v| < Z'_0 = |v'|$. The two paths in H started at v and v', call them P and P' respectively, lead to instances of processes Z_k and Z'_k, each a copy of the simple harmonic urn embedded process Z_k. Until P and P' meet, the urn processes they instantiate are independent. Thus it suffices to take Z_k, Z'_k to be independent and show that they eventually cross with probability 1, so that the underlying paths must meet. To do this, we consider the process $(H_k)_{k\in\mathbb{Z}_+}$ defined by $H_k := \sqrt{Z'_k} - \sqrt{Z_k}$ and show that it is eventually less than or equal to 0.

For convenience we use the notation $W_k = (Z_k)^{1/2}$ and $W'_k = (Z'_k)^{1/2}$. Since $H_{k+1} - H_k = (W'_{k+1} - W'_k) - (W_{k+1} - W_k)$, we have that for $x < y$,

$$\mathbb{E}[H_{k+1} - H_k \mid W_k = x, W'_k = y] = \frac{1}{4y} - \frac{1}{4x} + O(x^{-2}) = -\frac{(y-x)}{4xy} + O(x^{-2}),$$

by the $\mathbb{E}[\kappa] = 0$ case of (40). Similarly,

$$\mathbb{E}[(H_{k+1} - H_k)^2 \mid W_k = x, W'_k = y] = \frac{1}{3} + O(x^{-1}),$$

from (40) again. Combining these we see that

$$\begin{aligned}2(y-x)\mathbb{E}[H_{k+1} - H_k \mid W_k = x, W'_k = y] - \mathbb{E}[(H_{k+1} - H_k)^2 \mid W_k = x, W'_k = y] \\ \leq -\frac{1}{3} + O(x^{-1}) < 0,\end{aligned}$$

for $x > C$, say. However, we know from Theorem 2.1 that W_k is transient, so in particular $W_k > C$ for all $k > T$ for some finite T. Let $\tau = \min\{k \in \mathbb{Z}_+ : H_k \leq 0\}$. Then we have that $H_k\mathbf{1}\{k < \tau\}$, $k > T$, is a process on $\mathbb{R}_+$ satisfying Lamperti's recurrence criterion (cf Proposition 7.1). Here $H_{k\wedge\tau}$ is not a Markov process but the general form of Proposition 7.1 applies (see [26, Thm. 3.2]) so we can conclude that $\mathbb{P}(\tau < \infty) = 1$. □

Lemma 8.6. *The in-graph of any individual vertex in H is almost surely finite.*

Proof. We work in the dual percolation model H'. As we have seen, the oriented paths through H' simulate the leaky simple harmonic urn via the mapping Φ. The path in H' that starts from a vertex over $(n + 1/2, 1/2)$ explores the outer boundary of the in-graph in H of a lift of the set $\{(i,0) : 1 \le i \le n\}$. The leaky urn a.s. reaches the steady state with one ball, so every oriented path in H' a.s. eventually joins the infinite path cycling immediately around the origin. It follows that the in-graph of any vertex over a co-ordinate axis is a.s. finite. For any vertex v of H, the oriented path from v a.s. contains a vertex w over an axis, and the in-graph of v is contained in the in-graph of w, so it too is a.s. finite. □

All that remains to complete the proof of Theorem 2.11 is to establish the two statements about the moments of $I(v)$. For $p < 2/3$, $\mathbb{E}[I(v)^p]$ is bounded above by $\mathbb{E}[A^p]$, where A is the area swept out by a path of the leaky simple harmonic urn, or equivalently by a path of the noisy simple harmonic urn with $\mathbb{P}(\kappa = 1) = 1$ up to the hitting time τ. $\mathbb{E}[A^p]$ is finite, by Theorem 2.9(ii). The final claim $\mathbb{E}[I(v)] = \infty$ will be proven in the next section as equation (60), using a connection with expected exit times from quadrants.

9 Continuous-time models

9.1 Expected traversal time: proof of Theorem 2.14

Proof of Lemma 2.12. A consequence of Dynkin's formula for a continuous-time Markov chain $X(t)$ on a countable state-space S with infinitesimal (generator) matrix $Q = (q_{ij})$ is that for a function $g : \mathbb{R}_+ \times S \to \mathbb{R}$ with continuous time-derivative to be such that $g(t, X(t))$ is a local martingale, it suffices that

$$\frac{\partial g(t,x)}{\partial t} + Q(g(t,\cdot))(t,x) = 0, \tag{56}$$

for all $x \in S$ and $t \in \mathbb{R}_+$: see e.g. [37, p. 364]. In our case $S = \mathbb{C} \setminus \{0\}$, $X(t) = A(t) + iB(t)$, and for $z = x + iy \in \mathbb{C}$,

$$Q(f)(z) = \sum_{w \in S, w \neq z} q_{zw}[f(w) - f(z)] = |x|[f(z + \mathrm{sgn}(x)i) - f(z)] + |y|[f(z - \mathrm{sgn}(y)) - f(z)].$$

Taking $f(x+iy) = g(t, x+iy)$ to be first $x\cos t + y \sin t$ and second $y \cos t - x \sin t$ we verify the identity (56) in each case. Thus the real and imaginary parts of M_t are local martingales, and hence martingales since it is not hard to see that $\mathbb{E}|A(t) + B(t)| < \infty$. □

To prove Theorem 2.14 we need the following bound on the deviations of τ_f from $\pi/2$.

Lemma 9.1. *Suppose $\varepsilon_n > 0$ and $\varepsilon_n \to 0$ as $n \to \infty$. Let $\phi_n \in [0, \pi/2]$. Then as $n \to \infty$,*

$$\mathbb{P}\left(\left|\tau_f - \frac{\pi}{2} + \phi_n\right| \ge \varepsilon_n \mid A(0) = n\cos\phi_n, B(0) = n \sin\phi_n\right) = O(n^{-1}\varepsilon_n^{-2}),$$

uniformly in (ϕ_n).

Proof. First note that $M_0 = ne^{i\phi_n}$ and, by the martingale property,

$$\mathbb{E}[|M_t - M_0|^2] = \mathbb{E}[|M_t|^2] - |M_0|^2 = \mathbb{E}[A(t)^2 + B(t)^2] - n^2.$$

We claim that for all $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$,

$$\mathbb{E}[A(t)^2 + B(t)^2] - n^2 \le \frac{t^2}{2} + 2^{1/2}nt. \tag{57}$$

Since $M_t - M_0$ is a (complex) martingale, $|M_t - M_0|^2$ is a submartingale. Doob's maximal inequality therefore implies that, for any $r > 0$,

$$\mathbb{P}\left(\sup_{0\le s\le t} |M_s - M_0| \ge r\right) \le r^{-2}\mathbb{E}[|M_t - M_0|^2] \le 2t(t+n)r^{-2},$$

by (57). Set $t_0 = (\pi/2) - \phi_n + \theta$ for $\theta \in (0, \pi/2)$. Then on $\{t_0 < \tau_f\}$, $A(t_0) + iB(t_0)$ has argument in $[\phi_n, \pi/2]$, so that M_{t_0} has argument in $[2\phi_n - (\pi/2) - \theta, \phi_n - \theta]$. All points with argument in the latter interval are at distance at least $n \sin\theta$ from M_0. Hence on $\{t_0 < \tau_f\}$,

$$\sup_{0 \le s \le t_0} |M_s - M_0| \ge |M_{t_0} - M_0| \ge n \sin\theta.$$

It follows that for $\varepsilon_n > 0$ with $\varepsilon_n \to 0$,

$$\begin{aligned}\mathbb{P}(\tau_f > (\pi/2) - \phi_n + \varepsilon_n) &\le \mathbb{P}\left(\sup_{0 \le s \le (\pi/2) - \phi_n + \varepsilon_n} |M_s - M_0| \ge n \sin\varepsilon_n \right) \\ &= O(n^{-1} (\sin\varepsilon_n)^{-2}) = O(n^{-1} \varepsilon_n^{-2}).\end{aligned}$$

A similar argument yields the same bound for $\mathbb{P}_n(\tau_f < (\pi/2) - \phi_n - \varepsilon_n)$. It remains to prove the claim (57). First note that

$$\mathbb{E}[A(t+\Delta t)^2 + B(t+\Delta t)^2 - (A(t)^2 + B(t)^2) \mid A(t) = x, B(t) = y] = (|x| + |y|)\Delta t + O((\Delta t)^2),$$

and $(|x| + |y|)^2 \le 2(x^2 + y^2)$. Writing $g(t) = \mathbb{E}[A(t)^2 + B(t)^2]$, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) \le \sqrt{2} g(t)^{1/2},$$

with $g(0) = n^2$. Hence $g(t)^{1/2} \le n + 2^{-1/2} t$. Squaring both sides yields (57). □

A consequence of Lemma 9.1 is that τ_f has finite moments of all orders, uniformly in the initial point:

Lemma 9.2. *For any $r > 0$ there exists $C < \infty$ such that* $\max_{n \in \mathbb{N}} \mathbb{E}_n[\tau_f^r] \le C$.

Proof. By Lemma 9.1, we have that there exists $n_0 < \infty$ for which

$$\sup_{x>0, y>0 : |x+iy| \ge n_0} \mathbb{P}(\tau_f - t > 2n_0 \mid A(t) + iB(t) = x + iy) \le 1/2. \tag{58}$$

On the other hand, if $|A(t) + iB(t)| < n_0$, we have that $\tau_f - t$ is stochastically dominated by a sum of n_0 exponential random variables with mean 1. Thus by Markov's inequality, the bound (58) holds for *all* $x > 0, y > 0$. Then, for $t > 1$, by conditioning on the path of the process at times $2n_0, 4n_0, \ldots, 2n_0(t-1)$ and using the strong Markov property we have

$$\mathbb{P}_n(\tau_f > 2n_0 t) \le \prod_{j=1}^{t-1} \sup_{x_j > 0, y_j > 0} \mathbb{P}(\tau_f - 2n_0 j > 2n_0 \mid A(2n_0 j) + iB(2n_0 j) = x_j + iy_j) \le 2^{1-t},$$

by (58). Hence $\mathbb{P}_n(\tau_f > t)$ decays faster than any power of t, uniformly in n. □

Proof of Theorem 2.14. For now fix $n \in \mathbb{N}$. Suppose $A(0) = Z_0 = n$, $B(0) = 0$. Note that $A(\tau_f) = 0$, $B(\tau_f) = Z_1$. The stopping time τ_f has all moments, by Lemma 9.2, while $\mathbb{E}_n[|M_t|^2] = O(t^2)$ by (57), and $\mathbb{E}_n[|M_{\tau_f}|^2] = \mathbb{E}_n[Z_1^2] < \infty$. It follows that the real and imaginary parts of the martingale $M_{t \wedge \tau_f}$ are uniformly integrable. Hence we can apply the optional stopping theorem to any linear combination of the real and imaginary parts of $M_{t \wedge \tau_f}$ to obtain

$$\mathbb{E}_n[Z_1(\alpha \sin\tau_f + \beta \cos\tau_f)] = \alpha n$$

for any $\alpha, \beta \in \mathbb{R}$. Taking $\alpha = \cos\theta$, $\beta = \sin\theta$ this says

$$n\cos\theta = \mathbb{E}_n[Z_1 \sin(\theta + \tau_f)] = \mathbb{E}_n[(Z_1 - \mathbb{E}_n Z_1)\sin(\theta + \tau_f)] + \mathbb{E}_n[Z_1]\mathbb{E}_n[\sin(\theta + \tau_f)],$$

for any θ. By Cauchy–Schwarz, the first term on the right-hand side here is bounded in absolute value by $\sqrt{\mathbb{V}\mathrm{ar}_n(Z_1)}$, so on re-arranging we have

$$\left| \mathbb{E}_n[\sin(\theta + \tau_f)] - \frac{n \cos \theta}{\mathbb{E}_n[Z_1]} \right| \leq \frac{(\mathbb{V}\mathrm{ar}_n(Z_1))^{1/2}}{\mathbb{E}_n[Z_1]} \leq \frac{(\mathbb{E}_n[\Delta_1^2])^{1/2}}{\mathbb{E}_n[Z_1]},$$

and then using (27) and (28) we obtain, as $n \to \infty$,

$$|\mathbb{E}_n[\sin(\theta + \tau_f)] - \cos \theta| = O(n^{-1/2}), \tag{59}$$

uniformly in θ. This strongly suggests that τ_f is concentrated around $\pi/2, 5\pi/2, \ldots$. To rule out the larger values, we need to use Lemma 9.1. We proceed as follows.

Define the event $E_n := \{|\tau_f - (\pi/2)| < \varepsilon_n\}$ where $\varepsilon_n \to 0$. From the $\theta = -\pi/2$ case of (59) we have that $\mathbb{E}_n[\sin(\tau_f - (\pi/2))] = O(n^{-1/2})$. Since $\sin x = x + O(x^3)$ as $x \to 0$ we have

$$\begin{aligned} \mathbb{E}_n[\mathbf{1}_{E_n} \sin(\tau_f - (\pi/2))] &= \mathbb{E}_n[\tau_f \mathbf{1}_{E_n}] - \frac{\pi}{2} + O(\varepsilon_n^3) + O(\mathbb{P}_n(E_n^c)) \\ &= \mathbb{E}_n[\tau_f] - \frac{\pi}{2} + O(\varepsilon_n^3) + O((\mathbb{E}_n[\tau_f^r])^{1/r}(\mathbb{P}_n(E_n^c))^{1-(1/r)}), \end{aligned}$$

for any $r > 1$, by Hölder's inequality. Here $\mathbb{E}_n[\tau_f^r] = O(1)$, by Lemma 9.2, so that for any $\delta > 0$, choosing r large enough we see that the final term in the last display is $O(n^{\delta-1}\varepsilon_n^{-2})$ by Lemma 9.1. Hence for any $\delta > 0$,

$$O(n^{-1/2}) = \mathbb{E}_n[\tau_f] - \frac{\pi}{2} + O(n^{\delta-1}\varepsilon_n^{-2}) + O(\varepsilon_n^3) + \mathbb{E}_n[\mathbf{1}_{E_n^c} \sin(\tau_f - (\pi/2))],$$

and this last expectation is $O(n^{-1}\varepsilon_n^{-2})$ by Lemma 9.1 once more. Taking $\varepsilon_n = n^{-1/4}$ yields (7). Next, from the $\theta = 0$ case of (59)) we have that $\mathbb{E}_n |1 - \cos(\tau_f - (\pi/2))| = O(n^{-1/2})$. This time

$$\mathbb{E}_n [|1 - \cos(\tau_f - (\pi/2))| \mathbf{1}_{E_n}] = \mathbb{E}_n[|\tau_f - (\pi/2)|^2 \mathbf{1}_{E_n}] + O(\varepsilon_n^4).$$

Following a similar argument to that for (7), we obtain (8). □

9.2 Traversal time and area enclosed: proofs of Theorems 2.10 and 2.15

Our proofs of Theorems 2.10 and 2.15 both use the percolation model of Section 2.4.

Proof of Theorem 2.15. The asymptotic statement in the theorem is a consequence of Theorem 2.14. Thus it remains to prove the exact formula. For $x > 0$, and $y \geq 0$, let $T(x, y)$ denote $\mathbb{E}[\tau_f \mid A(0) = x, B(0) = y]$. Also, set $T(0, y) = 0$ for $y > 0$. Note that $T(n, 0) = \mathbb{E}_n[\tau_f]$. Conditioning on the first step shows that for $x > 0$ and $y \geq 0$,

$$T(x, y) = \frac{1}{x + y} + \frac{x}{x + y} T(x, y + 1) + \frac{y}{x + y} T(x - 1, y)\,.$$

For fixed x, $T(x, y) \to 0$ as $y \to \infty$. Indeed, for $y \geq 1$ the time to make x horizontal jumps is stochastically dominated by the sum of x exponential random variables with mean $1/y$.

We now consider the percolation model restricted to the first quadrant. More precisely, we consider the induced graph on the set of sites (x, y) with $x \geq 0$ and $y > 0$, on a single sheet of $\mathcal{R}$. Let $I(x, y)$ denote the expected number of sites in the in-graph of (x, y) in this restricted model. This count includes the site (x, y) itself. For $x > 0$ we also set $I(x, 0) = 0$. Considering the two possible directed edges into the site (x, y), we obtain

$$I(x, y) = 1 + \frac{y}{x + y + 1} I(x + 1, y) + \frac{x}{x + y - 1} I(x, y - 1)\,.$$

Dividing through by $(x+y)$ we have

$$\frac{I(x,y)}{x+y} \;=\; \frac{1}{x+y} + \frac{y}{x+y}\left(\frac{I(x+1,y)}{(x+1)+y}\right) + \frac{x}{x+y}\left(\frac{I(x,y-1)}{x+(y-1)}\right).$$

We now claim that for each fixed y, $I(x,y)$ is bounded as $x \to \infty$. Indeed, the number of sites in the in-graph of (x,y) is at most y plus y times the number of horizontal edges in this in-graph. The number of horizontal edges may be stochastically bounded above by the sum of y geometric random variables with mean $1/x$, so its mean tends to 0 as $x \to \infty$.

We see that $I(y,x)/(x+y)$ and $T(x,y)$ satisfy the same recurrence relation with the same boundary conditions; their difference satisfies a homogeneous recurrence relation with boundary condition 0 at $x=0$ and limit 0 as $y \to \infty$ for each fixed x. An induction with respect to x shows that the difference is identically zero. In particular, taking $x=m$ and $y=0$, for any $m \geq 1$, we find

$$I(0,m) \;=\; m\,T(m,0)\,.$$

The union of the in-graphs of the sites $(0,m)$, for $1 \leq m \leq n$, is the set of all sites (x,y) with $x \geq 0$ and $y > 0$ that lie under the oriented path of the dual percolation graph H' that starts at $(-1/2, n+1/2)$. Each of these sites lies at the centre of a unit square with vertices $(x \pm 1/2, y \pm 1/2)$, and the union of these squares is the region bounded by the dual percolation path and the lines $x = -1/2$ and $y = 1/2$. Reflecting this region in the line $y = x + 1/2$ we obtain a sample of the region bounded by a simple harmonic urn path and the co-ordinate axes. The expected number of unit squares in this region is therefore $\sum_{m=1}^{n} I(0,m)$, so we are done. □

Proof of Theorem 2.10. The argument uses a similar idea to the proof of Theorem 2.15, this time for the percolation model on the whole of $\mathcal{R}$. Choose a continuous branch of the argument function on $\mathcal{R}$. Let $I_+(v)$ denote the expected number of points w with $\arg(w) > 0$ in the in-graph of v in H, including v itself if $\arg(v) > 0$. Arguing as before, if the projection of v to $\mathbb{Z}^2$ is (x,y), then $I_+(v)$ satisfies the boundary condition $I_+(v) = 0$ for $\arg(v) \leq 0$, and the recurrence relation

$$I_+(v) = 1 + \frac{|y|}{|x+\operatorname{sgn}(y)|+|y|} I_+(v + (\operatorname{sgn}(y), 0)) + \frac{|x|}{|x|+|y-\operatorname{sgn}(x)|} I_+(v + (0, -\operatorname{sgn}(x))),$$

where on the right-hand side I_+ is evaluated at two of the neighbours of v in the graph G. Setting $J_+(v) := I_+(v)/(|x(v)|+|y(v)|)$, we have a recurrence relation for J_+:

$$J_+(v) = \frac{1}{|x|+|y|} + \frac{|y|}{|x|+|y|} J_+(v + (\operatorname{sgn}(y), 0)) + \frac{|x|}{|x|+|y|} J_+(v + (0, -\operatorname{sgn}(x)).$$

The same recurrence relation and boundary conditions hold for $T_+(\overline{v})$, where $T_+(w)$ is the expected time to hit the set $\arg z \geq 0$ in $\mathcal{R}$ in the fast embedding, starting from a vertex w. Here, $\overline{v}$ is the vertex of G at the same distance from the origin as v, satisfying $\arg(\overline{v}) = -\arg(v)$. The reasoning of the previous proof shows that $T_+(\overline{v}) = J_+(v)$ for all vertices v with $\arg v \leq \pi/2$, and the argument may be repeated on the subsequent quadrants to show by induction that $T_+(\overline{v}) = J_+(v)$ for all vertices v. We therefore have the lower bound

$$J_+(v) = T_+(\overline{v}) \geq \left\lfloor \frac{\arg(v)}{\pi/2} \right\rfloor \inf_n \mathbb{E}_n[\tau_f]\,.$$

The asymptotic expression (7), together with trivial lower bounds for small n, implies that $\inf_n \mathbb{E}_n[\tau_f] > 0$. Therefore, as v varies over the set of vertices of G with a given projection

(x, y), both $J_+(v)$ and $I_+(v)$ tend to infinity with $\arg(v)$. Note that $I_+(v)$ is a lower bound for $I(v)$, and $I(v)$ depends only on the projection (x, y). It follows that

$$\mathbb{E}[I(v)] = \infty. \tag{60}$$

Recall that the oriented path of H' starting at $(m + \frac{1}{2}, -\frac{1}{2})$ explores the outer boundary of the in-graph of the set S of vertices with $\arg(v) = 0$ and $x \leq m$, and that it can be mapped via Φ onto a path of the leaky simple harmonic urn. Let A denote the area swept out by this path up until time τ (the hitting time of $\{(x, y) : |x| + |y| = 1\}$). The mapping Φ from the vertices of G' to $\mathbb{Z}^2$ can be extended by affine interpolation to a locally area-preserving map from $\mathcal{R}$ to $\mathbb{R}^2 \setminus (0, 0)$. So A is equal to the area swept out by the dual percolation path until its projection hits the set $\{(\pm\frac{1}{2}, \pm\frac{1}{2})\}$. Since the expected number of points in the in-graph that it surrounds is infinite, we have $\mathbb{E}[A] = \infty$. □

9.3 Exact formulae for expected traversal time and enclosed area

In this section we present some explicit, if mysterious, formulae for the expected area enclosed by a quadrant-traversal of the urn process and the expected quadrant-traversal time in the fast embedding. We obtain these formulae in a similar way to our first proof of Lemma 3.3, and they are reminiscent, but more involved than, the formulae for the Eulerian numbers. There is thus some hope that the asymptotics of these formulae can be handled as in the proof of Lemma 4.5, which gives a possible approach to the resolution of Conjecture 2.13.

Lemma 9.3. $\mathbb{E}_n[\text{Area enclosed}]$ *and* $\mathbb{E}_n[\tau_f]$ *are rational polynomials of degree* n *evaluated at* e*:*

$$\mathbb{E}_n[\text{Area enclosed}] = \sum_{i=1}^{n}\sum_{x=1}^{i} \frac{i^{n-x-i} i!(-1)^{n-i}}{(n-i)!(i-x)!}\left(\mathrm{e}^i - \sum_{k=0}^{i-1}\frac{i^k}{k!}\right) \tag{61}$$

$$\mathbb{E}_n[\tau_f] = \sum_{i=1}^{n}\sum_{x=1}^{i} \frac{i^{n-x-i-1} i!(-1)^{n-i}}{(n-i)!(i-x)!}\left(\mathrm{e}^i - \sum_{k=0}^{i-1}\frac{i^k}{k!}\right). \tag{62}$$

Proof. The expected area enclosed can be obtained by summing the probabilities that each unit square of the first quadrant is enclosed. That is,

$$\mathbb{E}_n[\text{Area enclosed}] \;=\; \sum_{x=1}^{n}\sum_{y=1}^{\infty} \mathbb{P}_n((x, y) \text{ lies on or below the urn path}).$$

In terms of the slow continuous-time embedding of Section 3, (x, y) lies on or below the urn path if and only if $\sum_{j=1}^{y-1} j\zeta_j < \sum_{i=x}^{n} i\xi_i$. Let

$$R_{n,x,y} \;=\; \sum_{i=x}^{n} i\xi_i \;-\; \sum_{j=1}^{y-1} j\zeta_j\,,$$

so that

$$\mathbb{E}_n[\text{Area enclosed}] \;=\; \sum_{x=1}^{n}\sum_{y=1}^{\infty} \mathbb{P}\left(R_{n,x,y} > 0\right)\,.$$

The moment generating function of $R_{n,x,y}$ is

$$\mathbb{E}\left[\exp\left(\theta R_{n,x,y}\right)\right] \;=\; \prod_{i=x}^{n} \frac{1}{1 - i\theta} \prod_{j=1}^{y-1} \frac{1}{1 + j\theta} \;=\; \sum_{i=x}^{n} \frac{\alpha_i}{1 - i\theta} + \sum_{j=1}^{y-1} \frac{\beta_j}{1 + j\theta}\,,$$

where

$$\alpha_i = \frac{i^{n-x+y-1} i! (-1)^{(n-i)}}{(i+y-1)!(i-x)!(n-i)!} \,.$$

Now the density of $R_{n,x,y}$ at $w > 0$ is

$$\sum_{i=x}^{n} \alpha_i \frac{\exp(w/i)}{i} \,,$$

so that $\mathbb{P}\left(R_{n,x,y} > 0\right) = \sum_{i=x}^{n} \alpha_i$. Therefore

$$\mathbb{E}_n[\text{Area enclosed}] \;=\; \sum_{x=1}^{n} \sum_{y=1}^{\infty} \sum_{i=x}^{n} \frac{i^{n-x+y-1} i! (-1)^{(n-i)}}{(i+y-1)!(i-x)!(n-i)!} \,.$$

The series converges absolutely so we can rearrange to obtain (61). By the first equality in Theorem 2.15, we find that $\mathbb{E}_n[\tau_f]$ is also a rational polynomial of degree n evaluated at e. After some simplification we obtain (62). □

A remarkable simplification occurs in the derivation of (62) from (61), so it is natural to try the same step again, obtaining

$$\frac{1}{n}(\mathbb{E}_n[\tau_f] - \mathbb{E}_{n-1}[\tau_f]) = \sum_{i=1}^{n} \sum_{x=1}^{i} \frac{i^{n-x-i-2} i! (-1)^{n-i}}{(n-i)!(i-x)!} \left(\mathrm{e}^i - \sum_{k=0}^{i-1} \frac{i^k}{k!} \right) .$$

In light of Theorem 2.14 and Conjecture 2.13, we would like to prove that this expression decays exponentially as $n \to \infty$. Let us make one more observation that might be relevant to Conjecture 2.13. Define $F(i) = \sum_{x=1}^{i} \frac{i!}{(i-x)! i^x}$, which can be interpreted as the expected number of distinct balls drawn if we draw from an urn containing i distinguishable balls, with replacement, stopping when we first draw some ball for the second time. We have already seen, in equation (62), that

$$\mathbb{E}_n[\tau_f] = \sum_{i=1}^{n} F(i) \frac{(-1)^{n-i} i^{n-i-1}}{(n-i)!} \left(\mathrm{e}^i - \sum_{k=0}^{i-1} \frac{i^k}{k!} \right);$$

perhaps one could exploit the resemblance to the formula

$$\mathbb{E}_n[1/Z_{k+1}] \;=\; \sum_{i=1}^{n} \frac{(-1)^{n-i} i^{n-i-1}}{(n-i)!} \left(\mathrm{e}^i - \sum_{k=0}^{i} \frac{i^k}{k!} \right),$$

but we were unable to do so.

10 Other stochastic models related to the simple harmonic urn

10.1 A stationary model: the simple harmonic flea circus

In Section 3, we saw that the Markov chain Z_k has an infinite invariant measure $\pi(n) = n$. We can understand this in the probabilistic setting by considering the formal sum of infinitely many independent copies of the fast embedding. Here is an informal description of the model. At time 0, populate each vertex of $\mathbb{Z}^2$ with an independent Poisson-distributed number of fleas with mean 1. Each flea performs a copy of the process $(A(t), B(t))$, independently of all the other fleas. Let $N_t(m, n)$ denote the number of fleas at location (m, n) at time t.

As we make no further use of this process in this paper, we do not define it more formally. Instead we just state the following result and sketch the proof: compare the lemma in [16, §2], which the authors attribute to Doob.

Lemma 10.1. *The process $\{N_t(m,n) : m,n \in \mathbb{Z}\}$ is stationary. That is, for each fixed time $t > 0$, the array $N_t(m,n)$, $m,n \in \mathbb{Z}$ consists of independent Poisson(1) random variables. The process is reversible in the sense that the ensemble of random variables $N_t(m,n), 0 \le t \le c$ has the same law as the ensemble $N_{c-t}(m,-n)$, $0 \le t \le c$, for any $c > 0$.*

This skew-reversibility allows us to extend the stationary process to all times $t \in \mathbb{R}$.

To see that the process has stationary means, note that the expectations $\mathbb{E}[N_t(m,n)]$ satisfy a system of coupled differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[N_t(m,n)] = -(|m|+|n|)\mathbb{E}[N_t(m,n)] + |m|\mathbb{E}[N_t(m,n-\mathrm{sgn}(m))] + |n|\mathbb{E}[N_t(m+\mathrm{sgn}(n),n)],$$

the solution to which is simply $\mathbb{E}[N_t(m,n)] = 1$ for all t, m and n.

To establish the independence of the variables $N_t(a,b)$ when $t > 0$ is fixed, we use a Poisson thinning argument. That is, we construct each variable $N_0(m,n)$ as an infinite sum of independent Poisson random variables $N(m,n,a,b)$ with means

$$\mathbb{E}[N(m,n,a,b)] = \mathbb{P}((A(t),B(t)) = (a,b) \mid (A(0),B(0)) = (m,n)).$$

The variable $N(m,n,a,b)$ gives the number of fleas that start at (m,n) at time 0 and are at (a,b) at time t. Then $N_t(a,b)$ is also a sum of infinitely many independent Poisson random variables, whose means sum to 1, so it is a Poisson random variable with mean 1. Moreover, for $(a,b) \neq (a',b')$, the corresponding sets of summands are disjoint, so $N_t(a,b)$ and $N_t(a',b')$ are independent.

10.2 The Poisson earthquakes model

We saw how the percolation model of Section 2.4 gives a static grand coupling of many instances of (paths of) the simple harmonic urn. In this section we describe a model, based on 'earthquakes', that gives a dynamic grand coupling of many instances of simple harmonic urn processes with particularly interesting geometrical properties.

The earthquakes model is defined as a continuous-time Markov chain taking values in the group of area-preserving homeomorphisms of the plane, which we will write as

$$\mathfrak{S}_t : \mathbb{R}^2 \to \mathbb{R}^2\,, \quad t \in \mathbb{R}.$$

It will have the properties

- $\mathfrak{S}_0$ is the identity,
- $\mathfrak{S}_t(0,0) = (0,0)$,
- $\mathfrak{S}_t$ acts on $\mathbb{Z}^2$ as a permutation,
- $\mathfrak{S}_s \circ \mathfrak{S}_t^{-1}$ has the same distribution as $\mathfrak{S}_{s-t}$, and
- for each pair $(x_0,y_0) \neq (x_1,y_1) \in \mathbb{Z}^2$, the displacement vector

  $$\mathfrak{S}_t(x_1,y_1) - \mathfrak{S}_t(x_0,y_0)$$

  has the distribution of the continuous-time fast embedding of the simple harmonic urn, starting at $(x_1 - x_0, y_1 - y_0)$.

In order to construct $\mathfrak{S}_t$, we associate a unit-rate Poisson process to each horizontal strip $H_n := \{(x, y) \in \mathbb{R}^2 : n < y < n+1\}$, and to each vertical strip $V_n := \{(x, y) \in \mathbb{R}^2 : n < x < n+1\}$, (where n ranges over $\mathbb{Z}$). All these Poisson processes should be independent. Each Poisson process determines the sequence of times at which an *earthquake* occurs along the corresponding strip. An earthquake is a homeomorphism of the plane that translates one of the complementary half-planes of the given strip through a unit distance parallel to the strip, fixes the other complementary half-plane, and shears the strip in between them. The fixed half-plane is always the one containing the origin, and the other half-plane always moves in the anticlockwise direction relative to the origin.

Consider a point $(x_0, y_0) \in \mathbb{R}^2$. We wish to define $\mathfrak{S}_t(x_0, y_0)$ for all $t \geq 0$. We will define inductively a sequence of stopping times ε_i, and points $(x_i, y_i) \in \mathbb{R}^2$, for $i \in \mathbb{Z}_+$. First, set $\varepsilon_0 = 0$. For $i \in \mathbb{N}$, suppose we have defined (x_{i-1}, y_{i-1}) and ε_{i-1}. Let ε_i be the least point greater than ε_{i-1} in the union of the Poisson processes associated to those strips for which (x_{i-1}, y_{i-1}) and $(0,0)$ do not both lie in one or other complementary half-plane. This is a.s. well-defined since there are only finitely many such strips, and a.s. there is only one strip for which an earthquake occurs at time ε_i. That earthquake moves (x_{i-1}, y_{i-1}) to (x_i, y_i). Note that $\varepsilon_i - \varepsilon_{i-1}$ is an exponential random variable with mean $1/(\lceil |x_{i-1}| \rceil + \lceil |y_{i-1}| \rceil)$, conditionally independent of all previous jumps, given this mean. Since each earthquake increases the distance between any two points by at most 1, it follows that a.s. the process does not explode in finite time. That is, $\varepsilon_i \to \infty$ as $i \to \infty$. Define $\mathfrak{S}_t(x_0, y_0)$ to be (x_i, y_i), where $\varepsilon_i \leq t < \varepsilon_{i+1}$. The construction of $\mathfrak{S}_t$ for $t < 0$ is similar, using the inverses of the earthquakes.

Note that we cannot simply define $\mathfrak{S}_t$ for $t > 0$ to be the composition of all the earthquakes that occur between times 0 and t, because almost surely infinitely many earthquakes occur during this time; however any bounded subset of the plane will only be affected by finitely many of these, so the composition makes sense locally.

The properties listed above follow directly from the construction. For $(x_0, y_0), (x_1, y_1) \in \mathbb{Z}^2$, the displacement vector $(\Delta x_t, \Delta y_t) = \mathfrak{S}_t(x_1, y_1) - \mathfrak{S}_t(x_0, y_0)$ only changes when an earthquake occurs along a strip that separates the two endpoints; the waiting time after t for this to occur is exponentially distributed with mean $1/(|\Delta x_t| + |\Delta y_t|)$, and conditionally independent of $\mathfrak{S}_t$ given $(\Delta x_t, \Delta y_t)$.

The model is spatially homogeneous in the following sense. Fix some $(a, b) \in \mathbb{Z}^2$ and define

$$\tilde{\mathfrak{S}}_t(x, y) \;=\; \mathfrak{S}_t(x+a, y+b) - \mathfrak{S}_t(a, b)\,.$$

Then $\tilde{\mathfrak{S}}_t$ has the same distribution as $\mathfrak{S}_t$.

Lemma 10.2. *Define an oriented polygon Γ by the cyclic sequence of vertices*

$$((x_1, y_1), \ldots, (x_{n-1}, y_{n-1}), (x_n, y_n), (x_1, y_1))\,, \qquad (x_i, y_i) \in \mathbb{Z}^2\,.$$

The signed area enclosed by the polygon Γ_t, given by

$$(\mathfrak{S}_t(x_1, y_1), \mathfrak{S}_t(x_2, y_2), \ldots, \mathfrak{S}_t(x_n, y_n), \mathfrak{S}_t(x_1, y_1))\ ,$$

is a martingale.

Proof. For convenience we write $(x_i(t), y_i(t)) = \mathfrak{S}_t((x_i, y_i))$. The area enclosed by the oriented polygon Γ_t is given by the integral $\frac{1}{2}\int_{\Gamma_t} x\mathrm{d}y \,-\, y\mathrm{d}x$, which we can write as

$$\frac{1}{2}\sum_{i=1}^{n}\left(x_i(t)y_{i+1}(t) - x_{i+1}(t)y_i(t)\right)\ ,$$

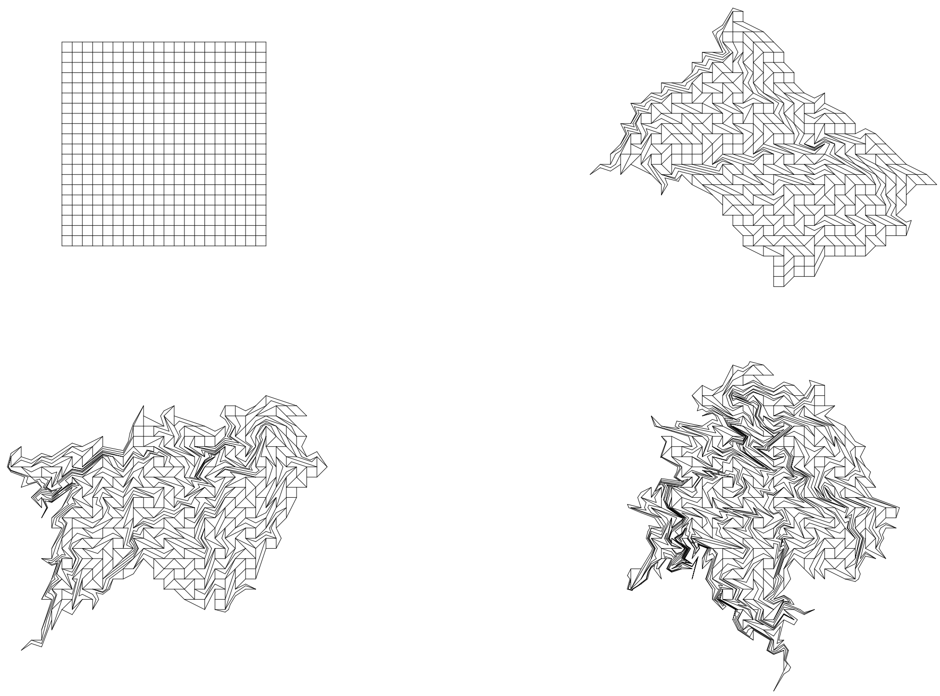

Figure 3: A simulation of $\mathfrak{S}_t$, shown at times $t = 0, \pi/4, \pi/2, 3\pi/4$ acting on a 20×20 box.

where (x_{n+1}, y_{n+1}) is taken to mean (x_1, y_1). So it suffices to show that each term in this sum is itself a martingale; let us concentrate on the term $x_1(t)y_2(t) - x_2(t)y_1(t)$, considering the first positive time at which either of $(x_1(t), y_1(t))$ or $(x_2(t), y_2(t))$ jumps. There appear to be at least 36 cases to consider, depending on the ordering of $\{0, x_1, x_2\}$ and $\{0, y_1, y_2\}$, but we can reduce this to four by taking advantage of the spatial homogeneity of the earthquakes model, described above. By choosing (a, b) suitably, and replacing $\mathfrak{S}$ by $\mathfrak{S}'$, we can assume that $x_i, y_i > 0$, for $i = 1, \ldots, n$. Furthermore, swapping the indices 1 and 2 only changes the sign of $x_1(t)y_2(t) - x_2(t)y_1(t)$, so we may also assume that $x_1 \leq x_2$. Suppose that the first earthquake of interest is along a vertical line. Then with probability x_1/x_2 it increments both y_1 and y_2 and otherwise it increments only y_2. The expected jump in $x_1(t)y_2(t) - x_2(t)y_1(t)$ conditional on the first relevant earthquake being parallel to the y-axis is therefore

$$\frac{x_1}{x_2}\left((x_1(y_2+1) - x_2(y_1+1)) - (x_1y_2 - x_2y_1)\right) + \frac{x_2 - x_1}{x_2}\left((x_1(y_2+1) - x_2y_1) - (x_1y_2 - x_2y_1)\right) = 0\,.$$

A similar argument shows that the expected jump in $x_1(t)y_2(t) - x_2(t)y_1(t)$ conditional on the first relevant earthquake being parallel to the x-axis is also zero. □

10.3 Random walks across the positive quadrant

In this section we describe another possible generalization of the simple harmonic urn that has some independent interest. We define a discrete-time process $(A_n, B_n)_{n\in\mathbb{Z}_+}$ on $\mathbb{R}^2$ based on the distribution of an underlying non-negative, non-arithmetic random variable X with $\mathbb{E}[X] = \mu \in (0, \infty)$ and $\mathbb{V}\mathrm{ar}[X] = \sigma^2 \in (0, \infty)$. Let $X_1, X_2, \ldots$ and $X_1', X_2', \ldots$ be independent copies of X. Roughly speaking, the walk starts on the horizontal axis and takes jumps $(-X_i', X_i)$ until its first component is negative. At this point suppose the walk is at $(-r, s)$. Then the walk starts again at $(s, 0)$ and the process repeats. We will see (Lemma 10.4) that in the case when $X \sim U(0, 1)$, this process is closely related to the simple harmonic urn and is consequently transient. It is natural to study the same question for general distributions X. It turns out that the recurrence classification depends only on μ and σ^2. Our proof uses renewal theory.

We now formally define the model. With X, X_n, X'_n as above, we suppose that $\mathbb{E}[X^4] < \infty$. Let $(A_0, B_0) = (a, 0)$, for $a > 0$. Define the random process for $n \in \mathbb{Z}_+$ by

$$(A_{n+1}, B_{n+1}) = \begin{cases} (A_n - X'_n, B_n + X_n) & \text{if } A_n \geq 0 \\ (B_n, 0) & \text{if } A_n < 0. \end{cases}$$

Theorem 10.3. *Suppose $\mathbb{E}[X^4] < \infty$. The walk (A_n, B_n) is transient if and only if $\mu^2 > \sigma^2$.*

Set $\tau_0 := -1$ and for $k \in \mathbb{N}$,

$$\tau_k := \min\{n > \tau_{k-1} : A_n < 0\}.$$

Define $T_k := \tau_k - (\tau_{k-1} + 1)$. That is, T_k is the number of steps that the random walk takes to cross the positive quadrant for the kth time.

Lemma 10.4. *If $X \sim U(0,1)$ and the initial value a is distributed as the sum of n independent $U(0,1)$ random variables, independent of the X_i and X'_i, then the distribution of the process $(T_k)_{k\in\mathbb{N}}$ coincides with that of the embedded simple harmonic urn process $(Z_k)_{k\in\mathbb{N}}$ conditional on $Z_0 = n$.*

Proof. It suffices to show that $T_1 = \tau_1$ has the distribution of Z_1 conditional on $Z_0 = n$ and that conditional on τ_1 the new starting point $A_{1+\tau_1}$, which is B_{τ_1}, has the distribution of the sum of τ_1 independent $U(0,1)$ random variables. Then the lemma will follow since the two processes (τ_k, B_{τ_k}) and (Z_k) are both Markov. To achieve this, we couple the process (A_n, B_n) up to time τ_1 with the renewal process described in Section 4. To begin, identify a with the sum $(1 - \chi_1) + \cdots + (1 - \chi_n)$. Then for $k \in \{1, \ldots, N(n) - n\}$, where $N(n) > n$ is as defined at (15), we identify X'_k with χ_{n+k}. For $m \leq \tau_1$ we have

$$A_m = a - \sum_{i=1}^{m} X'_i = n - \sum_{i=1}^{n+m} \chi_i ,$$

so in particular we have $A_{N(n)-n-1} \geq 0$ and $A_{N(n)-n} < 0$ by definition of $N(n)$. Hence $\tau_1 = N(n) - n$ has the distribution of Z_1 by Lemma 4.1. Moreover, $A_{1+\tau_1} = B_{\tau_1}$ is the sum of the independent $U(0,1)$ random variables X_i, $i = 1, \ldots, \tau_1$. □

Thus by Theorem 2.1, in the case where X is $U(0,1)$, the process (A_n, B_n) is transient, which is consistent with Theorem 10.3 since in the uniform case $\mu = 1/2$ and $\sigma^2 = 1/12$. To study the general case, it is helpful to rewrite the definition of (A_n, B_n) explicitly in the language of renewal theory. Let $S_0 = S'_0 = 0$ and for $n \in \mathbb{N}$ set $S_n = \sum_{i=1}^n X_i$, $S'_n = \sum_{i=1}^n X'_i$. Define the renewal counting function for S'_n for $a > 0$ as

$$N(a) := \min\{n \in \mathbb{Z}_+ : S'_n > a\} = 1 + \max\{n \in \mathbb{Z}_+ : S'_n \leq a\}.$$

Then starting at $(A_0, B_0) = (a, 0)$, $a > 0$, we see $\tau_1 = N(a)$ so that $B_{\tau_1} = S_{N(a)}$. To study the recurrence and transience of (A_n, B_n) it thus suffices to study the process $(R_n)_{n\in\mathbb{Z}_+}$ with $R_0 := a$ and R_n having the distribution of $S_{N(x)}$ given $R_{n-1} = x$. The increment of the process R_n starting from x thus is distributed as $\Delta(x) := S_{N(x)} - x$. It is this random quantity that we need to analyse.

Lemma 10.5. *Suppose that $\mathbb{E}[X^4] < \infty$. Then as $x \to \infty$, $\mathbb{E}[|\Delta(x)|^4] = O(x^2)$ and*

$$\mathbb{E}[\Delta(x)] = \frac{\sigma^2 + \mu^2}{2\mu} + O(x^{-1})$$
$$\mathbb{E}[\Delta(x)^2] = \frac{2x\sigma^2}{\mu} + O(1).$$

Proof. We make use of results on higher-order renewal theory expansions due to Smith [39] (note that in [39] the renewal at 0 is not counted). Conditioning on $N(x)$ and using the independence of the X_i, X_i', we obtain the Wald equations:

$$\mathbb{E}[S_{N(x)}] = \mu\mathbb{E}[N(x)]; \quad \mathbb{V}\mathrm{ar}[S_{N(x)}] = \sigma^2\mathbb{E}[N(x)] + \mu^2\mathbb{V}\mathrm{ar}[N(x)].$$

Assuming $\mathbb{E}[X^3] < \infty$, [39, Thm. 1] shows that

$$\begin{aligned} \mathbb{E}[N(x)] &= \frac{x}{\mu} + \frac{\sigma^2 + \mu^2}{2\mu^2} + O(x^{-1}), \\ \mathbb{V}\mathrm{ar}[N(x)] &= \frac{x\sigma^2}{\mu^3} + O(1). \end{aligned}$$

The expressions in the lemma for $\mathbb{E}[\Delta(x)]$ and $\mathbb{E}[\Delta(x)^2]$ follow.

It remains to prove the bound for $\mathbb{E}[|\Delta(x)|^4]$. Write $\Delta(x)$ as

$$S_{N(x)} - x = (S_{N(x)} - \mu N(x)) + (\mu N(x) - \mu\mathbb{E}[N(x)]) + (\mu\mathbb{E}[N(x)] - x). \tag{63}$$

Assuming $\mathbb{E}[X^2] < \infty$, a result of Smith [39, Thm. 4] implies that the final bracket on the right-hand side of (63) is $O(1)$. For the first bracket on the right-hand side of (63), it follows from the Marcinkiewicz–Zygmund inequalities [14, Cor. 8.2, p. 151] that

$$\mathbb{E}[(S_{N(x)} - \mu N(x))^4] \leq C\mathbb{E}[N(x)^2],$$

provided $\mathbb{E}[X^4] < \infty$. This last upper bound is $O(x^2)$ by the computations in the first part of this proof. It remains to deal with the second bracket on the right-hand side of (63). By the algebra relating central moments to cumulants, we have

$$\mathbb{E}[(\mu N(x) - \mu\mathbb{E}[N(x)])^4] = \mu^4(k_4(x) + 3k_2(x)^2),$$

where $k_r(x)$ denotes the rth cumulant of $N(x)$. Again appealing to a result of Smith [39, Cor. 2, p. 19], we have that $k_2(x)$ and $k_4(x)$ are both $O(x)$ assuming $\mathbb{E}[X^4] < \infty$. (The fact that [39] does not count the renewal at 0 is unimportant here, since the rth cumulant of $N(x) \pm 1$ differs from $k_r(x)$ by a constant depending only on r.) Putting these bounds together, we obtain from (63) and Minkowski's inequality that $\mathbb{E}[(S_{N(x)} - x)^4] = O(x^2)$. □

To prove Theorem 10.3 we basically need to compare $\mathbb{E}[\Delta(x)]$ to $\mathbb{E}[\Delta(x)^2]$. As in our analysis of $\tilde{Z}_k$, it is most convenient to work on the square-root scale. Set $V_n := R_n^{1/2}$.

Lemma 10.6. *Suppose that $\mathbb{E}[X^4] < \infty$. Then there exists $\delta > 0$ such that as $y \to \infty$,*

$$\begin{aligned} \mathbb{E}[V_{n+1} - V_n \mid V_n = y] &= \frac{\mathbb{E}[\Delta(y^2)]}{2y} - \frac{\mathbb{E}[\Delta(y^2)^2]}{8y^3} + O(y^{-1-\delta}), \\ \mathbb{E}[(V_{n+1} - V_n)^2 \mid V_n = y] &= \frac{\mathbb{E}[\Delta(y^2)^2]}{4y^2} + O(y^{-\delta}), \\ \mathbb{E}[|V_{n+1} - V_n|^3 \mid V_n = y] &= O(1). \end{aligned}$$

Proof. The proof is similar to the proof of Lemma 7.7, except that here we must work a little harder as we have weaker tail bounds on $\Delta(x)$. Even so, the calculations will be familiar, so we do not give all the details. Write $\mathbb{E}_x[\,\cdot\,]$ for $\mathbb{E}[\,\cdot \mid R_n = x]$ and similarly for $\mathbb{P}_x$. From Markov's inequality and the 4th moment bound in Lemma 10.5, we have for $\varepsilon \in (0, 1)$ that

$$\mathbb{P}_x(|\Delta(x)| > x^{1-\varepsilon}) = O(x^{4\varepsilon-2}). \tag{64}$$

We have that for $x \geq 0$,

$$\mathbb{E}[V_{n+1} - V_n \mid V_n = x^{1/2}] = \mathbb{E}_x[R_{n+1}^{1/2} - R_n^{1/2}] = \mathbb{E}_x[(x + \Delta(x))^{1/2} - x^{1/2}].$$

Here we can write

$$(x + \Delta(x))^{1/2} - x^{1/2} = [(x + \Delta(x))^{1/2} - x^{1/2}]\mathbf{1}\{|\Delta(x)| \leq x^{1-\varepsilon}\} + R_1 + R_2, \tag{65}$$

for remainder terms R_1, R_2 that we define shortly. The main term on the right-hand side admits a Taylor expansion and analysis (whose details we omit) in a similar manner to the proof of Lemma 7.7, and contributes to the main terms in the statement of the present lemma. The remainder terms in (65) are

$$R_1 = [(x + \Delta(x))^{1/2} - x^{1/2}]\mathbf{1}\{\Delta(x) > x^{1-\varepsilon}\}, \quad R_2 = [(x + \Delta(x))^{1/2} - x^{1/2}]\mathbf{1}\{\Delta(x) < -x^{1-\varepsilon}\}.$$

For the second of these, we have $|R_2| \leq x^{1/2}\mathbf{1}\{\Delta(x) < -x^{1-\varepsilon}\}$, from which we obtain, for $r < 4$, $\mathbb{E}_x[|R_2|^r] = O(x^{4\varepsilon+(r-4)/2})$, by (64). Taking ε small enough, this term contributes only to the negligible terms in our final expressions. For R_1, we have the bound

$$|R_1| \leq C(1 + |\Delta(x)|)^{(1/2)+\varepsilon}\mathbf{1}\{\Delta(x) > x^{1-\varepsilon}\}$$

for some $C \in (0, \infty)$ not depending on x, again for ε small enough. An application of Hölder's inequality and the bound (64) implies that, for $r < 4$, for any $\varepsilon > 0$,

$$\mathbb{E}_x[|R_1|^r] \leq C(\mathbb{E}_x[(1 + |\Delta(x)|)^4])^{\frac{r(1+2\varepsilon)}{8}}(\mathbb{P}_x(\Delta(x) > x^{1-\varepsilon}))^{1-\frac{r(1+2\varepsilon)}{8}} = O(x^{6\varepsilon+(r-4)/2}).$$

It is now routine to complete the proof. $\square$

Proof of Theorem 10.3. For the recurrence classification, the crucial quantity is

$$2y\mathbb{E}[V_{n+1} - V_n \mid V_n = y] - \mathbb{E}[(V_{n+1} - V_n)^2 \mid V_n = y] = \mathbb{E}[\Delta(y^2)] - \frac{\mathbb{E}[\Delta(y^2)^2]}{2y^2} + O(y^{-\delta}),$$

by Lemma 10.6. Now by Lemma 10.5, this last expression is seen to be equal to

$$\frac{\mu^2 - \sigma^2}{2\mu} + O(y^{-\delta}).$$

Now [26, Thm. 3.2] completes the proof. $\square$

Remarks. (i) To have some examples, note that if X is exponential, the process is recurrent, while if X is the sum of two independent exponentials, it is transient. We saw that if X is $U(0,1)$ the process is transient; if X is the square-root of a $U(0,1)$ random variable, it is recurrent.

(ii) Another special case of the model that has some interesting features is the case where X is exponential with mean 1. In this particular case, a calculation shows that the distribution of T_{k+1} given $T_k = m$ is negative binomial $(m+1, 1/2)$, i.e.,

$$\mathbb{P}(T_{k+1} = j \mid T_k = m) = \binom{j+m}{m}2^{-m-j-1}, \quad (j \in \mathbb{Z}_+).$$

Since $\mu^2 = \sigma^2$, this case is in some sense critical, a fact supported by the following branching process interpretation.

Consider a version of the gambler's ruin problem. The gambler begins with an initial stake, a pile of m_0 chips. A sequence of independent tosses of a fair coin is made; when the coin comes up heads, a chip is removed from the gambler's pile, but when it comes up tails, a chip is added to a second pile by the casino. The game ends when the gambler's original pile of chips is exhausted; at this point the gambler receives the second pile of chips as his prize. The total number of chips in play is a martingale; by the optional stopping theorem, the expectation of the prize equals the initial stake. As a loss leader, the casino announces that it will add one extra chip to each gambler's initial stake, so that the game is now in favour of the gambler. Suppose a gambler decides to play this game repeatedly, each time investing his prize as the initial stake of the next game. If the casino were to allow a zero stake (which of course it does not), then the sequence of augmented stakes would form an irreducible Markov chain S_k on $\mathbb{N}$. Conditional on $S_k = m$, the distribution of S_{k+1} is negative binomial $(m + 1/2, 1/2)$. So by the above results, this chain is recurrent. It follows that with probability one the gambler will eventually lose everything.

We can interpret the sequence of prizes as a Galton–Watson process in which each generation corresponds to one game, and individuals in the population correspond to chips in the gambler's pile at the start of the game. Each individual has a Geo(1/2) number of offspring (i.e., the distribution that puts mass 2^{-1-k} on each $k \in \mathbb{Z}_+$), being the chips that are added to the prize pile while that individual is on top of the gambler's pile, and at each generation there is additionally a Geo(1/2) immigration, corresponding to the chips added to the prize pile while the casino's bonus chip is on top of the gambler's pile. This is a critical case of the Galton–Watson process with immigration. By a result of Zubkov [44], if we start at time 0 with population 0, the time τ of the next visit to 0 has pgf

$$\mathbb{E}[s^\tau] = \frac{1}{s} + \frac{1}{\log(1-s)} .$$

Since this tends to 1 as $s \nearrow 1$, we have $\mathbb{P}(\tau < \infty) = 1$. In fact this can be deduced in an elementary way as follows. The pgf of the Geo(1/2) distribution is $f(s) = 1/(2-s)$, and its n^{th} iterate, the pgf of the n^{th} generation starting from one individual, is $f(s) = (n-(n-1)s)/((n+1)-ns)$. In particular the probability that an individual has no descendants at the n^{th} generation is $n/(n+1)$. If $S_0 = 1$, then $S_k = 1$ if and only if for each $j = 0, \dots, k-1$ the bonus chip from game j has no descendants at the $(k-j)^{\text{th}}$ generation. These events are independent, so

$$\mathbb{P}(S_k = 1 \mid S_0 = 1) = \prod_{j=0}^{k-1} \frac{k-j}{k-j-1} = \frac{1}{k+1},$$

which sums to ∞ over $k \in \mathbb{N}$ so that the Markov chain is recurrent (see e.g. [1, Prop. 1.2, §I]).

The results of Pakes [33] on the critical Galton–Watson process with immigration show that the casino should certainly not add two bonus chips to each stake, for then the process becomes transient, and gambler's ruin will no longer apply.

Acknowledgements

The authors are grateful to John Harris, Sir John Kingman, Iain MacPhee and James Norris for helpful discussions, and to an anonymous referee for constructive comments. Most of this work was carried out while Andrew Wade was at the Heilbronn Institute for Mathematical Research, Department of Mathematics, University of Bristol.

References

[1] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd edition, Springer.

[2] Aspandiiarov, S. and Iasnogorodski, R. (1997). Tails of passage-times and an application to stochastic processes with boundary reflection in wedges. *Stoch. Proc. Appl.* **66**, 115–145.

[3] Aspandiiarov, S., Iasnogorodski, R., and Menshikov, M. (1996). Passage-time moments for nonnegative stochastic processes and an application to reflected random walks in a quadrant. *Ann. Probab.* **24**, 932–960.

[4] Athreya, K.B. and Karlin, S. (1968). Embedding of urn schemes into continuous time Markov branching process and related limit theorems. *Ann. Math. Statist.* **39**, 1801–1817.

[5] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*, Springer. (Dover reprint 2004.)

[6] Bóna, M. (2004). *Combinatorics of Permutations*, Chapman and Hall / CRC press.

[7] Churchill, R.V. (1937). The inversion of the Laplace transformation by a direct expansion in series and its application to boundary-value problems. *Math. Z.* **42**, 567–579.

[8] Cox, D.R. (1962). *Renewal Theory*, Methuen.

[9] Feller, W. (1941). On the integral equation of renewal theory. *Ann. Math. Statist.* **12**, 243–267.

[10] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I. 3rd edition, Wiley.

[11] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II. 2nd edition, Wiley.

[12] Fisch, R., Gravner, J., and Griffeath, D. (1991) Cyclic cellular automata in two dimensions. pp. 171–185 in: *Spatial Stochastic Processes* (K.S. Alexander and J.C. Watkins, eds.), Birkhäuser.

[13] Flajolet, P., Gabarró, J., and Pekari, H. (2005). Analytic urns. *Ann. Probab.* **33**, 1200–1233.

[14] Gut, A. (2005). *Probability: A Graduate Course*, Springer.

[15] Harris, T.E. (1952). First passage and recurrence distributions. *Trans. Amer. Math. Soc.* **73**, 471–486.

[16] Hoffman, J.R. and Rosenthal, J.S. (1995). Convergence of independent particle systems. *Stochastic Process. Appl.* **56**, 295–305.

[17] Hutton, J. (1980). The recurrence and transience of two-dimensional linear birth and death processes. *Adv. Appl. Probab.* **12**, 615–639.

[18] Janson, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110**, 177–245.

[19] Jensen, U. (1984) Some remarks on the renewal function of the uniform distribution. *Adv. Appl. Probab.* **16**, 214–215.

[20] Johnson, N. and Kotz, S. (1977). *Urn Models and Their Applications: an Approach to Modern Discrete Probability Theory*, Wiley, New York.

[21] Karlin, S. and Taylor, H.M. (1981). *A Second Course in Stochastic Processes*, Academic Press.

[22] Kesten, H. (1976). Recurrence criteria for multi-dimensional Markov chains and multi-dimensional linear birth and death processes. *Adv. Appl. Probab.* **8**, 58–87.

[23] Kingman, J.F.C. (1999). Martingales in the OK Corral. *Bull. Lond. Math. Soc.* **31**, 601–606.

[24] Kingman, J.F.C. and Volkov, S.E. (2003). Solution to the OK Corral model via decoupling of Friedman's urn. *J. Theoret. Probab.* **16**, 267–276.

[25] Kotz, S. and Balakrishnan, N. (1997). Advances in urn models during the past two decades. pp. 203–257 in: *Advances in Combinatorial Methods and Applications to Probability and Statistics*, Stat. Ind. Technol., Birkhäuser.

[26] Lamperti, J. (1960). Criteria for the recurrence or transience of stochastic processes I. *J. Math. Anal. Appl.* **1**, 314–330.

[27] Lamperti, J. (1962). A new class of probability limit theorems. *J. Math. Mech.* **11**, 749–772.

[28] Lamperti, J. (1963). Criteria for stochastic processes II: passage-time moments. *J. Math. Anal. Appl.* **7**, 127–145.

[29] MacPhee, I.M. and Menshikov, M.V. (2003). Critical random walks on two-dimensional complexes with applications to polling systems. *Ann. Appl. Probab.* **13**, 1399–1422.

[30] Mahmoud, H.M. (2009). *Pólya Urn Models*, CRC press.

[31] Menshikov, M.V., Asymont, I.M., and Iasnogorodskii, R. (1995). Markov processes with asymptotically zero drifts. *Probl. Inf. Transm.* **31**, 248–261; translated from *Problemy Peredachi Informatsii* **31**, 60–75 (in Russian).

[32] Menshikov, M.V., Vachkovskaia, M., and Wade, A.R. (2008). Asymptotic behaviour of randomly reflecting billiards in unbounded tubular domains. *J. Statist. Phys.* **132**, 1097–1133.

[33] Pakes, A.G. (1971) On the critical Galton-Watson process with immigration, *J. Austral. Math. Soc.* **12**, 476–482.

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

[35] Pemantle, R. and Volkov, S. (1999). Vertex-reinforced random walk on Z has finite range. *Ann. Probab.* **27**, 1368–1388.

[36] Pittel, B. (1987). An urn model for cannibal behaviour. *J. Appl. Probab.* **24**, 522–526.

[37] Robert, P. (2003). *Stochastic Networks and Queues*, English edition, Springer.

[38] Smith, W.L. (1954). Asymptotic renewal theorems. *Proc. Roy. Soc. Edinburgh Sect. A* **64**, 9–48.

[39] Smith, W.L. (1959). On the cumulants of renewal processes. *Biometrika* **46**, 1–29.

[40] Stone, C. (1965). On moment generating functions and renewal theory. *Ann. Math. Statist.* **36**, 1298–1301.

[41] Ugrin-Šparac, G. (1990). On a distribution encountered in the renewal process based on uniform distribution. *Glasnik Mate.* **25**, 221–233.

[42] Watson, R.K. (1976). An application of martingale methods to conflict models. *Oper. Res.* **24**, 380–382.

[43] Williams, D. and McIlroy, P. (1998). The OK Corral and the power of the law (a curious Poisson kernel formula for a parabolic equation). *Bull. Lond. Math. Soc.* **30**, 166–170.

[44] Zubkov, A.M. (1972). Life-periods of a branching process with immigration. *Theor. Probab. Appl.* **17**, 174–183.